PREFACE

Understanding how worker well-being is distributed across the population is of paramount importance. With such knowledge policy makers can devise efficient strategies to improve social welfare. This volume contains 13 chapters on topics enhancing our comprehension of inequality across workers. The issues addressed deal directly with the economic institutions that affect individual and family earnings distributions. The themes explored include job training, worker and firm mobility, minimum wages, wage arrears, unions, collective bargaining, unemployment insurance, and schooling. Among the questions answered are: To what extent do greater work hours of women mitigate the widening family earnings distribution? To what extent does deunionization widen the distribution of earnings? Do computers really cause a widening of the earnings distribution? How would the Russian wage distribution change if one accounted for wage arrears? How much of job creation and job destruction comes about because of business relocation? To what extent does maternal education increase children's education? Why do increases in the minimum wage fail to substantially decrease employment as economic theory would predict? And, to what extent do job skills matter for low-income workers?

The widening dispersion of earnings in the United States and other economies over the last 30 years is now well documented. Not only has this dispersion grown for individual wage earners, but family earnings has become more dispersed, as well. However, understanding family earnings dispersion is complicated because labor force participation decisions of husbands and wives are interrelated. In the first chapter, John Pencavel examines US family earnings inequality between 1926 and 1995. First, he shows that earnings inequality among all couples has increased over the sample period. Concomitant with this increased disparity is higher earnings inequality both among men and women, but women's higher labor hours have had a mitigating impact on inequality among couples. Pencavel studies to what extent changes in earnings inequality (for husbands and wives separately) are driven by changes in employment.

Decisions regarding where and how much to work are in part potentially related to job creation and destruction. In the second chapter, David Neumark, Junfu Zhang, and Brandon Wall present a new data source – the National Establishment Time Series (NETS) – which offer rich possibilities for studying employment dynamics by tracking business establishment relocations that contribute to regional job creation/destruction. The authors first assess the quality and the measurement accuracy of the data by comparing the California extract to alternative data sources along various dimensions. Then they decompose employment changes into components related to (i) changes in the size of existing firms; (ii) changes due to birth and death of establishments; and (iii) changes due to relocation of firms (into and out of California). The chapter provides evidence that the highly debated phenomenon of business relocation accounts only for a small share of the overall job creation/destruction process, and the chapter derives policy conclusions.

In part, earnings dispersion has been widening because of increased training. In the next chapter, Alison L. Booth Mark L. Bryan examine who pays for training. It is one of a new genre of articles to find that corporations finance general training, counter to the prediction of the commonly applied human capital model that assumes no fixed costs of job mobility. Booth and Bryan use the British Household Panel Survey (BHPS) to show that despite the most work-related training being general the preponderance of training is paid by the firm despite most work-related training being general. This result is consistent either with credit constraints or substantial fixed costs of mobility, and implies that general training is more specific than once thought.

Not widely studied is earnings dispersion in Russia. In the next chapter, Hartmut Lehmann and Jonathan Wadsworth use the Russian Longitudinal Monitoring Survey for the years 1994–1998 to assess the effects of wage arrears on wage inequality. Specifically, using various econometric techniques, Lehmann and Wadsworth estimate what the wage distribution would look like if all workers had been paid their full contractual wage on time, i.e., if there were no arrears. This counterfactual series suggests that wage inequality would have been some 30% larger if workers had been paid in full. Moreover, since wage arrears affect men more than women, the gender pay gap would have been around 10% higher than the observed gap. On the other hand, both regional pay differentials and sectoral differentials would have been narrower in the absence of arrears. In short, wage arrears widen the observed earnings distribution. Thus, one must take arrears into account when making policy recommendations based on the overall wage distribution.

Most of the current literature argues that one can attribute the widening of the earnings distribution over the last 30 years to skill-biased technological change (SBTC). Many have argued that in part SBTC comes about because computers have become particularly useful in the workplace. In the next chapter, Michael J. Handel examines four possible mechanisms by which computers can affect skill demand. However, he finds none of these potential causal links between computers and wages to be strong and that the individual computing wage premium is negligible. In addition, he argues the timing and magnitude of the increase in computer usage appear inconsistent with the rise in inequality. He concludes that computers have done little to change the US wage structure in the last 20–30 years.

Where migrants locate geographically is important not only to the migrants but also to policy makers seeking to control a particular area's economic growth. One commonly observed phenomenon regarding locational choice is immigrant clustering, whereby immigrants of a particular racial, ethnic, or religious ilk locate in areas populated by similar inhabitants. In other words, a location has significant externalities based on its "ethnic capital". In the next chapter, Thomas Bauer, Gil S. Epstein, and Ira N. Gang perform an empirical study of Mexican migration to various US locations. Their innovation is to get at a location's ethnic capital by showing how an area's migrant "stock" and migrant "flow" affect the probability of migration. The significance and size of the effects vary according to the migrant's legal status and whether the migrant is a "new" or a "repeat" mover.

Understanding low-skilled jobs is important for policy makers seeking to alleviate poverty. The next chapter by Rucker C. Johnson analyzes wage growth prospects of former and current welfare recipients. He finds job markets for these workers have many of the same features as typical labor markets for the mainstream population. As such, jobs differ in their prospects for wage growth. Some jobs allow for wage increases and further job advances over the course of employment, while others do not. Similarly, low-skilled workers differ in their abilities and skills. Using the Women's Employment Survey along with the Michigan Employer Survey containing data from 1997 to 2004, Johnson finds that even in low-skilled jobs, workers sort based on ability. Workers with greater relative skills (such as knowledge of the computer) gravitate toward jobs with greater skill requirements and achieve a larger wage growth. Those relatively skilled workers initially in less desirable jobs move to better ones, so that turnover is smaller when able workers initially attain relatively more skilled jobs with higher wage growth. From a policy perspective these results question welfare reform that concentrates solely on job placement rather than training because in the end skills are found to be important even for current and former welfare recipients.

Unemployment insurance (UI) is a mechanism of government's mandate to ease a workers downside risk of unemployment. Typically employers and/or employees are required to pay into a central fund from which workers can draw if they later become unemployed. Rates are set using the ''law of large numbers'' that implies that the reported losses will be based on the underlying probability of the loss. In the long run the premium for each worker and firm should reflect the expected loss equally across all the insured. However, in the short run or with imperfect rating schemes cross-subsidization can occur. In the next chapter, by using 1986–1996 Canadian data that link firms, workers and claimants, Miles Corak and Wen-Hao Chen compute cross-subsidization benefits across industries, provinces, and firms, as well as the dead-weight losses of the Canadian UI system. They find significant transfers from cyclical to non-cyclical industries and significant transfers to industries with high separation rates and low wages. Also there is significant cross-subsidization between firms in subsidizing as well as receiving industries.

The widening earnings dispersion observed in many developed countries is now well documented. Also well documented is the decline of union membership in Britain. Given that unions tend to equalize wages, one can ask how much of the increase in Britain's wage dispersion is caused by declining union representation. In the next chapter, John T. Addison, Ralph W. Bailey, and W. Stanley Siebert utilize the 1983 General Household Survey (GHS) data and the 1995 Labor Force Survey (LFS) to answer this question. They find that the large decline in union density accounts for little of the increase in earnings variation in the private sector, either for men or women. However, in the public sector, although union density declined less precipitously, earnings dispersion has more or less held steady. The difference, they argue, results because public sector unions organized relatively skilled workers. As such, changes in the composition of unionized workers are important in understanding earnings dispersion.

Of course, the future of any nation lies in the human capital acquisition of its children. But in the underdeveloped world, acquiring human capital often costs the household dearly, so that overall levels of education remain relatively low. This is particularly true for girls in Nepal where literacy rates are particularly depressed. Thus, understanding the factors affecting children's education is important to get these types of countries on a path to higher plateaus of development. Development economists often model household behavior and test their models with data, so they can ascertain the factors that enhance the probability children get more education. However, one problem with empirical work is the type sampling techniques used to gather data. In particular, most surveys in developing countries are two-stage stratified samples of households in which the first stage samples villages, and the second samples households from within each village. However, households within each village often have similar characteristics, so that ignoring these cluster fixed effects is likely to result in biased estimates. In the next chapter, Diane Dancer and Anu Rammohan utilize a household Nash bargaining model to obtain derived demand curves for children's education. They then employ a cluster fixed-effects model using the Nepal Demographic Household Survey. As might be expected, they find that boys receive more education than girls, and that higher maternal education (both primary and secondary) more greatly affects the schooling of girls. Greater household wealth equally increases education of male and female children.

One important and often debated question is the effect of raising the minimum wage. At least for Britain and the United States a number of studies found only a meager detrimental impact on employment. These weak employment effects have served to justify small increases in the minimum wage. In the next chapter, Sara Lemos uses monthly data for Brazil from 1982 to 2000 to show that increases in the minimum wage raise not only wages but also prices. This sets off a wage–price inflationary spiral, but with little effect on employment. One implication is such inflationary pressures mitigate the power of using minimum wage increases as a tool help alleviate poverty.

In the next chapter, Cary Deck and Amy Farmer present a series of experiments to test how final offer and conventional arbitration affect bargaining outcomes. They consider the impact that the choice of dispute–resolution mechanism, conventional or final offer arbitration, has on settlement. They formally show that final offer arbitration can favor the informed party by shifting the contract zone toward more profitable allocations. Laboratory results confirm this result. Nonetheless, settlement is positively correlated with the width of the contract zone, which suggests that the location of the contract zone in final offer arbitration generates more disputes.

The final paper is purely theoretical. It analyzes why productivity is only weakly correlated with the business cycle, contrary to the implications of real business cycle models. In RBC models, positive productivity shocks raise the demand for labor, leading to higher levels of employment. In the model here, firms reduce their hiring standards in order to achieve their desired level of employment. As a result, firms increase the proportion of low-ability workers in their workforce, which moderates the observed change in productivity.

As with past volumes, we aimed to focus on important issues and to maintain the highest levels of scholarship. We encourage readers who have prepared manuscripts that meet these stringent standards to submit them to *RLE* via the IZA website (http://www.iza.org/index_html?lang=en&mainframe=http%3A//www.iza.org/en/webcontent/index_html) for possible inclusion in future volumes. For insightful editorial advice in preparing this volume, we thank Paul G. Althaus, Ann Bartel, Andrea H Beller, Mike Bognanno, Holger Bonin, Marco Castillo, Ludo Cuyvers, Andy Dickerson, Bruce Fallick, Gary S. Fields, Belton M. Fleisher, Alessandra Guariglia, Peter Haan, Todd Idson, Murat F. Iyigun, Peter Kuhn, David MacPherson Lena Nekby, Trond Petersen, Patrick Puhani, Barbara Rossi, Shannon Seitz, Wendy Sigle-Rushton, Curtis Simon, Konstantinos Tatsiramos, and Phanindra V. Wunnava.

Solomon W. Polachek
Oliver Bargain
Editors

LIST OF CONTRIBUTORS

John T. Addison	Department of Economics, University of South Carolina, USA; Universidade de Coimbra/GEMF, Portugal; IZA, Bonn, Germany
Ralph W. Bailey	Department of Economics, University of Birmingham, UK
Thomas Bauer	RWI Essen, Ruhr-Universität Bochum, IZA, Bonn, Germany; CEPR, London, UK
Alison L. Booth	University of Essex, the Australian National University, CEPR and IZA
Mark L. Bryan	Institute for Social and Economic Research (ISER), University of Essex, UK
Wen-Hao Chen	Family and Labour Studies, Statistics Canada
Miles Corak	Family and Labour Studies, Statistics Canada; IZA, Bonn, Germany
Diane Dancer	Econometrics and Business Statistics, University of Sydney, Australia
Cary Deck	Department of Economics, University of Arkansas, Fayetteville, AR, USA
Gil S. Epstein	Department of Economics, Bar-Ilan University, IZA, Bonn, Germany; CReAM, London, UK
Amy Farmer	Department of Economics, University of Arkansas, Fayetteville, AR, USA
Ira N. Gang	Department of Economics, Rutgers University; IZA, Bonn, Germany; CReAM, London, UK
Michael J. Handel	Department of Sociology, Northeastern University, Boston, MA, USA

Rucker C. Johnson	Goldman School of Public Policy, University of California, Berkeley, CA, USA
Hartmut Lehmann	University of Bologna; IZA, Bonn, Germany; CERT, Heriot-Watt University, Edinburgh, UK
Sara Lemos	Economics Department, University of Leicester, UK
David Neumark	Department of Economics, University of California-Irvine, USA; Public Policy Institute of California, San Francisco, CA, USA; NBER, Cambridge, MA, USA; IZA, Bonn, Germany
John Pencavel	Department of Economics, Stanford University, USA
Anu Rammohan	Department of Economics, University of Sydney, Australia
Michael Sattinger	Department of Economics, State University of New York at Albany, USA
W. Stanley Siebert	Department of Commerce, Birmingham University, UK
Sumati Srinivas	College of Business and Economics, Radford University, Radford, VA, USA
Jonathan Wadsworth	Royal Holloway College, University of London, UK; Centre for Economic Performance, LSE; IZA, Bonn, Germany
Brandon Wall	Department of Economics, Stanford University, USA
Junfu Zhang	Department of Economics, Clark University, USA

EARNINGS INEQUALITY AND MARKET WORK IN HUSBAND–WIFE FAMILIES ☆

John Pencavel

ABSTRACT

Constructing pseudo-panel data from successive Current Population Surveys, this paper analyzes earnings inequality in husband and wife families over the life cycle and over time. Particular attention is devoted to the role of labor supply in influencing measures of earnings inequality. Compact and accurate descriptions of earnings inequality are derived that facilitate the analysis of the effect of the changing market employment of wives on earnings inequality. The growing propensity of married women to work for pay has mitigated the increase in family earnings inequality. Alternative measures of earnings inequality covering people with different degrees of attachment to the labor market are constructed. Inferences about the extent and changes in earnings inequality are sensitive to alternative labor supply definitions especially in the case of wives.

☆ This paper benefitted from Alan Krueger's and Rob Valletta's comments on a previous draft and from the criticism of two anonymous referees.

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 1–37

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26001-1

1. INTRODUCTION

To what extent do the increases in earnings inequality among individual American workers pose an issue for public policy? To answer this, we would want to know the extent to which changes in individual earnings translate into changes in income inequality in the households within which these earnings are pooled and shared. The link between the earnings of one household member and the income consumed of each household member depends not only on the magnitude of this individual's earnings but on whether other household members work for pay, and, if so, how many hours they work, on other sources of income, and on changing patterns of household formation and dissolution. Hence, the connection between the growth in inequality of individual pay and changes in income consumed by individuals (both those who work for pay and those who do not) is complex and involves a number of interrelated factors.

Some of these links are traced out in this paper which focuses on income in husband–wife families. First, we determine the extent to which changes in income inequality are attributable to changes in inequality in labor market earnings. Second, we examine changes in family earnings inequality and assess how increases in wives' employment have affected family earnings inequality. To address this, a simple and compact accounting framework is derived that describes the movements of family earnings inequality and that may be used to discriminate between the part played by husbands' earnings and that played by married women's employment in understanding movements in family earnings inequality.

We then turn to earnings inequality of wives and husbands separately and ascertain how changes in earnings inequality are affected by differences in the degree to which the husbands and wives work in the labor market. Again, a simple expression is derived that links earnings inequality to the employment–population ratio. This inspires the more general question: are inferences about differences and changes in earnings inequality sensitive to variations among people in their commitment to market work? Imagine the population being censored in increasing degrees by the extent of their market work: have the changes in earnings inequality for these groups in the population been the same?

In addressing these questions, the analysis will recognize that income inequality varies over the lifetime: husband–wife incomes are more unequal among older couples than among younger couples. Furthermore, the past 30 years has seen an aging of the typical husband–wife couple induced in part by the postponement of age of first marriage. In 1967–1969, in almost

14 percent of all couples, the wives were aged between 20 and 25 years; by 1998–2000, only 5.4 percent were in this category. In 1998–2000, there were almost 10 percent more couples with wives aged above 36 years than there were in 1967–1969. By organizing the data by years since leaving school, we differentiate between two time effects on income inequality: the increase in income inequality associated with the aging of a household and the increase in income inequality that has occurred over time even among households of the same age.[1]

At the outset, some important restrictions on the analysis need to be noted. First, the data used in this paper are drawn from successive March Current Population Surveys and they do not constitute genuine panel data, which can record changes in the marital status of a given population. On the other hand, panel data have serious problems of nonrandom attrition with changes in marital status constituting one of the key reasons for losing individuals from the panel survey. The CPS allows the construction of pseudo-panels and, as the principal source of information about the U.S. labor force, the CPS provides a large and accurate characterization of the U.S. population.

Second, over the past 30 years or so, the number and attributes of married people have changed: many fewer adults are now married with spouse present and those who are married tend to be better schooled and older (relative to unmarried people) than they were in the 1960s. So married people at the end of our period are a more select group of the adult population.

Third, this paper focuses on incomes generated by the market so government taxes and transfers will be ignored. Of course, the presence of such taxes and transfers may well affect the level and structure of market incomes but this is neglected here. At the same time, the movement of pre-tax household income has followed closely the movement of post-tax household income even though there have been nonnegligible changes in the tax structure as in the Tax Reform Act of 1986 and the reform of the welfare system in the 1990s.[2]

We turn first to a description of the data and the methods underlying this research.

2. CONSTRUCTION OF THE DATA

There are different ways of examining the evolution of people's earnings over time. In a companion paper (Pencavel, 2006), husband and wife couples are organized by their year of birth and by their age.[3] In this paper, people are "born" when they have completed their schooling so each cohort is defined as the calendar year in which the cohort members left school and could have started their market work careers. Their "age" is measured by the years that have elapsed since schooling completion. Years since completion of schooling is called "experience".[4]

When a husband and wife are born in the same year and complete the same schooling, the couple's cohort and experience are the same whether defined by the husband's characteristics or the wife's. However, when the wife's year of birth and schooling differ from the husband's, their cohort and experience may not be the same. Because cross-classifying husbands and wives by the cohort and experience of each individual consumes many degrees of freedom, we define cohorts by 5-year intervals so that some couples of the same age would have to have large differences in schooling not to be in the same cohort and we index a family's "experience" by the years since the wife has left school. We organize the family's data by the wife's experience and cohort because the relationship between the employment and earnings of wives and family earnings inequality plays a special role in this analysis.

The Annual Demographic Supplements of the March Current Population Surveys for 1968 through to 2001 are used to sort husband and wife couples into cohorts defined by the estimated year of schooling completion and by the years of experience of the wife. Each cohort covers a 5-year interval from 1926–1930 to 1991–1995. Table 1 lists the resulting 294 cohort-experience cells we use. Each cell consists of no less than 1,000 husband–wife pairs.

Three components of family income are distinguished: the husband's earnings; the wife's earnings; and the interest, dividends, and rent received by the husband and wife. These components are measured before tax and transfers – the purpose is to examine the differences across families in the incomes generated by the market, not by the adjustments that governments make to these incomes – and they neglect the incomes of any other family members. For any experience x and cohort c cell, let $y_{Hi}(x,c)$ denote the annual earnings of the husband in household i, $y_{Wi}(x,c)$ the annual earnings of the wife in household i, and $y_{Ni}(x,c)$ the annual nonlabor income (the sum of dividends, interest, and rent) of household i.[5] To be included, both husband and wife must be at least 20 years of age and not more than 60 years. To avoid the difficulties in measuring the labor returns to people who are self-employed, couples containing a self-employed worker are excluded.

For most husband–wife families, labor market earnings constitute the most important components of income and nonlabor income represents a

Table 1. Definitions of Cells by Cohort and Experience (Omitting Cells with Fewer than 1,000 Husband–Wife Pairs).

Cohort	Years of Schooling Completion	Minimum Years of Experience	Maximum Years of Experience	Number of Cells
1	1926–1930	39	44	6
2	1931–1935	33	44	12
3	1936–1940	28	44	17
4	1941–1945	23	44	22
5	1946–1950	18	43	26
6	1951–1955	13	42	30
7	1956–1960	8	40	33
8	1961–1965	3	37	35
9	1966–1970	2	32	31
10	1971–1975	2	28	27
11	1976–1980	2	23	22
12	1981–1985	2	18	17
13	1986–1990	3	13	11
14	1991–1995	3	7	5
All	1926–1995	2	44	294

An individual's experience is defined as the minimum of (1) her years of age minus her years of schooling minus 6 and (2) her years of age minus 17. Then an individual's cohort is defined as the calendar year in which her experience is zero.

relatively small part. Across these 294 cells, the average of the ratio of nonlabor income to total income is 0.045. Furthermore, for a study of income inequality across all husband–wife families, variations in nonlabor income are not important. This is illustrated in Fig. 1 where for two cohorts, cohorts 6 and 9, the Gini coefficients of income inequality are graphed: one Gini coefficient includes nonlabor income (this is marked as "incl N") and the other Gini coefficient excludes nonlabor income (marked as "excl N"). For each cohort, income inequality rises with experience and the more recent cohort, cohort 9, exhibits greater income inequality than the earlier cohort, cohort 6. The values of the Gini coefficient that includes nonlabor income is slightly higher at low years of experience and slightly lower at high years of experience than the values of the Gini coefficient that excludes nonlabor income. However, the movements in the two Gini coefficients are close. Across all 294 experience-cohort cells, the correlation coefficient between the Gini coefficient including nonlabor income and the Gini coefficient excluding nonlabor income is 0.993. In view of this, we shall simplify our analysis of inequality by neglecting nonlabor income and by concentrating on labor

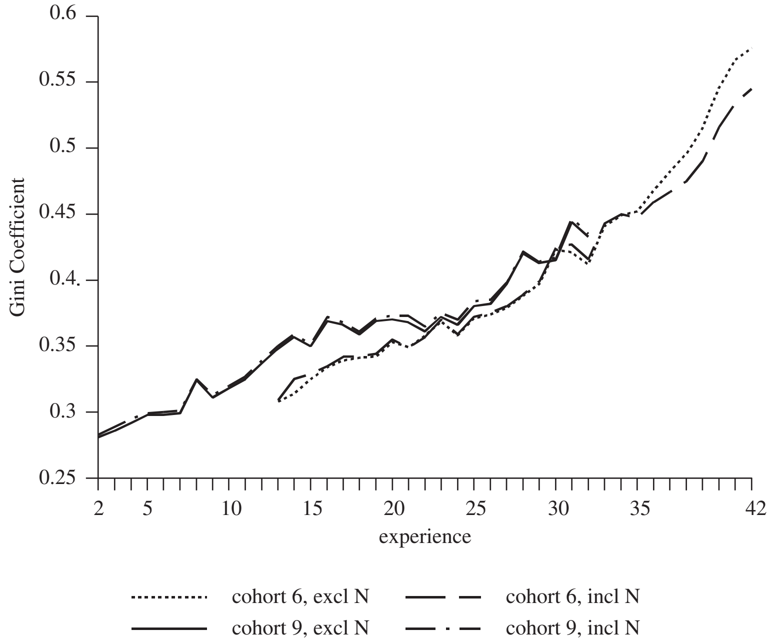


Fig. 1. Gini Coefficients for Cohorts 6 and 9: Including and Excluding Nonlabor Income. *Note:* "excl N" means excluding nonlabor income and "incl N" means including nonlabor income.

earnings. We shall call the sum of the husband's earnings and the wife's earnings "family" earnings: $y_i(x,c) = y_{Hi}(x,c) + y_{Wi}(x,c)$.[6]

3. AN EXPRESSION TO DESCRIBE MOVEMENTS IN FAMILY EARNINGS INEQUALITY

The purpose of this section is to derive an expression permitting reliable simulations of the movements of family earnings inequality. This expression is addressed to assessing the extent to which increases in wives' employment have affected family earnings inequality. The equation to be derived takes the following form:

$$G = \beta_0 + \beta_1 G_H + \beta_2 \frac{m_W E_W}{m_H E_H} + \beta_3 \left[\frac{m_W E_W}{m_H E_H}\right]^2 + u \qquad (1)$$

where G stands for the Gini coefficient of family earnings inequality, G_H is the Gini coefficient of husbands' earnings inequality, E_H and E_W are, respectively, the employment–population ratios of husbands and wives, m_H and m_W are the mean earnings of husbands and the mean earnings of wives among those husbands and wives employed for pay, and u is a term that incorporates other factors. The β's are parameters to be estimated. This expression is an approximation to an accounting framework. It will be shown that a very large fraction of the variations in family earnings inequality is removed by this linear (in the parameters) approximation and that Eq. (1) provides a compact means of discriminating between the roles of husbands' earnings inequality and married women's employment to describe the movements in family earnings inequality. We proceed to deriving and rationalizing Eq. (1).

If $\sigma_H\,(x,c)$ is the standard deviation of the earnings of husbands, $\sigma_W\,(x,c)$ the standard deviation of the earnings of wives, and $\sigma\,(x,c)$ the standard deviation of family earnings, then

$$\sigma^2(x,c) = \sigma_H^2(x,c) + \sigma_W^2(x,c) + 2r(x,c)\sigma_H(x,c)\sigma_W(x,c)$$

where $r(x,c)$ is the correlation coefficient between the earnings of the spouses. To reduce needless notation, we drop the cohort, c, and experience, x, identifiers. Let V denote the coefficient of variation in family earnings (i.e., $V = \sigma/\mu$, where μ stands for the mean of family earnings) and let V_j represent the coefficient of variation in j's earnings (i.e., $V_j = \sigma_j/\mu_j$), where $j = H$, W. Then the previous equation may be written

$$V^2 = (B_H)^2(V_H)^2 + (B_W)^2(V_W)^2 + 2rB_HB_WV_HV_W \tag{2}$$

where $B_H = \mu_H/\mu$ and $B_W = \mu_W/\mu$. So B_H and B_W are, respectively, each cell's average values of the shares of the husband's earnings and of the wife's earnings in family earnings. In Section 5 of this paper, expressions will be derived for $(V_H)^2$ and $(V_W)^2$ that involve the employment–population ratios of husbands and wives, respectively, but for now we concentrate on family earnings inequality, V^2 in the previous equation.

Descriptive statistics on all elements of Eq. (2) are contained in Table 2. The values of these variables describe all husband–wife households regardless of their labor market status. People who do not work in the market report zero earnings and such people are included in the statistics in Table 2. Thus the coefficient of variation of wives' earnings, V_W, is higher than that of husbands', V_H, principally because the employment–population ratio of wives has been much lower than that of husbands and, therefore, the frequency distribution of wives' earnings has a much higher spike at zero.

Table 2. Descriptive Statistics on Variables for 294 Experience-Cohort Cells.

	Variable	Mean	S.D.	Minimum	Maximum
1	V^2	0.601	0.253	0.234	1.600
2	$(V_H)^2$	0.820	0.406	0.256	2.598
3	$(V_W)^2$	1.982	0.727	0.832	5.252
4	V_H	0.882	0.206	0.506	1.612
5	V_W	1.386	0.244	0.912	2.292
6	$(B_H)^2$	0.580	0.087	0.408	0.765
7	$(B_W)^2$	0.061	0.028	0.016	0.131
8	B_H	0.760	0.057	0.639	0.875
9	B_W	0.240	0.057	0.125	0.361
10	r	0.057	0.065	−0.065	0.327
11	$\ln V$	−0.293	0.190	−0.727	0.235
12	$\ln V_H$	−0.151	0.225	−0.681	0.477
13	$\ln B_H$	−0.278	0.076	−0.449	−0.134

Similarly, r measures the correlation coefficient between husbands' earnings and wives' earnings in each cohort-experience cell among all husbands and wives, not merely among working husbands and working wives. r tends to be higher in recent cohorts principally because, in recent cohorts, the employment–population ratio of wives is much higher than in earlier cohorts.[7] When the wives' employment–population ratio is low, the relatively large number of zero values for wives' earnings inclines r to be low. As the wives' employment–population ratio rises and more women record positive earnings, so higher values of r are recorded. The frequency distribution of r is graphed in Fig. 2. Ninety-two percent of cells have values of r in the range of ± 0.15. With such values of r, an approximation of Eq. (2) is

$$V^2 = (B_H)^2(V_H)^2 + (B_W)^2(V_W)^2 \tag{3}$$

Confirmation that this is a good approximation is provided by values of

$$H = \text{abs}\left\{\frac{V^2 - (B_H^2)(V_H^2) - (B_W^2)(V_W^2)}{V^2}\right\}$$

where abs denotes the absolute value of the term in braces. H is simply a rearrangement of Eq. (2) that neglects the third term on the right-hand side. Low values of H suggest that Eq. (3) provides a good approximation to Eq. (2). H is graphed for five cohorts in Fig. 3. The only cases in which H exceeds 0.15 are for a few cells corresponding to young couples in the

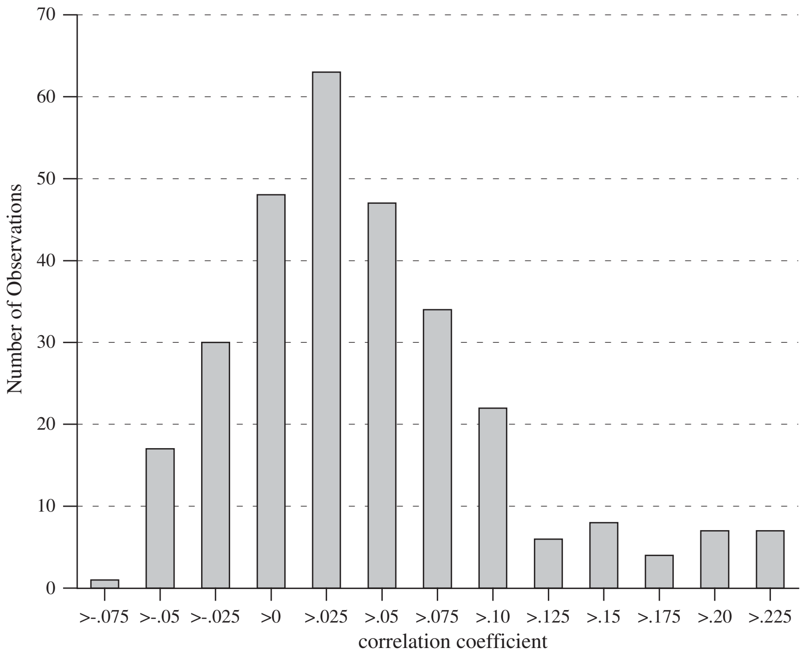


Fig. 2. Frequency Distribution of *r*.

most recent cohort. In most cases, H is less than 0.10. Hence, we shall proceed with the approximation given by Eq. (3).

After factoring $(B_H)^2(V_H)^2$ and taking logarithms, Eq. (3) can be rewritten as

$$\ln\ V = \ln\ V_H + \ln\ B_H + (0.5)\ln\left[1 + (B_W/B_H)^2(V_W/V_H)^2\right] \tag{4}$$

The left-hand side of this equation, the logarithm of the coefficient of variation in family earnings, is an indicator of the inequality in family earnings. The broad movements in $\ln V$ are similar to those of the Gini coefficient of family earnings, G, as is evident from the smoothed values of $\ln V$ and G shown in Figs. 4 and 5. The dispersion of family earnings rises sharply with experience for each cohort: for instance, following the data for the 1956–1960 cohort, according to both $\ln V$ and G, inequality at 37 years is about twice that observed 30 years earlier.[8] In addition, each cohort's family earnings inequality tends to lie above the previous cohort's inequality at any experience level: at 10 years of experience, the 1986–1990 cohort's values of $\ln V$ are 1.5 times and its values of G are 1.3 times those for the cohort entering the labor market 30 years earlier.[9]

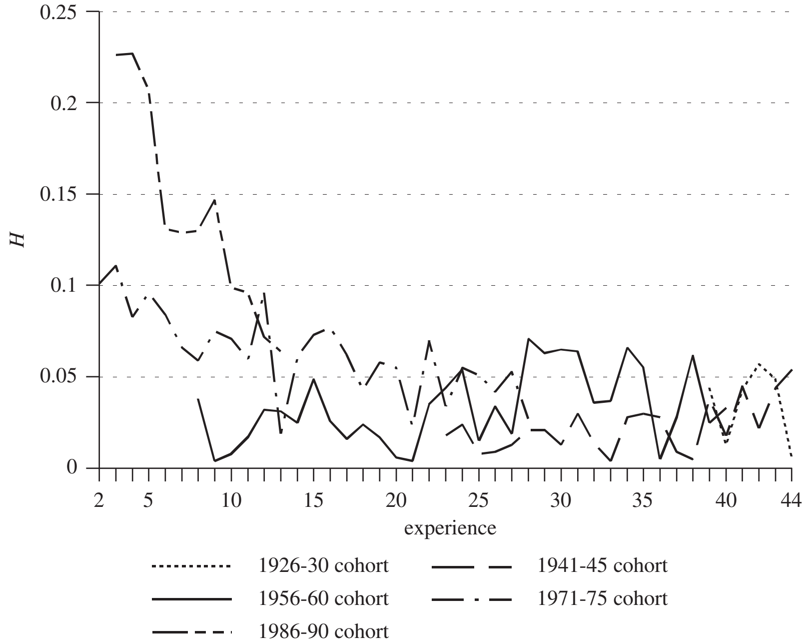


Fig. 3. Absolute Value of $\{V^2-[(B_H)^2(V_H)^2+(B_W)^2(V_W)^2]\}/V^2$ by Cohort and Experience.

Further approximations of Eq. (4) facilitate a better understanding of changes in family earnings inequality. First, given $B_H = \mu_H/\mu$, where μ_H is the mean of husbands' earnings and μ the mean of family earnings (including those not working for pay), if m_H and m_W denote, respectively, the mean earnings of husbands and the mean earnings of wives among those husbands and wives employed for pay and if E_H and E_W denote respectively the employment–population ratios of husbands and wives, then

$$\ln\ B_H = -\ln\left[1 + (m_W E_W)/(m_H E_H)\right] = -(m_W E_W)/(m_H E_H) \qquad (5)$$

The last step is an approximation and to assess the quality of this approximation form

$$-\ln\left[1 + (m_W E_W)/(m_H E_H)\right] + (m_W E_W)/(m_H E_H)$$

the frequency distribution of which is given in Fig. 6. All values are less than 0.10 and 95 percent are less than 0.06.

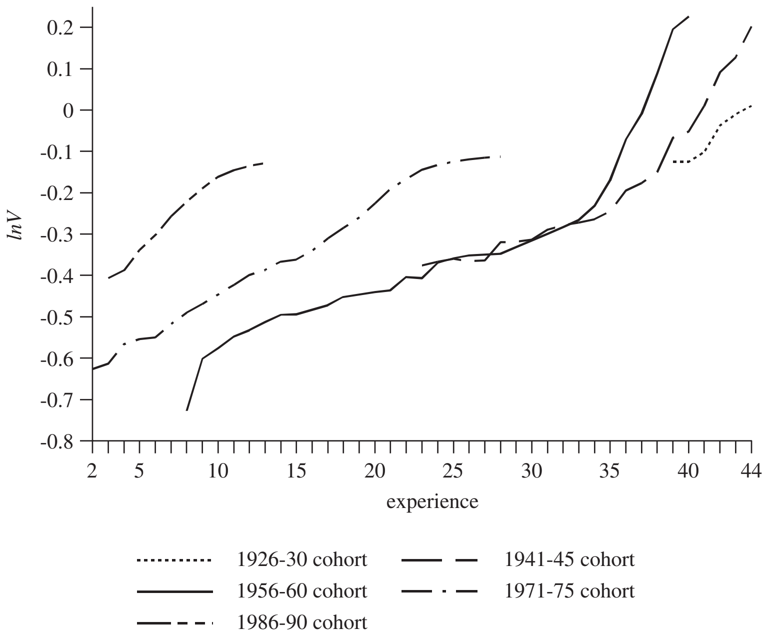


Fig. 4. Values of ln V by Cohort and Experience.

Now consider substituting $(B_W/B_H)^2(V_W/V_H)^2$ for $\ln[1+(B_W/B_H)^2(V_W/V_H)^2]$ in Eq. (4). To evaluate this, compute $\ln[1+(B_W/B_H)^2(V_W/V_H)^2]-(B_W/B_H)^2\,(V_W/V_H)^2$ whose values for all 294 cells are presented by the frequency distribution in Fig. 7. Over 90 percent of the cells have values between -0.075 and 0 with the mean being -0.032.

Replacing $\ln[1+(B_W/B_H)^2(V_W/V_H)^2]$ with $(B_W/B_H)^2(V_W/V_H)^2$ and using Eq. (5), Eq. (4) may be written approximately as

$$\ln\ V = \ln\ V_H - \frac{m_W E_W}{m_H E_H} + \frac{1}{2}\left[\frac{m_W E_W}{m_H E_H}\frac{V_W}{V_H}\right]^2 \tag{6}$$

Eq. (6) proposes a remarkably simple expression to describe movements in the dispersion of family earnings: approximately, the logarithm of the coefficient of variation of family earnings, ln V, equals the logarithm of the coefficient of variation of husbands' earnings, ln V_H, less a quadratic term involving $(m_W E_W)/(m_H E_H)$, the ratio of mean wives' earnings to mean husbands' earnings where these mean earnings are not conditional upon working for pay. The ratio of the coefficient of variation of wives' earnings

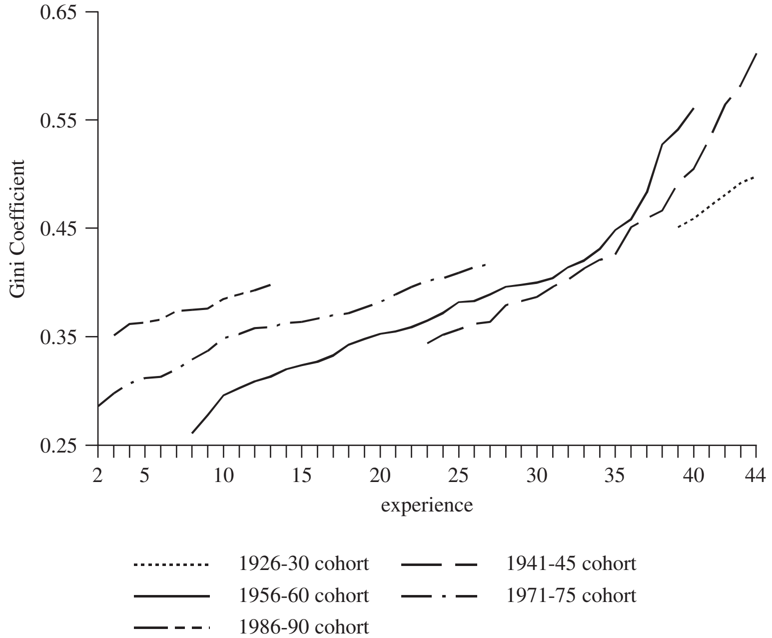


Fig. 5. Values of Gini Coefficients by Cohort and Experience.

to the coefficient of variation of husbands' earnings, V_W/V_H, also enters this expression, but the next step will involve treating this as parametric. Suppose $(V_W/V_H)^2 = k$ and, to move to the Gini coefficient as a more familiar indicator of inequality, suppose $\ln V = a_0 + a_1 G + u_1$ and $\ln V_H = b_0 + b_1 G_H + u_2$,[10] then Eq. (6) may be written as

$$G = \beta_0 + \beta_1 G_H + \beta_2 \frac{m_W E_W}{m_H E_H} + \beta_3 \left[\frac{m_W E_W}{m_H E_H}\right]^2 + u$$

where the stochastic term u incorporates the various approximations that have been made and the β's are parameters to be estimated. The above equation is Eq. (1), the expression introduced at the beginning of this section to describe variations in family earnings inequality in terms of variations in husbands' earnings inequality and in wives' relative employment and pay. Eq. (1) treats $(V_W/V_H)^2$ as parametric and incorporates it into the term β_3. Of course, $(V_W/V_H)^2$ is not fixed so the question is whether this assumption impedes an attempt to derive a useful compact description of the main empirical regularities in family earnings inequality.

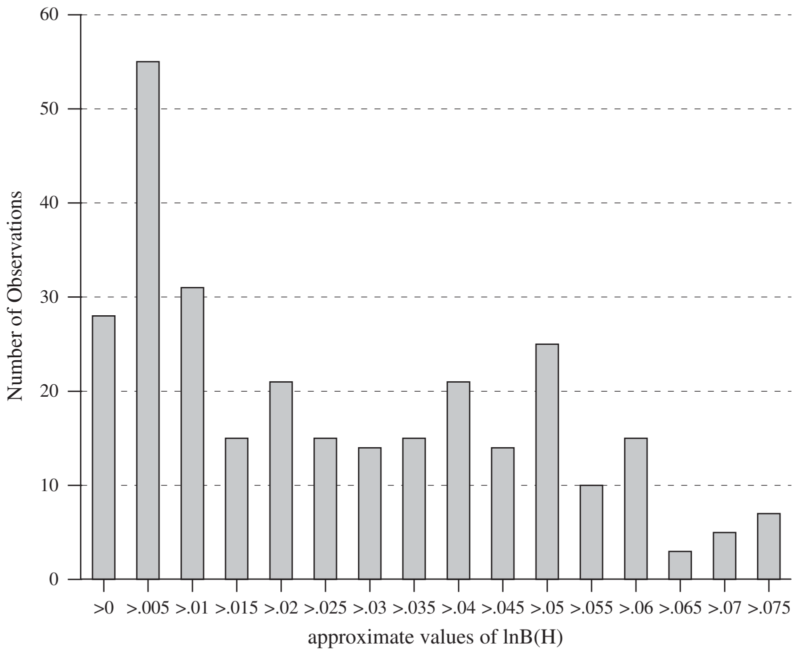


Fig. 6. $-\ln[1 + (m_W E_W)/(m_H E_H)] + (m_W E_\mathrm{W})/(m_H E_H)$.

The intuition behind Eq. (1) is that the inequality of family earnings will equal the inequality of husbands' earnings less an adjustment to take account of wives' relative contributions to family earnings. A major element of this adjustment reflects the fraction of people who do not work for pay, that is, the fraction of people with zero earnings. Holding constant the dispersion in husbands' earnings and husbands' employment, increases in the fraction of wives at work for pay (reductions in the fraction of wives with no earnings) will reduce earnings inequality among husband–wife families.[11] The effect of the increasing employment of wives in reducing family earnings inequality falls (if β_3 is positive as Eq. (6) suggests) as wives' employment–population ratio grows – this is the logic for the quadratic term in $(m_W E_W)/(m_H E_H)$. In addition, as wives' earnings rise relatively to husbands' earnings (i.e., as m_W rises relatively to m_H) so the importance of husbands' earnings inequality, G_H, in accounting for variations in family earnings inequality, G, falls.

How well does Eq. (1) describe the data on family earnings inequality? Column (1) of Table 3 presents the weighted least-squares of Eq. (1). In all,

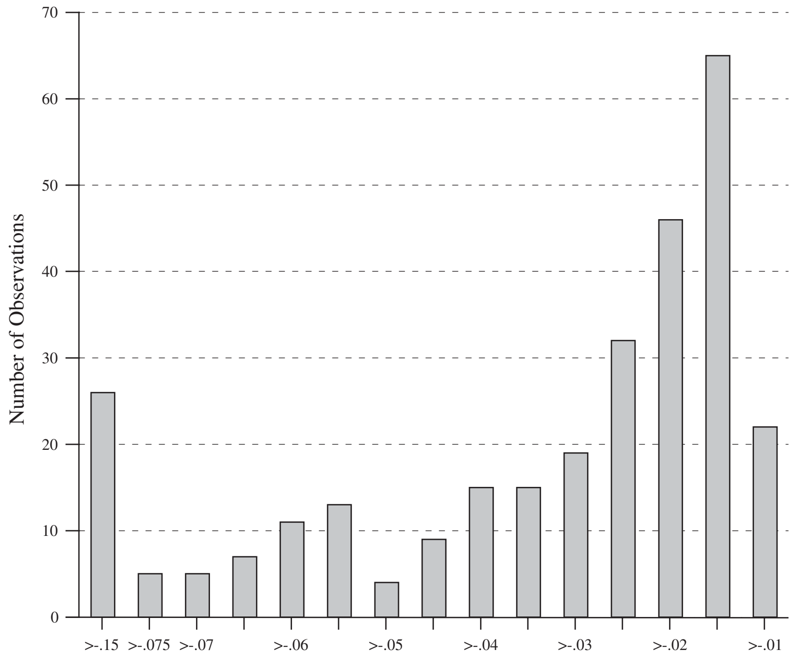


Fig. 7. $\ln\left[1+(B_w/B_H)^2(V_W/V_H)^2\right]-(B_w/B_H)^2(V_W/V_H)^2$. *Note:* Except for the first and last categories, each bar corresponds to an interval 0.005 wide. Thus, "$>-.075$" means "the interval from $-.075$ to $-.070$" and "$>-.07$" means "the interval from $-.070$ to $-.065$". The first category "$>-.15$" spans the interval from -0.15 to -0.75. The last category "$>-.01$" spans the interval from -0.01 to -0.005.

G_H, $(m_W E_W)/(m_H E_H)$, and $[(m_W E_W)/(m_H E_H)]^2$ remove 98 percent of the variance in G. In fact, as shown by the estimates in column (2) of Table 3, in describing the variations across these experience and cohort cells, G_H and $(m_W E_W)/(m_H E_H)$ alone remove over 97 percent of the variance in G. The estimates in columns (1) and (2) of Table 3 imply that increases in the wives' employment–population ratio or increases in the earnings of wives reduce family earnings inequality at observed values of the wives' employment and earnings.[12]

A component of the high correlations reported in columns (1) and (2) of Table 3 between family earnings inequality and the right-hand side variables arises because any measurement error in husbands' earnings will automatically be contained in family earnings. To determine the degree to which such measurement error inflates the R^2 values, consider fitting Eq. (1) to

Table 3. Describing Variations in the Gini Coefficient of Family Earnings.

	(1)	(2)	(3)	(4)
Constant	0.093	0.063	0.106	0.075
	(0.006)	(0.003)	(0.010)	(0.005)
G_H	0.798	0.784	0.773	0.759
	(0.007)	(0.007)	(0.012)	(0.012)
$(m_W E_W / m_H E_H)$	−0.292	−0.054	−0.303	−0.060
	(0.039)	(0.006)	(0.066)	(0.009)
$(m_W E_W / m_H E_H)^2$	0.360		0.369	
	(0.058)		(0.099)	
R^2	0.979	0.975	0.939	0.934
see	0.010	0.011	0.017	0.017

For both the estimates in columns (1) and (2) and for those in columns (3) and (4), the mean of G is 0.391 with a standard deviation of 0.067.

cells in which the left- and right-hand side variables are constructed from different underlying observations. Thus, in each cell, randomly allocate families into two groups: in one group, G, family earnings inequality, is formed and, in the second group, G_H and $(m_W E_W)/(m_H E_H)$ are formed. Now estimate Eq. (1) where the left- and right-hand side variables are constructed from different families in each cell. The results are contained in columns (3) and (4) of Table 3. The estimated coefficients in columns (3) and (4) are similar to those in columns (1) and (2) and the computed R^2 statistics for the equations in columns (3) and (4) are only a little below those in columns (1) and (2). The suggestion is that correlated measurement error is not a primary factor in accounting for the empirical performance of Eq. (1).[13]

4. THE RELATIVE IMPORTANCE OF HUSBANDS' EARNINGS AND WIVES' EMPLOYMENT

The estimates of Eq. (1) allow simulations of family earnings inequality to compute the relative importance for changes in family earnings inequality of increases in husbands' earnings inequality and of the increases in wives' relative earnings and employment. Simulations of Eq. (1) are hindered by the fact that, as Table 1 makes clear, observations on all years of experience are lacking for every cohort. This means a flexible and accurate description of the variables on the right-hand side of Eq. (1) outside their observed values is first required. That is, let $R_M = m_W/m_H$ and $R_E = E_W/E_H$ and express R_M and R_E, in addition to G_H, as general functions of experience and cohort as follows: $R_M = f(x,c)$, $R_E = g(x,c)$, and $G_H = h(x,c)$. Numerical expressions for $f(x,c)$, $g(x,c)$, and $h(x,c)$ allow Eq. (1) to be written as

$$\hat{G} = \hat{\beta}'_0 + \hat{\beta}'_1\, \hat{h}(x,c) + \hat{\beta}'_2[\hat{f}(x,c)\hat{g}(x,c)] + \hat{\beta}'_3[\hat{f}(x,c)\hat{g}(x,c)]^2 \tag{7}$$

where the circumflexes represent estimated values. Now family earnings inequality may be simulated for different values of years of experience and cohort. For instance, denote the cohort entering the labor market in 1941–1945, cohort 4, by c_4 and denote the cohort entering the labor market in 1971–1975, cohort 10, by c_{10}. Once the forms of $f(x,c)$, $g(x,c)$, and $h(x,c)$ are determined, we may ask what family earnings inequality would have looked like at each year of experience for cohort 10 if, say, relative earnings R_M and R_E had taken on their actual values for cohort 10 but husbands' earnings inequality, G_H, had remained at its cohort 4 values:

$$\begin{aligned}\hat{G}[x; G_H(4), R_M(10), R_E(10)] &= \hat{\beta}'_0 + \hat{\beta}'_1\, \hat{h}(x,c_4) + \hat{\beta}'_2[\hat{f}(x,c_{10})\hat{g}(x,c_{10})]\\ &\quad + \hat{\beta}'_3[\hat{f}(x,c_{10})\hat{g}(x,c_{10})]^2\end{aligned}$$

In this expression, relative earnings and employment assume the values associated with cohort 10 while the inequality of husbands' earnings assumes the values associated with cohort 4. Therefore, the simulated values of family earnings inequality describe the impact of the change in the relative earnings and relative employment of wives holding constant husbands' earnings inequality. The impact on family earnings inequality of changes in husbands' earnings inequality holding constant wives' relative earnings and employment can be assessed using

$$\begin{aligned}\hat{G}[x; G_H(10), R_M(4), R_E(4)] &= \hat{\beta}'_0 + \hat{\beta}'_1\, \hat{h}(x,c_{10}) + \hat{\beta}'_2\left[\hat{f}(x,c_4)\hat{g}(x,c_4)\right]\\ &\quad + \hat{\beta}'_3\left[\hat{f}(x,c_4)\hat{g}(x,c_4)\right]^2\end{aligned}$$

so that $R_M = m_W/m_H$ and $R_E = E_W/E_H$ assume the values of an earlier cohort, cohort 4, while G_H assumes values associated with a recent cohort, cohort 10.

To implement these counterfactuals, we require an accurate description of the experience and cohort patterns in wives' relative employment and earnings and husbands' earnings inequality; that is, we need to fit $R_M = f(x,c)$, $R_E = g(x,c)$, and $G_H = h(x,c)$. After investigating the implications of alternative specifications, we specified the variations in R_M, R_E, and G_H by means of a fully interacted quintic function of years of experience and a linear function of cohort. The weighted least-squares estimates are presented in Table 4. Ninety-two percent of the variations in wives' relative earnings and relative employment and almost 99 percent of the variations in husbands' earnings inequality are removed by these combinations of years of experience and cohort. Empirical estimates of $f(x,c)$, $g(x,c)$, and $h(x,c)$ have now been determined.

Consider Fig. 8 where the wide lines describe actual observations on G for cohorts 4 and 10. The actual observations for cohort 4 are from 23 to 40 years of experience and those for cohort 10 are from 2 to 28 years of experience. The continuous dotted line plots the values implied for G for the 1941–1945 cohort from the estimates of $R_M = f(x,c)$, $R_E = g(x,c)$, and $G_H = h(x,c)$ in Table 4 and the continuous solid line plots the values implied for G for the 1971–1975 cohort from the estimates of $R_M = f(x,c)$, $R_E = g(x,c)$, and $G_H = h(x,c)$ in Table 4. Clearly, within the sample years, the implied series for G does little more than smooth the raw data. The values for G for the 1941–1945 cohort are always below those for the 1971–1975 cohort: in the years that overlap, the Gini coefficient is an average of 0.05 higher for the later cohort. To what extent is that increase attributable to greater earnings inequality among husbands and to what extent is it affected by the growth in wives' employment and earnings?

Table 4. Weighted Least-Squares Estimates of Earnings Inequality, Relative Employment, and Relative Earnings as Functions of Experience and Cohort (Estimated Standard Errors in Parentheses).

	Left-Hand Side Variable		
	G_H	$R_E = E_W/E_H$	$R_M = m_W/m_H$
Constant	0.170	1.029	0.535
x	−0.030	−0.255	−0.159
x^2	$0.469(10)^{-2}$	0.024	0.015
x^3	$-0.220(10)^{-3}$	$-0.999(10)^{-3}$	$-0.618(10)^{-3}$
x^4	$0.442(10)^{-5}$	$0.189(10)^{-4}$	$0.119(10)^{-4}$
x^5	$-0.320(10)^{-7}$	$-0.137(10)^{-6}$	$-0.895(10)^{-7}$
c	0.016	$-0.557(10)^{-2}$	$0.822(10)^{-3}$
cx	$0.239(10)^{-2}$	0.017	0.015
cx^2	$-0.282(10)^{-3}$	$-0.167(10)^{-2}$	$-0.154(10)^{-2}$
cx^3	$0.116(10)^{-4}$	$0.730(10)^{-4}$	$0.683(10)^{-4}$
cx^4	$-0.209(10)^{-6}$	$-0.148(10)^{-5}$	$-0.144(10)^{-5}$
cx^5	$0.159(10)^{-8}$	$0.114(10)^{-7}$	$0.115(10)^{-7}$
R^2	0.985	0.922	0.924

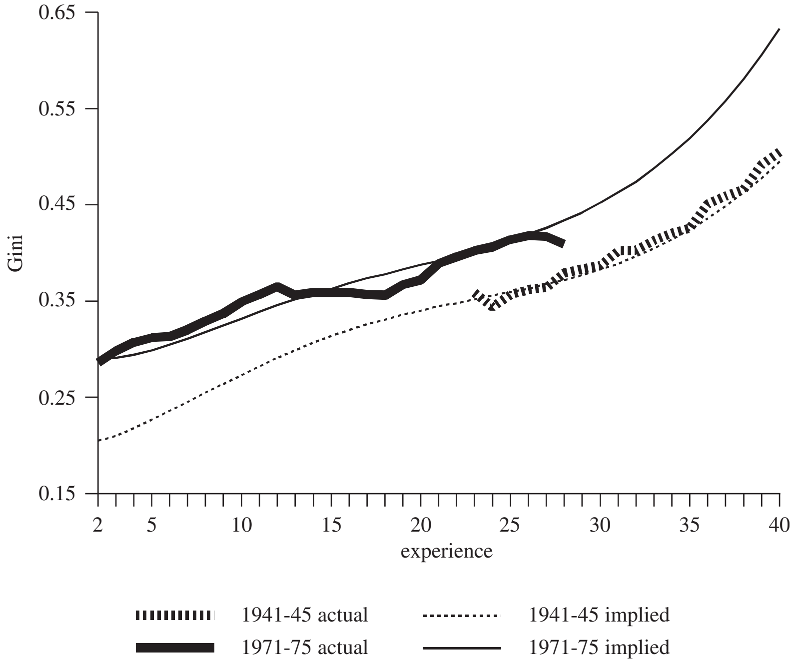


Fig. 8. Gini Coefficients by Experience: Actual and Implied for the 1941–1945 and 1971–1975 Cohorts. *Note:* The wide lines plot the actual observations on the Gini coefficients for the 1941–1945 and 1971–1975 cohorts. The continuous dotted line plots the values implied for the Gini coefficients for the 1941–1945 cohort from the estimates of $R_M = f(x,c)$, $R_E = g(x,c)$, and $G_H = h(x,c)$ in Table 4 and the continuous solid line plots the values implied for the Gini coefficients for the 1971–1975 cohort from the estimates of $R_M = f(x,c)$, $R_E = g(x,c)$, and $G_H = h(x,c)$ in Table 4.

The implied values of G for cohorts 1941–1945 and 1971–1975 in Fig. 8 are reproduced in Fig. 9. These are the dotted and continuous lines, respectively, in Fig. 9. Fig. 9 also presents some simulations of G corresponding to different assumptions about husbands' earnings inequality, G_H, and about the relative employment and earnings of wives. Thus the series denoted $GH(4),RE(10),Rm(10)$ plots the Gini coefficients when wives' relative employment and relative earnings assume their implied values for the 10th cohort whereas husbands' earnings inequality assumes its implied values for the 4th cohort. This series is the lowest of the four lines graphed in Fig. 9. This means that family earnings inequality is least when husbands' earnings inequality takes on its values for the early cohort and when wives'

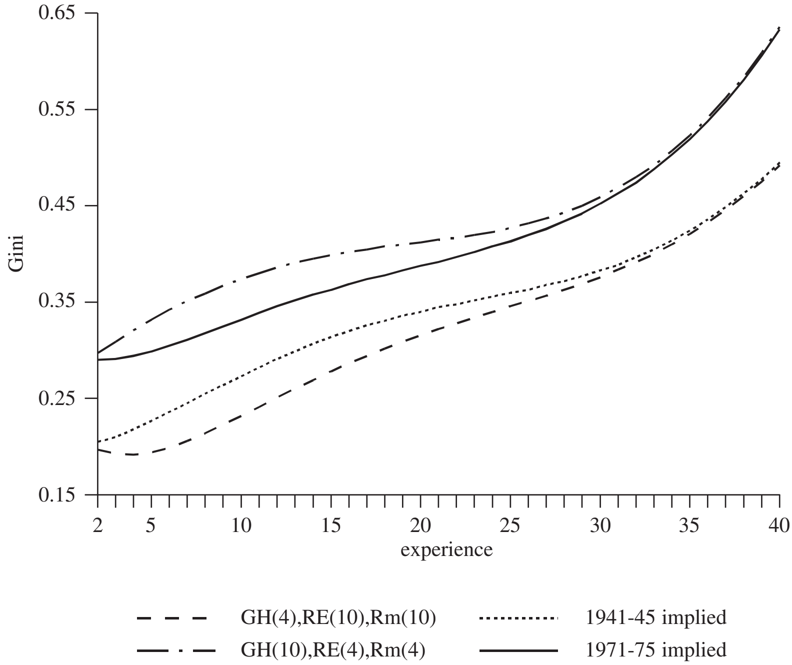


Fig. 9. Gini Coefficients by Experience: Implied and Extrapolated for the 1941–1945 and 1971–1975 Cohorts. *Note:* The dotted line ("1941–45 implied") plots the values implied for the Gini coefficients for the 1941–1945 cohort from the estimates of $R_M = f(x,c)$, $R_E = g(x,c)$, and $G_H = h(x,c)$ in Table 4 and the continuous line ("1971–75 implied") plots the values implied for the Gini coefficients for the 1971–1975 cohort from the estimates of $R_M = f(x,c)$, $R_E = g(x,c)$, and $G_H = h(x,c)$ in Table 4. The other two lines graph simulations of the Gini coefficients corresponding to different assumptions about husbands' earnings inequality and relative employment and relative earnings. The series denoted $GH(4),RE(10),Rm(10)$ plots the values of the Gini coefficients when husbands' earnings inequality assumes its implied values for the 4th cohort and relative employment and relative earnings assume their implied values for the 10th cohort. The series denoted $GH(10),RE(4),Rm(4)$ plots the values of the Gini coefficients when husbands' earnings inequality assumes its implied values for the 10th cohort and relative employment and relative earnings assume their implied values for the 4th cohort.

employment and earnings take on their values for the later cohort. Expressed differently, relatively low husbands' earnings inequality and relatively high wives' employment and earnings contribute to lower family earnings inequality. The differences are most evident at younger years of experience.

The greatest values of inequality in family earnings in Fig. 9 correspond to the line $GH(10),RE(4),Rm(4)$ where wives' relative employment and relative earnings assume their implied values for the 4th cohort whereas husbands' earnings inequality assumes its implied values for the 10th cohort. In other words, higher values of husbands' earnings inequality and lower values of wives' employment and earnings result in greater family earnings inequality. The fact that the simulated series $GH(10),RE(4),Rm(4)$ is closer to the implied series for cohort 10 ("1971–75 implied") and the fact that the simulated series $GH(4),RE(10),Rm(10)$ is closer to the implied series for cohort 4 ("1941–45 implied") indicates that variations in family earnings inequality are more closely tied to the movements in husbands' earnings inequality than to variations in wives' relative employment and earnings.

A summary of the contributions of the changes in G_H and $(m_W E_W)/(m_H E_H)$ to the change in G from the 1941–1945 to the 1971–1975 cohorts by experience is given in Table 5. Over this period, G increased by 0.079 for the youngest couples (as shown in the first line of Table 5). G_H increased by more than this, namely, by 0.100. The relative increase in wives' employment and earnings partially offset the increase in G_H and contributed −0.021 to the change in G. At each experience level, the increase in wives' employment and earnings offset the increases in husbands' earnings inequality and induced family earnings inequality to rise by less than it would otherwise have done. The relative contribution of wives' employment and earnings is especially marked at between 6 and 20 years of experience. Although the greater employment and earnings of wives has attenuated the growth in family earnings inequality, these movements are more than offset by increases in husbands' earnings inequality.[14]

Table 5. Changes in the Gini Coefficient from the 1941–1945 to the 1971–1975 Cohorts by Experience.

Experience	Change in G	Change in G_H	Change in $(m_W E_W)/(m_H E_H)$
< 6	0.079	0.100	−0.021
6–10	0.064	0.104	−0.040
11–15	0.053	0.092	−0.039
16–20	0.048	0.077	−0.029
21–25	0.050	0.068	−0.018
26–30	0.062	0.072	−0.010
31–35	0.084	0.088	−0.004

5. EARNINGS INEQUALITY FOR HUSBANDS AND WIVES SEPARATELY

We turn now from describing family earnings inequality to the inequality of husbands' earnings and wives' earnings separately with particular attention to the way in which market work patterns affect inferences about changes in inequality. Eq. (2) above relates a measure of the dispersion in family earnings, namely, V^2, the square of the coefficient of variation of family earnings, to the same dispersion indicator for husbands' earnings and wives' earnings, respectively, $(V_H)^2$ and $(V_W)^2$. This section will establish a useful relationship for $(V_H)^2$ and $(V_W)^2$ separately that will allow us to assess the impact on earnings inequality of changes in the fraction of individuals (husbands and wives, in turn) in market employment.

Consider the variance in earnings for, say, husbands when some husbands work for pay and some do not. If σ_H is the standard deviation of husbands' earnings (including both those with positive earnings and those with zero earnings), then

$$\sigma_H^2 = E_H s_H^2 + E_H(1 - E_H)m_H^2 \tag{8}$$

where E_H is the employment–population ratio of husbands, s_H^2 is the variance of earnings among husbands employed for pay, and m_H^2 is the square of mean earnings of those husbands employed.[15] Most research on earnings inequality focuses on s_H or another metric of inequality among those with positive earnings only. However, Eq. (8) can be rearranged to derive a relationship between earnings inequality among workers only and earnings inequality among all people.

Converting to a scale-invariant measure of dispersion, let $V_H (= \sigma_H/\mu_H)$ denote the coefficient of variation of earnings among all husbands and $V_H^e = (s_H/m_H)$ the coefficient of variation of earnings among those husbands with positive earnings. Then Eq. (8) may be rewritten with the coefficient of variation as the indicator of inequality:

$$(V_H)^2 = E_H(V_H^e)^2(q_H)^2 + E_H(1 - E_H)(q_H)^2$$

where $q_H = m_H/\mu_H$, the ratio of mean earnings among workers only to the mean earnings of all workers. Given $\mu_H = E_H m_H$, $q_H = (E_H)^{-1}$ and a convenient expression is arrived at that relates earnings inequality among all husbands to two variables, namely, earnings inequality among those husbands working for pay and the fraction of husbands at work for pay:

$$(V_H)^2 = (E_H)^{-1}(V_H^e)^2 + (E_H)^{-1} - 1 \tag{9}$$

Necessarily, $(V_H)^2$ exceeds $(V_H{}^e)^2$: the dispersion of earnings including those with zero earnings exceeds the dispersion of earnings excluding nonworkers. Eq. (9) suggests that, neglecting selection, increases in employment reduce earnings inequality among all people with the size of the effect increasing with earnings inequality among workers.[16]

Of course, an equation analogous to Eq. (9) holds for wives:

$$(V_W)^2 = (E_W)^{-1}(V_W^e)^2 + (E_W)^{-1} - 1 \tag{10}$$

and, if Eqs. (9) and (10) are substituted into Eq. (2), an expression is derived that relates family earnings inequality to earnings inequality among husbands and among wives, separately, and to their employment–population ratios:

$$\begin{aligned} V^2 = {} & (B_H)^2[(E_H)^{-1}(V_H^e)^2 + (E_H)^{-1} - 1] \\ & + (B_W)^2[(E_W)^{-1}(V_W^e)^2 + (E_W)^{-1} - 1] + \psi \end{aligned} \tag{11}$$

where $\psi = 2rB_HB_WV_HV_W$. In this way, the analysis in this section of earnings inequality for husbands and wives separately relates closely to the analysis of family earnings inequality in the previous section. Using the coefficient of variation as the measure of dispersion, Eq. (11) indicates that the impact of changes in, say, husbands' earnings inequality among workers on family earnings inequality depends on the fraction of husbands at work, E_H, and on the importance of husbands' earnings in family earnings, B_H.

Eq. (9) is an exact relationship that holds for the square of the coefficient of variation as a measure of inequality. In fact, other measures of earnings inequality among all people also exhibit the implied relationship with the employment–population ratio and earnings inequality among workers.[17] Thus consider Fig. 10 which plots the relationship for the 1961–1965 cohort among G_H (the Gini coefficient for husbands' earnings among all husbands both those working for pay and those not), $G_H{}^e$ (the Gini coefficient for husbands' earnings only among those husbands working for pay), and E_H (employment–population ratio of husbands). G_H and $G_H{}^e$ rise with experience although their values are closer at earlier years when E_H is higher than it is in later years when E_H is lower.

As for wives, let G_W be the Gini coefficient for wives' earnings among all wives (those at market work and those not), $G_W{}^e$ the Gini coefficient for wives' earnings only among those wives working for pay, and E_W the employment–population ratio of wives. For the 1961–1965 cohort, as shown in Fig. 11, the employment–population ratio for these wives falls with experience, then rises, and then falls again. Earnings inequality among workers as measured by $G_W{}^e$ tends to follow the opposite pattern rising with experience

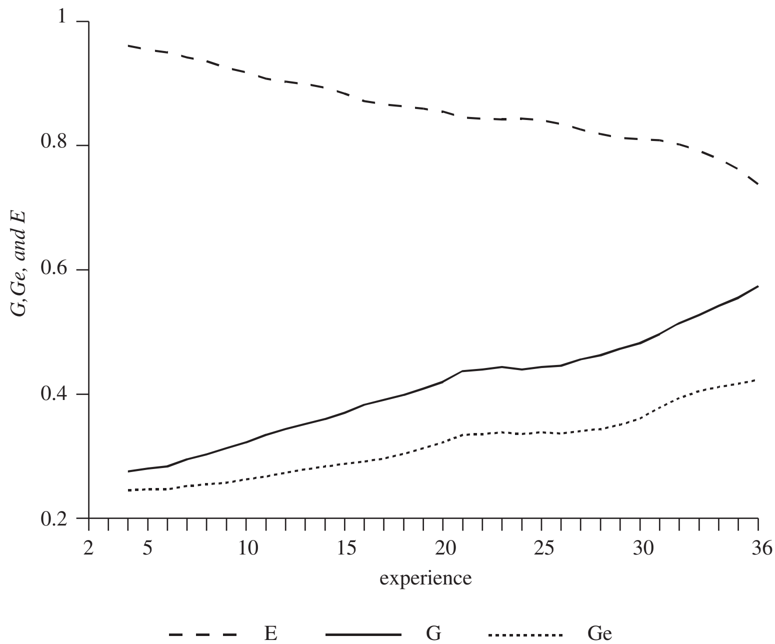


Fig. 10. Husbands: Values of G_H, G_H^e, and E_H by Experience for the 1961–1965 Cohort.

and then falling. Unlike husbands, for at least this cohort, earnings inequality among wives does not rise monotonically with experience. Earnings inequality among all wives – both those working for pay and those not working for pay – as measured by G_W is greatest when wives' employment–population ratio is least and earnings inequality is least when wives' employment–population ratio is highest. Of course, changes in the employment–population ratio will also affect the dispersion of earnings among workers only; that is, $\partial(G_H^e)/\partial E_H$ and $\partial(G_W^e)/\partial E_W$ are unlikely to be zero. Whatever their signs, the suggestion in Figs. 10 and 11 is that the relationship between the employment–population ratio and earnings inequality among all individuals is negative.

The preceding two graphs depict the relationships for a single cohort with respect to experience. Turning to the behavior of earnings inequality over time, the strong upward trends in the employment–population ratio of wives would suggest the possibility of quite different inferences about the movement of earnings inequality over time depending on whether the zero earnings of nonworking wives are included in the computation of inequality.

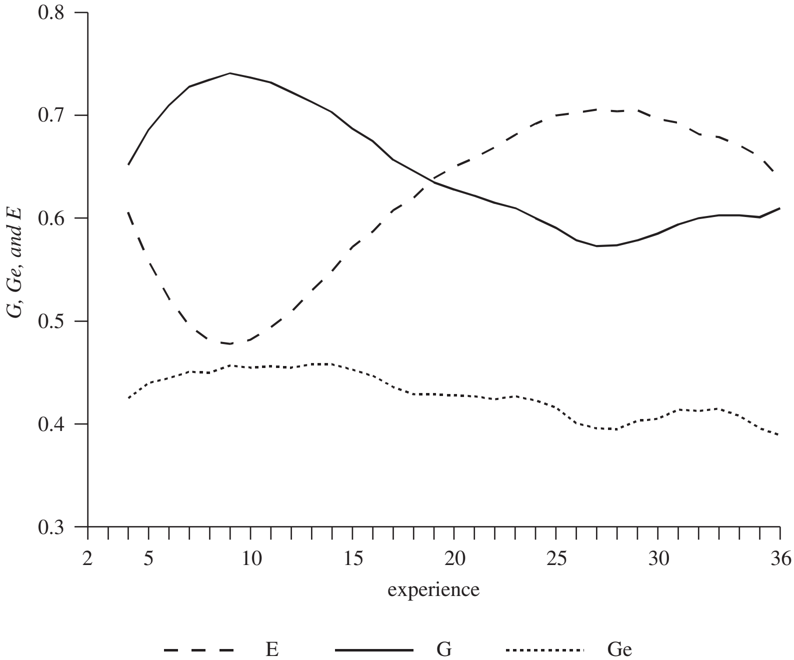


Fig. 11. Wives: Values of G_W , $G_W{}^e$, and E_W by Experience for the 1961–1965 Cohort.

Fig. 12 graphs these two measures of earnings inequality, G_W and $G_W{}^e$, as a function of cohort for wives with 10, 20, and 30 years of experience. The G_W series are shown as thin lines while the $G_W{}^e$ series are depicted with thicker lines. For wives with 20 or 30 years of experience, the series on $G_W{}^e$ suggests small changes in earnings inequality over time. The series on $G_W{}^e$ for wives with 10 years of experience (i.e., younger women) falls with cohort before turning upwards. The corresponding series on G_W all fall steeply over time indicating that the rising employment of women more than offsets trends in earnings inequality among workers. If we were to measure earnings inequality by taking account of the zero earnings of nonworkers, we would conclude that earnings inequality among all wives fell over time.

Should we be concerned with earnings inequality among workers only or earnings inequality among all people, workers and nonworkers? Some would respond that the answer depends on the reasons why people do not work. That is, perhaps some women seek and fail to obtain work at prevailing rates of pay because they are rationed out of jobs. Suppose these

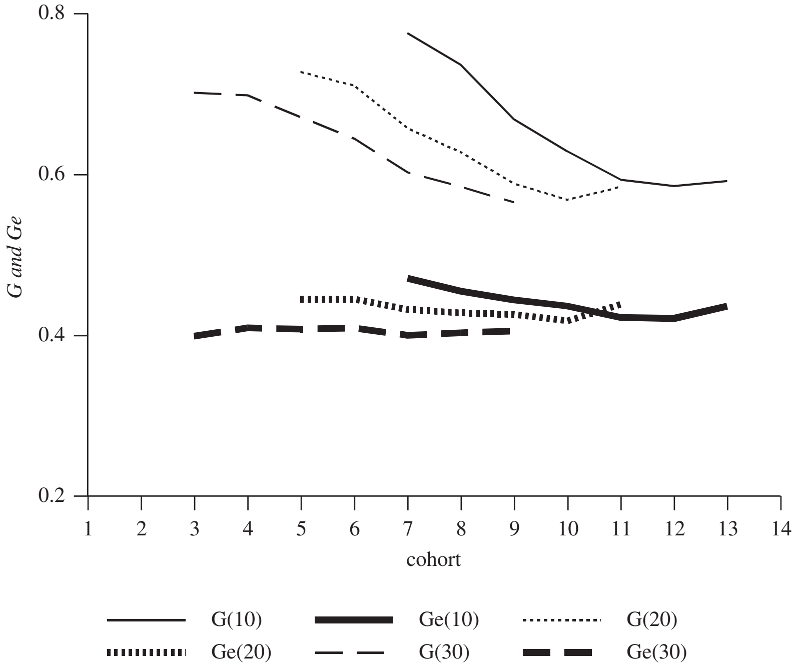


Fig. 12. Values of G and G^e for Wives by Cohort for 10, 20, and 30 Years of Experience.

women are those recorded as unemployed. Then a compromise between choosing between G_W and $G_W{}^e$ as a measure of inequality is to count not all nonworking wives in the computation of earnings inequality but simply those nonworking wives who report being unemployed. Let $G_W{}^U$ be the Gini coefficient of earnings inequality that includes all wives who are working for pay and all wives who are recorded as unemployed. The measurement of this series $G_W{}^U$ across cohorts is shown in Fig. 13 where it is compared with $G_W{}^e$. Naturally each series on $G_W{}^U$ lies above the corresponding series on $G_W{}^e$ (that omits the unemployed). However, because the unemployed represent a small fraction of those wives not employed for pay, each $G_W{}^U$ series is close to $G_W{}^e$. (Note the scale of Fig. 13's vertical axis.)

The contrast between the cross-cohort patterns in G_W and $G_W{}^e$ for wives – one measure of earnings inequality that includes those not working for pay and the other restricted to wives working for pay – raises the general question of whether additional differences in labor supply behavior affect inferences

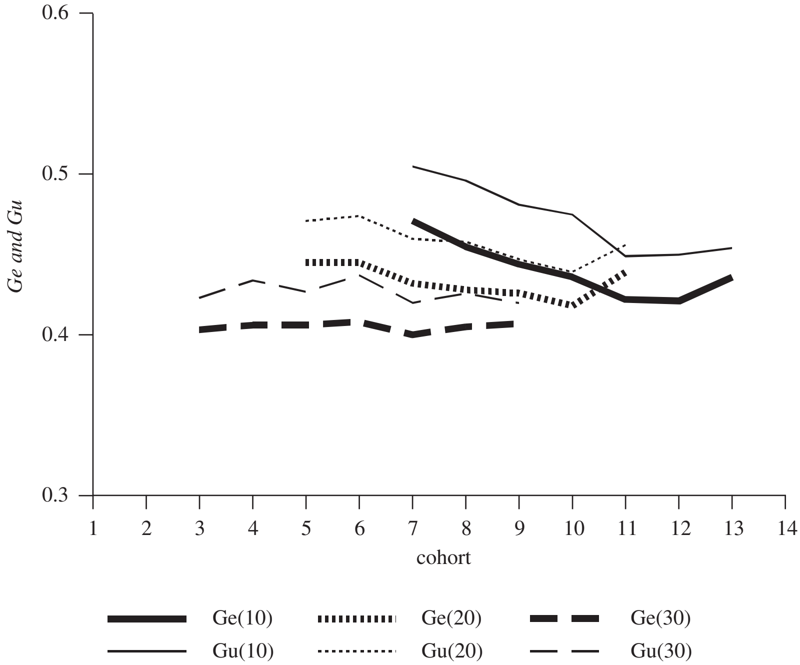


Fig. 13. Values of G^e and G^U for Wives by Cohort for 10, 20, and 30 Years of Experience.

about changes in annual earnings inequality. To this effect, in addition to earnings inequality among (a) all wives (including those not working for pay) and (b) all working wives, construct measures of earnings inequality among those (c) working at least 700 annual hours, (d) working at least 1,400 h per year, and (e) working at least 1,800 h per year. For each of these groups, form the Gini coefficient of earnings inequality for wives and husbands separately. Across experience and cohort cells, the mean values and standard deviations of these Gini coefficients of annual earnings inequality among wives are reported in column (1) and those for husbands in column (3) of Table 6. Naturally, dispersion falls when inequality is measured over an increasingly selective group of wives. To determine whether the cross-cohort movements in earnings inequality vary across these different groups of wives – as suggested by the contrast between G_W and $G_W{}^e$ in Fig. 12 – regress each Gini coefficient on experience fixed effects and a linear cohort trend. The weighted least-squares estimates of the coefficients on the cohort trend for wives are reported in column (2) and those for husbands in column (4) of Table 6.[18]

Table 6. Mean Values, Standard Deviations, and Estimates of the Cohort Trend in Earnings Inequality for Wives and Husbands by the Extent of their Market Work.

		Wives		Husbands	
		(1)	(2)	(3)	(4)
		Mean (S.D.)	Trend estimates	Mean (S.D.)	Trend estimates
1	All (including nonworkers)	0.653	−0.0277	0.439	0.0185
		(0.076)	(0.0007)	(0.084)	(0.0004)
2	All workers	0.377	0.0035	0.319	0.0183
		(0.027)	(0.0005)	(0.042)	(0.0005)
3	Working $\geq$ 700 h	0.375	0.0036	0.309	0.0178
		(0.023)	(0.0004)	(0.299)	(0.0005)
4	Working $\geq$ 1,400 h	0.292	0.0132	0.299	0.0173
		(0.032)	(0.0005)	(0.041)	(0.0005)
5	Working $\geq$ 1,800 h	0.278	0.0154	0.296	0.0173
		(0.033)	(0.0005)	(0.042)	(0.0005)

In columns (1) and (3), standard deviations are in parentheses beneath mean values. The entries in columns (2) and (4) under "trend estimates" are the estimated values of weighted least-squares coefficients attached to a linear cohort trend in regressions in which the Gini coefficient of earnings inequality is regressed on experience fixed effects and a linear cohort trend. In columns (2) and (4), estimated standard errors are in parentheses beneath estimated coefficients. For all but two cases, there are 294 experience-cohort cells. For wives working at least 1,400 annual hours and wives working at least 1,800 annual hours, cells with less than 300 underlying observations on these wives were omitted and this results in 279 experience-cohort cells for wives in line 4 and 263 experience-cohort cells in line 5.

According to the estimates in column (2), while there are negative trends in annual earnings inequality for all wives (including nonworkers), there are strong positive trends in inequality for full-time working wives (those working at least 1,800 h per year).[19] As increasingly selective work hours criteria are applied to the wives so the trends in earnings inequality tend to become more positive. The cohort trend variable implies that, for wives working at least 1,800 h, the Gini coefficient increases by 0.03 over 10 years whereas for all working wives the cohort trend implies an increase in the Gini coefficient of about one-fifth of this. Clearly quite dissimilar movements in inequality are implied for wives who work different market hours.[20]

The corresponding estimates on the cohort trend for husbands in column (4) of Table 6 (again allowing for experience fixed effects) are positive and, in most instances, do not vary across the different types of husbands as much as those of wives. Whereas inequality trends are negative for all wives (including nonworkers), they are positive for all husbands (including non-workers). Of course, while wives exhibit strong positive trends in employment–population ratios, husbands have small declines and this accounts for the difference in these trends.[21]

One possible interpretation of these findings for wives – strong positive trends in earnings inequality for wives who work long hours and much smaller positive trends in earnings inequality for all working wives – is that wives' work hours have changed systematically at different points in the earnings distribution. Suppose the strong positive trends in inequality among those wives working full-time (at least 1,800 h) approximates the increase in the dispersion of hourly earnings facing all workers. So hourly earnings have increased more at high wage levels than at low wage levels. This increase in hourly earnings inequality will be mapped into corresponding changes in annual earnings inequality if work hours do not change systematically at different points in the earnings distribution. However, suppose wives' labor supply curves are not only positively sloped with respect to hourly wages but also they are more wage-elastic at low wage levels than at high wages. This is compatible with income effects being greater at higher wage and hours levels. In this case, even though hourly earnings have increased more at high wage levels than at low wage levels, the positive work hours response is greater for wives with low wages and low earnings than for those with high wages and high earnings. In effect, greater increases in work hours for those with low hourly earnings than the increases in work hours for those at high hourly earnings has the effect of making the trends in annual earnings inequality smaller than the trends in hourly earnings inequality.[22] So have wives' work hours changed systematically at different points in the earnings distribution?

To answer this, examine changes in median hours worked at different points of the annual earnings distribution. That is, for wives in each cohort-experience cell, construct the median annual hours worked, h^M, at each π percentile of the annual earnings y_W distribution: $h^M \; y_W = \pi$ from the 5th percentile to the 95th percentile. How has $h^M y_W = \pi$ changed over time? To answer this, regress by weighted least-squares $h^M y_W = \pi$ on experience fixed effects and on a linear cohort trend.

The estimated coefficients on the linear cohort trend for wives for each fifth percentile are graphed by the solid line in Fig. 14.[23] All estimated trend coefficients are positive and the largest trend increases in median hours worked are for wives between the 15th and 25th earnings percentile. Above the 25th earnings percentile, the trend coefficients fall in value into the upper tail of the earnings distribution. The estimate of 93.82 for wives at the 25th earnings percentile implies that, over a 10-year period, annual hours worked

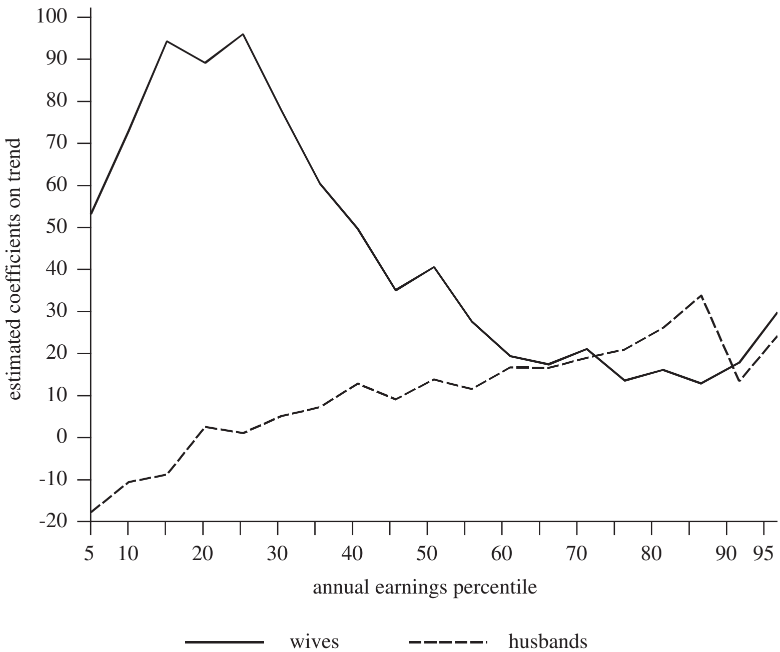


Fig. 14. Estimated Cohort Trends in Median Hours Worked at Different Earnings Percentiles.

increased by 188 h. By contrast, the estimate of 12.98 for wives at the 75th earnings percentile implies that, over a 10-year period, annual hours worked for these women by merely 26 h. So, among working wives, hours worked have increased more for those with low earnings than for those with high earnings. This differential trend has caused earnings inequality among all working wives to increase less than among full-time working wives.[24]

This analysis is undertaken also for husbands. The trends in median work hours for husbands at different levels of annual earnings are given by the dashed line in Fig. 14. These trends tend to be smaller for husbands than for wives. The estimated trends for husbands at the 35th earnings percentile or less are not significantly different from zero (applying conventional criteria) whereas most of the trends at higher earnings levels are greater than zero. Nevertheless these are small trends: the largest trend is estimated for husbands at the 85th earnings percentile and the coefficient of 32.83 implies a 10 year change of 66 annual hours which is less than three percent of the median (of 2,309).

Hence, the labor supply dimension is a very important aspect of the changes in earnings inequality for wives but a much smaller component of the changes in earnings inequality for husbands. Quite different inferences about changes in earnings inequality – decreases in inequality among all wives (including nonworkers), mild increases for all working wives, and substantial increases for full-time working wives – apply to wives according to their work behavior.

If inferences about changes in earnings inequality over time are affected by the labor supply behavior of individuals, are inferences about life cycle earnings inequality similarly sensitive? Figs. 15 and 16 address this question by plotting the estimated fixed experience effects when Gini coefficients of annual

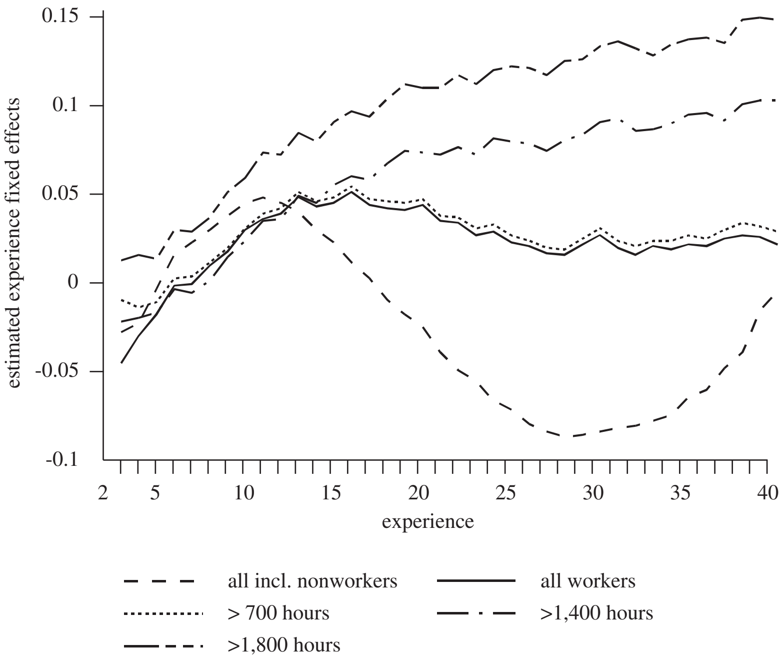


Fig. 15. Annual Earnings Inequality for Wives Against Years since Leaving School by the Extent of their Work Behavior. *Note:* This figure graphs the estimated experience fixed effects from regression equations in which the Gini indicator of annual earnings inequality of wives is related to experience fixed effects and cohort fixed effects. The reference group is 2 years of experience. Five regression equations are estimated, each equation describing a different group of wives depending on their work behavior.

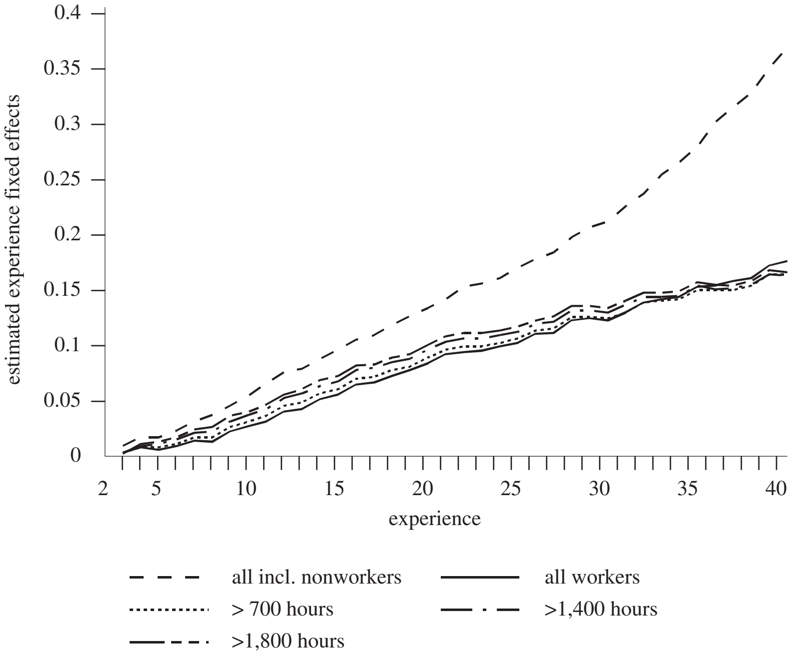


Fig. 16. Annual Earnings Inequality for Husbands Against Years since Leaving School by the Extent of their Work Behavior. *Note:* This figure graphs the estimated experience fixed effects from regression equations in which the Gini indicator of annual earnings inequality for husbands is related to experience fixed effects and cohort fixed effects. The reference group is 2 years of experience. Five regression equations are estimated, each equation describing a different group of husbands depending on their work behavior.

earnings inequality are related to experience and cohort fixed effects.[25] Fig. 15 pertains to wives and Fig. 16 to husbands. For wives, the experience fixed effects increase with years of experience for full-time workers but, as wives with fewer work hours are incorporated, so this pattern is attenuated. When annual earnings inequality among all wives including nonworkers is analyzed, the experience effects reflect how employment probabilities change with experience. For husbands (Fig. 16), earnings inequality increases with experience as it does for full-time working wives. However, the magnitude of the increases is greater for husbands than it is for wives. So, inferences about life cycle patterns of annual earnings inequality are also affected by the labor supply behavior of the individuals whose earnings are being described.

6. CONCLUSIONS

This paper has analyzed earnings inequality in husband and wife families over the life cycle and over time by assembling pseudo-panel data from over 30 years of the Current Population Survey. The role of market work in understanding differences in earnings inequality has received special attention.

Family earnings inequality has increased over calendar time and increases with years since leaving school. A compact and accurate description relates the Gini coefficient of family earnings to variations in husbands' earnings inequality and to wives' relative employment–population ratio and mean earnings. Changes in family earnings inequality over time have been driven principally by changes in husbands' earnings inequality. This supports the research effort into understanding the growth in wage inequality among men. Of secondary importance, family earnings inequality has been affected by the growth of wives' earnings: if the employment of wives had not increased, the cohort growth in family earnings inequality would have been greater. The growing propensity of married women to work for pay has mitigated the increase in family earnings inequality. The growth in the correlation between the earnings of husbands and that of wives has also contributed to the growth in family earnings inequality but its direct role is smaller than the other factors listed in this paragraph.

The second part of the paper examines the link between labor supply and earnings inequality among husbands and wives separately. Measures of earnings inequality covering people with different degrees of attachment to the labor market have been presented. Inferences about the extent and changes in earnings inequality are sensitive to alternative labor supply definitions especially in the case of wives. Trends in earnings inequality among wives are negative if the population consists of all wives, both workers and nonworkers. Alternatively, among full-time workers, earnings inequality for wives has trended upwards. Market work hours for wives have increased most for those at the lower part of the earnings distribution. These links among wives between changes in earnings inequality and labor supply decisions warrant more research.

Finally, note that the research in this paper is directed toward understanding movements in the income inequality of families and in the income inequality of husbands and wives. A related but distinct issue concerns the inequality of consumption across and within families. Movements in consumption inequality do not mirror those in earnings inequality[26] and we conjecture that a decomposition analogous to that used in this paper may help understand how movements in the inequality of family consumption relate to movements in the consumption inequality of husbands and wives.

NOTES

1. Previous research tends to place a heavier emphasis on cross-section data to draw inferences about the effects of the growth of market employment of wives on family earnings inequality. See, for instance, the recent careful research of Daly and Valletta (2006) who reach conclusions fully compatible with those in this paper. Earlier work includes that of Smith (1979) using the 1960 and 1970 Censuses of Population and Lehrer and Nerlove (1981) who use data from the National Survey of Family Growth. Cancian, Danziger, and Gottschalk (1993) draw on March CPS data from 1968 to 1988. Cancian and Reed (1998) argue that the distribution of family income in 1979 and 1989 would have shown greater dispersion without the earnings of wives. A similar conclusion was reached by Gronau (1982) for Israel. Hyslop's (2001) research is confined to husbands and wives both of whom work for pay and he covers the 6 years from 1979 to 1985.
2. See Congressional Budget Office (2005).
3. This companion paper uses a different decomposition of family earnings inequality from the one used here and it does not take up the issue of the extent to which inferences about changes in earnings inequality for wives and husbands separately are affected by differences in market work behavior.
4. In the research reported in this paper, experience is defined as the minimum of (1) current age minus years of schooling minus 6 and (2) age minus 17. Other definitions were investigated for earlier cohorts with minimal consequences for the empirical regularities reported here.
5. For a very small number of families, the sum of interest, dividends, and rent is negative. This arises because some report negative rent (i.e, the payment, not the receipt, of rent). In these instances, the sum of dividends, interest, and rent was set to zero. This happens so infrequently that nothing of any consequence follows from this.
6. To address the changing top-coding of income in the CPS, we use an imputation procedure to generate a measure of earnings for people whose earnings are above the top-coded level. Information on the earnings structure of people just below the top-coded earnings level is used to infer earnings of those people above the top-coded level. The appendix of Pencavel (2006) describes this procedure. In addition, many of the results reported in this paper on income inequality were confirmed for measures that do not use information on the earnings of all people such as ratios of earnings at different percentiles. The principal results in this paper are independent of the particular measure of income inequality used and are not affected by the issue of top-coding. Basically this is because, in most years, only a very small fraction of husband–wife families have their earnings top-coded. Nevertheless, for some measures of inequality in some years, top-coding may have profound effects on inferences about earnings inequality (see Burkhauser, Butler, Feng, & Houtenville, 2004).
7. The simple correlation coefficient between r and the wives' employment-population ratio across these 294 cells is 0.614. An associated reason for a rising

value of r is the growth in assortative mating by skill, that is, the growing propensity for well-educated men and well-educated women to marry (see Pencavel, 1998).

8. For this cohort, $\ln V$ at 8 years is about -0.70 and at 37 years it is about 0 so $\ln[V(1956\text{–}1960, 37)/\ln V(1956\text{–}1960, 8)] = 0.70$ and $\exp(0.70) = 2.014$. Similarly, G at 37 years is almost twice that at 7 years.

9. At 10 years of experience, $\ln V$ for the 1986–1990 cohort is -0.141 and for the 1956–1960 cohort it is -0.575 so $\ln[V(1986\text{–}1990, 10)/\ln V(1956\text{–}1960, 10)] = 0.434$ and $\exp(0.434) = 1.543$. The experience and cohort effects graphed in Figs. 4 and 5 are similar to those for another indicator of inequality, namely, the variance of the logarithm of income. The latter is a common measure of inequality among workers. However, the logarithmic transform is less appealing when incomes are zero and when some adjustment to zero incomes is required to make the measure more meaningful.

10. The correlation coefficients between G and $\ln V$ and between G_H and $\ln V_H$ are both 0.93. Figs. 4 and 5 show how similar are the variations in G to those in $\ln V$.

11. Strictly, whether increases in the fraction of wives at work for pay will reduce family earnings inequality depends on the values of the β_2 and β_3 coefficients. We shall see shortly, however, that the estimated values of these coefficients justify the statement in the text for almost all the observed values of $(m_W E_W)/(m_H E_H)$ in the cohort-experience cells.

12. $\partial G/\partial E_W$ and $\partial G/\partial m_W$ are negative for almost all the values of the right-hand side variables. It is only at some high values of $(m_W E_W)$ for which this is not the case. This is the consequence of the excessive curvature placed on the relationship by the quadratic term. Indeed, when the quadratic expressions in Eq. (1) are replaced with a more flexible functional form (one in which a series of dummy variables indicate various categories of $(m_W E_W)/(m_H E_H)$, $\partial G/\partial E_W$, and $\partial G/\partial m_W$ are always negative. When the Gini coefficients are replaced by other indicators of inequality (one is the variance of the logarithm of income and the other is the logarithm of the coefficient of variation of income), again a very large fraction (well over 90 percent) of the variation in family earnings dispersion is removed by a least-squares linear combination of right-hand side variables.

13. Note that the right-hand side variables of Eq. (1) involve not only the employment-population ratios of wives to husbands but also their relative earnings. One may inquire into the role of relative earnings by fitting

$$G = \gamma_0 = \gamma_1 G_H + \gamma_2(E_W/E_H) + \gamma_3(E_W/E_H)^2 + \varepsilon \tag{1'}$$

which omits relative earnings from the second two right-hand side variables. The weighted least squares estimates (with estimated standard errors in parentheses) of the γ parameters are as follows: $\gamma_1 = 0.799(0.006)$, $\gamma_2 = 0.108(0.046)$, and $\gamma_3 = -0.126(0.033)$ with an R^2 of 0.985. While the explanatory power of the right-hand side variables of Eq. (1′) is as great as those in column (1) of Table 3, the signs of the coefficient estimates of γ_2 and γ_3 are opposite those estimated for β_2 and β_3 above. Hence, the presence of relative earnings makes a meaningful difference to the estimates.

14. The greater sensitivity of changes in husbands' earnings inequality for family earnings inequality may be inferred directly from the estimates of Eq. (1) in Table 3. Using the estimates in column (1) and evaluating the estimates at sample mean

values, the elasticity of G with respect to G_H is 0.880 whereas the elasticity of G with respect to $(m_W E_W)/(m_H E_H)$ is -0.045.

15. Let $D = 1$ if $y_H > 0$ and $D = 0$ if $y_H = 0$. Then $\sigma_H^2 = \mathscr{E}(y_H^2) - [\mathscr{E}(y_H)]^2 = \mathscr{E}[(Dy_H)^2] - [\mathscr{E}(Dy_H)]^2 = p\mathscr{E}(y_H^2|D = 1) - p^2[\mathscr{E}(y_H|D = 1)]^2$ where $p = \text{prob}(D = 1)$. Adding and subtracting $p[\mathscr{E}(y_H|D = 1)]^2$ and recognizing that $s_H^2 = \mathscr{E}(y_H|D = 1) - [\mathscr{E}(y_H|D = 1)]^2$ and that $m_H^2 = [\mathscr{E}(y_H|D = 1)]^2$, Eq. (8) in the text is derived.

16. In other words, $\partial[(V_H)^2]/\partial(E_H) = -(E_H)^2[(V_H{}^e)^2+1] < 0$.

17. The correlation coefficient between $(V_H)^2$ and G_H is 0.91 and that between $(V_W)^2$ and G_W is 0.89.

18. Equations were also fitted with cohort fixed effects instead of a linear cohort trend. In most cases the linear cohort trend provides a good approximation to the cohort patterns. The exception is for wives working at least 1,400 h and for wives working at least 1,800 h. For these groups, the positive trend is much stronger for recent cohorts than is implied by a linear trend.

19. The Gini coefficients for women (not just wives) reported in Table 1 of Katz and Autor's (1999) survey describe the weekly earnings of those working 35 or more hours per week and working at least 40 weeks per year. So their measurements resemble the series here of those wives working at least 1,400 annual hours.

20. The cohort trend variable increases by the value of unity for each successive cohort. Because each cohort is defined in 5-year intervals, a 10-year change means an increase in the value of the cohort trend of two. According to the estimates in Table 6, the estimated coefficient on the cohort trend for wives working at least 1,800 h is 0.0154 so over 10 years the Gini coefficient increases by 0.0308 (0.0154×2). By contrast, the estimated coefficient on the cohort trend for all working wives is 0.0035 so over 10 years the Gini coefficient increases by 0.0070(0.0035×2). The estimated coefficient on the cohort trend for all working wives is a little over one-fifth (precisely, 0.227) of the estimated coefficient for full-time working wives.

21. The estimates in columns (2) and (4) were also computed for two other indicators of earnings inequality: the ratio of the 90th percentile to the 10th percentile of earnings and the logarithm of the coefficient of variation of earnings. (However, when constructing earnings inequality among all people including zero earners, for many cells, the 10th earnings percentile corresponds to 0 earnings so this is not a useful indicator of inequality for all people.) The general cohort trends for these two other indicators of inequality were similar to those reported for the Gini coefficient in Table 6 with the exception that a negative trend was estimated for the ratio of the 90th to the 10th earnings percentile for all working wives and for wives working at least 700 h.

22. To be clear, write annual earnings, y, as the product of hourly earnings, w, and annual work hours, h:$y = wh$ and, for convenience, use the ratio of earnings at the 90th percentile to earnings at the 10th percentile as the indicator of inequality for cohort c at experience x:$y(9, c, x)/y(1, c, x)$. Form the change over cohorts in the logs of this measure of annual earnings inequality $\Delta\ln[y(9)/y(1)] = \Delta\ln w(9) - \Delta\ln w(1) + \Delta\ln h(9) - \Delta\ln h(1)$. Suppose the labor supply functions for workers at the 90th and 10th percentiles may be written as $\ln h(9) = \eta.\ \ln w(9)$ and $\ln h(1) = \mu.\ \ln w(1)$ so

$$\Delta\ln\ [y(9)/y(1)] = (1 + \eta)\Delta\ln w(9) - (1 + \mu)\Delta\ln w(1).$$

Even if $\Delta\ln w(9) - \Delta\ln w(1) > 0$ (i.e., hourly earnings at the 90th percentile of annual earnings increase more than hourly earnings at the 10th percentile), annual earnings inequality will increase less than hourly earnings inequality if the elasticity of labor supply of low wage workers, μ, is sufficiently larger than the elasticity of labor supply of high wage workers, η (assuming $\Delta \ln w(1) > 0$).

23. Again, the cohort trend variable takes the value of unity for the 1926–1930 cohort, two for the 1931–1935 cohort, and so on up to the value of 14 for the 1991–1995 cohort, so an increase in the value of this trend by unity corresponds to a 5-year change.

24. By conventional statistical criteria, each of these estimated coefficients for wives is significantly greater than zero.

25. In each figure, five weighted regression equations are fitted. One sample consists of all wives (husbands) including nonworkers, another involves all working wives (husbands), another includes wives (husbands) working at least 700 annual hours, a fourth sample is of wives (husbands) working at least 1,400 annual hours, and the last sample are wives (husbands) working at least 1,800 h. In each case, the dependent variable is the Gini measure of annual earnings inequality. The reference experience level is 2 years of experience.

26. See, for instance, the information for Britain contained in Lise and Seitz (2004).

ACKNOWLEDGMENTS

I am indebted to Niny Khor for her scrupulous research assistance. Support from the Smith-Richardson Foundation is gratefully acknowledged.

REFERENCES

Burkhauser, R. V., Butler, J. S., Feng, S., & Houtenville, A. J. (2004). Long term trends in earnings inequality: What the CPS can tell us. *Economics Letters*, *82*(2), 295–299.

Cancian, M., Danziger, S., & Gottschalk, P. (1993). Working wives and family income inequality among married couples. In: S. Danziger & P. Gottschalk (Eds), *Uneven tides: Rising inequality in America* (pp. 195–221). New York: Russell Sage Foundation.

Cancian, M., & Reed, D. (1998). Assessing the effects of wives' earnings on family income inequality. *Review of Economics and Statistics*, *80*(1), 73–79.

Congressional Budget Office (2005). United States, Historical Effective Tax Rates: 1979 to 2003, December, http://www.cbo.gov/ftpdoc.cfm?index = 7000&type = 1

Daly, M. C., & Valletta, R. G. (2006). Inequality and poverty in the United States: The effects of rising dispersion on men's earnings and changing family behavior. *Economica*, *73*(Feb.), 75–98.

Gronau, R. (1982). Inequality of family income: Do wives' earnings matter. *Population and Development Review*, *8*(Suppl.), 119–136.

Hyslop, D. R. (2001). Rising U.S. earnings inequality and family labor supply: The covariance structure of intrafamily earnings. *American Economic Review*, *91*(4), 755–777.

Katz, L. F., & Autor, D. H. (1999). Changes in the wage structure and earnings inequality. In: O. Ashenfelter & D. Card (Eds), *Handbook of labor economics*, (Vol. 3A, pp. 1463–1555). Amsterdam: Elsevier.

Lehrer, E., & Nerlove, M. (1981). The impact of female work on family income distribution: Black–white differentials. *Review of Income and Wealth*, *27*(4), 423–431.

Lise, J., & Seitz, S. (2004). *Consumption inequality and intra-household allocations*. Queen's University, Department of Economics, December.

Pencavel, J. (1998). Assortative mating by schooling and the work behavior of wives and husbands. *American Economic Review (Papers and Proceedings)*, *88*(2), 326–329.

Pencavel, J. (2006). A life cycle perspective on changes in earnings inequality among married men and women. *Review of Economics and Statistics*, *88*(2), 232–242.

Smith, J. P. (1979). The distribution of family earnings. *Journal of Political Economy*, *87*(5, Part 2), S163–S192.

EMPLOYMENT DYNAMICS AND BUSINESS RELOCATION: NEW EVIDENCE FROM THE NATIONAL ESTABLISHMENT TIME SERIES

David Neumark, Junfu Zhang and Brandon Wall

ABSTRACT

We analyze and assess new evidence on employment dynamics from a new data source – the National Establishment Time Series (NETS). The NETS offers advantages over existing data sources for studying employment dynamics, including tracking business establishment relocations that can contribute to job creation or destruction on a regional level. Our primary purpose in this paper is to assess the reliability of the NETS data along a number of dimensions, and we conclude that it is a reliable data source although not without limitations. We also illustrate the usefulness of the NETS data by reporting, for California, a full decomposition of employment change into its six constituent processes, including job creation and destruction stemming from business relocation, which has figured prominently in policy debates but on which there has been no systematic evidence.

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 39–83

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26002-3

1. INTRODUCTION

Employment growth is a major goal of economic policy at both the national and regional levels. Changes in employment are driven by job creation and job destruction, which in turn are made up of six dynamic processes including the birth, death, growth, contraction, and relocation of business establishments. This "demographic" characterization of business establishment and employment dynamics emphasizes that employment change in an economy is the net result of six influences – three that create jobs (births, expansions, and in-migration) and three that destroy jobs (deaths, contractions, and out-migration). Ultimately, we need to understand all six of these dynamic processes to characterize employment change in an economy, and to identify the job creation and destruction processes on which it might be the most productive for policymakers to focus in encouraging employment growth.[1] Moreover, the fact that employment change is the net result of potentially large gross changes – for example, overall expansion of jobs at existing establishments and overall contraction of jobs at other existing establishments – suggests that what often appear as relatively moderate overall changes in employment over time may mask potentially volatile gross job flows. This implies that relatively small changes in any of the gross flows can lead to sharp changes in net job growth.

But tracking a large population of business establishments across time and space, including births, deaths, and relocations, is difficult and costly, and thus data have not been available with which to fully capture the underlying processes of employment dynamics. Primarily for this reason, although the importance of understanding the job creation–destruction process has long been widely recognized (e.g., Schumpeter, 1942, Chapter 7), systematic empirical research on this topic did not start until quite recently as researchers began to develop appropriate data sources. However, this research has continued to face significant limitations imposed by the data.

In this paper we help to introduce a new data source – the National Establishment Time Series (NETS) – which we believe is the first data set that permits a full decomposition of the sources of employment change in regions of the U.S. economy, and which offers other advantages relative to the existing data sources. Our primary emphasis is on assessing the reliability of the NETS data along numerous dimensions. We provide this detailed assessment because the NETS data should prove useful to researchers in many fields, but it is a new data source of unknown quality, and we know that there are inherent difficulties in tracking business establishments – especially new establishments and those that relocate. In general, we conclude that the

NETS is a reliable data source although not without limitations, and we provide some guidance on its use. We also illustrate the usefulness of the NETS by using data for the entire state of California to fully decompose employment change into its six constituent processes, documenting the importance of each in contributing to employment change and its volatility. Because a principal advantage of the NETS data is the tracking of business establishment relocations, we focus on the role of relocation in employment dynamics. This analysis contributes hard evidence to a policy debate over business relocation that has been entirely speculative and reliant upon anecdotal evidence.

2. THE NETS DATABASE

2.1. Overview

The NETS database is a new longitudinal file based on recent Dun and Bradstreet (D&B) data. It is a long-term project of Walls & Associates in conjunction with D&B. We currently have access to an extract of this data set that covers all business establishments that were ever located in California between 1989 and 2002, and their respective parent headquarters (regardless of location).[2,3]

The version of the NETS database that we use begins with 14 cross-sectional files of the full Data Universal Numbering System (DUNS) Marketing Information (DMI) file for each year from 1990 through 2003, each of which covers the previous year. From here on, we refer to the year covered by the data, i.e., 1989–2002 for the full sample period. The primary purpose of D&B's data collection effort is to provide information on businesses to the business community, in order to enhance their decision making by constructing a set of "predictive indicators" (e.g., the D&B Rating and PayDex scores), and for marketing purposes. The DMI file for each year is constructed from an ongoing effort to capture each business establishment in the United States in each year (including nonprofits and the public sector). The DMI file is based on a multi-layered process incorporating many data sources.

D&B strives to identify and assemble information on all business establishments, through a massive data collection effort, including over 100 million telephone calls from four calling centers each year, as well as obtaining information from legal and court filings, the newspapers and electronic news services, public utilities, all U.S. Secretaries of State, government registries and licensing data, payment and collections information, company filings and news reports, and the U.S. Postal Service.[4] Particular efforts are devoted to identifying the births and deaths of establishments. For every establishment identified, D&B assigns a DUNS number as a means of tracking the establishment. It should be pointed out that since around 1990 the DUNS has been adopted by many government agencies in the United States and also internationally has become the standard means of tracking businesses.[5]

Although the goal of D&B is not to collect and organize data for scholarly research, it does have an incentive to ensure the accuracy of its data, because inaccuracies would hurt D&B's business and might even result in lawsuits. D&B has established a sophisticated quality control system and engages in extensive quality and consistency checks.[6] Thus, the data in each cross-section should provide high quality "snapshots" of business establishments.

Walls & Associates entered into a collaboration with D&B with a very different purpose – namely, to provide a dynamic view of the U.S. economy using the data from the D&B archives (Walls & Associates, 2003). This requires linking the D&B cross-sections into a longitudinal file that tracks every establishment from its birth, through any physical moves it may make, capturing any changes of ownership, and recording the establishment's death if it occurs. This is a multi-stage process, the most important steps of which include merging the data files, imputing data when data are not reported,[7] eliminating duplicate records, merging records on establishments for which the DUNS number changes (which happens occasionally) yet which appear to cover the same establishment, and identifying establishment relocations.

The resulting NETS database includes the following variables that are of particular importance to this research: current business name; current establishment location (street addresses and phone numbers); FIPS county codes in each year; type of location (single location, headquarters, branch) in each year; employment in each year; and four-digit standard industrial classification (SIC) codes that are also disaggregated to an eight-digit level by D&B.[8] Because the NETS reports establishment location for every year, it is possible to infer moves through changes of address.

One highly desirable feature of the NETS database is that it covers essentially all establishments. This reflects the fact that it is designed to capture the universe rather than a sample of establishments. Over the sample period of 1989–2002, the database includes information each year on between 1.2 and 1.8 million establishments in California providing about 15–18 million jobs. In total, more than 3.5 million establishments are covered in our extract of the NETS database.

As the preceding discussion indicates, the data construction effort – including both the cross-sectional files and the longitudinal linking that tracks establishments over time – is a massive and complicated one. For this reason, we have undertaken a good deal of investigation to document and examine the quality of the NETS data in order to assess their reliability, potential limitations, and how these limitations might affect results of various analyses.

2.2. Classification of Relocations, Births, and Deaths

A central question for using the NETS data to calculate decompositions of employment change into its constituent processes is how D&B distinguishes whether an establishment at a new location previously existed elsewhere – and hence will be labeled a relocation in the longitudinal file – or instead is a new establishment. Clearly, the correct classification of relocations is critically important in estimating the contributions of births, deaths, and business relocations to job creation and destruction.

In thinking about classifying relocations, a key point is the centrality of the DUNS number to D&B's data system. It is the DUNS number, after all, that allows D&B to attach information on credit histories and marketing databases, which is what its clients value. Consequently, DUNS numbers are unique, and D&B never recycles numbers. If an establishment closes, its DUNS number goes into an "out of business or inactive" file, where it remains permanently unless that business reopens. Each time D&B updates establishment information, it attempts to contact the establishment based on the previous location information on the establishment. Moves can be indicated in a number of ways. Frequently, there is a forwarding address or telephone number, or continuing email contact that allows D&B to identify a new location. (In addition, business establishments sometimes notify D&B of their move.) Most importantly, any establishment that cannot be contacted at the previous year's address or telephone number also goes into the "out of business or inactive" file, and before any "new" establishment can be given a DUNS number, it must be checked against this file, and if there are indications of a match, follow-up investigation is undertaken. For example, if an establishment belonging to a multi-unit firm cannot be found, D&B contacts the headquarters to determine whether a relocation has occurred. In any case in which D&B finds that the establishment previously existed elsewhere, it assigns its existing DUNS number. Finally, if a new establishment is identified whose characteristics do not match those of an existing establishment, D&B contacts the establishment to verify its start date, and assigns a new DUNS number. With these procedures, the longitudinal file should correctly identify relocations of establishments and distinguish them from births of new establishments (and deaths of others), although, of course, one cannot rule out the possibility of occasional errors of a move being classified as a death in one location and a birth in another, which would lead to an undercount of relocating establishments.[9] As a consequence, in our assessment of the NETS data we focus in part on accurate identification of relocating establishments.

An establishment relocation in the NETS data is identified by street address and zip code changes from one year to another. Both establishments that moved out of California and establishments that moved into California are included in the database, so we are able to track cross-state relocation. However, there are some limits to what this form of relocation can tell us about the dynamics of employment change, as other types of changes in employment might be viewed as sharing features of establishment relocation, or reflecting the same forces that drive relocation. First, if a California company sets up an establishment in another state, that establishment does not show up in our extract. That is, we can study establishments that ''move out'' but not those that ''branch out.''[10] The latter should not be regarded as equivalent to the former because branching out does not necessarily occur at the cost of creating an additional establishment within the state. Second, the NETS database only tracks physical establishment relocation. There are several other types of relocation that it does not capture by design. For instance, it does not allow us to determine when specific jobs or positions are shifted between two discrete locations of the same firm. This type of relocation, which also constitutes a relocation of jobs between establishments, will be observed in our data set as employment expansion at one establishment and contraction at another. Also, relocations that involve the consolidation of activities at two or more locations into a single location will often be missed, and will be reflected in one establishment growing and another closing. Despite these caveats, the NETS database enables empirical research that represents a significant step toward understanding the role of business relocation in job creation and destruction, especially given that the policy debate frequently refers to physical relocations of business establishments.

2.3. Advantages of the NETS

The NETS is not the first data set with which researchers can study employment dynamics, nor is this the first project to attempt to study this question using data from D&B. However, other data sources present important limitations in studying employment dynamics, and previous work using earlier D&B files has been criticized. (See Table 1 for a summary of alternative data sources and findings for the United States, and Neumark, Zhang, and Wall (2005a) for more detailed discussion of past research using these data sources).[11]

Compared with alternative data sources for the United States, the NETS has a few key advantages. First, from the perspective of fully characterizing employment change, the NETS captures business relocation. Unlike most other data sources described in Table 1, the NETS database tracks business address changes and identifies business moves over time within the entire country. As discussed below, this is important because business retention and attraction issues are often at the center of policy debates at the state (and local) level.

In addition, the NETS offers significant advantages in actually carrying out research on employment dynamics (or other topics). Access to the alternative data sources collected by federal and state government agencies is highly restricted because of confidentiality reasons, and hence requires a long and complex process of application and approval. As a practical matter, this has deterred many researchers from pursuing research with these data, and has clearly made it difficult to do research in a timely manner. In addition, again because of confidentiality, researchers working with these data sources are restricted in the geographic detail to which they can disaggregate in describing results. And this confidentiality extends to studying and certainly extends to identifying particular companies. With the NETS data, in contrast, none of these problems arise. The data are accessible and there are no confidentiality restrictions imposed on the users.[12]

3. ASSESSMENT OF NETS DATABASE

We use three strategies to assess the reliability of the NETS data. First, we compare the NETS data with alternative data sources that are publicly available to assess the accuracy of measurements of employment levels and changes. Second, we search business relocation cases reported in the media and check whether they are captured by the NETS data. And third, we use phonebooks and company web sites to try to identify business establishment births and assess the accuracy with which the NETS tracks such births.

In all cases, the reader is reminded that there are complexities involved in each of these measurement exercises, and it is not clear that any one

Table 1. Previous Data Sources for Studying Employment Dynamics in the United States.

Database	Description	Evaluations and Applications
Early D&B data	Annual establishment-level data collected by the credit rating company Dun & Bradstreet for their commercial uses.	Birch (1979, 1981, and 1987) uses the data to study the role of small firms in job creation. Allaman and Birch (1975) study geographic migration of businesses. Aldrich, Kalleberg, Marsden, and Cassell (1989) show that the D&B data used by Birch tracked new businesses poorly. Davis et al. (1996) noted that the D&B data overstate total employment compared to BLS or Census data.
U.S. Establishment and Enterprise Microdata (USEEM)	The U.S. Small Business Administration used the D&B data from the late 1970s and early 1980s to create this data file, which linked establishments cross-sectionally (with parent firms) and over time.	Research assessing this file points to some coding and related errors common to most data sets, problems with coverage of single owner–operator establishments in a few sectors, and overall higher counts of employment in small establishments compared to Census data. Overall, though, the research concludes that the coverage of the file was quite accurate and timely (MacDonald, 1985, p. 180). Audretsch (1995) provides some further analysis of the quality of the USEEM and uses the data to study industry evolution.
Census of Manufactures (CM)	The Census of Manufactures is one part of the Economic Census collected by the U.S. Census Bureau every five years. It is restricted to the manufacturing sector and covers establishments with five or more employees.	Dunne et al. (1989a, 1989b) use these data to study manufacturing plant turnover, growth, and their resulting employment flows.

Longitudinal Research Database (LRD)	Created by the Census Bureau, the LRD is a large micro database of establishment-level data constructed by combining information from the Census of Manufactures (CM) and the Annual Survey of Manufactures (ASM). It is restricted to the manufacturing sector and covers establishments with five or more employees.	A large amount of literature has been based on LRD. See, for example, Davis and Haltiwanger (1992) and Davis et al. (1996), which use the LRD data to study the process of job creation, job destruction, and employment reallocation.
Longitudinal Business Database (LBD)	The LBD is created by the Census Bureau. It covers almost all the nonfarm private economy, as well as some public sector activities, improving upon the coverage of LRD.	Jarmin and Miranda (2002) document the efforts of constructing the LBD at the Census Bureau. Foster (2003) uses LBD data to study establishment and employment dynamics in Appalachia. Recent studies using the LBD are Jarmin et al. (2005), and Dunne et al. (2005).
Longitudinal Employer-Household Dynamics (LEHD) data	Also based in the Census Bureau, the LEHD program links federal and state administrative data with the Bureau's censuses and surveys to create a longitudinal database of employers and their employees.	U.S. Census Bureau (2002) provides detailed documentation of the LEHD data. Benedetto et al. (2004) explore exploiting the LEHD to improve tracking of entry and exit of firms as well as administrative changes by following clusters of matched workers.
Unemployment Insurance (UI) data	State employment security agencies are authorized by law to collect employment and wage information on workers covered by unemployment insurance, which results in a large amount of data on both employers and employees. UI data cover employers in all sectors and all size categories (except no employees), but do not capture physical relocations of business establishments.	Leonard (1987) uses UI data to describe the nature and magnitude of structural and frictional shifts in employment across industries and establishments. Spletzer (2000) uses UI data from West Virginia to study the contribution of establishment births and deaths to employment growth.

particular data source is the "gold standard." Thus, our analysis does not focus solely on whether the NETS "measures up" to these other data sources, but instead discusses the strengths and weaknesses of each and the degree of correspondence between them. There is a lot more to learn about these measurement differences, and we suspect that the potential advantages of the NETS will spur further assessments that build on those we carry out here.

It is important to note that there was a dramatic change in the data collection process at D&B in 1991. In particular, on July 25, 1991, a federal court ruling allowed regional Bells to sell information they collected (United States *vs*. Western Elec. Co., 767 F. Supp. 308 (D.D.C., 1991)). In 1992, therefore, D&B started to use yellow pages to identify business units, which greatly expanded its database. This resulted in a significant surge in the number of establishments and jobs in the NETS data in that year, which we expect to have seriously mitigated earlier problems with coverage by the D&B data. Consistent with this, as we show below, the D&B data now detect *more* very small establishments than do other data sources. From this point on, we will drop the 1989–1991 data and focus on the 1992–2002 sample period.

3.1. Measurement of Employment Levels

One approach to assessing the NETS database is to compare its estimates of employment levels and changes with similar estimates from other sources. The data products that can be used for these comparisons are the Quarterly Census of Employment and Wages (QCEW), the Current Employment Statistics (payroll) survey (CES), and the Size of Business data (SOB). The QCEW and SOB are based on ES-202 data.[13] Consequently, these two sources exclude the self-employed, proprietors, domestic workers, unpaid family members, and some other groups. The CES covers all nonfarm payrolls. These data sets only provide aggregate statistics at various geographic, industry, or establishment size levels, and thus it is only at these levels that we can compare the data sources.[14]

We begin by comparing employment level measurements in the NETS to those in the QCEW and the CES. We use 1997–2000 for the comparison with QCEW data because earlier years were not readily available, and subsequent years use the North American Industry Classification System (NAICS) instead of SIC codes, and hence cannot be directly compared. We use the full sample period for the CES comparison. For both sources, we do this at the most disaggregated level at which QCEW data are publicly available for all

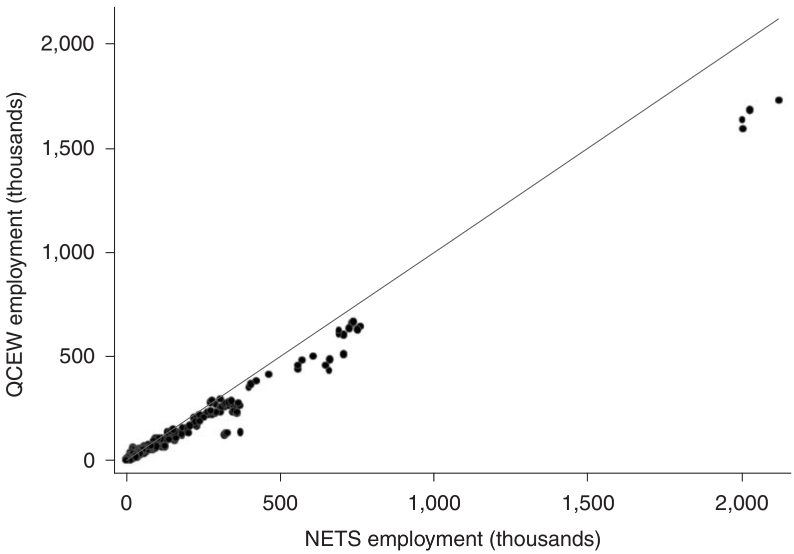


Fig. 1. NETS and QCEW Employment by County and Industry (1997–2000).

counties – by county and one-digit SIC industry.[15] Fig. 1 plots the data for the alternative measurements of employment by county and industry from the NETS and the QCEW. If the measurements agreed exactly, then they would all lie on a 45-degree line, which is drawn in the figure. It is clear from visual examination of the figure, as well as the very high computed correlation of 0.994, that employment levels in these two data sources correspond very closely. On the other hand, the points actually lie on a line that is flatter than the 45-degree line, implying higher employment levels in the NETS.[16] We return to this issue below. We constructed a similar figure for comparing employment measurements in the NETS with those in the CES data. It reveals a similar pattern, and also a high correlation (0.948).

To assess the quality of employment measurements in the NETS by establishment size, we also examined the correspondence between employment as measured by the NETS and the SOB; for this latter source employment can be measured by industry and size of establishment (as well as county, of course) as can also be done in the NETS database. Here the data correspond less well, and the correlation falls to 0.817. Looking at employment by establishment size shows that the greater discrepancies reflect the fact that the NETS database has much higher coverage of small establishments than does the SOB.

Part of the disparity in employment and the number of establishments indicated by the SOB and the NETS data sets for small establishments might be driven by the fact that business owners are typically excluded from coverage under the ES-202 UI system (although they are permitted to pay UI taxes and be covered).[17] This could be quite important for the smaller establishments in which single owners can represent a sizable share of total employment. So the comparison is more informative if we remove one employee from each establishment covered in the NETS database, and then reassign NETS establishments to size categories based on the adjusted employment levels. As shown in Fig. 2, the adjusted data on the size distribution of establishments in the two data sources indicates relatively similar distributions. But the NETS still captures more employment in the smallest size category, and the overall employment discrepancy between the two data sources is 5.0% (higher in the NETS).[18,19]

Finally, in Table 2 we attempt to account for differences between the NETS database and the SOB data described above by examining employment data for 1994–2002. The first two rows of the table indicate the total

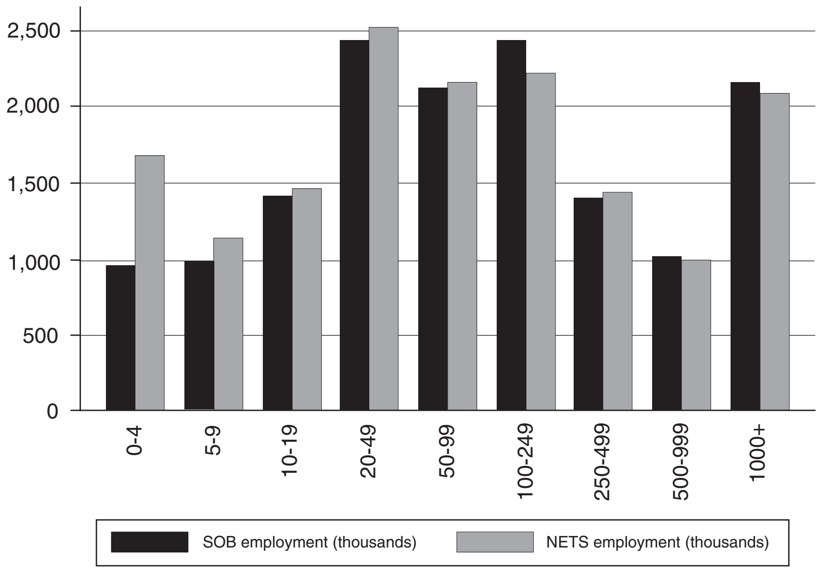


Fig. 2. NETS and SOB Employment by Size Category (2002), Subtracting One Employee from Each NETS Establishment, and Reassigning to New Establishment Size Categories.

Table 2. Accounting for the Discrepancies between NETS and SOB Employment (1994–2002).

	1994	1995	1996	1997	1998	1999	2000	2001	2002
1. NETS	16,371,012	16,241,156	16,314,659	16,546,553	16,512,479	16,864,781	17,666,262	18,149,748	17,527,918
2. Size of Business (SOB)[a]	12,696,157	13,047,314	13,312,913	13,739,592	14,257,229	14,642,495	15,144,896	14,997,165	14,967,297
3. Self-employed/ independent contractor (SE/IC)[b]	2,084,696	2,093,767	2,008,958	2,083,693	1,851,667	1,893,306	1,877,283	1,899,806	1,895,814
4. Size of Business + self-employed/ independent contractor (SOB + SE/IC)	14,780,853	15,141,081	15,321,871	15,823,285	16,108,896	16,535,801	17,022,179	16,896,971	16,863,111
5. Current Population Survey (CPS)	13,979,022	14,039,848	14,261,005	14,791,531	15,180,850	15,522,223	16,056,438	16,249,075	16,214,933
6. Row 1 – row 4	1,590,159	1,100,075	992,788	723,268	403,583	328,980	644,083	1,252,777	664,807
7. Row 5 – row 4	−801,831	−1,101,233	−1,060,866	−1,031,754	−928,046	−1,013,578	−965,741	−647,896	−648,178

[a]California Size of Business employment data includes individuals that are covered by unemployment insurance for the pay period that includes September 12th, regardless of the type of payroll. The self-employed and independent contractors, as well as several other worker categories, are excluded from unemployment insurance coverage (California Unemployment Insurance Code, Chapter 3, Article 2, Section 656).

[b]The number of self-employed and independent contractors is calculated by multiplying the weighted proportion of individuals reported in these categories in the February Contingent Work Supplement (CWS) to the Current Population Survey by the annual average of household employment in California. The CWS was compiled in 1995, 1997, 1999, and 2001. In this table, the 1995 CWS is used to calculate the level of self-employment and independent contractors in 1994 and 1995; the 1997 CWS is used for 1996 and 1997; the 1999 CWS is used for 1998 and 1999; and the 2001 CWS is used for 2000, 2001, and 2002.

employment levels reported in the NETS and the SOB, respectively, for each year. Since SOB data only include individuals earning wages that are covered by UI, several categories of workers that are reported in the NETS, most notably the self-employed and independent contractors, are excluded from the SOB by statute.

To estimate the number of individuals who are either self-employed or independent contractors, we use data reported in the Contingent Work Supplement (CWS) to the Current Population Survey (CPS) in years 1995, 1997, 1999, and 2001, as reported in row 3 of the table. In row 4, we combine the total employment reported in SOB with the number of self-employed and independent contractors from CPS to arrive at an approximate level of household employment in California (which we label SOB + SE/IC). As we see in row 6, our approximation of household employment in California falls short of the level of employment reported in the NETS database for each of the years we examined. It is instructive to note that while the NETS over-reports household employment in comparison to our approximation in row 4 of the table, this approximation itself overstates household employment when compared to the CPS employment figures for each year, as shown in the last row of the table. The differences between the NETS and the SOB + SE/IC series may be partly explained by some self-employed or independent contractors having multiple businesses – all of which should show up in the NETS, but not in the SE/IC series, where an individual is counted only once. On the other hand, this cannot account for the CPS versus SOB + SE/IC difference.[20]

Overall, these calculations suggest that the NETS estimate of employment (more accurately, the number of jobs) is higher than other sources because it uses a more comprehensive approach. Despite the remaining discrepancies – and note that there are discrepancies among any pair of data sets one chooses to compare – the NETS data appear to measure employment levels relatively accurately.

3.2. Measurement of Employment Change

Next, we turn to measurements of employment change. We first begin by documenting, in Figs. 3 and 4, the extent of rounding of employment in the NETS data. These figures show that for both smaller and larger establishments the distribution of the number of employees is disproportionately concentrated on numbers that are divisible by 5, 10, 100, and so on. While employment rounding may bias some of our estimations, it is not a

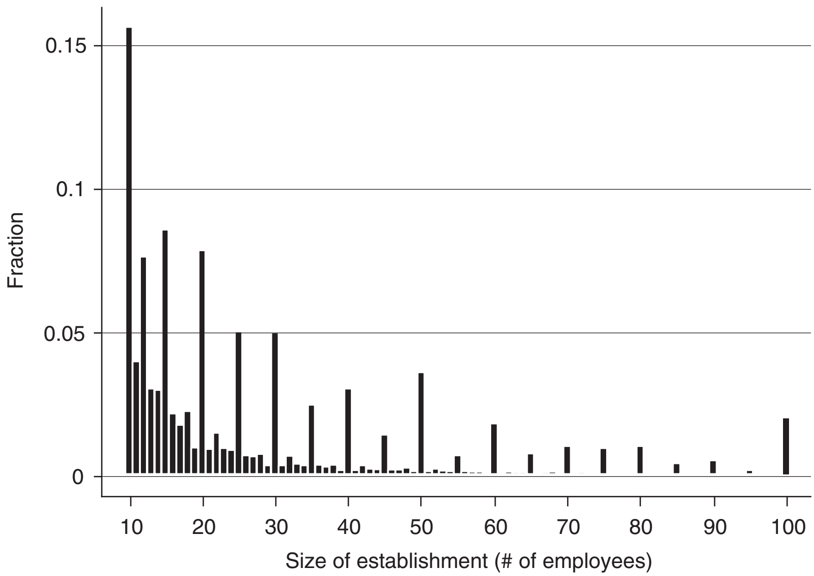


Fig. 3. Histogram of NETS Establishment Sizes, 10–100 Employees (2000).

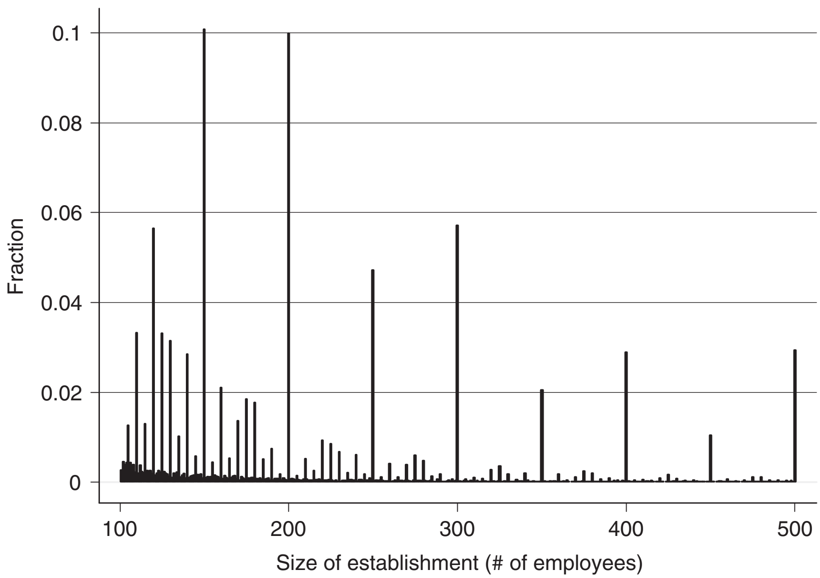


Fig. 4. Histogram of NETS Establishment Sizes, 101–500 Employees (2000).

particularly serious problem for the measurement of employment levels if we believe that employment numbers are rounded to the closest "salient numbers." In that case, our aggregate levels are unlikely to be biased appreciably, because some people round their numbers up and others round them down, and the establishment-level measurements may contain measurement error that is largely random (although non-classical). It does, however, mean that employment change is "sticky," and that our estimates likely underreport the frequency with which establishments change their levels of employment, thereby underestimating the degree of employment change caused by establishment expansion and contraction.

Another potential source of stickiness in the measurement of employment change in the NETS is imputed data. During 1993–2002, between 55% and 73% of each year's employment figures are actual data.[21] The remaining establishment records are imputed – either by D&B or by Walls & Associates – with the latter occurring when the D&B imputations were suspect, attempting to improve on the imputation by using time-series information on the establishment instead of only cross-sectional information. Imputed data are far less common for older establishments. Moreover, once actual employment data are provided for an establishment, they are very likely to be provided in all subsequent years. Both of these regularities indicate that imputation is a feature of establishments' earliest appearances in the database. More specifically, the establishments that are tracked for a relatively short period of time exhibit a bimodal distribution, with either no years with actual data or all years with actual data. But the establishments that are tracked in the data set for a longer period are much less likely to have no years without actual employment data, and conversely have relatively more years with actual data; and the mode is to have actual data for all years.[22]

Together, rounding and imputation of employment data result in infrequent year-to-year changes in employment. Table 3 illustrates this with regard to imputation, showing the incidence of employment change by type of employment data. As we would expect, it is far lower for imputed data. And overall, 7.6% and 16.3% of existing establishments reported a change in the number of employees in 1993 and 2002, respectively, and 19.6% and 14.1% of workers were at establishments that reported a change in the number employed in those years – numbers that we suspect are low.

The implication of these measurement problems is that the NETS data compare less favorably with other data sources when we look at employment changes, rather than employment levels, especially for high-frequency (short-term) changes. As shown in Fig. 5, the correspondence between NETS and QCEW yearly employment changes by industry and county is

Table 3. Share with Employment Changes from Previous Year, by Employment Imputation Type (1993–2002).

	Actual Figure (%)	D&B Estimate (%)	Walls Estimate (%)
1993	11.74	0.02	1.62
1994	8.01	0.02	2.00
1995	11.23	0.02	2.56
1996	13.05	0.02	3.17
1997	12.23	0.01	4.30
1998	11.85	0.00	4.65
1999	11.56	0.01	4.56
2000	8.24	0.01	2.82
2001	9.29	5.74	9.41
2002	8.36	29.14	0.01

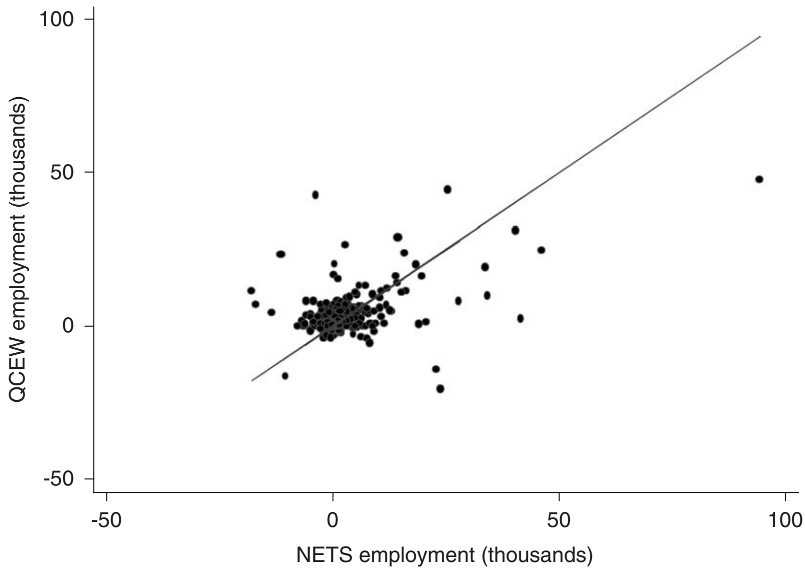


Fig. 5. QCEW/NETS Employment Changes, by Industry and County, One-Year Changes (1997–1998, 1998–1999, 1999–2000).

not very strong, with a correlation of only 0.528. However, if we look at employment changes over periods of at least a few years, this problem is substantially mitigated; for example, the correlation rises to 0.864 for changes over three-year intervals (Fig. 6).

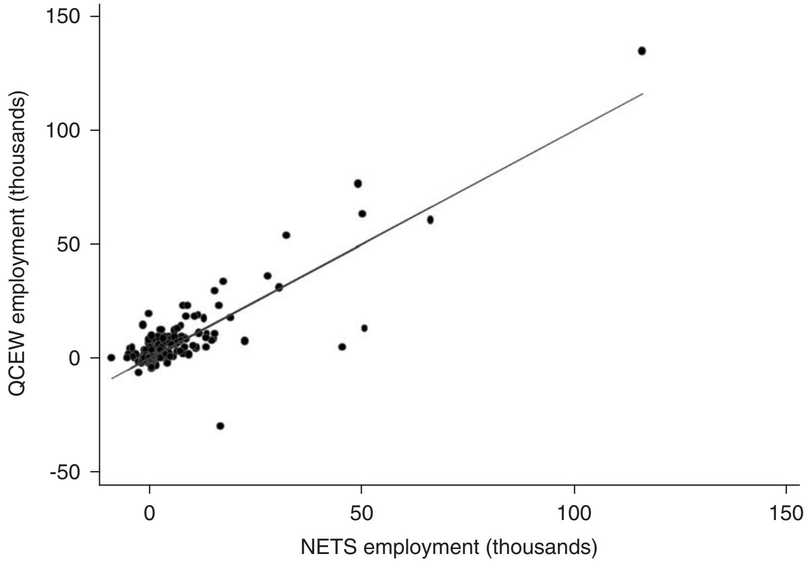


Fig. 6. QCEW/NETS Employment Changes, by Industry and County, Three-Year Changes (1997–2000).

This greater correspondence of employment changes over longer intervals is consistent with what we would expect based on the findings noted above regarding rounding and imputation. With rounding, the data will likely more accurately measure employment changes over a longer period, because rounding results in small changes being ignored but larger changes being measured. Similarly, we saw that imputation tends to be a feature of establishments' first appearance in the data set, whereas over time actual data are more likely to be reported, and hence employment changes are better measured. The implication of these findings is that the NETS database should not be used for measuring very short-term employment changes, but is more useful for measuring employment changes over periods of a few years or more. This does present a tradeoff, however, as an inability to focus on short-term changes inhibits our ability to observe high-frequency changes in job creation and destruction, such as over the business cycle.[23] We also note that making the unit of analysis for employment change longer affects what proportion of employment change we attribute to job creation and destruction versus establishment expansion and contraction, and to a much lesser extent relocation, a point to which we return below.

3.3. Tracking Business Relocations

A unique feature of the NETS data set is its ability to track establishment relocations. There are no other comprehensive data sets with which to compare measurements of geographic movement of establishments over time to such information in the NETS. Instead, we used Lexis-Nexis to search for business relocations involving California establishments, and conducted a detailed comparison of evidence on relocation in the NETS database to evidence found in these searches. Our search was not meant to be exhaustive; it was only intended to obtain a replicable sample of press coverage of specific business relocations.

We describe in detail our results from searching relocation reports in the *Los Angeles Times*, which has the largest circulation of any California newspaper. The *Los Angeles Times* has a regional bias in that it focuses on business moves in Southern California, especially the Los Angeles region. Business relocations in other regions are reported only if they are high profile or reflect a move between the Los Angeles region and the rest of the state.[24]

Using a carefully designed search algorithm,[25] we focused on 1,067 newspaper articles from the *Los Angeles Times* (1996–2000), from which we were able to identify 576 references to specific instances of business relocation, covering 452 unique relocation events. Of these, 237 business relocations were confirmed as valid moves by the NETS database. For the reported relocations not confirmed in the NETS, we undertook thorough efforts to independently verify whether there was in fact a relocation. It turns out to be very difficult to use other information sources to locate the establishments whose relocations are reported in the media, but for which there is not an obvious match in the NETS. Ideally, we would contact the establishment directly and confirm that the reported relocation occurred. However, this becomes very difficult when establishments (or often, businesses) can be acquired by other firms or for other reasons currently do business under a different name, or no longer exist. Naturally, these problems are more severe in trying to verify reports of relocation that are relatively old. Nonetheless, when possible we contacted the establishments directly. We also searched for company information using Hoovers.com[26] and Lexis-Nexis Company Information Search – web-based resources that track business addresses and would reveal new addresses for businesses that changed location.

Of the 215 relocations not found in the NETS, 47 were confirmed as "invalid" moves.[27] Of the remaining 168 reports of relocation that we could not locate in the NETS database, we were able to independently verify that 18 relocations indeed occurred. And not one of the 18 was a cross-state move. Despite our best efforts using the methods described above, we were unable to confirm the remaining 150 reports of relocation from Lexis-Nexis. And at least 91% of these businesses (136 out of 150) are captured by the NETS database with no relocation indicated. Furthermore, 92 (68%) of these establishments were still in existence through 2002, although we were only tracking relocations that were reported between 1996 and 2000. If these establishments had relocated, but not been tracked properly as relocations by NETS, then these establishments would have reported closing years close to the date of the relocation.

Thus, in total, 58.5% (237/{452 – 47}) of the valid business relocations that we identified from the *Los Angeles Times* could be found in our NETS data set. This rate of confirmation varies dramatically depending on the distance over which the relocation occurred. We are able to confirm only 27% (21/77) of within-city moves, whereas we are able to confirm 70% (177/252) of between-city, within-state moves, and 74% (37/50) of cross-state moves. It is neither surprising nor worrisome that the NETS detects only a relatively small share of within-city moves, because short-distance moves are much less significant for the scope of research for which this database is most useful. In fact, many within-city moves occur over such short distances that they could not be identified within the NETS database. For instance, several contacted establishments said that the moves had occurred, as indicated in the newspaper article, but the new location was adjacent to or "across the street" from the previous location. The NETS is designed to report only "significant moves," which are defined as moves where both the street address and zip code information change; this criterion was chosen to avoid mistaking the changing boundaries of zip codes for actual moves.[28]

We do not expect every relocation to appear in Lexis-Nexis, but we do expect all real relocations that are covered in the media to also appear in the NETS. Given the difficulty of checking whether reported cases actually occurred, it is impossible to quantify exactly what share of real relocation is captured in the NETS. But for moves crossing city or state boundaries, we estimate that the share is well over 75% and probably closer to 100%, based on the fact that most of the cases not captured by NETS cannot be independently confirmed as real relocations. Thus, we conclude that the NETS database does quite a good job of tracking business relocations, with a very low rate of false negatives, although our analysis probably pertains more to larger establishments that would be reported in the media. However, in contemplating the empirical results on establishment relocation discussed later in this paper, one might want to modestly adjust upward the job creation and destruction attributed to relocation.

3.4. Capture of New Business Establishments

Given the concern from earlier research regarding the ability of the D&B data to track new establishments, and the potential importance of establishment births in job creation, it is also important to assess how well the NETS tracks new business establishments. We do not have access to ES-202 data with which measurement of new establishments can be compared. We therefore first attempted to compare the NETS data to new establishments identified from phonebooks, following the earlier work by Birley (1984). Specifically, we selected a random sample of establishments from the 1999–2001 San Francisco Pacific Bell Business White Pages, and identified businesses that are initially *not* in the phonebook, but then show up in a later year, as a means of identifying an alternative list of new establishments, drawing a sample of 58 openings.[29]

Of these 58 openings, 52 (90%) of the establishments could be identified in the NETS database. Many listings were difficult to match because companies often do business under multiple names, and because of differences in spelling or abbreviations. Thus, this matching required that we also try to match using company name keyword search, phone number reverse lookup, address information, or alternate company names provided by workers whom we contacted at the particular establishment. While many of the NETS opening dates corresponded well with those indicated by the phonebook listings, many did not. Given the disagreements, we attempted to obtain each business establishment's opening year directly from the company, or through their website. We were able to obtain approximate start date information from 33 of these 52 establishments. This comparison revealed that many of the opening years that were indicated by changes in phonebook listings were inaccurate, with the phonebook method necessarily assigning these openings to 2000 or 2001, but the actual openings spread over a long span of years. In contrast, the NETS data match opening dates much more accurately, as indicated in Fig. 7.[30] And for those establishments that could not be reached directly, the NETS and phonebooks start dates were generally in agreement, because many of these were young establishments that failed subsequently, so there was not much scope for the start dates to differ.[31]

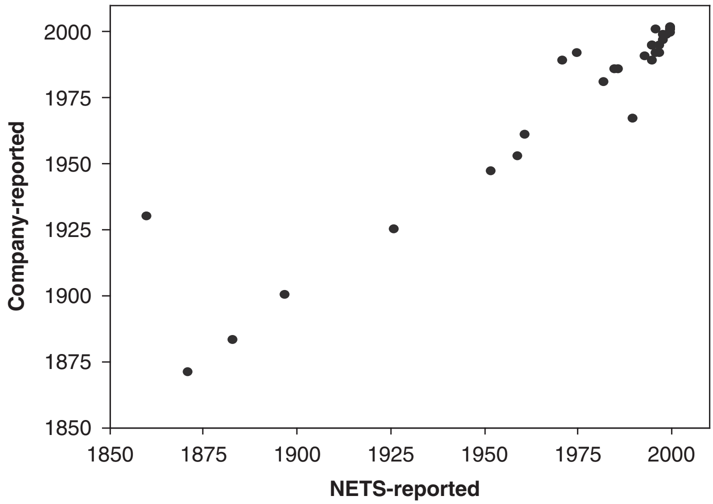


Fig. 7. Establishment Openings by Year, Company-Reported *vs*. NETS-Reported. *Note:* Figure covers establishments identified as opening in 2001 or 2002 from phonebooks, and identified in the NETS.

Given the inaccuracies in openings based on the appearance of businesses in the phonebook, we wanted to check another source for data on establishment openings.[32] To do this, we carried out a similar exercise for California biotech companies listed in the BioAbility database of U.S. biotech companies.[33] Because we are assessing how well the NETS captures births, we first chose companies that this database indicated were founded in our sample period for the NETS data (1992–2002), of which there were 300. To be more certain that we had the founding dates correct, we checked the BioAbility founding dates against company websites, retaining only the 161 cases for which the website also reported a founding date. Of these 161 cases, in 89% (142) the websites reported founding dates that corresponded exactly with the start year listed in the BioAbility database. If they did not match, we used the date from the company website, presuming that this was more accurate.

We then checked these founding dates against the appearance of these companies in the NETS to determine how well the NETS captures births. Only 8 of these 161 companies could not be located in the NETS database.[34] Of the remaining 153 records, 75% (114) of the start dates listed in the NETS corresponded exactly with the company start dates reported on the

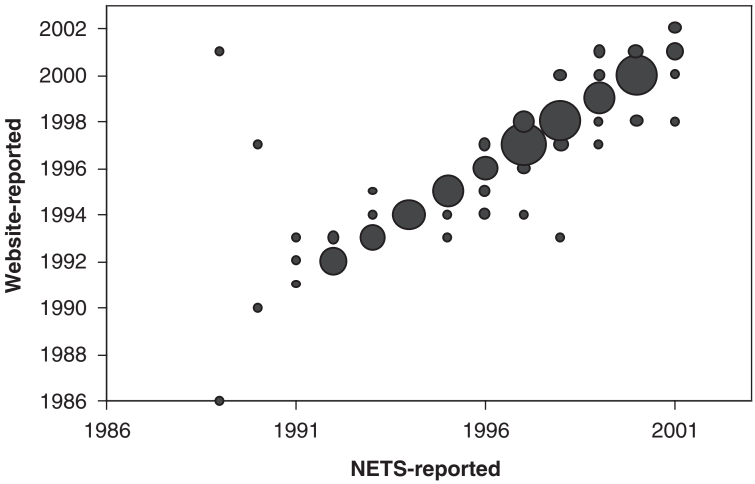


Fig. 8. Biotech Establishment Openings by Year, Reported on Company Website *vs*. NETS-Reported. *Note:* Areas of symbols are proportional to the number of observations with the same dates in both data sources.

web site, 88% (135) fell within one year, and 92% (141) of them fell within two years. The correspondence between the two data sources is graphed in Fig. 8. The correlation between NETS start dates and company website reported start dates was 0.87. Our two checks, then, indicate that the NETS tracks establishment births quite accurately, adding to the overall evidence of the reliability of the NETS data.

4. EMPLOYMENT DYNAMICS AND BUSINESS RELOCATION IN CALIFORNIA

4.1. The Policy Debate over Business Relocation

We illustrate the usefulness of the NETS data by addressing a substantive question regarding business establishment dynamics that figures prominently in policy debates, and for which the NETS database is uniquely suited. Specifically, business relocation is often cited as a source of job loss, especially at regional and state levels. While job loss due to business relocation is rarely precisely measured, it is often invoked in the rhetoric of policy debate and political campaigns. For example, when Kimberly-Clark moved its headquarters from Wisconsin to Texas in 1985, it sparked heavy criticism of Wisconsin's business climate, contributing to the governor losing his job in the following election (Dresang, 2002). Similarly, local governments work to attract companies to relocate to their jurisdictions (Klier & Testa, 2002). For example, in 2001, when Boeing announced that its headquarters would leave Seattle, cities like Chicago, Dallas, and Denver engaged in fierce competition to recruit the company.[35]

During the past decade, the debate over business relocation has been particularly prominent in policy discussions in California (see Neumark, Zhang, & Wall, 2005b). The media, business leaders, and politicians pay close attention to incidents of businesses leaving California, often citing them as a threat to the state's economy, and as evidence of a hostile business climate. For example, during the gubernatorial recall election in 2003, the public was inundated with criticism of California's business environment and stories of businesses leaving California (e.g., Vames, 2003). Candidates for Governor routinely referred to the state's "onerous business regulations and overtaxation" (Roberts, 2003) that were believed to push businesses away. After Arnold Schwarzenegger won the recall election and became Governor, he adopted an aggressive public relations strategy focusing specifically on business relocation, which in turn led to a response from other states (Tamaki, 2004).

Claims regarding the importance of business relocation that have arisen in this debate, however, have rarely relied on empirical evidence of relocation behavior. Rather, they generally rely on surveys that elicit subjective assessments from employers (e.g., California Business Roundtable and Bain & Company, 2004). One earlier study tried to measure actual relocation activity, based on data on manufacturing plants from the Los Angeles Department of Water and Power and several large utility companies in Southern California (Bules & Associates, 1992). But this study – like most of the public debate – focused only on businesses leaving the state, as if traffic moves in only one direction.

Moreover, the debate is often framed as if relocation is the key determinant of employment change, and hence a barometer of the "business climate." Yet the formation of new business establishments, the death of existing ones, and employment changes at continuing establishments, also affect employment change. Thus, the debate over business relocation ignores five of the six components of employment change (births, deaths, expansions, contractions, and in-migration), focusing only on the loss of jobs from establishments that move out of the state. The NETS data can obviously fill in many of the gaps in understanding the importance of business relocation (in both directions) in employment change, and more generally in identifying which processes – births, deaths, expansions, contractions, moves in, and moves out – drive employment change.

4.2. Aggregate-Level Evidence on Employment Dynamics

This section presents our findings from the NETS data regarding business relocation in California and employment dynamics more generally. The analysis is at the aggregate state level. In the next section we consider some analyses, focused more exclusively on business relocation, at the industry level. In this section and the next, in each case we state our key result and then provide some detailed discussion.

(1) *California generally loses establishments and jobs due to business relocation, but the impact is negligible.*

As Table 4 shows, in every year during the 1992–2002 sample period, some establishments left California, taking jobs away; at the same time, others moved into California, bringing jobs into the state. Measured by either the number of business establishments or the number of jobs, California experienced a net loss owing to relocation in every year. However, compared with the size of its overall economy, California's net loss from relocation has to be considered negligible. In terms of number of establishments lost to other states, the worst years are 1993 and 1994. In each of these years, California lost about 750 establishments to other states, which amounted to 0.05% of the total number of establishments in California. The job numbers tell a similar story. In terms of job loss from relocation, 1994 and 1997 represent the worst years. In these years, business relocation cost 0.1% of California jobs. Another way to see that these job change numbers are negligible is to compare them to ongoing employment changes that the state experiences. For example, from July 1990 to January 1993, employment in California fell by 6.1%, while from December 1997 to December 2000, employment in California grew by 8.2%.[36] These comparisons suggest that whether during an upturn or a downturn, business relocation simply does not play a major role in employment change.

(2) *Establishments are much more likely to move locally than across state boundaries.*

While establishment moves are quite common, most of these moves are within state. Out of 255,838 cases of establishment relocation originating in California during 1993–2002, 246,283 (or 96.3%) were moves within

Table 4. Business Relocation and its Effect on Employment in California, 1992–2002.

A. *By number of establishments*

Year	Moved In	Moved Out	Net Effect	Total Number of Establishments	Net Loss as % of Total
1993	612	1,364	−752	1,532,256	0.049
1994	534	1,285	−751	1,515,142	0.050
1995	519	1,104	−585	1,497,623	0.039
1996	489	835	−346	1,521,247	0.023
1997	504	763	−259	1,518,940	0.017
1998	545	676	−131	1,492,105	0.009
1999	582	669	−87	1,461,135	0.006
2000	802	828	−26	1,519,325	0.002
2001	752	1,032	−280	1,644,230	0.017
2002	731	999	−268	1,814,938	0.015

B. *By number of jobs*

Year	Moved In	Moved Out	Net Effect	Total Number of Jobs	Net Loss as % of Total
1993	13,853	27,094	−13,241	16,266,713	0.081
1994	8,977	25,452	−16,475	16,371,012	0.101
1995	14,136	28,224	−14,088	16,241,156	0.087
1996	13,158	18,352	−5,194	16,314,659	0.032
1997	11,073	28,209	−17,136	16,546,553	0.104
1998	15,098	16,709	−1,611	16,512,479	0.010
1999	18,893	23,437	−4,544	16,864,781	0.027
2000	15,589	16,994	−1,405	17,666,262	0.008
2001	18,586	23,916	−5,330	18,149,748	0.029
2002	12,656	16,551	−3,895	17,527,918	0.022

California. While cross-state moves draw a lot of attention, they are rare. In fact, 35.4% of all the moves originating in California occurred within a city and 78.5% of the moves did not go beyond the county boundary.[37] As a result, the impact of relocation on employment at the local level, while still modest, is more pronounced than its effect on state employment. In 1993, though less than 0.01% establishments moved out of California, 0.4% of establishments moved outside their own county, and 1.2% of establishments moved beyond their own city. The employment changes associated with these moves represented 0.1%, 0.6%, and 1.5% of total California employment, respectively. Of course, the preponderance of within-state moves may reflect the unique economic geography and size of California. The state has numerous quite different regional economies, and relatively few population centers near borders with other states that would permit "local" moves that nonetheless cross state lines.

(3) *Employment growth is primarily driven by expansion, contraction, births, and deaths.*

Table 5 presents decompositions of employment growth over three-year periods during 1992–2002. For each period, in the top panel we show California employment in the starting year, in the ending year, the overall net change, and then the number of jobs created or eliminated by each process of employment dynamics. The bottom panel shows the decomposition of employment change. In principle, we can decompose annual employment changes in the same way. But as noted earlier, year-to-year employment changes are not as reliable in the NETS data because of rounding and imputation.

Table 5 shows that in every three-year period the expansion of existing establishments always creates more jobs than are lost through the contraction of existing establishments. This is perhaps not surprising, because at any time we expect that the surviving business establishments tend to be those that are growing rather than shrinking. The net effect of births and deaths of establishments on overall employment change is positive in some years and negative in others. This tends to reflect the business cycle. In boom years many new establishments are created, and at the same time existing establishments are less likely to go out of business. As a result, jobs created by new establishments outnumber jobs eliminated by establishments that close in such years. Conversely, during slower economic times business formation is lower and more businesses tend to close, resulting in a net loss of jobs because new businesses do not suffice to cover the loss of those that die. For example, during 1995–1998, establishment deaths in California cut 454,000 jobs more than the number of jobs created through establishment births. But during the next three years, from 1998–2001, business establishment births and deaths resulted in a net gain of 848,000 new jobs.

The table also provides a comparison of the contribution of relocation to employment change with the contributions of other sources of employment change. The bottom rows of Table 5 indicate just how small the role of business relocation is. As the last row shows, the employment loss from relocation ranges from about 6,000 to 44,000, averaging around 20,000 per year. But the employment changes from the expansion–contraction processes and the birth–death processes are much greater, often by a factor of 20 or more. In other words, employment changes in California are primarily

Table 5. Decomposition of Employment Growth in California.

	1992–1995	1993–1996	1994–1997	1995–1998	1996–1999	1997–2000	1998–2001	1999–2002
A. Employment change								
Starting employment	16,394,151	16,266,713	16,371,012	16,241,156	16,314,659	16,546,553	16,512,479	16,864,781
Ending employment	16,241,156	16,314,659	16,546,553	16,512,479	16,864,781	17,666,262	18,149,748	17,527,918
Change	−152,995	47,946	175,541	271,323	550,122	1,119,709	1,637,269	663,137
Job creation								
Expansion	1,134,603	1,220,681	1,480,284	1,742,557	1,874,193	1,933,519	1,934,525	1,862,952
Birth	2,641,169	2,915,369	2,716,969	2,456,024	2,317,230	2,776,719	3,488,940	3,092,281
Move in	34,327	37,993	41,994	37,355	46,076	49,515	45,268	42,277
Job destruction								
Contraction	1,102,839	965,717	1,030,221	994,987	973,018	901,333	1,134,032	1,410,608
Death	2,781,915	3,086,093	2,965,193	2,909,694	2,648,325	2,682,980	2,640,929	2,870,695
Move out	78,340	74,287	68,292	59,932	66,034	55,731	56,503	53,070
B. Employment change decomposition								
Employment change =	−152,995	47,946	175,541	271,323	550,122	1,119,709	1,637,269	663,137
(Expansion−contraction)	31,764	254,964	450,063	747,570	901,175	1,032,186	800,493	452,344
+(Birth−death)	−140,746	−170,724	−248,224	−453,670	−331,095	93,739	848,011	221,586
+(Move in−move out)	−44,013	−36,294	−26,298	−22,577	−19,958	−6,216	−11,235	−10,793

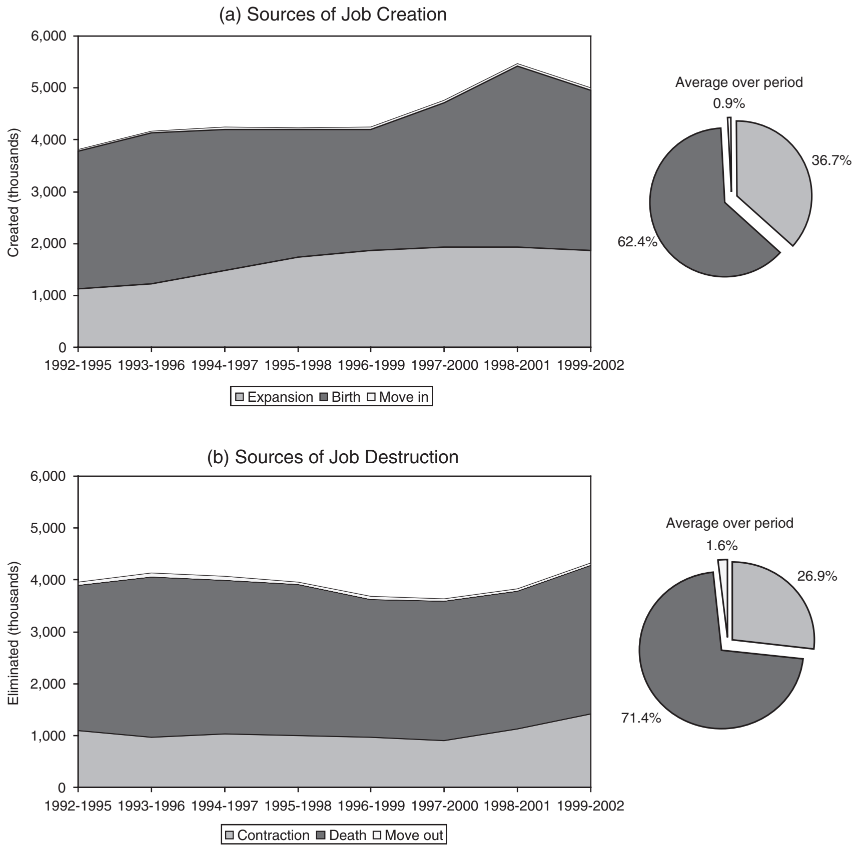


Fig. 9. Sources of Job Creation and Destruction.

driven by expansion–contraction and birth–death processes, rather than by relocation.

The relative importance of different sources of employment change is illustrated more clearly in Fig. 9. The two panels display the sources of job creation and destruction, respectively, in each three-year period during 1992–2002. The top panel shows that in each period the birth of new business establishments is the major source of job creation, while the expansion of existing establishments is also important. The number of jobs created by business establishments that moved to California is trivial compared to the number of jobs created by the other two sources. Likewise, the bottom panel shows that the death of establishments is the major factor in job destruction. Contraction at existing establishments is also substantial but less important. Finally, business relocation out of California again contributes only minimally. But recall that, on net, it is expansion minus contraction that generally yields the largest share of employment growth, and is always positive.

The decomposition of the sources of employment change is informative about the potential for each of the underlying processes to lead to more dramatic variation in employment. Given that births and deaths contribute large gross flows into and out of employment, a quite modest change in the balance between births and deaths could lead to large shifts in net employment growth. In contrast, the very low gross job flows associated with relocation imply that even if the rate of mobility out of the state doubled, and establishments completely ceased to move into the state, there would be little impact on net employment change.

(4) *Decompositions of the sources of employment change are sensitive to the interval over which the change is measured, but regardless, the contribution of business establishment relocation is negligible.*

There is a potential caveat to the results reported in Table 5 and Fig. 9. The magnitude of gross job creation and destruction, as well as its decomposition, is dependent on the interval length. First, as the interval gets shorter (for example, one year versus two, or a quarter versus a year), we might expect the gross flows to become larger because more employment changes due to temporary fluctuations are captured, although the opposite could occur (for example, as the interval length approaches zero). Second, the longer the interval chosen, the greater the contribution of births and deaths to gross flows. To see this most simply, note that all establishments in existence during a period are born and die during that period as the period gets infinitely long. Recall that we concluded that the NETS was not as reliable when looking at employment changes over shorter time intervals. Nonetheless, Table 6 illustrates that changes in the interval length (from one to ten years) do not affect the relative ranking of the contribution of each process of employment change to either job creation or job destruction, nor do they change the conclusion that the contribution of relocation is minimal.

This issue is also relevant when we compare employment change decompositions to what we get from other data sets, although because other data sets do not have information on business relocation, we have not focused on such comparisons. Results from ES-202 data (for the earlier 1990–1994 period) are available at frequencies of one, two, and three years, for West

Table 6. Employment Change Decomposition (1992–2002), Various Interval Lengths of Observation.

	1 year	2 years	3 years	5 years	10 years
Expansion	39.3%	37.9%	35.6%	33.5%	26.7%
Birth	59.8%	61.2%	63.5%	65.5%	72.2%
In-migration	0.8%	0.9%	0.9%	1.0%	1.1%
Gross creation	17,096,718	15,847,399	13,514,768	13,000,185	10,160,780
Contraction	32.1%	29.7%	27.5%	25.3%	20.7%
Death	66.4%	68.8%	70.9%	73.1%	77.5%
Out-migration	1.4%	1.5%	1.7%	1.6%	1.9%
Gross destruction	15,962,951	14,713,632	11,759,171	11,866,418	9,027,013
Net change	1,133,767	1,133,767	1,755,597	1,133,767	1,133,767

Note: For three-year intervals the analysis is limited to 1992–2001, which can be divided into periods of 3 years length.

Virginia (Spletzer, 2000).[38] As shown in Table 7, for the ES-202 data for West Virginia about 40% of gross job creation is attributed to births, compared with nearly 60% in the NETS, and the corresponding numbers for the contribution of deaths to job destruction are 41% and 66%. Part of the explanation for the smaller shares of job creation and destruction attributed to expansion and contraction in the NETS may stem from the "stickiness" in employment change, discussed earlier, resulting from rounding and imputation of employment in the NETS. Of course, the data sources cover two very different states in periods with little overlap, which may also help account for the differences. Note, though, that as we extend the interval to two and to three years, the discrepancies between these two data sources lessen, although some differences remain. The better match as we move to the three-year interval is consistent with our earlier conclusions that employment change measures in the NETS become more accurate as the window lengthens.

4.3. Industry-Level Evidence on Business Relocation

We next consider three analyses of business relocation and employment dynamics at the industry level. There are at least three reasons to examine

Table 7. Comparisons with ES-202 Data for West Virginia.

	NETS, CA, 1992–2002 (%)	ES-202, WV, 1990–1994 (%)	NETS, CA, 1992–2002 (%)	ES-202, WV, 1990–1994 (%)	NETS, CA, 1992–2001 (%)	ES-202, WV, 1990–1994 (%)
	1 year	1 year	2 years	2 years	3 years	3 years
Share of job creation						
Expansion	39.3	60.2	37.9	51.1	35.6	44.2
Birth	59.8	39.8	61.2	48.9	63.5	55.8
In-migration	0.8	N.A.	0.9	N.A.	0.9	N.A.
Share of job destruction						
Contraction	32.1	59.4	29.7	47.3	27.5	39.8
Death	66.4	40.6	68.8	52.6	70.9	60.2
Out-migration	1.4	N.A.	1.5	N.A.	1.7	N.A.

Note: The estimates for West Virginia using the ES-202 data are from Spletzer (2000).

business relocation and other sources of employment change by industry. First, if employment change differs by industry, the composition of jobs can be shifting and this is masked by focusing exclusively on overall job changes. The quality of jobs may vary across industries along a number of dimensions. Most notable, perhaps, is variation in pay. Indeed, average earnings in different industries in California vary considerably. In 2000, average annual pay was about \$60,000 in finance, insurance, and real estate, and \$58,000 in manufacturing. In contrast, in retail average pay was only \$22,000.[39] Thus, if a manufacturing job leaves the state and a retail job comes to the state, we might not want to view these as offsetting because, on average, a high-paying job has been replaced by a low-paying job.

Second, we would expect that business establishments in some industries are more mobile than in other industries; for example, we would expect that businesses that sell tradable goods may find it much easier to relocate to outside California but to maintain their customer base inside the state. In this case we may understate the importance of relocation when we look at the totals because we are averaging over industries where relocation is a viable strategy and industries where it is not.

Third, we have motivated the analysis of relocation in part based on attention to the issue on the part of the media, business leaders, and politicians. Although we have shown that relocation is a very minor contributor to job change, it is possible that relocation is important, and receives a good deal of attention, not because it constitutes a large flow of jobs but because it can reveal the "tip of the iceberg." That is, there may be a general problem with the health of industry X in California, but relocation behavior in the industry is most easily observable by the media and others.[40] Thus, it is important to ask whether changes in each of the sources of net job growth by industry – expansions minus contractions, births minus deaths, and relocations – are positively correlated rather than uncorrelated.

To assess these questions, we examine business establishment dynamics and employment change by industry, as reported in Table 8. We focus on one-digit SIC industries, although we break up the two largest industries – services and manufacturing – into sets of low-wage and high-wage two-digit sub-industries, based on whether average annual pay (in 2000) was above or below average annual pay for the one-digit industry.[41] We do this because, relative to other one-digit industries, manufacturing and services each include high-paying and low-paying sub-industries,[42] implying that small employment changes in the one-digit industry could mask larger shifts from higher-paying to lower-paying jobs, or vice versa. For each industry, in the first four columns we show the total employment change over the period 1992–2002,

Table 8. Business Establishment Dynamics and Employment Change by Industry, 1992–2002.

SIC codes	Major industry title	Net Employment Change, 1992–2002: Total	Expansion–contraction	Birth–death	Move	Net Change, Share of 1992 Employment: Total (%)	Expansion–contraction (%)	Birth–death (%)	Move (%)	Average Annual Pay, 2000[a]
01–97	All industries[b]	1,104,192	850,749	312,597	−59,154	6.7	5.2	1.9	−0.4	$41,182
15–17	Construction	43,252	136,366	−89,936	−3,178	5.6	17.8	−11.7	−0.4	$40,360
20–39	Manufacturing	−218,996	136,247	−340,583	−14,660	−8.9	5.5	−13.8	−0.6	$57,695
	High-wage manufacturing[c]	−26,763	44,354	−65,024	−6,093	−2.6	4.4	−6.4	−0.6	$91,278
	Low-wage manufacturing[c]	−192,233	91,893	−275,559	−8,567	−13.3	6.3	−19.0	−0.6	$35,953
40–49	Transportation and public utilities	98,406	23,980	83,878	−9,452	11.4	2.8	9.8	−1.1	$47,278
50–51	Wholesale trade	42,600	126,812	−77,735	−6,477	4.5	13.3	−8.1	−0.7	$48,935
52–59	Retail trade	184,508	78,444	106,556	−492	7.3	3.1	4.2	0.0	$21,815
60–67	Finance, insurance, and real estate	114,011	81,421	55,721	−23,131	9.6	6.9	4.7	−1.9	$60,163
70–89	Services	954,064	338,065	616,479	−480	15.2	5.4	9.8	0.0	$41,372
	High-wage services[d]	423,900	154,817	264,196	4,887	17.8	6.5	11.1	0.2	$54,484
	Low-wage services[d]	530,164	183,248	352,283	−5,367	13.6	4.7	9.0	−0.1	$29,690

[a] *Source:* Quarterly Census of Employment and Wages (QCEW), downloaded from http://www.calmis.cahwnet.gov/file/es202/cew-select.htm (viewed on March 1, 2006). 2000 was the last year QCEW classified industries by SIC code; it has shifted to the NAICS since then. For 2% of the establishments whose SIC code changed over time, we regard them as belonging to the industry in which they are classified for the most number of years. In the event that an establishment is classified in two industries for an equally long period of time, the more recent of the two industries is chosen. High-wage (and low-wage) industries represent a grouping of SIC2 sub-sectors which fall above (or below) the average salary for that industry.

[b] Excludes unclassified establishments (SIC 99).

[c] SIC Codes: high-wage manufacturing (28, 29, 35, 36, 38); low-wage manufacturing (20, 21, 22, 23, 24, 25, 26, 27, 30, 31, 32, 33, 34, 37, 39).

[d] SIC Codes: high-wage services (73, 78, 81, 87, 89); low-wage services (70, 72, 75, 76, 79, 80, 82, 83, 84, 86, 88).

and decompose the total change into three sources: expansions minus contractions, births minus deaths, and in-migration minus out-migration.[43] The next four columns report the same figures on a percentage basis, showing the overall change in employment and the separate components as percentages of 1992 total employment in the industry. Finally, the last column reports average annual pay (as of 2000) in each industry. Our findings are as follows:

(5) *Job loss due to relocation has tended to occur in higher-paying industries.*

We first consider the relationship between relocation and the quality of jobs as measured by annual average pay in the industry. This analysis focuses on the fourth and eighth columns of Table 8, which show the levels of employment change due to relocation, and these changes as shares of 1992 employment, and the ninth column, which reports average pay.

There is evidence indicating that relocation cost more jobs in higher-paying than in lower-paying industries, suggesting that relocation may have had a negative impact – although modest – on the composition of jobs. In particular, net job loss attributable to relocation was highest in the finance, insurance, and real estate industry, where relocation cost nearly 2% of 1992 employment; and this is the highest paying one-digit industry. Similarly, the contribution of relocation to job loss was higher in manufacturing – although the difference relative to other industries is much less pronounced – and manufacturing ranks second among one-digit industries in terms of average annual pay. Nonetheless, it is interesting that the job loss due to out-migration was the same in high-wage and low-wage manufacturing, indicating that there is not a tendency to lose the higher-wage jobs in this industry via relocation. Overall, though, the correlation between average annual pay and the net change in employment due to relocation is -0.23, indicating that there was more job loss in industries where pay is high; using percentage changes in employment, the correlation is -0.36.[44] This evidence implies that relocation – although it has a small influence – did tend to result in the substitution of lower-wage for higher-wage jobs over this period.

(6) *Job loss from interstate relocation is similar in "footloose" industries and other industries.*

Next, we look at differences in relocation by industry, with an emphasis on asking whether it is more significant in tradable-goods industries in which gross relocation rates are significantly higher. Table 8 shows that most private industries lost jobs due to relocation; the only exception is high-wage services. At the same time, as in the aggregate, the contribution of relocation to total employment change within industries is relatively small. In all industries except two, job loss due to business relocation over the decade is less than 1% of initial employment. (The comparable aggregate figure, shown in the first row of Table 8, is 0.4%.) Although California experienced a large loss in manufacturing jobs during 1992–2002 (as did many other states in the nation), job loss due to relocation in this industry is not a major contributor, accounting for a loss of 0.6% of jobs, versus 0.4% for all industries combined. Instead, business closures are far more important in explaining the decline in manufacturing jobs. Of course the NETS does not track moves overseas, which would be regarded as closures. From the perspective of simply accounting for job loss, the distinction may be irrelevant. But from the perspective of policy it is quite important. The finding that manufacturing job loss is *not* due to moves of manufacturing establishments to other states undermines arguments that the problem facing manufacturing in the state is a bad business climate relative to other states, although given the relatively small role of relocation we would learn more from comparisons across states of births, deaths, expansions, and contractions.

Moreover, the evidence indicates that relocation does not loom particularly larger in industries that are more "footloose." Although not reported in Table 8, we compared the distributions across industries of job changes due to in- and out-migration to the overall distribution of employment across industries. This calculation revealed that manufacturing (both high-wage and low-wage) is footloose. Its 1992 employment share was 15.1%, but it contributed 32.1% of job gains due to in-migration, and 29.6% of job loss due to out-migration. In contrast, for example, retail trade had a 1992 employment share of 15.4%, but contributed only 7.7% of in-migration and 5.3% of out-migration. However, while manufacturing is, as we would expect, more footloose, as already noted the *net* effect of relocation in manufacturing is still quite negligible, accounting for a loss of only 0.6% of jobs over 1992–2002. Thus, relocation does not appear particularly more significant when we focus attention on the footloose manufacturing industry. Furthermore, the substantial job loss due to out-migration of business establishments in finance, insurance, and real estate is again worth noting because this is not a particularly footloose industry. In-migration is disproportionately low, contributing only 5.7% of the total job gains due to in-migration, relative to the industry's 7.3% share of 1992 employment. In contrast, this industry contributed 17.3% of total job loss attributable to out-migration.

(7) *Relocation does not appear to be an indicator of more substantial problems of job creation or destruction stemming from births, deaths, expansions, and contractions.*

Finally, we ask whether industries experiencing job loss due to relocation were also experiencing job loss due to either the excess of deaths over births, or of contractions over expansions. If so, one might argue that media and policy attention focused on relocations is detecting more widespread problems, and that perhaps the focus is on relocations because these are most easily observable. This could potentially be quite significant. We just noted that the correlation between job changes due to relocation and average industry pay is −0.23. Relocation is small or negligible relative to job change from expansions, contractions, births, and deaths. But if there are similar correlations between net job change due to births minus deaths, or expansions minus contractions, and industry pay, this would indicate larger-scale substitution of jobs in low-paying industries for jobs in high-paying industries.

Some of the numbers in Table 8 suggest that this may not be the case. For example, job loss due to relocation is the same (as a share of 1992 employment) in high-wage and low-wage manufacturing, but job loss due to more deaths than births is much higher in low-wage manufacturing (19% of 1992 employment, versus 6.4% for high-wage manufacturing). Similarly, although the job loss due to relocation is most pronounced in finance, insurance, and real estate, this industry also had robust net job creation due to expansions minus contractions, and a net job creation rate in about the middle of the pack due to births minus deaths. This overall impression is confirmed by looking at correlations across industries between the percentage changes in jobs due to each of the three net processes, which gauge whether trends in employment due to expansions minus contractions, births minus deaths, and relocations are similar or not. We find a correlation of −0.01 between the percentage changes in employment due to relocation and due to expansions minus contractions, 0.06 between the percentage changes due to relocation and due to births minus deaths, and 0.07 between the percentage changes due to relocation and the combination of births, deaths, expansions, and contractions. Thus, the patterns of job loss (or gain) in relocation are largely uncorrelated with the patterns generated by the other two net processes.[45]

5. CONCLUSIONS

We assess and present findings from a newly constructed longitudinal database covering business establishments – the NETS. The NETS database is particularly well-suited to study the underlying dynamics of employment change, specifically the processes of business establishment expansion and contraction, births and deaths, and relocation. As such, it builds on earlier research on this topic using the Longitudinal Research Database to study manufacturing, and numerous other data products based on the ES-202 data to study all sectors of the private economy. However, the NETS has some important advantages, including capturing business relocation, more complete coverage, and the ability to disaggregate to a fine geographic level, as well as ease of access and the absence of confidentiality restrictions.

Since the NETS is based on Dun & Bradstreet data – which have been criticized in the past – we devote a great deal of attention to assessing the quality of the NETS. Overall, we conclude that the NETS data are quite reliable and in many respects comparable to more frequently used administrative and Census data. The NETS captures new businesses and start dates quite accurately. Coverage of business moves in the NETS is good, which enables researchers to tackle a source of job creation and destruction that has been understudied. One limitation is that because data are often initially imputed for new establishments, and there is considerable rounding of employment, short-term (such as one-year) employment changes are not measured very accurately; use of somewhat longer intervals mitigates this problem.

Partly as an illustration of the value of the NETS data, and partly out of substantive interest, we use the data to study employment dynamics in California. We provide overall decompositions of the sources of employment change in the state, focusing particular attention on the empirical importance of business relocation into and out of the state, which has figured prominently in policy discussions.

We find that the birth–death and expansion–contraction processes of business establishments are responsible for nearly all gross job creation and destruction, and that cross-state business relocation is virtually a negligible factor. Cross-state business relocation resulted in a net job loss to California in every year during the period 1993–2002. However, compared to the size of the California economy, the net loss from relocation is trivial. This limited effect of business relocation on employment is also found within each industry, although business establishments in certain industries (such as manufacturing) are much more likely to move in both directions. There is some evidence that relocation – though relatively unimportant empirically – has cost the state higher-paying jobs. Finally, trends in job loss from relocation do not appear to be indicative of related trends in job loss from the other dynamic processes driving employment. Overall, given the small role played by relocation in job growth, these findings imply that a policy focus on business relocation is badly misdirected, and unlikely – even if successful at attracting new businesses and retaining old ones – to contribute visibly to job growth. To the extent that policy has any role to play, the evidence suggests that efforts to foster the formation of new businesses and to help existing businesses survive and grow would be better placed, unless relocation behavior is inordinately responsive to policy.

NOTES

1. Strictly speaking, whenever we refer to "employment" measured at the establishment level, we should refer to "jobs," because workers can have jobs with more than one employer. But since "job change" usually conveys a different meaning than "employment change," we usually refer to employment instead.

2. An observation in the NETS data is an "establishment." An establishment is a business or industrial unit at a single physical location that produces or distributes goods or performs services, for example, a single store or factory. Many firms own or control more than one establishment, and those establishments may be located in different geographic areas and may be engaged in different industries. We will sometimes refer to an establishment as a "business," reserving the word "firm" to refer to what may be collections of many establishments. While the NETS database is based on information collected at the establishment level, it also uses the Data Universal Numbering System (DUNS) to indicate the relationships among establishments in multi-establishment firms.

3. Data are available for the entire country, but cost precluded purchasing the entire file. A two-year license for the California file we use in this paper costs $15,000; a similar license for the entire file would cost in the range of $200,000.

4. See http://mddi.dnb.com/mddi/story.aspx (viewed on April 28, 2005). The information from the U.S. Postal Service includes the National Change of Address database of all changes of address in the United States.

5. See, for example, http://www.dnb.co.in/whoduns.htm (viewed on May 11, 2005).

6. See http://www.dnb.com/us/about/db_database/dnbinfoquality.html (viewed on April 28, 2005).

7. The file indicates when data are imputed.

8. There is also a detailed cross-walk between these 8-digit codes and the North American Industry Classification System (NAICS).

9. Most of the information in this paragraph was supplied by Don Walls and confirmed by him with D&B (personal communications, May and July, 2005).

10. With the full national NETS database one can do both of these.

11. There is also related data available for studying employment and business establishment dynamics in other countries. The work with which we are aware focuses to a large extent – although not exclusively – on the role of small firms or establishments in job creation, as originally considered for the United States by Birch (1987) and Allaman and Birch (1975), and then reconsidered by Davis, Haltiwanger, and Schuh (1996). Examples of this work include: Davidsson, Lindmark, and Olofsson (1998),

who study a plant-level data set for Sweden compiled by the authors based on the registers of all business establishments from Statistics Sweden; Broersma and Gautier (1997), who use data on a sample of manufacturing firms in the Netherlands; Baldwin and Picot (1995), who use longitudinal data on Canadian manufacturing establishments from Statistics Canada's Census of Manufactures; and Bednarzik (2000), who reports some results for 10 European countries based on Eurostat data on small- and medium-sized enterprises. The only study we have found that looks at relocation decisions is a paper by van Dijk and Pellenbarg (2000) studying firm relocation in the Netherlands, based on data from Chambers of Commerce throughout the country. It appears that the business register data (as in Sweden) most closely parallel the type of data available in the NETS and other U.S. datasets.

12. Although the NETS data are expensive, the time plus money costs appear quite favorable, compared to using the alternative data sources, which as noted are very difficult to access, and also require payment of "substantial user costs" to the Census Bureau (see http://webserver01.ces.census.gov/index.php/ces/1.00/research-guidelines, viewed on March 22, 2006); these can range into tens of thousands of dollars per year.

13. The ES-202 program, formally known as the Covered Employment and Wages program, is a joint effort of the Bureau of Labor Statistics (BLS) and state employment security agencies. Using quarterly data collected by the state agencies, BLS summarizes employment and wage data for workers covered by state UI laws and for civilian workers covered by the Unemployment Compensation for Federal Employees (UCFE) program. The ES-202 program provides a comprehensive and accurate source of data on employment and wages, by industry, at the national, state, and county levels (Bureau of Labor Statistics, 1997, Chapter 5).

14. The publication of employment/establishment data is sometimes withheld in order to protect the identity of cooperating employers. For example, QCEW data are suppressed if there are fewer than three establishments in a cell, or if a single employer makes up more than 80% of the employment in that cell.

15. Data at finer levels of industry disaggregation are often suppressed at the county level for reasons of confidentiality.

16. The points that are farthest off the line, at high employment levels, are for service-related industries in Los Angeles. However, these points actually lie quite close to a regression line through the data.

17. This same point was noted earlier in work by MacDonald (1985).

18. Unfortunately, CES does not include a series for the number of establishments. Therefore, a similar comparison of average establishment size is impossible. Because CES is periodically benchmarked to UI universe counts, we would not expect that results would be much different.

19. In addition, apparently in the NETS data separate lines of business (industry) at the same physical site are sometimes reported as separate establishments, which would tend to create a higher count of establishments of smaller sizes. However, this seems unlikely to be important for very small establishments that are unlikely to operate in more than one industry. Also, note that this should not play a role in the measurement of employment levels, unless there is double counting.

20. We carried out a similar comparison using alternative estimates of the self-employed provided by the Census' Nonemployer Statistics, available for some years.

Estimates using this method are somewhat closer to the NETS employment levels for the most years. This method likely explains more of the gap between the NETS and the SOB data because it allows self-employed individuals to report multiple businesses, as does the NETS database. (See Neumark et al., 2005a, for details.)

21. A very small share of these, 1–2%, are reported by D&B as "bottom of range" rather than actual data, and seem to indicate cases where the responded provided a range for employment rather than a single number. However, judging by the variation in these observations, they behave like actual data and are treated as such in this discussion.

22. More detail is given in the working paper version of this study (Neumark et al., 2005a).

23. This does not imply that the NETS data are uninformative about high-frequency changes, just that researchers using these data for this purpose need to be cognizant of the measurement error and to think about how it may affect their estimates and conclusions.

24. We also searched the *Kiplinger California Letter*, a concise bi-monthly business newsletter that has a section specifically reserved for business relocation reports. Comparing the two sources we found that, as expected *Kiplinger California Letter* provides more balanced coverage of business moves in different regions in the state.

25. We experimented with several search terms. Our final choice of search term – "ATL2(RELOCAT!) AND BUSINESS AND (MOVE OR MOVING OR MOVED) AND (SECTION ("BUSINESS") OR SECTION ("METRO"))" – was guided by a desire to exclude irrelevant articles, which we assessed through repeated searches and screening of articles. A detailed appendix describing how we arrived at this search term is available from the authors.

26. Hoovers.com utilizes the same raw data provided by the DMI file as the NETS database. However, the search mechanism is very flexible, sometimes making it easier to locate establishments that could not be found through company keyword searches in the NETS database.

27. Five moves turned out to be consolidations of businesses because the establishment at the destination already existed before the move; 17 cases were planned moves but did not occur later; 12 of the establishments at "destination" were new branches instead of relocated businesses; and 13 moves involved establishments such as schools and nonprofits that are not the focus of our research.

28. Our search for reports of business relocation in the *Kiplinger California Letter* (1996–2001) revealed similar results. Of the 79 incidents of relocation we identified in this search, 12 were found to be misreports of establishment relocation. Of the remaining 67 media observations of relocation, 35 (55%) were confirmed in the NETS database. In addition, three cases were confirmed in Hoovers.com, but occurred too recently to be found in the NETS, and 29 cases could neither be confirmed nor denied.

29. We randomly chose to start with listings beginning with "B." We have no particular reason to expect a relationship between the name of the business and the likelihood of its inclusion in the NETS database, although we cannot rule this out. Preliminary investigation suggested that business establishments that use initials in their name (such as "B & G auto rental") may change names from year-to-year, so we instead began drawing our sample with telephone listings that started with "Ba." We chose enough observations to get approximately 60 new establishments, which

required 313 listings that appeared in the phonebook at least once between 1999 and 2001. We excluded 35 records from the analysis because businesses from outside the area code presumably have to pay to be listed in the business white pages, meaning that the appearance or disappearance of such businesses would often occur for reasons unrelated to opening (or closing). These records indicate 58 openings (and 61 closings). There were also three records for which the listing appeared in 1999, was absent in 2000, and reappeared in 2001 with the same name and phone number. This is one indication – more are described below – that the phonebooks do not provide an accurate means of tracking openings and closings.

30. Even if the NETS data were completely accurate, we would not expect an exact correspondence with the start dates obtained from our efforts to contact businesses directly. In our phone calls, we often talked to employees who had limited tenure and did not know the founding date, in which case we were only able to obtain information that provided a lower bound for the number of years that a particular establishment had been in business.

31. As another check on the NETS data, we also attempted to locate NETS records for business establishments that were listed in the San Francisco phonebook for all three years. If NETS records indicate opening or closing years within 1999–2001, then we might be concerned that NETS is inaccurately reporting the timing or incidence of openings or closings. We randomly chose 72 of the 156 records in the phonebooks in all three years, and we were able to locate 66 (92%) in the NETS database, which represents a slightly higher percentage than those that we could identify from the earlier subset of phonebook-inferred openings. Of these records, according to the NETS data all but 6 continued to exist through 2002, and only one record indicated a closing by 2001, indicating a close correspondence between the NETS and phonebook data for continuing establishments.

32. For the same reason, coupled with the difficulty of verifying information directly with businesses that have closed, we deemed the phonebook method inappropriate for assessing the ability of the NETS data to identify establishment closings.

33. This is a database of more than 2,000 U.S. biotech companies (based on a relatively narrow definition of biotech) maintained by BioAbility, a biotech consulting firm. See http://www.bioability.com/us_biotech_companies.htm (viewed on September 14, 2005).

34. Six of these eight were founded in 2001 or 2002. Hoovers.com (based on the same raw data from D&B) listed a company record and DUNS number for all six of these records, indicating either that they were established too late to be included in our extract of the NETS database or that they were picked up by D&B with a modest delay after their establishment.

35. See, for example, http://www.conway.com/ssinsider/incentive/ti0106.htm (viewed on May 2, 2005).

36. See http://www.calmis.cahwnet.gov/file/lfhist/cal$shlf.xls (viewed on May 3, 2005).

37. As discussed in the previous section, within-city moves may be undercounted in the NETS, in which case these percentages would be even higher.

38. Published results from this source as well as the Business Employment Dynamics (BED) and LDB (a longitudinal establishment file at the U.S. Bureau of Labor Statistics) are only available on a quarterly basis for the United

States as a whole, and that is a higher frequency than we can study with the NETS data.

39. This is partly because more workers in the retail sector work part-time, although clearly hourly pay is much lower in retail as well.

40. We are grateful to a referee for raising this issue.

41. We omit industry-specific results for mining, for agriculture, forestry, and fishing, and for government; the first two are small, and none of the three is generally subject to relocation as we normally think about it. These industries are, however, included in the first row of Table 8.

42. This is reflected in far greater variance in average pay across two-digit industries within these one-digit industries, than in other one-digit industries.

43. We also examined changes over shorter time intervals averaged over this decade, and the qualitative conclusions were the same.

44. The first measure is more meaningful as a simple description of the relationship between job change due to relocation and industry pay. These correlations are computed treating high-wage and low-wage manufacturing as two data points, and similarly for services. The computed correlations are much more negative if we do not disaggregate manufacturing and services into high- and low-wage sub-industries, which masks somewhat higher job loss from relocation in low-wage than in high-wage manufacturing, and job gains from relocation in high-wage services coupled with job loss in low-wage services.

45. The correlations of the absolute (rather than percentage) changes are more positive, but these are driven by variation in the size of industries, and are less reflective of how similar relocation trends are to trends in births minus deaths, expansions minus contractions, or their combination. Furthermore, the larger positive correlation between employment change due to relocation and due to the other dynamic processes is in large part driven by high-wage services, which is the one (sub)-industry that experienced employment growth due to relocation and also experienced relatively large employment increases due to the other dynamic processes (especially births minus deaths).

ACKNOWLEDGMENTS

We are grateful to Don Walls, an anonymous referee, and seminar participants at PPIC, the NBER Summer Institute, the Hudson Institute, and the 4th Annual IZA/SOLE Transatlantic Meetings of Labor Economists for valuable comments. Any views expressed are those of the authors alone, and not of the Public Policy Institute of California.

REFERENCES

Aldrich, H., Kalleberg, A., Marsden, P., & Cassell, J. (1989). In pursuit of evidence: Sampling procedures for locating new businesses. *Journal of Business Venturing*, *4*, 367–386.

Allaman, P. M., & Birch, D. L. (1975). *Components of employment change for states by industry group, 1970–1972*. Working Paper No. 5, Joint Center for Urban Studies of MIT and Harvard University.

Audretsch, D. B. (1995). *Innovation and industry evolution*. Cambridge: The MIT Press.

Baldwin, J., & Picot, G. (1995). Employment generation by small producers in the Canadian manufacturing sector. *Small Business Economics*, *7*, 317–331.

Bednarzik, R. W. (2000). The role of entrepreneurship in U.S. and European job growth. *Monthly Labor Review*, *123*, 3–16.

Benedetto, G., Haltiwanger, J., Lane, J., & McKinney, K. (2004). *Using worker flows to measure firm dynamics*. Unpublished Manuscript, University of Maryland.

Birch, D. L. (1979). *The job generation process*. Unpublished report prepared by the MIT Program on Neighborhood and Regional Change for the Economic Development Administration, U.S. Department of Commerce, Washington, DC.

Birch, D. L. (1981). Who creates jobs? *The Public Interest*, *65*, 3–14.

Birch, D. L. (1987). *Job creation in America: How our smallest companies put the most people to work*. New York: Free Press.

Birley, S. (1984). Finding the new firms. *Proceedings of the Academy of Management Meetings*, *47*, 64–68.

Broersma, L., & Gautier, P. (1997). Job creation and job destruction by small firms: An empirical investigation for the Dutch manufacturing sector. *Small Business Economics*, *9*, 211–224.

Bules & Associates. (1992). *California industry migration study*. Prepared for Los Angeles Department of Water and Power, PG&E, San Diego Gas and Electric Company, Southern California Edison Company, and Southern California Gas Company, San Francisco, CA, October.

California Business Roundtable and Bain & Company. (2004). *California competitiveness project*. February.

Davidsson, P., Lindmark, L., & Olofsson, C. (1998). The extent of overestimation of small firm job creation – An empirical analysis of the regression bias. *Small Business Economics*, *11*, 87–100.

Davis, S. J., & Haltiwanger, J. (1992). Gross job creation, gross job destruction, and employment reallocation. *Quarterly Journal of Economics*, *107*, 819–863.

Davis, S. J., Haltiwanger, J. C., & Schuh, S. (1996). *Job creation and destruction*. Cambridge, MA: The MIT Press.

Dresang, J. (2002). Firm's relocation unseated a governor. In *Milwaukee Journal Sentinel* on October 13, available online at http://www.jsonline.com/bym/news/oct02/87385.asp (viewed June 1, 2005).

Dunne, T., Klimek, S., & Roberts, M. J. (2005). *Entry and exit in geographic markets*. Unpublished manuscript, University of Oklahoma.

Dunne, T., Roberts, M., & Samuelson, L. (1989a). Plant turnover and gross employment flows in the U.S. manufacturing sector. *Journal of Labor Economics*, *7*, 48–71.

Dunne, T., Roberts, M., & Samuelson, L. (1989b). The growth and failures of U.S. manufacturing plants. *Quarterly Journal of Economics*, *104*, 671–698.

Foster, L. (2003). *The Appalachian economy, establishment and employment dynamics, 1982–1997: Evidence from the Longitudinal Business Database*. Unpublished manuscript, Center for Economic Studies, U.S. Bureau of the Census.

Jarmin, R. S., Klimek, S. D., & Miranda, J. (2005). *The evolution of retail markets in metropolitan, micropolitan and rural regions*. Unpublished manuscript, Center for Economic Studies, U.S. Census Bureau.

Jarmin, R. S., & Miranda, J. (2002). *The Longitudinal Business Database*. Unpublished manuscript, Center for Economic Studies, U.S. Census Bureau, Suitland, Maryland.

Klier, T., & Testa, W. (2002). Location trends of large company headquarters during the 1990s. *Economic Perspectives* (pp. 12–26), Federal Reserve Bank of Chicago.

Leonard, J. S. (1987). In the wrong places at the wrong time: The extent of frictional and structural unemployment. In: K. Lang & J. S. Leonard (Eds), *Unemployment and the structure of labor markets* (pp. 141–163). New York: Blackwell.

MacDonald, J. M. (1985). Dun & Bradstreet business microdata: Research applications, and the detection and correction of errors. *Journal of Economic and Social Measurement*, *13*, 173–185.

Neumark, D., Zhang, J., & Wall, B. (2005a). *Employment dynamics and business relocation: New evidence from the national establishment time series*. NBER Working Paper 11647.

Neumark, D., Zhang, J., & Wall, B. (2005b). Are businesses fleeing the state? Interstate business relocation and employment change in California. *California Economic Policy*, *1*(4).

Roberts, T. (2003). Execs' ad to politicos: Enough. *Silicon Valley/San Jose Business Journal*, September 26, available online at http://sanjose.bizjournals.com/sanjose/stories/2003/09/29/story2.html (viewed on January 2, 2007).

Schumpeter, J. A. (1942). *Capitalism, socialism, and democracy*. New York: Harper & Row.

Spletzer, J. R. (2000). The contribution of establishment births and deaths to employment growth. *Journal of Business and Economic Statistics*, *18*, 113–126.

Tamaki, J. (2004). Battle of the state billboards. *Los Angeles Times*, October 11, p. c. 1.

U.S. Bureau of Labor Statistics. (1997). *BLS Handbook of Methods*, Department of Labor. Available online at http://www.bls.gov/opub/hom/home.htm (viewed on March 2, 2006).

U.S. Census Bureau. (2002). *The Longitudinal Employer-Household Dynamics Program: Employment dynamics estimates project versions 2.2 and 2.3*. Suitland, MD.

Vames, S. (2003). Businesses ponder leaving California: For some firms, California Isn't golden state anymore. *The Wall Street Journal Online*, August 20, available online at http://www.caltax.org/Vames-BusinessesPonderLeavingCalifornia8-20-03.pdf (viewed on January 2, 2007).

van Dijk, J., & Pellenbarg, P. H. (2000). Firm relocation decisions in The Netherlands: An ordered logit approach. *Papers in Regional Science*, *79*, 191–219.

Walls & Associates, (2003). *NETS: National Establishment Time-Series database*. Oakland, CA.

WHO PAYS FOR GENERAL TRAINING IN PRIVATE SECTOR BRITAIN?

Alison L. Booth and Mark L. Bryan

ABSTRACT

We use new training data from the British Household Panel Survey to explore the degree to which the data are consistent with the predictions of human capital theory. According to the raw data, most work-related training is general and is paid for by employers. Our fixed effects estimates reveal that employer-financed training is associated with higher wages both in the current and future firms, with some evidence that the impact in future firms is larger. These results are consistent with human capital theory with credit constraints, and with the relatively recent literature on training in imperfectly competitive labour markets.

1. INTRODUCTION

For many years it was thought that human capital theory, based on the assumption of a perfectly competitive labour market, fully explained who would pay for general training. The consensus was that any stylised facts that diverged from the predictions of this model could be explained by

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 85–123

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26003-5

imperfections such as credit market constraints. More recently, however, a number of published papers have challenged this orthodoxy. These papers show that, if the labour market is actually characterised by oligopsonistic wage-setting, some of the predictions of the human capital model are overturned. In particular, the wage returns to general training may be less than the productivity returns and firms may find it profitable to pay for training even though it is general.

In this paper we briefly summarise the main predictions of the various human capital theories for wages and cost-sharing and then confront these with important new data from using waves 8 to 10 of the British Household Panel Survey (BHPS). One of our findings is that employers do indeed pay the explicit costs of training that is general. We have several pieces of evidence for this. First, from the raw data we know that most work-related training is viewed by its recipients as general and that most is directly paid for by employers. Second, we have evidence from our wage equations that employer-financed training has a statistically significant positive impact on wages in the subsequent firm conditional on changing firm, even after controlling for unobservable heterogeneity. The fact that employers pay the direct costs of training that is transferable across employers is inconsistent with orthodox human capital theory without credit constraints. It is, however, consistent with some of the relatively recent training literature that assumes imperfectly competitive labour markets. It is also consistent with the hypothesis that firms offer credit-constrained workers binding training contracts whereby firms pay for general training and workers repay this 'loan' by receiving a post-training wage below their marginal product.[1]

The remainder of our paper is set out as follows. In Section 2 we briefly outline the hypotheses and their predictions as to who pays for general training and the returns to training (at both the training firms and at subsequent firms). In the following section we describe the data source and the novel features of the training questions. We also investigate the association between training characteristics and training types. In Section 4 we present and interpret our estimates of the impact of the various forms of training on wages. The final section concludes.

2. HYPOTHESES

According to standard human capital theory, workers in perfectly competitive labour markets will pay for general work-related training by receiving low training wages. They will reap the returns to this investment by receiving

Table 1. Predictions of Human Capital Theory.

Row No.	Model	Who Pays	Divergence between Wages (w) and Marginal Productivity (MP) at Training Firm	Transferability of Training
[1]	Perfect competition, general training	Worker	None	Fully transferable
[2]	As above but with credit constraints	Sharing	w > MP during training and w < MP after training	Transferable but wage returns elsewhere greater than returns at firm providing training
[3]	Perfect competition, specific training	Sharing	w > MP during training and w < MP after training	Non-transferable
[4]	Perfect competition, mix of general and specific training	Sharing	w > MP during training and w < MP after training	Partially transferable; wage returns elsewhere less than returns at firm providing training
[5]	Oligopsonistic labour market, general training	Firm	w < MP during and after training, implying rents for the firm	Fully transferable, wage returns elsewhere greater than or equal to returns at firm providing training

higher wages afterwards and their post-training wages will be the same across firms, ceteris paribus (Becker, 1964). These predictions are summarised in the first row of Table 1.

Workers who cannot afford to accept low wages during general training will be adversely affected by any credit market constraints that disbar them from borrowing to finance their investment. However, should the firm be willing to act as lender, it can pay workers more than their marginal product during training and less afterwards. The firm would only agree to such a contract if some mechanism can be devised to bind workers to the firm post-training until the loan has been paid back. A binding contract – such as an apprenticeship contract or a minimum employment guarantee – is one means of doing so. The predictions of this model are that firms will pay for general training, and workers' wages will be above marginal productivity during training and below marginal productivity after training. The magnitude of this wedge will reflect the degree of cost-sharing. The training will be transferable across firms, and after changing employers workers should get a greater return to their training than they received in the firm that provided the training and the loan. These predictions are summarised in the second row of Table 1.

Now consider the predictions of the pure specific training model. For specific human capital, it is efficient for the firm and the worker to share both the costs and the returns of the training investment (Kuratani, 1973; Hashimoto, 1981). Consequently workers' wages will be above marginal productivity during training and below marginal productivity after training, and the magnitude of this wedge will reflect the degree of cost-sharing. The training will not be transferable across firms by definition (in contrast to the model of credit constrained workers seeking general training outlined above). These predictions are summarised in the third row of Table 1.

Now suppose that, while the labour market is perfectly competitive, training comprises a mix of general and specific components. Here workers will finance their general training and firms will share the costs of the specific training. Since there will be some sharing of costs, wages at the training firm will be greater than productivity during training and less than productivity after training. Wages at subsequent firms will reflect returns only to the general component of training, and consequently will be less than wages at the training firm (in which there is some return to the worker to the shared investment in specific training). These predictions are summarised in the fourth row of Table 1.

Now consider a labour market characterised by oligopsonistic wage-setting, as in the 'new' training literature.[2] It can be shown that the wage 'compression' associated with imperfectly competitive labour markets may increase the incentive for firms to invest in general training, provided that post-training productivity is increasing in training intensity at a faster rate than are wages (Stevens, 1994; Acemoglu & Pischke, 1999b; Booth & Zoega, 2004). However, the amount of training provided in equilibrium may be sub-optimal from the viewpoint of society. The predictions of this model are that the firm may finance general training, and that the wages at the training firm will be less than marginal product. According to the contracting model of Loewenstein and Spletzer (1998) there may be a greater wage return to training in future firms than in the current firm depending on whether or not a minimum wage guarantee binds in the current job. If it does bind, the employer can extract rents from providing general training. According to other models – see for example, Stevens (1994) and Acemoglu and Pischke (1999b) – the wage returns to training will be the same across all firms in that sector. These predictions are summarised in the fifth row of Table 1.

Next we consider the impact of asymmetric information on the predictions of the orthodox human capital model. Asymmetry of information about the value of firm-provided training (for example, where the firm providing general training knows its value but other firms do not) can affect the transferability of training in an otherwise competitive labour market with identical ability workers. If outside firms assign a value of zero to the training – as they might if they have no information – such training is in effect specific to the training firm. Consequently the firm may be willing to share in the costs of its provision and the pay returns in other firms will be non-existent or small. The predictions of this model are as for Row [3] of Table 1. However, a formal qualification associated with a training course is a means of conveying to outsiders the value of the employer-provided general training. For this reason one would expect accredited training to have a larger impact on wages in future firms than non-accredited training ceteris paribus. One would also expect it to be financed by the individual, since it is transferable or general. The predictions of the model with accreditations for training are therefore the same as for Row [1] of Table 1 – the individual will pay and will get all the pay returns.

An alternative approach is found in the asymmetric information model of Acemoglu and Pischke (1998), in which workers are characterised by heterogeneous abilities. Here training is rewarded more in the current firm than in outside firms because the current firm will pay higher wages to retain high ability workers, whereas low-ability workers will be dismissed. Some of the high ability workers who need to leave their jobs will be treated as low ability workers in the outside market. Since training and ability are complements, training will be valued less for workers who have been laid off or who have quit. Consequently in the outside market these workers will receive lower returns to their training. The predictions of this model are as for Row [4].

In this paper we attempt to distinguish between these hypotheses by (i) using information about who pays directly for work-related training, and (ii) comparing the pay returns to such training at the current and subsequent firms. Of course, the predictions of some of these hypotheses are observationally equivalent. Two models predict that transferable training might have bigger returns to subsequent firms than to the training firm, as inspection of Rows [2] and [5] shows. On the other hand, some predictions are quite distinct. For example, while both models in Rows [1] and [5] predict that training is transferable, the first predicts workers pay for it while the fifth predicts firms do.

Our data have a number of advantages, compared to other surveys, in allowing us to address these predictions. This will be shown in the following section, where we describe the data and provide a picture of the various forms of work-related training that take place in Britain.

3. THE DATA

The BHPS is a nationally representative random-sample survey of private households in Britain. Although information on work-related training was collected in the first seven waves, it was fairly limited, and focused on training receipt, type and total duration in the previous year.[3] However, the training questions were expanded from Wave 8 – conducted in 1998 – onwards. Respondents are now asked how many training schemes or courses they started in the past year, and detailed information is then collected on the three longest events (or all events if there were fewer than three). These data shed new light on the nature of each event as a human capital investment. First, we know the duration (what we term intensity) of each event. Second, we know its type or purpose (defined by the same categories as in previous waves) and, third, where it took place. Fourth, individuals are asked how the event was financed, enabling us to identify who pays the explicit costs. Finally, we know whether or not the event led to a qualification. We do not, however, know the date at which the training event occurred within a given year, or the wages an individual received.

We use these new data from Waves 8 to 10 of the BHPS for individuals who are either original members of the panel or who joined the panel subsequently. Our sample consists of private sector full-time employees aged between 16 and 65 years with valid information on our main variables and who did not report more than a calendar year of training.[4] Our analysis covers any training (whether employer-provided or not) received by individuals, and excludes spells of full time education (only 2.1% of our final sample had undergone any full time education in the previous year). Respondents were specifically asked to exclude leisure courses. We drop observations where there is missing information on the place, type, duration or financing of the training event, or on whether or not it led to a qualification (330 training events were dropped). This leaves us with 8,316 person-years, for which training was received in 2,575 (or 31% of) cases.

The precise form of the new training questions is given in Appendix A. Individuals were asked to report the total number of training courses/events in the past 12 months, and then questioned in detail about the three longest. In total 5,272 events were reported and we have details of the 4,317 longest (so there is some truncation of the training history data).[5] Table 2 shows

Table 2. Attributes of Work-Related Training, BHPS 1998–2000.

	All [1]	Men [2]	Women [3]
A. Any training			
1998	0.307	0.296	0.326
1999	0.303	0.297	0.316
2000	0.318	0.307	0.339
B. Number of training events			
Total number reported	2.050	2.067	2.021
Number where data collected	1.677	1.666	1.694
C. Training type:			
Induction (help get started in current job)	0.123	0.122	0.126
Current skills (improve skills in current job)	0.864	0.868	0.859
Future skills (to prepare for future jobs)	0.589	0.587	0.592
General skills (to develop general skills)	0.845	0.848	0.842
D. Location[a]			
Workplace	0.363	0.362	0.364
Employer training centre	0.173	0.173	0.172
Private training centre	0.198	0.222	0.160
College	0.173	0.153	0.205
Home	0.030	0.029	0.031
Other	0.063	0.061	0.068
E. Financing method			
None (no fees)	0.267	0.272	0.260
Self (respondent or family paid)	0.088	0.076	0.107
Employer	0.617	0.624	0.605
Other (including New Deal and TEC)	0.037	0.038	0.036
F. Accreditation			
Proportion qualified	0.420	0.405	0.443
Proportion accredited	0.225	0.222	0.229
G. Training intensity (days)	12.64	12.37	13.07
N	8,316	5,379	2,937

Notes: [a] See Table A2, Appendix A for a breakdown of training location by training type.

that, for each of the available waves, roughly 30% of individuals received training. The (conditional) mean number of reported training events is 2.05. We also know (but do not show in the table) that over half those receiving training (52%) experienced one event only, 24% participated in two events, 12% in three events and 12% in more than three events. If we exclude those events for which detailed information was not collected, the conditional mean number of events is 1.68. Note that we present in Table 2 descriptive statistics for our entire sample as well as for the sample disaggregated into men and women. In our econometric specification, we test the hypothesis that observations on men and women can be pooled, as discussed in Section 4 below.

3.1. Training Type

What *types of training* do individuals report? Respondents are asked to specify the purpose of each event experienced in the past 12 months, using 5 non-mutually exclusive categories: (i) to help them get started in the current job, (ii) to increase their skills in the current job, (iii) to improve their skills in the current job, (iv) to prepare for future job(s), and (v) to develop general skills.[6] We redefine the first category as *induction training*.[7] Since it is difficult to see any distinction between categories (ii) and (iii) other than differences of interpretation, we combine training to increase/improve skills in the current job into a single type – *skills in the current job*. Panel C of Table 2 shows the proportions in each of the four categories.[8] Unsurprisingly induction training is relatively infrequent, being reported for only 12% of events. Training events are viewed as increasing/improving current skills in nearly 85% of cases and future skills in 59% of cases. Some 85% of events are viewed as improving general skills. There is comparatively little variation in these figures across gender.

There are several problems to be considered before using these data. First, there is some overlap of the training categories; in particular for training for *current job skills* and *general skills*, where the correlation coefficient of the two indicator variables is 0.75. Only 5.7% of events are described as general skills training only (i.e. with no other categories cited). So it is not possible to construct meaningful separate variables for each of these types, since respondents typically view their training as falling into a number of different categories, as Table A1 in Appendix A reveals.[9] For this reason we drop the separate general training indicator in our subsequent analyses, although we would remind the reader that 85.4% of *training for current job skills* is viewed by respondents as general.

A second potential problem relates to respondents' interpretation of the question. Campanelli et al. (1994) note, from a study of both linguistic and survey data, that the interpretation of the term "training" varies across groups in the population, in particular employers, employees, and training researchers.[10] They emphasise that individuals in the general population typically interpret training as referring to "that which happens in formal courses". This is our focus of interest in the present study, rather than on less formal training that is harder to measure. Inspection of the BHPS questions provided in Appendix A reveals that the training data elicited in the BHPS is of this more formal nature.

3.2. Training Location

Panel D of Table 2 reports the proportions of training events taking place at different locations, of which we distinguish six – the current or former workplace;[11] the employer's training centre; a private training centre; a higher or further education college, adult education centre or university; at home; and other unspecified locations. Some 36% of training takes place in the workplace, and a further 37% in a training centre (either employer-based or private), while 17% is college-based. Women are less likely to train in private training centres and more likely to train in college. Cross-tabulations of training type and location, reported in Appendix A Table A2, show that there is little difference between the location patterns of induction and current job skills training, with nearly 80% taking place in workplaces or training centres and 14% in colleges. On the other hand, training for future job skills is less likely to occur in workplaces or in training centres (70%) while nearly a quarter takes place at college (21%) or at home (3.6%).

3.3. Training Finance

Panel E of Table 2 shows four non-mutually exclusive categories for financing of training: no fees; the respondent or their family paid; the employer or future employer paid; or it was financed in some other way. Because of the small number of cases of training financed by schemes such as the New Deal and Training for Work, we combine them with the residual category of 'other' training. The raw figures indicate that women are half as likely again as men to finance their own training (10.7% of their courses are self-financed compared to 7.6% for men). For both genders, the employer is reported as financing just over 60% of events.

The substantial proportion of individuals reporting *no fees* is interesting and may suggest economic naivety on the part of respondents, since it is unlikely that any training activity is truly costless. At a minimum there will be some loss of production while individuals are in training (in the absence of pure learning-by-doing, which is anyway not captured in the BHPS training questions). It seems likely that individuals not paying for training themselves, and who see no evidence of the employer paying, may report that no-one pays. In Table A3 of Appendix A evidence is presented that individuals tend to report no fees when training is internal to the employing organisation and that in fact the costs are borne by the employer. In our multivariate analysis we therefore combine the *no fees* and *employer finance* categories, which together account for nearly 90% of training finance.[12]

In Table A4 of Appendix A we report cross-tabulations of financing method for induction, current job skills and future job skills training. Induction and current job skills training are financed in a similar fashion – mainly by the employer (some 90%, including 'no fees'). For future job skills training the balance is marginally tilted away from employer finance (85%) towards self financing (11%).

One implication of human capital theory is that the firm and workers may share the cost of specific training. If employees do not understand the idea that cost sharing can take the form of a lower wage during training, these raw data will only reveal any direct sharing. Table A5 in Appendix A shows that there is very little sharing; for example, for those events paid for by the individual or their family, only 3.4% were also financed by the employer. To test for the other possibility – that wages are lower during training – we also perform a test below using our multivariate wage model, of whether training to be received in the future is associated with lower wages. We find no evidence that wages are lower during training spells using this indirect method (see footnote 21).

3.4. *Qualifications and Accredited Training*

We noted in Section 2 that, where there is asymmetry of information about the value of firm-provided training, the award of a qualification upon completion of a training course may signal to alternative employers the value, and verify the receipt, of newly acquired human capital. In the BHPS questionnaire, individuals were asked if each training event was intended to lead to a qualification, and if so whether any of a list of recognised qualifications had actually been obtained in the previous year. This latter measure of accredited training will, of course, be subject to right censoring when training is in progress but the qualification not yet obtained. Panel F of Table 2 shows that women are more likely to undertake training leading to qualifications than men. Table A4 of Appendix A shows that 17% of accredited training courses are self-financed, compared to only 3% of courses, which do not lead to qualifications. Nevertheless, it is striking that 78% of accredited courses are still paid for by the employer (including the 'no fees' category).

3.5. Correlation of Training Measures

As a means of parsimoniously further describing the data, we report in Table 3 the marginal effects from some simple conditional probits. These estimates show the ceteris paribus association between the characteristics discussed above and the type of training (conditional on training being received). The marginal effects show the increase in the expected probability relative to the base case of training financed by a scheme or 'other' means,

Table 3. Training Types and Other Training Characteristics – Probit Analysis.

	Induction Training		Current Skills Training		Future Skills Training	
	Men (1)	Women (2)	Men (3)	Women (4)	Men (5)	Women (6)
Workplace	0.096***	0.013	0.052**	−0.009	−0.016	0.023
	(2.75)	(0.37)	(2.08)	(0.25)	(0.39)	(0.44)
Employer training centre	0.105***	0.032	0.032	−0.010	0.010	0.047
	(2.63)	(0.80)	(1.19)	(0.24)	(0.22)	(0.86)
Private training centre	0.057	0.055	0.055**	−0.041	0.024	−0.013
	(1.57)	(1.32)	(2.21)	(0.98)	(0.55)	(0.23)
College	0.050	0.023	0.006	−0.046	0.065	0.045
	(1.27)	(0.57)	(0.23)	(1.11)	(1.35)	(0.78)
Home	0.031	−0.074	−0.023	0.017	0.140**	−0.091
	(0.58)	(1.30)	(0.56)	(0.31)	(1.97)	(1.03)
Employer paid	−0.043	−0.046	0.107***	0.132***	−0.072	−0.202***
	(1.24)	(1.01)	(3.11)	(2.84)	(1.34)	(2.95)
Self/family paid	−0.062*	−0.081**	−0.166***	−0.110**	−0.015	−0.126
	(1.95)	(2.12)	(3.95)	(2.24)	(0.23)	(1.51)
Accredited	0.065***	0.007	−0.052***	−0.045**	0.137***	0.186***
	(4.53)	(0.37)	(3.62)	(2.40)	(6.37)	(6.82)
Observations	2,689	1,628	2,689	1,628	2,689	1,628
R^2	0.02	0.01	0.12	0.10	0.03	0.04

Notes: (1) Asymptotic *z*-statistics in brackets. (2) *Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level.

occurring in an 'other' place, and not leading to qualifications. Thus for men, if training takes place in the workplace or at an employer's training centre it is 10 percentage points more likely to include an element of induction than training in the base category of another place. For men induction training is also positively associated with gaining qualifications. For women, on the other hand, induction training typically does not follow a systematic pattern, with the exception of being less likely to be self-financed. (Of course very little induction training, 5%, is self financed.) For both men and women, current skills training is strongly associated with employer financing (an 11–13 percentage point increase) and negatively associated with self-financing (an 11–16 percentage point decrease). Training which leads to qualifications is much more likely to be for future skills (14 percentage points for men, 19 percentage points for women) and less likely to be for current skills.

3.6. Training Intensity

Training intensity is reported in Panel G of Table 2. Respondents are asked for the total time, in hours, days, weeks or months, devoted to training since 1st September of the previous fieldwork year. We converted the responses to days. Mean intensity for men is 12.4 days and for women 13.1 days. Fig. A1 in Appendix A graphs the distribution of intensity (conditional on receiving training and truncated at 30 days) for the three types of training.

Table 4 reports the estimates from simple regressions of intensity on the various training characteristics.[13] The base category is purely 'general' training not leading to qualifications, financed by a scheme or 'other' means, and occurring in an 'other' place. Compared to this base, induction training is ceteris paribus associated with a statistically significant increase in expected intensity of nearly 8 days for both men and women. Similarly, training for future skills is associated with a 3–4 day increase in intensity, while training for current skills does not significantly affect intensity relative to the base case. Turning to training location, both college and home training are associated with much higher intensity (14–29 days). However, it should be stressed that training at home accounts for only 3% of events. For men the finance dummies both have large, negative and significant coefficients, reflecting the relatively long duration of training financed by official schemes (in the base case), though the effect is not particularly evident for women. Finally, training which leads to qualifications is associated with about a 13-day increase for men and women.

Table 4. Training Intensity (Days) and Training Characteristics – OLS Estimates.

	Men	Women
Induction	7.872***	7.618***
	(4.43)	(3.17)
Current skills	−1.689	−0.093
	(0.93)	(0.04)
Future skills	3.245***	4.089**
	(2.70)	(2.45)
Workplace	0.564	3.715
	(0.22)	(1.12)
Employer training centre	−1.757	0.355
	(0.65)	(0.10)
Private training centre	−3.996	1.504
	(1.52)	(0.42)
College	14.793***	14.362***
	(5.09)	(3.89)
Home	16.244***	29.720***
	(3.84)	(5.40)
Employer paid	−20.080***	−7.484*
	(6.42)	(1.73)
Self/family paid	−19.363***	−0.326
	(5.32)	(0.07)
Accredited	12.201***	13.541***
	(9.24)	(7.53)
Constant	23.653***	4.747
	(5.81)	(0.88)
Observations	2,689	1,628
R^2	0.15	0.15

Notes: (1) *t*-statistics in brackets. (2) *significant at 10% level; **significant at 5% level; ***significant at 1% level.

3.7. Summary

The picture of work-related training emerging from our descriptive analysis of the raw data is as follows. Most recipients view it as general, and it takes place at the workplace or a training centre. Small proportions of training events are for induction purposes, or take place at college or at home. However, they are among the longest events. The direct costs of most training events are paid by employers.[14] About 40% of training events lead to qualifications and these events tend to be longer.

We now turn to our multivariate fixed effects estimates of the impact of training on wages. We focus in particular on employer-financed training to increase or improve skills in the current job, and investigate the transferability of such training across employers. We also investigate the degree to which accredited training is more transferable than non-accredited training.

4. WAGE LEVELS

The new training data enable us to test the predictions of various different models of training and the labour market that are summarised in Table 1. The descriptive statistics presented in the previous section already allow a partial evaluation of these models. They indicate that a large majority of training (85%) is regarded by recipients as general, that an even larger proportion (89%) is viewed as either employer-financed or entailing no fees, and that there is almost no explicit cost sharing.

These figures cast some doubt on the predictions of the simple human capital model that assumes a competitive labour market. We further explore the implications of the different models by investigating, in a multivariate framework, the impact on wages of training incidence, the number of training events, and training intensity. Most studies simply examine the impact of training incidence (and sometimes intensity) on wages, but not the number of events.[15]

4.1. Empirical Model

Suppose the hourly wage is determined by:

$$w_{ijt} = x'_{ijt}\beta + T'_{it}\alpha + D'_t\gamma + \mu_i + \upsilon_{ij} + \varepsilon_{ijt} \tag{1}$$

where w_{ijt} is the natural logarithm of the real (1998 prices) hourly wage of individual i in job j at time t; x_{ijt} is a vector of individual and firm characteristics influencing the wage and associated with parameter vector β; T_{it} is a vector containing various measures of the amount of training accumulated from the start of the sample period in Wave 8, and is associated with parameter vector α; and D_t are year-specific dummy variables with associated parameter vector γ. Unobservable characteristics which affect the individual's wage are decomposed into a permanent effect μ_i, an employer match specific component υ_{ij} and a transitory effect ε_{ijt}. As we discuss below, equation (1) is estimated as a fixed effects (FE) model, where μ_i is treated as the fixed effect and υ_{ij} is approximated by dummy variables. Theoretical priors, as well as initial testing, supported the FE approach.[16]

We divide the variables into training undertaken with the *current* employer (including training received in the past year) and training undertaken with *previous* employers. Including training accumulated with *previous* employers allows us to test the joint hypothesis of no depreciation, constant returns and the transferability of such training across employers.[17] The general human capital model typically ignores potential skills obsolescence and assumes constant returns for tractability, but in any empirical implementation it is worth keeping these assumptions in mind.[18] We return to them below when considering the functional form that training should take in the wage equation. In an extension, we also estimate a version of the equation, which controls for the time of job change. This checks whether the effects being assigned to training in the previous job may actually be picking up differential effects from older training.

In our specification, we separately include the three types of relevant training – employer-financed current job skills training, self-financed current job skills training and a residual category of other forms of training. We focus on current job skills training because we are primarily interested in skills investments intended to have a direct effect on productivity and in the subsequent portability of these skills. For brevity, in the text below we drop the 'current job skills' qualification and simply distinguish this training by its finance method and when it took place (current or previous employer).

In another specification we distinguish between training that leads to qualifications (which we term accredited training) and training that does not. We do this because – as noted in Section 2 – a formal qualification associated with a training course is a means of conveying to outsiders the value of the employer-provided general training. In the BHPS data, we know if each training event leads to a qualification, and if so, whether it has been obtained at the interview date. However, there is unfortunately no follow-up information on accredited training for which the qualification has yet to be acquired at the interview date. Since we do not know whether these events are ongoing or ended in either success or failure, we risk misclassifying them. If we classify them as accredited training, then the category will include longer, ongoing spells of training which did lead to qualifications, but also events which ultimately ended in failure. On the other hand, if we classify them as non-accredited training, then our accredited training category only includes completed accredited events, but in the non-accredited category we conflate some events which did eventually lead to qualifications with events never intended to. There are merits in both approaches, but we chose to include all training leading to qualifications, whether or not obtained, in the accredited category. We experimented with the alternative classification and the results were qualitatively the same.

The difficulty in separating incomplete from complete spells of training also means we cannot *directly* test one prediction of the orthodox human capital model, namely that wages are lower during training (although we do investigate this indirectly, as reported below). The prediction has been contested by Bishop (1997).[19] Moreover Loewenstein and Spletzer (1998) find little evidence in favour of it using NLSY data. In addition a severe selectivity problem arises in testing this, since unobservable ability is likely to raise starting wages and affect the amount of training provided. Ideally, one would like to observe two identical individuals starting the same job and receiving different amounts of training. Sicilian (2001) presents a partial solution to the problem using US data on pairs of individuals starting the same or similar jobs, and finds that training does reduce starting wages for his sample.

We performed a rough test of the prediction that wages are lower during training by including training to be received in the *next* year in our current wage equations. While the fixed-effect framework solves the problem of time-invariant unobserved ability, our test does assume that the current wage is the same as that to be paid during the following year's training spell, which could take place any time between one and twelve months after the current wage is observed.[20] Our test also assumes that next year's training is not a response to a current unobserved productivity shock (which would not be removed by the estimation procedure). Notwithstanding these caveats, if wages are lower during training, the coefficient on the additional regressor – next period's training – should be negative. In fact the estimate was positive though insignificant, suggesting that wages are not lowered during training.[21] Thus, for our sample, employees do not appear to be paying for training indirectly through reduced earnings.

4.2. Endogeneity of Training

We now consider in detail how the unobservable components of Eq. (1) may also be related to training receipt, and therefore how the measured returns to training may be biased by, for example returns to mobility. The unobserved individual-specific effect μ_i may capture some aspect of individual ability or motivation that reduces the cost of training and which will therefore be positively associated with the receipt of training. Furthermore, μ_i will also reflect the stock of training acquired in the pre-sample period. If past and future training are correlated (either negatively or positively) then T_{it} will be correlated with μ_i (see Loewenstein & Spletzer, 1998, p. 160). We can eliminate this source of bias by specifying (1) as a fixed-effects (FE) model, so that identification uses only within-individual variation and all time-invariant effects are removed.

The employer-match effect υ_{ij} is also likely to be correlated with variables in T_{it}. First, where the employer-match is good, more training may occur since expected tenure will be longer. We might expect this to bias upward the estimated returns to training in the current job, but not returns in the previous job. Second, the measures of training obtained with the current and previous employers and υ_{ij} change whenever the employer changes (indeed, the previous job training measure alters *only* when the individual changes employer). If job mobility is non-random – for example if workers move only when they get a better wage offer – the measured returns to training will be biased by an amount reflecting the average return to mobility among movers.

While we do not observe υ_{ij}, which would enable us to control fully for these effects, we can approximate υ_{ij} by an employer-specific effect υ_j that is constant across individuals. This will control for the average effects of job mobility (or increase in match quality) amongst individuals who move, and a priori we expect these effects to be positive. Bias will remain insofar as individuals who receive more training than the average also tend to experience higher than average returns to mobility. This might be the case if, for example, highly trained workers were better at job search, or purchased more information about higher paying jobs, or had lower mobility costs than workers with less training. While we suspect this is unlikely to be a major cause of bias, we are unable to explore this any further with our available data.[22]

We model υ_j by including a step dummy variable taking the value one throughout the duration of a new job (if an individual changes jobs), and zero otherwise; and another similar dummy capturing a second new job (since a maximum of two job changes can be observed). The base case is the first job observed in the panel. This is similar to the approach of Loewenstein and Spletzer (1998) in their analysis of the returns to training between jobs.

A final potential source of bias is the possibility that individuals with high wage growth (as reflected in a high ε_{ijt}), even after controlling for job changes and all other observables, get more training. While noting this possibility, we might expect it to affect the return to current and previous job training in the same way, and much of our discussion below turns on the difference between these coefficients. An alternative to our methods would be to control for endogeneity by instrumenting training. However, we would require as many instruments as there are training variables (a minimum of five – the most aggregated specification). Since it is difficult to suggest even one variable, which is correlated with training but not wages, in practice identification would be through functional form assumptions alone.

4.3. Functional form of training

The effect of training on wages may depend on how much training has already been received (for example, there might be diminishing returns to training). It may only be the first event that matters for wages; at the other extreme, all training events may have the same impact (see Arulampalam & Booth, 2001). Similarly, the return to training may or may not be in proportion to the length of the event. Furthermore the skills acquired may depreciate (particularly in a period of rapid technological change). We therefore investigated the appropriate functional form for the cumulative training measures T_{it} by estimating an equation for wage growth between Waves 8 and 10 in which the training received in each wave was entered separately.[23] We tried three alternative measures of training – incidence, event counts and total intensity per wave – with counts and intensity entering linearly, quadratically, and as logs and square roots. We were unfortunately unable to obtain robust results that clearly distinguished between the different models, apparently for two reasons. First, the cell sizes of the disaggregated training measures were rather small.[24] Second, measurement error in the training variables is likely be more important when they are disaggregated, resulting in more downward bias of their coefficient estimates.

Because of the difficulty in determining functional form with any precision, we simply define the elements of T_{it} as the cumulative total of training received since Wave 8 according to the three different measures: incidence, event counts and intensity. Assuming that measurement error is uncorrelated over training events and years, while training is positively correlated, the cumulative variable will be a cleaner indicator of training received: the summing process increases the share of the variance accounted for by the true training indicator, which is reinforced, and decreases the share due to measurement error, which, intuitively, tends to cancel out.[25] The coefficient estimates should therefore be subject to less downward bias. We estimate separate equations for incidence, event counts and intensity. The results reported below are qualitatively similar for all three measures, suggesting some robustness to possible mis-specification of functional form.

As already noted, permanent individual-specific effects on wages are removed during estimation. These effects also include any time-invariant unobservables affecting female participation in the labour market. Insofar as this endogenous selection is determined only by permanent unobservables, selection bias in the estimates is also eliminated. Since the FE model controls for time-invariant individual-specific heterogeneity affecting wages, we tested the hypothesis that observations on men and women could be pooled in estimation of (1). The test yielded an F statistic with an implied p-value of 0.36. Therefore (1) was estimated on the pooled sample of men and women.

4.4. Training and Employer Changes

Table 5 reports the number of individuals (men and women combined) receiving each type of training for our sample of individuals with valid information on all variables in Eq. (1). The first row shows the number who had any training at their different employers. So 1269 individuals received employer-financed training at some point during the three waves, 127 undertook self-financed training and 425 experienced other training. The second row shows the number observed to undertake training and then change employers: 152 for employer-financed training, 12 for self-financed training and 60 for other training. In order to identify the effect of training in previous jobs it is clearly important to have reasonable numbers of such observations.[26] This is not the case for self-financed training, and we therefore combine the current and previous employer event counts into a single category.

Table 5. Individual Receipt of Training.

	Employer-Paid	Self-Paid	Other
(a) Number of individuals receiving training (at any point over sample period)			
Any training – all employers	1,269	127	425
Any training – previous employers	152	12	60
Accredited training – all employers	629	109	283
Accredited training – previous employers	78	11	44
Non-accredited training – all employers	877	22	169
Non-accredited training – previous employers	88	1	20
(b) Accumulated training (conditional on any training event over sample period)			
Mean accumulated events	2.41	1.36	1.42
Mean accumulated intensity (days)	22.7	35.3	29.4

Note: The total number of individuals in the sample is 3,333.

The third to the sixth rows distinguish accredited and non-accredited training within the all and previous employer categories. For our training type of primary interest – employer-financed training – we have observations on over 70 individuals in both of the previous employer cells. But for other training, the disaggregated number of observations with previous employers is quite low (44 individuals experienced accredited training and 20 non-accredited training). In our estimation, we therefore maintain the aggregate category, distinguishing only between previous and current employer. We also keep the combined self-financed training category, as above. The relatively small cell sizes should be borne in mind when interpreting our more disaggregated results. The penultimate row in the table – in panel (b) – reports the mean number of events accumulated by individuals over the sample period, conditional on receiving at least one such event. The final row shows the conditional mean accumulated intensity. So individuals who received any employer-financed training received on average 2.4 events, which lasted a total of 22.7 days.

4.5. Wages and Training Incidence and Event Counts

Table 6 reports the key coefficient estimates of the fixed effect model when training is measured by incidence and event counts. Table C1 of Appendix C reports estimates of the remaining coefficients, including year dummies for waves 9 and 10, and using the specification reported in Column (2) of Table 6. These estimates are similar across specifications. Definitions of the variables are given in Appendix B.[27]

Columns (1) and (2) present the estimated coefficients when the incidence and count variables are included separately.[28] Both sets of results tell a similar story. Employer-financed training received with former employers raises current wages more than training undertaken with the current employer, and both effects are statistically significant. Having received any employer-financed training with previous employers is associated with 7.8% higher expected wages subsequently, whereas incidence of training with the current employer is associated with only 2.4% higher expected wages. The difference in impact is significant at the 5% confidence level.[29] The results from both specifications also indicate that training in the residual category ("other training") increases wages but only when it was received with previous employers. There is no evidence that self-financed training has any effect on wages.

Column (3) reports the estimates when employer-financed training is disaggregated according to whether it is accredited or not, and is entered into

Table 6. The Effect of Training on Wages: Fixed Effects Model.

	Incidence (1)	Counts (2)	Counts (3)	Intensity (4)	Intensity (5)
Employer financed, current skills training					
Current employer	0.0240**	0.0104***		0.0005***	
	(2.28)	(2.69)		(2.75)	
Accredited			0.0191**		0.0005***
			(2.55)		(2.74)
Non-accredited			0.0075*		0.0002
			(1.66)		(0.38)
Previous employer	0.0779***	0.0243**		0.0012**	
	(3.38)	(2.41)		(2.34)	
Accredited			0.0529***		0.0015***
			(2.91)		(2.70)
Non-accredited			0.0115		−0.0028
			(0.94)		(1.21)
Self-financed, current skills training					
Current and previous	0.0245	0.0148	0.0142	0.0005	0.0005
Employers	(0.80)	(0.72)	(0.69)	(1.10)	(1.10)
Other training					
Current employer	0.0227	0.0190*	0.0189*	0.0001	0.0001
	(1.33)	(1.77)	(1.76)	(0.44)	(0.42)
Previous employer	0.0759**	0.0408**	0.0396**	0.0009*	0.0009*
	(2.22)	(2.08)	(2.02)	(1.94)	(1.88)
Employer match 1	0.0170	0.0295**	0.0287**	0.0322***	0.0348***
	(1.26)	(2.24)	(2.18)	(2.64)	(2.83)
Employer match 2	0.0470*	0.0660***	0.0646***	0.0716***	0.0765***
	(1.92)	(2.73)	(2.67)	(3.15)	(3.33)
Observations	7,167	7,167	7,167	7,167	7,167
Number of individuals	3,333	3,333	3,333	3,333	3,333
R^2 – within	0.15	0.15	0.15	0.15	0.15
R^2 – between	0.07	0.07	0.06	0.06	0.06
R^2 – overall	0.05	0.05	0.05	0.05	0.05

Notes: (1) *t*-statistics in parentheses (2) * Significant at 10%; ** Significant at 5%; *** Significant at 1%. (3) Other controls: experience and experience squared, tenure and tenure squared, local unemployment rate and dummies for charity sector, 1 digit industry, region, marital status, firm size, fixed and temporary contracts, trade union coverage, highest educational qualification lagged one year, 1 digit occupation and year dummies. See also Table C1.

the equation as event counts. The results indicate that only accredited training has a statistically significant effect at the 5% level. Again the point estimate for training acquired with previous employers is substantially larger than for that received with the current employer (the difference is not quite statistically significant at the 5% confidence level). An additional accredited training event with a previous employer raises wages by 5.3%, whereas a similar event with the current employer raises wages by 1.9%.

The estimated coefficients on the ‘‘employer match’’ dummy variables indicate that an employer change is generally associated with an improved unobserved match of 2.5–3.0%. It was argued earlier that this match component is likely to be correlated with the measures of previous employer training. As a check, we therefore re-estimated the equations omitting the two match dummies. The coefficients on the previous employer training variables were larger and much more precisely estimated. For example, an event of previous accredited training is expected to increase wages by 6.4% ($t = 3.62$), compared to 5.3% ($t = 2.91$) when the dummies are included (reported in Table 6). As expected, controlling for job mobility therefore appears to matter when estimating the returns to training with previous employers.

4.6. Wages and Training Intensity

The estimates of the model when intensity (in days) is used as the training measure are reported in columns (4) and (5) and show a similar pattern. The results in column (4) indicate that time spent training with previous employers has more than twice the effect on current wages as time spent training with the current employer (though the difference is just insignificant at the 5% confidence level). Thus a trainee undergoing the sample mean of about 12 days (per year) of training could expect to receive a wage boost of nearly 1.5% with a future employer, ceteris paribus. Column (5) shows the estimates when training is distinguished by accreditation status. One explanation of the higher return to accredited training over non-accredited training seen in column (3) is that accredited events are longer (as noted in Section 3 and illustrated in Table 4). The results in column (5) show that, even after controlling for intensity, accredited training still has a higher return (the return to non-accredited training is not statistically significant). Again, accredited training received with previous employers has a larger effect than that received with the current employer (the difference is significant at the 8% confidence level).

Next we consider a further question about our results, and that is the following. Might the larger coefficients associated with training received from previous employers actually be capturing an effect due to the age of the training – perhaps because there is a time lag before productivity rises? Our simple functional form tests did not provide any evidence of skills appreciation or depreciation. Nevertheless, as a further check, we re-estimated the equations in Table 6, interacting previous employer training with a dummy variable indicating that the job changed just before wave 9 (rather than just before wave 10). This interaction term therefore captures any additional effect from older training acquired with previous employers. Table 7 contains the results from this estimation.

None of the interaction coefficients is statistically significant and they tend to be small relative to the main effects of previous employer training. While the main effects are less statistically significant than in Table 6, in general they are of similar magnitude. For example, in column (3), an additional training event from a previous employer is associated with 4.3% ($t = 1.7$) higher wages (compared to 5.3% ($t = 2.9$) in Table 6), and the additional older training effect is estimated as 1.8% but is not significant ($t = 0.5$). In column (1) the main effect from training incidence with the previous employer is 9.3% ($t = 2.8$), compared to 7.8% ($t = 3.4$) in Table 6, and the additional effect from older training is actually negative (–2.5%) but again not statistically significant ($t = 0.7$). From these results, we conclude that our evidence of higher returns from training with previous employers is not an artefact of the age of this training.

5. DISCUSSION

Our results suggest that training explicitly financed by the employer is associated with higher wages in the current firm.[30] However, we also find that such employer-financed training received with previous employers has a statistically significant positive impact on wages paid by the current employer, even after controlling for unobservable heterogeneity and the average returns to job mobility.[31] Furthermore, this effect appears larger than the impact of training at the present employer on current wages.

Our findings that most training is viewed by its recipients as general need not, of course, be taken at face value. It could be that the training comprises both specific and general components and respondents have simply not perceived this. However, if that were the case we would expect – see row [4] of Table 1 – that the direct training costs would be shared by both parties and furthermore that the wage returns elsewhere would be less than the returns at the firm providing the training. This is not what the data show.

The fact the employers pay for training that is transferable across employers is inconsistent with orthodox general human capital theory, but consistent with several other hypotheses, as inspection of rows [2] and [5] of

Table 7. The Effect of Older Previous Employer Training: Fixed Effects Model.

	Incidence (1)	Counts (2)	Counts (3)	Intensity (4)	Intensity (5)
Employer financed, current skills training					
Current employer	0.0243**	0.0102***		0.0005***	
	(2.31)	(2.63)		(2.75)	
Accredited			0.0188**		0.0005***
			(2.51)		(2.71)
Non-accredited			0.0074		0.0002
			(1.63)		(0.40)
Previous employer	0.0926***	0.0176		0.0012	
	(2.82)	(1.30)		(1.64)	
Employer changed wave 9	−0.0258	0.0129		0.0001	
	(0.65)	(0.71)		(0.05)	
Accredited			0.0432*		0.0017**
			(1.72)		(2.22)
Employer changed wave 9			0.0179		−0.0004
			(0.54)		(0.38)
Non-accredited			0.0073		−0.0068**
			(0.44)		(2.03)
Employer changed wave 9			0.0087		0.0074
			(0.38)		(1.64)
Self-financed, current skills training					
Current and previous Employers	0.0240	0.0152	0.0145	0.0005	0.0005
	(0.78)	(0.74)	(0.71)	(1.10)	(1.12)
Other training					
Current employer	0.0227	0.0187*	0.0186*	0.0001	0.0001
	(1.33)	(1.74)	(1.74)	(0.44)	(0.43)
Previous employer	0.0616	0.0297	0.0313	0.0015	0.0014
	(1.08)	(0.84)	(0.89)	(0.59)	(0.56)
Employer changed wave 9	0.0230	0.0139	0.0099	−0.0006	−0.0005
	(0.35)	(0.35)	(0.25)	(0.23)	(0.21)
Employer match 1	0.0167	0.0299**	0.0291**	0.0320***	0.0352***
	(1.23)	(2.27)	(2.21)	(2.61)	(2.85)
Employer match 2	0.0484**	0.0624**	0.0614**	0.0714***	0.0735***
	(1.96)	(2.53)	(2.49)	(3.13)	(3.18)
Observations	7,167	7,167	7,167	7,167	7,167
Number of individuals	3,333	3,333	3,333	3,333	3,333
R^2 – within	0.15	0.15	0.15	0.15	0.15
R^2 – between	0.07	0.07	0.06	0.06	0.06
R^2 – overall	0.05	0.05	0.05	0.05	0.05

Notes: (1) *t*-statistics in parentheses. (2) * Significant at 10%; ** Significant at 5%; *** Significant at 1%. (3) Other controls: experience and experience squared, tenure and tenure squared, local unemployment rate and dummies for charity sector, 1 digit industry, region, marital status, firm size, fixed and temporary contracts, trade union coverage, highest educational qualification lagged one year, 1 digit occupation and year dummies. (4) The base case for the time of employer change is wave 10.

Table 1 makes clear. First consider row [2]. Our evidence that the returns to training between employers exceed the returns with the current employer is consistent with the (otherwise) perfectly competitive general human capital model with credit constraints. If workers cannot borrow freely on the capital markets and binding training contracts (such as apprenticeships) are possible, then firm financing of general training may act as a loan to workers which is repaid by setting the post-training wage below their marginal product. In this case, although the firm merely acts as a banker to workers, it still finances training while it takes place. On termination of the contract workers are free to earn their full marginal product with a rival employer.

Now consider row [5] of Table 1. Our findings that the returns to training between employers exceed the returns with the current employer are also consistent with the model of Loewenstein and Spletzer (1998), based on imperfectly competitive labour markets. In their model, training is determined within long-term contracts, including minimum wage guarantees, in an environment of uncertainty. If the wage guarantee binds, the employer can earn rents by providing general training, and the worker can only receive the full return by switching employers. Our results from the British labour market corroborate theirs, particularly since we have more detail on individual training spells, including whether training is accredited, and given the higher frequency of training in Britain (almost three times that of the USA).

Our findings that the returns to training at future employers exceed the returns with the current employer also fit the hypothesis advanced by Hart and Ritchie (2002), in the context of returns to tenure. They suggest that returns to general experience are assessed only at the point of job change, whereas the returns to firm-specific performance occur throughout the lifetime of a job.[32] In the context of our paper, individuals' training experiences might be translated into higher earnings predominantly at periodic points of evaluation (such as at internal or external promotion procedures or job changes).

We also find that accredited employer-financed training has a bigger impact on wages with both the current and future employers than non-accredited training, and that only accredited training is transferable between employers. This result perhaps vindicates the policy initiatives of various governments to encourage accreditation of training where appropriate. The fact that employers pay for highly portable accredited training is again inconsistent with simple human capital theory in the absence of credit constraints.

Finally, there is no indication that self-financed training to develop current skills has any effect on wages, although the cell sizes for this relatively uncommon form of training are quite small. This finding is consistent with the view that firms, and not individuals, are better placed to evaluate the returns to training.

6. CONCLUSION

Our analysis shows that employers do indeed pay for training that is general. We have several pieces of evidence for this. First, from the raw data we know that most work-related training is viewed by its recipients as general and that most is paid for by employers. Second, we find that that employer-financed training increases wages both in the current and future firms, with evidence that the impact in future firms is larger, especially for accredited training.

What are the implications of our results for theory? The fact the employers pay for training that is transferable across employers is inconsistent with orthodox human capital theory with no credit constraints. However, it *is* consistent with the relatively recent training literature based on the assumption of imperfectly competitive labour markets. It is also consistent with the hypothesis that firms offer credit-constrained workers binding training contracts whereby firms pay for general training and workers repay this 'loan' by receiving a post-training wage below their marginal product.

NOTES

1. We use the terms *firm* and *employer* interchangeably and reserve the term *job* for a particular function or set of duties within a firm. The theories being tested concern employers not jobs.

2. See inter alia, Katz & Ziderman, 1990; Stevens, 1994, 1996; Chang & Wang, 1996; Loewenstein & Spletzer, 1998; and Acemoglu & Pischke, 1999b.

3. Respondents were also asked – in a separate body of questions not explicitly linked to training – about any new qualifications they had obtained.

4. We exclude public sector workers from our analysis. In preliminary training and wage equations estimated for public and private sector workers, we rejected the hypothesis that the two sectors could be pooled.

5. As indicated in the questionnaire (Appendix A), details of the most recent of these events were collected first.

6. Note that 'current job' might be interpreted by respondents as being for either the current employer or the current set of duties or 'job' at a single employer. Hence a change of 'job' might be construed by respondents either narrowly as a change of duties at the one employer, or more broadly as a change of employer. A change of employer always implies a change of job, but a change of job does not necessarily imply a change of employer.

7. Median job tenure for induction training events is 6 months. The event counts show that 72% of individuals receiving induction training undergo only one spell of induction, whereas for training as a whole 52% of individuals receive only a single spell.

8. The question on training type was also asked at waves prior to Wave 8 – although in a different part of the questionnaire. For full-time private sector workers, average training incidence for men was 35.0% in waves 1–7 (with a standard error (SE) of 1.0 percentage point) and for women was 35.0% (SE = 1.3), as compared to 30.0% (SE = 1.2) for men and 32.7 (SE = 1.7) for women in waves 8–10. The difference appears statistically significant for men. This reported decrease may be due to a change in the order of the questions: in waves 8–10 the training questions follow those about education much more closely and respondents are specifically asked to exclude previously mentioned full-time educational courses. For these reasons we recommend caution in using the training data to examine questions of human capital formation across Waves 1–10 without taking proper account of this.

9. Table A1 shows that the modal combination is non-induction training for current and future job skills (43% of events). Within induction training, the modal combination of types is training for current and future job skills (71% of induction events). Note that the table omits the general skills category for brevity. The top row represents the 5.7% of events, which were reported as being for general skills only.

10. Barron, Berger and Black (1997) use US data from a matched survey to compare the employer's response about training with the responses of the worker who received the training. They find substantial measurement error in the training variables, and that firms tend to report more training than workers.

11. Only 5% of workplace based events took place in the former rather than the current workplace.

12. In the multivariate equations analysing the association of training characteristics with training type and intensity, as well as in the wage equations, we also tested for differences between the *no fees* and *employer finance* categories by including them as separate regressors. The evidence was mixed: for men (but not women), the *no fees* category was somewhat less strongly associated than was *employer finance* with a training event being for current skills. Both categories had a similar association with training intensity, for both men and women. In the wage equations, only when training was measured by event counts (rather than by incidence or intensity) was there evidence that *employer financed* training had a different – and larger – effect on wages than *no fees* training. On balance the two categories of training – *no fees* and *employer finance* – seemed to have comparable effects. We therefore maintained our combined category in our reported estimates. This also has the advantage of alleviating the problems of small cell size in the wage equations.

13. Since the characteristics are only defined if training actually occurs, the estimates show the associations between them and training intensity *conditional* on training taking place. An OLS rather than a tobit estimator is therefore appropriate.

14. A number of studies apart from our own (Ryan, 1980; Acemoglu & Pischke, 1999a; Leuven & Oosterbeek, 1999) also show that firms incur significant financial costs in providing general training.

15. Exceptions are Lillard and Tan (1992), Arulampalam, Booth and Elias (1997), Blundell, Dearden, and Meghir (1999) Blundell et al. (1999) and Arulampalam and

Booth (2001). While Lillard and Tan (1992) note the importance of multiple training occurrences, they treat these as exogenous when examining the impact of training on economic outcomes. They also note (p. 31) that multiple training occurrences within a period are typically not known from US survey data. The NLS data for young men, for example, contain training information for every survey period, but multiple sources of training are not known within each period; data about sources and types of training are available only for the longest event. Thus Lillard and Tan use as their "events" measure of training the accumulated sum of all training events, where there is only one event measured at each wave. Booth and Bryan (2005) focussed on training incidence and event counts. Unlike the analysis presented here, their more parsimonious set of wage equations did not include training intensity or examine effects due to the timing of training receipt.

16. While FE is consistent when μ_i and the x_{ijt} are correlated, consistency of random effects estimation hinges on orthogonality of μ_i and the x_{ijt} (Wooldridge, 2002: Chapter 10). A Hausman test showed that this was rejected by the data and therefore we report only the fixed effects estimates.

17. Thus training events with previous employers might have an insignificant effect on current wages if (i) the training were received such a long time ago that skills have depreciated due to obsolescence; or (ii) the training was not transferable, or (iii) if there are diminishing returns to the number of courses.

18. See Johnson (1970) and Arulampalam et al. (1997) for analyses of skills obsolescence.

19. Bishop (1997) remarks on the conspicuous absence of evidence that on-the-job wage growth is substantially raised by training or that wages are lower during the training period. He conjectures that there are institutional barriers in the US labour market that prevent firms and workers from sharing the costs.

20. In the British context, pay negotiations typically occur at annual intervals and pay awards are also usually made once a year even for workers not covered by collective bargaining.

21. The coefficient on a dummy variable for the incidence of employer-paid training in the following year was 0.008, with *t*-ratio 0.65.

22. A similar effect, in a model with heterogeneous returns to training, would occur if workers with higher returns tended to change jobs more often. But this also requires some reasoning as to why mobility costs should be lower for workers with higher returns.

23. To our knowledge, little previous work has been done into the functional form of training in wages equations. An exception is Frazis and Loewenstein (2005).

24. For example, only 50 individuals were observed to receive training in wave 9 before changing employers. In addition, the estimates of the 2-year wage growth equation only used observations on the first and last wave (except for the training variables). By contrast, the fixed effects model (1) estimated below to derive the main results used observations from all three waves.

25. The raw data show that training experiences are positively correlated. For example, individuals who received training in one year are 30 percentage points more likely to receive training the next year (conditional probability 0.48) than individuals who had no training in the first year (conditional probability 0.18).

26. The total variation is greater than suggested by the table since some individuals change jobs more than once.
27. We also estimated all our wages models with occupation omitted. The estimated coefficients barely changed; e.g. the largest change is for the coefficient on accredited training counts with the current employer, where the coefficient falls from 0.019 (Column 3, Table 6) to 0.018 with occupation omitted.
28. To allow some flexibility of functional form, we also estimated a specification including both incidence and count variables for all training categories. The estimates suggested that only employer-financed training significantly affects wages, mainly through incidence of training with previous employers, which increases expected current wages by nearly 10%. However, a drawback of this specification is that, by construction, incidence and counts are highly correlated. The correlation coefficient of incidence and counts of employer-paid skills training is 0.78 for training with the current employer and 0.88 for training with the previous employer.
29. As we discussed further below, the fact that the returns to training with future employers exceed the returns with the current employer suggests that a large part of the training is transferable.
30. Unfortunately we cannot determine from our data whether or not wages rise as fast as productivity; but our evidence does not wholly support Bishop (1997), who argues that firms and workers are unable to share the costs and benefits of training.
31. Indeed, since any specific component of training will be lost when a worker moves between employers, our estimates may well be a lower bound on the returns to training across employers.
32. Hart and Ritchie argue that this occurs because it is more efficient to evaluate the returns to general training at the point of job change. The reason is that it (i) simplifies performance-assessment processes (lowering costs), (ii) provides scale economies because filling vacancies or processing promotions offers involves groups of individuals, and (iii) simplifies within-job wage assessment by confining attention only the job-specific elements.

ACKNOWLEDGEMENT

The support of the Leverhulme Trust (under Award F/00213C "Work-related Training and Wages of Union and Non-union Workers in Britain") is gratefully acknowledged. Thanks to Wiji Arulampalam, Bruce Chapman, Bob Hart and seminar participants at the University of Essex and the Australian National University for helpful comments, and to the Editor and an anonymous referee.

REFERENCES

Acemoglu, D., & Pischke, J. S. (1998). Why do firms train? Theory and evidence. *The Quarterly Journal of Economics*, *113*(1), 79–119.

Acemoglu, D., & Pischke, J. S. (1999a). Beyond Becker: Training in Imperfect Labour Markets. *Economic Journal, Features,* F112-142.

Acemoglu, D., & Pischke, J.-S. (1999b). The structure of wages and investment in general training. *Journal of Political Economy*, *107*(3), 539–572.

Arulampalam, W., & Booth, A. L. (2001). Learning and earning: Do multiple training events pay? A decade of evidence from a cohort of young British men. *Economica*, *68*(271), 379–400.

Arulampalam, W., Booth, A. L., & Elias, P. (1997). Work-related training and earnings growth for young men in Britain. *Research in Labor Economics*, *16*, 119–147.

Barron, J. M., Berger, M. C., & Black, D. A. (1997). *On-the-job training*. W.E. Upjohn Institute for Employment Research.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. New York: National Bureau of Economic Research.

Bishop, J. H. (1997). What do we know about employer training. *Research in Labor Economics*, *16*, 19–88.

Blundell, R., Dearden, L., & Meghir C. (1999). *Work-related training and earnings*. Working Paper, Institute of Fiscal Studies.

Booth, A. L., & Bryan, M. L. (2005). Testing some predictions of human capital theory: New training evidence from Britain. *Review of Economics and Statistics*, *LXXXVII*(2), 391–394.

Booth, A. L., & Zoega, G. (2004). Is wage compression a necessary condition for firm-financed general training? *Oxford Economic Papers*, *56*, 88–97.

Campanelli, P., & Joanna, C., with contributions from L. McAulay, A. Renouf and R. Thomas (1994). *Training: An Exploration of the Word and the Concept*. Employment Department Research Series No. 30, July.

Chang, C., & Wang, Y. (1996). Human capital investment under asymmetric information: The Pigovian conjecture revisited. *Journal of Labor Economics*, *14*(July), 505–519.

Frazis, H., & Loewenstein, M. A. (2005). Reexamining the returns to training: Functional form, magnitude, and interpretation. *Journal of Human Resources*, *40*(2), 453–476.

Hart, R. A. and Ritchie, F. (2002). *Tenure-based Wage Setting*, Mimeo, University of Stirling, June.

Hashimoto, M. (1981). Firm-specific human capital as a shared investment. *American Economic Review*, *71*(June), 475–482.

Johnson, T. (1970). Returns from investment in human capital. *American Economic Review*, *60*(4), 546–560.

Katz, E., & Ziderman, A. (1990). Investment in general training: The role of information and labour mobility. *Economic Journal*, *100*, 1147–1158.

Kuratani, M. (1973). *A theory of training, earnings and employment in Japan*. Ph.D. Dissertation, Columbia University.

Leuven, E., & Oosterbeek, H. (1999). Demand and supply of work-related training: Evidence from four countries. *Research in Labor Economics*, *18*, 303–330.

Lillard, L. A., & Tan, H. W. (1992). Private sector training: Who gets it and what are its effects? *Research in Labor Economics*, *13*, 1–62.

Loewenstein, M. A., & Spletzer, J. R. (1998). Dividing the costs and returns to general training. *Journal of Labor Economics*, *16*(1), 142–171.

Ryan, P. (1980). The cost of job training for a transferable skill. *British Journal of Industrial Relations*, *8*, 334–352.

Sicilian, P. (2001). On-the-job training and starting wages. *Journal of Labor Research*, *22*(4), 809–816.

Stevens, M. (1994). A theoretical model of on-the-job training with imperfect competition. *Oxford Economic Papers*, *46*, 537–562.

Stevens, M. (1996). *Transferable training and poaching externalities*. In: A. L. Booth, & D. J. Snower (Eds), *Acquiring skills: Market failures, their symptoms and policy responses*, (Chapter 2). Cambridge: Cambridge University Press.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

APPENDIX A. FORM OF TRAINING QUESTIONS IN THE BHPS, WAVES 8–10

(Apart from the full-time education you have already told me about:) Have you taken part in any other training schemes or courses at all since September 1st last year or completed a course of training which led to a qualification? Please include part-time college or university courses, evening classes, training provided by an employer either on or off the job, government training schemes, Open University courses, correspondence courses and work experience schemes.

EXCLUDE LEISURE COURSES

INCLUDE CONTINUING COURSES STARTED BEFORE SEPTEMBER 1st 1997

D69. How many training schemes or courses have you done since September 1st 1997, including any that are not finished yet?

EXCLUDE FULL-TIME COURSES ALREADY MENTIONED WRITE IN NUMBER

IF MORE THAN THREE TRAINING SCHEMES OR COURSES
PLEASE COMPLETE THE GRID FOR THE THREE LONGEST
FILL IN DETAILS FOR EACH TRAINING SCHEME OR COURSE IN
GRID STARTING WITH THE MOST RECENT
MOST RECENT COURSE/TRAINING = ONE THAT ENDED MOST
RECENTLY OR IS STILL CONTINUING

I would like to ask some details about all of the training schemes or courses you have been on since September 1st last year, (other than those you have already told me about), starting with the most recent course or period of training even if that is not finished yet.

Where was the main place that this course or training took place? (Write in place.)

Was this course or training. . .
To *help* you get started in your current job?.......
To *increase* your skills in your current job for example by learning new technology?..............
To *improve* your skills in your current job?........
To *prepare* you for a job or jobs you might do in the future?.............
To *develop* your skills generally?........................

Since September 1st last year how much time have you spent on this course or training in total?
Hours.....................1
Days.......................2
Weeks.....................3
Months...................4
Other (*SPECIFY*)..5

Which statement or statements on this card describe how any fees were paid, either for the course or for examinations?
No fees.........................01
Self/family....................02
Employer/future emp...03
New Deal scheme........05
Training for work,
Youth/Emp training/
TEC.............................06
Other arrangement *(SPECIFY)*

Was there a course or qualification designed to lead directly to a qualification, part of a qualification, or no qualification at al?

Did you actually get any qualification from this course or training since September 1st last year?
Please look at this card and tell me whether you obtained any of these qualifications from this course or training since September 1st last year. (LIST)

How many subjects did you get?

Table A1. Combinations of Reported Training Categories.

Type of training			Frequency	Proportion	Standard Error
Induction	Current skills	Future skills			
0	0	0	245	0.057	0.004
0	0	1	278	0.064	0.004
0	1	0	1,405	0.325	0.007
0	1	1	1,856	0.430	0.008
1	0	0	33	0.008	0.001
1	0	1	29	0.007	0.001
1	1	0	91	0.021	0.002
1	1	1	380	0.088	0.004
			4,317		

Table A2. Location of Training by Type.

Type	Frequency	Workplace	Employer Training Centre	Private Training Centre	College	Home	Other
Induction	533	0.386	0.191	0.195	0.167	0.019	0.041
Current skills	3,732	0.383	0.181	0.204	0.144	0.025	0.063
Future skills	2,543	0.336	0.166	0.194	0.208	0.036	0.060

Table A3. Financing of Training by Location.

Location	Frequency	No Fees	Employer Paid	Self/Family Paid	Other Payment
Workplace	1,567	0.432	0.555	0.005	0.013
Employer training centre	745	0.348	0.643	0.005	0.005
Private training centre	856	0.091	0.818	0.056	0.042
College	745	0.094	0.486	0.340	0.103
Home	130	0.092	0.531	0.369	0.023
Other place	274	0.212	0.664	0.066	0.073

Table A3 presents evidence that finance category 'no fees' tends to be reported for training locations where it is likely that training expenditure is not visible to employees. Thus, training reported in the first two rows takes place directly within the sphere of the employing organisation (either in the firm or at an employer training centre), while the remainder occurs in external locations. Abstracting from other factors which may be associated with how training at different locations is financed, the drop in the proportion of no fees reports (more than 25 percentage points) between the top two rows and the rest of the table is notable, as is the difference of 8 percentage points between workplace training and that occurring at the employer's training centre. These raw data suggest that for most training events where no fees are reported the employer in fact pays. Also noteworthy in the table is that self financing of college and home-based training is very prevalent, characterising about 35% of events.

Table A4. Financing of Training by Type and Accreditation Status.

Type	Frequency	No Fees	Employer Paid	Self/Family Paid	Other Payment
Induction	533	0.276	0.623	0.058	0.054
Current skills	3,732	0.276	0.647	0.053	0.033
Future skills	2,543	0.247	0.607	0.109	0.047
Accredited	1,811	0.158	0.618	0.171	0.025
Non-accredited	2,506	0.346	0.616	0.028	0.012

Table A5. Combinations of Finance Methods.

Finance method	Frequency	Proportion of Events also Financed by:			
		No Fees	Employer Paid	Self/Family Paid	Other Payment
No fees	1,154	1	0.0095	0	0
Employer paid	2,662	0.0041	1	0.0049	0.0026
Self/family paid	379	0	0.0343	1	0.0211
Other payment	159	0	0.0440	0.0503	1

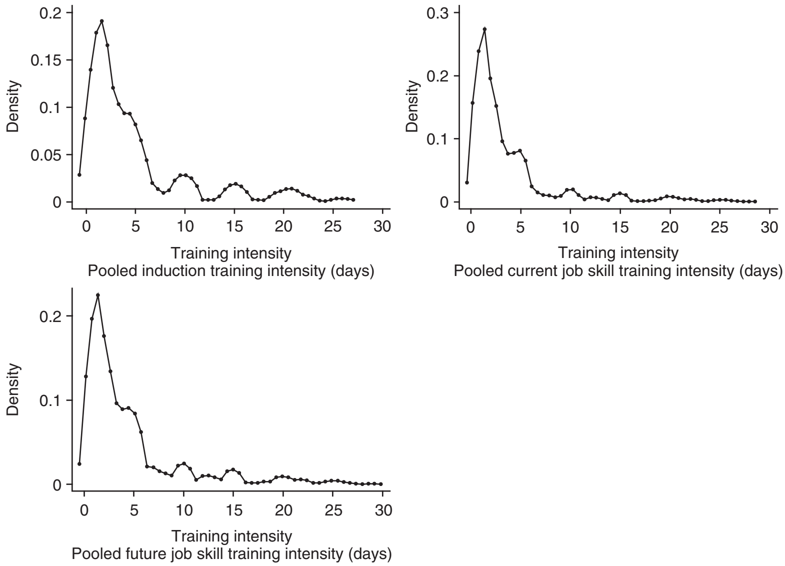


Fig. A1. Distribution of training intensity.

Although the distributions have similar shapes their means differ substantially, in particular the mean intensity is 20.6 days for induction training but only 11.5 days for current job skills training (see Fig. A1) The percentiles reinforce the contrast: the respective medians are 4.0 and 2.0 days and a quarter of all induction events last 17.5 days or longer compared to 5 days or longer for skills training. These results indicate that while induction training is a relatively infrequent event, it often involves a substantial investment on the part of firms and workers.

APPENDIX B. DEFINITION OF VARIABLES

Variable	Definition
Demographics:	
Married	Legally married or living in a couple as partners at interview date
Region of residence	Eleven regional dummy variables: East Midlands (base), Greater London, South East , South West, East Anglia, West Midlands, North West (incl. Manchester), Yorkshire and Humberside, North, Wales and Scotland
Highest Educational Level:	
No qualification (base)	Respondent does not report any academic qualification
O level/GCSE	Highest educational qualification is one or more "Ordinary"-level qualifications (later replaced by General Certificate of Secondary Education), taken at end of compulsory schooling at age 16
A level	Highest educational qualification is one or more "Advanced"-level qualifications, representing university entrance-level qualification, taken typically at age 18
Vocational qualification	Higher vocational qualifications (e.g., HNC, HND, teaching and nursing)
First degree	First (bachelors-level) university degree
Postgraduate degree	Higher university degree
Labour market history:	
Experience	Total experience since labour market entry (years)
Tenure	Time in current job (years)
Workplace size	
Size 1–24 (base)	Firm size: fewer than 25 employees at the establishment (base)
Size 25–49	Firm size: 25–49 employees at the establishment
Size 50–99	Firm size: 50–99 employees at the establishment
Size 100–199	Firm size: 100–199 employees at the establishment
Size 200–499	Firm size: 200–499 employees at the establishment
Size 500–999	Firm size: 500–999 employees at the establishment
Size 1,000 plus	Firm size: 1,000 or more employees at the establishment

APPENDIX B. (*Continued*)

Variable	Definition
Occupation	
Professional	Professional occupation (from the Standard Occupational Classification)
Managerial	Managerial occupation
Non-manual	Associate professional and technical occupations, clerical and sales occupations
Skilled manual	Craft and related, personal and protective service occupations, and plant and machine operatives.
Unskilled (base)	Other semi-skilled and unskilled occupations
Industry	Ten one-digit Standard Industrial Classification dummy variables: agriculture, forestry and fishing , energy, extraction, metal goods, other manufacturing, construction, distribution, hotels and catering, transports, banking and finance, other services. Base is other services.
Charity	Works in a non-profit organisation
Fixed-term contract	Job covered by a fixed-term contract
Temporary contract	Job is seasonal, agency, casual or other non-pemanent job
Trade union covered	Recognised trade union/staff association at workplace covering type of job
log (hourly wage)	= (usual gross pay per month)/[(usual standard weekly hours) + 1.5*(usual paid overtime weekly hours)] *(12/52)
Unemployment rate	Local unemployment rate. The geographic unit is 306 matched job centres and travel-to-work areas (source is National On-line Manpower Information Service)

APPENDIX C

Table C1. Fixed-Effect Wage Equation – Coefficient Estimates not Reported in Table 6, Column (2).

Variable	Mean	Coefficient	*t*-stat	Variable	Mean	Coefficient	*t*-stat
Experience (years)	16.86	0.1008***	3.86	Fixed-term contract	0.02	−0.0508*	1.66
Experience squared		−0.0009***	6.69	Temporary contract	0.01	−0.0698**	2.22
Tenure (years)	4.35	0.0002	0.10	Unemployment rate	0.04	−0.0606	0.18
Tenure squared		0.0001	0.78	Agric, forests, fishing	0.01	−0.1050**	2.13
O-level/GCSE	0.23	0.0718	1.62	Energy and water	0.02	−0.0531	1.30
A-level	0.16	0.0760*	1.73	Extraction, chemicals	0.05	0.0550*	1.93
Vocational qual	0.28	0.0864**	2.21	Metal goods	0.14	0.0051	0.23
First degree	0.11	0.2489***	4.05	Other manufacturing	0.13	−0.0127	0.56
Postgraduate degree	0.03	0.3034**	2.39	Construction	0.04	0.0288	0.95
Married/cohabiting	0.73	0.0021	0.13	Dist, hotels, catering	0.21	−0.0495**	2.55
Manager	0.19	0.1353***	5.50	Transports	0.08	−0.0738***	2.88
Professional	0.07	0.1074***	3.91	Banking & finance	0.20	0.0514***	2.61
Non-manual	0.35	0.1038***	4.47	London	0.09	0.0750	0.61

Skilled manual	0.34	0.0548**	2.52	South-East	0.21	−0.0051	0.05
Charity sector	0.04	−0.0133	0.43	South-West	0.09	−0.0734	0.65
Wave 9 dummy	0.33	−0.0223	0.82	East Anglia	0.05	−0.1332	0.81
Wave 10 dummy	0.34	−0.0625	1.21	West Midlands	0.09	−0.0914	0.73
Estab size 25–49	0.13	−0.0439***	3.55	North-West	0.11	−0.1750	1.41
Estab size 50–99	0.12	0.0134	0.93	Yorkshire	0.09	−0.1151	1.01
Estab size 100–199	0.12	0.0470***	3.23	North	0.06	−0.3682*	1.90
Estab size 200–499	0.15	0.0197	1.36	Wales	0.05	0.1782	1.21
Estab size 500–999	0.08	0.0335*	1.94	Scotland	0.07	0.1176	0.62
Estab size $\geqslant$1000	0.08	0.0649***	3.34	Constant		0.2974	0.71

Notes: * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

WAGE ARREARS AND INEQUALITY IN THE DISTRIBUTION OF PAY: LESSONS FROM RUSSIA

Hartmut Lehmann and Jonathan Wadsworth

ABSTRACT

Many developing and transition countries, and even some in the industrialized West, experience periods in which a substantial proportion of the workforce suffer wage arrears. We examine the implications for estimates of wage gaps and inequality using the Russian labor market as a test case. Wage inequality grew rapidly as did the incidence of wage arrears in Russia in the 1990s. Given data on wages and the incidence of wage arrears we construct counterfactual wage distributions, which give the distribution of pay were arrears not present. The results suggest that wage inequality could be some 30 percent lower in the absence of arrears. If individuals in arrears are distributed across the underlying wage distribution, as appears to be the case in Russia, we show that it may be feasible to use the wage distribution for the subset of those not in arrears to estimate the underlying population wage distribution parameters.

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 125–155

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26004-7

1. INTRODUCTION

Many countries in the developing world, those undergoing the transition from planned to market economic systems and even those in the industrialized West, experience periods in which a substantial proportion of the workforce suffer wage arrears.[1] For any research based on wage distributions, such as estimation of wage inequality, gender pay gaps or the returns to education, failure to account for wage arrears can have important implications, as we show below. Russia is particularly interesting in this regard, since it experienced well-documented increases in both the incidence of wage arrears and wage inequality over the first decade of the transition period. Moreover, the availability of data on both these issues facilitates exploration of the linkages between the two that is not always possible in other countries.

One contributory factor toward inequality in the wage distribution, in any country, could be the presence of wage arrears. If in any given month some workers receive only part of their normal wage, or no wage at all, then wage inequality will be higher, or in exceptional cases lower, than otherwise. Wage inequality in Russia following the end of central planning rose much more than in Central and Eastern European (CEE) countries undergoing transition. The Gini coefficient for wages in Russia rose from 0.22 before transition to around 0.5 in 1996 (Flemming & Micklewright, 1997) and has remained around this level ever since. Wage inequality in Russia is also very high by international standards.[2] Wage arrears were also a pervasive feature of Russian economic life over the 1990s. Lehmann, Wadsworth, and Acquisti (1999), show that around 65% of the workforce was owed money at the height of the problem in 1998. Moreover, the withholding of wage payments was systematic and concentrated heavily on sub-sections of the workforce in certain regions and industries (see e.g. Earle & Sabirianova, 2002; Lehmann et al., 1999; Desai & Idson, 2000). Despite their prevalence, most studies of wages, however, tend to ignore the effect of wage arrears on the earnings distribution.[3]

In what follows, we use data for Russia, to try to estimate what the wage distribution would have looked like if all workers had been paid the full contractual wage on time. By establishing the parameters of the underlying distribution it is then possible to adjust estimates of any between-group differences based on the observed wage distribution. Using Russian Longitudinal Monitoring Survey (RLMS) data, we apply several different imputation methods to generate predicted wages for those in arrears and construct counterfactual estimates of the underlying wage distributions.

We find similar results across the various imputation methods. Earnings dispersion may have been some 30% lower if workers had been paid in full. Since, on average, women seem to be less affected by wage arrears (Lehmann et al., 1999), the mean gender gap is larger in the counterfactual distributions compared with the observed gender pay gap. We also look at pay gaps across other quantiles of the earnings distribution, which cannot be done in the presence of large-scale wage arrears. We then look at how wage arrears affect estimates of the returns to education and relative wage distributions by region and industry.

In the next section we discuss the rationale for constructing counterfactual wage distributions. The subsequent section outlines the various methods used to construct counterfactual wage distributions, while Section 4 discusses data issues. Section 5 analyses earnings inequality in Russia and the decomposition of its change over time, followed by the counterfactual results. Section 7 then concludes.

2. ECONOMIC REALITY IN RUSSIA AND THE CONSTRUCTION OF COUNTERFACTUAL WAGE DISTRIBUTIONS

Is a wage distribution that assumes payment of wages in full and on time for all employees a realistic counterfactual to pursue? Of course, economic welfare depends on the actual distribution of earnings, but a comparison of the actual and counterfactual will provide a means of estimating the cost of arrears. Also, if wage arrears are a problem of irregular pay and not of permanently withheld wages, then we have a strong rationale for constructing counterfactual wage distributions, since there is less concern over possible general equilibrium effects concerning any trade-off between the elimination of wage arrears and employment.[4] We believe that evidence garnered from various sources on the dynamic nature of the arrears process provides this rationale. Aggregate data from Russian Statistical Office (Goskomstat) indicate that since 1996 the stock of wage arrears has been approximately stable, equivalent to the aggregate wage bill for two months. At the same time, there is strong evidence in the RLMS data, the principal source of data in this paper, which supports the hypothesis of wage arrears as a problem of irregular pay rather than that of permanently withheld wages. Lehmann et al. (1999), use the RLMS to document the existence of simultaneous inflows into and outflows from wage arrears. In the data we analyze below, 10% of workers are in arrears at all four interview points and 20% never experience wage arrears.

These flow patterns allied to the fact that the stock of wage arrears is, approximately, in a steady state at the end of our sample period, suggest that the amount of contractual wages not paid to (some) workers is close to the amount of wage debts paid back to (some) workers in any month. Payroll data from a sample of 19 firms in a central Russian industrial city also seem to confirm this pattern in Fig. A1. At times, the stock of arrears in some firms rises, while falling in others. Moreover, the Figure indicates that wage arrears are eventually paid off, and at different rates across firms. It seems that most workers do get paid the wages owed to them eventually.

Given this, because the RLMS elicits information on wages received in the month of the survey, this window might be too narrow to obtain an estimate of the contracted earnings of workers affected by pay irregularities. For example, in an economy where all workers get paid monthly but the data window on earnings is the third week of the month, if we ask: "How much did you get paid in the third week of this month?" some workers will have been paid their monthly salary in this week, but many will have been paid in another week of the month. Estimation of monthly earnings on this weekly window will be certainly inefficient, or even misleading. If, in the Russian case, we had a window of, say, six months, we could obtain better estimates of the contracted monthly earnings of Russian workers. Since we do not have such a wide sample window in which to observe everyone paid in full at least once, then counterfactual distributions provide one way of estimating contracted monthly earnings.[5]

3. BUILDING COUNTERFACTUAL ESTIMATES OF THE EFFECTS OF WAGE ARREARS

The literature suggests several methods of building counterfactual mean estimates, Y^0, given membership of a treatment group, $T_i \varepsilon \{0,1\}$ essentially built on the conditional independence assumption (CIA) whereby assignment to the treatment group is ignorable conditional on a set of exogenous control variables, X, that are unaffected by the treatment and that $Y^0 \perp T/X$. Given the CIA and assuming there is overlap, or common support, in the X distributions of those in arrears and those not, if we take experience of wage arrears as the treatment and let the X variables influence the likelihood of being observed in arrears, then the counterfactual mean of the wage distribution for this treatment group equals the mean of the wage distribution for the no arrears control group, adjusted for differences in observable characteristics across the two groups.

In our case we are interested in not just the counterfactual mean but also the counterfactual distribution of wages, netting out the effect of arrears.

Counterfactual wage distributions have been applied to a variety of economic and statistical issues, e.g. minimum wages (DiNardo, Fortin, & Lemieux, 1996), item non-response (Biewen, 2001) and international differences in wage inequality (Blau & Kahn, 1996). Given the CIA assumption Imbens (2004) shows that it is possible to identify different quantiles of a counterfactual distribution. Fröhlich (2003) shows that either matching on observables or propensity score estimation can be used to estimate counterfactual density functions consistently in addition to counterfactual means, since $E[\xi_{y/x,t}(X,t)/X = r, \quad T = t] = \xi_{y/p,t}(r,t)$ is satisfied both for $\xi_{y/x,t}(X,t) = E[Y/X = x, \quad T = t]$ and the conditional density function $\xi_{y/x,t}(X,t) = f_{y/X,\ T=t}$. Once the counterfactual density is estimated the counterfactual quantiles can be recovered.[6]

If selection into the treatment group also depends on unobservables, then identification of the counterfactual densities, as with counterfactual means, requires data from before the treatment began in order to differentiate or net out any bias caused by unobservables. This generally requires the assumption that the bias caused by unobservables is constant over time. Whether researchers can ever be truly confident that treatment selection is observable, or that any bias from unobservables is constant, are moot points. We therefore produce a series of estimates that rely on the CIA, but which involve different sets of assumptions and look to compare the estimates based on the different methods.

We begin with a simple least-squares prediction and then use least squares with the addition of a random residual, both of which use parameters from a wage equation estimated on the sample without wage arrears to predict wages for those in arrears.[7] We then apply a different residual according to the method proposed by Juhn, Murphy, and Pierce (1993). We next provide counterfactual estimates of the wage distribution following the Kernel density approach pioneered by DiNardo et al. (1996). We then employ a variation of the exact matching techniques used by, among others, Heckman, Ichimura, and Todd (1997), and Kluve, Lehmann, and Schmidt (1999), to assign wages to those in arrears by matching their characteristics to the subsample of those who continue to be paid in full but who had a similar labor market pre-treatment history. The last method used matches on the propensity score rather than a vector of characteristics (for example, Lechner, 2002).

3.1. OLS Methods

Following Oaxaca (1973) we can estimate a wage equation using the sample of those without wage arrears. Using the vector of (consistently) estimated parameters from this equation and the observed characteristics of those in arrears we then predict wages, which those in arrears would receive if they had been paid in full. More formally, let B_{NW} be the vector of parameter estimates from the wage equation of the sample without wage arrears and let $X_{\mathrm{i,WA}}$ be a vector of individual and job-related characteristics that determine whether the *i*-th person experiences arrears. The set of covariates is based on those used by Lehmann et al. (1999) who used the same data set to examine the incidence of wage arrears.[8] The predicted wage of this individual, $Y_{\mathrm{i,WA}}$, will be

$$Y_{\mathrm{i,WA}} = B'_{\mathrm{NA}} X_{\mathrm{i,WA}} \tag{1}$$

Since this method gives only a mean prediction and the actual wage equals the sum of the predicted wage and a residual, $y = \hat{Y} + \hat{u}$, we can add a residual so as to proxy wage dispersion in full. We do this by first taking the standard error of the regression from the no arrears equation, σ_{NW}, and multiplying each individual observation by a, randomly assigned, standard normal random variable z_{i}. This random residual is then added to the predicted wage for the arrears sub-group and is given by

$$\varepsilon_{\mathrm{iWA}} = z_{\mathrm{i}} * \sigma_{\mathrm{NW}} \tag{2}$$

Table A1 gives the estimates from the OLS real wage equations for the no arrears group used to generate these estimates.

3.2. Juhn, Murphy and Pierce

Juhn et al. (1993) and Blau and Kahn (1996) have suggested that it may be worthwhile trying to take into account unobserved heterogeneity as measured by the percentile ranking of each individual in the residual wage distribution. With a simple transformation of the residual into the product of a standard normal residual, θ, and the residual standard deviation from the wage equation, σ, the predicted wage can be written as

$$Y_{\mathrm{i,WA}} = B'_{\mathrm{NW}} X_{\mathrm{i,WA}} + \sigma_{\mathrm{NW}} \theta_{\mathrm{WA}} \tag{3}$$

Applying this method in the context of wage arrears, the counterfactual is then the set of wages that would result if the no arrears wage coefficients and residual standard deviation were given to those currently in arrears. Since many of the observations on the dependent variable in the arrears sample are zero, this technique relies on the assumption of normality in the residuals estimated from this subset.[9] The method uses the standard residuals from the arrears regression to calculate counterfactuals. This standardized residual is usually interpreted as an individual's ranking in the residual wage distribution and as such a measure of unobserved relative skill. However, the outcome we analyze in Eq. (3) gives an individual's relative ranking in the residual arrears wage distribution, which is hard to interpret as a measure of unobserved skill, unless one is prepared to make the unlikely assumption that the size of non-payment reflects unobserved skill. The estimates from the equations for those not in arrears used to construct the counterfactuals are given in Table A1.

3.3. Kernel Density Counterfactuals

DiNardo et al. (1996), (hereafter DFL), have suggested that a broader insight may be obtained by taking into account the entire wage structure, allowing the returns to observables and unobservables to vary across the distribution of wages. The principle remains the same, to estimate the wages that those in arrears would receive had they been paid as those paid in full. Given the joint distribution of wages, w, and characteristics, x, the marginal distribution of wages conditional on x can be written $g(w) = \int f(w/x)h(x)\ \mathrm{d}x$. Following DFL, using Bayes' law, the counterfactual wage distribution if everyone were paid in full can be obtained by taking the observed wage distribution of the subset of those paid in full and reweighting by a parameter $\Phi(x)$, where $\Phi(x)$ reflects the relative incidence of arrears conditional on characteristics x, $\Phi(x) = \mathrm{Pr(No\ Arrears)}/\mathrm{Pr(No\ Arrears}/x)$. The weights are normalized to sum to one. So,

$$g(w) = \int \Phi(x) f^{\text{No Arrears}}(w/x) h(x/i = \text{No Arrears})\ \mathrm{d}x$$

The integral is approximated using Kernel density estimation, producing no predictions of individual wages, only the quantiles of the distribution. The numerator in $\Phi(x)$ is the sample proportion of those not in arrears in any year and the denominator is estimated by a logit regression conditional on a set of characteristics determining the incidence of arrears. The estimates from the logit equations used to construct these estimates (Table A2) confirm the dominance of location and firm characteristics in explaining arrears, as found in Lehmann et al. (1999).

3.4. Matching Estimators

If there were unobserved heterogeneity among those in arrears, then the preceding techniques would fail to account for this. The JMP approach and the DFL density approach perhaps come closest, however they implicitly assume that heterogeneity among those not in arrears is duplicated among those in arrears. If those not in arrears are different from those in arrears, the counterfactual estimates could be biased.

We therefore experiment with alternative approaches based on the matching estimator literature. The first technique follows Heckman et al. (1997) in that we also condition, non-parametrically, on "pre-treatment history" in order to minimize any biases arising from unobserved heterogeneity. This means conditioning on events before wage arrears began, together with a set of current observable, exogenous characteristics, in order to try and capture heterogeneity in the arrears population. Conditioning on a set of pre-treatment covariates is assumed to be sufficient to allow the assumption of assignment to the treatment group as random, such that unobservables may be ignored. Heckman et al. (1997) find that for this type of matching estimator to work well the same data set should be used for the control and treatment group, the groups should be in the same local labor markets and the data set should contain a rich set of variables relevant to the treatment decision.

Using the panel element of the RLMS we condition on labor market status one year earlier and if employed, the ranking in the wage distribution of those paid in full. If the individual was out of work one year earlier we create unemployed and inactive categories. If the individual was in arrears one year earlier we create a separate sub-category. We divide last year's wage distribution, excluding arrears, into deciles. We assign the wages of those currently paid in full to those in the treatment group, who were placed in the same decile a year ago when both treatment and control groups were paid in full. Those in arrears in both years are given the current wages of those not in arrears now that were in arrears one year earlier. Those in arrears now but non-employed a year ago are given the current wages of those non-employed a year ago but paid in full now. In each case, if more than one person can be matched with the individual we assign the average wage of the matched controls. In addition we match according to age (with a maximum allowed difference of ten years), gender, region (3 groups, Metropolitan Moscow and St. Petersburg, East and West) and qualifications (3 groups) in the current year. This strategy conforms broadly to the criteria set out by Heckman et al. (1997) required for a good performance of a matching estimator.[10]

The matching algorithm is shown in Box A1. Since this approach can only be used when there are at least two consecutive years of longitudinal data, we confine our estimates using this approach to 1996 and provide comparisons using the other counterfactual techniques estimated over the same sample. The approach assumes that individuals do not move much across the earnings distribution.[11] Fig. 2, which for those currently in arrears, plots the share coming from each wage decile in the previous year, also suggests that those in arrears are drawn from across the entire wage distribution.

3.5. Propensity Scores

The non-parametric matching approach omits around 10 percent of potential matches for whom a donor from the control group cannot be found. To avoid this lack of common support, we also employ propensity score matching, where all individuals are matched according to the closeness in the estimated probability of experiencing wage arrears. We use the matching algorithm suggested by Dehejia and Wahba (2002).[12] We estimate probit regressions, conditional on the same co-variates as used in the matching approach, take the predicted probability – the propensity score – and match, with replacement, those in arrears to those not with the nearest propensity score. We estimate two variants of the propensity score, one with pre-treatment variables included in the set of co-variates and one without.

4. DATA

Our main data source is the second phase of the Russian Longitudinal Monitor Survey (RLMS) a longitudinal panel of around 4,000 households across the Russian Federation. We use the surveys conducted in the autumn of 1994, 1995, 1996 and 1998, the period in which wage arrears first emerged and subsequently affected two-thirds of the workforce at the height of the problem in 1998. The data contains a set of demographic and establishment characteristics, together with information on the labor market activities of its sample. Despite its relatively small size, the advantage of this source is that we can track individual wages and the incidence of wage arrears over time. We restrict our sample to employees of working age and exclude the military.[13] The survey design does not follow individuals if they move, but does sample new occupants of the same address. There are around 10,000 individual observations in each wave, of which around 4,000 are in work and around 3,500 give wage-related information.

The survey questions dealing with wage arrears ask whether, conditional on being in work, an individual was owed money by the firm in the past month or was paid "in kind" with goods produced by the firm. This constitutes our sample of those in arrears in any wave. Some of those in arrears are paid some money, while others, around one half of those in arrears, receive nothing. The RLMS also asks for the total amount owed, together with the number of months since the worker was paid last, but does not give the dates of when arrears occurred so it is not possible to ascertain the dynamic history of the wage arrears process. It may be that some of those not in arrears are paid more than their monthly wage if arrears are paid back. There may also be some in arrears who were paid in full in the current month. However there is no way of ascertaining these issues from the data. Respondents, both those paid in full and those in arrears, are also asked to state the amount of *money* received from their employers after tax in the past month. These are total wage receipts and not contractual wages, on which there is no reliable information.[14] There is no distinction made between basic wages and bonus. This constitutes the "true" wage for those paid on time.

These wage responses are then deflated by a national price deflator indexed to 100 at January 1998. We remove outliers from that data, namely those earning in excess of 4,000 rubles a month, or less than 50 rubles if the respondents are not in arrears.[15] Since we are interested in the impact of arrears on the aggregate distribution, we do not construct gender-specific counterfactual wage distributions.[16] Standard errors around the quantiles of the observed and counterfactual distributions are generated using the bootstrap method.[17] We also use a smaller, Russian household survey data set, VTsIOM,[18] undertaken in 1993, in order to provide summary comparative evidence on pay from an earlier period when wage arrears were less prevalent, together with labor force survey data from Poland and Britain as benchmark comparisons. The former is a transition economy without wage arrears or a dominant oligarchy that followed a different restructuring process where more attention was given to sharing the costs of reform equally (Hellman, 1998). The latter is a Western economy where wage inequality had risen sharply just prior to the sample period.[19]

5. EARNINGS DISTRIBUTIONS AND INEQUALITY IN RUSSIA

The timing of the dramatic rise in inequality during the first years of transition, documented in Brainerd (1998), indicates that most of the rise in inequality occurred before the problem of wage arrears really began, though hyperinflation at the onset of reforms was probably not the sole contributing factor to the initial rise in inequality. However, as inflation subsided aggregate inequality remained high. The RLMS data indicate that inequality fell in regions with a low incidence of wage arrears, and rose most in regions with the largest increase in wage arrears. The Gini coefficient in the metropolitan areas, where arrears are lowest, fell from 0.39 to 0.35 between 1994 and 1998, but rose from 0.43 to 0.49 in the Far East, where arrears are highest. It seems important, therefore, to try to analyze to what extent wage arrears have affected the earnings distribution since payment problems began.

In order to demonstrate the effects of wage arrears on the wage distribution, Table 1 gives summary measures of the changes in real monthly wage distribution across our sample period. The VTsIOM data show that wage inequality was already higher in Russia than in Poland before wage

Table 1. Real Monthly Wage Distributions in Russia.

	1993 VTsIOM	1994 RLMS	1996 RLMS	1998 RLMS	1996 Poland	1996 Britain
Total						
Mean	916 (1014)	609 (656)	501 (659)	371 (494)		
90th	1,724 (23)	1,500 (19)	1,376 (14)	907 (11)		
50th	690 (28)	422 (13)	287 (34)	217 (11)		
10th	276 (20)	0	0	0		
90/10	6.25	n/a	n/a	n/a	2.70	8.55
90/50	2.5	3.55	4.79	4.18	1.83	2.20
50/10	2.5	n/a	n/a	n/a	1.48	3.89
Coefficient variable	1.11	1.11	1.32	1.33	0.62	0.80
Gini	0.407 (.009)	0.547 (.005)	0.637 (.006)	0.619 (.006)	0.239	0.387
% arrears	10 (0.6)	44.4 (0.8)	64.9 (0.9)	67.6 (0.8)	0	0
% no pay	0	19.3 (0.6)	34.6 (0.9)	28.1 (0.8)	0	0
No arrears						
Mean	944 (1,030)	808 (625)	896 (727)	629 (550)		
90th	1,724 (24)	1,718 (27)	1,802 (37)	1,273 (25)		
50th	690 (30)	625 (18)	677 (30)	484 (22)		
10th	276 (21)	188 (16)	229 (26)	187 (18)		
90/10	6.25	9.14	7.87	6.81		
90/50	2.5	2.75	2.66	2.63		
50/10	2.5	3.32	2.96	2.59		
Coefficient variable	1.12	0.77	0.81	0.87		
Gini	0.407 (.011)	0.420 (.005)	0.415 (.008)	0.428 (.009)		

Note: Wage data indexed to December 1997 prices. Wage observations for population of employees aged 18–69. Standard errors in brackets, based on bootstrapping over 100 replications. Inequality measures use delta method approximation using standard normal distribution. Standard errors of proportions are used in percentage rows.

arrears took off, indicative of the different restructuring paths pursued by the two transition countries. By 1996, the Gini coefficient on Russian wages was more than twice that observed in Poland and 60% higher than in Britain. The earnings distribution also widens over the first half of the sample period, while the evidence for the second half of the sample period is mixed. The coefficient of variation continues to increase, albeit more slowly, but the Gini coefficient and the ratio of the 90th to 50th wage quantiles falls back. The Table also shows that real average earnings fell markedly over the sample period, as a series of national economic crises left inflation soaring and nominal wages failing to keep pace. By 1998, around two thirds of employees were not receiving a wage complete or on time, and around 40% of these received nothing in the preceding month. The large number of zero wage observations means that any conventional measures of inequality based around logarithmic transformations will be of little use.

The inequality estimates are influenced strongly by wage arrears. Fig. 1 tracks the increasing skewness of the real monthly wage distribution as the incidence of arrears builds up. The bottom panel of Table 1 confirms that inequality is lower and rises by much less among those paid in full during the sample period. The Gini coefficient, for example, is around one third for the subset of those without wage arrears, in any period. Many individuals appear in low deciles solely because they are not paid at all or paid only part of their wages (Fig. 2).

5.1. Counterfactual Estimates

We now present our counterfactual estimates of the underlying wage distribution for the years 1994, 1996 and 1998. Table 2 summarizes details of the estimated distributions for the different methods used.[20] Fig. 3 graphs the counterfactual Kernel densities, the sum of the actual wage of those paid in full and the predicted wage of those in arrears. Table 2 confirms that the mean and various quantiles of the distributions are all higher using any of the counterfactual estimates. The bootstrapped standard errors indicate that all the imputed distributions lie within 2 standard errors of each other, with the exception of the OLSI estimates – though these do not contain a random residual and so would be expected to differ. The magnitudes of the estimated standard errors are also similar. In general then, the counterfactuals indicate that mean wages would have been around 30% higher in 1994 and around 60% higher in 1998 in the absence of wage arrears. Similarly, the estimated overall dispersion, as measured by the coefficient of variation, would be

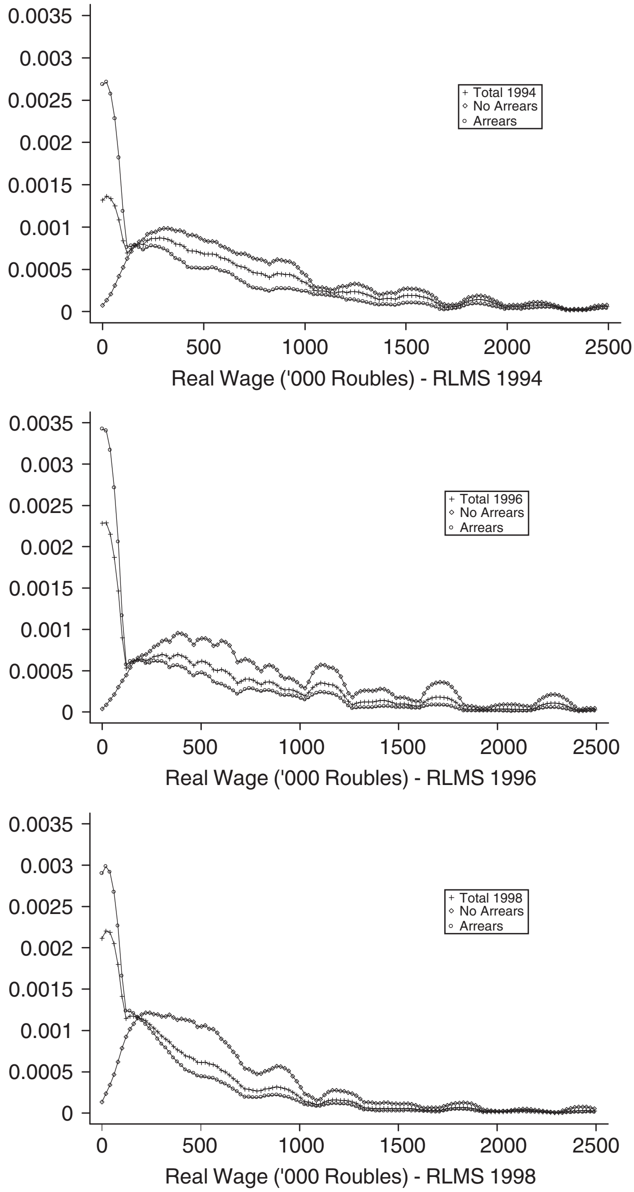


Fig. 1. Distribution of Real Wages in Russia.

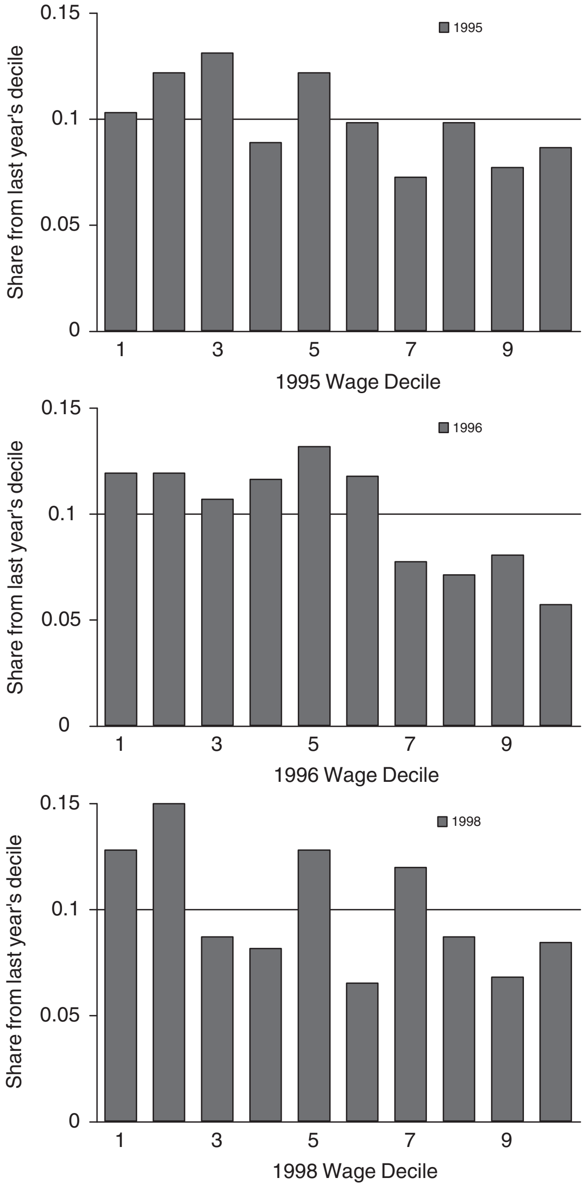


Fig. 2. Previous Wage Decile of Those in Arrears.

Table 2. Counterfactual Real Wage Distributions.

	Mean	90th P'centile	Median	10th P'centile	90/10	90/50	50/10	Coefficient Variable	Gini
(*1994*)									
Actual	629	1,538	451	0	N/a	3.4	N/a	1.04	0.532
OLS I	743 (12)	1,406 (32)	607 (14)	250 (8)	5.6	2.3	2.4	0.73 (.01)	0.365 (.006)
OLS II	816 (17)	1,672 (60)	613 (15)	190 (8)	8.8	2.7	3.2	0.88 (.03)	0.429 (.006)
JMP	815 (15)	1,688 (57)	625 (13)	203 (11)	8.3	2.7	3.1	0.81 (.02)	0.411 (.007)
DFL	805 (16)	1,719 (80)	625 (15)	188 (5)	9.1	2.8	3.3	0.82 (.01)	0.417 (.005)
PS I	832 (17)	1,818 (74)	625 (10)	188 (7)	9.7	2.9	3.3	0.81 (.02)	0.420 (.006)
(*1998*)									
Actual	384	907	242	0	N/a	3.7	N/a	1.30	0.605
OLS I	517 (14)	907 (23)	422 (11)	212 (10)	4.3	2.1	2.0	0.73 (.02)	0.337 (.008)
OLS II	594 (18)	1210 (40)	425 (12)	146 (7)	8.3	2.8	2.9	0.98 (.05)	0.443 (.009)
JMP	607 (18)	1211 (42)	451 (18)	145 (17)	8.4	2.7	3.1	0.90 (.03)	0.430 (.014)
DFL	588 (16)	1210 (38)	423 (13)	121 (11)	10.0	2.9	3.5	0.91 (.03)	0.433 (.009)
PS I	609 (22)	1247 (89)	434 (23)	127 (12)	9.8	2.9	3.4	0.91 (.03)	0.449 (.011)

Note: OLS I is OLS estimate without residuals, OLS II includes residuals, JMP is the Juhn–Murphy–Pierce decomposition, DFL is the DiNardo, Fortin, Lemieux technique, PS I is the estimate based on propensity score without conditioning on pre-treatment history. Actual values may vary from Table 1 due to missing observations on covariates used to construct counterfactuals. Bootstrapped standard errors in brackets are based on 300 replications. Sample sizes: 3,962 in 1994 and 3,336 in 1998.
Source: RLMS authors' calculations.

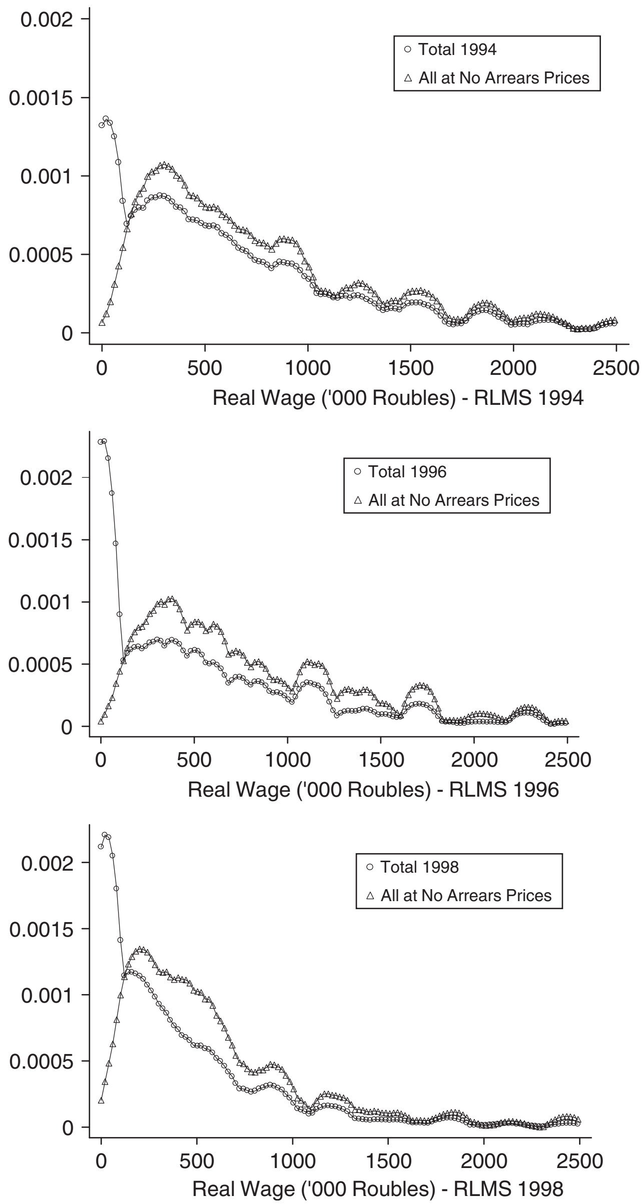


Fig. 3. Counterfactual Estimates of Wage Distribution in Absence of Arrears.

around 20% lower in 1994 and some 40% lower in 1998 in the absence of arrears. The counterfactual Gini coefficients are now similar to that observed in Britain around the same time but much higher than for Poland. Interestingly, the counterfactual Gini coefficients are also similar to those of the no arrears sub-group in Table 1.

Table 3 uses the panel element of the data in order to add estimates based on exact matching and a second propensity score estimator based on ‘‘pre-treatment history’’ included as additional regressors in the propensity score logit. We compare the results with those using the other methods for the year 1996, based on the sub-sample with valid pre-treatment histories. We also show the distribution of those in the sample who get paid in full and on time (column 2). The pattern of results follows that of Table 2. Mean wages would be around 60% higher and the wage distribution narrower by around 40% in the absence of wage arrears. Apart from the estimates based on simple OLS prediction (OLS I) all other counterfactual distributions have a similar spread as can be seen from the coefficients of variation and Gini coefficients.[21] Conditioning on pre-treatment history for the propensity score estimates (PSII), results in estimates within two standard errors of the propensity score estimates without pre-conditioning. This suggests that unobserved heterogeneity as captured by this method is not important for this sample. Note that the quantiles of the no arrears distribution again appear insignificantly different from the counterfactuals, a point to which we return later.

These counterfactual techniques can also be applied to wages observed over any combination of years to give estimates of the average wage distribution over a given interval. One advantage of pooling data across years is that we can net out the influence of unobservables in the prediction equations through random effects estimation of the wage of treatment equations. Rather than reveal the counterfactual distribution at a single year, it inevitably reveals a medium- run average wage distribution, which is based on the predictions of a much smaller proportion of the population who are never in arrears over successive years. Table A2, showing the results for pooling the years 1994 to 1996, suggests little difference between the pooled and the random effects estimates of the counterfactual distributions.

5.2. Gender, Region and Education Pay Gaps Revisited

We now examine the implications of these counterfactual estimates for pay gaps between various sub-groups of the workforce. If the incidence of wage

Table 3. Counterfactual Real Wage Distributions (1996).

	Actual	No Arrears	OLS I	OLS II	JMP	DFL	Match	PS I	PS II
Mean	512 (14)	897 (26)	762 (21)	858 (30)	860 (26)	845 (32)	889 (30)	887 (32)	861 (36)
90th	1,261 (66)	1,835 (125)	1,351 (65)	1,750 (79)	1,776 (64)	1,720 (71)	1,720 (114)	1,720 (130)	1,802 (129)
50th	339 (18)	688 (14)	635 (21)	630 (25)	674 (19)	630 (37)	688 (37)	688 (23)	631 (31)
10th	0	229 (10)	304 (20)	227 (10)	194 (24)	221 (22)	229 (9)	229 (12)	225 (18)
90/10	n/a	8.0	4.4	7.7	9.2	7.8	7.5	7.5	8.0
90/50	3.7	2.7	2.1	2.1	2.6	2.7	2.5	2.5	2.9
50/10	N/a	3.0	2.1	2.8	3.4	2.9	3.0	3.0	2.8
Coefficient variable	1.26 (.03)	0.79 (.02)	0.68 (.02)	0.88 (.05)	0.83 (.03)	0.83 (.03)	0.77 (.03)	0.81 (.03)	0.84 (.03)
Gini	0.617 (.008)	0.405 (.009)	0.332 (.011)	0.423 (.011)	0.411 (.012)	0.411 (.012)	0.392 (.012)	0.409 (.012)	0.423 (.014)

Note: See Table 3. PS II is an estimate based on propensity score conditioning on pre-treatment history. Sample size = 2,538, of which 1,351 are in arrears and 1,187 are paid in full and on time. Bootstrapped standard errors in brackets.
Source: RLMS authors' calculations.

arrears is concentrated on sub-groups of the population, then pay gaps estimated on the observed distribution may be misleading.[22] In Table 4 we compare gender pay ratios using the actual distribution, the no arrears distribution and the counterfactual distributions for the year 1996. The imputation methods are broadly in agreement with the exception of the propensity score based estimates, which show a narrowing of the gender pay gap rather than the expected widening when the incidence of arrears across gender is taken into account.[23] The observed distribution suggests a mean gender pay gap of around 20% (column 1). Since women are less likely to be observed with wage arrears, the counterfactual estimates, other than PSI and PSII, suggest that if everyone were paid in full there would be more dispersion in pay between men and women and the gender wage gap would be closer to 30%.

Table 5 gives mean and median wages of three educational categories (graduate, intermediate and primary) and median pay ratios of the first two groups relative to the primary educational category using the actual, the no arrears and all counterfactual distributions. This time all the imputation methods are in broad agreement. Since graduates are under-represented amongst the arrears group, the observed distribution suggests a higher

Table 4. Counterfactual Gender Wage Ratio (1996).

	Actual	No Arrears	OLS I	OLS II	JMP	DFL	Match	PS I	PS II
Men									
Mean	578	1,113	922	1,076	1,043	1,009	1,078	929	910
Median	344	917	803	803	839	803	917	688	688
90th	1,577	2,294	1,615	2,231	2,079	1,950	2,293	1,835	1,720
10th	0	344	351	279	322	252	321	229	203
Women									
Mean	459	752	633	723	714	704	687	847	821
Median	310	573	533	550	560	550	573	656	619
90^{th}	1,126	1,605	1,080	1,425	1,498	1,456	1,261	1,720	1,720
10^{th}	0	221	262	201	145	184	216	229	203
Gender ratio									
Mean	0.79	0.68	0.69	0.67	0.68	0.70	0.64	0.91	0.90
50th	0.90	0.62	0.66	0.68	0.67	0.68	0.62	0.95	0.90
90th	0.71	0.70	0.67	0.64	0.72	0.75	0.55	0.94	1.00
10th	n/a	0.64	0.75	0.72	0.45	0.73	0.67	1.00	1.00

Sample size = 2,193, of which 976 are male and 1,217 female.
Source: RLMS.

Table 5. Actual and Counterfactual Education Wage Ratios (1996).

	Actual	No Arrears	OLS I	OLS II	JMP	DFL	Match	PS I	PS II
Graduate									
Mean	594	944	823	923	917	907	902	904	852
Median	394	732	688	688	692	688	722	688	676
Intermed									
Mean	437	831	702	800	772	771	815	865	867
Median	248	631	573	573	573	563	642	653	630
Primary									
Mean	448	874	721	804	880	835	824	871	868
Median	229	581	585	569	688	607	574	676	631
Ratio: wrt primary									
Graduate	1.59	1.26	1.18	1.21	1.01	1.13	1.25	1.02	1.07
Intermed	1.08	1.09	0.98	1.01	0.83	0.93	1.12	0.96	1.00

Sample size = 2,193, of which 1,059 are graduate, 759 intermediate and 415 primary. Ratios are based on median values in each group.
Source: RLMS.

relative return to graduate education than the counterfactual estimates. There is less difference in the estimates of the relative returns for the intermediate group, since the incidence of arrears does not vary much compared with the default group.

We now turn to two dimensions that have the largest explanatory power in the incidence of wage arrears estimates, namely region and industry. We divide the sample into two areas: those living in Moscow and St. Petersburg (Metro), where the incidence of wage arrears is low and wages are high and those living outside the major metropolitan areas where wages are lower and the incidence of wage arrears is high. In Table 6, the actual distribution suggests that there is a 100% median wage gain from living in the metropolitan areas. Accounting for the skewed incidence of wage arrears by region reduces this regional wage premium to around 30%.

In Table 7 we aggregate industries into two sectors, production and services. Table A2 suggests that workers in the former are more likely to experience wage arrears than workers in the latter. The actual distribution suggests a median pay penalty in production relative to services. However, since the production sector is affected more by wage arrears, if everyone were paid in full this would be sufficient to generate a small pay premium for the production sector. Again the different imputation methods are broadly in agreement.

Table 6. Actual and Counterfactual Regional Wage Ratios (1996).

	Actual	No Arrears	OLS I	OLS II	JMP	DFL	Match	PS I	PS II
Metro									
Mean	758	1,073	971	1,072	1,034	972	1,074	1,000	971
Median	573	845	821	803	802	803	917	803	788
Other									
Mean	462	847	720	815	824	819	817	861	838
Median	275	654	588	588	650	573	642	676	631
Ratio:wrt other									
Metro	2.08	1.30	1.22	1.37	1.23	1.40	1.43	1.19	1.25

Sample size = 2,193, of which 332 are metropolitan, 1,702 are elsewhere. Ratios are based on median values in each group.
Source: RLMS.

Table 7. Actual and Counterfactual Industry Wage Ratios (1996).

	Actual	No Arrears	OLS I	OLS II	JMP	DFL	Match	PS I	PS II
Production									
Mean	491	991	796	884	910	876	897	887	884
Median	281	784	675	642	739	596	748	688	654
Services									
Mean	531	843	730	835	814	813	824	882	839
Median	344	676	603	630	619	653	654	688	631
Ratio									
wrt services	0.82	1.16	1.12	1.02	1.19	0.91	1.14	1.00	1.04

Sample size = 2,193, of which 975 are production and 1,059 services. Ratios are based on median values in each group.

One striking feature is that the parameters of the counterfactual wage distributions are very similar to the parameters of the observed wage distributions of those not in arrears. While this does not mean that experience of wage arrears is a random event as confirmed by evidence in Earle and Sabirianova (2002) and Lehmann et al. (1999), it does suggest that those in arrears are drawn from throughout the underlying wage distribution. Fig. 2 seems to confirm this. For those wishing to study aspects of wage differentials and inequality in Russia, it may, therefore, be feasible to use the subset of those not in arrears to estimate the population parameters, subject to an efficiency loss.

7. CONCLUSIONS

It seems apparent that estimates of wage inequality, and pay gaps in general, can be affected strongly in countries that experience bouts of wage arrears. Studies that fail to account for wage arrears can over-estimate wage inequality substantially in countries where arrears are eventually paid back. In countries where wage arrears are never paid back the actual wage distribution is more relevant for measuring inequality, assessing welfare costs and formulating appropriate policy responses. In countries where arrears are paid back, pay gaps across sub-groups of the population could be misleading if no account is taken of the differing incidence of wage arrears across these sub-groups. Russia in the 1990s, having both one of the highest levels of wage inequality and a large incidence of wage arrears, is a particularly interesting case. The large share of employees who receive no wages in any month also renders many conventional estimates of inequality based on logarithmic transformations inoperable.

Using imputation techniques that could be applicable to any data set for any country with information on wages and wage arrears, we show that in the absence of arrears average earnings would be some 20–50% higher, depending on the extent of arrears and that earnings dispersion would be lower by similar amounts if everyone were paid in full. This conclusion is broadly the same whatever imputation method is used. This would put Russian wage inequality back toward levels currently experienced in Western countries like Britain. In the absence of arrears, the gender pay gap could be around 10 percentage points higher than the observed gap, though the imputation methods are less in agreement in this regard. Regional pay differentials would become more compressed and sectoral differentials would narrow in the absence of wage arrears. In this particular study, it appears that those in arrears are drawn from throughout the underlying wage distribution. For those wishing to study wage differentials and inequality, it may for Russia, be feasible to use the subset of those not in arrears and get close to the true population parameters.

NOTES

1. A glance at the BBC web site: www.bbc.co.uk contains reports on unpaid wages in Argentina, Azerbaijan, Belarus, Bulgaria, Central African Republic, China, Colombia, Honduras, Iran, Kazakhstan, Kenya, Kosovo, Mexico, Niger and the Ukraine as well as Russia over the last 5 years. Following the introduction of the

national minimum wage in Britain in 1999, a recent report indicates that some 36% of firms were underpaying their minimum wage workers. A glance at the BBC web site: www.bbc.co.uk contains reports on unpaid wages in Argentina, Azerbaijan, Belarus, Bulgaria, Central African Republic, China, Colombia, Honduras, Iran, Kazakhstan, Kenya, Kosovo, Mexico, Niger and the Ukraine as well as Russia over the last 5 years. Following the introduction of the national minimum wage in Britain in 1999, a recent report indicates that some 36% of firms were underpaying their minimum wage workers http://news.bbc.co.uk/1/hi/business/2255947.stm.

2. Over the same period, the Gini indices for wages in CEE grew from levels in the range of 0.2–0.25 to levels in the range of 0.3–0.35. In Chile, the Gini coefficient is around 0.45 and in Turkey around 0.37.

3. Ogloblin (1999) is an exception, using a selection equation in his analysis of the mean gender pay gap in Russia.

4. With no trade-off between wage arrears and employment the counterfactual becomes the actual underlying wage distribution that would occur in the absence of arrears.

5. These are imperfect estimates, since the counterfactuals ignore the losses in earnings over time due to inflation, foregone interest and the costs of borrowing. However incidences of wage arrears in Russia were much higher after the hyper-inflations of the mid-1990s when inflation rates were back to single figures (Gimpelson, 2000). The RLMS data do not give the dates of when arrears occurred so it is not possible to ascertain the dynamic history of the wage arrears process needed to infer inflation, interest and borrowing costs.

6. Firpo (2004) demonstrates that it is possible to estimate the quantiles directly without first estimating the counterfactual distribution. For estimates of the conditional variance or quantiles of the distribution, Fröhlich (2003) shows that while matching on a set of covariates X is consistent, propensity score matching is not.

7. Imbens (2004) notes that the "debate concerning the practical advantage of the various estimators … is still ongoing with no firm conclusions yet reached."

8. The set of controls include individual controls for age, gender, education and tenure job-level controls for 1 digit industry, firm size and region (see Appendix Table A1). Lehmann et al. (1999) and Earle and Sabirianova (2000, 2002) find that job and location rather than individual characteristics are the more relevant predictors of the incidence of arrears.

9. This is not always the case in our data.

10. Sample size constraints prevent us from matching within all eight macro regions identified by the data and used in the OLS estimates. Also, while within regional mobility may be affected by arrears, the regions in the RLMS are so large as to make mobility between regions as a result of arrears unlikely.

11. The IZA discussion paper version of this paper presents one and four-year earnings transition matrices. While there is a degree of mobility, there is considerably less among those not in arrears.

12. See also Kluve, Lehmann, and Schmidt (2001). The literature stresses that there seems to be a bias vs. efficiency trade-off between non-parametric and propensity score matching. Smith and Todd (2001) show that estimates from different propensity score matching methods do not vary much as long as the conditioning variables satisfy the requirements set out by Heckman et al. (1997).

13. The RLMS is ambiguous on the nature of self-employment, referring instead to the extent of self-ownership in the enterprise where the individual works. We exclude only those who say they own between 51 and 100% of the enterprise.

14. A question on the contractual wage appears for the first time in 1998, but the responses given for those in arrears unfortunately hardly differ from the actual wage responses. Therefore, we cannot use this information.

15. This comprises less than 1% of those at the bottom of the no treatment group and less than 1% of those at the top of the wage distribution.

16. Given a sufficiently large sample this would, of course, be possible. In what follows we capture gender effects through a simple intercept dummy variable.

17. Imbens (2004) questions the validity of bootstrap-based standard errors in the case of matching. In practice, the subsequent tables show that the standard errors of the matching estimates do not differ markedly from the standard errors of the estimates derived from the other methods.

18. VTsIOM is the Russian acronym for the All-Union Center for the Study of Public Opinion.

19. The data for Poland are restricted to full-time workers only, though, as in Russia, part-time working amounts to less than 3% of the Polish workforce.

20. Other quantiles and moments of the distributions are available on request.

21. The OLS estimates without the added residual are used only as a benchmark to highlight the problem of distribution imputation based solely on predicted values from a wage equation and we do not recommend that this technique be used to estimate counterfactuals.

22. Table A3 gives marginal effects from logit estimates of the probability of being in arrears. The same estimates are used to generate the counterfactual kernel density estimates.

23. The differences for PSI and PSII relative to the other methods are not caused by the chosen parametric specification of the prediction equation, since the results are very similar across different specifications. Nor do the results vary significantly depending on the propensity score matching method. Results available on request.

ACKNOWLEDGMENTS

The authors are grateful to John DiNardo and Jochen Kluve for making their STATA routines available; to Boris Augurzky, Todd Idson, Alan Manning, Mark Schaffer, Christoph Schmidt, Jan Svejnar, Paul Walsh, and two anonymous referees, participants at the CEPR-WDI Conference on Transition in Moscow as well as seminar participants at the universities of Heidelberg, Heriot-Watt University, the London School of Economics, the University of Trento and the WDI, University of Michigan for valuable comments. The Fritz Thyssen Foundation provided financial support within the project "Economic Reform and the Microeconomics of Labor Market Adjustment in the Russian Federation." The authors are also grateful to IZA for providing the facilities for the authors to undertake research on this paper.

REFERENCES

Biewen, M. (2001). Item non-response and inequality measurement: Evidence from the German earnings distribution. *Journal of the German Statistical Society (Allgemeines Statistisches Archiv)*, *85*, 409–425.

Blau, F., & Kahn, L. (1996). International differences in male wage inequality. *Journal of Political Economy*, *104*(4), 791–836.

Brainerd, E. (1998). Winners and losers in Russia's economic transition. *American Economic Review*, *88*(5), 1094–1116.

Dehejia, R., & Wahba, S. (2002). Propensity score matching methods for non-experimental causal studies. *Review of Economics and Statistics*, *84*(1), 151–161.

Desai, P., & Idson, T. (2000). *Work without wages: Russia's nonpayment crisis*. Cambridge, MA: MIT Press.

DiNardo, J., Fortin, N., & Lemieux, T. (1996). Labour market institutions and the distribution of wages. 1979–1992. *Econometrica*, *64*(5), 1001–1044.

Earle, J. & Sabirianova, K. (2000). *Equilibrium wage arrears: A theoretical and empirical analysis of institutional lock-in*. IZA Discussion Paper 196.

Earle, J., & Sabirianova, K. (2002). Understanding wage arrears in Russia. *Journal of Labor Economics*, *20*(3), 661–707.

Firpo, S. (2004). *Efficient semiparametric estimation of quantile treatment effects*. University of British Columbia, Economics Department Discussion Paper 04-01, January.

Flemming, J., & Micklewright, J. (1997). *Income distribution, economic systems and transition*. London: EBRD.

Fröhlich, M. (2003). *Propensity score matching without conditional independence assumption*. Department of Economics, University of St. Gallen Working Paper.

Gimpelson, V. (2000). Politics of labor market adjustment. The case of Russia. In: J. Kornai, S. Haggard & R. Kaufman (Eds), *Reforming the state: Fiscal and welfare reform in post-socialist countries*. New York: Cambridge University Press.

Heckman, J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator. *Review of Economic Studies*, *65*(2), 261–294.

Hellman, J. (1998). Winners take all: The politics of partial reform in postcommunist transition world politics. *World Politics*, *50*(2), 203–234.

Imbens, G. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, *86*(1), 4–29.

Juhn, C., Murphy, K., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, *101*(3), 410–442.

Kluve, J., Lehmann, H., & Schmidt, C. (1999). Active labour market policies in Poland: Human capital enhancement, stigmatization or benefit churning? *Journal of Comparative Economics*, *27*(1), 61–89.

Kluve, J., Lehmann, H., & Schmidt, C. (2001). *Disentangling treatment effects of polish active labour market policies: Evidence from matched samples*. IZA Discussion paper 355, September.

Lehmann, H., Wadsworth, J., & Aquisti, A. (1999). Grime and punishment: Job insecurity and wage arrears in the Russian Federation. *Journal of Comparative Economics*, *27*(5), 595–617.

Lechner, M. (2002). Program heterogeneity and propensity score matching. *Review of Economics and Statistics*, *84*(2), 205–220.

Oaxaca, R. (1973). Male–female earnings differentials in urban labor markets. *International Economic Review*, *14*(5), 693–709.

Ogloblin, C. G. (1999). The gender earnings differential in the Russian transition economy. *Industrial and Labor Relations Review*, *52*(4), 602–627.

Smith, J. A., & Todd, P. E. (2001). Reconciling conflicting evidence on the performance of propensity-score matching methods. *American Economic Review*, *91*(2), 114–118.

APPENDIX

Payroll data from a sample of 19 firms in a central Russian industrial city is depicted in Fig. A1. The matching algorithm is shown in Box A1.

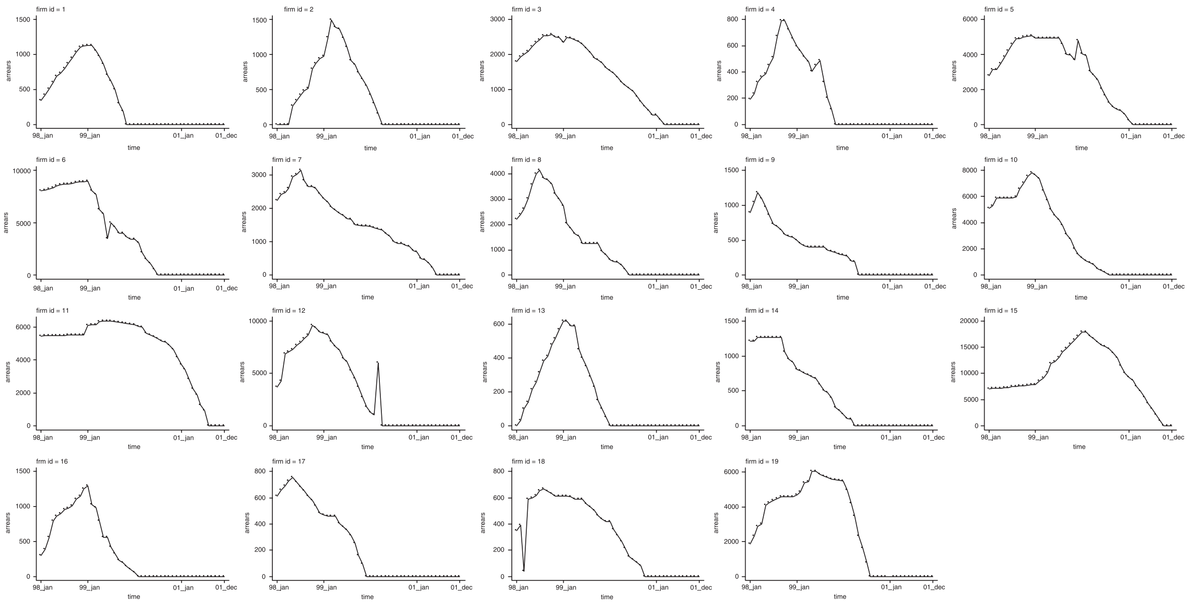


Fig. A1. Monthly Stock of Wage Arrears within Russian Firms (City of Ryazan – 1998–2001) *Source*: Authors' Calculations Based on CERT Regional Firm Data.

Box A1. Exact Matching – Algorithm and Scheme of Conditioning on Pre-Treatment History.

Exact matching algorithm

I. Condition on following possible pre-treatment labor market history:
– employed and fully paid and in *x*-th decile of wage distribution
– unemployed
– inactive
– employed and experiencing wage arrears (WA)

II. Match treated individuals to individuals with same pre-treatment history using following observable characteristics:
– gender
– region (4 categories)
– qualifications (6 categories)
– age (maximum allowed difference of 10 years – choose those controls that have the minimum age difference)

Assumption: these variables are not affected by the treatment (WA).
Because treated are more than potential controls, matching is done with replacement.

III. Assign wage of matched control to treated individual, or assign average of wages of matched controls

Scheme of Conditioning on pre-treatment history by example

Pre-treatment period	*Treatment period*
Potential Control 1 in 95 Employed and fully paid and in 2nd decile of wage distribution	Potential Control 1 in 96 Employed and fully paid
Treated 1 in 95 Employed and fully paid and in 2nd decile of wage distribution	Treated 1 in 96 In wage arrears
Potential Control 2 in 95 Unemployed	Potential Control 2 in 96 Employed and fully paid
Treated 2 in 95 Unemployed	Treated 2 in 96 In wage arrears

Table A1. OLS Log Real Weekly Wage Estimates for those not in Arrears.

	1994	1996	1998
Female	−0.430 (0.033)**	−0.446 (0.048)**	−0.417 (0.047)**
Age	0.056 (0.009)**	0.057 (0.012)**	0.052 (0.012)**
Age2	−0.001 (0.000)**	−0.001 (0.000)**	−0.001 (0.000)**
University	0.512 (0.051)**	0.251 (0.070)**	0.456 (0.076)**
Technical	0.302 (0.049)**	0.084 (0.069)	0.193 (0.074)**
PTU 1	0.090 (0.055)	−0.094 (0.080)	−0.042 (0.081)
PTU 2	0.052 (0.065)	0.004 (0.093)	−0.035 (0.100)
Other Quals.	0.052 (0.059)	−0.125 (0.089)	−0.061 (0.090)
North West	0.088 (0.072)	−0.063 (0.104)	−0.165 (0.112)
Central	−0.349 (0.052)**	−0.313 (0.069)**	−0.311 (0.070)**
Volga	−0.509 (0.054)**	−0.528 (0.078)**	−0.462 (0.081)**
Caucasus	−0.479 (0.060)**	−0.310 (0.090)**	−0.438 (0.086)**
Urals	−0.229 (0.056)**	−0.232 (0.078)**	−0.297 (0.079)**
Western Siberia	0.119 (0.065)	0.278 (0.098)**	0.281 (0.100)**
East	−0.014 (0.068)	−0.098 (0.112)	−0.178 (0.101)
State	−0.115 (0.034)**	−0.162 (0.051)**	−0.229 (0.050)**
Agriculture	−0.271 (0.094)**	−0.352 (0.143)**	−0.190 (0.109)
Manufacturing	0.084 (0.062)	0.149 (0.091)	−0.028 (0.079)
Construction	0.303 (0.081)**	0.459 (0.131)**	0.120 (0.132)
Energy	0.331 (0.072)**	0.423 (0.108)**	0.313 (0.096)**
Transport	0.287 (0.070)**	0.373 (0.102)**	0.196 (0.088)**
Retail	0.073 (0.069)	0.162 (0.095)	0.163 (0.081)**
Finance	0.411 (0.121)**	0.634 (0.145)**	0.248 (0.130)
Health/Education	−0.098 (0.058)	0.052 (0.087)	−0.186 (0.076)**
Firm size 11–50	0.040 (0.063)	0.038 (0.094)	0.044 (0.093)
Firm size 51–100	0.093 (0.072)	0.048 (0.109)	0.110 (0.105)
Firm size 101–500	0.176 (0.064)**	0.127 (0.101)	0.117 (0.101)
Firm size 501–1000	0.277 (0.068)**	0.171 (0.106)	0.403 (0.105)**
Firm size missing	0.109 (0.064)	−0.027 (0.090)	0.090 (0.093)
Job Tenure 1–2 years	0.076 (0.053)	0.177 (0.080)**	0.112 (0.076)
2–5 years	−0.026 (0.048)	0.252 (0.068)**	0.117 (0.067)
5–10 years	0.021 (0.052)	0.107 (0.077)	0.183 (0.076)**
10–20 years	0.081 (0.051)	0.201 (0.077)**	0.292 (0.082)**
20+ years	0.224 (0.060)**	0.243 (0.089)**	0.215 (0.092)**
Constant	5.470 (0.190)**	5.635 (0.255)**	5.373 (0.268)**
N	2,213	1,019	1,091
R^2	0.31	0.31	0.31

Standard errors in parentheses. ** Significant at 5%. Default region is metropolitan Moscow & St. Petersburg. Default industry is other services.

Table A2. Counterfactual Average Real Wage Distributions (1994–1996).

	Mean	90th P'tile	Median	10th P'tile	90/10	90/50	50/10	Coefficient Variable	Gini
Actual	535	1,250	375	0	N/a	3.3	N/a	1.11	0.555
OLS I	731 (22)	1,215 (58)	625 (19)	360 (12)	3.4	1.9	1.7	0.58 (.02)	0.284 (.009)
OLS I_ RE	746 (26)	1,246 (64)	641 (21)	370 (12)	3.4	1.9	1.7	0.58 (.02)	0.282 (.010)
OLS II	842 (24)	1,674 (55)	633 (19)	248 (11)	6.8	2.6	2.6	0.85 (.03)	0.401 (.008)
OLS II_RE	861 (30)	1,670 (63)	654 (20)	252 (10)	6.6	2.6	2.6	0.85 (.03)	0.400 (.009)
JMP	758 (13)	1,536 (36)	581 (11)	213 (12)	7.2	2.6	2.7	0.80 (.02)	0.400 (.007)
JMP_RE	750 (13)	1,518 (39)	573 (10)	219 (14)	6.9	2.6	2.6	0.80 (.02)	0.400 (.009)
DFL	753 (16)	1,562 (80)	573 (15)	188 (5)	8.3	2.7	3.0	0.82 (.01)	0.405 (.007)
DFL_RE	732 (16)	1,562 (82)	530 (16)	181 (7)	8.6	2. 9	2.9	0.86 (.01)	0.416 (.009)

Note: RE = counterfactual based on random effects regressions for prediction equations.

Table A3. Logit Estimates of Probability of Being in Arrears (Marginal Effects).

	1994	1996	1998
Female	−0.070 (0.019)**	−0.037 (0.021)	−0.018 (0.018)
Age	0.012 (0.005)**	0.007 (0.006)	0.008 (0.005)
Age2	−0.0002 (0.00006)**	−0.0001 (0.0001)	−0.0001 (0.0001)
University	0.030 (0.029)	−0.084 (0.031)**	−0.086 (0.029)**
Technical	0.031 (0.028)	−0.030 (0.029)	−0.061 (0.029)**
PTU 1	−0.007 (0.030)	0.001 (0.033)	−0.049 (0.032)
PTU 2	0.018 (0.036)	−0.091 (0.043)**	−0.029 (0.038)
Other Quals.	0.031 (0.032)	0.054 (0.033)	−0.044 (0.035)
North West	0.204 (0.042)**	0.326 (0.047)**	0.382 (0.046)**
Central	0.070 (0.034)**	0.119 (0.037)**	0.151 (0.034)**
Volga	0.122 (0.034)**	0.278 (0.039)**	0.319 (0.036)**
Caucasus	0.083 (0.039)**	0.247 (0.044)**	0.218 (0.040)**
Urals	0.126 (0.035)**	0.257 (0.039)**	0.259 (0.036)**
Western Siberia	0.145 (0.039)**	0.333 (0.044)**	0.299 (0.042)**
East	0.252 (0.039)**	0.429 (0.049)**	0.358 (0.043)**
State	0.079 (0.019)**	0.051 (0.022)**	0.109 (0.019)**
Agriculture	0.262 (0.045)**	0.216 (0.057)**	0.074 (0.042)
Manufacturing	0.071 (0.034)**	0.156 (0.042)**	0.162 (0.031)**
Construction	0.152 (0.042)**	0.142 (0.055)**	0.183 (0.048)**
Energy	−0.063 (0.041)	0.047 (0.046)	0.057 (0.037)
Transport	−0.055 (0.039)	−0.067 (0.047)	−0.022 (0.036)
Retail	−0.105 (0.042)**	−0.143 (0.048)**	−0.175 (0.038)**
Finance	−0.254 (0.098)**	−0.444 (0.111)**	−0.338 (0.078)**
Health/Education	−0.110 (0.032)**	0.081 (0.040)**	0.130 (0.030)**
Firm size 11–50	0.062 (0.038)	−0.031 (0.045)	−0.031 (0.041)
Firm size 51–100	0.021 (0.043)	0.056 (0.046)	−0.042 (0.046)
Firm size 101–500	0.007 (0.038)	0.094 (0.042)**	0.030 (0.041)
Firm size 501–1,000	0.074 (0.041)	0.072 (0.045)	0.009 (0.043)
Firm size missing	0.042 (0.039)	−0.040 (0.044)	−0.015 (0.040)
Job Tenure 1–2 years	0.007 (0.032)	0.039 (0.037)	−0.027 (0.032)
2–5 years	0.066 (0.028)**	0.005 (0.031)	−0.024 (0.043)
5–10 years	0.069 (0.030)**	0.053 (0.034)	−0.027 (0.027)
10–20 years	0.089 (0.030)**	0.074 (0.034)**	0.025 (0.032)
20+ years	0.102 (0.035)**	0.106 (0.038)**	0.031 (0.036)
Rural	0.207 (0.025)**	0.197 (0.030)**	0.182 (0.027)**
N	3,962	2,884	3,336
Log *L*	−2,448	−1,590	−1,831

Standard errors in parentheses. ** Significant at 5%.

COMPUTERS AND THE WAGE STRUCTURE

Michael J. Handel

ABSTRACT

A leading explanation for the growth of wage inequality is that greater use of information technology increased the demand for human capital. This paper identifies four different explanations for the relationships between computers, skills, and wages: computer-specific human capital, greater general human capital among computer users, greater general human capital for both users and nonusers due to contextual effects, and skill-biased changes in the job composition of the workforce. The paper tests the first three explanations and finds little support for them once pre-computer and other job characteristics are adequately controlled. This conclusion receives further support from a comparison of the timing of inequality growth and computer diffusion and from analyses of the contribution of computer use to overall inequality growth using DiNardo, Fortin, and Lemieux's (1996) reweighting standardization technique.

1. INTRODUCTION

Wage inequality in the United States has grown dramatically since the late 1970s (Katz & Murphy, 1992; Levy & Murnane, 1992; Gottschalk, 1997;

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 157–198

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26005-9

Katz & Autor, 1999). Considerable debate persists, however, over the reasons for this growth.

Increased returns to education led many to argue that skill requirements were rising as a result of the spread of new technology, drawing attention to the possible role of computers, which remains the leading explanation of rising inequality. In this view, consistent with longstanding theories of capital-skill complementarity, computers are a form of skill-biased technological change (SBTC) that increases the relative demand for more educated workers and raises their wages relative to the less educated (Katz & Murphy, 1992; Krueger, 1993; Berman, Bound, & Griliches, 1994; Autor, Katz, & Krueger, 1998).

However, other studies indicate problems with explanations of inequality growth based on the increase in computer-driven skill demand (DiNardo & Pischke, 1997; Howell & Wieler, 1998; Mishel & Bernstein, 1998; Entorf, Gollac, & Kramarz, 1999; Haisken-DeNew & Schmidt, 1999; Card & DiNardo, 2002). A recent, intensive examination of the issue concluded, "Overall, the evidence linking rising wage inequality to SBTC is surprisingly weak" (Card & DiNardo, 2002, p. 776).

Alternative explanations of inequality growth have focused on institutional changes, broadly conceived, such as deunionization, declining real value of the minimum wage, deregulation, sectoral shifts from manufacturing to services, and the role of the recession of the early 1980s in altering the relative bargaining power of capital and labor, pay norms, and informal rent-sharing bargains to the disadvantage of those in the lower half of the wage distribution (DiNardo, Fortin, & Lemieux, 1996; Fortin & Lemieux, 1997; Howell, 1997; Galbraith, 1998).

One unresolved question even within the SBTC framework is the precise mechanism by which computers affect skill demand. This paper identifies four distinct mechanisms by which computers may affect wages and inequality that are implicit in the SBTC literature and tests three of them, as well as noting some concerns associated with the fourth.

An additional limitation of most SBTC studies is that they do not directly address whether computers account for a large proportion of the growth of overall inequality. Two issues are relevant: whether the temporal trends in inequality growth and computer diffusion are consistent and the magnitude of possible computer effects on changes in the overall dispersion of wages. This paper examines both the issues of temporal consistency and whether computers can account for a large share of inequality growth.

The rest of the paper is organized as follows. Section 2 identifies the different causal arguments embedded in the SBTC literature. Sections 3 and

4 describe the data and test three explanations of computers' effects on wages. Section 5 considers whether computers can account for a large share of inequality growth by comparing temporal patterns of the inequality growth and computer diffusion and using more formal standardization techniques. Section 6 concludes. Results suggest that the effects of computers on wages and human capital requirements are modest at best and the spread of computer use at work is an unlikely candidate for explaining a large part of inequality growth.

2. EXISTING RESEARCH

Initial research on inequality growth simply inferred technological change from the dramatic growth in the returns to education in the 1980s, pointing to what appeared to be the contemporaneous growth in computer use for support (Katz & Murphy, 1992). Subsequent studies sought to provide more direct evidence of links between computers, wages, and inequality growth, and eventually produced a range of causal accounts that are still not fully sorted out and continue to be debated.

The first direct evidence of a connection between computer use and inequality was Krueger's (1993) widely cited study that found a wage premium associated with computer use on the job in the United States, net of standard human capital variables, on the order of 17–19% in the 1980s. After considering different specifications, Krueger concluded that actual returns to computer use likely ranged from 10% to 15%. Computer use also explained about 40% of the 0.01 increase in the return to years of education between 1984 and 1989. The robustness of these results to a number of sensitivity tests strengthened the interpretation that computer skills specifically are highly rewarded in the labor market. Krueger (1993) concluded that government policy could moderate inequality through increased public funding for computer training programs until computer skills were sufficiently common that the wage premium fell.

In this view, the increased demand for *computer-specific human capital* increased wages for more skilled workers, presumably because the training and knowledge needed to operate computers is costly, difficult to acquire, and relatively scarce.

However, skepticism grew after DiNardo and Pischke's (1996, 1997) analysis of German data found that many other job characteristics, such as the use of calculators, telephones, and pens or pencils at work or even sitting down while working were associated with wage premiums comparable to computer use when each was entered individually in a standard wage equation. DiNardo and Pischke argued that the size of the coefficients were so similar and many of the variables so removed from what are conventionally considered scarce, productivity-enhancing skills that the results could not be taken at face value. The measured effects of pencils and computers likely reflected associations with some unobserved aspect of either human capital or occupational position, rather than returns to specific, identifiable skills *per se*.

Krueger (2000) replied that while the German results were suggestive, spurious results for other job characteristics need not imply a similar problem for computer use. Indeed, estimated returns to computer use in the German data were larger than those associated with the other job characteristics when all were entered jointly in a single model. Likewise, there was no evidence indicating the German results generalized to the United States (Krueger, 2000).

In support of the last point, one might note that Germany did not experience rising returns to education or inequality growth as did the United States, undoubtedly partly due to labor market institutions that tend to equalize wages across skill groups (Gottschalk & Smeeding, 1997; Freeman & Katz, 1994). It is possible that German labor market institutions also dampened returns to computer use. Whether or not DiNardo and Pischke's (1997) results generalize to the United States remains unknown.

The computer wage premium and its role in inequality growth remains an object of study and debate, as has the more general proposition that computer skills specifically play an important role in labor market dynamics, e.g., early retirement decisions of older workers (Entorf et al., 1999; Cappelli & Carter, 2000; Goss & Phillips, 2002; Black & Lynch, 2004; Borghans & ter Weel, 2004; Lee & Kim, 2004; Friedberg, 2003; Dickerson & Green, 2004).

However, this research continues to wrestle with the problem of controlling for unobserved heterogeneity. There is a strong probability that computer use is associated with other characteristics that are usually unobserved yet causally related to wages, yielding regression coefficients that reflect partly the effects of the unobservables as well as any true effects of computer skills themselves. Preexisting differences between individuals, jobs, or firms may account for both higher wages and the likelihood of using a computer at work.

For example, one might find that the manager of a construction firm uses a computer at work and receives higher pay than a carpenter working at the firm. This would not be surprising, but the pay differential would not be due to computers since this kind of difference long predates the introduction of computers. The disparity in computer use does not necessarily imply anything about relative skill requirements, but may merely reflect the fact that computers are not much use for carpentry or most other manual tasks, while they are very useful for office work.

This points to a significant ambiguity in the concept of *complementarity* between computers and skilled workers. If complementarity means merely that computers are more useful in certain kinds of jobs than others, the concept is uncontroversial. No one disputes that computers are associated more with office work than manual or service work, which are often less skilled. But this does not reflect the skill required to use computers as much as the nature of the technology itself and the kinds of tasks it can handle. Computers are better at internal symbol manipulation than external object manipulation, particularly in physical work environments that are unstandardized and require visual perception or manual dexterity that is not easily codified. Borghans and ter Weel (2002, pp. 152–153) confirm that computer use is positively associated with job tasks involving reading and math and negatively associated with those requiring physical exertion. Indeed, the association of computers and work roles is so significant that three-digit occupation accounts for 40% of the variance in computer use and one-digit occupation accounts for nearly 30%, while education, the usual measure of skill, accounts for less than 15% (author's calculation).[1]

The real issue is not whether computer use *accompanies* certain kinds of job tasks more than others, but whether the introduction of computers *increases* skill requirements within jobs. Presumably, the point of SBTC is not that computers are simply good markers for jobs whose high skill and status antedate computer diffusion and whose tasks happened to be more suited to computer assitance. If only this weaker form of complementarity holds, meaning computers are associated with but have no causal impact on skills, then workers observed to use computers in recent data sets would receive similar wage premiums in both pre- and post-computer labor markets. In this case, increased computer use would not be a contributor to inequality growth.[2]

Identifying a causal role for computers in increasing wage disparities, complementarity in the stronger sense, requires effectively controlling for pre-computer sources of wage inequality that may be picked up by computer use in later years simply because computers proved more useful in certain jobs than others. Controlling for pre-computer characteristics is especially important since the strong version of capital-skill complementarity asserts that new technologies increase rewards for the already well rewarded. In other words, the theory itself predicts nonrandom assignment.

Therefore, it is critical to control for pre-computer base-level rewards in testing for computer-induced changes to avoid correlation of the predictor with the error term.

Substantively, a key issue raised by this concern is whether the knowledge needed to operate computers represents a large increment to human capital requirements. There are reasons to doubt that this is the case. Researchers reported short training times for the most commonly used computer skills, such as word processing, even early in the computer diffusion process, when few users had prior experience with computers at school or work (Goldstein & Fraser, 1985; Levin & Rumberger, 1986; Carroll, 1997). Software usability has also improved over time, particularly with the development and spread of graphical user interfaces, which partly reflected competitive pressures in the software product market to improve user friendliness in order to increase sales (Carroll, 1997). However, even though far more people know how to use computers and computers themselves have become easier to use over time, the measured returns to computer use did not decline between 1984 and 1993, but rose slightly and remained substantially unchanged through 2001 (Autor, Katz, & Krueger, 1997, pp. 1187, and author's calculations).

The original computer premium study implicitly acknowledged the potential for omitted variable bias in settling on an estimated computer wage premium below the 17–19% that resulted from adding a computer indicator to a standard wage equation (Krueger, 1993). But even a computer wage premium of 10–15% is similar in magnitude to the returns to one and a half years of schooling, which seems too high on intuitive grounds. The true magnitude and wage implications of computer-specific human capital remain unsettled.

The perceived problems with the original theory of computer-driven SBTC prompted a number of alternatives that do not rest on claims that computer skills per se are particularly scarce or complex. While the causal mechanisms are usually somewhat implicit, they can be classified into three broad categories. Two of them, like the computer-specific human capital argument, imply computers affect the task content of jobs (within-job effects), while the third implies computers alter the job or occupational composition of employment (between-job effects). These hypotheses represent distinct arguments but are not mutually exclusive, and some studies implicitly invoke more than one.

The first alternative to the computer-specific human capital argument holds that computer use requires higher levels of *general human capital among computer users* because computers transform work into a more knowledge-intensive activity in various ways. For example, considering the technology in a narrow sense, the claim is that computers reduce the need for physical strength, intuition, tacit skills, and routine cognitive activity, while increasing users' needs for formal literacy, numeracy, and higher-order, abstract, symbolic, and procedural reasoning skills (Zuboff, 1988; Levy & Murnane, 1996; Levy, Beamish, Murnane, & Autor, 1999; Autor, Murnane, & Levy, 2002; Fernandez, 2001).

More broadly, computers also make it feasible to reverse the division of labor for less skilled jobs by reintegrating more skilled tasks into previously narrow jobs designed according to Scientific Management principles (Zuboff, 1988; Autor et al., 2002; Fernandez, 2001). A commonly cited example is customer service representatives who are given responsibility for an integrated customer database and upgraded into low-level account managers (Murnane & Levy, 1996). Other examples might be secretaries whose jobs are upgraded to administrative assistants with the addition of simple bookkeeping and similar tasks as a result of office software and forklift operators given clerical tasks related to computerized inventory control.

More broadly still, it is argued that computers give more workers access to information and therefore make it more rational for employers to adopt participative management techniques, often called "high performance workplace practices," which involve greater restructuring of work roles involving the downward delegation of decision making, problem solving, and quality-control responsibilities to less skilled workers closer to the point of production (Zuboff, 1988; Siegel, 1999; Autor et al., 2002; Fernandez, 2001; Bresnahan, Brynjolfsson, & Hitt, 2002; Bartel, Ichniowski, & Shaw, 2000; Shaw, 2002).

According to this view, what is most important is that for various reasons computers require users to have more general cognitive skills, such as problem solving and intellectual flexibility, compared to otherwise similar nonusers; the human capital embodied in knowledge of specific software applications themselves is less central.[3]

Testing this alternative hypothesis can be problematic because at first sight it suggests model specifications similar to the original computer wage premium literature, with simply a different interpretation of the estimated effects of computer use (general vs. computer-specific human capital). Operationally distinguishing the two explanations, as well as determining whether observed associations are causal, remains problematic. The data used for this paper have a rich set of measures relating to job literacy requirements as well as computer use that help to distinguish and test both general and computer-specific human capital versions of SBTC. Models below test whether indicators of general human capital fully mediate the effects of computer use on wages.

A second, less developed, alternative to the initial computer premium literature argues that computers transform the workplace so thoroughly that they increase the demand for *general human capital among nonusers* as well as computer users. The exact reasoning behind this view remains somewhat elusive and empirical research in this vein remains thin. Piecing together various hints, the claim appears to be that computerization of organizations generates more information for users and nonusers to process, interpret, and use for creating product and process innovations (Autor et al., 1998; Bresnahan, 1999; Bresnahan et al., 2002). "This raises the marginal product of skilled workers' ideas, even if those workers never see a computer" (Bresnahan, 1999, p. F410). Perhaps another mechanism by which nonusers might be affected is if computerization stimulates adoption of high performance work practices for jobs throughout a firm regardless of whether the jobs involve computer use.

This is a contextual effects argument insofar as it argues that a computerized work environment requires employees to have more skills than otherwise even if they do not use a computer themselves. Bresnahan (1999, p. F400) goes so far as to concede the entire substance of DiNardo and Pischke's (1997) case against the computer wage premium literature. He argues for what he sees as the neglected importance of what he calls organizational computing, such as corporate accounting systems, which are traditionally associated with mainframe or other large-scale systems, as opposed to the individual productivity software associated with personal computers (e.g., word processing). The claim is that the contextual, organization-level, effect of computers on skill requirements almost entirely dominates the individual-level skill effects of actual computer use. The argument is provocative, but much vaguer on the precise causal mechanism through which computerization affects the work of nonusers.

As with the initial computer premium literature, existing research dealing with the general human capital and organization-level effects of computers is not conclusive. Some of the important works are qualitative case studies that are suggestive but whose generality is unknown (e.g., Zuboff, 1988; Levy & Murnane, 1996; Shaw, 2002). Some quantitative studies suggest small effects of computers on general cognitive skill requirements of jobs (Keefe, 1991; Fernandez, 2001). Empirical studies of high performance work systems also suggest modest skill and wage impacts; a review of 18 studies and 87 coefficients suggest wage effects on the order of 0–5% (Handel & Levine, 2004).

Levy and Murnane (1996) show that computers shift accountants' work time away from routine calculations toward more purely accounting tasks, but this represents a more efficient utilization of skills demanded long before the era of computers rather than an increase in their required level of complexity. Likewise, managers in a computerized organization may have higher quality information for decision making, but managers were always hired for their decision-making skills. It is not obvious that computers would raise hiring standards significantly in these circumstances. Unless computers require greater human capital investment it is unclear why, under standard assumptions, the gains resulting from computer investment and increased skilled worker productivity would accrue to labor rather than capital or consumers.

The final mechanism by which computers are claimed to increase the demand for skill is by *altering the relative numbers of more- and less-skilled jobs* irrespective of any impact on the skill content of tasks within jobs. Computer diffusion may increase the share of more skilled jobs that are directly computer-related (e.g., programmers, systems analysts, software writers) and noncomputer occupations that involve analysis and decision making based on the additional information computers generate (e.g., accountants, production planners). Computers may also substitute for or eliminate through automation various less skilled, blue-collar and clerical jobs (Berman et al., 1994; Levy & Murnane, 1996; Autor et al., 2002).

As with the other proposed mechanisms of computer-driven SBTC, various empirical studies question whether computers and related microelectronics technology explain observed shifts in the occupational composition of employment (Doms, Dunne, & Troske, 1997; Howell & Wieler, 1998; Handel, 2000). One problem is the issue of two-way causation. While an exogenous increase in computer use may raise the demand for managers, professionals, and other nonproduction workers, it is also clear that an exogenous increase in the demand for these kinds of workers will raise the demand for computers, which have become standard office equipment in the last 25 years. It is hard to see that many increases in white-collar office workers would fail to be accompanied by growth in the stock of computers, not to mention desks and office chairs. Because there has been a secular increase in employment shares for managerial, professional, and other nonproduction workers that began long before the advent of computers (Melman, 1951; Kaplan & Casey, 1958), any estimate of the effects of computers on the growth of nonproduction occupations would need to be purged of the association resulting from causal forces working in the opposite direction.

The analyses that follow test the first three explanations of the effects of computer use on wages (computer-specific human capital, general human capital, contextual effects) and also go beyond the framework of these approaches by considering the effect of computers on the overall distribution of wages. Detailed consideration of the role of computers in altering the distribution of workers across jobs is left for a future paper.

3. DATA

This paper uses the January 1991 supplement to the Current Population Survey (CPS). These data are contemporaneous with those used in the original study that initiated this debate (Krueger, 1993) and with the most rapid period of inequality growth. Unlike the October CPS supplements used in previous studies, this data include seven indicators of noncomputer tasks workers perform on the job in addition to computer use. The variables measure how often workers read or use different kinds of materials on the job (e.g., news articles, forms, letters, diagrams, manuals), write text to be read by others, use math or arithmetic, and use a computer or terminal at work. Unlike the October series, this survey also asked how often workers performed each task (never, less than once per week, one or more times per week, every day), not simply whether or not they performed them. Unlike the October supplements, the January 1991 survey collected information only from the target respondent, rather than permitting proxy responses from other household members, which increases the reliability of the information collected.

The reading, writing, math, and computer items can be interpreted within a human capital framework as measuring workers' cognitive skills. Alternatively, these eight variables may be seen as proxies for occupational position, as DiNardo and Pischke argue, though none are so plainly lacking in overt skill content as their pencil use or "sit while working" variables as to rule out a human capital interpretation as well.

The measures are used to test the computer-specific human capital SBTC hypothesis by replicating DiNardo and Pischke's (1997) comparison of the returns to computer and noncomputer job characteristics and by using them as controls for usually unobserved job characteristics. The measures are also used to test the general human capital SBTC hypothesis by examining whether the noncomputer cognitive skills mediate the relationship between computer use and wages.

The sample is restricted to wage and salary workers, age 18–65, who report earning between $1.50 and $250 per hour in current dollars (Krueger, 1993). Workers paid by the hour are assigned their reported hourly wage. The hourly wage for salaried employees is calculated by dividing reported weekly earnings by reported usual hours worked. The dependent variable in all wage regressions is the log hourly wage.

Models also use three measures of pre-computer job characteristics from the late 1960s and early 1970s to control for characteristics that might otherwise be confounded with computer use. Mean log earnings in 1969 for three-digit occupations and three-digit industries are calculated from a special 1970 Decennial Census sample ($n = 109{,}605$). These two variables capture human capital and institutional characteristics of jobs, such as qualifications, positional rank, and occupation and industry rents (Krueger & Summers, 1987), that pre-date the rapid growth in information technology that began in the early 1980s. The 1970 Census file includes both the 1970 and 1980 Census occupation and industry-coding schemes, which permits mean earnings for 1969 to be merged easily onto the January 1991 CPS data, which used the 1980 coding schemes.[4]

A third control for pre-computer differences is a scale constructed from six measures of cognitive skill requirements drawn from the *Dictionary of Occupational Titles* (DOT) (1977 edition). The DOT contains ratings of job skill requirements conducted by job analysts at the U.S. Department of Labor's Employment and Training Administration. The scale used here is a standardized sum of six standardized variables that measure cognitive skill requirements at the three-digit occupational level. They are measures of general human capital ("General Educational Development"), occupation-specific training time ("Specific Vocational Preparation"), the job's complexity of involvement with data, and the typical job incumbent's rank in the national distribution of three aptitudes (verbal, numerical, general intelligence). This group of variables is very similar to those loading on the substantive complexity factor extracted by Miller, Treiman, Cain, and Roos (1980, pp. 177ff.) in their thorough study of the DOT. Not surprisingly, there is considerable overlap among the six variables, which makes them well-suited to be combined in a single measure. The first principal component accounts for nearly 88% of the variance and the reliability coefficient of the additive scale used in analyses below, calculated after merging the scores onto the 1991 CPS data, is also very high (Cronbach's $\alpha = 0.97$).

These three variables – occupational and industry mean earnings in 1969 and the DOT scale of substantive complexity – measure job rewards and cognitive skill demands prior to both the broad diffusion of computers and the large rise in inequality. Indeed, roughly 80% of the 1977 DOT job descriptions were simply carried over unchanged from the previous edition of the DOT published in 1965, which underscores the degree to which they measure longstanding occupational differences (Miller et al., 1980, pp. 159ff.). These variables are added to regression models to control for usually unobserved, pre-existing human capital requirements and institutional labor market characteristics of jobs that may be associated with the subsequent introduction of computers and whose omission may bias coefficient estimates of the effect of computer use on changes in the wage structure.

The correlations in Table 1 show a strong continuity in occupational characteristics across time, illustrating the potential power of these baseline measures. The correlation between mean occupational pay levels in 1969 and 1991 is 0.87 (r_{32}), as is the correlation between industry pay across these years (r_{54}). The correlation between occupational complexity as measured by the DOT for the 1960s–1970s and as measured in 1991 is also very high ($r_{87} = 0.85$). (The 1991 measure is the occupational mean of the additive scale composed of the seven noncomputer job characteristics in the January 1991 CPS.)[5]

Table 1 also suggests that the introduction of computers was not random with respect to longstanding differences in job characteristics and rewards. The correlation between computer use in 1991 and mean occupational earnings in 1969 equals the correlation of computer use with individuals' wages in 1991 ($r_{91} = r_{93} = 0.33$), and is similar to the correlation between computer use and mean occupational wages in 1991 ($r_{92} = 0.38$). The correlations between computer use in 1991 and mean industry wages in 1969 ($r_{95} = 0.25$) and 1991 ($r_{94} = 0.30$) are also quite similar. Computer use in 1991 is associated with wages in 1991, but it is almost as strongly associated with wages in 1969, when it could not have a causal influence on wages. This suggests the possibility that the contemporaneous association of computers and wages reflects large, pre-computer baseline wage differentials more than computer-induced pay gaps, a proposition tested more formally below.

Likewise, the correlation between computer use in 1991 and the DOT occupational complexity measure (r_{98}), which was based heavily on 1960s job descriptions, is 0.48, which is higher than the association between computers and current wages ($r_{91} = 0.33$) and similar in magnitude to the correlation between computer use and the seven-task scale derived from the January 1991 CPS, whether measured at the individual level ($r_{96} = 0.49$) or the occupational level ($r_{97} = 0.53$).

Clearly, there is a high degree of persistence in both job characteristics and rewards by occupation and industry. Further, computer use is about as closely correlated with pre-computer differences between jobs, which cannot reflect the effects of computerization, as it is with current differences, which

Table 1. Correlations Between Contemporary and Pre-Computer Job Characteristics.

		1	2	3	4	5	6	7	8
1	Individual wage – 1991								
2	Mean occupation wage – 1991	0.6472							
3	Mean occupation wage – 1969	0.5622	0.8686						
4	Mean industry wage – 1991	0.5097	0.5376	0.4635					
5	Mean industry wage – 1969	0.4424	0.4945	0.5173	0.8679				
6	Individual-level seven-task scale – 1991	0.4190	0.4845	0.4555	0.3220	0.2803			
7	Mean occupation seven-task scale – 1991	0.4982	0.7698	0.7241	0.4009	0.3504	0.6294		
8	Occupational complexity – 1977 (DOT)	0.4673	0.7241	0.6735	0.2961	0.2215	0.5313	0.8461	
9	Computer use – 1991	0.3341	0.3768	0.3319	0.3014	0.2531	0.4907	0.5264	0.4798

Note: All wages are in log form. Values for mean seven-task scale are occupation-level means. Computer use is measured dichotomously at the individual level.

are presumed to result from computerization. This supports the intuition that one needs to control for pre-computer differences to avoid possible coefficient bias in estimates of computers' incremental effect.

However, it should be noted that the variables available as controls are measured with error because it is only possible to capture these qualities at the occupational and industrial level rather than at the level of the individual job. The correlation between the 1970 Census respondents' own earnings, on the one hand, and the mean earnings in their occupation (0.61) and industry (0.47), on the other hand, are not perfect. Therefore, when the aggregate-level measures are added to models using the 1991 CPS, they will capture a great deal of baseline of pre-computer job-level variation, but some portion of this variation will also likely remain unmeasured and potentially correlated with other regressors.

Individual fixed effects models might address this problem, but also have some well-known disadvantages. There may be few cases that change user status in a short panel and the user status will be measured with some error in both periods, particularly in the case of proxy reporting. Even a modest amount of misclassification can increase the proportion of error variance to the point that the downward coefficient bias produces more misleading estimates than standard cross-sectional estimates (Freeman, 1984). Fixed effects models would be desirable because they can often be taken as lower-bound estimates of the true effects (Freeman, 1984). Panel data with computer use are unavailable for the U.S. However, since cross-sectional data are usually taken as upper-bound estimates, the approach used here can be considered favorable to the SBTC hypothesis.

To control for heterogeneity in firm characteristics that might be associated with computer use, analyses below also include the size of the 1991 CPS respondents' employer, which has usually been absent from studies using CPS data. This is accomplished by using the short panel nature of the CPS to match approximately one-quarter of cases found in both the January 1991 and March 1991 files, the latter containing both wage and firm size data. Research using the October CPS series cannot link information on computer use and firm size because there is no overlap between the CPS rotation groups covered by the October and March supplements. The correlation between computer use and firm size in 1991 is 0.20.

The CPS Outgoing Rotation Group's (ORG) annual merged files are also used to estimate a time series for the variance of log wages (1979–2001) to compare trends in wage inequality and computer diffusion.[6] The sample in all years is wage and salary workers age 18–65 earning between $1.50

and $250 per hour in constant 1984 dollars. Sample sizes are roughly 150,000–180,000 for each year.

The October 1984 and 1989 CPS supplements are also used to assess the contribution of computer use to overall inequality by adjusting rates of computer use in 1989 to levels observed in 1984 using the method described in DiNardo et al. (1996) and calculating a counterfactual distribution of wages. Sample restrictions follow Autor et al. (1998). Sample sizes are about 13,700 for each year.

4. WAGE AND SKILL IMPLICATIONS OF COMPUTERS

4.1. Computer-Specific and General Human Capital among Users

This section tests the claims that the additional computer-specific and general human capital required by computer use are associated with large labor market returns.

Following common practice (Krueger, 1993; DiNardo & Pischke, 1997), Table 2 enters each of the eight job task items from the January 1991 CPS individually in a standard wage equation of the form,

$$\ln \quad W_i = \mathbf{X}_i\beta + \mathbf{Z}_i\alpha + \varepsilon_i$$

where W_i = hourly wage for individual i; $\mathbf{X}_i$ = vector of control variables;[7] $\mathbf{Z}_i$ = dummies for frequency with which individual i performs particular job task; and ε_i = error term.

Table 2 shows that when entered individually, each of the eight job tasks is associated with very large wage differentials of roughly comparable magnitude. The coefficients for computer use tend to be in the upper end of the range of estimates, but are not exceptional. Nor are the estimates for computer use peculiar to this sample or low relative to other estimates.[8] The results indicate that those who perform *any* of the eight tasks every day earn roughly 21% more per hour than those who never perform them, while the corresponding figures for those who perform any of the tasks once or more per week and less than once a week are about 17% and 14%, respectively. The computer measures and about half the noncomputer variables also reduce the size of the education coefficient by comparable amounts.

These results confirm that DiNardo and Pischke's (1997) results are not restricted to Germany and generalize to the United States. Following their

Table 2. OLS Regression Estimates of the Effects of Eight Job Tasks on Ln (Wage), Job Tasks Entered Individually.

Model	No Task Variables	Use PC or Terminal	Read or Use:					Write Memos or Reports	Use Math or Arithmetic
			Letters	Instruction Manuals	Articles or Reports	Forms	Diagrams, Blueprints		
Education (years)	0.0886 (0.0015)	0.0770 (0.0017)	0.0735 (0.0017)	0.0821 (0.0017)	0.0784 (0.0017)	0.0814 (0.0017)	0.0845 (0.0016)	0.0760 (0.0017)	0.0848 (0.0017)
Task frequency									
< once per week		0.1528 (0.0194)	0.1375 (0.0135)	0.1377 (0.0116)	0.1283 (0.0135)	0.1176 (0.0170)	0.1598 (0.0132)	0.1698 (0.0139)	0.1647 (0.0180)
≥ once per week		0.1467 (0.0160)	0.1886 (0.0122)	0.1825 (0.0117)	0.1641 (0.0125)	0.1445 (0.0141)	0.1982 (0.0141)	0.1972 (0.0123)	0.1389 (0.0152)
Every day		0.2343 (0.0090)	0.2624 (0.0101)	0.2007 (0.0105)	0.1826 (0.0103)	0.1945 (0.0102)	0.1877 (0.0107)	0.2420 (0.0099)	0.1487 (0.0106)
R^2	0.417	0.454	0.454	0.441	0.439	0.439	0.443	0.450	0.431
N	14438	11465	11404	11403	11466	11460	11368	11456	11466

Note: Standard errors in parentheses. All reported coefficients significant at 0.05 level. All models include variables for experience, experience^2, part-time status, union status, female, black, other non-whites, resident of metropolitan area, married, married*female, veteran, and three region dummies. The omitted category for job-task variables is "never use."

argument, the magnitude and similarity of the estimated returns to these eight tasks suggest that in addition to any true returns, they may be picking up common, unmeasured variation in human capital, occupational position, or firm characteristics.

Supporting this interpretation is the fact that the first principal component accounts for a sizable 50% of the total variance in a principal components analysis that includes all variables except the more specialized "diagrams, plans, and blueprints" and no other component has an eigenvalue greater than 1. The reliability coefficient for a simple additive scale composed of these items is also substantial (Cronbach's $\alpha = 0.83$). This strongly suggests that all job task variables, including computer use, are not simply measuring returns to separate and distinct skills, but are proxying for some common, unobserved variable.

Some of the specific coefficient values also support this interpretation. For example, the results imply that those who read or use letters every day earn about 30% more than those who never do so.[9] More than likely this reflects other factors, such as an individuals' occupational status or general abilities, at least as much as a return to a specific ability to read or use letters.

Similar sorts of results from Krueger's (1993) original study were also awkward from a human capital perspective. The computer premium varied by the specific type of computer task, but the pattern was not easily interpretable. Using e-mail at work received the highest premium (0.149) above the basic return to any form of computer use, while spreadsheet use was rewarded only half as much (0.079), and programming and computer-aided design software use brought no additional reward beyond the basic computer premium (Krueger, 1993, p. 41f.). These relative magnitudes do not reflect the likely actual differences in the costs of acquiring such skills as implied by human capital theory and cast doubt on the reliability of the estimates.

Table 3 introduces additional controls to further test for upward bias in the computer coefficient. The first column replicates the model for computer use in Table 2, but collapses the four categories of computer usage frequency into a binary variable that contrasts all categories of users with nonusers to facilitate comparison with previous studies. The computer coefficient (0.20) remains larger than previous estimates (Krueger, 1993; Autor et al., 1997).[10]

Model 2 adds three controls for pre-computer differences among jobs that may be correlated with both computer use and current wages: mean earnings by occupation and by industry in 1969, and the DOT substantive complexity scale from about the same period.[11] These variables control for job characteristics that are well prior to any significant influence of

Table 3. Effects of Computer Use on Ln (Hourly Wage) Controlling for Other Job Characteristics.

	1	2	3	4	5	6
Education (years)	0.0682** (0.0018)	0.0470** (0.0043)	0.0403** (0.0045)	0.0380** (0.0043)	0.0376** (0.0020)	0.0382** (0.0020)
Use PC or terminal	0.2003** (0.0095)	0.0844** (0.0127)	0.0473** (0.0117)	0.0362** (0.0117)	0.0199 (0.0106)	0.0468** (0.0098)
1969 Occupational earnings (ln)		0.1844** (0.0288)	0.1591** (0.0283)	0.1519** (0.0296)	0.1478** (0.0127)	0.1470** (0.0129)
1969 Industry earnings (ln)		0.2180** (0.0329)	0.2174** (0.0319)	0.2060** (0.0305)	0.2092** (0.0136)	0.2136** (0.0138)
Early 1970s occupational complexity (DOT)		0.0623** (0.0163)	0.0451* (0.0185)	0.0395* (0.0191)	0.0396** (0.0077)	0.0370** (0.0081)
Other covariates (1991)			Yes	Yes	Yes	Yes
Seven job tasks (dummies) (1991)				Yes		
Seven job tasks (scale) (1991)					0.0485** (0.0084)	
Seven job tasks (scale) (1991): users assigned nonuser means						0.0371** (0.0093)
N	9309	9120	8964	8684	8684	8458
Adj. R^2	0.4508	0.5294	0.5622	0.5680		
R^2					0.5661	0.5603

Note: Standard errors in parentheses. All models include controls listed in note for Table 1. "Other covariates" are firm size, firm and occupation tenure, hourly worker status, government employee, management/supervisory training. Models 5 and 6 include the seven job task variables for 1991 in the form of an additive scale with a reliability correction ($\alpha = 0.814$). These models also correct the DOT scale for reliability ($\alpha = 0.97$).
*Significant at 5%.
**Significant at 1%.

computers. This avoids concerns over potential endogeneity that might arise from using individuals' own occupation and industry as regressors (Krueger, 1993), which are not included in any models. After including controls for pre-computer job characteristics, the estimated premium for computer use drops from 0.20 to 0.08. Roughly 60% of the measured returns to computer use reflect the fact that computers are found disproportionately in occupations and industries that were already highly skilled and well paid in the pre-computer era (ca. 1970).

This effect is not restricted to this particular year or data set. The corresponding estimates for the September 2001 CPS supplement are virtually identical. The average computer premium across all October/September supplements for 1984–2001 drops from 0.199 to 0.087 when pre-computer job characteristics are controlled (results not shown).

To control for additional job-level characteristics not captured by the three aggregate-level variables, Model 3 adds an extended set of controls, most of which are only available using the January 1991 and linked March 1991 data – firm size, hourly worker status, government worker, firm tenure, tenure in current occupation, and whether the respondent received managerial or supervisory training. When these variables are added the computer premium drops to 4.8%. Model 4 adds the seven noncomputer job task variables from the January 1991 CPS supplement as dummy variables and the premium drops to 3.7%. Finally, combining the seven noncomputer job characteristics into a standardized scale and applying a reliability correction ($\alpha = 0.814$) reduces the effects of computer use to insignificance (Model 5). There do not appear to be any rewards for computer use after controlling for pre-computer and contemporaneous noncomputer job characteristics.

Since both the computer-specific and general human capital theories imply computer users receive a wage premium relative to nonusers, the absence of any significant premium in Model 5 would seem to suggest that neither theory is supported once pre-computer and relevant contemporaneous job characteristics are controlled. However, the seven noncomputer job tasks in Model 4 and Model 5 have a somewhat ambiguous status. It is reasonable to include them in the models insofar as they represent job-level characteristics unrelated to computer use that are not captured by the three coarser, aggregate-level measures. However, these measures of reading, writing, and math frequency may be also partially endogenous to computer use as implied by the general human capital account of computer-driven SBTC, in which case they absorb some of the total effect of computer use.

One way to correct for this possibility and test whether general human capital mediates the effect of computer use on wages is to replace scale values for computer users with values for comparable nonusers. For example, computer users can be assigned scale values equaling the means of nonusers in their three-digit occupation, on the grounds that this represents a reasonable estimate of what their job requirements would be if they did not use computers. Insofar as users must perform the seven noncomputer tasks more frequently than nonusers and are rewarded for it, the coefficient for computer use would reflect this fact and the seven-task scale would no longer over-control for job characteristics that may reflect the effects of computer use. However, it should also be noted that this strategy may reintroduce some problems of unobserved heterogeneity because it is possible that some part of the difference in scale values between users and nonusers reflects job-level variation within occupations that antedates or is otherwise not causally related to computer use. With these considerations in mind, the results for Model 6 indicate that when scale values for computer users are replaced with occupational means for nonusers, the effect of computer use returns to a level comparable to Model 3 (4.7%), which did not include any controls for the seven noncomputer job tasks. Taken together, the results in Table 3 suggest that there are no returns to computer-specific human capital and relatively modest returns to the general human capital that may be associated with computer use.

These conclusions are reinforced by results from further tests that add two other computer-related variables to the last three models in Table 3: a dummy for whether an individual reported their computer skills were good enough for their current job, and a dummy for whether an individual received computer training after obtaining their current job. These variables measure computer-specific human capital. If there is a genuine return to computer skills per se, one would expect those who report their computer skills are inadequate would suffer a wage penalty and those who received computer training would reap positive returns, all else equal.

In addition, these models include variables for whether individuals reported their reading, writing, and math skills were good enough for their current job and whether they received training in four noncomputer areas after obtaining their job: management/supervisory skills, (noncomputer) occupation-specific technical skills, reading/writing/math skills, and other skills. These controls may also affect the estimated effects of computer use resulting from increased general human capital requirements (Table 3, Model 6).

Model 1 in Table 4 uses the additional variables to augment Model 4 in Table 3. The premium for computer use is less than 3% and those reporting their computer skills are inadequate for their job receive essentially no premium. Model 2 incorporates the seven noncomputer job tasks into a scale and applies a reliability correction, like Model 5 in Table 3. The coefficients for both computer use and inadequate computer skills are insignificant in this model. When computer users are assigned nonuser means for the seven-task scale, the computer premium is 3.8% and again significant, but smaller than the analogous results from Model 6 in Table 3 and the coefficient for inadequate computer skills remains insignificant. These results are unchanged when the adequacy of computer skills is interacted with dummies for frequency of computer use (results not shown).

Table 4. Effects of Computer Skills and Training and Other Skills and Training on Ln (Hourly Wage).

	1	2	3
Education (years)	0.0400**	0.0395**	0.0402**
	(0.0047)	(0.0022)	(0.0023)
Use PC or terminal	0.0279*	0.0117	0.0375**
	(0.0139)	(0.0127)	(0.0117)
1969 Occupational earnings (ln)	0.1604**	0.1566**	0.1567**
	(0.0302)	(0.0142)	(0.0145)
1969 Industry earnings (ln)	0.2050**	0.2092**	0.2123**
	(0.0284)	(0.0154)	(0.0156)
Early 1970s occupational complexity (DOT)	0.0374	0.0391**	0.0345**
	(0.0202)	(0.0086)	(0.0094)
Seven job tasks (dummies) (1991)	Yes		
Seven job tasks (scale) (1991)		0.0401**	
		(0.0096)	
Seven job tasks (scale) (1991): users assigned nonuser means			0.0292**
			(0.0112)
Skills good enough for job (1 = no)			
Computer skills	−0.0228*	−0.0224	−0.0186
	(0.0109)	(0.0114)	(0.0116)
Reading skills	−0.0310	−0.0244	−0.0262
	(0.0380)	(0.0438)	(0.0456)
Writing skills	0.0280	0.0243	0.0299
	(0.0275)	(0.0300)	(0.0316)
Math skills	0.0469	0.0509	0.0518
	(0.0322)	(0.0320)	(0.0334)
Training since hired: (1 = yes)			
Computer-related skills	−0.0083	−0.0097	−0.0153
	(0.0148)	(0.0127)	(0.0132)
Other technical skills specific to occupation	0.0465**	0.0469**	0.0506**
	(0.0149)	(0.0104)	(0.0105)
Managerial or supervisory skills	0.0558**	0.0585**	0.0656**
	(0.0154)	(0.0141)	(0.0143)
Reading, writing, or math skills	0.0005	−0.0047	0.0068
	(0.0169)	(0.0176)	(0.0180)
Other skills	−0.0281	−0.0291	−0.0287
	(0.0226)	(0.0165)	(0.0167)
N	7034	7034	6808
Adj R^2	0.5630		
R^2		0.5617	0.5564

Note: Robust standard errors in parentheses. All models include other controls present in Model 4 in Table 2. Models 2 and 3 include the seven job task variables for 1991 in the form of an additive scale with a reliability correction ($\alpha = 0.814$). These models also correct the DOT scale for reliability ($\alpha = 0.97$).

*Significant at 5%.

**Significant at 1%.

Those who say their reading, writing, or math skills are inadequate seem to bear no wage penalty, but this may reflect the lack of variation in these three variables, which is not an issue in the case of computer skills. While less than 3% reported that their reading, writing or math skills were not good enough for their current job, 21.4% said that their computer skills were inadequate. If computer skills are important one would expect that those with inadequate skills earn less than other users. This does not seem to be the case, though the possibility of measurement error in this kind of self-report item argues for caution. Nevertheless, it is notable that many at all levels of computer use acknowledge computer-skill deficits, yet incur no specific penalty.[12]

If computer skills were important one would expect that those who received computer training after being hired would earn more than others, all else held equal, but this does not appear to be the case in any model in Table 4. Clearly, there is a potential selection issue here that argues for caution. Those who received post-hire training may have had a computer skill deficit prior to training that is unobserved in the data. If training simply brought them to parity with those who already have the necessary skills, then the absence of measured returns in the cross-section may mask a real treatment effect for those receiving training. In the absence of panel data there is no way to test this possibility, but there are positive returns observed for other kinds of technical training and managerial/supervisory training. The absence of a positive effect for computer training, then, may simply reflect the short training times for the most commonly used computer skills, such as word processing, that were reported even early in the computer diffusion process, when few users had prior experience with computers at school or work and software did not reflect later improvements in usability (Goldstein & Fraser, 1985; Levin & Rumberger, 1986; Carroll, 1997). If the absence of measured returns to computer training reflects its brevity, then computer skills are unlikely to be so scarce and expensive as to garner large returns or account for a large part of the growth of inequality.

The preceding suggests that the concerns raised by DiNardo and Pischke (1997) generalize to the United States. Given the similarly high returns for the eight job task variables in the January 1991 CPS when they are entered individually into a standard wage equation, it seems implausible that the results reflect returns to separate and distinct skills. It is more likely that the greater part of their effects on wages reflect some common, unmeasured variation in human capital, occupational position, or firm characteristics. When measures of pre-computer and other job characteristics are added as controls, there seems to be no wage premium associated with computer skills narrowly conceived, that is, understanding how to operate the software and equipment. The absence of measured returns to computer training and the lack of penalties for self-reported computer skill deficits reinforce this conclusion. Taken together, these results suggest that computer skills *per se* are not as important in wage determination as previously argued.

Models also tested for whether computer use is associated with a wage premium mediated by general human capital by allowing the indicator for computer use to pick up the effects of reading, writing, and math tasks insofar as their frequency differed from those of the average nonuser within a computer user's occupation. Results suggest that the increased general intellectual requirements that may accompany computer use due to more knowledge-intensive tasks and restructured work roles is associated with only a modest wage premium on the order of 3–4% net of other job characteristics.

4.2. Contextual Effects of Computers on Users and Non-Users

The preceding still leaves open the possibility of contextual effects. Assigning nonuser skill means to computer users allows computer use to pick up the effects of computerization on skill requirements only under the assumption that computerization did not raise skill requirements for nonusers. If the presence of computers raised skill requirements of nonusers as well as users, this approach will underestimate the total effect of computers on within-job skill requirements.

Autor et al. (1998, p. 1190) try to test a contextual effects argument by estimating the effects of changing computer use within industries on changes in educational composition within industries, in a rough effort to convert worker-level information on computer use into an organization-level, contextual variable. However, if the estimated effects of computer use on skills or wages are upwardly biased when the data is used in its original form for individual-level models, transforming the same data into industry-level percentages may not solve the problem.

Table 5 uses CPS supplements for October 1984, 1993, and 1997 and September 2001 to update and elaborate on Autor et al.'s (1998) industry-level regressions. Average annual changes in the percentages of workers with different levels of education (high school, some college, at least four years of college) are regressed on average annual changes in the percentages of workers using computers within an industry for different time periods.

Table 5. Effects of Changes in Percentage of Computer Users on Employment Share of Educational Groups within Industries.

Employment Share	1984–2001	1984–1989	1989–1993	1993–1997	1997–2001	1971–1976
High school						
Δ Computer use	−0.088	−0.125**	−0.232**	−0.053	−0.010	−0.507**
	(0.065)	(0.033)	(0.052)	(0.044)	(0.038)	(0.141)
Constant	−0.343**	0.037	−0.396*	−0.446**	−0.280**	0.946**
	(0.113)	(0.112)	(0.155)	(0.077)	(0.066)	(0.236)
N	223	222	221	221	222	219
R^2	0.027	0.074	0.089	0.007	0.000	0.078
Constant-only	−0.488**	−0.267**	−0.932**	−0.494**	−0.287**	0.066
Total Δ (in %)	−8.31	−1.34	−3.73	−1.98	−1.15	0.33
Some college						
Δ Computer use	0.013	0.063**	0.038	0.007	0.005	−0.062
	(0.036)	(0.023)	(0.040)	(0.034)	(0.034)	(0.093)
Constant	0.487**	0.100	1.452**	0.168**	0.130*	0.710**
	(0.065)	(0.069)	(0.111)	(0.057)	(0.062)	(0.161)
N	223	222	221	221	222	219
R^2	0.001	0.033	0.004	0.000	0.000	0.003
Constant-only	0.509**	0.255**	1.541**	0.175**	0.134*	0.602**
Total Δ (in %)	8.66	1.27	6.16	0.70	0.54	3.01
College+						
Δ Computer use	0.053	0.097**	0.113*	0.096*	0.033	0.244*
	(0.060)	(0.027)	(0.045)	(0.043)	(0.041)	(0.100)
Constant	0.214*	0.082	0.015	0.250**	0.221**	0.292
	(0.103)	(0.062)	(0.120)	(0.073)	(0.057)	(0.149)
N	223	222	221	221	222	219
R^2	0.013	0.066	0.041	0.033	0.005	0.042
Constant-only	0.301**	0.318**	0.276**	0.337**	0.246**	0.715**
Total Δ (in %)	5.11	1.59	1.10	1.38	0.99	3.57

Note: Independent and dependent variables measured as annual average change in percentages. Figures in constant-only rows are from baseline models in which average annual changes in education shares are regressed on a constant only. College + includes all those with 16 or more years of education (1971–1989) or at least a Bachelors degree (1993–2001). For the final column, the independent variable is the average annual change in computer users within industries between 1984 and 2001. All models are weighted by the average percentage of workers in industries across the two years compared. Robust standard errors in parentheses.

*Significant at 5%.

**Significant at 1%.

(Similar regressions using the percentage of workers with less than high school as the dependent variable are omitted for space reasons and because they showed consistently no impact of changes in computer use, similar to Autor et al.'s (1998, p. 1190) original results.)

The results in Table 5 present a rather mixed picture. The total change and average annual rate of change in the dependent variables are reported below the R^2 row for each set of results. The share of high school graduates across all industries declined by over eight percentage points between 1984 and 2001, implying an annual rate of change of about -0.488 percentage points. The full-period increases for workers with some college and four years of college or more are 8.7 percentage points and 5.1 percentage points, respectively.

The regression results indicate that changes in computer use are not significantly related to any of these changes for the full period of 1984–2001. Analyses by sub-period indicate that the spread of computer use is significantly associated with changes in industry educational composition in the expected direction between 1984 and 1993, the period covered by Autor et al. (1998), but for the most part not significantly associated with changes in educational composition between 1993 and 2001. Where computer effects are significant, they suggest mostly decline in high school educated workers and increase in workers with four years of college or more, rather than changes in the percentage of workers with some college.

The final column provides a robustness check on this exercise by regressing average annual changes in industry-level education shares between 1971 and 1976 on the average annual change in computer use between 1984 and 2001. These results are some of the strongest of all. Future changes in computer use are associated with much greater changes in the use of high school and college educated workers in the early 1970s than at any time in the period 1984–2001. Since the changes in educational composition between 1971 and 1976 could not actually result from the surge in computer use that occurred in the following decades, the results in the final column suggest strongly that computer diffusion may proxy for other changes within industries that upgrade educational levels. In particular, results in Table 5 that seem to suggest a causal role for computer in educational upgrading within industries may be driven by exogenous changes in occupational composition that increase both education levels and the demand for computers.

The contextual effects explanation also contains certain implications worthy of further investigation. Specifically, this account implies that the presence of computers increases the skill required of nonusers and increases the skills required of users over and above any effects of computer use itself, e.g., due to the effects of computers on the quantity and quality of information in workplace (Bresnahan, 1999).

One way to test whether the spread of computers increased the skill demands on nonusers is to re-run the regressions in Table 5 on nonusers only.

In results not shown here, such models do not show any effects of computer use on the education levels of nonusers. However, a problem with these kinds of panel models is that the spread of computer use itself is likely to leave an increasingly less-educated group within the nonuser sample across successive years, which would make any educational upgrading effect among nonusers difficult to detect. Cross-sectional, individual-level models avoid this problem and can also be used to test whether working in a computerized environment increases the skills of users beyond the effects of working with a computer oneself.

Table 6 presents results of a series of descriptive regressions of worker's own education on computer and other skill-related variables that give another perspective on these issues. As a baseline, Column 1 shows that computer use was associated with nearly two years of education in 1991, controlling for background variables.[13] This effect drops to a half-year differential after controls are added for 1970 levels of education within the respondent's occupation and industry and the DOT occupational complexity scale (Column 2). Omitting the pre-computer controls, Column 3 suggests that the percentage of computer users in an industry has a significant association with workers' education independent of whether respondents use a computer themselves. However, this contextual effect becomes negative once the educational level of industry in 1970 is re-entered into the model (Column 4), though insignificant once occupational characteristics in 1970 are added (Column 5).

In the middle panel the analyses are repeated for users and nonusers separately, with similar results. In particular, it is significant that the intensity of computer use in an industry seems to have no association with nonuser education levels once 1970 industry education levels are controlled. Again, when a similar model is applied to the 1970 Census file in the bottom panel, the level of computer use in the industry in 1991 is strongly associated with an individual's education, even after controlling for occupational complexity as measured by the DOT.

To test for the possibility that Bresnhan's (1999) concept of organizational computing is more relevant to the issue of SBTC than personal computing, all models in Table 6 were reestimated using as the contextual predictor the percentage of computer systems analysts, computer scientists, and programmers within industries, rather than the percentage of computer users. This alternative variable is not significantly associated with worker education levels in any model for the 1991 sample.

In sum, the evidence for contextual effects is rather weak. Insofar as increased computer use is associated with educational upgrading within

Table 6. Regression of Years of Education on Computer Use and Percentage of Computer Users in Industry.

	1	2	3	4	5	6
	All Workers					
Computer use	1.927**	0.499**	1.471**	1.461**	0.534**	
	(0.023)	(0.022)	(0.078)	(0.120)	(0.058)	
Users in industry (%)			0.022**	−0.010*	−0.003	
			(0.003)	(0.004)	(0.002)	
Industry education (1970)		0.175**		0.749**	0.209**	
		(0.008)		(0.104)	(0.032)	
Occupational education (1970)		0.678**			0.670**	
		(0.010)			(0.033)	
Occupational complexity (DOT)		0.034			0.047	
		(0.019)			(0.051)	
N	46,058	45,117	46,058	46,033	45,117	
R^2	0.166	0.420	0.195	0.271	0.420	
	Users – 1991			**Nonusers – 1991**		
Users in industry (%)	0.016**	−0.015**	−0.006**	0.029**	−0.000	0.002
	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)
Industry education (1970)		0.819**	0.253**		0.617**	0.151**
		(0.051)	(0.047)		(0.052)	(0.033)
Occupational education (1970)			0.657**			0.640**
			(0.051)			(0.029)
Occupational complexity (DOT)			0.047			0.092
			(0.082)			(0.060)
N	23,899	23,894	23,363	22,159	22,139	21,754
R^2	0.066	0.197	0.384	0.086	0.134	0.286
	Census Sample – 1970					
Users in industry (%)	0.041**	0.019**				
	(0.004)	(0.003)				
Occupational complexity (DOT)		1.466**				
		(0.135)				
N	120,480	116,169				
R^2	0.194	0.377				

Note: Dependent variable is own years of education. Percentage of computers users in industry refers to 1991. All models include controls for age, gender, race, marital status and its interaction with gender, part-time status, and region (except models for 1970 sample). Standard errors in parentheses.

*Significant at 5%.

**Significant at 1%.

industries, the effects are mostly restricted to 1984–1993 and some of the strongest associations are with educational upgrading between 1971 and 1976, years that pre-date the growth in computer use, educational differentials, and overall inequality. Cross-sectional models suggest that computers do not have a spillover effect on the skill requirements of nonusers nor is there strong evidence that working in a computerized environment increases the skill requirements of users beyond any effects of working with a computer personally.

5. COMPUTERS AND THE GROWTH OF INEQUALITY

The preceding analysis examined the effects of computers on wages and skills for computer users and nonusers. However, this kind of analysis does not address directly the magnitude of computers' contribution to the growth of overall wage inequality.

Relatively early in the debate over SBTC but unnoticed until more recently and not much emphasized, Howell (1995) observed that the timing of inequality growth did not seem to track trends in computer investment in an obvious manner, with inequality growing rapidly between 1980 and 1988 and remaining flat between 1988 and 1992. Subsequent research confirmed this observation (Katz & Autor, 1999, pp. 1484ff.).

More recent data further reinforces this concern. Fig. 1 plots the variance in log wages for 1979–2001 using the CPS ORG files, and the percentage of workers who use computers at work using the October CPS supplements.[14] Inequality growth was most rapid in the early 1980s. The variance of log wages rose about 23% between 1979 and 2001, but slightly more than half the total growth occurred very early in the period during two years of deep recession (1981–1983), when the jobless rate reached a post-war high of 9.7 percent (Economic Report of the President: 2003, Table B-42). Inequality growth moderated steadily as the decade progressed and there was essentially no change in the overall inequality between 1988 and 2001.[15] Yet computer use and its estimated effects on wages and industry skill composition continued to rise (Autor et al., 1998, pp. 1187, 1194).

The pattern in Fig. 1 is not one that would be expected if the growth in computer use were driving the growth in inequality. When computer use was

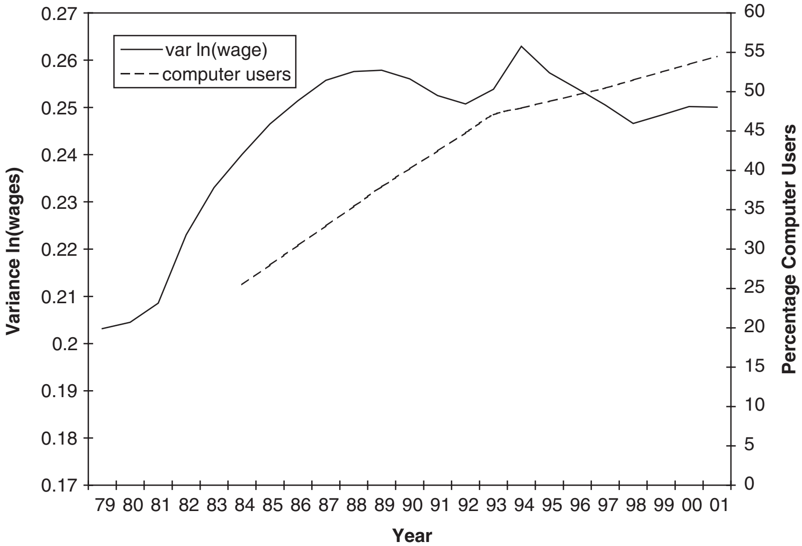


Fig. 1. Trends in the Variance of ln(wages) and Percentage of Computer Users. *Note*: Variance Ln(wage) is author's calculations based on bottom 95% of the weekly earnings distribution to eliminate top-coded cases consistently in all years.

at relatively low levels, inequality grew most rapidly and reached near-peak levels. As computer use continued to rise, inequality stabilized or declined. Inequality ceased growing in 1989, six to eight years before computer use reached 50% of the workforce, the point at which computer use would have had maximal impact on inequality.[16] In fact, if it is computer use that is rewarded in the labor market, its continued growth since the late 1990s may well have been equalizing as the ostensible wage gains extended to an ever-larger majority of the work force.

The temporal pattern of inequality growth is also not consistent with patterns of IT investment, which accelerated dramatically in the late 1990s relative to earlier years and even delivered the long-awaited productivity payoff in the late 1990s (Karoly & Panis, 2004; Oliner & Sichel, 2000, 2003; Greenspan, 2000). From the beginning of the debate over SBTC, there was the awkward fact that computers had little detectable impact on productivity in the 1980s, but were said to have had a dramatic impact on inequality during that time. More recent developments compound the problem by presenting the opposite set of circumstances, dramatic IT-led productivity growth in the late 1990s occurred in tandem with relatively little change in overall inequality.

This creates difficulties for SBTC explanations based on between-job shifts in the composition of employment. If automation and other effects of IT accelerated the rate of labor productivity growth in skill-biased fashion, one would expect shifts in the job composition of employment to have had significant inequality impacts in the late 1990s. SBTC theories need to explain within the terms of their argument how computers not only had large inequality and small productivity effects in the early- and mid-1980s, but also large productivity and small inequality effects in the late 1990s. If productivity trends are an indicator of IT effectiveness, the temporal pattern of inequality growth is essentially the reverse of what would be expected.

As years pass since the SBTC explanation of inequality growth was initially advanced, information technology has continued to grow but inequality has not. Consistent with the principal alternatives to the SBTC thesis, the pattern of inequality growth corresponds most obviously to macroeconomic conditions, particularly the onset and lingering effects of the deep recession and trade shocks of the early 1980s, which upset previous wage norms and perhaps institutionalized the new differentials as an enduring part of the wage structure. There is a growing sense that the recent rise in inequality seems more like an episode than an ongoing trend and reflects changing balance of bargaining power, social norms, and government policy more than secular trends in technology (Atkinson, 1997, pp. 303ff.; Piketty & Saez, 2003, pp. 33ff.; Card & DiNardo, 2002, p. 774).

Nevertheless, even if other forces dominated changes in the wage distribution, it is possible that the spread of computers contributed a disequalizing impact in the 1980s, when both inequality and computer use rose rapidly, and it would be useful to determine its magnitude. This is not captured by most regression studies, which only consider the effects of computers on subgroups defined by user status and education level rather than their impact on all components of the variance. A characteristic can have a disequalizing impact that appears large from its regression coefficient, but whose effect on overall inequality is modest or even equalizing depending on the proportion of the population to which it applies or if it implies higher wages for those in the lower part of the distribution, as in the case of unions (Freeman, 1980; DiNardo et al., 1996).

Indeed, although computer use is positively associated with higher-wage characteristics, such as education, it is also associated with a lower-wage characteristic, female gender (Weinberg, 2000). Regression estimates of the computer wage premium imply that the spread of computers narrowed the gender gap. Wage regressions that control for computer use show the gender gap widening between 1984 and 1989, while those that omit computer use and allow its effect to be picked up by other variables show almost no change in the gender gap (Krueger, 1993, p. 52). By implication, the disproportionate spread of computers to women helped stabilize inequality across gender groups, even as it may have contributed to inequality growth across education groups.

In addition, reanalysis of the October 1989 data shows that the residual variance from a standard wage equation is 10% lower for computer users than nonusers, implying that a rising share of computer users between 1984 and 1989 also moderated growth of within-group inequality, an important source of inequality growth in the 1980s (author's calculations).[17] In short, even assuming computer use is causally associated with higher wages, its net effect on overall inequality is indeterminate *à priori* because equalizing impacts may partly or fully offset disequalizing effects.

One way to evaluate the net effect of increased computer use on overall inequality is to adjust the level of computer use in later years to 1984 levels and then compare the wage distributions for the adjusted and unadjusted samples. In a series of papers, DiNardo, Fortin, and Lemieux (DiNardo et al., 1996; DiNardo & Lemieux, 1997; Fortin & Lemieux, 1997) develop and apply a procedure to investigate the effects on inequality of declining unionization rates, declining real minimum wage, and industry deregulation. Their method adjusts CPS sample weights at time t_1 so that the incidence of a given characteristic, such as unionization, is the same as at time t_0, while other differences in variance components, such as within-group inequality, are left unchanged. The difference in inequality levels between the original and adjusted t_1 samples is a measure of the effect of changing unionization rates on overall inequality.

To calculate reweighting factors, the probability of computer use at work conditional on a vector of noncomputer characteristics, **X**, was estimated for the 1984 and 1989 CPS samples using a logit model.[18] The probabilities from the 1989 logit were used to derive the denominator of the reweighting factors, while the 1984 logit coefficients were applied to the 1989 sample (i.e., $\hat{b}_{84}\,\mathbf{X}_{89}$) to derive the numerator. The CPS sample weights were multiplied by the resulting adjustment factors to obtain predicted probabilities of computer use for the 1989 sample assuming the 1984 relationships predicting computer use remained in effect. Specifically, CPS sample weights were adjusted by multiplying them by factors equal to:

$$\frac{\mathrm{pr}\left(\text{use computers} = 1 | \mathbf{X}_{89},\ \ t_{\beta} = 84\right)}{\mathrm{pr}\left(\text{use computers} = 1 | \mathbf{X}_{89},\ \ t_{\beta} = 89\right)} \quad \text{for computer users and}$$

$$\frac{\mathrm{pr}\left(\text{use computers} = 0 | \mathbf{X}_{89},\ \ t_{\beta} = 84\right)}{\mathrm{pr}\left(\text{use computers} = 0 | \mathbf{X}_{89},\ \ t_{\beta} = 89\right)} \quad \text{for noncomputer users}$$

where $\mathbf{X}$ is the vector of control variables whose distributions are to remain fixed at 1989 levels and t_{β} indexes the year whose coefficients are applied to the 1989 sample. Applying these factors effectively adjusts the 1989 sample's group-specific rates of computer use downward to 1984 levels, where groups are defined by the variables in $\mathbf{X}$, while the distribution of these noncomputer characteristics and the structure of wages remain as observed in 1989. The same adjustment was also applied to the computer use CPS supplements for 1993, 1997, and 2001.

Conceptually, this method is analogous to other, more familiar decomposition and standardization techniques (e.g., Kitagawa, 1955; Oaxaca, 1973; Blinder, 1973), but while they typically decompose differences in means or rates into shares attributable to differences in the characteristics and the returns to them, this technique permits decomposition of changes in variances or other measures of dispersion.

Comparing wage inequality in the actual and adjusted 1989 samples answers the question, *What would be the level of inequality in 1989 if rates of computer use within groups remained at their 1984 level, but the other components of the variance remained as observed in 1989 (i.e., the distribution of noncomputer characteristics (***X***), the returns to those characteristics except insofar as they were affected by the proportions of computer users within the groups they define, the returns to computer use, and the levels of within-group inequality)?*[19] After adjusting the rates of computer usage to 1984 levels, the returns to computer use can also be adjusted to 1984 levels by subtracting the growth in the computer premium from the wages of users. This exercise assumes that the observed wage premium associated with computer use is unbiased net of the control variables, which is a generous assumption given by the previous results.

The results of this analysis, presented in Table 7, indicate that the spread of computer use at work between 1984 and 1989 had a very slight equalizing impact on the overall wage distribution. The actual variance of log wages in 1989 (0.3222) was 6% higher than in 1984 (0.3039). Reweighting the 1989 sample to reflect the lower rates of computer use prevailing in 1984 yields

Table 7. Estimated Actual and Counterfactual Variance of Log Wages.

	1984	1989	1993	1997	2001
Raw values –					
No reweighting					
All	0.3039	0.3222	0.3076	0.3192	0.3290
Men	0.3057	0.3278	0.3191	0.3281	0.3425
Women	0.2435	0.2735	0.2744	0.2839	0.2920
Rates of computer use (1984)					
All		0.3254	0.3100	0.3260	0.3260
Men		0.3296	0.3169	0.3269	0.3390
Women		0.2760	0.2822	0.3023	0.2921
Rates of and returns to computer use (1984)					
All		0.3236	0.3070	0.3240	0.3253
Men		0.3277	0.3134	0.3245	0.3383
Women		0.2740	0.2791	0.3003	0.2914

Note: Figures are calculated from the October Current Population Survey for each year using the sample deletions and wage definition in Krueger (1993). Top-coded cases in 1984 and 1997 are assigned estimated values (from Krueger, 1993, p. 56).

a variance estimate for 1989 that is over 7% higher than in 1984 (0.3254), implying a lower rate of computer use in 1989 would have raised inequality slightly above its observed level. Since the variance for both men and women are also higher in the reweighted sample, the equalizing impact of increased computer use probably reflects the lower within-group variance of computer users more than the equalizing effects of computers on the gender wage gap.

This impression is reinforced when the returns to computer use are adjusted down to 1984 levels, which lowers the variance in the reweighted sample (0.3236). This indicates that the rise in the computer premium over this period had a net disequalizing impact, which is not what one would expect if the equalizing effect of computers on the gender gap dominated. Nevertheless, what is most striking is the limited impact of any sort which greater computer use had on the variance of log wages compared to the actual increase between 1984 and 1989.

Results for later years are similar. Adjusting computer use and the returns to computer use for 1993–2001 to levels prevailing in 1984 has very small effects on wage inequality and in the most cases the effect is inappropriately signed according to the expectations of the theory of skill-biased technological change.

6. DISCUSSION AND CONCLUSION

The literature on skill-biased technological change implies at least four ways in which computers might affect wage levels and inequality: the addition of computer-specific human capital requirements, increases in general human capital requirements for users, contextual effects that raise general human capital requirements for both users and nonusers, and changes in occupational composition regardless of any within-job effects. However, analyses suggest that the additional skills due to the spread of computers do not seem to have been as scarce, expensive, and important in the growth of overall wage inequality as the SBTC account holds.

Seven measures of noncomputer job content are associated with high returns similar to computers when entered individually into a standard wage equation, suggesting all reflect some common, unobserved aspect(s) of human capital or occupational position that antedate or are otherwise causally unrelated to computers. Analyses confirmed that computer use is strongly associated with occupational work roles in particular and that there is strong persistence in occupational characteristics over time, indicating the importance of controlling for these longstanding differences. Including pre-computer and other characteristics in models reduces the returns to computer use *per se* to insignificance, though there may be a small wage premium to general human capital occasioned by computer use on the order of 3–4%, net of other job characteristics. There are no returns to computer training in the cross-section and no penalty for self-reported computer skill deficits, though possible selection issues and measurement error argue for caution in interpreting these results. Computers also do not appear to have contextual effects on the human capital of either users or nonusers. Using computers at work oneself probably involves only modest increases in skill requirements for most workers and working in a computerized environment in itself has no effect.

Taking a wider perspective, the timing of inequality growth casts doubt on the role of computers as the prime mover. Inequality rose more rapidly within a few short years in the recessionary early 1980s, prior to the widespread use of computers at work, and essentially stopped rising by 1989, even as computer use and investment continued to grow, quite rapidly in the case of investment. Adjusting rates of computer use for 1989–2001 to 1984 levels, while holding other variables constant suggests that the spread of computers did not even play a secondary role in raising inequality when all components of its contribution to the overall variance are taken into account, even assuming simple estimates of the returns to computer use are unbiased. The timing of inequality growth suggests that the decline of institutional protections and traditional wage norms account for a larger portion of the rise in inequality than growth in the cognitive demands of work resulting from the increased use of computers.

This does not rule out the possibility that the computers and other information technology may have altered the skill and wage distributions through between-job compositional shifts. This is a separate and difficult question because of the problem of two-way causation and is not addressed here, except insofar as the timing of inequality growth, computer investment, and productivity growth also call into question between-job effects of computer diffusion.

The very high estimates of the returns to computer use that continue to influence discussions of SBTC should always have been the object of some skepticism, certainly if interpreted as reflecting returns to computer-specific human capital. One suspects that computer training for most jobs is mostly a matter of weeks or a few months if further learning-by-doing is included, while a wage premium of 10–15% is equivalent to roughly one and a half years of education. This seems quite implausible even as an estimate of the general cognitive skills occasioned by the introduction of computers within jobs.

This suggests two points that deserve recognition in the debates over SBTC. The extraordinary rise of computer technology in the past 25 years understandably attracts attention. However, studies of labor markets need to distinguish the internal complexity of these products and the skills of high-level users (computer scientists, systems analysts, programmers) from the skills most users require. Most workers do not program or perform high-level systems troubleshooting at work. They may not need much beyond good keyboard skills, knowledge of a limited set of operating systems and application-program operations, and some modest increment to other cognitive skills.

Related to this is another consideration that is absent from most discussions of the impact of high technology on labor markets. The prevailing assumption seems to be that workers must adjust to technology. While no doubt true, it neglects the fact that product market dynamics also force the technology to adjust to users. Technology that is hard to use is at a competitive disadvantage, all else equal. If word processing required the skills to program in FORTRAN or C, there would be far fewer word processors. The field of human factors and the actual history of computer software suggest that ease of use is an important consideration in product development, most notably the development of the graphical user interface, whose icons and pull-down menus replaced command lines with pictures (Carroll, 1997, pp. 67ff.; Staggers, 2000; Margono & Schneiderman, 1987). There are some complexities to the process, notably the tendency for software to become feature-rich, hence more complex, even as core functions are simplified. Actual data on computer training times would go a long way toward clarifying their impact on the cognitive complexity of work. In the absence of such data, the preceding provides some caution against accepting too quickly the all too easy equation of high technology and high skill requirements.

NOTES

1. These results are from a regression of frequency of computer use (never, < once per week, one or more times per week, every day) on occupation and education only, using Stata's –areg- procedure. The data are the January 1991 supplement to the Current Population Survey, described further in the next section.
2. Computers may also alter skill demands by changing the relative numbers of high- and low-skill jobs without affecting their task content, an alternative mechanism of IT-based SBTC discussed further below.
3. In addition to cognitive skills, high-performance work practices also increase the demand for teamwork, autonomy, and discretionary effort, but the latter remain peripheral to SBTC research because they involve interpersonal skills and issues of worker motivation associated more with efficiency wages than returns to human capital.
4. I would like to thank Libbie Stephenson of the UCLA ISSR Data Archives for making this data available to me.
5. Likewise, 1991 wages at the individual level correlate almost as closely with the three aggregate-level measures from the late 1960s (occupational pay, industry pay, occupational complexity) as with the analogous aggregate-level variables constructed from the January 1991 CPS supplement itself (see Column 1, Table 1).
6. I thank Daniel Feenberg of the National Bureau of Economic Research for making available the CPS ORG files.
7. All models in Tables 1–3 control for years of education, experience, experience^2, and for female, black, other non-white, part-time status, union status, resident of metropolitan area, married, married*female, veteran, and region.
8. If the computer variable is dichotomized to make it comparable to the October CPS supplements the coefficient is 0.215, which is a bit *higher* than comparable estimates for both October 1989 (0.188) and October 1993 (0.203) (Krueger, 1993; Autor et al., 1997). This may reflect the fact that the October supplements permit proxy responses from other household members, whereas the January 1991 supplement did not.

9. Percentage wage differences associated with a unit difference in predictors are calculated using the formula: e^b-1.

10. This result differs slightly from that reported in note 4 because Table 1 used rotation groups 4 and 8 from the January 1991 supplement, but Table 2 and Table 3 use January 1991 rotation groups 2 and 6 to enable matching with earners in the March 1991 supplement.

11. All models using the aggregate variables use robust standard errors with a correction for clustering within three-digit occupation.

12. After reading this result some suggested that there is no observed wage penalty because it is the highly skilled with the most demanding jobs, who are the most aware of their computer-skill deficits. However, the correlation between education and self-reported computer-skill deficit is negative ($r = -0.16$, $p<0.01$), supporting the interpretation above.

13. All models in Table 6 control for age, gender, race, marital status and its interaction with gender, part-time work status, and region.

14. Because the CPS did not inflation-adjust the top-code value for weekly earnings between 1979 and 1988, all variances are calculated on the bottom 95% of the weekly earnings distribution, which eliminates all top-coded cases in all years and the progressive negative bias in the variance that would result from using full samples with increasing proportions of top-coded high earners. As a check, variances were also estimated from samples truncated at the 99th percentile for years in which this cutpoint eliminates all top-coded values (1979–1980, 1989–1993) and the results were very similar. Results also did not differ substantively when using the full sample and imputing a top-code value suggested by Autor et al. (1997, p. A1).

15. The one significant exception is the transitory and rather anomalous increase in inequality in 1994, observed by others, and generally attributed to changes in the content and administration of the CPS, though still imperfectly understood (Katz & Autor, 1999, p. 1485; Card & Dinardo, 2002, p. 748; cf. Bernstein & Mishel, 1997).

16. If $y = bx+c$ and $\mathrm{var}(y) = b^2 \times \mathrm{var}(x)$, then $\mathrm{var}(y)$ is maximized when $x = 0.5$, all else equal. Simple weighted CPS tabulations indicate that computer users were 50.5% of all workers in 1997. However, logistic regressions show that proxy respondents in the October CPS somewhat underreport computer use at work for other household members, controlling for other characteristics (education, gender, race, one-digit occupation and industry, marital status and its interaction with gender, part-time status, metropolitan residence, and region of residence). When the probability of computer use is adjusted for proxy reporting conditional on these variables, rates of computer use are about two percentage points higher in 1993 and 1997 than simple tabulations suggest, implying computer use reached 50% of the work force around January 1995.

17. The lower residual inequality among computer users does not seem to reflect selection processes that might lead to greater homogeneity among users after controlling for observable variables because the residual variance among users *declined* between 1984 and 1989 even as they accounted for a larger share of the work force, while residual variance among nonusers increased, as shown in the table below. The top panel uses a suggested imputed value for top-coded cases in 1984 (Krueger, 1993,

p. 56) and the lower panel uses truncated samples in both the years to eliminate top-coded cases. In both panels the residual variance declines for users and increases for nonusers.

Residual Variance among Computer Users and Nonusers

Variance	All	Users	Nonusers
Top-code values imputed, 1984			
1984	0.1731	0.1717	0.1673
1989	0.1770	0.1608	0.1779
Top 2.67% of cases deleted, both years			
1984	0.1582	0.1522	0.1548
1989	0.1630	0.1432	0.1663

18. The predictors were years of education, experience, experience^2, and dummies for female, black, other non-whites, part-time status, union status, one-digit occupation, one-digit industry, metropolitan residence, marital status and its interaction with female, veteran status, and region of residence.

19. This comparison between actual and counterfactual wage distributions describes the effects of the changing percentage of computer users on the wage distribution and assumes no spillover effects on nonusers, which previous results suggest is reasonable. It also assumes that changing rates of computer use did not influence the returns to other characteristics except perhaps by changing the user/nonuser composition of groups. Thus, if computers raised the returns to education through between-job effects, such as automation and displacement of less educated workers, as opposed to within-job effects that increase skill demands by putting computers on some workers' desks, the impact will not be attributed to computers here, which is also the case with other studies dealing with the computer use premium. However, since the measured returns to computer use likely overestimate actual returns, as argued above, there is also a potential offsetting bias in favor of finding a large effect of computer diffusion on the growth of inequality.

ACKNOWLEDGEMENT

I thank Christopher Winship, Christopher Jencks, Lee Rainwater, Michael Piore, John DiNardo, Jörn-Steffen Pischke, Thomas Lemieux, Richard Freeman, John Budd, Maury Gittleman, and Greg Weyland for helpful discussions. I thank the Center for Industrial Competitiveness, University of Massachusetts – Lowell and the Levy Economics Institute for supporting this research. I am responsible for any remaining problems.

REFERENCES

Atkinson, A. B. (1997). Bringing income distribution in from the cold. *The Economic Journal*, *107*, 297–321.

Autor, D. H., Katz, L. F., & Krueger, A. B. (1997). *Computing inequality: Have computers changed the labor market?* National Bureau of Economic Research Working Paper 5956, Cambridge, MA.

Autor, D. H., Katz, L. F., & Krueger, A. B. (1998). Computing inequality: Have computers changed the labor market? *Quarterly Journal of Economics*, *113*, 1169–1213.

Autor, D. H., Murnane, R. J., & Levy, F. (2002). Upstairs, downstairs: Computers and skills on two floors of a large bank. *Industrial and Labor Relations Review*, *55*, 432–447.

Bartel, A. P., Ichniowski, C., & Shaw, K. (2000). *New technology, human resource practices, and skill requirements: Evidence from plant visits in three industries*. Paper presented at the American Economics Association Meetings, January 2001.

Berman, E., Bound, J., & Griliches, Z. (1994). Changes in the demand for skilled labor within U.S. manufacturing: Evidence from the annual survey of manufactures. *Quarterly Journal of Economics*, *109*, 367–397.

Bernstein, J., & Mishel, L. (1997). Has wage inequality stopped growing? *Monthly Labor Review*, *December*, 3–16.

Black, S. E., & Lynch, L. M. (2004). What's driving the new economy: The benefits of workplace innovation. *The Economic Journal*, *114*, F97–F116.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, *8*, 436–455.

Borghans, L., & ter Weel, B. (2002). Do older workers have more trouble using a computer than younger workers? *Research in Labor Economics*, *21*, 139–173.

Borghans, L., & ter Weel, B. (2004). Are computer skills the new basic skills? The returns to computer, writing, and math skills in Britain. *Labour Economics*, *11*, 85–98.

Bresnahan, T. F. (1999). Computerisation and wage dispersion: An analytical reinterpretation. *The Economic Journal*, *109*, F390–F415.

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics*, *117*, 339–376.

Cappelli, P. & Carter, W. (2000). *Computers, Work Organization, and Wage Outcomes*. National Bureau of Economic Research Working Paper 7987. Cambridge, MA.

Card, D., & DiNardo, J. (2002). Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics*, *20*, 733–783.

Carroll, J. M. (1997). Human-computer interaction: Psychology as a science of design. *Annual Review of Psychology*, *48*, 61–73.

Dickerson, A., & Green, F. (2004). The growth and valuation of computing and other generic skills. *Oxford Economic Papers*, *56*, 371–406.

DiNardo, J. E. & Pischke, J.-S. (1996). *The returns to computer use revisited: Have pencils changed the wage structure too?* National Bureau of Economic Research Working Paper no. 5606, Cambridge, MA.

DiNardo, J. E., & Pischke, J.-S. (1997). The returns to computer use revisited: Have pencils changed the wage structure too? *Quarterly Journal of Economics*, *112*, 290–303.

DiNardo, J. E., & Lemieux, T. (1997). Diverging male wage inequality in the United States and Canada, 1981–1988: Do institutions explain the difference? *Industrial and Labor Relations Review*, *50*, 629–651.

DiNardo, J. E., Fortin, N. M., & Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach. *Econometrica*, *64*, 1001–1044.

Doms, M., Dunne, T., & Troske, K. R. (1997). Workers, wages, and technology. *Quarterly Journal of Economics*, *112*, 253–290.

Entorf, H., Gollac, M., & Kramarz, F. (1999). New technologies, wages, and worker selection. *Journal of Labor Economics*, *17*, 464–491.

Fernandez, R. M. (2001). Skill-biased technological change and wage inequality: Evidence from a plant retooling. *American Journal of Sociology*, *107*, 273–320.

Fortin, N. M., & Lemieux, T. (1997). Institutional change and rising wage inequality: Is there a linkage? *Journal of Economic Perspectives*, *11*, 75–96.

Freeman, R. B. (1980). Unionism and the dispersion of wages. *Industrial and Labor Relations Review*, *34*, 3–23.

Freeman, R. B. (1984). Longitudinal analyses of the effects of trade unions. *Journal of Labor Economics*, *2*, 1–26.

Freeman, R. B., & Katz, L. F. (1994). Rising wage inequality: The U.S. vs. other advanced countries. In: R. B. Freeman (Ed.), *Working under different rules* (pp. 29–62). New York: Russell Sage.

Friedberg, L. (2003). The impact of technological change on older workers: Evidence from data on computer use. *Industrial & Labor Relations Review*, *56*, 511–529.

Galbraith, J. K. (1998). *Created unequal: The crisis in American pay*. New York: Free Press.

Goldstein, H., & Fraser, B. S. (1985). *Training for work in the computer age: How workers who use computers get their training*. National Commission for Employment Policy. Washington, DC.

Goss, E. P., & Phillips, J. M. (2002). How information technology affects wages: Evidence using internet usage as a proxy for IT skills. *Journal of Labor Research*, *23*, 463–474.

Gottschalk, P. (1997). Inequality, income growth, and mobility: The basic facts. *Journal of Economic Perspectives*, *11*, 21–40.

Gottschalk, P., & Smeeding, T. M. (1997). Cross-national comparisons of earnings and income inequality. *Journal of Economic Literature*, *35*, 633–687.

Greenspan, A. (2000). *The revolution in information technology*. Talk before the Boston College Conference on the New Economy, Boston, MA.

Haisken-DeNew, J. P., & Schmidt, C. M. (1999). *Money for nothing and your chips for free? The anatomy of the PC wage differential*. Discussion Paper No. 86. Bonn: Institute for the Study of Labor (IZA).

Handel, M. J. (2000). *Is there a skills crisis? Trends in skill requirements, technology, and wage inequality in the United States*. Jerome Levy Economics Institute Working Paper No. 295.

Handel, M. J., & Levine, D. I. (2004). The effects of New Work practices on workers. *Industrial Relations*, *43*, 1–43.

Howell, D. R. (1995). Collapsing wages and rising inequality: Has computerization shifted the demand for skills? *Challenge*, *38*, 27–35.

Howell, D. R. (1997). *Institutional failure and the American worker: The collapse of lowskill wages*. Public Policy Brief No. 29. Levy Economics Institute.

Howell, D. R., & Wieler, S. S. (1998). Skill-biased technological change and the wage collapse in the United States: A critical perspective. *Eastern Economic Journal*, *24*, 343–366.

Kaplan, D. L., & Casey, M. C. (1958). *Occupational trends in the United States: 1900–1950*. Washington, DC: United States Department of Commerce. Bureau of the Census. Working Paper Number 5.

Karoly, L. A., & Panis, C. W. A. (2004). *The 21st century at work*. Santa Monica, CA: RAND.

Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics*, *107*, 35–79.

Katz, L. F., & Autor, D. H. (1999). Changes in the wage structure and earnings inequality. In: O. Ashenfelter & D. Card (Eds), *Handbook of labor economics*, (Vol. 3, pp. 1463–1555). New York: Elsevier.

Keefe, J. H. (1991). Numerically controlled machine tools and worker skills. *Industrial and Labor Relations Review*, *44*, 503–519.

Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association*, *50*, 1168–1194.

Krueger, A. B. (1993). How computers have changed the wage structure: Evidence from microdata, 1984–1989. *Quarterly Journal of Economics*, *108*, 33–61.

Krueger, A. B. (2000). *The digital divide in educating African-American students and workers*. Working Paper #434. Industrial Relations Section. Princeton University. Available at http://www.irs.princeton.edu.

Krueger, A. B., & Summers, L. H. (1987). Reflections on the inter-industry wage structure. In: K. Lang & J. Leonhard (Eds), *Unemployment and the structure of labor markets* (pp. 17–48). New York: Basil Blackwell.

Lee, S.-H., & Kim, J. (2004). Has the Internet changed the wage structure too? *Labour Economics*, *11*, 119–127.

Levin, H. M., & Rumberger, R. W. (1986). Education and training needs for using computers in small businesses. *Educational Evaluation and Policy Analysis*, *8*, 423–434.

Levy, F., & Murnane, R. J. (1992). U.S. earnings levels and earnings inequality: A review of recent trends and proposed explanations. *Journal of Economic Literature*, *30*, 1333–1381.

Levy, F., & Murnane, R. J. (1996). With what skills are computers a complement? *American Economic Review: Papers and Proceedings 86* (pp. 258–262).

Levy, F., Beamish, A., Murnane, R. J., & Autor, D. (1999). *Computerization and skills: Examples from a car dealership*. Manuscript, Massachusetts Institute of Technology.

Margono, S., & Schneiderman, B. (1987). A study of file manipulation by novices using commands vs. direct manipulation. In: *Proceedings of the Twenty-sixth Annual Technical Symposium* (pp. 154–159), ACM: Washington, DC.

Melman, S. (1951). *The rise of administrative overhead in the manufacturing industries of the United States, 1899–1947*. Oxford Economic Papers (pp. 63–112).

Miller, A. R., Treiman, D. J., Cain, P. S., & Roos, P. A. (1980). *Work, jobs, and occupations: A critical review of the dictionary of occupational titles*. Washington, DC: National Academy Press.

Mishel, L., & Bernstein, J. (1998). Technology and the wage structure: Has technology's impact accelerated since the 1970s? *Research in Labor Economics*, *17*, 305–355.

Murnane, R. J., & Levy, F. (1996). *Teaching the new basic skills*. New York: Free Press.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, *14*, 693–709.

Oliner, S. D., & Sichel, D. E. (2000). The resurgence of growth in the late 1990s: Is information technology the story? *Journal of Economic Perspectives*, *14*, 3–22.

Oliner, S. D., & Sichel, D. E. (2003). Information technology and productivity: Where are we now and where are we going? *Journal of Policy Modeling*, *25*, 477–503.

Piketty, T., & Saez, E. (2003). Income inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, *118*, 1–39.

Shaw, K. (2002). By what means does information technology affect employment and wages? The value of IT, HRM practices, and knowledge capital. In: N. Greenan, Y. L'Horty & J. Mairesse (Eds), *Productivity, inequality, and the digital economy*. Cambridge, MA: MIT Press.

Siegel, D. S. (1999). *Skill-biased technological change: Evidence from a firm-level survey*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Staggers, N. (2000). Comparing response time, errors, and satisfaction between text-based and graphical user interfaces during nursing order tasks. *Journal of American Medical Informatics Association*, *7*, 164–176.

Weinberg, B. A. (2000). Computer use and the demand for female workers. *Industrial and Labor Relations Review*, *53*, 290–308.

Zuboff, S. (1988). *In the age of the smart machine: The future of work and power*. New York: Basic Books.

THE INFLUENCE OF STOCKS AND FLOWS ON MIGRANTS' LOCATION CHOICES☆

Thomas Bauer, Gil S. Epstein and Ira N. Gang

ABSTRACT

We examine the determinants of a current migrant's location choice emphasizing the relative importance and interaction of migrant stocks and flows. We show that both stocks and flow have significant impacts on the migrant's decision of where to locate. The significance and size of the effects vary according to legal status and whether the migrant is a "new" or a "repeat" migrant.

1. INTRODUCTION

A characteristic of international migration is the clustering of immigrants in ethnic communities. Prominent examples are the concentration of Turks in

☆Earlier versions of this paper were presented at the NEUDC Conference, Ithaca, NY; IZA Transatlantic Labor Economics Workshop, Ammersee, Germany; ESRC Workshop, Trade, Migration and Labour Markets, University of East Anglia, Norwich; Centre for Economic Development and Institutions (CEDI) Inaugural Conference, Brunel University; as well as at several university workshops.

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 199–229

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26006-0

Germany, Tamils in Switzerland, Moroccans in the Netherlands and Belgium, Italians in Argentina, Greeks in Australia, and Ukrainians in Canada. Clustering may be very narrow, such as when immigrants from a town or region are concentrated in a specific foreign town or region. For example, Macedonians from Skopje have come to make up a notable part of the population of Gothenburg, Sweden. In the United States, noticeable clusters of Mexican immigrants exist in California, Texas, Florida, and Chicago. Fifty-eight percent of migrants from Guanajuato, the Mexican state with the highest emigration rate to the US, go to California and another 23 percent to Texas.

The prevailing explanation for immigrant clusters is the existence of beneficial network externalities when previous immigrants provide shelter and work, assistance in obtaining credit, and/or generally reduce the stress of relocating to a foreign culture (see Gottlieb, 1987; Grossman, 1989; Marks, 1989; Church & King, 1983; Carrington, Detragiache, & Vishwanath, 1996; Chiswick & Miller, 1996; Zahniser, 1999; Munshi, 2003). Ethnic networks, however, might also be associated with negative externalities. Disadvantageous network externalities may arise if immigration is subject to adverse selection, or if increases in immigrant concentration increases competition for jobs and lower immigrants' wages. Under certain conditions the tendency to cluster may lower incentives to learn the language of the host country, which in turn may "trap" migrants in poverty (Bauer, Epstein, & Gang, 2005). These negative network externalities limit the benefits immigrants can obtain from clustering.

Several studies investigate the determinants of location choice by immigrants in the United States. Bartel (1989) finds that post-1964 migrants to the US tend to locate in cities with a high concentration of immigrants of similar ethnicity. She further shows that highly skilled migrants are less geographically concentrated and rely less on the location of fellow compatriots. Similarly, Jaeger (2000), who differentiate between immigrants of different admission statuses, finds that immigrants tend to locate where former immigrants of the same ethnicity are concentrated.

Migrants consider several factors in making their decisions about where to move, including the clustering of compatriots and similar folk in various localities. Ties of kinship, friendship, and village, link migrants, former migrants, and non-migrants in the home and host country. This paper contributes to the literature by investigating the differing effects of "stock" and "flow" factors on migrants' location decisions. Stock factors measure the degree to which migrants may view a US location as (ethnically) hospitable and the availability of information about specific locations. We characterize two types of stock factors, an ethnic goods component and a village migration history component. Our flow factor measures the tendency of migrants to follow the paths of very recent migrants from their own villages.

These factors offer different information to a potential migrant. The ethnic goods component sends signals to the migrant about the possibility of living in a culturally similar environment, i.e., speaking his native language, listening to his music, reading his own newspapers, and eating ethnic food. The ethnic goods factor reduces the monetary and psychic costs of migrating. The village migration history component largely captures information about the host region received in the home village. This includes, for example, information on the labor and housing market, and information on specific employers in a region. In addition, the migrant may be able to count on contacts in a specific location established by former migrants from the same village. This factor reflects the probability of receiving help from compatriots. The flow factor represents potential herd behavior by migrants, a sort of "peer emulation effect." Following the argument by Epstein (2002), migrants may choose a location on the supposition that recent migrants had information that he does not have. We examine the relative importance of migrant stocks and flows as explanations of immigrant location choice, also accounting for several other determinants.

Until the appearance of the paper by Polachek and Horvath (1977) much of migration theory treated migration as an individual investment decision. Family members other than the household head are not always explicitly considered. However, other members are clearly influential in migration decisions. Polachek and Horvath (1977) established the foundations for models of location choice that take into consideration all the different type of considerations. They do so by adopting a life cycle approach used in human capital theories of earnings accumulation, accounting for household considerations in both a general theoretical and empirical model. More importantly, migration was analyzed within a nonstochastic framework and remigration was endogenously explained.

We describe our data and define and characterize the variables we employ in Section 2. Section 3 presents our empirical results, while Section 4 offers a theoretical model explaining our results. Section 5 concludes.

2. THE GEOGRAPHIC DISTRIBUTION OF MEXICAN MIGRANTS IN THE US

In absolute numbers, the US is the world's largest country of immigration; Mexico is the world's major country of emigration; migration from Mexico to the United States is the largest sustained flow of migration in the world. Empirical evidence suggests that there exist strong network effects in Mexican migration (Bustamante, 1998; Munshi, 2003; Winters, de Janvry, & Sadoulet, 2001). We explore the stock and flow effects of clustering on migrants' location choices using individual and village level data on Mexican-US migration available through the Mexican Migration Project.[1] The data comprise more than 7,000 households in 52 communities in the states of Colima, Guanajuato, Guerrero, Jalisco, Michoacán, Nayarit, San Luis Potosí, and Zacatecas. The data set provides information on the socioeconomic characteristics of household heads, such as age, education and marital status, their migration histories including information on year of migration, costs of border crossing, documentation and location in the United States. In calculating our flow variable and one of our stock variables, we use an event-history file containing detailed labor and family histories of each household head, such as information on the number of trips to the United States, the duration of each trip, and related information, for each year from the birth of the household head until the year of the survey.

We calculate for each year t ($t = 1,\ldots,T$) the cumulative migration experience (in months) for each migrant I ($i = 1,\ldots,N$) from the Mexican community m ($m = 1,\ldots,M$) in each US location j ($j = 1,\ldots,J$).[2] The cumulative migration experience of community m in US location j, EXP_{mjt}, is

$$EXP_{mjT} = \sum_{t=1}^{T} \sum_{i=1}^{N} M_{mjit} \tag{1}$$

where M_{mjit} is a dummy variable that takes the value 1 if an individual i in the Mexican community m is in US location j at year t.

We define the *Village Migration Experience*, VME_{mjt}, as the cumulative migration experience for each migrant i from the Mexican community m in each US location *j, relative to* the total experience of that village in the US. The measure captures the Mexican village's migration experience in a US location at the time a person makes his migration decision, and is calculated as

$$\textit{Village Migration Experience} = VME_{mjt} = \frac{EXP_{mjt}}{\sum_{j=1}^{J} EXP_{mjt}} \times 100 \tag{2}$$

In addition to the migration experience of a particular Mexican village, we use the *Mexican Share of the Total Population* in a US location (see

Appendix A for a description of the calculation of this variable). This second stock variable disregards specific village information, instead capturing the concentration of ethnic goods in a location relative to other locations. Adding this second stock variable helps distinguish a generalized stock effect from village-specific links.[3]

We also examine the impact of the *flow* of migrants during the year before an individual migrates, calculated as the *Change in Village Migration Experience*, $FLOW_{mjt}$, in the year before an individual migrates,

$$FLOW_{mjt} = VME_{mjt} - VME_{mj(t-1)} \tag{3}$$

where $VME_{mjt} \geq VME_{mj(t-1)} \geq 0$. We visualize that the person makes his/her decision at the end of period t. This enables us to see how the relative flow of migrants between t–1 and t affects the probability of migrating to a particular location at time t. Since we are interested in the flow to a certain destination relative to other locations, we present define the flow variable in relative terms.

In an uncertain environment, networks provide information about the host locations. Although knowledge of current labor market conditions may deteriorate over time, migrants who have returned several years ago may still provide key links and support for new migrants, such as arranging a coyote to smuggle them across the border, provide information about alternative locations, or simply telling stories about their experiences and passing on knowledge. More recent migrants with current first hand information about job opportunities are likely to help their community members find jobs. Others from the broader ethnic group set the tone and atmosphere of living in locations away from home.

Our village migration experience variable and our flow variable are scaled by the village's experience in the United States, making them relative measures reflecting the influence of the village network on location choice. The spread of migrants across the US has an important impact on the utility a migrant obtains from the network; there are both positive and negative network effects. As the concentration of migrants' increase, their wages decrease; however, as geographic mobility is high and US labor markets are highly integrated, wages and network effects are relative. A similar argument can be put forward when considering the attitude of the local population toward immigrants.

To control for other factors that may affect the utility levels associated with a US location, we include several variables capturing the economic and social characteristics of a location in the multivariate analysis. These factors include, for each US location, population size, the consumer price index to capture cost of living differences, and the unemployment rate in those locations. Though the unemployment rate is sometimes problematic in migration studies, the literature often assumes that the probability of choosing a particular location decreases with the unemployment rate in this location (see the discussion in Jaeger, 2000).[4] A detailed description of the variables used in the empirical analysis is given in Appendix A.

Migration costs affect location choice. Most Mexican migrants have a very low income in their home village. Therefore, the cost of migrating may be an important factor in determining the specific location to which to migrate. To control for these costs we include road mileage from the migrant's origin village in Mexico to the alternative US locations.[5] We also examine US location specific fixed effects in order to control for time constant determinants of the location choice.

The covariates just discussed are US location specific, as dictated by our desire to examine determinants of migrants' location choice, and by the conditional logit model we discuss in the next section. In addition we use several individual specific variables and examine how these individual dimensions interact with our stock and flow effect variables. We look at the interaction of the location specific variables with skill level, legal status, and whether it is someone's first trip to the US or their last trip (as recorded in the data). Migrants with six or fewer years of schooling are assumed to be unskilled; those with more than 6 years are considered skilled. Migrants report themselves whether they migrated legally (documented) or illegally (undocumented). We expect the migrant's use of the information provided by the stock of previous migrants or their inclination to follow the flow will vary depending on these factors. In particular, we expect the impacts of stocks and flows to vary between the first-time an individual migrates to the US and repeat movers.

Table 1 presents a description of the data we use in our analysis. For the first migration, we have information on 1739 individuals from 47 Mexican villages who migrated to 43 different locations in the US The geographic unit in the US varies – some are cities, some are parts of a county, and some are counties – but they are generally recognizable as sensible divisions (see Appendices B and C for a list of the locations). We assume that each person has the possibility of going to each of these 43 locations, but does not consider other locations.[6] This generates 74,777 observations – each person may or may not go to each of the 43 locations. For the last migration, we have 1,561 individuals from 47 Mexican villages going to 46 US locations, resulting in 71,806 observations. Unskilled migrants dominate, comprising

Table 1. Descriptive Statistics, Means of US Recipient Locations.

		First Migration	Last Migration
Unemployment rate (in %)		7.103	7.310
		(3.309)	(3.413)
CPI		85.203	110.821
		(31.979)	(30.624)
Total population (in 100,000)		13.351	14.066
		(18.867)	(19.216)
Miles		1459.956	1431.984
		(527.774)	(510.941)
Mexican share of population (in %)		5.511	5.568
		(6.476)	(6.163)
Village migration experience (in %)		1.986	1.870
		(7.622)	(7.563)
Flow (in %)		0.878	0.442
		(46.054)	(26.340)
Unskilled legal	(Observations)	3,784	22,908
	(Individuals)	88	498
Unskilled illegal	(Observations)	46,268	30,360
	(Individuals)	1,076	660
Skilled legal	(Observations)	5,289	11,040
	(Individuals)	123	240
Skilled illegal	(Observations)	19,436	7,498
	(Individuals)	452	163
Total	(Observations)	74,777	71,806
	(Individuals)	1,739	1,561
Number of mexican Villages		47	47
Number of US locations		43	46

Note: Standard deviations in parentheses.

67 percent of first time migrants and 74 percent of last time migrants. On the other hand, 88 percent of first time migrants are undocumented, while only 46 percent of repeat migrants are undocumented, indicating that Mexicans obtain US residence permits over time.

Table 1 further shows that Mexicans make up about 5.5 percent of the population of the US locations in our sample. The highest concentration could be observed in Laredo, Texas, where 24.2 percent of the residents are of Mexican origin (Appendix B). Laredo has the highest unemployment rate in our sample (more than 16 percent), a very small local population and is very close to Mexico. Though the city is small and has a high unemployment rate, many appear to migrate there, because it is close to the border. The *Village Migration Experience* variable averages 1.9 percent. It reaches a maximum of 29.2 percent in Los Angeles, followed by Chicago with 9.2 percent (Appendix B). The migration flow appears to be about twice as large for first time migrants than for repeat migrants. Each of our locations has, on average, an unemployment rate of 7.1 percent, a population of 1.35 million, and is approximately 1,460 miles away from the sending village in Mexico.

Figs. 1 and 2 describe some typical patterns of our two stock variables. In Fig. 1 we plot the Herfindahl index of the concentration of the US migration experience of nine typical Mexican villages for the time period covered in our sample. The index is given by

$$HERF_{mt} = \sum_{j=1}^{J} \left(\frac{VME_{mjt}}{100} \right)^2 \tag{4}$$

with $0 \leq HERF_{mt} \leq 1$. Higher values of $HERF_{mt}$ indicate a higher concentration of the migration experience of a Mexican village. The villages differ in the concentration of their migration experience. Compared to the other villages depicted in Fig. 1, concentration is relatively low in communities 36 and 38 in the Mexican State S.L.P., community 46 in Zacatecas, and community 33 in Colima.[7] In most of the villages the concentration of the migration experience is increasing over time and flattens out at the end of the sample, indicating some kind of quadratic pattern, though most of the villages do not reach a turning point. Only in community 36 do we observe the concentration of the migration experience increasing at the very beginning of the sample period, reaching a maximum and then decreasing. In contrast to all other communities we observe a U-shaped pattern in community 52 in Oaxaca. Note that we find such a pattern only in two communities. In terms of US locations, Los Angeles County is the location with the highest average value of migration experience.

Fig. 2 shows the development of our second stock variable, the share of the Mexican population, in six US locations for the period covered in our sample. We display the Imperial Valley, Chicago, Houston, and Miami for their geographical dispersion and generic interest. In these five US locations the share of the Mexican population is increasing. The sixth US location is Laredo, Texas, which has the highest average share of Mexican population in our sample. In Laredo, the share of the Mexican population shows a U-shaped pattern over time; it decreases until 1982 and then increases.

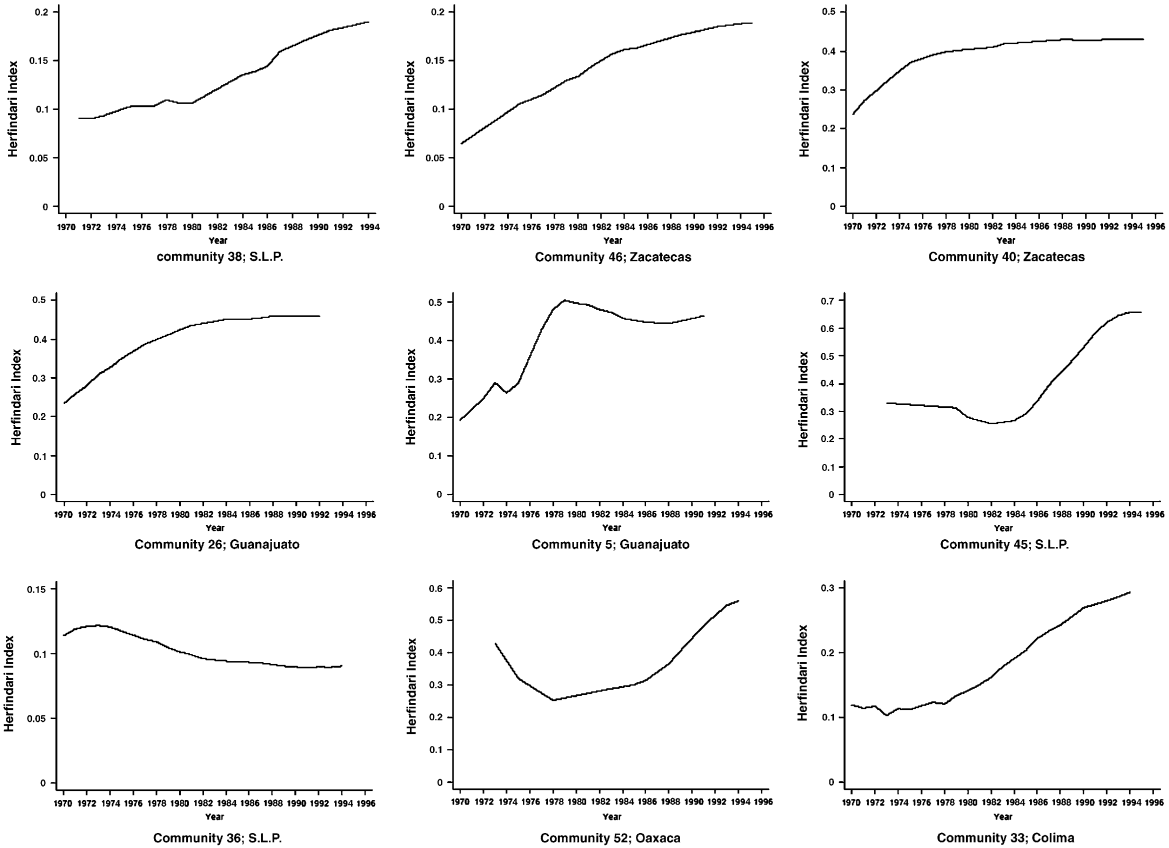


Fig. 1. Concentration of Mexican Migrants in the US.

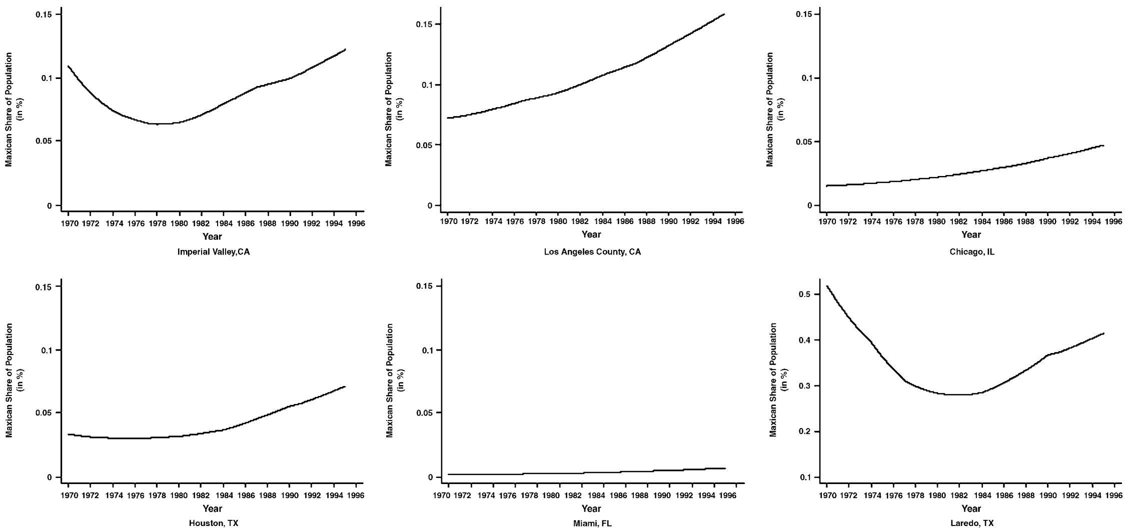


Fig. 2. Mexican Share of Population in Selected US Locations.

3. MULTIVARIATE ANALYSIS

3.1. Econometric Approach

To analyze the determinants of the location choice of Mexican migrants to the US, we estimate a conditional logit model (McFadden, 1984).[8] Each Mexican migrant *i,* who is assumed to maximize his utility, faces a choice among J alternative US communities. Assume that the utility of choosing location j is given by

$$U_{ij} = X_j\beta + \varepsilon_{ij} \tag{5}$$

where X_j is a vector of the characteristics of the US location *j,* including stock and flow effects, and ε_{ij} is an error term that is assumed to be independent and identically distributed. The probability that an individual i chooses location j is given by

$$\Pr(U_{ij} > U_{ik}) \quad \text{for all } k \neq j \tag{6}$$

Let Y_i be a random variable that takes the values 0 and 1, indicating the location choice made by the migrant. The probability that individual i chooses the US location j can then be written as

$$\Pr(Y_i = j) = \frac{\exp(X_j\beta)}{\sum_{j=1}^{J} \exp(X_j\beta)} \tag{7}$$

where X_j is a vector of characteristics of the US communities in our sample and β a parameter vector to be estimated. Eq. (7) can be estimated using maximum likelihood. Note that our sample is restricted to individuals who actually migrated at some point in time to the US. The analysis does not consider migration within Mexico.

As discussed in Section 2, our regressors include two measures of the effect of the stock of migrants, i.e., the Mexican share of the total population in US location j and the migration experience of a Mexican village m in the US location j, VME_{mjt}, a measure of the flow of migrants, $FLOW_{mjt}$, the population size, the consumer price index (cpi), and the unemployment rate in US location j, as well as the cost of migration proxied by the road mileage distance between Mexican village m and US location j. The existing theory (see, for example, Epstein, 2002) shows that we should expect the stock variables, *Village Migration Experience* and the *Mexican share of the population* to have an inverted U-shape relationship with respect to the probability of migrating to a certain location. As the stock of migrants in a location increases, the probability of a new migrant moving to that location increases at a decreasing rate, because positive network effects decrease and negative network externalities increase as the number of immigrants increase. Eventually, a turning point is reached after which a further increase in the stock of migrants will decrease the probability of a new migrant moving to that location. Hence, our specification of Eq. (7) includes both a linear and a squared term for the two stock variables. All other variables enter linearly.

We analyze the determinants of location choice both with and without US location fixed effects. Accounting for location fixed effects controls for the influence of time invariant heterogeneity. For example, one might argue that climate is an important determinant of location choice – especially in a study of the migration of persons from Mexico to US locations as climatically diverse as Laredo, TX, and New York City, NY.

In our empirical analysis we consider several specifications of Eq. (7). As individuals may have migrated more than once to the US, we divide our analysis into two parts: first and last migration. In the former we consider only the location decision made by the Mexican migrants at his/her first time migrating to the US while the latter consider only the location decisions made at his/her last time migrating to the US, conditional that he/she migrated to the US at least once before. For both specifications we estimate an overall (constrained) equation and an unconstrained equation. In the latter all variables considered in the basic specification are fully interacted with four dummy variables, one for unskilled illegal migrants, one for unskilled legal migrants, one for skilled illegal migrants, and one for skilled illegal migrants.

Estimation Results – Without US Location Fixed Effects

The second and seventh columns in Table 2 (first migration) and Table 3 (last migration) present the results for the constrained model (i.e., where we do not account for variation due to skill or legal status); columns 3–6 and 8–11 present the results for the unconstrained model; columns 2–6 do not include US location fixed effects, while columns 7–11 do. Consider the results for the constrained specification for the first migration decision. The Mexican share in the population of a US location appears to have an inverted U-shaped effect on the probability of choosing a particular location. Evaluated at the sample mean of a Mexican population share of 5.51 percent, the average marginal effect of an increase of the population share by

Table 2. Conditional Logit Analysis of Mexican Migrant's Location Choices: First Migration.

	Constrained	Unconstrained				Constrained	Unconstrained			
		Unskilled		Skilled			Unskilled		Skilled	
Variables		Illegal	Legal	Illegal	Legal		Illegal	Legal	Illegal	Legal
Unemployment rate	−0.033**	0.0004	0.091	−0.169**	0.001	−0.032	0.003	0.094	−0.157**	−0.016
	(0.015)	(0.012)	(0.066)	(0.036)	(0.061)	(0.023)	(0.026)	(0.070)	(0.039)	−0,066
CPI	0.002	0.010	0.021	0.024	−0.154**	0.029	0.043*	0.047	0.036	−0.129**
	(0.016)	(0.022)	(0.068)	(0.032)	(0.055)	(0.018)	(0.024)	(0.067)	(0.031)	(0.056)
Total population	0.012**	0.009**	0.021**	0.019**	0.013**	0.048**	0.041**	0.053**	0.046**	0.043**
	(0.001)	(0.001)	(0.005)	(0.003)	(0.005)	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)
Distance in miles (in 1,000)	0.044	0.044	0.196	0.362*	0.701*	−3.571**	−3.645**	−3.142**	−3.685**	−2.554**
	(0.085)	(0.103)	(0.416)	(0.187)	(0.381)	(0.623)	(0.628)	(0.772)	(0.650)	(0.749)
Share of Mexican population	0.154**	0.125**	0.209**	0.283**	0.266**	−0.071	−0.082*	−0.063	−0.056	−0.025
(in %)	(0.019)	(0.022)	(0.083)	(0.057)	(0.093)	(0.048)	(0.049)	(0.089)	(0.067)	(0.091)
Share of Mexican population	−0.008**	−0.006**	−0.011**	−0.017**	−0.014**	0.001	0.002	0.001	−0.001	−0.0001
(in %)Squared	(0.001)	(0.001)	(0.005)	(0.004)	(0.006)	(0.001)	(0.001)	(0.003)	(0.003)	(0.004)
Village migration experience	0.110**	0.113**	0.123**	0.100**	0.137**	0.092**	0.093**	0.098**	0.088**	0.120**
(in %)	(0.004)	(0.006)	(0.025)	(0.009)	(0.019)	(0.005)	(0.006)	(0.025)	(0.009)	(0.183)
Village migration experience	−0.087**	−0.093**	−0.127**	−0.071**	−0.127**	−0.067**	−0.070**	−0.099**	−0.058**	−0.108**
(in %)squared/ 100	(0.006)	(0.008)	(0.038)	(0.011)	(0.028)	(0.006)	(0.008)	(0.038)	(0.011)	(0.025)
Flow	0.233**	0.227**	0.154*	0.292**	0.234**	0.212**	0.201**	0.193**	0.251**	0.222**
	(0.025)	(0.031)	(0.089)	(0.056)	(0.117)	(0.025)	(0.023)	(0.094)	(0.053)	(0.112)
US location fixed effects	No	No				Yes	Yes			
Log-likelihood	−4025.0	−3975,40				−3743.70	−3706,70			
Pseudo-R^2	0.385	0,392				0.428	0,4333			

Note: Observations: 74,777. Standard errors in parentheses.

*Statistically significant at least at 10% level.

**Statistically significant at least at the 5% level.

Table 3. Conditional Logit Analysis of Mexican Migrant's Location Choices: Last Migration.

	Constrained	Unconstrained				Constrained	Unconstrained			
		Unskilled		Skilled			Unskilled		Skilled	
Variables		Illegal	Legal	Illegal	Legal		Illegal	Legal	Illegal	Legal
Unemployment rate	−0.031**	−0.001	−0.004	−0.200**	−0.092**	−0.055**	0.018	−0,045	−0.227**	−0.155**
	(0.015)	(0.021)	(0.027)	(0.064)	(0.045)	(0.027)	(0.032)	(0.038)	(0.071)	(0.055)
CPI	−0.038**	0.002	−0.063**	−0.035	−0.071**	−0.005	0.034	−0.041	−0.031	−0.072**
	(0.013)	(0.022)	(0.025)	(0.039)	(0.032)	(0.018)	(0.025)	(0.023)	(0.042)	(0.036)
Total population	0.00003	0.004*	−0.010**	0.023**	0.009**	0.033**	0.048**	0.030**	0.061**	0.047**
	(0.001)	(0.002)	(0.002)	(0.004)	(0.004)	(0.011)	(0.013)	(0.012)	(0.013)	(0.013)
Distance in miles (in 1,000)	0.471**	0.213	0.746**	−0.068	0.755**	−2.547**	−3.116**	−2.363**	−2.927**	−2.088**
	(0.101)	(0.140)	(0.206)	(0.341)	(0.296)	(0.725)	(0.737)	(0.762)	(0.811)	(0.802)
Share of Mexican population	0.201**	0.119**	0.268**	0.312**	0.382**	0.012	−0.075	0.013	0.043	0.103
(in %)	(0.022)	0.029	(0.044)	(0.096)	(0.076)	(0.057)	(0.063)	(0.067)	(0.114)	(0.094)
Share of Mexican population	−0.008**	−0.005**	−0.011**	−0.017**	−0.019**	0,0004	0.003*	−0.0001	−0.003	−0.004
(in %)squared	(0.001)	(0.001)	(0.002)	(0.007)	(0.005)	(0.001)	(0.001)	(0.002)	(0.006)	(0.005)
Village migration experience	0.149**	0.129**	0.208**	0.096**	0.133**	0.135**	0.112**	0.197**	0.088**	0.124**
(in %)	(0.006)	(0.009)	(0.010)	(0.015)	(0.014)	(0.006)	(0.009)	(0.011)	(0.015)	(0.014)
Village migration experience	−0.133**	−0.110**	−0.209**	−0.070**	−0.116**	−0.117**	−0.091**	−0.197**	−0.062**	−0.104**
(in %)squared/102	(0.008)	(0.012)	(0.015)	(0.020)	(0.018)	(0.008)	(0.011)	(0.015)	(0.019)	(0.018)
Flow	0.374**	0.357**	0.376**	0.451**	0.429**	0.383**	0.356**	0.438**	0.391**	0.414**
	(0.046)	(0.064)	(0.091)	(0.140)	(0.145)	(0.046)	(0.063)	(0.091)	(0.131)	(0.134)
US location fixed effects	No	No				Yes	Yes			
Log-likelihood	−3446.05	−3370,000				−3242.7	−3170,5			
Pseudo-R^2	0,423	0,436				0,457	0,47			

Note: Observations: 74,777. Standard errors in parentheses.

*Statistically significant at least at 10% level.

**Statistically significant at least at the 5% level.

one percent is 0.15.[9] Simulations we performed show the predicted effect of the share of Mexicans in the population of an average US location on the probability of choosing that location peaks at a population share of about 10 percent.[10]

Our other stock variable, the migration experience of a Mexican village, also follows an inverted U-shaped pattern, i.e., an increase in the share of a village's migration experience in a particular US location relative to its total Mexican migration experience increases the probability of choosing a particular US location at a decreasing rate. At a *Village Migration Experience* (VME) of approximately 63 percent the impact peaks, declining afterwards. While most cities in most times are on the uphill side of this turning point, we do observe four US locations where the value of VME_{mjt} exceeds 63 percent: Los Angeles County, Orange County and San Diego County in California and Chicago.[11] The coefficient on the variable capturing flow effects is significantly positive. The average marginal effect for this variable is calculated to be 0.0053, indicating that a 1 percent increase in the flow of migrants to a specific US location in the last year increases the probability that a migrant chooses this location on average by 0.53 percent.

For the constrained model without US location fixed effects, and for the four subgroups considered in the unconstrained model, the Mexican share in the population of a US location has an inverted U-shaped pattern. It appears that the Mexican stock in a US location is more important for unskilled as compared to skilled workers. Whereas the probability of choosing a US location peaks at a Mexican population share of approximately 10 percent for the latter, it reaches a maximum for skilled workers at a population share of 8 percent. Comparing legal and illegal migrants, however, no clear pattern emerges.

As in the constrained model, the estimated inverted U-shaped pattern for the village experience variable is much flatter than the respective pattern for the Mexican population share. However, in contrast to the Mexican population share, important differences between legal and illegal migrants appear. For illegal migrants, the effect of village migration increases the average probability of choosing a US location up to a share of 61 percent for unskilled, and 71 percent for skilled. For legal migrants this variable reaches its maximum effect at a share of 48 percent for unskilled and a share of 53 percent for skilled migrants.

The flow of migrants significantly affects all sub-groups considered. It further appears that there are no significant differences of the estimated flow effect between the different groups. Finally, the response of illegal migrants is more sensitive to changes in the migration flow before their migration decision as compared to legal migrants. However, as already noted above, these differences are not statistically significant.

Overall, these results indicate that legal and skilled migrants are less dependent on the stock of migrants when deciding on the location. The results further suggest that village-specific links, captured by the migration experience of a village, are on average relatively more important for the location choice of a migrant than ethnic goods, captured by the Mexican population share.

The estimation results for the last migration decision, not accounting for the possibility of US location fixed effects, are reported in Table 3. As for the first migration decision, both stock variables appear to have an inverted U-shaped pattern on the probability of choosing a US location and the pattern of the effect is much flatter for the village migration experience as compared to the share of the Mexican population in a US location. Comparing the different groups differentiated in the unconstrained model does not give a significantly different picture than the one obtained in Table 2. Comparing the first and last migration decision, however, it appears that both flow and stock effects are slightly more important for the last migration decision: the peaks are at a higher probability level and a higher share for the two stock variables. The effect of the flow variable on the probability of choosing a US location is steeper for the last as compared to the first migration decision.

Let us now consider what effects US location characteristics have on migrant location choice, still not accounting for the possibility of US location fixed effects. In the constrained model, the unemployment rate in a US location has a negative effect on the probability of choosing a location. However, only for the migrants' first trip is this effect statistically significant. In the unconstrained model, the effect of the unemployment rate on the location decision of a migrant is unclear for his/her first trip. According to the results reported in Table 2, the unemployment rate has a significant negative impact on the location decision of skilled illegal migrants and an unexpected significant positive impact on unskilled legal migrants. For the last trip of a migrant, the unemployment rate in the US location j affects only the location choices of skilled migrants on a statistically significant level; an increase in the unemployment rate in a US location decreases the probability that a skilled Mexican migrates there by 0.4 percent for illegal migrants and by 0.2 percent for legal migrants. Cost of living differences as captured by the consumer price index do not seem to drive location choice. Where the CPI is significant, it lowers the probability of moving to a location. We also examined a specification omitting the CPI, which left our other coefficient estimates essentially unchanged.

The probability that migrants choose a particular US location increases with the total population in that location for the first trip. For the last trip the total population has a positive effect on the location choice of unskilled illegal and skilled migrants, and a negative effect on unskilled legal migrants. This result reflects preferences for moving to regions with relatively large labor markets. The distance between the home community and the US location has a negative impact on illegal migrants and a positive impact on documented migrants on their first trip; the estimated coefficients are, however, not statistically significant at the 5 percent-level. For the last migration decision the distance to the US location shows an unexpected pattern. For the constrained model and for unskilled workers in the unconstrained model the coefficient of the distance variable is significantly positive indicating that a higher distance increases the probability of choosing a US location. It might be that this variable captures some other effects of characteristics of the US locations we did not control for in our specification.

Estimation Results – With US Location Fixed Effects[12]

A difficulty in our analysis is that there are probably unobserved region specific factors that determine migrants' location choice. For example, there may be variations among US locations with respect to resource endowments, cultural influences on legal and political arrangements, climate, and so on. To the extent that these factors are time invariant, we can control for these time invariant factors by including location specific fixed effects. Our results when we do this are seen in Tables 2 and 3, columns 7–11.

With a notable exception the inclusion of US location fixed effects does not change our results. Comparing the estimations for first migration in Table 2 without (columns 2–6) and with (columns 7–11) fixed effects, *Village Migration Experience* and our flow variable continue to be significant with approximately the same impact. The *Mexican Population Share*, however, is now not significant, small, and generally has a negative impact. Perhaps this share is slow to change and its effect is now absorbed by the location dummies. In Table 3, we observe the same phenomenon. On the bright side, the inclusion of US location fixed effects helps clear up some anomalies we noticed. In particular, distance is now negative and significant in its impact on location choice.

Our empirical results show that both of our stock measures and the flow of immigrants have significant effects on the migrant's decision about where to migrate. We should and cannot neglect these effects when analyzing location choice. These results confirm and extend other results on the importance of networks in location choice (for example, Jaeger, 2000; Winters et al., 2001). However, the choice of network variable can make a difference in ones conclusions. The Mexican share of the population becomes insignificant when US location fixed effects are included. Without accounting for these fixed effects, the Mexican share of the population is significant and portrays an inverse U shape. The village stock externality effect is larger and more robust, and also exhibits an inverse U shape, and not the simple positive linear effect as often presented in the literature.

4. A POSSIBLE EXPLANATION: HERD EFFECTS AND MIGRATION NETWORKS

One possible explanation for the results presented above is the relative importance and interaction of herd behavior (flows) and network externalities (stocks) in determining migration behavior. Both motivations give rise to immigrant clustering, a phenomenon observed in a wide variety of migration destinations. The theory we develop below builds on the work of Epstein (2002).

Let us first consider network externalities. Consider individual *j's* utility from migrating to a certain location, $U_j\,(\cdot)$. $U_j\,(\cdot)$ is a function of two variables: (i) the wage that the migrant will receive by migrating to the new location; and (ii) the stock of immigrants from the same origin who previously migrated to the new location, N. From the above discussion, the migrant's utility increases with the migrant's wage and increases with network externalities:

$$\frac{\partial U_j\left(w_j, N\right)}{\partial w_j} > 0 \quad \text{and} \quad \frac{\partial U_j\left(w_j, N\right)}{\partial N} > 0$$

Assume a normal downward-sloping demand function for workers in the host location, $q^d\left(w_f\right)$ and an upward-sloping supply function workers, $q^s(N_L, N)$. In equilibrium demand equals supply: $q^d\left(w_f\right) = q^s(N_L, N)$. In equilibrium wages are given by $w_f^*(N)$. Note that the equilibrium wage decreases as the stock of immigrants increases. The stock of migrants (the network effect) affects utility in two ways: directly via positive externalities and indirectly via negative externalities on the wages. The "old" migrants (the stock of immigrants) who are already in the host location prefer that the maximum number of migrants coming to this location will be such that their utility is maximized. That is, the marginal increase in the migrants' utility from externalities equals the marginal effect of the decrease in wages because of the additional migrant.

Denote by N_I the optimal stock of immigrants in the sense that this is the preferred stock of migrants who have previously migrated to this host location. Thus if the stock of immigrants exceeds N_I, further increasing the stock of immigrants raises non-wage network benefits, however wages also decrease. The effect of the increase in non-wage benefits is smaller than the effect of the decrease in wages and the utility of the immigrants who had previously migrated to this location decreases.

We may still observe migrants deciding to migrate to a location in which the stock of migrants has already exceeded N_I. Thus, the probability that an individual chooses to migrate to a location where the stock of immigrants already exceeds N_I is positive. This probability however, decreases as the stock of immigrants already in the host location increases. We conclude,

> Given network externalities, the probability an individual migrates to a certain location has an inverse U shape relationship with regard to the stock of immigrants already in the host location.

Now let us consider *herd behavior*. Following Epstein (2002) migration decisions are made sequentially, with people contemplating emigration at a given stage in their lives. Individuals respond to signals or information packets about host location possibilities. An individual receives a signal with probability p and with probability q this signal is true. The individual also observes the behavior of previous migrants. Potential migrants cannot, however, observe the information signal that was the basis for previous migrants' decisions. Given the information available, each individual chooses a location to which to migrate. The structure of the game and Bayesian rationality are common knowledge. Three assumptions govern individuals' actions: (a) An individual, who does not receive a signal and observes that everybody else has chosen to stay home, will also choose not to migrate. (b) An individual who is indifferent between following his or her own signal and copying someone else's choice will follow his or her own signal. (c) An individual who is indifferent between following more than one of the previous migrants' decisions will choose to randomize his or her decision with equal probabilities assigned to the different alternatives.

Under this framework it can be shown[13] that if an individual receives a signal to migrate to a specific place, he will follow this signal. If a second individual receives the same signal he will follow *individual 1* and if he does not receive a signal he will also follow *individual 1*, since the first migrant decided to migrate only because he had a signal and *individual 2* can see it as if he himself received this signal. Now consider the case where *individuals 1* and *2* migrated to one location, *location 1*, and *individual 3* receives a signal to immigrate to *location 2*. Since the first individual migrated it is clear he had a signal. Therefore, *individuals 1* and *3* had different signals. Since *individual 2* also migrated to the same location as *individual 1,* he may have received a signal to migrate to *location 1* or he may not have received a signal at all. Using a Bayesian Rule it can be shown that the probability of immigrating to *location 1* is higher than that of *location 2* and thus *individual 3* will follow the first two migrants instead of following the signal he received. On average there are *1+p* signals for *location 1* while there is only 1 signal for *location 2*. Epstein (2002) shows that

> As the number of individuals that have already migrated to some certain location increases, the probability of a new individual migrating to the same location increases. Individuals will migrate following the herd (flow) while disregarding their own private information.

As we can see from the empirical results, the theory of herd verses network externality may explain the behavior of the Mexican migrants in their location choices in the US.

5. CONCLUSION

Immigrant clustering is an important phenomenon to study for a number of reasons. The process by which immigrants decide where to locate is one that is not clearly understood, though there is much research on the subject. Standard economic theory argues that there are significant externalities, or "ethnic capital," of which immigrants wish to take advantage. They move to where members of their community, generally defined, had previously gone, planning to avail themselves of these externalities. In this paper we emphasize the different information content in different types of networks by examining two stock influences and a flow influence on the migration location decision.

Although previous studies have highlighted the role of networks on migration, no one has studied the potentially different impacts of migration stocks vs. migration flows. The paper argues that the relationship between the stock of migrants and the location choice of new migrants follows an inverted U, while because of herding the relationship between migration flows and location choice is positive.

We use data from the Mexican Migration Project to investigate the location decision of Mexican migrants in the US We distinguish between two types of stock effects, capturing general ethnic goods available in a US location (*Mexican share of the total population* in a particular US location) on the one hand, and origin village connections and the history of the migration experience of a village in different US locations (*Village Migration Experience*, the tendency of residents of a given Mexican village to migrate to a particular US location, measuring information available in the sending village of a given US location) on the other hand. These two variables help us to distinguish a generalized stock effect from village-specific links. The flow effect is measured using the flow of migrants to a particular US location during the year prior to the migration decision of an individual.

We show that both stock externalities and the flow have a significant effect on the migrant's location decision, though the "cultural goods" stock effect disappears once we control for location fixed effects. Moreover, the significance and size of the effects vary according to the legal status of the migrant and whether the migrant is a "new" or a "repeat" migrant. The estimated stock effects show an inverse U-shaped pattern, not a linear positive effect as often presented in the literature. The results indicate that village-specific links are relatively more important for the location decision of a migrant than the availability of ethnic goods. Furthermore, legal and skilled migrants appear to be less dependent on the migration stock than illegal and unskilled migrants. Flow effects have significant positive effects on the location decision of a migrant. Our estimations indicate, however, that there are no significant differences in these flows between different types of migrants.

Although a number of studies have underscored the importance of networks for location choices, the argument that immigrant clustering could be explained by herd behavior has been recently introduced to the migration literature (Epstein, 2002). Networks and herds reflect different types of information. Migrants might be motivated to choose a location to benefit from the network externalities it has to offer. However, because of herd effects, the migrant may choose a location on the supposition that recent migrants had information that he does not have. Migrants may choose to follow the flow and migrate to the location recent migrants have been observed to choose. Our empirical results indicate that network externalities and herd effects can both be present and influence emigration location decisions. The network externalities and herds' story is one of many interpretations. The fact is that both stocks and flows affect the decision of the migrant of where to go; this is the most important message of this paper.

NOTES

1. See Appendix A for a description of the data and its reliability.
2. We do not discount months over time, or for those who have returned to their village in Mexico. Although their knowledge of current labor market conditions may deteriorate, they provide key links and support for the network. The differential impact of more recent migrants is captured by our *flow* measure, while our *Mexican share of the total population* variable captures the generalized impact of having other Mexicans around in a US location.
3. We thank Julie Phillips for making this variable available to us.
4. Our empirical analysis treats the network for each Mexican community as exogenous to the individual migration decision. We feel this is the appropriate specification. However, one could argue that unobserved autocorrelated fluctuations in US local labor markets may draw migrants together to one location even if previously migrants from that community had gone elsewhere. Local unemployment rates in the US receiving communities also control for this.
5. In addition to road mileage, we also examined hours by car and the actual migration costs expressed by the migrant himself. All three cost variable yield similar results in our estimations.
6. Under the conditional logit formulation we apply, it is assumed the potential availability of other location choices will not affect the coefficient estimates.
7. The data set does not provide names for the Mexican villages.
8. Bartel (1989) and Jaeger (2000) also use this model to study the location choice of migrants in the United States.
9. The marginal effects of a change in the characteristics X_j of a US location j on the probability that a Mexican migrant will choose location j are given by the derivative of Eq. (15) with respect to the characteristics X_j, Note that these marginal effects will vary with the characteristics of a US location j, which leads to a very large number of marginal effects to interpret. Therefore, we follow the simplifying approach chosen by Jaeger (2000) and calculate average effects of a change in the characteristics X on $Pr(Y_i = j)$, i.e., $\partial Pr(Y_i = j)/\partial X_j = \left[(1/J)(1-(1/J))\right]\hat{\beta}$, where $J = 43$ for the first migration decision and $J = 47$ for the last migration decision. Hence, to obtain average marginal effects, the coefficients reported in Table 3 have to be multiplied by 0.0227 and those in Table 4 by 0.0208.
10. In particular, we calculated $Pr(Y_i = 1) = \exp(\beta' X_j)/1 + \exp(\beta' X_J)$ using sample means for X_j for all variables except the variable of interest and assuming that the location specific fixed effects are zero.
11. This only happened in certain years and does not show up in the Appendix tables.
12. We thank the referee for leading us to this analysis.
13. See Epstein (2002) for the formal proof and generalization.

ACKNOWLEDGMENTS

We thank Julie Phillips for discussions and helping us with the data. We thank David Jaeger, John Landon-Lane, Jackline Wahba, Yoshi Yonezawa and Myeong-Su Yun for their comments. This paper was partly written while Gil Epstein and Ira Gang were visiting IZA, Bonn. Ira Gang thanks the Rutgers Research Council for partial support.

REFERENCES

Bartel, A. P. (1989). Where do the new US immigrants live? *Journal of Labor Economics*, *7*(4), 371–391.

Bauer, T., Epstein, G. S., & Gang, I. N. (2005). Enclaves, language, and the location choice of migrants. *Journal of Population Economics*, *18*(4), 649–662.

Bustamante, J. A. (1998). Some thoughts on perceptions and policies. In: *Migration between Mexico and the United States: Mexico—US Binational Migration Study*. Mexican Ministry of Foreign Affairs and US Commission on Immigration Reform. www.utexas.edu/lbj/uscir/binational

Carrington, W. J., Detragiache, E., & Vishwanath, T. (1996). Migration with endogenous moving costs. *American Economic Review*, *86*(4), 909–930.

Chiswick, B. R., & Miller, P. W. (1996). Ethnic networks and language proficiency among immigrants. *Journal of Population Economics*, *9*(1), 19–35.

Church, J., & King, I. (1983). Bilingualism and network externalities. *Canadian Journal of Economics*, *26*(2), 337–345.

Epstein, G. S. (2002). *Informational cascades and decision to migrate*. IZA Working Paper 445.

Gottlieb, P. (1987). *Making their own way: Shorthorn blacks' migration to Pittsburgh, 1916–30*. Urbana: University of Illinois Press.

Grossman, J. R. (1989). *Land and hope: Chicago, black southerners, and the great migration*. Chicago: University of Chicago Press.

Jaeger, D. A. (2000). Local labor markets, admission categories, and immigrant location choice. http://dajaeg.people.wm.edu/research/ last accessed, May 22, 2006.

Marks, C. (1989). *Farewell – we're good and gone: The great black migration*. Bloomington: Indiana University Press.

Massey, D. S., Alarcón, R., Durand, J., & Gonzáles, H. (1987). *Return to Azlan: The social process of international migration from Western Mexico*. Berkley and Los Angeles: University of California Press.

McFadden, D. (1984). Econometric analysis of qualitative choice models. In: Z. Griliches & M. D. Intriligator (Eds), *Handbook of econometrics, Vol. II*. Amsterdam: North-Holland.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *Quarterly Journal of Economics*, *118*, 549–599.

Polachek, S. W., & Horvath, F. W. (1977). A life cycle approach to migration: Analysis of the perspicacious peregrinator. In: R. Ehrenberg (Ed.), *Research in labor economics* (Vol. 1, pp. 103–149).

Winters, P., de Janvry, A., & Sadoulet, E. (2001). Family and community networks in Mexico–US migration. *Journal of Human Resources*, *36*, 159–184.

Zahniser, S. S. (1999). *Mexican migration to the United States: The role of migration networks and human capital accumulation*. New York, N.Y.: Garland Publishing Inc.

Appendix A – Additional References

Donato, K. M., Durand, J., & Massey, D. S. (1992). Stemming the tide? Assessing the deterrent effects of the immigration reform and control act. *Demography*, *29*(2), 139–157.

Massey, D. S., Goldring, L., & Durand, J. (1994). Continuities in transnational migration: An analysis of nineteen Mexican communities. *American Journal of Sociology*, *99*, 1,492–1,533.

Massey, D. S., & Zenteno, R. (1999). A validation of the ethnosurvey: The case of Mexico–US migration, *International Migration Review*, 766–793.

APPENDIX A. DATA DESCRIPTION

Data basics: The Mexican Migration Project is an ongoing collaborative research project that was originally based at the University of Pennsylvania and the University of Guadalajara. The American base is now at the Office of Population Research, Princeton University. The data are available to users at http://mmp.opr.princeton.edu/databases/dataoverview-en.aspx. The Project combines techniques of ethnographic fieldwork and representative survey sampling in its data collection. Interviews are generally conducted in December–January when sojourner US migrants often return to Mexico, supplemented with surveys of out-migrants located in the United States.

Each year since 1987, two to five additional communities in these states are surveyed, selected based on their diversity in size, ethnic composition and economic development, not because they were known to contain return migrants. Each community is surveyed only once. 200 households in each community are interviewed, though in smaller communities fewer households are chosen. We use the MMP52 version of the data, as some of the complementary data we use is not available to us for the MMP71 or the MMP93. In particular, we cannot recreate Mexican share of the US population for later years as the MMP is unable to make the necessary coding available for translating outside data into geographical areas consistent with the MMP geographical areas. Massey, Alarcón, Durand, and Gonzáles (1987); Massey, Goldring, and Durand (1994); and Massey and Zenteno (1999) provide details and some data analysis. Massey and Zenteno (1999) show that the data are a source of reasonably representative retrospective data on documented and undocumented migration to the United States.

There are a few serious problems with the data. The interviews were free ranging, with the questioners following a semi-structured format. While the questioners tried to cover core questions, this process left many missing observations. Moreover, while the sample may be representative in a particular survey year, it will not be representative across time since it is retrospective and people are surveyed only once. To be included a migrant must have a link to a household in Mexico. It is impossible to know how important the "missing" information is for the analysis, but it may potentially severely bias the results. Also, as the data has been collected over a 20-year period there are issues with deflating wages, relative price changes, and the like.

We know if individuals ever migrated to the US, whether they were legal or not, how many times they worked in the US, the aggregate time spent in the US, when they made their first trip and when they made their last trip, how long was each of these trips, whether they were currently working in the US, their wages and occupations in the US, as well as information on the socioeconomic characteristics of the household members such as age, education, and marital status. The MMP also contains more detailed migration information on household heads that have migrated.

In constructing our village migration experience and flow variables we make use of the migration event history file of the data. This file provides detailed information on each migratory experience of all heads of household, including detailed information on the first and last trip to the US such as year and duration of the trip, the documentation used, the state and city of residence, performed occupation, and hourly wage, as well as some basic information on each border crossing. See Donato, Durand, and Massey (1992) for a more detailed description of the event-history file.

Mexican share of population: This variable has been obtained from the US Census Bureau for the censual years 1970, 1980, and 1990. A second-degree polynomial equation was estimated to these three data points to estimate the size of the Mexican foreign-population in each area during the inter-censual years. To estimate the Mexican foreign-born population in the years 1991–1995, it has been assumed that the annual growth rate during this period is the same as the annualized constant growth rate in each area between 1980 and 1990. The size of the Mexican foreign-born population is then divided by the *Total Population* in a US location. *Source*: We thank Julie A. Phillips for making this variable available to us.

Village migration experience and flow: These variables were calculated as indicated in the text from the event history file. *Source: MMP 52.*

Unemployment rate: The most recent information on the number unemployed and the size of the civilian labor force at the county level was obtained for the years 1974 and 1976–1996 from the Bureau of Labor Statistics, Local Area Unemployment Statistics Division. For the early 1970s, no information by county is available although information on unemployment for the censual years 1960 and 1970 is available. For the years 1971–1973, the assumption was made that unemployment rates in a county follow the same trends as that of the state. An estimate of the unemployment rate for 1975 was obtained by averaging the unemployment rates for 1974 and 1976. *Source: MMP 52.*

Total population: Data were obtained from Census publications, e.g., the CPS and County and City Yearbook, for the following years: 1970, 1974, 1976, 1977, 1980, 1984, 1986, 1987, 1990, and 1991. The population for the intercensual years was estimated by assuming an exponential growth function. To estimate the population between 1992–1995, the constant growth rate that prevailed between 1980 and 1991 was applied. *Source: MMP 52*.

Migration costs: We collected data on three measures of migration costs. For *Miles* and *Hours* we entered in the main town in the Mexican state in which the origin village is located and the main town in the US location into *Mapquest* (http://www.mapquest.com). For *Actual Costs* the data come from the *MMP 52*. Since the actual cost data was very sketchy, we decided not to use it. Trials with the *Hours* and the *Actual Costs* data yielded similar results to those when we used *Miles*.

Skilled vs. unskilled, legal vs. illegal: All migrants with less than 7 years of schooling are considered to be unskilled; those with more than 6 years of schooling are considered to be skilled. Undocumented migrants are labeled illegal, documented migrants *legal. Source: Mexican Migration Project 52*.

US communities: Imperial Valley, CA; Lower San Joaquin, CA; Middle San Joaquin, CA; Upper San Joaquin, CA; Salinas-Monterey-Santa Cruz, CA; Sacramento Valley, CA; Ventura-Oxnard-Simi, CA; Santa Barbara, CA; Napa-Sonoma, CA; Los Angeles County, CA; Orange County, CA; San Francisco Urban Area, CA; San Jose Urban Area, CA; Riverside-San Bernardino, CA; San Diego County, CA; Rio Vista, CA; Abilene, TX; Austin, TX; Beaumont-Port Arthur, TX; Brownsville, TX; Bryan-College, TX; Corpus Christi, TX; Dallas-Ft.Worth, TX; El Paso, TX; Galveston, TX; Houston, TX; Laredo, TX; McAllen, TX; Odessa-Midland, TX; San Antonio, TX; Victoria, TX; Chicago, IL; Las Cruces, NM; Tucson, AZ; Phoenix, AZ; Denver-Boulder, CO; Reno, NV; Las Vegas, NV; Omaha, NE; New York City, NY; Washington, DC, WA; Miami, FL; Atlanta, GA.

APPENDIX B. DESCRIPTIVE STATISTICS BY US RECEIVING COUNTY: FIRST MIGRATION

	Unemployment Rate	Total Population (in 100,000)	Miles	Mexican Share of Population (in %)	Village Migration Experience (in %)
Imperial Valley, CA	9.767	8.407	1828.160	8.137	1.703
	(2.491)	(2.603)	(160.834)	(1.533)	(5.863)
Lower San Joaquin, CA	9.573	4.278	1828.160	6.201	0.532
	(2.267)	(0.812)	(160.834)	(1.059)	(1.421)
Middle San Joaquin, CA	10.217	5.328	1828.160	7.957	2.496
	(2.192)	(0.857)	(160.834)	(1.773)	(3.548)
Upper San Joaquin, CA	11.752	7.940	1828.160	5.687	3.731
	(2.183)	(1.558)	(160.834)	(1.148)	(8.151)
Salinas-Monterey-Santa Cruz, CA	7.538	10.260	1996.099	7.474	2.934
	(2.031)	(1.676)	(160.467)	(0.513)	(4.492)
Sacramento Valley, CA	7.632	16.661	1996.099	3.134	2.377
	(2.162)	(3.029)	(160.467)	(0.534)	(3.661)
Ventura-Oxnard-Simi, CA	7.213	5.356	1608.970	7.465	1.875
	(1.314)	(0.957)	(160.619)	(0.858)	(3.747)
Santa Barbara, CA	6.056	3.121	1608.970	5.937	1.320
	(0.944)	(0.352)	(160.619)	(2.144)	(2.899)
Napa-Sonoma, CA	6.392	4.005	1996.099	2.256	0.893
	(1.899)	(0.676)	(160.467)	(1.050)	(2.686)
Los Angeles County, CA	6.866	77.237	1608.970	10.041	29.241
	(1.283)	(7.279)	(160.619)	(2.079)	(24.917)
Orange County, CA	4.856	19.621	1608.970	5.638	4.932
	(1.127)	(2.995)	(160.619)	(2.391)	(9.727)
San Francisco Urban Area, CA	5.586	33.490	1996.099	2.099	1.206
	(1.710)	(2.163)	(160.467)	(0.465)	(3.185)
San Jose Urban Area, CA	5.646	13.021	1996.099	3.713	2.595
	(0.966)	(1.345)	(160.467)	(0.921)	(6.343)
Riverside-San Bernardino, CA	7.241	17.184	1608.970	4.952	0.856
	(2.315)	(5.212)	(160.619)	(1.701)	(1.689)
San Diego County, CA	6.533	19.309	1608.970	5.490	5.184
	(1.488)	(3.758)	(160.619)	(1.374)	(12.782)
Rio Vista, CA	7.515	2.478	1996.099	1.737	0.067
	(1.397)	(0.594)	(160.467)	(0.245)	(0.314)
Abilene, TX	4.694	1.125	940.678	1.351	0.181
	(1.995)	(0.089)	(149.496)	(0.458)	(0.882)
Austin, TX	4.307	6.984	940.678	1.674	0.209
	(1.418)	(1.505)	(149.496)	(0.575)	(0.989)
Beaumont-Port Arthur, TX	8.069	3.671	940.678	0.561	0.091
	(3.559)	(0.136)	(149.496)	(0.102)	(0.581)
Brownsville, TX	11.093	2.125	621.961	21.255	1.313
	(2.788)	(0.440)	(134.766)	(5.922)	(2.760)
Bryan-College, TX	3.807	0.957	940.678	1.421	0.026
	(1.155)	(0.236)	(149.496)	(0.482)	(0.141)
Corpus Christi, TX	7.032	3.290	621.961	4.124	0.380
	(2.548)	(0.288)	(134.766)	(2.181)	(1.852)

APPENDIX B. (*Continued*)

	Unemployment Rate	Total Population (in 100,000)	Miles	Mexican Share of Population (in %)	Village Migration Experience (in %)
Dallas-Ft.Worth, TX	4.366	31.542	940.678	2.216	2.705
	(1.370)	(5.828)	(149.496)	(1.028)	(6.872)
El Paso, TX	9.263	4.906	1036.457	20.713	0.074
	(2.087)	(0.794)	(154.082)	(3.998)	(0.224)
Galveston, TX	6.904	1.997	940.678	1.835	0.128
	(3.031)	(0.165)	(149.496)	(0.668)	(1.079)
Houston, TX	5.317	30.412	940.678	3.791	3.782
	(2.430)	(5.510)	(149.496)	(0.969)	(8.975)
Laredo, TX	16.013	1.448	621.961	24.189	0.037
	(4.430)	(0.330)	(134.766)	(6.569)	(0.234)
McAllen, TX	14.252	2.946	621.961	22.815	1.030
	(4.765)	(0.743)	(134.766)	(6.316)	(2.370)
Odessa-Midland, TX	5.230	2.052	1036.457	3.399	0.125
	(2.748)	(0.340)	(154.082)	(0.319)	(0.669)
San Antonio, TX	5.668	11.520	940.678	6.203	1.369
	(1.621)	(1.486)	(149.496)	(2.580)	(3.491)
Victoria, TX	5.259	1.722	940.678	1.458	0.300
	(1.577)	(0.148)	(149.496)	(0.623)	(1.123)
Chicago, IL	6.398	73.705	2033.580	2.461	9.197
	(1.896)	(1.345)	(149.848)	(0.766)	(17.617)
Las Cruces, NM	7.715	1.029	1298.042	11.449	0.080
	(1.030)	(0.237)	(152.066)	(2.955)	(0.315)
Tucson, AZ	5.462	5.320	1238.160	4.378	0.153
	(1.437)	(1.003)	(160.834)	(1.437)	(0.816)
Phoenix, AZ	5.401	15.676	1238.160	2.592	0.762
	(1.368)	(3.832)	(160.834)	(0.836)	(2.612)
Denver-Boulder, CO	5.240	7.615	1605.164	1.328	0.240
	(1.468)	(2.462)	(144.481)	(0.438)	(0.616)
Reno, NV	5.629	1.923	1524.070	1.550	0.263
	(1.230)	(0.444)	(160.925)	(1.557)	(1.770)
Las Vegas, NV	6.909	5.533	1524.070	1.421	0.365
	(1.672)	(1.822)	(160.925)	(0.606)	(1.173)
Omaha, NE	4.398	1.766	1687.938	6.046	0.095
	(1.133)	(1.787)	(149.981)	(7.250)	(0.347)
New York City, NY	7.246	73.383	2596.999	0.205	0.375
	(2.371)	(2.263)	(129.604)	(0.188)	(1.413)
Washington D.C., WA	7.344	6.581	2386.269	0.085	0.059
	(1.922)	(0.489)	(132.258)	(0.033)	(0.296)
Miami, FL	6.954	16.210	1926.681	0.324	0.066
	(1.849)	(2.128)	(132.039)	(0.121)	(0.293)
Atlanta, GA	5.073	10.963	1749.061	0.208	0.042
	(1.586)	(0.681)	(132.803)	(0.282)	(0.315)
Total	7.103	13.351	1459.956	5.511	1.986
	(3.309)	(18.867)	(527.774)	(6.476)	(7.622)

Observations per US county: 1,739; Total observations: 74,777.

APPENDIX C. DESCRIPTIVE STATISTICS BY US RECEIVING COUNTY: LAST MIGRATION

	Unemployment Rate	Total Population (in 100,000)	Miles	Mexican Share of Population (in %)	Village Migration Experience (in %)
Imperial Valley, CA	10.019	10.887	1805.511	9.201	1.463
	(2.265)	(3.172)	(128.499)	(1.665)	(5.367)
Lower San Joaquin, CA	10.889	4.991	1805.511	7.142	0.625
	(2.350)	(0.870)	(128.499)	(1.391)	(1.179)
Middle San Joaquin, CA	11.381	6.092	1805.511	9.139	2.618
	(2.165)	(0.940)	(128.499)	(2.066)	(3.107)
Upper San Joaquin, CA	12.188	9.324	1805.511	6.776	4.410
	(2.032)	(1.689)	(128.499)	(1.479)	(10.385)
Salinas-Monterrey-Santa Cruz, CA	7.774	11.598	1973.522	7.932	2.798
	(1.684)	(1.591)	(128.294)	(0.689)	(4.014)
Sacramento Valley, CA	7.802	19.363	1973.522	3.429	2.095
	(1.773)	(3.299)	(128.294)	(0.575)	(2.715)
Ventura-Oxnard-Simi, CA	6.951	6.126	1586.325	8.177	2.670
	(1.375)	(0.914)	(128.400)	(1.111)	(5.449)
Santa Barbara, CA	5.831	3.436	1586.325	7.992	1.890
	(1.094)	(0.386)	(128.400)	(2.866)	(3.975)
Napa-Sonoma, CA	5.901	4.563	1973.522	3.264	1.463
	(1.557)	(0.668)	(128.294)	(1.340)	(4.713)
Los Angeles County, CA	6.949	83.756	1586.325	11.916	30.545
	(1.596)	(7.899)	(128.400)	(2.315)	(24.412)
Orange County, CA	4.602	22.118	1586.325	7.887	4.562
	(1.251)	(2.998)	(128.400)	(2.882)	(8.758)
San Francisco Urban Area, CA	5.233	35.446	1973.522	2.539	1.175
	(1.348)	(2.393)	(128.294)	(0.623)	(2.580)
San Jose Urban Area, CA	5.300	14.100	1973.522	4.421	2.213
	(1.077)	(1.289)	(128.294)	(1.173)	(6.033)
Riverside-San Bernardino, CA	7.476	22.107	1586.325	6.448	0.920
	(2.032)	(6.221)	(128.400)	(2.143)	(1.453)
San Diego County, CA	5.977	22.546	1586.325	6.794	5.801
	(1.454)	(3.890)	(128.400)	(1.690)	(15.218)
Rio Vista, CA	7.018	3.014	1973.522	1.930	0.087
	(1.393)	(0.657)	(128.294)	(0.299)	(0.328)
Abilene, TX	5.543	1.174	918.260	1.731	0.252
	(1.642)	(0.068)	(133.864)	(0.635)	(1.200)
Amarillo, TX	4.847	1.847	1157.821	1.616	0.107
	(1.194)	(0.146)	(134.139)	(0.886)	(0.521)
Austin, TX	4.752	8.237	918.260	2.098	0.177
	(1.248)	(1.510)	(133.864)	(0.721)	(1.066)
Beaumont-Port Arthur, TX	9.060	3.682	918.260	0.588	0.097
	(2.799)	(0.100)	(133.864)	(0.093)	(0.601)
Brownsville, TX	12.328	2.451	604.505	20.797	0.794
	(2.155)	(0.390)	(126.000)	(4.099)	(1.978)

APPENDIX C. (*Continued*)

	Unemployment Rate	Total Population (in 100,000)	Miles	Mexican Share of Population (in %)	Village Migration Experience (in %)
Bryan-College, TX	4.071	1.121	918.260	1.817	0.020
	(1.094)	(0.201)	(133.864)	(0.650)	(0.082)
Corpus Christi, TX	8.169	3.465	604.505	3.613	0.223
	(2.152)	(0.229)	(126.000)	(1.414)	(1.188)
Dallas-Ft.Worth, TX	5.114	36.443	918.260	3.214	3.142
	(1.164)	(5.902)	(133.864)	(1.324)	(8.955)
El Paso, TX	10.156	5.551	1012.657	20.871	0.059
	(1.460)	(0.772)	(133.075)	(2.936)	(0.153)
Galveston, TX	7.803	2.117	918.260	1.705	0.160
	(2.234)	(0.150)	(133.864)	(0.430)	(1.345)
Houston, TX	6.120	34.397	918.260	4.741	3.309
	(1.939)	(4.840)	(133.864)	(1.258)	(8.157)
Laredo, TX	16.912	1.721	604.505	24.742	0.024
	(3.217)	(0.331)	(126.000)	(4.995)	(0.138)
McAllen, TX	17.086	3.533	604.505	22.232	0.579
	(3.884)	(0.701)	(126.000)	(4.306)	(1.445)
Odessa-Midland, TX	6.289	2.229	1012.657	3.531	0.048
	(2.363)	(0.259)	(133.075)	(0.360)	(0.332)
San Angelo, TX	5.015	1.395	1012.657	4.359	0.111
	(1.137)	(0.105)	(133.075)	(0.760)	(0.472)
San Antonio, TX	6.147	12.716	918.260	5.716	0.815
	(1.364)	(1.427)	(133.864)	(1.677)	(2.190)
Victoria, TX	5.897	1.831	918.260	2.055	0.252
	(1.338)	(0.131)	(133.864)	(0.875)	(1.174)
Chicago, IL	6.521	74.754	2011.131	3.163	8.350
	(1.456)	(1.295)	(134.142)	(0.873)	(16.917)
Tucson, AZ	5.030	6.102	1215.511	4.733	0.111
	(1.246)	(0.934)	(128.499)	(1.250)	(0.551)
Phoenix, AZ	5.084	18.915	1215.511	3.198	0.402
	(1.110)	(3.869)	(128.499)	(1.038)	(1.352)
Denver-Boulder, CO	5.456	8.437	1583.517	1.680	0.211
	(1.098)	(3.109)	(127.860)	(0.695)	(0.422)
Pueblo, CO	8.673	1.244	1583.517	0.857	0.158
	(2.370)	(0.017)	(127.860)	(0.410)	(1.058)
Reno, NV	5.403	2.289	1501.444	3.058	0.368
	(1.006)	(0.438)	(128.591)	(2.160)	(2.306)
Las Vegas, NV	6.426	7.201	1501.444	1.976	0.251
	(1.501)	(2.100)	(128.591)	(0.718)	(0.934)
St. Louis, MO	5.614	14.196	1549.915	0.069	0.063
	(1.324)	(0.434)	(134.021)	(0.017)	(0.210)
Omaha, NE	3.992	1.551	1665.505	5.543	0.157
	(1.168)	(1.679)	(134.353)	(6.134)	(0.460)
New York City, NY	7.469	73.164	2578.455	0.390	0.265
	(2.047)	(1.464)	(120.590)	(0.266)	(1.148)

APPENDIX C. (*Continued*)

	Unemployment Rate	Total Population (in 100,000)	Miles	Mexican Share of Population (in %)	Village Migration Experience (in %)
Washington D.C., WA	7.281	6.264	2367.655	0.114	0.053
	(1.733)	(0.393)	(123.929)	(0.034)	(0.256)
Miami, FL	7.307	17.999	1908.071	0.436	0.056
	(1.630)	(2.156)	(123.611)	(0.142)	(0.194)
Atlanta, GA	5.388	11.564	1730.293	0.484	0.068
	(1.075)	(0.725)	(124.443)	(0.460)	0.404)
Total	7.310	14.066	1431.984	5.568	1.870
	(3.413)	(19.216)	(510.941)	(6.163)	(7.563)

Observations per US county: 1,561; Total observations: 71,806.

WAGE AND JOB DYNAMICS AFTER WELFARE REFORM: THE IMPORTANCE OF JOB SKILLS☆

Rucker C. Johnson

ABSTRACT

I use data from employers and longitudinal data from former/current recipients covering the period 1997 to early 2004 to analyze the relationship between job skills, job changes, and the evolution of wages. I analyze the effects of job skill requirements on starting wages, on-the-job training opportunities, wage growth prospects, and job turnover. The results show that jobs of different skill requirements differ in their prospects for earnings growth, independent of the workers who fill these jobs. Furthermore, these differences in wage growth opportunities across jobs are important determinants of workers' quit propensities (explicitly controlling for unobserved worker heterogeneity). The determinants and consequences of job dynamics are investigated. The results using a multiplicity of methods, including the

☆This research was supported in part by grants from the Charles Stewart Mott Foundation, Ford Foundation, Joyce Foundation, the National Institute of Mental Health (R24-M51363), and the Office of the Vice-President for Research at the University of Michigan.

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 231–298

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26007-2

estimation of a multinomial endogenous switching model of wage growth, show that job changes, continuity of work involvement, and the use of cognitive skills are all critical components of the content of work experience that leads to upward mobility. The results underscore the sensitivity of recipients' job transition patterns to changes in labor market demand conditions.

1. INTRODUCTION

Much of the current welfare reform debate centers around opposing views regarding the job and wage dynamics, and potential for wage growth, for former/current welfare recipients. There is consensus that initial wages are likely to be low for low-skilled workers. Some analysts think that low-wage jobs represent a port of entry into higher-paying jobs, whereas others are concerned that entry-level jobs simply represent the first in a succession of "dead-end" jobs (Connolly & Gottschalk, 2000; Edin & Lein, 1997).

Few studies analyze whether jobs differ in their prospects for earnings growth (independent of the worker who fills the job), and the existing evidence lacks a consensus. A further issue that remains elusive is whether serial correlation in wage increases is attached to jobs or to workers. It is difficult to sort out, for example, whether persistently low wages are a greater reflection of a lack of on-the-job training and other human capital investment opportunities, as opposed to the worker's learning and earnings ability. Two prominent studies (Topel, 1991; Topel & Ward, 1992), based on the time series properties of within-job wage changes of men, conclude that heterogeneity in permanent rates of wage growth among jobs is empirically unimportant. Their direct evidence seems to show that jobs do not in fact differ in their prospects for wage growth. However, it remains unclear whether these models and empirical estimates apply to less-skilled workers.

There is scant empirical evidence concerning the job and wage dynamics that accompany initial employment at low wages. Analyses that have focused on the wage growth of less-skilled workers have not distinguished between within-job wage growth and between-job wage growth. Understanding the mechanics of wage growth for less-skilled workers and assessing the relative contributions of different sources of wage growth (returns to general work experience, job tenure, and improvements in job matches) are as important as the estimates of the overall rate of wage growth. Labor market "success" is usually reduced to a single indicator measured at a point in time, such as whether employed, wage rate, or earnings. Employment activities within the firm, such as job skills used, on-the-job training, promotion activity, and the consequences of training and promotion, are typically unmeasured. This paper makes strides to bridge this gap by analyzing employment experiences of representative samples of former/current welfare recipients using both individual-level and employer survey data.

This paper addresses the following set of research questions. Do jobs of differing skill requirements exhibit differential wage growth opportunities independent of the workers who fill these jobs? What is the skill content of work experience that leads to upward mobility? How do those characteristics contrast with those prevalent in dead-end jobs? Are differences in wage growth opportunities across jobs (independent of wage levels) an important determinant of workers' quit propensities? Do jobs (as opposed to workers in them) have different turnover behavior? How much of wage growth depends on job transitions, and how much is accounted for by the accumulation of tenure and experience?

The study of these questions is relevant to our understanding of less-skilled labor markets and may help inform the development of policy initiatives designed to facilitate the transition of disadvantaged workers into steady-living wage jobs. The importance of analyzing the returns from holding a steady job versus the return from switching jobs, as well as how these returns may depend on the skill requirements of the job, is evidenced by two contrasting views of the effects that turnover has for workers. One view is based on the belief that the labor market experiences of low-skilled workers are often characterized by cycling through a series of low wage, unstable, dead-end jobs. Proponents of this view argue that this results in a waste of human capital because the job instability prevents workers from developing skills or behaviors that might lead to higher-paying jobs. An alternative view posits that through the job search process workers gain knowledge about their aptitudes, skills, and interests that lead to better job matches as they move from job to job and up the job ladder. This view is supported by several studies that show that, on average, job mobility accounts for the dominant share of wage growth among young men (Topel & Ward, 1992). The findings of this paper reveal that the skill content of work experience is a critical determinant of which one of these viewpoints becomes a reality for former welfare recipients.

I analyze unique longitudinal individual-level and firm-level survey data over a seven-year period (1997 to early 2004) to provide a complimentary evidence from both the supply and demand side. Both data sets were dministered after the implementation of welfare reform in Michigan, and the same set of detailed questions about job tasks/work skills were asked in each survey.

A primary goal of this paper is to investigate the effects of skill requirements of jobs on starting wages, on-the-job training opportunities, wage growth prospects (likelihood of within-job pay increases and promotion within the firm, and voluntary inter-firm job mobility), and job turnover. There are two key features of my empirical analysis that differentiate it from earlier studies and allow for the possibility of new insight. First, because jobs differ in the learning opportunities they provide, I explore how differences in these opportunities generate heterogeneity of wage-growth rates among jobs that have different job skill requirements. I provide evidence of heterogeneity across workers and jobs in the experience-earnings profile – its steepness (in return to experience) and its discontinuities (due to wage changes associated with job change) – and document systematic differences in expected wage changes with job mobility that depend on reason for and type of job change. Second, the interrelationship between wage growth prospects and job turnover behavior will be examined using both the employer survey and longitudinal individual-level survey data. I will investigate how wage growth and the types of jobs held (job skill requirements) are associated with job turnover. The analysis contributes to our understanding of the nature of the job mobility and wage growth process for less-skilled workers, and highlights the importance of jointly considering both processes. The analysis also underscores the sensitivity of former/current welfare recipients' job transition patterns to changes in local labor market demand conditions.

This paper consists of four parts. In the next section, I briefly review related research on wage growth and job turnover. Section 3 describes the data sets and the definitions of the key variables. Section 4 discusses the estimation strategy, model specification, and central results. In the final section, I summarize the findings and discuss their policy significance.

2. RELATED STUDIES

The rapid development and diffusion of new technologies in the workplace over the past several decades, coupled with globalization, has led to growing concerns that these innovations have displaced less-skilled jobs that were once a good source of career earnings paths and replaced them with dead-end, high-turnover service and retail jobs.[1] Recent research has documented the growing importance of cognitive skills in wage determination, for all workers, including less-educated workers (Murnane, Levy, & Willett, 1995; Jencks & Phillips, 1998; Tyler, Murnane, & Willett, 1999). However, the explanation of increasing returns to dimensions of skill not proxied by educational attainment has not resolved the puzzle as to which particular job skills have become relatively more valued in the labor market (Krueger, 1993; DiNardo & Pischke, 1997). Most analyses of earnings have relied on survey data that have limited information on the characteristics of the jobs individuals hold. Because little attention has been given to the skills required, we currently have little systematic knowledge of the evolution of job assignments and resulting effects on wages, particularly in less-skilled labor markets.

Studies of women who have left AFDC find low-paying jobs to be the norm, and there is little wage growth in the first several years after leaving welfare (Harris, 1996; Riccio, Fredlander, & Freedman, 1994; Pavetti, Holcomb, & Duke, 1995; Cancian, Haveman, Meyer, & Wolfe, 2000). Burtless (1995), using NLSY data, showed that women with low levels of schooling and low AFQT scores had lower rates of wage growth with age than did other women and conjectured that these low rates of wage growth reflect recipients' low skill levels.

Loeb and Corcoran (2001), on the other hand, claim that AFDC recipients have low rates of wage growth with age because they work fewer years and are more likely to work part-time than are nonrecipients. They report that wage growth per years actually worked is similar for AFDC recipients and nonrecipients (roughly 6% for every year of full time work), and that wage growth is slow when individuals work part-time. Gladden and Taber (2000) find no significant differences in wage growth with experience by educational attainment.

Neither Loeb and Corcoran (2001) nor Gladden and Taber (2000), however, consider dimensions of skill not proxied by educational attainment and experience. Their estimates include both individuals in jobs that require only soft skills who may gain little from work experience, and those in jobs requiring hard skills (e.g., reading, writing, math, or computer skills) who may experience significant gains from work experience.

The wage premium associated with particular job skills reflects a combination of the cost of acquisition, quasi-rent due to rising demand, and the extent to which the skill can be signaled to the external labor market (Green, 1998). The premium arises because workers can credibly threaten to quit for higher wages elsewhere. Krueger (1993) documented that computer users earn higher pay than nonusers. It remains unclear, however, to what extent their higher wages are due to computing skills, or whether people with higher abilities are selected to use computers and would have received higher pay even in the absence of computer usage (DiNardo & Pischke, 1997). Computer skills have received the bulk of the attention in the literature on U.S. wage inequality. Apart from computer skills, there has been little analysis of the link between other job skills (such as reading/writing and math) and the wage growth process and job dynamics for less-skilled workers. In this paper, I will analyze these relationships.

The extant evidence on whether jobs differ in their prospects for earnings growth (independent of the worker who fills the job) is limited. Topel (1991) and Topel and Ward (1992) analyze the time-series properties of within-job wage changes of men and conclude that heterogeneity in permanent rates of growth among jobs is empirically unimportant. Their results, however, are based on weak tests that fail to reject the hypothesis that within-job wages evolve as a random walk. An important implication of the result for job turnover, if indeed true, is that the current wages, along with experience and seniority, are sufficient statistics for future wages and the value of the job. Thus, this would predict that job separations should decline as a function of the wage level and not as a function of wage growth. However, Topel and Ward's (1992) own job turnover analysis contradicts this prediction and reveals that jobs offering higher wage growth are significantly less likely to end in worker-firm separations than jobs offering lower wage growth. This finding not only implies that the source of wage growth must have a firm-specific component, but it also implies heterogeneity of wage growth among jobs.

Other work analyzing serial correlation in wage increases (Abowd & Card, 1989; Baker, 1997) have yielded mixed results, but the most recent of these studies conducted by Baker (1997) provides a strong evidence in support of the wage profile heterogeneity model. To tackle the related issue of whether serial correlation in wage increases is attached to jobs or to workers, the approach taken in this paper (using longitudinal data of a sample of former/current welfare recipients) estimates the effects of job skills and explicitly controls for unobserved worker heterogeneity by contrasting recipients' wage growth and turnover rates in jobs held of differing skill requirements.

In human capital and job matching models, wage growth over a career reflects accumulation of experience, growth in seniority within a given firm, and movement toward better job matches (Altonji & Shakotko, 1987). The returns to job tenure (relative to job mobility) is an increasing function of the accumulation of job/firm-specific skills (i.e., skills acquired that are valued within the firm, but less easily transferable to other jobs/employers) and the quality of the job match. The proportion of on-the-job training opportunities that are job/firm-specific rises with the skill-level/education requirements of the job (Simpson, 1992). As a result, the human capital model predicts job changes to be a more important source of wage growth for less-skilled workers.[2]

Compared to the voluminous empirical literature on wage growth via human capital accumulation, much less work has been done on wage growth via job changes.[3] Altonji and Williams (1997) after surveying alternative estimates of wage growth reach a consensus estimate of on-the-job wage growth of 1.1% per year. Moreover, this is likely an upper bound since the Altonji–Williams estimate is based on the worker being continuously employed for 10 years. The on-the-job wage growth component appears to account for a small fraction of overall wage growth, which suggests that job mobility may be the most important component in earnings growth.

Topel and Ward (1992) and Loprest (1992) highlight the importance of job mobility (that is, job-to-job transitions) to early career wage growth, estimating that job changes account for roughly one-third of total wage growth during the first 10 years in the market. These studies, however, are based on samples of better-educated workers. Studies that have focused on the wage growth of less-skilled workers have not distinguished between within-job wage growth and between-job wage growth.[4] One exception is Connolly and Gottschalk (2000) who find that high school dropouts experience both lower wage growth within-jobs and lower wage growth in starting wages across jobs than do females with more education. Royalty (1998) and Holzer and LaLonde (2000) show that the kinds of job-to-job changes that have potentially positive effects on the earnings of young workers are relatively infrequent among young, less-educated women, while job-to-nonemployment changes occur more frequently among this group.

Few previous studies adequately take into consideration unobservable differences between job changers and stayers and the endogenous determination of mobility (i.e., the self-selection problem).[5] Moreover, with the exception of Antel (1986) and Garcia-Perez and Sanz (2004), these studies do not distinguish between voluntary and involuntary separations when computing average mobility returns and job turnover. In this paper, I estimate a multinomial endogenous switching model of wage growth to attempt to address the endogeneity between job transitions and wage growth. The analysis explores the relationship between turnover and expected wage growth opportunities, and examines differences in job skill requirements that link these two dynamic processes.

3. DATA DESCRIPTION AND DEFINITIONS OF KEY VARIABLES

3.1. The Women's Employment Survey (WES)

The Women's Employment Study drew a random sample of single mothers who received cash assistance in February 1997 in an urban Michigan county. To be eligible for the sample, the women had to reside in this county, be U.S. citizens between the ages of 18 and 54, and be either Caucasian or African-American. Interviews were conducted in Fall 1997, Fall 1998, Fall 1999/Winter 2000, Fall 2001/Winter 2002, and Fall 2003/Winter 2004. The response rate was 86% for the first wave ($N = 753$), 92% for the second wave ($N = 693$), 93% for the third wave ($N = 632$), 91% for the fourth wave ($N = 577$), and 92% ($N = 532$) for the fifth wave of this panel study. Roughly 80 months of data are available for respondents.

The sample was drawn as the transition from the old welfare system to the new one was being implemented. Whereas all respondents received cash assistance in February 1997, about one-quarter had left welfare by Fall 1997, one-half by Fall 1998, 70% by Fall 1999, and 75% by Fall 2001.

I utilize many measures not available in other studies, including information about respondents' work histories, welfare histories, basic job skills, hourly wage of their main job, number of hours worked in this job, and whether they received employer-provided health benefits. Human capital variables include years of schooling, years of full-time and part-time work experience, occupation in which recipient has previous work experience, and number and type of job tasks ever performed on a daily basis in any previous job held. Type of job tasks include reading/writing paragraph-length material, arithmetic, use of computer, supervising co-workers, keeping a close watch on gauges/dials/instruments, filling out forms on a daily basis, and use of client/customer communication skills on a daily basis.[6] The health-related measures I use include physical limitations, mental health problems, child health problems, and experiences of severe domestic abuse.

3.2. The Michigan Employer Survey (MES)

In Fall 1997 (during the same period the initial wave of WES was underway), Harry Holzer administered a telephone survey to 900 establishments in three large metropolitan areas in Michigan. The employers surveyed were drawn from a sample that was stratified ex-ante by establishment size, so that the sample roughly represents the distribution of the workforce across establishment size categories. The survey was administered to the individual responsible for entry-level hiring, and to all establishments that had hired someone within the past two years. Conditional on meeting these criteria, response rates averaged over 70% (Holzer, 1999). In Fall 1999, a follow-up survey of these firms was conducted, yielding a response rate of 70%.

Each employer was asked a series of questions about the characteristics of the most recently filled job that did not require a college degree. Because the firms are represented in proportion to the number of workers they employ, this sample of recently filled noncollege jobs constitutes a representative sample of the jobs that are available in the local labor markets over a period of several months (Holzer, 1996). Employers were also asked a similar series of questions about the characteristics of jobs previously (within the past two years of the survey) filled by welfare recipients. Questions focused on: (1) the hourly wage, hours, and health benefits offered in the job; (2) the occupation/position in which this worker was hired; (3) the credentials and skills employers sought and the hiring criteria used; (4) the daily task requirements of the job (where the job task measures are identical to those used in WES); (5) the wage growth prospects of the job (including provision of on-the-job training, chance of within-job pay increases, and chance of promotion within the firm assuming good performance); and (6) job performance and job tenure of the recently hired worker.

Given the high response rates and extensive survey instruments, these data sets provide complimentary evidence from the supply and demand side on the relationships between job skill requirements, and the wage and job dynamics of former/current recipients in the post-welfare reform era.[7]

3.3. Job Skill Variables

The MES and the WES contain the same sets of questions about job tasks/work skills. WES collected information from each respondent about whether she performed each of these job tasks on a daily basis in a job(s) held between waves, as well as whether she had ever performed these tasks on any job previously held. I use this information to construct a job task work history for each respondent. I compute a measure of experience using each of these job skills for every individual and build a dynamic measure of job skill use. Suppose that a worker reports having no prior work experience using computer skills as of the Wave 1 interview, then reports using computer skills on a job(s) held between Waves 1 and 2, and also reports using computer skills on a job(s) held between Waves 2 and 3. Hence, between 1997 (Wave 1) and 1998 (Wave 2), the computer-use indicator in the wage regression changes from zero to one. Furthermore, between 1998 (Wave 2) and 1999 (Wave 3), experience using computer skills also changes from zero to one.

I measure workers' ability to perform tasks based on their having done so on a previous job, even though previous job skill experience may not accurately reflect current abilities. Because previously acquired skills may depreciate during periods of nonwork (Mincer & Ofek, 1982; Corcoran, Duncan, & Ponza, 1983; Stratton, 1995), I focus on respondents' job skills used within the year prior to the employment outcome. In the analyses that follow, I measure years of job skill experience like a tenure-skill measure – i.e., the number of consecutive years using the relevant job skill. I also tried, alternatively, measuring years of job skill experience as a pure experience-skill measure – i.e., the cumulative number of years in which a worker ever used the relevant job skill. This alternative way of measuring job skill experience did not qualitatively change any of the underlying findings reported in this paper.[8]

Now, consider an individual who reports having prior work experience using computer skills at Wave 1, but reports not using computer skills on a job(s) held between Waves 1 and 2, and then reports using computer skills on a job(s) held between Waves 2 and 3. I count an accumulated year of experience using a particular job skill only if the job skill has been used in consecutive periods. Thus, this individual is not counted as having accumulated an additional year of experience using the particular job skill over the period because of her intermittent job skill use.

3.4. Job-Transition Pattern Variables

Using the WES, I characterize employment patterns and the extent of job stability and job mobility between waves, using retrospective questions from each wave on job tenure, monthly job/employment history, and reported reason for job separation (if any occurred). The wages, hours, and health benefits of the most recent job are recorded at each interview (given the individual has worked at some point between interviews).[9] Therefore, I count job separations over the period between two interviews.[10] If a person is between jobs at the time of an interview, the separation is assigned to the interview year when she starts her next job. I distinguish job separations both by whether they were voluntary or involuntary (i.e., due to being laid-off or fired), and by whether they were followed by a nonemployment spell of four or more weeks.

I define three patterns of job transitions: job stability, job mobility, and job instability. Individuals whose the current/most recent job at wave t was the same as that held at the previous wave are denoted as experiencing job stability. Job mobility occurs when respondents made a voluntary job change without experiencing any involuntary separations or transitions into nonemployment. I distinguish between job instability that is due to being laid-off or fired from instability that results from an employee-initiated job-to-nonemployment transition.[11,12] I define a "transition" as a job-to-job transition if the job change was voluntary and the interval between jobs was less than four weeks. Conversely, I define a transition as being into nonemployment only if the spell of nonwork lasts four or more weeks, or if the job change results from being laid-off or fired. Nonemployment spells of more than a month are less likely to be the result of nonemployment chosen in order to search for a new job more intensively, and are more likely to be the result of nonmarket/nonsearch reasons.[13] Royalty (1998) and Gladden and Taber (2000) use similar definitions of job transitions.

4. ESTIMATION STRATEGY, MODEL SPECIFICATION, AND RESULTS

I begin by using the MES to estimate the determinants of the starting wage earned on jobs recently filled by former/current welfare recipients, with particular emphasis on the effects of job skill requirements. To examine the determinants of wage growth prospects, I next estimate a series of probit equations of whether the job provides on-the-job training opportunities, a chance of merit-based pay increases, the likelihood of promotion (ordered probit: poor, fair, good, excellent), and whether a promotion was received within the past year (since date of hire), respectively, using MES.

The longitudinal aspect of the WES is then exploited to take into account unobserved heterogeneity on (i) the effects of various job skills on the wage profile, (ii) the effects of different job transition patterns (job stability, job mobility, job instability) on wage growth, and (iii) the propensity to change jobs. I also identify the returns to various job skills. Are workers who use a given set of job skills better paid than workers who do not use these skills? If the answer is positive, I examine whether workers using these skills received higher pay before using these skills on the job, or received higher pay as soon as they started using these skills on the job, or finally, received higher pay once they had sufficient experience using these skills on the job.

In the models estimated below, I conceptualize a job in terms of its production aspects (inputs) as a collection of tasks. Job tasks are not independent of the workers who perform the tasks. Thus, disentangling person-specific and job-specific effects has implications for whether low-wage jobs are inherently dead-end – and if so, which kinds of jobs? A job can be defined by the technological investment opportunity it provides a worker

(Lazear, 1995; Rosen, 1972). On-the-job training typically provided in non-college jobs are not firm-specific (i.e., training received, which is valued within the firm but less easily transferable to other jobs), but rather consist of general and occupation-specific training. These opportunities may be of especial importance for low-skilled workers, affecting both their probability of experiencing wage growth within jobs and the probability of experiencing wage growth via job changes.[14] If at all training costs are paid by the employers, and the skill enhancement programs are, at least to some degree, portable, then we would expect the workers to bear some portion of the costs by receiving lower starting wages (Parent, 1999).

4.1. Wage Analysis Using MES

4.1.1. Specification

Consider the following log starting wage equation augmented with a set of job task/skill variables:

$$\begin{aligned}\ln(\text{STARTWAGE})_{ijt} = \beta_0\text{HSGRAD}_{it} + \beta_1\text{PRIOREXP}_{it} \\ + \beta_2\text{SKILLCERT}_{it} + \beta_3\text{JOBSKILL}_{ijt} \\ + \beta_4\text{JOBHOURS}_{ijt} + \beta_5\text{OJT}_{ijt} + \Gamma\mathbf{Z}_{ijt} + \varepsilon_{ijt} \quad (1)\end{aligned}$$

where STARTWAGE represents the real starting hourly wage of person i in job j at time t; HSGRAD, PRIOREXP, and SKILLCERT are variables indicating whether the individual possesses a high school diploma/GED, prior occupation-specific work experience, and training/skill certification, respectively; JOBSKILL is a vector of job skill/task variables; JOBHOURS indicates whether the job is part-time; on-the-job training (OJT) indicates whether on-the-job training opportunities are provided; and $\mathbf{Z}$ represents a vector of firm characteristics. I include OJT to test whether workers pay for formal OJT by accepting lower starting wages.

I am particularly interested in estimating the effects of the set of job skills. The inclusion of employee characteristics used in conventional human capital specifications – specifically, possessing high school diploma, previous occupation-specific work experience, and skill or training certification – may lead to an underestimate of the impact of job skills because the output of schooling presumably includes many of the observed job skills. In addition, it is not clear whether occupation dummies are appropriate variables to include in the regressions that follow, because possessing particular job skills may enable workers to qualify for jobs in higher paying occupations. Thus, I present several alternative specifications of the model using MES in

Columns (1–3) of Table 1. The specifications differ in whether they control for conventional human capital characteristics and/or differences across occupation. This helps to determine whether particular job skills are associated with higher pay because they are associated with higher paying occupations or because, within occupation those with more job skills receive higher pay. This also helps identify whether education is associated with higher pay because it is associated with the possession of essential job skills that are associated with higher pay. Column (1) includes only the set of job skill variables as measures of skill; the specification in Column (2) includes controls for conventional human capital variables (but not occupation); both conventional human capital variables and occupation controls are included in Column (3). Because the inclusion of occupation variables in such a regression is likely to lead to an underestimate of the impact of job skills, I emphasize the regression results from specification (2).

4.1.2. MES Results

Columns (1–3) of Table 1 show the results obtained by estimating the starting wage equation using MES. The mean and median starting wage in jobs previously filled by former/current welfare recipients was \$6.75 and \$6.50, respectively. As can be seen from specification (2), possessing a training or skill certification increases the starting wage by 8%; neither the possession of a high school diploma nor previous experience in the particular line of work significantly affected the starting wage after the set of job skill variables were included. Use of reading/writing skills is associated with a 12.7% higher starting wage; while use of math and customer communication skills are both linked with lower pay. The likely reason for the negative coefficients on the use of math and customer communication skills is that these activities are negatively correlated with other unobserved activities using valued skills. Thus, where math and/or customer communication skills are very important, workers are not using other more highly valued skills. Another explanation is that math (including making change) and customer communication skills have a relatively low supply price, as they are more easily learnable, with an effectively zero/low cost of acquisition. It is also likely that computers have increased the value of some skills (e.g., reading/writing), while decreased the value of others (e.g., arithmetic, see, Levy and Murnane's, 1996, work examining with what skills are computers a complement). Somewhat surprisingly, jobs that required the use of computer skills did not pay significantly higher starting wages than those that did not require these skills. The set of job skill/task variables are not simply capturing attachment to specific occupations (e.g., fast-food jobs (math/customer communication),

Table 1. Determinants of Starting Wages using MES. Dependent Variable: Log of Real Starting Hourly Wages ($1999). (Robust Standard Errors in Parentheses).

Explanatory Variables	Mean	(1)	(2)	(3)
Human capital variables				
High school Diploma/GED	0.8239	–	0.0197	0.0229
			(0.0249)	(0.0244)
Prior occupation-specific work experience	0.5074	–	−0.0257	−0.0131
			(0.0285)	(0.0271)
Training/skill certification	0.3743	–	0.0801***	0.0953***
			(0.0309)	(0.0291)
Job skill variables				
Reading/writing	0.4771	0.1357***	0.1273***	0.1011***
		(0.0271)	(0.0284)	(0.0276)
Computer	0.4060	0.0399*	0.0334	−0.0353
		(0.0290)	(0.0291)	(0.0298)
Math	0.6327	−0.1005***	−0.0952***	−0.1090***
		(0.0271)	(0.0266)	(0.0279)
Customer communication	0.7399	−0.1006***	−0.1049	−0.0469*
		(0.0282)	(0.0274)	(0.0339)
Occupation				
(Reference category: service)				
Sales	0.1996	–	–	0.1357***
				(0.0358)
Clerical	0.2067	–	–	0.2492***
				(0.0385)
Blue-collar	0.1767	–	–	0.1920***
				(0.0441)
Other job characteristics				
Part-time	0.2500	−0.1084***	−0.1156***	−0.0859***
		(0.0306)	(0.0311)	(0.0316)
On-the-job training	0.6371	–	−0.0369*	−0.0355*
			(0.0243)	(0.0230)
Firm characteristics				
% Employees unionized	16.1507	0.0025***	0.0025***	0.0021***
		(0.0004)	(0.0004)	(0.0004)
Firm Size (Reference category: ≥100 employees)				
1–9 employees	0.2071	−0.0760*	−0.0772*	−0.0891**
		(0.0400)	(0.0399)	(0.0369)
20–99 employees	0.3636	−0.0588*	−0.0518*	−0.0678**
		(0.0325)	(0.0321)	(0.0307)
R^2		0.2266	0.2400	0.3223
Sample size		505	505	505

Note: Regressions also include metropolitan area dummies and a constant term. The mean and median wage for this sample of jobs filled by former/current welfare recipients is $6.75 and $6.50, respectively. See Section 3 for description of MES.

*Statistically significant at the 0.10 level (one-tailed test).

**Statistically significant at the 0.05 level.

***Statistically significant at the 0.01 level.

clerical jobs (reading/writing)), since the pattern of results is similar when occupation variables are included. Among the occupations, the results indicate that service jobs – the occupation in which recipients are disproportionately concentrated – offered the lowest starting pay, while clerical jobs offered the highest starting pay.

Part-time jobs are associated with 11.6% lower starting wages; while both larger firms and firms with greater fractions of unionized employees pay higher starting wages. I also find evidence that workers pay for part of their training programs by accepting lower starting wages. The starting wage estimates reveal that provision of OJT opportunities lowers starting wages by 3.7%. Furthermore, the potential effect of job-match or individual heterogeneity biases will be to underestimate the effect of OJT on the starting wage since higher ability (and better matched) individuals are likely to be paid more and receive more training. Thus, this estimate of the impact of OJT may be considered a lower bound.

The emphasis of the remainder of my empirical analysis is on modeling the process of wage changes resulting in the current hourly earnings (as opposed to modeling wage levels), because a fundamental question that needs closer investigation concerns earnings dynamics that accompany initial employment at low wages. Employers report that jobs filled by previously hired recipients that require both reading/writing and computer skills were more likely to offer potential wage increases for merit, greater chances for promotion (with good performance), and were more likely to offer formal job training opportunities. Recipients who received formal job training and worked in jobs requiring reading/writing and computer skills experienced almost twice the number of formal job training hours relative to those holding jobs that require only soft skills.[15] This suggests that a lack of cognitive skills may not only affect the kinds of jobs some recipients can get, but, because of fewer OJT opportunities, may also affect their potential for wage growth.

In columns 1–4 of Table 2, I present estimates from a series of probit equations of whether the employer reports that the job provides OJT opportunities, a chance of merit-based pay increases, the likelihood of promotion (ordered probit: poor, fair, good, excellent), and whether a promotion was received within the past year (since date of hire), respectively.[16,17] As shown in the first column, 63.7% of the sample of jobs recently filled by former/current welfare recipients provided some type of OJT (not including training that was remedial).[18] The results indicate that the probability that a given job offers OJT increases by eight percentage points if the job requires reading/writing skills, and increases by 4.4 percentage points if the job

Table 2. Determinants of Wage Growth Prospects.

Explanatory Variables	Dependent Variables (Robust Standard Errors in Parentheses)						
	(1)		(2)		(3)	(4)	
	Provision of on-the-job training (probit estimates)		Offers chance of within-job pay increase (probit estimates)		Employer-reported promotion prospect (1 = poor, 2 = fair, 3 = good, 4 = excellent) (ordered probit estimates)	Received promotion since date of hire (probit estimates)	
	Mean	dF/dx	Mean	dF/dx	Coefficient	Mean	dF/dx
Work performance-related variables							
Absenteeism problem						0.4203	−0.0385
							(0.0386)
Work attitude problem						0.1884	0.0510
							(0.0692)
Job skill-related problem						0.1304	−0.0866
							(0.0287)
On-the-job training			0.6371	0.1221***	0.3166***	0.6860	0.0588*
					(0.1234)		(0.0392)
Remedial training						0.2657	−0.0788**
							(0.0338)
Job skill variables							
Reading/writing	0.4771	0.0803*	0.4771	0.0202	0.0665	0.5217	0.0255
		(0.0460)		(0.0441)	(0.1103)		(0.0456)
Computer	0.4060	0.0438	0.4060	0.0711*	0.3894***	0.3140	0.1761***
		(0.0466)		(0.0438)	(0.1295)		(0.0722)
Math	0.6327	−0.0017	0.6327	0.0356	0.1160	0.5652	−0.1149**
		(0.0442)		(0.0451)	(0.1139)		(0.0526)
Customer communication	0.7399	−0.0419	0.7399	−0.0455	−0.0820	0.7053	0.0366
		(0.0484)		(0.0474)	(0.1497)		(0.0446)

Human capital variables							
Job tenure (months)						7.4	0.0062**
							(0.0032)
High school Diploma/ GED						7101	0.0027
							(0.0446)
Prior occupation-specific work experience						0.4976	−0.0057
							(0.0452)
Training/skill certification						0.4300	−0.0086
							(0.0428)
Sales					0.3610**	0.1836	−0.0152
					(0.1521)		(0.0519)
Clerical					−0.1047	0.1884	−0.0135
					(0.1709)		(0.0470)
Blue-collar					0.2644*	0.1691	0.0285
					(0.1744)		(0.0683)
Other job characteristics							
Part-time	0.2500	−0.0955**	0.2500	−0.1123**	0.0140	0.3768	−0.0542*
		(0.0472)		(0.0498)	(0.1215)		(0.0388)
Firm characteristics							
% Employees unionized	16.1507	0.0001	16.1507	−0.0032***	−0.0034**	19.2121	−0.0004
		(0.0006)		(0.0006)	(0.0017)		(0.0007)
Firm size (Reference category: ≥100 employees)							
1–9 employees	0.2071	−0.0531	0.2071	0.1209**	−0.1600	0.1594	0.1732**
		(0.0609)		(0.0512)	(0.1642)		(0.1082)
20–99 employees	0.3636	0.0168	0.3636	0.0676*	0.0012	0.3285	0.1208**
		(0.0490)		(0.0430)	(0.1248)		(0.0653)
Log-likelihood		−377.7363		−297.3637	−580.7765		−63.2523
Observed Fraction providing OJT		0.6371					
Predicted problem of OJT (eval at sample means)		0.6395					

Table 2. (*Continued*)

Explanatory Variables	Dependent Variables (Robust Standard Errors in Parentheses)						
	(1)		(2)		(3)	(4)	
	Provision of on-the-job training (probit estimates)		Offers chance of within-job pay increase (probit estimates)		Employer-reported promotion prospect (1 = poor, 2 = fair, 3 = good, 4 = excellent) (ordered probit estimates)	Received promotion since date of hire (probit estimates)	
	Mean	d*F*/d*x*	Mean	d*F*/d*x*	Coefficient	Mean	d*F*/d*x*
Obsvd fraction offer chance of W/in-job pay increase				0.7036			
Predicted problem of within-job pay increase				0.7238			
Sample Size	587		550		502	207	

Note: Regressions also include controls for metropolitan area, starting hourly wages, and employee human capital characteristics. 43.8% of employers reported excellent promotion prospects; 33.7% reported good promotion prospects; 13.9% reported fair promotion prospects, and 8.6% reported poor promotion prospects. The mean length of time represent the derivative of the probability of the outcome with respect to a unit-change in the explanatory variable (discrete change of dummy variable from 0 to 1), evaluated at the sample means.

*Statistically significant at the 0.10 level (one-tailed test).

**Statistically significant at the 0.05 level.

***Statistically significant at the 0.01 level.

requires computer skills (though the latter coefficient is not statistically significant). On the other hand, part-time jobs are 9.6 percentage points less likely to provide OJT.

As shown in the second column, according to employer reports, 70% of the jobs recently filled by welfare recipients offered chances for within-job pay increases (above cost of living increases) assuming good performance. The results show that jobs that provide OJT are 12.2 percentage points more likely to offer within-job wage growth opportunities. The impact of the use of reading/writing skills on the potential of within-job pay raises becomes insignificant after the inclusion of OJT, suggesting that one of the primary ways reading/writing skills affects wage growth prospects is through the provision of more OJT opportunities. Although computer skills did not significantly affect the starting wage (Table 1), jobs that require computer skills are 7.1 percentage points more likely to offer potential merit-based pay increases. On the other hand, part-time jobs are 11.2 percentage points less likely to offer chances of merit-based pay increases. While larger firms and firms with greater fractions of employees that are unionized offered higher starting wages (Table 1), these firms offer fewer chances for within-job merit-based pay increases.[19]

As shown in the third column, employers reported that, assuming good performance, 43.8% of the jobs recently filled by welfare recipients offered excellent promotion prospects, 33.7% offered good, 13.9% offered fair, and 8.6% offered poor promotion prospects. I estimate an ordered probit regression, where the dependent variable takes on the values: 1 = poor, 2 = fair, 3 = good, 4 = excellent. The same general pattern of results emerges: use of computer skills and OJT are associated with greater upward mobility prospects. I include a set of occupation dummy variables to control for differences in the structure of promotion opportunities across occupations. As expected, sales and blue-collar occupations have greater promotion prospects than service and clerical jobs.

Working in firms with smaller fractions of unionized employees is positively associated with promotion receipt, possibly because unionized firms are more likely to base promotion on seniority than are nonunionized firms (Abraham & Medoff, 1985). Unions are associated with flatter age-earnings profiles. Given that the sample is comprised of relatively young workers, seniority rules may hamper the promotion prospects of those who are unionized.

The results presented in columns 2 and 3 of Table 2 are based upon employer reports of the potential wage growth prospects, while the last column presents results from estimating a probit equation on actual receipt of a promotion since being hired with the firm. 44.7% of employers reported that former/current recipients' work performance was about the same as other workers that have previously filled the position; 16.5% reported recipients' work performance was much better, 25.7% reported recipients' work performance was a little better; while 9.2% and 3.9% reported recipients' work performance was a little worse and much worse, respectively, than other workers. 42% of employers reported previously hired recipients had absenteeism problems, 18.9% reported work attitude problems, and 13% reported previously hired recipients had job skill-related problems. I include these indicator measures of poor work performance, based upon employer reports, in the model of actual promotion receipt.

The mean length of time that had elapsed since the date of hire was 7.4 months. Fifteen percent of recipients had received a promotion as of the survey interview date. Despite the relatively short period of time that had elapsed since the date of hire, the results are instructive. Most of the significant predictors of the probability of promotion are the same for employer reports of promotion prospects as for actual promotion receipt. Job skill-related work performance problems significantly reduce the probability of promotion. The OJT (not including that which is remedial) significantly increases the probability of promotion receipt, while a remedial OJT is negatively associated with promotion receipt (this is likely picking up worker job-skill deficiencies). Use of computer skills is significantly associated with promotion receipt, while use of math skills is negatively associated with promotion receipt (likely explanation for negative association previously discussed). I do not find significant differences in promotion receipt across occupation groups, after the inclusion of the set of job skills. Job tenure is significantly related to promotion receipt. The effect of job tenure and company training on promotion likelihood suggests that the acquisition of job-specific skills resulted in promotion.

Working part-time is associated with a significant reduction of promotion rates, as is working in large firms. This latter result is counter-intuitive since we would expect larger workplaces to have greater availability of opportunities for upward mobility (Idson, 1989). Given the short length of time that had elapsed since the date of hire for this sample of relatively young workers, seniority rules may have hampered the promotion prospects of those who were working in large firms, due to the more structured organization of jobs that generally accompanies larger firms.

The results presented up to this point cannot be used to determine decisively between competing explanations – in particular, whether the estimated effects of different job skills (e.g., reading/writing, computer) reflect the true return to the job skill (i.e., job skill affecting wage profile), or whether the relationship between use of a set of job skills and wage growth is purely the result of job sorting (selection of abler workers/high-ability types). It remains unclear how the use of different sets of job skills affects the earnings profile, since unobservable worker characteristics are not directly controlled for here. Controlling for unobservable worker heterogeneity is important because workers using a particular job skill that is associated with higher wage growth may have experienced greater wage growth in the absence of the use of that skill (if unobserved fixed worker quality is driving results). Thus, in the next section, I use the longitudinal data on former/current welfare recipients to control for unobservable worker heterogeneity to isolate the return to job skills.

4.2. Wage Growth Analysis using WES

4.2.1. WES Sample Descriptive Statistics

Overall work experience accumulated masks heterogeneity in job transition patterns, which may have significant effects on wage growth trajectories. In particular, while the most respondents worked in for the most of the months over the five years of the panel (the mean number of months worked is roughly 40 months),[20] and much of this accumulated experience working in full-time jobs, job instability was the most common employment pattern between successive waves. Roughly half of the respondents experienced job instability, while 27.4% experienced job stability and 20.2% experienced job mobility between successive waves.[21,22] The worsening economic conditions in 2001 increased the risk of job loss. Among individuals who experienced job separations between waves, separations resulting from being laid-off or fired increased from 21.3% to 27.9% between 1998–1999 and 1999–2001.

There was a significant amount of within-person changes in job skills used over the period. In estimating the wage growth models that follow, I include differences in job skills used, changes in job hours, and occupation transitions to account for the heterogeneity in wage growth. I am interested in the relationship between job transition patterns and wage growth. I examine the mean wage growth associated with different job transition patterns – job stability, voluntary job mobility, and job instability. I investigate the extent to which average wage growth masks heterogeneity in within- and between-job wage growth, and examine whether differences in job skill requirements can explain the observed heterogeneity in wage growth.

4.2.2. WES Wage Specification and Estimation Strategy

My estimation model assumes human capital characteristics (job task attributes) affect not only wage levels, but also the process of wage growth (e.g., via learning ability or differences in human capital investment opportunities across jobs). Low wages may be a greater reflection of a worker's learning ability (or lack of OJT opportunities) as well as their earning ability – e.g., individuals who have more ability and motivation may learn more from work experience.

Consider the following log wage equation augmented with job-skill variables:

$$\ln(\mathrm{WAGE})_{ijt} = \Gamma \mathbf{Z}_{ijt} + \beta_0 \mathrm{EXP}_{it} + \beta_1 \mathrm{JOBSKILL}_{ijt} + \beta_2 (\mathrm{EXP\ using\ JOBSKILL})_{ijt} + \alpha_i + u_{ijt} \tag{2}$$

where WAGE represents the real hourly wage of person i in job j at time t; $\mathbf{Z}$ is a vector of educational attainment, demographic variables, health-related variables, county unemployment rate, and other controls; EXP is years of full-time and part-time work experience (entered separately, with quadratic terms); JOBSKILL and EXP using JOBSKILL is a vector of job-skill variables and the corresponding years of experience using these job skills, respectively.

I include both the vector of job-skill variables and measures of the number of years of experience using the these job skills to allow the use of job skills to affect both the wage level and wage growth (i.e., the slope of the wage-experience profile). For example, the latter may capture the potentially enhancing productivity of computer usage or the greater provision of OJT opportunities in jobs requiring particular skills. I decompose returns to various job skills into a constant and a part related to experience.

Note that the error term in the above equation contains a time-invariant person-specific effect, α_i. If less-able or less-motivated workers are less likely to work in jobs requiring valued job skills, estimates of returns to job skills that fail to control for α_i may be biased toward finding larger effects. I present the WES cross-sectional wages estimates of Eq (2) (which do not control for unobserved heterogeneity) in Appendix Table A1. I use the cross-sectional estimates as a benchmark to compare with the fixed effect estimates. The overall pattern of the WES cross-sectional results are similar to those yielded using employer reports. The fundamental problem with the cross-sectional results is that, despite the extensive set of controls, the measure of particular skills in the workplace may be positively correlated with unobserved characteristics that also generate wage premia, causing the job-skill coefficients to be upwardly biased.

I explore two different ways of assessing the likely size and significance of this bias by exploiting the longitudinal aspect of WES. First, to control for unobserved worker characteristics, I estimate a first-difference fixed effect wage equation of the following form (augmented with job transition-pattern variables):

$$\begin{aligned}\Delta \ln(\text{WAGE})_{i(t-1,t)} = {} & \beta_0(\Delta\text{EXP})_{i(t-1,t)} + \beta_1(\text{JOBTRNSITN})_{i(t-1,t)} \\ & + \beta_2(\Delta\text{EXP} * \text{JOBTRNSITN})_{i(t-1,t)} \\ & + \beta_3(\Delta\text{JOBSKILL})_{i(t-1,t)} \\ & + \beta_4(\Delta\text{EXP using JOBSKILL})_{i(t-1,\,t)} \\ & + \Gamma(\Delta\mathbf{Z})_{i(t-1,t)} + \Delta u_{i(t-1,t)} \end{aligned} \tag{3}$$

Because the person-specific time-invariant effect (α_i) has been differenced out, equation (3) can be estimated by OLS and is a consistent estimation method for identifying the effects of time-varying characteristics.[23] In estimating the first-difference fixed effect model, many of the terms in **Z**, such as education, sex, and race, have also been eliminated since they do not vary with time.

In the first-difference specification, I include job transition pattern variables and control for occupation transitions using a one-dimensional occupation index. The inclusion of these variables enables me to isolate the true return of job skills independent of the effects of job changes that may have led to the change in job skills used (for a given worker).

The JOBTRNSITN vector captures whether the individual experienced job stability, job mobility, a voluntary job separation with an intervening spell of nonemployment, or an involuntary job separation, between the most recent job of successive waves. The change in work experience and job transition variables are entered separately and interacted with each other in the first-difference specification. The sum of the relevant job transition and work experience terms along with their interactions, captures the sum of the returns to experience and returns to tenure for individuals who experienced job stability; and captures the sum of the returns to experience and the change in the job match component for individuals who experienced the relevant type of job change.

We expect wage growth to be higher for individuals who experience job mobility relative to those who experience job instability. Individuals are presumed to voluntarily change jobs because they expect a wage gain, while individuals who experience job instability (particularly, resulting from being laid-off/fired) may lose job-specific human capital and matching capital because employers use the stability of potential workers' employment histories as a signal for good matches (Gladden & Taber, 2000). We also expect that returns to job stability (i.e., individuals whose current/most recent job in wave *t* was the same as that held in the previous wave) will be higher than the returns to job instability.

The job mobility decisions are likely endogenous with respect to wage changes. One reason individuals stay in the same job is because they work in jobs with more potential wage growth opportunities. This produces a downward bias on the estimated effects of job mobility (relative to job stability), since the counterfactual – the wage growth of the individual would have experienced had she stayed in the same job – is not observed. In this way, the estimates of the gains to job mobility (relative to job stability) may be considered lower bound estimates. For this precise reason, in the final empirical section of this paper, I estimate a multinomial endogenous switching model of wage growth to better address the endogeneity of job mobility, which is described at the end of Section 4 and Appendix A.

In light of the prevalence of occupation changes among our sample, I include a control for occupation-transition characteristics. I create a one-dimensional occupation index that is designed to capture the amount of human capital needed to work in different occupations. I detail in Appendix B the derivation of the occupation index. My construction of the index is adapted from that previously developed by Sicherman and Galor (1990).[24]

4.2.3. Mean Wage Growth

Table 3 shows the distribution of annual within-job real wage growth and the distribution of annual real wage growth with voluntary job mobility and with job instability. On average, real wages grew 4.1% per year for individuals who remained in the same job, but by 7.3% per year for individuals working full-time on the same job, and not at all for individuals working part-time on the same job. The mean wage gain for workers who experienced voluntary job mobility was 6.2%. The selected sample of individuals who experienced voluntary job mobility is not representative of all workers, and thus their mean wage growth does not represent that which a random worker would experience if she changed jobs, but rather represents the expected wage growth conditional on voluntarily changing jobs.[25] In terms of the underlying economic variables of standard wage models, these results suggest that the improvement in job match, for those who find successful job matches, is comparable to the gains from returns to work experience and tenure; and, thus job changes are an important source of wage growth.

Table 3. Distribution of Average Annual Real Wage Growth (in natural logs), by Type of Job Transition.

	Within-Job Wage Growth, Full-Time Job	Within-Job Wage Growth, Part-Time Job	Wage Growth w/Vol Job Mobility	Wage Growth w/Invol Job Change	Wage Growth w/EE-initiated Job Instab
	(1)	(2)	(3)	(4)	(5)
Annual wage growth					
Mean	0.073*	−0.014	0.062***	0.007	0.024**
Median	0.042	0.012	0.051	−0.006	0.032
Cumulative distribution:					
−0.10	0.108	0.259	0.197	0.277	0.220
0 (percent non-positive)	0.280	0.452	0.384	0.518	0.401
+0.10	0.693	0.716	0.605	0.712	0.633

All wages have been converted to real wages (1999 dollars) using the Consumer Price Index for All Urban Consumers (CPI-U).
Source: Women's Employment Survey, 1997 – early 2004.
*Statistically significant at 10% level.
**Statistically significant at 5% level.
***Statistically significant at 1% level (two-tailed tests).

The results in Table 3 reveal the importance of differentiating between job changes resulting from voluntary job mobility and those resulting from job instability. Annual wage growth was nonexistent among individuals who experienced involuntary job separations, and 2.4% among those who experienced employee-initiated job-to-nonemployment transitions. Table 3 also reveals that average wage growth masks substantial heterogeneity in wage growth within each of the job transition patterns. Large fractions of individuals experienced real wage declines, particularly those who experienced job instability. Part of the declines in real wages, however, is likely due to measurement error (Gottschalk, 2002).

4.2.4. First-Difference Fixed Effect Results

The first-difference estimates are presented in Table 4. The first column reports estimates of a model that includes only the job transition and standard work experience variables, while the second column shows the full model. Results from the parsimonious specification indicate that an additional year of full-time work experience with job stability is associated with 4.8% increase in pay, and an additional year of full-time work experience accompanied by a voluntary job change is associated with 10.3% increase. This evidence suggests that job mobility is a critical component of the wage growth process for these less-skilled women. On the other hand, the return to an additional year of full-time work experience that includes an involuntary job separation is small and statistically insignificant. Accumulated part-time work experience had a negligible effect on wage growth (this was true with any of the job transition patterns).

The average annual amount of full-time work experience accumulated for individuals who experienced job instability was roughly five months – or only 44% of the amount accumulated by individuals working full-time continuously over the year. The indirect effect of job instability on wage growth through its effect on the loss of potential full-time work experience accumulation is, therefore, estimated to be a wage loss of 2.7% relative to the rate of annual within-job wage growth (0.56*0.048), and a loss of 5.8% relative to the rate of annual wage growth occurring with voluntary job mobility (0.56*0.103).

The model in column (2) estimates the effects of job skills. I find that when workers change from not using reading/writing skills to using these skills on a daily basis, their wage increases immediately by 4.7%. Furthermore, workers earn an additional 4.6% wage premium with each additional year of experience using reading/writing skills (over and above the return to general work experience). In contrast, in the longitudinal dimension, the

Table 4. First-Difference Fixed Effect Wage Estimates.

Explanatory Variables	(1)	(2)
Dependent Variable: First-Difference of Log of Real Hourly Wages ($1999)		
Human capital variables		
ΔFull-time work experience	0.0484***	0.0150
	(0.0103)	(0.0159)
ΔPart-time work experience	0.0286	0.0115
	(0.0242)	(0.0272)
Job mobility	0.0996*	0.0533
	(0.0617)	(0.0612)
Involuntary job instability	0.0662	0.0069
	(0.0523)	(0.0508)
Employee-initiated job instability	0.0180	−0.0341
	(0.0348)	(0.0367)
(ΔFull-time work experience)*(Job mobility)	−0.0453	−0.0175
	(0.0413)	(0.0406)
(ΔFull-time work experience)*(Employee-initiated job instability)	−0.0187	0.0067
	(0.0341)	(0.0355)
(ΔFull-time work experience)*(Involuntary job instability)	−0.0920**	−0.0445
	(0.0418)	(0.0414)
(ΔPart-time work experience)*(Job mobility)	−0.0972*	−0.0852
	(0.0551)	(0.0553)
(ΔPart-time work experience)*(Employee-initiated job instability)	0.0026	0.0453
	(0.0455)	(0.0451)
(ΔPart-time work experience)*(Involuntary job instability)	−0.1159*	−0.0524
	(0.0648)	(0.0635)
Return to (FTExp+Tenure) w/job stability	0.0484***	0.0150
Return to (FTExp+ΔJobMatch component) w/Job mobility	0.1027***	0.0509*
Return to (FTExp+ΔJobMatch component) w/InvolJobInstability	0.0226	−0.0226
Return to (FTExp+ΔJobMatch component) w/EE-InitiatedJobInstability	0.0477***	−0.0124
Job skill variables		
ΔReading/writing		0.0472**
		(0.0198)
ΔExperience using reading/writing		0.0461***
		(0.0154)
ΔComputer		−0.0071
		(0.0212)
ΔExperience using computer		−0.0119
		(0.0189)
ΔMath		0.0213
		(0.0183)
ΔExperience using math		0.0086
		(0.0173)

Table 4. (*Continued*)

Explanatory Variables	(1)	(2)
Dependent Variable: First-Difference of Log of Real Hourly Wages ($1999)		
ΔGauges/dials/instruments		0.0496***
		(0.0189)
ΔExperience using gauges/dials/instruments		0.0095
		(0.0161)
ΔCustomer communication		−0.0815***
		(0.0251)
ΔExperience using customer communication		0.0184
		(0.0194)
ΔOccupation index		0.1164***
		(0.0231)
ΔUnion		0.0948***
		(0.0288)
ΔFull-time		0.0361*
		(0.0188)
ΔUnemployment rate		−0.0009
		(0.0061)
Observations	1,844	1,822
R^2	0.0261	0.0825

Note: Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.

wage premium associated with computer skills disappears (both the immediate returns as well as the returns to computer usage experience). This finding suggests that the large and significant effects of computer skills observed in the cross-sectional results do not reflect the true return of computer skills (i.e., the productivity enhancing effect of computers in the workplace), but rather is a result of the job sorting process through which abler workers (i.e., workers with greater ability) are systematically selected into the jobs requiring computer skills. Unobserved but compensated characteristics of the workers matter.

This evidence contrasts with the common interpretation given to the results found in Krueger (1993) (albeit for a different population), that the computer-use wage differential reflects the true return to computer use or skill. These results highlight the importance of using longitudinal data to isolate the true return to job skills, which was difficult to address by Krueger (1993) or DiNardo and Pischke (1997) using only cross-sectional information on workers. Entorf, Gollac, and Kramarz (1999) find similar results for the effects of computer usage on wages using panel data in France. An explanation for the estimated negligible effects of computer skills in the first-difference fixed effect model could be the small number of workers changing status from nonuser to user during the sample period. However, since the standard errors do not increase much when we move from the cross-sectional to the fixed effect model, it appears that there is a sufficient number of workers changing status from nonusers to users in order to identify the effects of computer use on wages.

I also find that the first-difference estimates of the effects of having job responsibilities that include keeping a close watch over gauges, dials, or instruments of any kind are larger in magnitude and significance than the cross-sectional estimates. We see that when workers change from not having these job responsibilities to carrying out these job tasks on a daily basis, their wage increases immediately by 5%. When workers job task responsibilities change from requiring customer communication skills (i.e., daily direct communication between worker and customers/clients) to not requiring the use of these skills on a daily basis, their wage increases immediately by 8.2% (likely reasons for negative coefficient on customer communication skills discussed above). The return to experience using customer communication skills and the return to math skills are small and statistically insignificant.

The estimated effects of job skills are robust to the inclusion/exclusion of the change in occupation index measure, designed to control for occupation transitions. There is a strong significant relationship between change in occupation index and wage growth, as expected.

The effect of unionism remains in the longitudinal dimension, as the first-difference estimates show that when workers change union status from nonunionized to unionized (vice versa), they receive 9.5% higher (lower) pay on average. The first-difference fixed effect estimates reveal that changing from part-time to full-time (and vice versa) work hours increased wages immediately by 3.6%. The hours' effect appears to also operate through the flatter wage profile associated with part-time work experience. Changes in the local unemployment rate, which capture changes in local labor market demand conditions, had small and insignificant effects on wage growth after the inclusion of the work experience and job transition variables. As we will see in the job turnover analysis that follows, however, changes in local labor market demand conditions impact job transition patterns, which we have shown affect wage growth.

4.2.5. Double-Difference Model Results

It is possible that workers using a particular job skill that is associated with higher wage growth may have experienced greater wage growth in the absence of the use of that skill if unobserved fixed worker quality is driving the first-difference results. The second approach I use to evaluate the magnitude and significance of potential bias from unobserved heterogeneity involves estimating a double-difference equation. This procedure is equivalent to estimating the determinants of changes in wage growth rates (between Wave 1–2 vs. Wave 2–3 vs. Wave 3–4 vs. Wave 4–5) for a given worker to isolate the return to job skill. The general pattern of results from the double-difference model was similar to the first-difference results reported in Table 4 (results available from author upon request). In fact, the double-difference estimates indicate even larger effects of the usage of reading/writing skills on wage trajectories.

4.3. Job Turnover Analysis

The evidence presented in this paper has shown that jobs of different skill requirements differ in their prospects for wage growth. I now extend this analysis to study the effects of the skill requirements of jobs (via their effect on wage growth prospects) on job turnover. I am interested in the relationship between job transition patterns and wage growth. I have shown how average wage growth masks heterogeneity in within- and between-job wage growth. The first-difference estimates highlighted job mobility as a critical component of the wage growth process. This motivates the investigation of the determinants of job dynamics (job-to-job transitions – job mobility; job-to-nonemployment transitions – job instability) and the role of wage growth prospects in predicting job turnover. I also examine to what extent the worsening economic conditions in 2001–2002 affected job transition patterns.

4.3.1. Model of Job Turnover

Drawing on the key aspects of job search theory and human capital theory, I model job dynamics as on-the-job search with the wage offer distribution as the central factor that drives job transitions. Assume that while on the job, workers sample outside job offers in each period from a known wage offer distribution. Following Connolly and Gottschalk (2000), I extend the standard search model to include three key features of a job offer:

(1a) starting wage;
(2a) opportunity for wage growth on a job;
(2b) chance of upward mobility (promotion) within firm; and
(3) chance of job leading to job offers from superior wage offer distributions in future.

Assume workers have imperfect information about features (2a), (2b), and (3) of the job offer – workers learn about these characteristics of the job in the first several months of the job.

Other things being equal, we expect increases in job characteristics (1), (2a), and (2b) to reduce the hazard of leaving a job/employer, while we expect increases in job characteristic (3) to increase the hazard of leaving a job/employer.[26] Note, however, that job characteristic (3) is not observed by the econometrician. Assume individuals currently working in high wage-growth jobs, individuals working in jobs offering a high chance of upward mobility, and individuals working in jobs requiring more cognitive skills, are all more likely to receive job offers from the superior wage-offer distribution in the future.[27] This will result in a countervailing effect on the hazard of leaving a job/employer. For example, if skills acquired on a job become more valued in outside jobs/firms (than in the firm in which they are acquired), then high within-job wage growth could lead to higher quit rates.[28] Thus, high-learning jobs may or may not have lower quit rates than low-learning jobs –the expected effect is not clear as a matter of theory, it depends on which effect dominates in a particular type of job.[29]

The value of the present job depends on both the expected wage path and the uncertainty around that wage path. A central prediction of economic models of turnover is that, conditional on the present wage, quits will decline in the level of expected wage growth, and will increase in the value of outside opportunities. Factors that increased the present value of earnings on the job will be negatively associated with quits, while factors that increase the present value of earnings on alternative jobs and the arrival rate of alternative offers will be positively associated with quits holding the option value of jobs and the arrival rate of new information constant.

Within a search framework, local labor market conditions will affect the frequency and quality of job offers given a level of search intensity. Increases in the local availability of jobs will increase work by women through increases in the frequency of job offers and stability of employment. Labor market conditions affect wage levels and the probability of finding and keeping employment.

I use three different approaches to analyze job turnover. First, using the WES, I estimate a dependent competing-risks hazard model of job turnover, distinguishing between involuntary job separations (due to being laid-off/fired), voluntary job-to-job transitions, and employee-initiated job-to-nonemployment transitions. Drawing from the wage growth results above, I use the set of job skills/tasks to proxy for the effects of earnings growth prospects on job turnover. Given the results from the wage growth analysis of the large-and-significant returns to reading/writing skills, we would expect individuals working in jobs requiring reading/writing to have lower rates of job-to-nonemployment transitions, but potentially greater rates of job-to-job transitions (job mobility), if potential wage growth is an important factor affecting job turnover behavior. Similarly, we expect individuals working in jobs that require more cognitive skills to have lower layoff rates, and turnover rates overall, because these workers accumulate greater levels of firm-specific human capital as a result of greater OJT provision (Devereux, 2000).

It is difficult to sort out "person" from "job" effects using the first approach that analyzes job turnover because unobserved worker heterogeneity is not directly controlled. My second empirical strategy that analyzes job turnover involves estimation of a fixed-effect Cox proportional hazard model (known as the fixed-effect partial likelihood model, Chamberlain, 1985) using WES. I analyze the impact of explicitly taking into account unobserved heterogeneity on the propensity to change jobs.

The analysis using WES cannot distinguish between inter-firm and intra-firm job mobility for Waves 1–4 (Fall 1997–2001) because the WES survey questions on job tenure for these waves only refer to length of time worked in the position held, not employer tenure. However, for the period spanning Winter 2002–2004 when information was collected on employer tenure, I find that the lion's share of job-to-job transitions occurred between firms rather than promotion within firms. Thus, my third and final empirical strategy to analyze turnover utilizes MES. Using MES, I estimate a hazard model of worker-firm separations for the sample of jobs previously filled by former/current welfare recipients. The model includes direct measures of wage growth prospects (both chances of within-job pay raises and chances of promotion) from employer reports, as well as starting wage, whether job provides OJT, job skill requirements, employee and firm characteristics. The differences in specification between the WES and MES analyses of turnover due to the different variables at disposal in each data set, allow new and different insights from each analysis.

4.3.2. WES Job Transition Summary Statistics

The first sample I use in my analysis of job transitions consists of the 653 respondents that were employed at some point between Waves 1 and 5 of the WES. The 653 respondents held a total of 2,416 primary jobs over this period (Fall 1997 to Winter 2004). Of these jobs, 321 (13%) were censored because they were still in progress at the Wave 5 interview. Fifty-nine percent ($N = 1{,}418$) of the jobs ended in transitions to nonemployment, while

28% ($N = 677$) ended via voluntary job changes. Furthermore, the overwhelming majority of these voluntary job-to-job changes were between firms rather than due to promotion within the same employer. As observed in WES during the 2002–2004 period, less than 10% of women experienced promotions within the firm while working at their most recent employers.

In Table 5, I present means on overall monthly transitions out of jobs, as well as those into nonemployment and other jobs. The results are also shown separately by educational attainment. The results show that the monthly probability of a transition out of a job averages about 7.1% for the WES sample. The median job duration is seven months; only about a third (32.6%) of jobs lasted a year or more. Examining job turnover rates by education, we see significantly higher monthly transition rates out of jobs among high school dropouts relative to those with a high school diploma or GED, and especially higher turnover relative to those with some post-secondary education. These differences in turnover rates by education are driven by differences in the incidence of job-to-nonemployment transitions by education. Job-to-job transition rates do not differ significantly by

Table 5. Job Transition Summary Statistics by Education.

	All	Dropout	GED	HS	Post HS
Job turnover					
Monthly incidence rate	0.071	0.096	0.075	0.066	0.056
Duration of job (months):					
25th percentile	3	2	3	3	4
Median	7	5	7	9	10
75th percentile	17	12	14	17	23
One-year survival probability	0.326	0.240	0.292	0.354	0.411
Job-to-job turnover					
Monthly incidence rate	0.023	0.024	0.024	0.022	0.023
Duration of job (months):					
25th percentile	12	12	12	13	12
Median	28	26	26	27	30
75th percentile	63	56	76	80	61
One-year survival probability	0.739	0.735	0.702	0.756	0.739
Job-to-nonemployment turnover					
Monthly incidence rate	0.048	0.072	0.051	0.045	0.034
Duration of job (months):					
25th percentile	4	3	3	4	6
Median	10	7	10	12	16
75th percentile	32	19	28	34	56
One-year survival probability	0.450	0.337	0.426	0.477	0.564

education. Job transitions observed in the sample are disproportionately comprised of job-to-nonemployment transitions, as opposed to voluntary job-to-job transitions, which are associated with wage gains (see wage growth estimates).[30]

4.4. Dependent Competing-Risks Model of Job Turnover Using WES.

4.4.1. Specification

A competing-risks hazards model is used to distinguish between three types of job discontinuations: voluntary job-to-job mobility, employee-initiated job-to-nonemployment, and involuntary (employer-initiated) job-to-nonemployment transition. By using a competing-risks model of turnover, I can allow the determinants of job transitions to vary between job spells that end by a voluntary job change and those that end by a movement out of the workforce or end involuntarily (laid-off/fired). This allows me to test whether the variables I use to explain job duration have different effects on the propensity to leave jobs for different reasons.

Nearly all women in the sample experience more than one job spell over the observation period. The durations of jobs contributed by the same woman may be correlated because of unobserved individual characteristics that influence the duration of each of a woman's job spells. If ignored, the correlation between observations may lead to underestimation of standard errors owing to a reduction in the effective sample size. Random effects are therefore incorporated in the model to allow for unobserved heterogeneity between women. These random effects are defined at the individual level and represent unobserved individual-level characteristics that influence the hazard of a job ending at each month of a given job spell, and for each job spell.

I analyze determinants of monthly job transitions in the three-way competing-risks framework (where risks are voluntary job change, involuntary job separation, and employee-initiated movement to nonwork) using a multinomial probit specification of the hazard. Specifically, I specify the hazard – i.e., the probability that woman i leaves job j for reason $r = 1, 2, 3$, during month t, given that the woman has remained in the job the previous $(t-1)$ months—as

$$h_{ijr}(t) = \beta_r \mathbf{X}_{ij}(t) + \Gamma_r \text{TenureDummies} + u_{ir} + \varepsilon_{ijr}(t) \quad (r = 1, 2, 3) \tag{4}$$

In this model, the variables in $\mathbf{X}$ are individual-specific and do not differ across alternatives. Estimated coefficients will therefore represent the effect of a given variable on the value of new job relative to its effect on the value on the current job, or the effect of this variable on the value of nonemployment relative to its effect on the value of the current job. The level of **X** is set at wave T for months between wave T and $(T+1)$, $(T = 1, 2, 3, 4)$, for all variables except the county monthly unemployment rate, which corresponds with the observation month, and the set of job task variables which correspond with the job tasks performed on jobs held between wave T and $(T+1)$. I use a similar set of explanatory variables (vector **X**) as was used in the WES wage model. In order not to place restrictions on the functional form of the relationship between job tenure and the hazard of leaving a job, I enter 10 dummy variables consisting of eight monthly dummies for the first eight months, a dummy for months 9–12, and an annual variable for year two (tenure greater than two years is the reference category). Woman-level unobserved variables are represented by a woman-specific random effect, u_{ir}, which is assumed to follow a normal distribution with zero mean.

A common yet very restrictive assumption in the analysis of competing risks is to assume that the latent survival times are independent, conditional on covariates. In this context, this involves treating the woman-specific random effects, $\boldsymbol{u}_i = (u_{i1}, u_{i2}, u_{i3})$, as uncorrelated across the different types of discontinuation. This means that a woman with a higher propensity toward leaving a job for reason r does not tell us anything about her propensity toward leaving a job for any of the other reasons. This assumption of Independence of Irrelevant Alternatives (IIA), however, is unlikely to be true if certain characteristics of job transition types make two of them more similar than the others. Dependency between competing risks and shared unobserved risk factors may be accommodated by permitting the random effects $\boldsymbol{u}_i$ to be correlated across different types of discontinuation.

In light of these considerations, a multinomial probit specification is utilized to allow for flexible correlation structures across alternatives. An assumption of joint normality on the errors (trivariate normal distribution – (u_{i1},u_{i2},u_{i3}) in (1) implies a multinomial probit for the estimation of these turnover equations. The residual terms $\varepsilon_{ijr}(t)$ are assumed to be uncorrelated and to follow standard normal distributions. (footnote: The estimation was carried out using full-information likelihood, as implemented in aML (see Lillard and Panis (2000) for details of the estimation procedure).

I first estimate this version of the model of job transitions, which specifies unobserved worker heterogeneity as a random effect, and then estimate a fixed-effect Cox proportional hazard model (discussed in the next section) to examine the effects of explicitly controlling for unobserved heterogeneity on the propensity to leave jobs.

I have tested the hypothesis that the three-way dependent competing-risks model is inappropriate and that a simple duration model or a two-way independent competing-risks model is preferred, which does not distinguish between involuntary job changes and employee-initiated movements out of the workforce, or voluntary job changes, and/or the dependency between the types of discontinuation. The data reject these options. I highlight a few of the results of the competing-risks model of *monthly* job transitions below.[31]

4.4.2. Competing-Risks Hazard Results

The competing-risks hazard estimates are displayed in Table 6. An inspection of the correlation coefficients of the multinomial probit shows the relevance of the multinomial probit specification of the hazard in correctly estimating the probability of voluntary and involuntary job discontinuation. Recall that the woman-specific random effects allow for the possibility that some component of the unobservable value of a new job or of nonemployment may persist over time for the same individual. Table 6 presents the estimated standard deviations of the woman-specific random effects in the hazards equations for the different types of job discontinuation and the correlations between these random effects. These results provide evidence that indicate that time-persistent individual unobservables and the dependency between risks are both important in these job turnover equations. The estimated standard deviations of each random effect and the correlations between them are significantly different from zero. The economic interpretation of the correlations across job discontinuation types is that the worker differentiates between the alternative routes of discontinuation, and that women with an above-average probability of discontinuation via voluntary job-to-job changes tend also to have a below-average probability of discontinuation via job-to-nonemployment transitions (either voluntarily or through being laid-off/fired). This is an important result that emerges from the richer model, which violates the restrictive assumptions of independent competing-risks models of turnover that have been used in previous research, and suggests ignoring unobservable correlations across alternatives may lead to erroneous inferences of the determinants of job dynamics.

Both the involuntary job separation hazard and employee-initiated job-to-nonemployment hazard remain high through the first seven months of the job before gradually declining over the remainder of the job spell. This familiar pattern of the hazard has been observed in previous work (see, for example, Farber (1998), or Holzer and LaLonde (2000)). The process of gaining information about the quality of the job match early in jobs, as well as worker heterogeneity, are common explanations of the pattern. The

Table 6. Dependent Competing-Risks Hazard Model of Job Turnover (MNP) with Random Effects.

Explanatory Variables	Multinomial Probit Coefficient Estimates		
	Involuntary job separation	EE-initiated job-to-nonemployment transition	Voluntary job-to-job transition
	(1)	(2)	(3)
Job tenure (reference category: Tenure >2 years)			
Month 1	0.3669 ***	0.1040	0.0036
	(0.1008)	(0.0651)	(0.0658)
Month 2	0.5469 ***	0.2354 ***	−0.1642 **
	(0.1070)	(0.0654)	(0.0749)
Month 3	0.4572 ***	0.2303 ***	−0.1535 **
	(0.1051)	(0.0696)	(0.0755)
Month 4	0.3947 ***	0.2809 ***	−0.2726 ***
	(0.1105)	(0.0688)	(0.0804)
Month 5	0.3233 ***	0.2070 ***	−0.1631 **
	(0.1230)	(0.0733)	(0.0803)
Month 6	0.4562 ***	0.1416 *	−0.0426
	(0.1191)	(0.0748)	(0.0793)
Month 7	0.3964 ***	0.1808 **	−0.3814 ***
	(0.1232)	(0.0787)	(0.0962)
Month 8	−0.0397	0.0558	−0.1302
	(0.1581)	(0.0929)	(0.0981)
Months 9–12	0.2965 ***	0.2250 ***	0.1344 **
	(0.0892)	(0.0531)	(0.0540)
Year 2	0.2397 ***	0.0679	0.0469
	(0.0891)	(0.0514)	(0.0449)
Labor market demand conditions			
Unemployment rate	0.0621 ***	−0.0241 **	0.0258 ***
	(0.0164)	(0.0095)	(0.0095)
Human capital variables			
HS Grad/GED (reference category: HS dropout)	−0.3175 ***	−0.1155 *	−0.0336
	(0.1212)	(0.0594)	(0.0474)
Some post-secondary education	−0.4121 ***	−0.2163 ***	−0.0062
	(0.1484)	(0.0668)	(0.0522)
Years of full-time work experience	−0.0143	−0.0112	0.0102 **
	(0.0164)	(0.0077)	(0.0046)
Years of part-time work experience	−0.0115	−0.0022	0.0006
	(0.0200)	(0.0081)	(0.0055)
Job skill variables			
Reading/writing	0.0056	−0.0644 *	0.0802 *
	(0.0701)	(0.0408)	(0.0425)

Table 6. (*Continued*)

Explanatory Variables	Multinomial Probit Coefficient Estimates		
	Involuntary job separation	EE-initiated job-to-nonemployment transition	Voluntary job-to-job transition
	(1)	(2)	(3)
Computer	−0.0395	0.0613	−0.0019
	(0.0737)	(0.0435)	(0.0454)
Math	0.2594 ***	0.0077	0.0534
	(0.0727)	(0.0408)	(0.0471)
Gauges/dials/instruments	−0.0763	0.0182	−0.0054
	(0.0610)	(0.0351)	(0.0362)
Customer communication	−0.2981 ***	−0.0812 *	−0.0124
	(0.0723)	(0.0423)	(0.0526)
Supervise co-workers	−0.1548 **	−0.0802 **	−0.0820 **
	(0.0673)	(0.0401)	(0.0404)
Other job characteristics			
Ln(Wage)	0.1406	0.0229	−0.1513 **
	(0.0982)	(0.0552)	(0.0591)
Health insurance	0.0408	−0.1026 ***	0.0085
	(0.0558)	(0.0348)	(0.0381)
Full-time	0.0329	−0.0461	−0.0124
	(0.0603)	(0.0393)	(0.0426)
Occupation (Reference category: Service)			
Professional/managerial/ technical	−0.0172	0.0585	−0.0746
	(0.1258)	(0.0710)	(0.0679)
Sales	−0.0616	0.0530	0.0136
	(0.0902)	(0.0527)	(0.0494)
Clerical	0.1266	−0.0383	0.0263
	(0.1099)	(0.0658)	(0.0636)
Operator	0.2848 ***	0.1604 ***	−0.1032
	(0.0930)	(0.0584)	(0.0644)
Craft	0.0694	0.1873	−0.1973
	(0.2106)	(0.1345)	(0.1321)
Laborer	0.2074	0.1654 **	0.0830
	(0.1268)	(0.0719)	(0.0920)
Demographic variables			
Black	0.1332	0.0219	0.0388
	(0.1317)	(0.0601)	(0.0428)
Age 25–34	−0.0628	−0.0972	−0.0514
	(0.1559)	(0.0725)	(0.0518)
Age ≥ 35	−0.0438	−0.2025 **	−0.1490 **
	(0.2221)	(0.0981)	(0.0731)

Table 6. (*Continued*)

Explanatory Variables	Multinomial Probit Coefficient Estimates		
	Involuntary job separation	EE-initiated job-to-nonemployment transition	Voluntary job-to-job transition
	(1)	(2)	(3)
Married/cohabiting	−0.0132	0.0543	0.0058
	(0.0789)	(0.0411)	(0.0402)
Child 0–2 years	−0.0009	0.0459	0.1320 **
	(0.0771)	(0.0437)	(0.0517)
Child 3–5 years	0.1098 *	−0.0887 ***	0.0394
	(0.0612)	(0.0339)	(0.0425)
Health-related variables			
Pregnant	−0.0090	0.2388 ***	−0.1881 ***
	(0.0968)	(0.0486)	(0.0632)
Work-limiting (physical) health condition	0.0709	0.0922 **	−0.0567
	(0.0727)	(0.0417)	(0.0473)
Child health problems	0.0648	0.0658	0.0582
	(0.0822)	(0.0467)	(0.0572)
Mental health condition	0.1950 ***	0.2133 ***	0.0496
	(0.0654)	(0.0404)	(0.0446)
Domestic violence (past year)	0.0755	0.0889 **	−0.0538
	(0.0748)	(0.0445)	(0.0520)
Lack access to a car	0.2856 ***	0.2170 ***	−0.0214
	(0.0688)	(0.0452)	(0.0432)
Constant	−3.9924 ***	−1.6566 ***	−2.0815 ***
	(0.3118)	(0.1562)	(0.1601)
Standard Deviations and pairwise correlations for woman-level random effects			
Involuntary job separation equation	1.0494 ***		
	(0.1170)		
EE-initiated job-to-nonemployment equation	−0.1347 *	0.5079 ***	
	(0.0795)	(0.0295)	
Voluntary job-to-job equation	−0.3323 **	−0.4207 **	0.1409 ***
	(0.1550)	(0.1765)	(0.0419)

Note: Asymptotic standard errors in parentheses. Significance: *10%;**5%; ***1%.

job-to-job hazard follows a noticeably different pattern as it declines gradually through the seventh month, before increasing significantly between months nine through twelve, and declining thereafter. This pattern may be the result of the fact that the majority of the jobs this less-educated sample of women is able to obtain lack career ladders and/or provide limited learning opportunities that can increase wages, and thus for them, job changes are a more important source of wage growth than for other workers (see wage growth estimates).

One of the most insightful results of the turnover analysis is the sensitivity of these women's job transition patterns to changes in labor market demand conditions. The results indicate significant effects of the monthly unemployment rate, which is used as a measure of local labor market demand conditions. We find that a one percentage-point increase in the local unemployment rate increases the hazard of being laid-off/fired by about 12%. On the other hand, a one percentage-point increase in the local unemployment rate decreases the probability of an employee-initiated transition into nonemployment by about 4%. The differential effect of the unemployment rate by type of job separation is expected. One reason for the latter result is that it decreases in job availability increases the costs of job-to-nonemployment transitions (or, alternatively stated, increase the value of maintaining employment) by decreasing the expected (monthly) re-employment probability.

The results from the analysis indicate that individuals with lower levels of education have higher transition rates out of jobs. By distinguishing transition rates from jobs into nonemployment from transitions to new jobs, I find that the higher job transition rates for the least-educated individuals result primarily from higher rates of both involuntary job separations and employee-initiated job-to-nonemployment transitions.

Individuals working in jobs requiring reading/writing on a daily basis are significantly more likely to experience voluntary job changes (job mobility), which are associated with wage gains, and have significantly lower transition rates into nonemployment. My previous analyses of the determinants of wage growth have revealed that individuals working in jobs requiring reading/writing on a daily basis experienced significantly higher wage growth (as well as wage levels) between waves across all job transition types (job stability, job mobility, and job instability), and that a primary route of advancement was through changing jobs. Thus, the present evidence of lower quit rates into nonemployment and higher job-to-job transition rates among individuals working in jobs requiring reading/writing (on the order of about 15%) is consistent with the following story. Individuals working in jobs requiring more cognitive skills and in jobs providing more learning opportunities, and thus more wage growth, are also more likely to receive job offers from superior wage offer distributions in the future (controlling for the wage). This has the effect of both reducing the likelihood of voluntary transitions out of the labor force and increasing the hazard of voluntarily leaving the current job for another.

Individuals working in blue-collar occupations (operator/laborer) were more likely to experience involuntary job separations. Full-time work experience is positively associated with job-to-job transitions (job mobility), while part-time work experience has insignificant effects on both job-to-nonemployment and job-to-job transition rates. The wage of the job as of the most recent wave is negatively associated with job-to-job transitions. Individuals working in jobs providing employer-sponsored health insurance have lower rates of employee-initiated job-to-nonemployment transitions.

Given the relatively high prevalence of health-related conditions among this sample of women, I also include a set of health-related variables in the model to attempt to better understand the causes of the high incidence of job instability. (The sample means for these health-related measures are displayed in the last set of rows of the first column of Appendix Table B1). The results indicate that individuals with physical health limitations had higher job-to-nonemployment transition rates. The results also indicate that individuals with mental health conditions, mothers with children who have health problems, and women who suffered domestic violence within the past year, had higher rates of job-to-nonemployment transitions. As expected, child bearing is a significant predictor of job turnover as we see that being pregnant and having pre-school aged children (0–2 years old) each significantly increases rates of job turnover.

4.5. Fixed-Effect Cox Proportional Hazard Model Using WES

4.5.1. Empirical Strategy

Do jobs (as opposed to workers in them) have different turnover behavior? It is very difficult to sort out "person" from "job" effects in the above analysis of job turnover. With wages and other characteristics held constant, individuals working in jobs requiring particular job skills are shown to have significantly lower job turnover rates than individuals working in jobs not requiring these job skills. Why? There are two possible explanations: (1) jobs requiring more cognitive skills reduce worker's propensity to quit the job by providing greater learning opportunities (human capital investment opportunities – training (informal/formal)), thereby offering more potential to experience within-job wage growth; (2) the job turnover – job skills relationship reflects a selection effect whereby workers are sorted by ability resulting in unobserved worker quality differences across jobs of different skill requirements (e.g., underlying unobserved heterogeneity among workers affecting the propensity to quit, such as "stick-to-it-iveness"). The analysis below seeks to disentangle these two possible causes – i.e., whether the skill requirements of jobs affect job turnover behavior of workers (job-specific effect), or whether differences in job turnover rates across jobs of different skill requirements are being driven by unobserved worker characteristics (i.e., person-specific effects are observed indirectly via the types of jobs individuals hold).

My empirical strategy involves exploiting the longitudinal dimension of the WES data to estimate a fixed-effect Cox proportional hazard model (known as the fixed-effect partial likelihood approach Chamberlain, 1985) of job turnover, distinguishing between involuntary job separations, voluntary job-to-job transitions, and employee-initiated job-to-nonemployment transitions. I have information on two or more complete job spells for almost all of the WES respondents (so selection bias should not be a concern), along with job skills used during periods in which different jobs were held. In essence, the fixed effect partial likelihood approach uses only information about the rank ordering of job spell lengths within individuals, and asks how that ordering may depend on variations in the explanatory variables. There is a significant amount of within-person changes across the periods in job skills used.[32] Use of Cox's fixed-effect partial likelihood approach eliminates all individual-specific factors (and thus the selectivity bias) by comparing job turnover behavior of the same worker in jobs held of differing skill requirements, thereby isolating the behavioral impact of the skill content of jobs. I relate the results from the turnover analysis with those from the wage growth analysis, and I use the job skill requirements to proxy the role of a worker's wage growth prospects in predicting turnover.

To be more specific, suppose that for worker i we have n_i spells (ordered by their increasing length) and that the duration for each spell is denoted t_{ij}, where j stands for the spell number. Assuming all spells for the same person are independently distributed given her heterogeneity parameter, I can write the hazard functions as

$$\lambda_{ij}(t) = \exp(\beta' \mathbf{X}_{ij}(t) + a_i)\lambda_{i0}(t),\ j = 1, \ldots, n_i; i = 1, \ldots, N \tag{5}$$

Then it can be shown that the partial log-likelihood function is equal to[33]

$$L_p = \sum_{i=1}^{N} \sum_{j=1}^{n_{i-1}} \ln \left(\frac{\exp\left(\beta' \mathbf{X}_{i(y)}(t_{i(j)})\right)}{\sum_{k=j}^{n_i} \exp\left(\beta' \mathbf{X}_{i(k)}\left(t_{i(j)}\right)\right)} \right) \tag{6}$$

where the denominator corresponds to the risk set of worker *i*. Note that both α_i and λ_{i0} do not appear in Eq. (6). Although all biases caused by unobserved individual heterogeneity are removed using Chamberlain's extension of Cox's partial likelihood method, the problem of biases caused by unobserved job-match heterogeneity remains.

4.5.2. Fixed-Effect Hazard Results

The exponentiated coefficients from the fixed-effect Cox proportional hazard model are presented in Table 7. I find that workers' probability of being laid-off/fired is 26% lower when working in jobs requiring reading/writing and computer skills on a daily basis relative to their probability while working in jobs not requiring these skills.[34] On the other hand, workers' probability of being laid-off/fired is significantly higher when working in jobs requiring use of arithmetic (including making change) on a daily basis relative to their probability while working in jobs not requiring these skills. These results are consistent with employer incentives to concentrate layoffs on workers with the lowest levels of firm-specific human capital (Devereux, 2000). The previous findings using MES documented significantly greater provision of OJT opportunities in jobs requiring reading/writing, which may signal greater firm investments in the worker and enable workers in these jobs to accumulate greater levels of firm-specific human capital.

The results also indicate that workers' probability of voluntary movements out of the workforce is significantly less likely when individuals work in jobs requiring reading/writing skills, which serve to proxy wage growth opportunities. The results also indicate that when individuals work in jobs that require supervisory responsibilities, their job turnover rates decline relative to their turnover rates when working in jobs not requiring these responsibilities. This finding is consistent with predictions from the theory of turnover, if skills accumulated on the job that lead to supervisory responsibilities are fairly firm-specific.

The estimated effects of labor market demand conditions are robust to explicit controls for unobserved individual heterogeneity. The economic downturn significantly increased workers' probability of being laid-off/fired; and significantly reduced workers' quit rate into nonemployment, indicating that workers are less likely to quit if jobs are scarce in their local community. We find that a one percentage-point increase in the local unemployment rate increases the hazard of being laid-off/fired by 7.7%, and decreases the probability of an employee-initiated transition into nonemployment by 7.1%.

Table 7. Exponentiated Coefficients from Fixed-Effect Cox Proportional Hazard Model of Job Turnover Using WES.

Explanatory Variables	Job Turnover: With Control for Heterogeneity (Workers w/2 or more Spells)		
	Involuntary job separation	EE-initiated job-to-nonemployment transition	Voluntary job-to-job transition
	(1)	(2)	(3)
Job skill variables			
Reading/writing	0.7407*	0.8382*	1.1225
	(0.1607)	(0.1036)	(0.1783)
Computer	0.7236*	1.2916*	1.0178
	(0.1664)	(0.1711)	(0.1475)
Math	2.0219***	1.1118	1.0124
	(0.5228)	(0.1420)	(0.1500)
Gauges/dials/instruments	0.8082	0.9947	1.0170
	(0.1638)	(0.1098)	(0.1433)
Customer communication	0.7477	1.0184	0.9735
	(0.1849)	(0.1471)	(0.1828)
Supervise co-workers	0.8081	0.7547**	0.7927
	(0.1907)	(0.0915)	(0.1279)
Labor market demand conditions			
Unemployment rate	1.0769*	0.9292**	1.0051
	(0.0566)	(0.0287)	(0.0374)
Other job characteristics			
Ln(Wage)	1.3882	1.1095	0.6367**
	(0.5997)	(0.2180)	(0.1208)
Health insurance	1.2582	0.8817	0.9218
	(0.2455)	(0.0998)	(0.1222)
Union	1.3421	0.8855	0.8785
	(0.3920)	(0.1796)	(0.1633)
Full-time	1.0814	0.8791	0.8827
	(0.2413)	(0.1020)	(0.1187)
Human capital variables			
Years of full-time work experience	2.4355***	2.5440***	1.1486
	(0.7274)	(0.3924)	(0.1960)
Years of part-time work experience	2.3733***	2.6084***	1.0965
	(0.7600)	(0.4165)	(0.2198)
Pregnant	0.9580	1.2365*	0.7277*
	(0.2208)	(0.1502)	(0.1353)
Log-likelihood	−239.2687	−989.3010	−492.3998
Observations	29,485	29,485	29,485
Subjects (Jobs)	2,415	2,415	2,415
Failures	289	1,128	677

Robust standard errors in parentheses. * Significant at 10% (one-tailed test); ** significant at 5%; *** significant at 1%.

4.6. Results from Job Turnover Analysis using MES

The results from the hazard model of worker-firm separations using MES are presented in Table 8. The overall pattern of the MES turnover results are similar to those yielded using WES. The results indicate that, controlling for starting wages, jobs that offer greater wage growth opportunities have significantly lower turnover rates. In particular, jobs that provide chances for merit-based within-job pay raises (above cost of living increases), jobs that have good or excellent chances of promotion upward mobility (assuming good performance), and jobs requiring reading/writing skills on a daily basis, all have substantially lower turnover rates in a given week. Jobs that provide OJT opportunities also had significantly lower turnover rates in a given week – these effects became statistically insignificant only after the inclusion of the variables capturing employer reports of wage growth prospects of the job. On the other hand, individuals working in jobs requiring computer skills and individuals with prior occupation-specific work experience have significantly higher turnover rates than individuals that do not possess these skills or experience. These results are not inconsistent, however, with results from the previous analyses above, since acquiring computer skills and experience may enable individuals to sample from better (outside) wage offer distributions in the future, thereby increasing job mobility. The effects of the other job skills are statistically insignificant. As was found using WES, high school graduates had significantly lower turnover rates than high school dropouts. The results indicate, as expected, that having work performance-related problems (either problems with absenteeism, work attitude, or job skills) significantly increases the hazard of worker-firm separations.

After inclusion of the starting wage, wage growth, and job skill variables, the effects of occupation variables became insignificant. The MES results indicate that jobs that provided employer-sponsored health insurance benefits had significantly lower turnover rates. This finding is not necessarily inconsistent with the previous finding that women in the WES sample were more likely to experience involuntary job separations when working in jobs that offer employer-provided health benefits, because we cannot disaggregate employer-initiated and employee-initiated job separations in the MES turnover analysis. Somewhat surprisingly, part-time jobs did not have significantly higher turnover rates after the inclusion of the variables that affect wage growth prospects.

The MES results indicate that firms that were neither within 0.3 miles of public transit nor within 30 min of downtown, had significantly higher rates

Table 8. Worker-Firm Separation Hazard Estimates Using MES.

Explanatory Variables	Worker-Firm Separation
Job characteristics	
Starting wage	0.1431**
	(0.0639)
Health insurance	−0.6451**
	(0.2864)
Offers chance of within-job pay raise (assuming good performance)	−0.9819***
	(0.2890)
Offers good/excellent promotion prospects (assuming good performance)	−0.6478***
	(0.2753)
Offers on-the-job training	−0.2879
	(0.2725)
Part-time	0.1275
	(0.3145)
Job skill variables	
Reading/writing	−0.5434**
	(0.2517)
Computer	0.6136**
	(0.3154)
Math	0.1973
	(0.2767)
Customer communication	−0.1084
	(0.3594)
Occupation (Reference category: Service)	
Sales	0.1831
	(0.3832)
Clerical	−0.0354
	(0.3893)
Blue-collar	0.0739
	(0.4817)
Employee characteristics	
High school Diploma/GED	−0.6451**
	(0.3140)
Prior occupation-specific work experience	0.4886*
	(0.2581)
Training/skill certification	−0.2691
	(0.2512)
Work performance-related problem	0.7677***
	(0.2445)
Firm characteristics	
% Employees unionized	−0.0079*
	(0.0054)

Table 8. (*Continued*)

Explanatory Variables	Worker-Firm Separation
Firm size (Reference Category: $\geq$100 Employees)	
1–19 employees	0.5853*
	(0.3466)
20–99 employees	0.4446*
	(0.3213)
Not within 0.3 miles of public transit nor within 30 min of downtown	0.3042*
	(0.2424)
Job tenure	
Ln(Tenure)	0.5644*
	(0.4543)
$(\text{Ln(Tenure)})^2$	−0.1175
	(0.1369)
Log-likelihood	−396.8297
Observations	3,694
Subjects (Jobs)	418
Failures	106

Note: Regressions also include metropolitan area dummies and a constant term. Robust standard errors are in parentheses.

*Statistically significant at the 0.10 level (one-tailed test).

**Statistically significant at the 0.05 level.

***Statistically significant at the 0.01 level.

of job turnover. This result suggests that job accessibility may play a role in predicting job turnover. The results on the other variables indicate that large firms ($\geq$100 employees) and firms with greater fractions of unionized employees have lower job turnover rates.

4.6.1. Multinomial Endogenous Switching Model of Wage Growth

A final aim of this paper is to analyze wage differentials between job stayers and voluntary and involuntary job movers after taking into account the endogeneity of these job mobility decisions. The final empirical specification involves the estimation of a multinomial endogenous switching model of wage growth to attempt to address the endogeneity between job transitions and wages/wage growth. In particular, I specify a multinomial switching regression model, which allows the joint estimation of a quadrivariate selection process that accounts for the type of job transition and four wage change equations conditional on each type of transition with the appropriate selection corrections. These estimates are then used to predict a woman's change in wages for the four potential job transition types – job stability, voluntary job mobility, employee-initiated job instability, and involuntary (employer-initiated) job instability. By comparing potential wage growth in each transition type, I am able to estimate the relative returns of job stability, mobility, and the costs of having a spell of nonemployment. I also investigate whether wage differentials with job transitions vary significantly by job skill requirements.

The details of the endogenous switching model estimation procedures follow the work of Garcia-Perez and Sanz (2004) and are contained in Appendix B. The first-stage selection process of the type of job transition experienced between waves is specified as a multinomial probit model. The estimation of the model is highly computationally intensive and is estimated using aML. The exclusion restrictions used for identification of the model involve the exclusion of the following variables from the wage change equation: the presence of pre-school aged children, marital status, whether became pregnant over the year, whether experienced domestic violence over the year, mental or physical health conditions, and whether any children with a health condition.

I estimate four wage change equations, one for each job transition type, to allow the marginal effects of the explanatory variables of the woman's wage growth to depend on the type of job transition. For example, the effects of changes in work experience reflect the sum of the returns to experience and returns to tenure for individuals who experienced job stability; and they capture the sum of the returns to experience and the change in the job match component for individuals who experienced the relevant type of job change. To test for the endogeneity of the switching model, the parameters of interest are the covariances of the error term of each wage change equation with the error term of the selection equations.

The results are presented in Tables 9 and 10.[35] I first inspected the results of the correlation structure of the error terms and the likelihood ratio test for the endogenous switching model with respect to the exogenous one, which is a restricted case. The likelihood ratio test together with the correlation parameters reveal that there is evidence of non-random selection, and thus, if we omit the effects of unobservables, predicted wage growth with job mobility, job stability, and the wage penalty with job instability would be inconsistently estimated.

Table 10 presents predicted wage returns to voluntary job mobility, evaluated at sample means and evaluated at different levels of selected job skill variables. A comparison of the results previously presented with those shown in Table 10 reveal that if we do not consider the self-selection

Table 9. Endogenous Switching Model Estimates of Wage Growth.

	Dependent Variable: First-Difference of Log of Real Hourly Wages ($1999)			
	Wage change equation w/job stability	Wage change equation w/vol. job mobility	Wage change equation w/ee-initiated job instability	Wage change equation w/ involuntary job instability
ΔFull-time work experience	0.0004	−0.0353	0.0337	−0.0136
	(0.0407)	(0.0415)	(0.0274)	(0.0485)
ΔPart-time work experience	−0.0264	−0.0569	0.0358	−0.0373
	(0.0448)	(0.0441)	(0.0348)	(0.0570)
ΔReading/writing	0.0461*	0.0732**	−0.0148	0.0583*
	(0.0290)	(0.0332)	(0.0243)	(0.0367)
ΔExperience using reading/writing	0.0459	0.0773*	0.0233	0.0184
	(0.0442)	(0.0513)	(0.0398)	(0.0562)
ΔComputer	−0.0100	−0.0219	0.0332	0.0231
	(0.0377)	(0.0360)	(0.0274)	(0.0430)
ΔExperience using computer	0.0240	0.0396	−0.0642	−0.0388
	(0.0550)	(0.0746)	(0.0452)	(0.0722)
ΔMath	0.0867**	0.0128	0.0257	−0.0194
	(0.0354)	(0.0425)	(0.0269)	(0.0407)
ΔExperience using math	−0.0038	−0.0262	0.0323	0.0057
	(0.0417)	(0.0500)	(0.0335)	(0.0523)
ΔGauges/dials/ instruments	−0.0262	0.0296	0.1149***	0.0027
	(0.0270)	(0.0343)	(0.0207)	(0.0420)
ΔExperience using gauges/dials/ instruments	0.0121	0.0671	0.0133	−0.0215
	(0.0447)	(0.0545)	(0.0393)	(0.0552)
ΔCustomer communication	−0.0168	−0.0523	−0.1154***	−0.0972**
	(0.0390)	(0.0448)	(0.0263)	(0.0391)
ΔExperience using customer communication	0.0010	0.0143	0.0660**	0.0381
	(0.0473)	(0.0506)	(0.0293)	(0.0537)
ΔSupervise co-workers	0.0301	−0.0096	−0.0458*	−0.0435
	(0.0337)	(0.0407)	(0.0244)	(0.0434)
ΔExperience supervising co-workers	−0.0324	0.0057	−0.0414	−0.0407
	(0.0496)	(0.0931)	(0.0493)	(0.0939)
ΔOccupation index		0.0959***	0.1250***	0.1350***
		(0.0331)	(0.0281)	(0.0319)
ΔUnion		0.0248	0.1646***	0.1284**
		(0.0464)	(0.0345)	(0.0544)
ΔFull-time	−0.0512	0.0601	0.0362	0.0530
	(0.0340)	(0.0398)	(0.0238)	(0.0379)
	0.0131	0.0098	−0.0034	−0.0053

Table 9. *(Continued)*

	Dependent Variable: First-Difference of Log of Real Hourly Wages ($1999)			
	Wage change equation w/job stability	Wage change equation w/vol. job mobility	Wage change equation w/ee-initiated job instability	Wage change equation w/ involuntary job instability
ΔUnemployment rate	(0.0156)	(0.0147)	(0.0070)	(0.0149)
Constant	−0.0359	0.1933*	0.0968	−0.0739
	(0.0672)	0.1057)	(0.0712)	(0.1097)
Correlations of error terms across job turnover (T.O.) and wage change equations				
Correlation (T.O. mobility equation, wage J-stability equation)	−0.1574			
	(0.3508)			
Correlation (T.O. EE instability equation, wage J-stability equation)	−0.4634***			
	(0.1727)			
Correlation (T.O. layoff equation, wage J-stability equation)	0.4692**			
	(0.2112)			
Correlation (T.O. mobility equation, wage J-mobility equation)		−0.4158***		
		(0.1486)		
Correlation (T.O. EE instability equation, wage J-mobility equation)		−0.6009***		
		(0.1797)		
Correlation (T.O. layoff equation, wage J-mobility equation)		0.3543		
		(0.2854)		
Correlation (T.O. mobility equation, wage EE Instability equation)			−0.6304**	
			(0.2949)	

Table 9. (*Continued*)

	Dependent Variable: First-Difference of Log of Real Hourly Wages ($1999)			
	Wage change equation w/job stability	Wage change equation w/vol. job mobility	Wage change equation w/ee-initiated job instability	Wage change equation w/ involuntary job instability
Correlation (T.O. EE instability equation, wage EE Instability equation)			−0.6020*** (0.1089)	
Correlation (T.O. Layoff equation, wage EE instability equation)			−0.0033 (0.3032)	
Correlation (T.O. mobility equation, wage layoff equation)				−0.3491 (0.6143)
Correlation (T.O. EE instbility equation, wage layoff equation)				−0.8710*** (0.1662)
Correlation (T.O. layoff equation, wage layoff equation)				−0.1527 (0.2149)
Log-likelihood	−2421.13	−2476.61	−2611.21	−2341.84

Note: Asymptotic Standard Errors in Parentheses. Significance: ***1%; **5%; *10% (one-tailed test).

problem we will considerably underestimate the wage returns to job mobility. The results indicate the estimated wage differentials are largest when we use involuntary job instability as the comparison group, as we find wage returns to mobility of 29.5%. Furthermore, workers who experience voluntary job changes without intervening spells of nonemployment earn around 22.9% more than if they had stayed at the same job. After accounting for differences in work experience accumulated over the period, there are not significant wage differences between voluntary job mobility and employee-initiated job instability.

Table 10. Multinomial Endogenous Switching Model Estimates of Wage Return to Voluntary Job Mobility: Evaluated at Sample Mean and at Different Levels of Selected Job Skill Variables.

		Wage Returns to Voluntary Job Mobility
All evaluated at sample means	*Counterfactual*	
	Job stability	0.2291
	EE-initiated job instability	0.0155
	Involuntary job instability	0.2950
Job skills		
Additional years of experience using reading/writing	Job stability	0.2461
	EE-initiated job instability	0.0426
	Involuntary job instability[a]	0.4052
No use of reading/writing on job	Job stability	0.2147
	EE-initiated job instability	−0.0114
	Involuntary job instability	0.2696

[a]This involuntary job instability counterfactual estimate assumes the woman is unable to secure a job requiring reading/writing skills immediately following layoff. Other variables were held at sample means.

These wage differentials with job transitions between waves vary significantly by the skill content of work experience, in much the same ways that the previous analyses have shown. In particular, wage returns to mobility (relative to job stability and instability) tend to be the largest for jobs requiring the use of reading/writing. The results indicate that when an individual's job-to-job changes involve the accumulation of additional experience using reading/writing skills, wage returns are 24.6% higher than returns experienced by continual usage of those job skills while holding the same job. On the other hand, the wage penalty is 40.5% for being laid-off or fired from a job that required reading/writing skills and failing to secure a job requiring these skills following the lay-off (relative to the wage gains with job mobility while accumulating experience using these skills).

5. SUMMARY AND CONCLUSION

In this paper, I used survey data from employers and longitudinal data from former/current welfare recipients covering the period 1997-early 2004 to analyze the relationship between job skills, job changes, and the evolution of wages. The results indicate that average wage growth masks considerable heterogeneity in within- and between-job wage growth. Differences in job skill requirements explain a significant portion of the observed heterogeneity in wage growth. I provide evidence that jobs of different skill requirements differ in their prospects for earnings growth, independent of the workers who fill these jobs. This contradicts some previous studies that have concluded that heterogeneity in permanent rates of wage growth among jobs is empirically unimportant (Topel, 1991; Topel & Ward, 1992; Abowd & Card, 1989). I have shown that, in terms of wage differentials, reading/writing skills, in particular, substantially increase wages not only through mere use, but also via experience using these skills, because these jobs offer more on-the-job training opportunities (formal/informal), and thus greater wage growth potential. This result was robust to explicit controls for unobserved heterogeneity related to wage levels and wage growth, as evidenced in the first-difference fixed effect and double-difference wage growth estimation results.

Computer usage is associated with relative pay differentials over non-computer-users. The association of computer usage with higher pay remains, even after controlling for many other sources of pay variation, thus replicating the similar findings of Krueger (1993) and others. However, unlike previous studies, the evidence here suggests that the large and significant effects of computer skill observed in the cross-sectional results do not reflect the true return of computer skills (i.e., the productivity enhancing effect of computers in the workplace), but rather is a result of the job sorting process through which workers with greater ability are systematically selected into jobs requiring computer skills. In the longitudinal dimension, the wage premium associated with computer skills disappears (both the immediate returns as well as the returns to computer usage experience). These results highlight the importance of using longitudinal data to isolate the true return to job skills, which was difficult to address by Krueger (1993) or DiNardo and Pischke (1997) using only cross-sectional information on workers.

Results from the wage growth analysis identified job mobility as a critical component of the wage growth process in the less-skilled labor market of this sample of less-educated women. My analysis of the determinants of job transitions underscored the importance of potential wage growth as an important factor affecting job turnover behavior. The turnover analysis underscores the sensitivity of these women's job transition patterns to changes in labor market demand conditions, which ultimately affect wage growth. This work highlights the importance of *jointly* considering processes of turnover along with wage growth when analyzing the labor market experiences of less-skilled workers. The results from the analyses of wage and job dynamics taken together, suggest that jobs requiring more cognitive skills (e.g., reading/writing) reduce worker's (firm's) propensity to quit (layoff/discharge) by providing greater learning opportunities (human capital investment opportunities – firm-specific training (formal/informal)), thereby offering more potential to experience wage growth. The results from the job turnover analyses, which suggest an important role of wage growth prospects in predicting job turnover, are robust to explicit controls for unobserved heterogeneity, as evidenced in the fixed-effect Cox proportional hazard model estimation results.

The results show that factors that predict future wage growth reduce quits as predicted by economic theory. The findings are consistent with an economic model in which workers compare the long-run value of employment opportunities when making quit decisions, which supports recent theoretical work by Munasinghe (2000) on the relationship between wage growth and job turnover. Because of data limitations, most previous work has relied on the assumption that, together with tenure and experience, the wage is a sufficient statistic for future wages. The results of this paper are inconsistent with that assumption both in that the analysis shows an additional predictor of wage growth – job skill requirements (independent of the workers who fill these jobs) – and in that this predictor helps explain quit behavior. The results therefore also point to the importance of developing good longitudinal data sets with information about firm characteristics, job skill requirements, and wages in order to improve our understanding of the wage growth process, particularly for less-skilled workers.

The results have important implications for welfare reform. TANFs work participation mandates have shifted the focus of welfare-to-work programs away from education and training and toward immediate job placement. As this study demonstrates, however, job skills profoundly affect the wage-experience profile, and thus, remain a central ingredient that will determine welfare recipients' ability to attain economic self-sufficiency.

Because most welfare-to-work programs have focused narrowly on job placement, we unfortunately have limited knowledge about how to design and implement programs that promote job retention and job advancement. Analyses that inform and evaluate the likely effects of various post-employment services is an important topic for future research.

The focus of this paper was to analyze the effects of job skills on the wage growth process and job turnover behavior of former/current welfare recipients. Ultimately, an important direction of future research will be to investigate whether particular skills have rising or falling value, analyzed separately by education and gender. This will provide the type of labor market information that may illuminate and inform policy with respect to the skill-supplying institutions.

NOTES

1. The broader questions of the extent to which the employment problems of the working poor emanate from job skill deficiencies versus a deterioration of job quality for less-educated workers, is a related issue but one beyond the scope of this paper.

2. On the other hand, there is a countervailing effect because job matching is likely a more important component of the earnings of high-skilled workers (e.g., Barron, Berger, & Black, 1997, show that employers spend more resources trying to make good matches for high-skilled workers), which may act to increase the value of their job changes. It is not clear, as a matter of theory, whether job changes are more important for less-skilled workers. Indeed, skill-level differences in the importance of the relationship between job changes and wage growth are borne out empirically. For example, Bartel (1980) finds that less-educated workers had the largest proportion of earnings gains occurring between jobs.

3. A few exceptions are Loprest (1992), Keith and McWilliams (1997), and Abbott and Beach (1994).

4. Notable studies focusing on the wage growth of less-skilled workers include Connolly and Gottschalk (2000), Gladden and Taber (2000), Loeb and Corcoran (2001), Burtless (1995), Moffitt and Rangarajan (1989), Card, Michalopoulos, and Robins (2001).

5. Notable exceptions include (Antel, 1986; Topel, 1991; Mincer, 1986; Bartel & Borjas, 1981).

6. The job task questions were developed from Harry Holzer (1996).

7. Michigan's welfare policies are quite similar to those of many other states. For example, women in Michigan who worked part-time at minimum wage jobs were at the median for monthly net income among 12 states that contained a large portion of the nation's population and about half of the 1998 caseload (Acs, Coe, Watson, & Lerman, 1998 Acs et al., 1998). While the study uses data only from Michigan, the policy and economic conditions in Michigan are broadly representative of the majority of the TANF caseload.

8. For the job turnover analysis, I use information collected at each wave of the WES on the set of job skills used on jobs held over roughly the past year. Some individuals may have used a job skill on a job held in a given year, but not on all the jobs held that are analyzed in that year. Consequently, job skills used do not correspond to jobs held perfectly in all cases (i.e., they are not perfectly aligned). However, I do not believe this to cause a serious mismeasurement issue.

9. Since only about 10% of the sample did not work between waves (and thus lack wage information), selection bias should not be a major concern.

10. Job change and employer change are used interchangeably here due to insufficient data information to distinguish between the two. However, for the period spanning Winter 2002–2004 when information was collected on job and employer

tenure, I find that the lion's share of job-to-job transitions occurred between firms rather than promotion within firms.

11. Involuntary job separations resulting from being laid-off and separations resulting from being fired are grouped together here due to insufficient data information to distinguish between the two.

12. I compared the total number of job transitions with information on the total number of jobs held between waves, as well as information on jobs held concurrently and could account for nearly all primary job changes.

13. This is confirmed empirically in the WES data from self-reported reasons for job separations.

14. See, for example, Lynch (1991) for evidence on the effects of on-the-job training on wage growth and job mobility patterns of female workers.

15. Previous research has documented that most employer-provided training is short and intensive, concentrated during the first four weeks of the job spell (Lynch 1991). Thus, the observed differences in the amount of hours of job training are not likely to be driven by potential differences in job turnover rates between these jobs.

16. Employer reports of potential wage increases for merit and chances for promotion are likely upward-biased, since employers may consider it more socially acceptable to claim that they are willing to offer chances of upward mobility. Still, the differences in these reports provide useful comparisons of the potential for wage growth and chance for promotion in jobs of different skill.

17. The reported coefficients in Columns 1, 2, and 4 of Table 2 are the derivative of the probability with respect to a one-unit change in the particular variable, where the derivatives are evaluated at the sample means of the independent variables.

18. It is important to note that the OJT variable in the Employer Survey reflects the formal aspect of the process by which workers accumulate human capital – certainly, a significant portion of training and the process by which workers accumulate skills is informal, and is thus not captured by the OJT measure. The set of job tasks may pick up the effect of informal training opportunities.

19. The model includes the starting wage and conventional human capital variables as controls, though the coefficient estimates of these variables are suppressed in Table 2. One would not think of wages as an exogenous variable in this setting, but it is of interest to know whether individuals with relatively high starting wages are more likely to receive raises or to be promoted.

20. Only 14.3% of the respondents did not work between Waves 1 and 2, 11.9% did not work between Waves 2 and 3, and only 12.1% did not work between Waves 3 and 4. The WES data contains wage information for the most recent job of each respondent as of the survey interview dates of Waves 1–5, given the individual worked sometime between waves. Since only a small fraction of the sample did not work between waves (and thus lack wage information), selection bias should not be a major concern.

21. These results are consistent with those of Royalty (1998) and Holzer and LaLonde (2000), who found that job-to-nonemployment changes were more frequent than were job-to-job changes among young women with low levels of schooling.

22. Women who were working at Wave 1 were asked if they expected to be working in their current job less than six months, six months to one year, one to two years, or over two years. Sixty-three percent of those working at Wave 1 expected to

be working in the same job at Wave 2, but only 38% actually still worked at the same jobs at Wave 2. The primary reason reported for job separations between Waves 2 and 3 were: 21.3% fired/laid-off; 21.3% job-related quit (includes dissatisfaction with current job, such as inadequate pay, poor working conditions, suboptimal hours, poor job match); 10.3% child care concern; 9.4% health problem; 7.6% transportation problem; 2.7% family problem/pressure; 27.4% other. The large proportion reporting non-job-related reasons (57.4%) is consistent with the substantial job instability experienced by these women. Twenty percent of the women changed from working part-time to full-time on their primary job; 13.5% changed from full-time to part-time; 22.2% remained part-time; and 44.1% remained full-time between successive waves.

23. It is possible that the effects of job skills estimated with equation (3) will be biased if workers that do not use valued job skills have unobserved characteristics that lower not only their wage levels but also their rates of wage growth. For example, if the correct specification is

$$\begin{aligned}\ln(\mathrm{WAGE})_{ijt} = \Gamma \mathbf{Z}_{ijt} + \beta_0 \mathrm{EXP}_{it} + \beta_1 \mathrm{JOBSKILL}_{ijt} \\ + \beta_2 \mathrm{EXP\ using\ JOBSKILL})_{ijt} + \alpha_i + \gamma_i t + u_{ijt}\end{aligned} \qquad (2')$$

where the unobserved heterogeneity components can be decomposed into a time-invariant person-specific intercept term (α_i) and a person-specific growth term (γ_i). In this case, to eliminate bias on the estimated return to various job skills, I estimate a double-difference model to account for the person-specific growth effect. This is equivalent to estimating the determinants of changes in wage growth rates (between Wave 1–2 vs. Wave 2–3 vs. Wave 3–4 vs. Wave 4–5) for a given worker. In this model, the estimated return to job skills is identified by contrasting wage growth experienced over a period when the set of job skills used changes for a given worker. This specification is tested to evaluate potential bias from unobserved heterogeneity related to levels of wage growth.

24. Loprest (1992) also controls for occupation transitions using a one-dimensional occupation index in her analysis of wage growth (though her occupation index differs from that developed here). See Shaw (1987) and Sicherman and Galor (1990) for empirical work on occupational mobility.

25. Mincer (1986) pointed out that using all stayers as a comparison group presents selectivity bias, since the within-job wage growth for the type of worker prone to job change/loss may be different from the within-job wage growth in the economy as a whole. The differences may not be entirely controlled by observable characteristics. To control them, Mincer suggested using the following year's job changers/losers as the comparison group. However, because labor market demand conditions change significantly over the period analyzed, this is not a viable strategy here.

26. Baker, Gibbs, and Holmstrom (1994) document considerable variation in wages, as well as in their growth rate, within job grades, suggesting that the prospect of promotion is not the only means of providing incentives that firms use. Abowd et al. (1999) provide evidence showing that starting pay differentials and compensation growth profiles are negatively correlated across jobs; employers offering greater opportunities for compensation growth offer lower starting pay.

27. Empirical evidence supports this assumption. See, for example, the evidence of Baker et al. (1994) showing that those who experience the largest wage growth within a given job level also get promoted rapidly. They find the relationship between wage growth and time to promotion is uniformly negative. Furthermore, they find that promotees are drawn from all parts of the wage distribution within a given job level, suggesting that promotions are determined by factors other than the wage level.

28. As noted by Prendergast (1996), this may be caused by the common bureaucratic rules within firms where each job classification has a wage range that cannot be violated (e.g., job may have 6 grades). Workers who are at the top of their wage grade are generally impeded from future increases, constraining wage growth. Therefore, we would expect that job-to-job transition rates are accelerated by being at the top of a wage grade

29. Jovanovic and Nyarko (1996) stepping stone mobility model predicts that labor will flow from occupations with flat learning curves and into occupations where learning curves are steep, *as long as learning is sufficiently transferable between occupations.*

30. Royalty (1998) and Holzer and LaLonde (2000) find similar turnover patterns for non-college educated, young women in the NLSY using similar definitions of job transitions. However, the job turnover rates among jobs held by the WES sample of respondents are higher than that observed by Royalty, (1998; see her Figures 5 and 7) or Holzer and LaLonde (2000). For example, Holzer and LaLonde (2000) estimate an average weekly transition probability out of a job of about 2% in their sample of less-skilled (noncollege graduates) young workers. As a crude approximation, a 2% weekly transition rate translates into a median job duration of nine months; contrasted with the seven month median job duration found in the WES sample. Similarly, Royalty (1998) reports average annual job mobility (i.e., job-to-job turnover) and job instability (i.e., job-to-nonemployment turnover) rates of 18% and 28%, respectively, among noncollege educated women in the first year of job tenure; significantly lower turnover rates than observed in the WES sample. Additionally, the turnover rates observed in the WES sample are significantly higher than the 39% average annual turnover estimated by Holzer et al. (2001) from employer survey evidence of jobs recently filled by former/current recipients. (Potential sources of bias in their estimates are acknowledged and discussed in their paper).

31. To compute the implied marginal effects of explanatory variables on the hazard from the estimated coefficients of the multinomial probit model, I follow procedures developed in Stern (1989).

32. For example, one-third of workers who did not use reading/writing skills in a previous period did use these skills the next period; conversely, roughly 30% of women who were observed using reading/writing skills in the previous period did not use these skills in the subsequent period – the high prevalence of job instability was a factor that contributed to the latter pattern. Similarly high degrees of changes in job skills used are observed across periods for other job skills examined.

33. See Lancaster, 1990, pp. 268–271. Note that a censored spell must be at least as long as the smallest completed spell in order to contribute anything to the likelihood function.

34. Note that the effect of failing to control for heterogeneity is to bias the coefficients toward zero in a partial likelihood framework (see Lancaster, 1990, p. 304).

However, to the extent that workers with lower quit propensities work in jobs requiring more skills, one would expect that the coefficient should actually be smaller in absolute value when controlling for fixed effects. Thus, it appears that controlling for unobserved heterogeneity in Cox's partial likelihood framework may actually lead to an increase in the estimated coefficient if the effect mentioned above is more than counterbalanced by the removal of a substantial bias toward zero.

35. The first-stage estimates mirror the patterns of results shown for the dependent competing-risks hazard model of job turnover, and are available upon request.

36. Loprest (1992) also controls for occupation transitions using a one-dimensional occupation index in her analysis of wage growth (though her occupation index differs from that developed here). See Shaw (1987) and Sicherman and Galor (1990) for empirical work on occupational mobility.

REFERENCES

Abbott, M. G., & Beach, C. M. (1994). Wage changes and job changes of Canadian women: Evidence from the 1986-87 Labour Market Activity Survey. *Journal of Human Resources*, *29*(2), 429–460.

Abowd, J. M., & Card, D. (1989). On the covariance structure of earnings and hours changes. *Econometrica*, *57*(2), 411–445.

Abowd, J., Kramarz, F., & Margolis, D. (1999). High wage workers and high wage firms. *Econometrica*, *67*(March), 251–333.

Abraham, K., & Medoff, J. (1985). Length of service and promotions in union and nonunion work groups. *Industrial Labor Relations Review*, *38*(3), 408–420.

Acs, G., Coe, N., Watson, K., & Lerman, R. (1998). *Does work pay? An analysis of the work incentives under TANF*. Washington, DC: Urban Institute.

Altonji, J. G., & Shakotko, R. A. (1987). Do wages rise with job seniority? *Review of Economic Studies*, *54*(July), 437–459.

Altonji, J. G., & Williams, N. (1997). *Do wages rise with job seniority? A reassessment*. NBER Working Paper Series #6010, April.

Antel, J. J. (1986). Human capital investment specialization and the wage effects of voluntary labor mobility. *Review of Economics and Statistics*, *68*, 477–483.

Baker, M. (1997). Growth-rate heterogeneity and the covariance structure of life-cycle earnings. *Journal of Labor Economics*, *15*(2), 338–375.

Baker, G., Gibbs, M., & Holmstrom, B. (1994). The wage policy of a firm. *Quarterly Journal of Economics*, *109*(4), 921–955.

Barron, J. M., Berger, M., & Black, D. (1997). Employer search, training, and vacancy duration. *Economic Inquiry*, *35*(1), 167–192.

Bartel, A. P. (1980). Earnings growth on the job and between jobs. *Economic Inquiry*, *18* (January), 299–306.

Bartel, A. P., & Borjas, G. J. (1981). Wage growth and job turnover: An empirical analysis. In: S. Rosen (Ed.), *Studies in Labor Markets*. Chicago: University of Chicago Press.

Burtless, G. (1995). Employment prospects of welfare recipients. In: D. S. Nightingale & R. H. Haveman (Eds), *The work alternative: Welfare reform and the realities of the job market*. Washington, DC: The Urban Institute Press.

Cancian, M., Haveman, R., Meyer, D., & Wolfe, B. (2000). *Before and after TANF: The economic well-being of women leaving welfare*. Special report no. 77. Madison: University of Wisconsin, Institute for Research on Poverty.

Card, D., Michalopoulos, C., & Robins, P. K. (2001). *Measuring wage growth among former welfare recipients*. Unpublished manuscript, January.

Chamberlain, G. (1985). Heterogeneity, omitted variable bias, and duration dependence. In: J. J. Heckman & B. Singer (Eds), *Longitudinal analysis of labor market data* (pp. 3–38). Cambridge: Cambridge University Press.

Connolly, H., & Gottschalk, P. (2000). *Returns to tenure and experience revisited: Do less educated workers gain less from work experience?* Boston College Working Paper, November.

Corcoran, M., Duncan, G. J., & Ponza, M. (1983). A longitudinal analysis of white women's wages. *Journal of Human Resources*, *18*(4), 497–520.

Devereux, P. J. (2000). Task assignment over the business cycle. *Journal of Labor Economics*, *18*(1), 98–124.

DiNardo, J., & Pischke, J-S. (1997). The returns to computer use revisited: Have pencils changed the wage structure too? *Quarterly Journal of Economics*, *112*, 291–303.

Edin, K., & Lein, L. (1997). *Making ends meet*. New York: Russell Sage Foundation.

Entorf, H., Gollac, M., & Kramarz, F. (1999). New technologies, wages, and worker selection. *Journal of Labor Economics*, *17*(3), 464–490.

Farber, H. S. (1998). *Mobility and stability: The dynamics of job change in labor markets*. Working Paper #400, Princeton University, Industrial Relations Section, June.

Garcia-Perez, J. I., & Sanz, Y. R. (2004). *Wage changes through job mobility in Europe: A multinomial endogenous switching approach*. Unpublished manuscript.

Gladden, T., & Taber, C. (2000). Wage progression among less skilled workers. In: R. M. Blank & D. Card (Eds), *Finding work: Jobs and welfare reform*. New York: Russell Sage Foundation.

Gottschalk, P. (2002). *Downward nominal wage flexibility: Real or measurement error*? Boston College Working Paper, May.

Green, F. (1998). *The value of skills*. Unpublished manuscript, October.

Harris, K. (1996). Life after welfare: Women, work, and repeat dependency. *American Sociological Review*, *61*, 407–426.

Holzer, H. (1996). *What employers want: Job prospects for less-educated workers*. New York: Russell Sage Foundation.

Holzer, H. J., & LaLonde, R. J. (2000). Employment and job stability among less-educated young workers. In: D. Card & R. Blank (Eds), *Finding jobs: Work and welfare reform*. New York: Russell Sage Foundation.

Holzer, H. J., Stoll, M. A. & Wissoker, D. (2001). *Job performance and retention among welfare recipients*. Discussion paper # 1237-01 presented as part of the Institute for Research on Poverty at the University of Wisconsin-Madison, August.

Idson, T. (1989). Establishment size differentials in internal mobility. *Review of Economics and Statistics*, *71*(4), 721–724.

Jencks, C., & Phillips, M. (1998). *The black-white test score gap*. Washington, DC: Brookings Institution Press.

Johnson, R. C., & Corcoran, M. (2003). The road to economic self-sufficiency: Job quality and job transition patterns after elfare reform. *Journal of Public Policy Analysis and Management*, *22*(4), 615–639.

Jovanovic, B., & Nyarko, Y. (1996). *Stepping stone mobility*. NBER Working Paper Series, Working Paper 5651.

Keith, K., & McWilliams, A. (1997). Job mobility and gender-based wage growth differentials. *Economic Inquiry*, *35*(April), 320–333.

Krueger, A. B. (1993). How computers have changed the wage structure: Evidence from microdata, 1984–1989. *The Quarterly Journal of Economics*, *108*(1), 33–60.

Lancaster, T. (1990). *The econometric analysis of transition data*. Econometric Society Monographs no. 17. Cambridge: Cambridge University Press.

Lazear, E. P. (1995). A jobs-based analysis of labor markets. *The American economic review: Papers and proceedings of the hundredth and seventh annual meeting of the American economic association*, Washington, DC, January 6–8, (May), 260–265.

Levy, F., & Murnane, R. J. (1996). With what skills are computers a complement? *American Economic Review*, *86*, 258–262.

Lillard, L. A., & Panis, C. W. A. (2000). AML Multilevel Multiprocess Statistical Software, Release 1.0. Los Angeles, CA: EconWare.

Loeb, S., & Corcoran, M. (2001). Welfare, work experience, and economic self-sufficiency. *Journal of Public Policy Analysis and Management*, *20*(1), 1–20.

Loprest, P. J. (1992). Gender differences in wage growth and job mobility, In: *The American Economic Review: Papers and proceedings of the hundred and fourth annual meeting of the American economic association* (Vol. 82(2), pp. 526–532), May 2.

Lynch, L. M. (1991). The role of off-the-job vs. On-the-job training for the mobility of women workers. *American Economic Review*, *81*(2), 151–156.

Mincer, J. (1986). Wage changes in job changes. *Research in Labor Economics*, *8*(pt. A), 171–197.

Mincer, J., & Ofek, H. (1982). Interrupted work careers: Depreciation and restoration of human capital. *Journal of Human Resources*, *17*(1), 3–24.

Moffitt, R., & Rangarajan, A. (1989). The effect of transfer programmes on work effort and human capital formation: Evidence from the U.S. In: A. Dilnot & I. Walker (Eds), *The economics of social security*. Oxford University Press.

Munasinghe, L. (2000). Wage growth and the theory of turnover. *Journal of Labor Economcis*, *18*(2), 204–220.

Murnane, R., Levy, F., & Willett, J. (1995). *The growing importance of cognitive skills in wage determination*. NBER Working Paper.

Parent, Daniel. (1999). Wages and mobility: The impact of employer-provided training. *Journal of Labor Economics*, *17*(2), 298–317.

Pavetti, L., Holcomb, P., & Duke, A. (1995). *Increasing participation in work and work-related activities: Lessons from five state welfare reform demonstration projects*. Washington, DC: Urban Institute.

Prendergast, C. (1996). *What happens within firms? A survey of empirical evidence on compensation policies*. NBER Working Paper Series, Working Paper 5802.

Riccio, J., Friedlander, D., & Freedman, S. (1994). *GAIN: Benefits, costs and three-year impacts of a welfare- to-work program*. New York: Manpower Demonstration Research Corporation.

Rosen, S. (1972). Learning and experience in the labor market. *Journal of Human Resources*, *7*(3), 326–342.

Royalty, A. (1998). Job-to-job and job-to-nonemployment turnover by gender and education level. *Journal of Labor Economics*, *16*, 392–443.

Shaw, K. (1987). Occupational change, employer change, and the transferability of skills. *Southern Economic Journal*, *53*(5), 1283–1298.
Sicherman, N., & Galor, O. (1990). A theory of career mobility. *Journal of Political Economy*, *98*(1), 169–192.
Simpson, W. (1992). *Urban structure and the labor market*. New York: Oxford University Press.
Stratton, L. (1995). The effect interruptions in work experience have on wages. *Southern Economic Journal*, *61*(4), 955–970.
Stern, S. (1989). Rules of thumb for comparing multinomial logit and multinomial probit coefficients. *Economic Letters*, *31*(3), 235–238.
Topel, R. (1991). Specific capital, mobility, and wages: Wages rise with job seniority. *Journal of Political Economy*, *99*(1), 145–176.
Topel, R., & Ward, M. (1992). Job mobility and the careers of young men. *Quarterly Journal of Economics*, *107*(2), 439–479.
Tyler, J. H., Murnane, R. J., & Willett, J. B. (1999). *Do the cognitive skills of school dropouts matter in the labor market*? NBER Working Paper Series, Working Paper 7101.

APPENDIX A. MULTINOMIAL ENDOGENOUS SWITCHING MODEL OF WAGE GROWTH

Potential wage changes for a given job transition type can be represented by

$$\Delta \mathrm{Wage}_j = \beta_j \mathbf{X}_j + \varepsilon_j,\ j = 0, 1, 2, 3 \tag{A.1}$$

where j denotes the type of job transition – job stability, voluntary job mobility, employee-initiated job instability, and involuntary job instability. The selection mechanism is described through a latent variable model that captures the propensity of experiencing each of the transition types. We only observe the realization

$$I = k \Leftrightarrow I_k > \max\{I_j\},\ j = 0, 1, 2, 3 \tag{A.2}$$

That is, the worker will be observed experiencing job transition type k if the total value associated with this transition is greater than the value of any alternative transition type. This latent variable model may be interpreted as a reduced form approach, where supply and demand side effects interact and cannot be disentangled. This implies the behavior of workers and the functioning of the labor market jointly determine what job transition type is observed, I_j. The estimated coefficients of the explanatory variables therefore capture the joint effect of the preferences of the worker and employer's preferences with regard to the worker's characteristics. Thus, we have

$$\Delta \mathrm{Wage} = \Delta \mathrm{Wage}_k,\ \mathit{if}\ I_k = \max\{I_j\},\quad j = 0, 1, 2, 3 \tag{A.3}$$

I assume that I_j depends on observable characteristics (Z) and unobservable factors captured by woman-specific random effects (u_{ij}) and a random error component (v_{ijt}):

$$I_{ijt} = \alpha_j \mathbf{Z}_j + u_{ij} + v_{ijt} \tag{A.4}$$

In order to jointly estimate the wage change equations and job transition selection process, the likelihood function has to add the information relevant to the wage process and take account of the endogeneity of the job transition selection process. The selection process of the type of job transition experienced between waves is specified as a multinomial probit model with women-level random effects to allow a flexible correlation structure across alternative job transition types. Following Garcia-Perez and Sanz (2004), I estimate the endogenous switching model by full maximum likelihood. The estimation of the model is highly computationally intensive and is estimated using aML.

For ease of exposition, assume below there are only three types of potential job transition types. The likelihood function to be estimated has the following form:

$$\begin{aligned} L\left(\beta_{j^*}, \alpha_j, \sigma^2_{\varepsilon_j}, \sigma^2_{u_j}, \sigma_{u_j u_k}, \sigma_{\varepsilon_j u_{j^*}} \middle| \Delta\text{Wage}, \mathbf{X}, \mathbf{Z}, I^*\right) = \\ \prod_{\substack{I_1^*>0,\\ I_0^*>0}} \left[\varphi\left(\Delta\text{Wage}_0\right)\Phi\left(I_1^*>0, I_0^*>0 | \Delta\text{Wage}_0\right)\right] \\ \prod_{\substack{I_2^*>0,\\ I_0^*>0}} \left[\varphi\left(\Delta\text{Wage}_1\right)\Phi\left(I_2^*>0, I_0^*>0 | \Delta\text{Wage}_1\right)\right] \\ \prod_{\substack{I_3^*>0,\\ I_2^*>0}} \left[\varphi\left(\Delta\text{Wage}_2\right)\Phi\left(I_2^*>0, I_1^*>0 | \Delta\text{Wage}_2\right)\right] \end{aligned} \tag{A.5}$$

where the term $\varphi\left(\Delta\text{Wage}_j\right)$ denotes the density function of wage changes ($j = 0,1,2$) and $\Phi\left(I^* | \Delta\text{Wage}_j\right)$ is the cumulative distribution function of the bivariate selection process conditional on wage changes.

For each worker I observe one wage change and I have to predict the potential or counterfactual wage change for the alternative job transition types not observed. To illustrate how I compute the relative wage return to voluntary job mobility, the expected wage change experienced by voluntary job changers is described as:

$$E\left(\Delta\text{Wage}_0 | I_1^*>0, I_0^*>0\right) = \beta_0 \mathbf{X}_0 + \frac{\sigma_{\varepsilon_0}}{\left(1-\rho^2_{u_1^* u_0^*}\right)}\left(\theta_{01}\lambda_1 + \theta_{00}\lambda_0\right) \tag{A.6}$$

where θ_{00} and θ_{01} are functions of the correlations between the error terms of the wage change and job transition selection equations:

$$\theta_{00} = \left(\rho_{\varepsilon_0 u_0^*} - \rho_{\varepsilon_0 u_1^*}\rho_{u_1^* u_0^*}\right),\ \theta_{01} = \left(\rho_{\varepsilon_0 u_1^*} - \rho_{\varepsilon_0 u_0^*}\rho_{u_1^* u_0^*}\right) \tag{A.7}$$

If the selection process is not endogenous then these correlations between the error term of the wage change equation and the error term of the selection equation will be zero and therefore the estimated parameters θ_{00} and θ_{01} will also be zero. The terms λ_0 and λ_1 control the bivariate process of the probability of experiencing a voluntary job-to-job change relative to remaining in the same job and relative to job-to-nonemployment transition:

$$\lambda_0 = \phi\left(\frac{\alpha_0^* \mathbf{Z}}{\sigma_{u_0^*}}\right)\left(1 - \Phi\left(\frac{-\alpha_0^* \mathbf{Z}}{\sigma_{u_0^*}}\right)\right)^{-1}, \lambda_1 = \phi\left(\frac{\alpha_1^* \mathbf{Z}}{\sigma_{u_1^*}}\right)\left(1 - \Phi\left(\frac{-\alpha_1^* \mathbf{Z}}{\sigma_{u_1^*}}\right)\right)^{-1} \tag{A.8}$$

Thus, the returns to voluntary job mobility can be obtained by taking the difference between the wage equations for the observed job transition type and each of the counterfactuals, which can be computed in the same way.

Table A1 Cross-section OLS wages regressions using WES.

Dependent Variable: Log of Real Hourly Wages ($ 1999)				
Explanatory Variables	Mean	(1)	(2)	(3)
Human capital variables				
High school grad/ GED (reference category: high school dropout)	0.3704		0.0277	0.0246
			(0.0210)	(0.0208)
Some post-secondary education	0.3406		0.0655***	0.0495**
			(0.0233)	(0.0230)
Years of full-time work experience	4.8077		0.0143***	0.0136***
			(0.0053)	(0.0052)
Full-time work experience squared	23.1138		−0.0004	−0.0003
			(0.0003)	(0.0003)
Years of part-time work experience	3.4115		−0.0037	−0.0020
			(0.0065)	(0.0065)
Part-time work experience squared	11.6385		0.0004	0.0004

Table A1. (*Continued*).

Dependent Variable: Log of Real Hourly Wages ($ 1999)				
Explanatory Variables	Mean	(1)	(2)	(3)
Job skill variables			(0.0004)	(0.0004)
Reading/writing	0.5064	0.0531***	0.0434***	0.0401**
		(0.0163)	(0.0167)	(0.0164)
Experience using reading/writing	0.5982	0.0260**	0.0201*	0.0190*
		(0.0114)	(0.0114)	(0.0111)
Computer	0.2689	0.0815***	0.0723***	0.0444**
		(0.0182)	(0.0185)	(0.0189)
Experience using computer	0.2626	0.0471***	0.0391***	0.0291**
		(0.0152)	(0.0147)	(0.0142)
Math	0.5867	0.0084	0.0141	0.0134
		(0.0180)	(0.0177)	(0.0172)
Experience using math	0.7380	−0.0259**	−0.0237**	−0.0214*
		(0.0116)	(0.0114)	(0.0110)
Customer communication	0.7186	0.0130	0.0207	0.0239
		(0.0154)	(0.0158)	(0.0156)
Experience using customer communication	1.0279	0.0229*	0.0177	0.0144
		(0.0120)	(0.0119)	(0.0116)
Gauges/dials/instruments	0.4155	−0.0745***	−0.0848***	−0.0827***
		(0.0225)	(0.0228)	(0.0225)
Experience using gauges/dials/instruments	0.4359	0.0323***	0.0286**	0.0273**
		(0.0113)	(0.0114)	(0.0112)
Occupation index	1.4123			0.1244***
				(0.0213)
Below 6th grade reading competency	0.1915		−0.0605***	−0.0500**
			(0.0224)	(0.0223)
Learning disability	0.1503		−0.0524**	−0.0458*
			(0.0238)	(0.0238)
Full-time	0.6001	0.0757***	0.0627***	0.0626***
		(0.0161)	(0.0168)	(0.0167)
Union	0.1153	0.1786***	0.1753***	0.1769***
		(0.0265)	(0.0260)	(0.0256)
Demographic variables				
Child 0–2 years	0.2974	−0.0542***	−0.0289	−0.0279
		(0.0174)	(0.0184)	(0.0186)
Child 3–5 years	0.3547	0.0014	0.0190	0.0176
		(0.0155)	(0.0161)	(0.0162)

Table A1. (*Continued*).

Dependent Variable: Log of Real Hourly Wages ($ 1999)				
Explanatory Variables	Mean	(1)	(2)	(3)
Married/cohabiting	0.2908	0.0411**	0.0459**	0.0467**
		(0.0193)	(0.0194)	(0.0195)
Black	0.5401	0.0213	0.0276	0.0303
		(0.0189)	(0.0193)	(0.0191)
Health-related variables				
Work-limiting (physical) health condition	0.2668	−0.0613***	−0.0590***	−0.0491***
		(0.0189)	(0.0189)	(0.0188)
Mental health condition	0.2980	−0.0359**	−0.0329*	−0.0349**
		(0.0173)	(0.0175)	(0.0174)
Domestic violence (past year)	0.1547	−0.0201	−0.0095	−0.0059
		(0.0199)	(0.0200)	(0.0197)
Labor market demand conditions				
Unemployment rate	6.18	−0.0069	−0.0066	−0.0076
		(0.0056)	(0.0057)	(0.0056)
Observations		2,558	2,405	2,396
R^2		0.1737	0.2001	0.2180

Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.
Note: Regressions also include a constant term. The median (mean) wage for this sample of former/current welfare recipients is $6.63 ($7.24).

APPENDIX B. DERIVATION OF OCCUPATION INDEX

In order to control for occupation transitions, I create a one-dimensional occupation index that is designed to capture the amount of human capital needed to work in different occupations (after required training is completed). My construction of the index is adapted from that previously developed by Sicherman and Galor (1990) in their analysis of occupation mobility[36].

The mean levels of human capital needed to work in the various occupations our sample of women are likely to work in are constructed by summing the weighted means of the levels of schooling, previous occupation-specific experience, previous training (or skill certification), and job skills required in order for a worker to be qualified to work in the different occupations. Using the 1997 Michigan Employer Survey (MES), these means by occupation are estimated from employer reports of the requirements of a sample of recently-filled non-college jobs, which constitute a representative sample of the jobs that are available to non-college educated workers in local labor markets over a period of several months (Holzer, 1996). The weights are the estimated coefficients of these variables (level of schooling, previous occupation-specific experience, previous training (or skill certification), and job skill requirements) in a wage regression. Specifically, using the sample of recently-filled non-college jobs from MES, the occupation index is derived by first estimating the following wage regression:

$$\begin{aligned}\ln(\mathrm{W}_{ijo}) = \mathbf{X}_{ijo}\beta + \alpha\mathrm{ED}_j + \tau\mathrm{POCCEXP}_j + \delta\mathrm{PTRAIN}_j \\ + \mu\mathrm{JOBSKILLS}_j + \varepsilon_{ijo}\end{aligned} \tag{B.1}$$

where **X** is a vector of observed characteristics, ED is the level of schooling required to be considered for hire, POCCEXP is the degree of previous occupation-specific experience necessary to be considered for hire, PTRAIN is whether the job requires previous formal training or skill certification, JOBSKILLS is a vector of job tasks required on the job, i indexes the individual and j the job.

The mean level of human capital needed to be fully qualified to work in occupation k is given by:

$$\overline{HC_k} = \alpha\overline{ED_k} + \tau\overline{POCCEXP_k} + \delta\overline{PTRAIN_k} + \mu\overline{JOBSKILLS_k} \tag{B.2}$$

The bar over variables in equation (B.2) signifies the mean level of the variable across the sample of non-college jobs in occupation group k. The change in occupation index due to occupation transition from occupation l to m, or equivalently, the vertical distance between occupations l and m is given by:

$$\Delta\mathrm{OCCINDEX}_{lm} = \overline{HC_l} - \overline{HC_m} \tag{B.3}$$

This occupation index results in the following hierarchical ranking of occupations for non-college educated workers:

(1) Professional/Managerial/Technical
(2) Clerical
(3) Craft
(4) Operative

(5) Service
(6) Sales
(7) Laborer

Data limitations do not allow a more detailed (3-digit) occupational ranking. This ranking is highly correlated with that obtained by the mean levels of schooling and the mean wages per occupation.

Occupation changing is common among the WES sample. At baseline (Wave 1), WES women are concentrated in relatively few occupations, and are least represented in occupations that have the highest probability of requiring previous occupation-specific experience to be considered for hire (Johnson & Corcoran, 2003). By far, service is the occupation containing the largest fraction of respondents, 41%, and followed by 22% working in sales. Using one-digit census-level occupation codes, the average fraction of respondents remaining in the same occupation between successive waves range from only 25–68%. The occupation transition patterns suggest both a significant amount of upward and downward occupation mobility. Some of the occupation changes may be the result of measurement error due to misclassification. The largest occupation transition among our sample is service to sales. This evidence of frequent occupation changing is consistent with human capital theory since individuals who have invested less in occupation-specific skills have less to lose when changing occupations.

FIRMS, INDUSTRIES, AND UNEMPLOYMENT INSURANCE: AN ANALYSIS USING EMPLOYER–EMPLOYEE DATA

Miles Corak and Wen-Hao Chen

ABSTRACT

Administrative data on the universe of employees, firms, and unemployment insurance (UI) recipients in Canada over an 11-year period are used to examine the operation of UI using the firm as the unit of analysis. Persistent transfers through UI are present at both industry and firm levels, and an analysis using firm fixed effect indicates that an important fraction of variation in them can be attributed to firm effects. Calculations of overall efficiency loss are very sensitive to the degree to which firm-level information is used. A full appreciation of how UI interacts with the labour market requires recognition of the characteristics and human resource practices of firms.

1. INTRODUCTION

The exploration of newly available administrative data in a number of countries has led to a growing realization that a careful study of the interaction

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 299–336

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26008-4

between employer and employee characteristics is needed to fully understand labour market outcomes. Abowd, Kramarz, and Margolis (1999) represent one example of the importance of analysing both the demand and supply side sides of the market. They relate wage determination, inter-industry wage differentials, firm-size wage effects, and human resource management to both firm and individual effects. The authors use large linked administrative data sets from France, but other examples of this sort are to be increasingly found. Research in the United States, Canada, and the Nordic countries, particularly in Denmark, has also underscored this general point (Baldwin, 1995; Haltiwanger, Lane, Spletzer, Theeuwes, & Troske, 1999). The objective of this paper is to adopt this theme by introducing a new focus on the impact and design of social policy and its interaction with the labour market. In light of this literature it may be that many of the consequences of unemployment insurance (UI) attributable to individual behaviour in fact reflect the demand side of the market, or in general there may be a need for greater awareness of the roles of both supply and demand to accurately understand the impact of UI on wages, employment, and unemployment.

It is certainly the case that the interaction between UI and the labour market has received extensive study in all industrialized countries. But the focus of a great many analyses has been on the supply side of the labour market, in part reflecting the importance of search theory as a framework to guide both the development of data and empirical analysis. Consequently, the impact of UI replacement rates and benefit entitlements on the duration of unemployment spells has been a major concern. For example, Atkinson and Micklewright (1991) offer an extensive survey of this literature, while at the same time stressing the need for a broader perspective on the relationship between UI and labour market transitions. Another literature places the focus on the demand side of the labour market and relies on implicit contract theory to examine the incentives for firms to change their hiring and lay-off decisions. Hamermesh (1993, 1990) offers an overview of this literature, one that dates back at least to Feldstein (1976). These analyses deal for the most part with the US since it is the only country to have made extensive use of experience rating. Our objective is to adopt this approach and to paint a picture of the Canadian UI programme from the perspective of firms and industries.

Indeed, these themes have a particular relevance to the Canadian experience. The Canadian UI programme has been a relatively significant aspect of the country's social security system, particularly in the aftermath of an important reform in 1971 that significantly increased coverage and benefits. Lin (1998) offers a legislative overview of the programme. This reform in the structure of benefits was to have been accompanied by changes in the financial structure that would introduce experience rated premiums. Kesselman (1983) describes the legislation and how the introduction of experience rating was delayed and eventually dropped. The economic analysis of the subsequent history of the programme has been framed almost entirely in terms of the labour supply effects – the impact on the aggregate unemployment rate and the duration of benefits. Corak (1994) offers a broad survey of this literature, one that has informed successive incremental restrictions in benefits during the 1970s and 1980s. Major changes in the programme were introduced in the 1990s in part by the growing realization that a very significant fraction of claimants have repeatedly relied on the programme in a predictable way (Corak, 1993a, 1993b; Gray & Sweetman, 2001; Lemieux & MacLeod, 1995, 2000). In a climate focused on deficit reduction this led to substantial reductions in the benefit rates and entitlements, but also to innovative reforms that introduced a measure of experience rating. Tellingly, these were made to the supply side of the labour market. A claw back of benefits from higher income recipients became effective in 1997 with the rate depending upon the individual's claim history. An "intensity" rule was also introduced in which benefit rates would be tied to the number of weeks of benefits collected in the past. The benefit rate would decline by one percentage point for every 20 weeks of benefits collected during the past five years beginning in 2001 (to a maximum of five percentage points for those having collected 100 weeks of benefits). These innovations, however, were retracted in 2001, just before the intensity rule was to take affect.

The evolution of Canadian policy reveals a distinct tendency to evaluate the programme solely from the supply side of the labour market. Since this tendency has in part to do with the data available to analysts our objective is to bring a new perspective to bear on the operation of UI by relying on large administrative data sets that link information from firms, workers, and individual claimants. We follow the framework in Anderson and Meyer (1993) and build upon work by Corak and Pyper (1995a) to document patterns in the flow of UI benefits and taxes and to explain – in an accounting sense – the nature of the resulting cross-subsidies. This falls short of examining the consequences of the lack of experience rating in the structure of premiums, something that is not possible in the Canadian context given the universal nature of the programme and the lack of variation in tax rates across firms. Rather our analysis should be thought of as documenting the extent of the subsidies that may induce such changes, or perhaps represent their outcome. We also examine what fraction of the variance in these cross-subsidies are industry-specific, region-specific, and firm-specific, and also offer estimates to the extent of the associated deadweight loss (DWL).

The analysis is conducted both at the industry and firm levels in order to document the between and within industry patterns of cross-subsidization. It should be noted, however, that cross-subsidization between firms and industries will exist even in a perfectly experience rated UI programme at any point in time. Certain firms or industries will suffer adverse shocks that necessitate benefit receipt while others will not: that is the nature of insurance. It is persistence in the pattern of cross-subsidization through time, not its existence at any point in time, which suggests a deviation from insurance principles and illustrates both the incentives for firms to change their behaviour, and the results of such changes. Therefore we pay particular attention to longitudinal issues.

Section 2 describes the data and offers an overview of major developments. Our administrative data covers the universe of employers, workers, and UI recipients from 1986 to 1996. These years span a complete business cycle. Patterns of transfers across broad industry categories and provinces are presented. An analysis at a finer industrial level is offered in Section 3 and an accounting explanation of the observed patterns offered. This involves decomposing industry-level measures of Benefit/Tax ratios into components due to separations (both temporary and permanent), benefit rates, benefit durations, and contributions (which are directly related to earning levels). Section 4 presents a firm level analysis and a decomposition of variance, and Section 5 offers estimates of the efficiency losses due to the observed patterns.

We find that the Canadian UI programme redistributes significant moneys between industries and provinces, and that these transfers have been long-standing. This will come as no surprise to many observers. The major flow of funds is from the service industries towards the resource sectors and construction, and from Ontario towards the provinces east of it. Industries receive a net positive transfer through UI because of higher than average layoff rates, and lower than average wages (and hence contributions). Large net positive transfers are also associated with higher than average temporary layoff rates. In addition we find that not only do the same industries receive a positive transfer year in – year out, but also the same firms. In fact, the transfers imposed through UI are heavily concentrated at the firm level. Only 6.25% of firms consistently receive a net positive transfer in each of 11 years, and while they account for 6.6% of all jobs they are responsible for 28% of all UI benefits paid and contribute only 3.6% of total UI taxes; over 22% of firms never receive a transfer, and they represent 48% of all jobs, but account for only 28% of UI benefits and 60% of contributions. Almost three-quarters of UI claims in the "always subsidized" firms are due to above average rates of temporary layoffs suggesting not only that the same small fraction of firms receive subsidies every year, but also that the same workers repeatedly use UI year after year with the same employers. While "always subsidized" firms tend to be concentrated in "always subsidized" industries (particularly in construction), a significant fraction of the firms in most industries are of this sort. That is, in addition to considerable between-industry cross-subsidization, the UI programme also entails considerable within-industry cross-subsidization. Analysis of variance indicates that almost 60% of explained variation in persistent cross-subsidies can be attributed to firm effects. Firm effects are much more important than geography or industry. As a consequence estimates of overall efficiency losses are very sensitive to the level of aggregation used. Calculations based upon firm-level information are five to fifteen times larger than those using industry and province-level information only.

2. DATA AND AN OVERVIEW

We use a series of administrative files associated with the Canadian tax system, the UI programme, and a longitudinal catalogue of enterprises developed by Statistics Canada. Appendix A offers a detailed description of the source files and the procedures used to create the analytical files. Together these files approach universal coverage of firms, workers, and UI recipients. We create firm-level information on the number of employees, UI contributions made (by both the employer and employees), number of UI claims, the amount of UI benefits collected, and the average duration of claims. The basic unit in the analysis is the "firm," which should be taken to mean all private or public sector enterprises that remit tax deductions on behalf of their employees to Revenue Canada. Each reporting unit to Revenue Canada (as the Canada Customs and Revenue Agency was referred to during the period under study) is assigned a payroll deduction account, and this account number is the basis for aggregating to the enterprise level and linking across the various data sets. Our analysis begins in 1986 because that is the first year in which data files containing the universe of yearly UI claimants is available to us, and ends in 1996 because of a break in the longitudinal consistency of the payroll deduction account numbers in the following year. As it turns out these years represent a complete business cycle beginning with the recovery from the 1981/1982 recession and ending with the recovery from the recession of the early 1990s. During 1986 the aggregate unemployment rate was 9.6%, the same rate experienced in 1996 after first falling to 7.5% in 1989 and peaking at 11.4% in 1993. The end year also corresponds to the last year before substantial changes in the structure of the programme occurred in 1997. Most notably these involved a change in coverage and eligibility to an hour-based scheme (as opposed to the number of weeks worked subject to a minimum number of hours), a claw back of benefits from higher income claimants, and the introduction of the worker experience rating as described above.

In covering the entire population of employers, employees, and UI claimants over an 11-year period our data are much more comprehensive than the US analysis by Anderson and Meyer (1993) and the Canadian by Corak and Pyper (1995a, 1995b) that are precursors to our study. Anderson and Meyer offer an aggregate analysis of 22 states covering about 55 per cent of UI-covered employment to establish the degree and persistence in cross-subsidies for major industries (two-digit SIC). However, their more disaggregated analysis exploring the underlying causes of these patterns relies on eight states accounting for between 5 and 20 per cent of the states' covered workers; their analysis at the firm level is based on two states using only large employers and about 10 per cent of covered workers over a four- to six-year period. The structure of the data used by Corak and Pyper (1995a) is similar to that used in our work, but more limited in nature. Their aggregate analysis covers the years 1986–1990, but because of underlying changes in the way in which industries were coded the more detailed industry and longitudinal firm analysis is restricted only from 1986 to 1988. In addition, their analysis falls short of examining the independent role of firms in determining the extent of cross-subsidization.

Table 1 provides an overview of the programme's operation between 1986 and 1996 using aggregates derived from our data, expressed in constant 1997 dollars. For the most part the UI programme was in deficit during the mid to late 1980s and early 1990s. The deficit was around \$1.8 billion in both 1989 and 1990, and was over \$2.5 billion in 1991. However, the system turned to surpluses after 1992, recording a peak surplus of \$8.2 billion in 1996. During this 11-year period the programme collected \$17.2 billion in premiums on average per year, while paying out about \$15.2 billions in benefits to 2.5 million claimants. These results are consistent with those in Lin (1998). Basically, the UI balance is quite sensitive to the business cycle, and as mentioned this period covers a complete cycle. While the average annual balance over this period is roughly a \$2 billion surplus, the yearly balances are quite different during the recovery and expansion of the early to mid-1990s than they were a decade earlier during the expansion following the 1981/1982

Table 1. Overview of the Canadian UI Program from Administrative Data: 1986–1996.

Year	Number of Firms	Total UI Benefits ($ millions)	Total UI Contributions ($ millions)	Account Balance ($ millions)	Total Jobs ('000s)	Total UI Claims ('000s)	Fraction of Claims due to Temporary Separations	Unemployment Rate (%)
1986	839,832	14,239	13,720	−519	19,211	2,612	0.47	9.6
1987	871,068	13,153	14,351	1,198	20,284	2,449	0.46	8.8
1988	895,058	13,723	15,087	1,364	21,193	2,492	0.46	7.8
1989	915,217	14,762	13,016	−1,746	21,746	2,578	0.47	7.5
1990	925,314	17,011	15,188	−1,823	21,308	2,767	0.48	8.1
1991	915,244	19,111	16,572	−2,539	20,165	2,780	0.50	10.3
1992	915,008	20,289	19,868	−421	19,271	2,913	0.51	11.2
1993	918,720	17,309	19,879	2,570	18,976	2,614	0.52	11.4
1994	926,873	12,821	20,947	8,126	19,460	2,315	0.52	10.4
1995	932,169	13,194	20,812	7,618	19,656	2,430	0.50	9.4
1996	935,029	11,445	19,636	8,191	19,647	2,323	0.53	9.6
Average		15,187	17,189	2,002	20,083	2,572	0.49	9.5

Note: All dollar figures are expressed in constant 1997 dollars.
Derivations by the authors using Statistics Canada Administrative Data.
The unemployment rate is obtained from the Labour Force Survey.

recession. Significant surpluses were recorded during the 1990s despite the average unemployment rate being higher than during the mid to late 1980s. Lin (1998) suggests that theses surpluses may be attributed to a number of factors. First, there was a rapid increase in tax revenue after 1991, due to the recovery of the economy but also to increases in premium rates (see Appendix B). Another factor has to do with the declining amount of benefits, most likely associated with legislated reductions in benefit rates and eligibility.[1] A final notable feature of the data in Table 1 is the significant fraction of claims due to temporary separations.[2] On average, half of UI claimants were separated from work temporarily, with a slight rise over the period.

Tables 2 to 4 present information similar to that offered in Corak and Pyper (1995a) but over a longer time horizon. The dollar value of net UI transfers (total benefits less taxes) by province and major industry are offered in Table 2. A positive value denotes a net transfer and negative denotes a surcharge.[3] Generally, provinces east of Ontario receive net transfers from the rest of Canada (except British Columbia and the two territories). Ontario alone contributes on average \$1.95 billion each year, while Quebec is the largest recipient (about \$960 million annually). At the industry level, UI funds were transferred from Services and the Public Sector to Construction: the latter receiving \$1.58 billion each year, with the former together contributing \$1.79 billion. The largest individual contributor is the Service Sector in Ontario, being surcharged \$805 million annually. On the other hand, Construction in Quebec received the largest transfer, an average of \$529 million.

Table 3 presents these transfers on a per-job basis. The primary sectors receive the greatest per-job transfers: \$4,735, \$2,005, and \$1,336 for Fishing, Forestry, and Construction. The per-job transfers are relatively smaller in the surcharged industries, the largest being \$519 in the Public Sector followed by \$419 in Transportation and \$391 in Finance. With respect to inter-provincial transfers, Newfoundland and Prince Edward Island receive transfers of \$1,782 and \$1,371 per job, respectively. On the other hand, the largest per-job contributor to UI is Ontario at \$251. The most notable recipients are those in goods producing industries in Atlantic Provinces. The largest per-job transfer is the Fishing industry in Newfoundland and PEI receiving with about \$6,800 annually per job. On the other hand, the Service industries as well as Mining and Manufacturing west of the Ottawa River pay substantial contributions on a per-job basis, the largest being the Public Sector in Ontario at about \$766 per job.

The Relative Benefit/Tax ratio (RBT) is presented in Table 4. This is defined as $\mathrm{RBT}_i = (B_i / T_i)/(B/T)$, a number greater than one indicating that

Table 2. UI Income Transfers across Industries and Provinces: Annual Averages, 1986–1996 (UI Benefits less UI Taxes Expressed in Millions of 1997 Dollars).

	Nfld	PEI	NS	NB	Quebec	Ontario	Man	Sask	Alberta	BC	NWT	Yukon	Outside Canada	Canada
Agriculture	6.89	14.77	12.10	18.48	66.11	26.65	6.10	12.58	2.38	53.41	0.05	0.07	−0.02	218.67
Forestry	22.94	2.25	19.44	42.79	106.52	7.89	1.05	3.07	4.00	63.42	0.45	0.09	0.01	273.37
Fishing	19.58	13.90	26.25	33.44	9.45	2.98	1.17	0.07	−0.03	5.76	0.09	0.01	0.00	113.43
Mining	2.51	0.23	−0.57	5.56	15.32	−20.13	−1.40	−3.88	−31.42	0.05	1.42	3.22	−0.07	−28.93
Manufacturing	178.00	26.64	64.75	95.05	134.36	−519.31	−20.81	−10.38	−36.87	−8.04	0.29	0.23	−0.85	−96.95
Construction	96.66	18.16	82.75	104.80	528.90	389.97	43.02	36.97	111.79	151.22	6.71	4.19	0.13	1,575.26
Transportation	15.58	3.45	−2.63	5.83	−94.83	−286.14	−32.57	−25.32	−46.52	−77.04	1.78	0.06	−0.27	−538.63
Trade	62.92	14.85	37.16	35.01	122.73	−288.44	−20.93	−18.28	−61.18	−36.47	0.62	0.94	−0.02	−151.08
Finance	2.68	0.40	−5.95	−0.86	−68.95	−268.72	−17.13	−13.25	−32.77	−51.26	0.20	0.00	−0.10	−455.70
Service	92.05	12.50	18.62	48.49	16.20	−804.79	−57.23	−45.24	−115.15	−105.14	3.33	3.69	−0.75	−933.41
Public admin	27.50	10.93	−36.98	−11.36	−136.66	−404.41	−41.05	−23.46	−121.94	−104.58	−1.99	0.16	−14.37	−858.23
Total	597.24	129.72	261.43	416.97	958.78	−1,950.77	−115.05	−69.55	−260.74	19.33	15.25	14.32	−16.93	

Note: Table entries are $B_i - T_i\,(B/T)$, where B_i represents total UI benefits received in sector i, T_i total contributions made and unsubscripted totals are for the entire country. Unclassified industries are included in the total.

Table 3. UI Income Transfer per Job: By Industry and Province, Annual Average (1986–1996) (UI Benefits less UI Taxes divided by Number of Jobs, Expressed in Millions of 1997 Dollars).

	Nfld	PEI	NS	NB	Quebec	Ontario	Man	Sask	Alberta	BC	NWT	Yukon	Canada
Agriculture	2,863	2,206	1,029	2,068	1,218	239	396	532	78	1,237	535	2,200	710
Forestry	5,422	4,645	2,653	4,953	3,202	544	1,166	1,480	701	1,092	1,749	1,053	2,005
Fishing	6,849	6,828	5,329	6,503	5,210	1,547	3,233	1,131	−279	1,339	713	972	4,735
Mining	535	2,395	−31	957	463	−425	−272	−263	−289	9	412	1,860	−111
Manufacturing	3,979	2,596	897	1,357	160	−384	−235	−257	−199	−25	241	566	−33
Construction	3,963	2,688	2,173	2,776	2,094	989	1,226	1,007	733	888	1,167	1,965	1,336
Transportation	767	610	−81	196	−312	−605	−538	−512	−386	−413	352	34	−419
Trade	1,102	996	352	430	138	−208	−164	−171	−169	−77	111	257	−42
Finance	237	123	−229	−51	−252	−521	−421	−396	−328	−357	77	48	−391
Service	899	424	97	385	9	−282	−212	−199	−145	−101	242	484	−126
Public admin	455	763	−417	−124	−390	−766	−486	−368	−681	−692	−97	24	−519
Total	1,782	1,371	438	860	197	−251	−158	−116	−127	13	239	498	

Table 4. Relative Benefit – Tax Ratios: By Industry and Province, Annual Average (1986–1996).

	Nfld	PEI	NS	NB	Quebec	Ontario	Man	Sask	Alberta	BC	NWT	Yukon	Outside Canada	Canada
Agriculture	10.86	10.01	4.87	8.63	4.69	1.73	2.16	2.99	1.22	4.74	3.37	8.67	0.20	3.18
Forestry	16.35	18.29	8.49	13.45	9.05	1.85	4.39	4.62	2.89	2.93	7.40	5.55	0.77	5.06
Fishing	25.54	27.42	21.35	22.15	17.99	3.31	16.09	7.67	0.69	4.68	6.40	4.34	0.61	14.76
Mining	1.35	7.32	1.02	1.79	1.36	0.68	0.78	0.76	0.74	0.98	1.38	3.58	1.50	0.90
Manufacturing	7.75	5.59	2.07	2.70	1.16	0.64	0.75	0.72	0.78	0.98	1.29	1.82	0.39	0.95
Construction	9.36	6.20	5.13	7.06	4.70	2.45	3.15	3.02	2.38	2.71	3.85	5.18	2.30	3.29
Transportation	1.80	1.75	0.93	1.19	0.72	0.46	0.54	0.52	0.59	0.60	1.47	1.02	0.29	0.61
Trade	3.28	3.09	1.65	1.82	1.24	0.65	0.71	0.69	0.71	0.88	1.27	1.47	2.52	0.92
Finance	1.32	1.17	0.74	0.94	0.73	0.42	0.48	0.51	0.59	0.60	1.13	1.09	0.37	0.56
Service	2.57	1.87	1.17	1.82	1.02	0.57	0.65	0.66	0.74	0.84	1.49	2.32	0.47	0.80
Public admin	1.49	1.85	0.63	0.90	0.64	0.37	0.55	0.65	0.45	0.44	0.89	1.04	0.28	0.55
All Industries	3.74	3.44	1.62	2.27	1.27	0.68	0.79	0.83	0.82	1.02	1.38	1.74	0.32	

the industry/province receives a subsidy and a value less than one indicating a surcharge. For example, the RBT should be interpreted as indicating that for every dollar of UI contributions from the Agriculture industry in Newfoundland $10.86 in UI benefits are received, while only 37 cents of benefits are received for every dollar of contribution from the Public Sector in Ontario. The patterns of cross-subsidization presented in Table 4 are consistent with those in Corak and Pyper (1995a) despite their use of a much shorter time horizon. In addition these data also paint the same general picture as those reported by Karagiannis (1986) who documents the patterns of cross-subsidization over the 1975–1982 period. Together these studies suggest that there is a long established and stable pattern of cross-subsidization in the Canadian UI programme that is little influenced by the business cycle and extends back at least to the years immediately following the introduction of the 1972 legislative changes.[4]

Details outlining the time series patterns of the RBT by province and industry are provided in the working paper version of this text.[5] In summary, developments in the RBT can be divided into three distinct types at the provincial level. Regardless of the magnitudes, the Atlantic Provinces as well as Quebec display a very similar pattern over time. The ratios are greater than one throughout the period, rising slightly during 1986–1989, dropping in 1990, then rising (with the exception of Newfoundland) since then. It is not known why there is a drop during 1990 in these regions. This may reflect the temporary suspension of the Variable Entrance Requirement (VER) between January and November of that year.[6] Historically these provinces have had much higher unemployment rates than the rest of the country. In contrast, the RBT in the provinces west of Ontario are below one, and generally declining through time. (British Columbia changed status from being subsidized to being surcharged.) Finally, in Ontario the evolution of the RBT is unique, with a value below 0.8 over the entire period representing the largest surcharge. There isn't a simple relationship between provincial variations in the RBT and the business cycle. This is expected because standardizing by the national ratio in RBT formula should remove cyclical effects.

Developments in the RBT by industry are, with a few exceptions, also relatively stable over time. Some industries always receive net transfers, while others always contribute, manufacturing being the sole exception. Cross-subsidization over the entire period is not, therefore, the result of a particularly bad few years requiring extensive readjustment and reliance on UI. Rather, it reflects a structural pattern in which some industries receive a net subsidy year-in and year-out, while others are repeatedly surcharged. In sum, it is something about the way in which employment is structured within provinces or about the way that industries operate that determines the pattern of persistent cross-subsidization embodied in the UI programme.

3. INTER-INDUSTRY PATTERNS IN DETAIL

An analysis at a finer industrial level allows a closer examination of the underlying causes of these persistent patterns. The results in Table 5 summarize the longitudinal patterns in the RBT for three-digit industries. The RBT is calculated for each of 228 industries defined according to SIC 1980 in each of 11 years. The distribution according to the number of years each industry had an RBT greater than one is concentrated at the two extremes: industries are either "never subsidized" or "always subsidized" over the 11 years under study. Nearly, 39% of industries never received a transfer over an 11-year period. The never subsidized industries account for 45% of all jobs, 34% of UI benefits, but contributed 61% of total UI contributions. In contrast, more than 30% of industries received a positive transfer in every year during 1986–1996, accounting for 32% of all employment, but 45% of total UI benefits and only 18.6% of total UI taxes.

We use the same decomposition method as Anderson and Meyer (1993) to develop an understanding of the underlying causes of the RBT in each

Table 5. Longitudinal UI Status of Industries, 1986–1996.

Number of Years in which RBT > 1	Number of Industries	Proportion of All Industries	Proportion of All Jobs	Proportion of All UI Benefits	Proportion of All Taxes Paid
0	88	38.6	45.0	34.0	61.2
1	12	5.3	1.9	1.9	2.5
2	9	3.9	3.5	3.2	3.2
3	6	2.6	3.2	3.5	3.6
4	9	3.9	4.1	3.1	3.1
5	5	2.2	1.6	1.2	1.2
6	7	3.1	2.7	2.2	2.1
7	5	2.2	1.7	2.0	1.7
8	9	3.9	1.6	1.1	0.9
9	2	0.9	1.4	1.2	1.0
10	7	3.1	1.1	1.5	1.0
11	69	30.3	32.3	45.0	18.6
Total	228	100	100	100	100

industry. Eq. (1) breaks the RBT into its constituent components.

$$\mathrm{RBT}_i = \frac{B_i/T_i}{B/T} = \frac{(n_i d_i b_i)/(t_i w_i)}{(ndb)/(tw)} = \left(\frac{n_i}{n}\right)\left(\frac{d_i}{d}\right)\left(\frac{b_i}{b}\right)\left(\frac{tw}{t_i w_i}\right) \tag{1}$$

where n_i represents the total number of UI claimants in industry i relative to the total number of jobs, d_i is the average duration (in weeks) of benefit recipient of these claims, b_i is the average weekly benefit amount, and $t_i w_i$ is the total premium paid by the employers and employees in the industry. Variables without subscripts represent the corresponding country-wide figures. As such an RBT greater than one can be attributed to: (a) an excessive relative number of claimants; (b) a longer benefit duration; (c) a higher benefit amount; and (d) a lower contribution. Since there is no experience rating in Canadian UI system t/t_i equals one. This implies that the value of the last term is governed by the relative earnings in the industry (w/w_i). Industries paying relatively lower wages will make relatively lower contributions, resulting in this term being greater than one and implying the industry is subsidized. Likewise industries paying higher than average wages will make relatively more contributions and the last term in Eq. (1) will be less than one, implying a tendency for the industry to be surcharged.

As an illustration, Table 6 shows the decomposition of the RBT ratio by major industry. The numbers in Columns (2–5) correspond to the four components of Eq. (1), their product being the RBT in Column (1). In Forestry, Fishing, and Construction, all the terms (with one small exception) contribute to the cross-subsidization of these industries, but a higher than average number of claimants is the major factor. The net subsidy in Agriculture is mainly caused by a higher value in Column 5 (meaning a lower tax contribution). For most surcharged industries, lower claim rates and/or higher contribution rates appear to be the leading causes of a lower RBT. In Mining and Manufacturing higher than average wages (and hence contributions) offset higher than average layoff and benefit rates leading both industries to be surcharged. Trade and Services pay a surcharge because lower claim rates dominate and override the fact that wages are lower than average. For the remaining surcharged industries (Transportation, Finance, and the Public Sector) both claim and contribution rates work together to reduce the RBT.

The relative claim rate can be considered as the sum of two parts: one for temporary separations (n_{ti}/n) and another for permanent separations (n_{pi}/n). These are illustrated in Columns (6) and (7), respectively. In all cross-subsidized industries the claim rate due to temporary separations is greater than that due to permanent separations.

Table 6. Causes of Cross-Subsidization by Major Industry, 1986–1996.

	RBT Ratio	Relative Incidence of Claims (n_i/n)	Relative Duration of Benefits (d_i/d)	Relative Benefit Rate (b_i/b)	Relative Taxes Paid (tw/t_iw_i)	Contribution of Separations	
						Temporary (n_{ti}/n)	Permanent (n_{pi}/n)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Agriculture	3.18	1.413	1.113	0.876	2.316	0.800	0.613
Forestry	5.06	2.335	1.107	1.297	1.513	1.412	0.923
Fishing	14.76	4.055	1.302	1.273	2.195	2.002	2.053
Mining	0.90	1.150	0.896	1.335	0.652	0.679	0.471
Manufacturing	0.95	1.342	0.895	1.063	0.748	0.791	0.552
Construction	3.29	1.967	0.989	1.283	1.315	1.079	0.888
Transportation	0.61	0.831	0.944	1.113	0.701	0.492	0.339
Trade	0.92	0.785	1.074	0.855	1.281	0.289	0.497
Finance	0.56	0.593	1.107	0.992	0.859	0.254	0.339
Service	0.80	0.760	1.002	0.864	1.219	0.384	0.376
Public admin	0.55	0.734	1.046	1.058	0.671	0.483	0.251

A complete tabulation of this sort for 228 industries defined at the three-digit SIC is presented in working paper version of our paper. The minority of industries (100 out of 228) have an RBT ratio greater than one. Of these 84 have a value between 1 and 3, and 16 have a value greater than 3. Fig. 1 summarizes this information by plotting each component against the RBT. The clearest positive relationship is, in the first instance, with the relative incidence of claims, and secondly with the relative taxes paid. The duration of benefits is also positively related with the RBT, but not as strongly. And there is an even weaker relationship between the RBT and the relative benefit rate. The covariance of the RBT ratio with each of the components described in Eq. (1) confirms the impression from these scatter plots. All covariances are positive and significantly different from zero, with the exception of the benefit rate. The relative claim and contribution rates have the strongest tie with the RBT ratio (0.79 and 0.46, respectively). The covariance between RBT and duration is also significant but with a smaller in magnitude (0.34), while that with the benefit rate is near zero (0.058).

With this in mind, Table 7 summarizes the information from the 228 industries by cross-classifying subsidized and surcharged industries by their relative claim and contribution rates. A large proportion of the subsidized industries (42 out of 100) tend to have both a higher than average separation rate and a lower than average wage rate. This is consistent with the theoretical prediction of an equilibrium under a system without experience

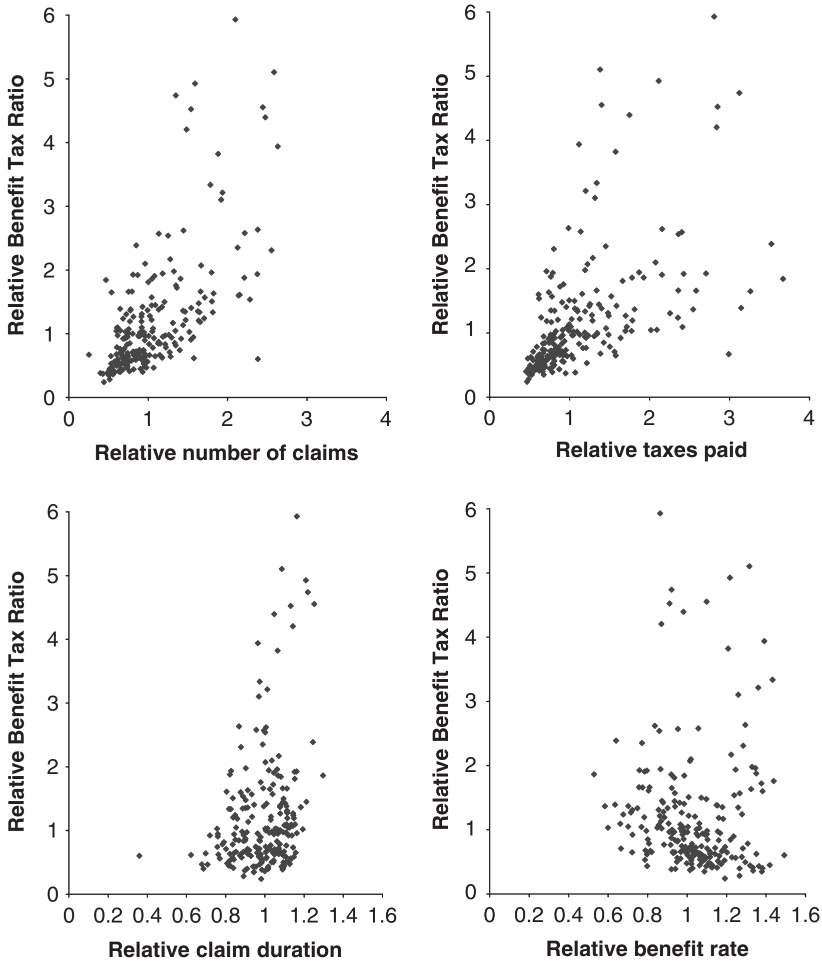


Fig. 1. The Relationship between the Relative Benefit Tax Ratio and Its Components for Industries at the Three-Digit SIC. *Note:* Three industries with relative benefit tax ratios greater than 10 have been suppressed. The definition of the relative benefit tax ratio and each of its four components are described in the text as Eq. (1).

rating. Hamermesh (1993) points out that if UI taxes are not tied to expected benefit receipt the programme offers a subsidy that presents an incentive for firms to increase layoffs and/or reduce wages. In spite of this, however, a significant proportion of subsidized industries are either low

Table 7. Distribution of Cross-Subsidization by Relative Separation Rate and Relative Contributions (1986–1996).

	High Lay-off Industries $(n_i/n) > 1$	Low Lay-off Industries $(n_i/n) < 1$
Net UI Recipients (RBT > 1)		
High wage industries $(tw/tw_i) < 1$	25	0
Low wage industries $(tw/tw_i) > 1$	42	33
Net UI Contributors (RBT < 1)		
High wage industries $(tw/tw_i) < 1$	28	76
Low wage industries $(tw/tw_i) > 1$	1	23

Note: Total Industries: 228.

layoff–low wage industries (33 out of 100) or high layoff–high wage industries (25 out of 100). In a similar vein, 76 out of the 128 surcharged industries (or almost 60%) are low layoff–high wage industries, but 28 (22%) are high layoff–high wage industries, and 23 (18%) are low layoff–low wage industries. Only one surcharged industry is classified as high layoff–low wage industry (Platemaking, Typesetting and Bindery with an RBT of 0.97).

In sum, a higher incidence of separation (especially temporary separations) as well as a lower than average wage rate are the major – though not exclusive – reasons for persistent inter-industry subsidies. This is consistent with theoretical predictions of firm behaviour under less than perfectly experience rated UI programs, and resonates with the fact that firms have much more ability to influence wages and layoff decisions than the other components in Eq. (1). In this sense it is not surprising that these two terms are important influences, in an accounting sense, of the RBT. That being said there remains considerable variation in these results even at the three-digit industry level, and it may therefore be important to model firm-level effects directly.

4. FIRM-LEVEL ANALYSIS

This challenge is taken up by examining firm-level patterns in the flow of UI benefits and contributions. Table 8 shows the distribution of firms by the number of years a positive transfer is received during the 11 years under study. The table contains two panels: one based on information for firms in operation for at least one year; another for those in operation all 11 years. There are about 2.2 million firms that operated in at least one of the 11 years

Table 8. Longitudinal UI Status of Firms: 1986–1996.

Number of Years Cross-Subsidized (RBT > 1)	Number of Firms	Percent of Firms	Percent of Jobs	Percent of UI Benefits Paid	Percent of UI Taxes Paid
A. Firms in Operation in All 11 Years					
0	70,275	22.1	48.1	28.4	60.3
1	42,645	13.4	10.4	6.8	10.8
2	37,016	11.6	6.7	5.0	6.2
3	31,730	9.97	5.2	4.2	4.3
4	26,118	8.21	4.6	4.0	3.6
5	21,292	6.69	4.0	3.9	2.9
6	17,458	5.49	3.1	3.2	2.0
7	14,621	4.59	3.0	3.3	1.7
8	12,595	3.96	2.9	3.9	1.8
9	11,725	3.68	2.5	3.7	1.4
10	12,853	4.04	2.9	5.7	1.5
11	19,889	6.25	6.6	27.9	3.6
Total	318,217	100	100	100	100
B. Firms in Operation for at Least One Year					
0	1,087,890	48.9	41.2	21.0	54.9
1	484,653	21.8	12.6	9.1	12.0
2	225,297	10.1	9.2	8.0	7.7
3	135,522	6.1	7.1	7.1	5.5
4	87,409	3.9	5.9	6.5	4.4
5	59,143	2.7	4.8	6.0	3.5
6	41,319	1.9	3.7	4.9	2.4
7	30,164	1.4	3.2	4.6	2.0
8	22,568	1.0	2.9	4.6	1.9
9	17,650	0.8	2.3	4.3	1.4
10	15,585	0.7	2.4	5.0	1.4
11	19,889	0.9	4.7	19.1	2.9
Total	2,227,089	100	100	100	100

under study, and almost 320,000 that operated during all 11 years. The underlying data used to develop the table reveals that these long-lived firms account for 71.4% of all job-years that existed over this period. They are the focus of our analysis for this reason but also because credible implicit contracts between employers and employees are most likely to have evolved in this sector. Of these firms more than one-fifth (22%) never received a subsidy. These "never subsidized" firms represent almost half of total employment, contributed over 60% of total UI taxes but received only about

28% of all benefits. At the other extreme, there is a small fraction of firms (6.25%) that received subsidies every year during this 11-year period. These "always subsidized" firms account for only 6.6% of all jobs, contributed only 3.6% of total UI taxes, but received fully 28% of all benefits. These firms represent less than 1% of all firms that ever existed during this period (see panel B of Table 8), but still account for about one-fifth of all UI benefits paid.

Table 9 provides a closer look at the characteristics of the never- and always-subsidized firms, focusing just on those firms operating in all 11 years. The first row shows the distribution of employees by firm size. More than half (54%) of jobs are in large enterprises (those with more than 500 jobs), while only 11% are with small firms (less than 20 jobs). This distribution is quite different for never- and always-subsidized firms. Mid-size enterprises (with between 20 and 500 jobs) account for 56% of the total in the always-subsidized firms, while nearly four-fifths of all jobs in never-subsidized firms are in large enterprises. The second row of the table presents information on the fraction of claims by type of separation. In never-subsidized firms the proportion of UI claims due to temporary and permanent layoff is about the same (each accounting for just over 40%), but over 70% of claims in always-subsidized firms are the result of temporary layoffs with only about one-fifth being due to permanent separations. In the context of the work by Corak (1993a, 1993b), Gray and Sweetman (2001), and Lemieux and MacLeod (1995, 2000) on the high degree of repeat UI use at the individual level this suggests that the same workers repeatedly use UI supported by employment with the same employers. The third and fourth rows of the table deal with the distribution of firms both across and within industries. Always-subsidized firms are not necessary concentrated in always-subsidized industries. For example, 24% and 11% of always-subsidized firms belong, respectively, to Services and Trade. This suggests that significant cross-subsidization also occurs within industries. The final rows of the Table 9 displays the distribution across and within provinces, and show, in the first instance, that both Quebec and Ontario consist of a significant portion of always subsidized and never subsidized firms. This reflects the absolute size of these provinces. Almost 38% of always-subsidized firms are located in Quebec, and a further 15% in Ontario; these percentages are almost exactly the same with respect to never-subsidized firms but reversed.

The within-industry distributions suggest that up to 35% of firms in the Forestry sector are always subsidized and about 30% in fishing. In contrast, 45% in Finance and about a quarter in Services and Mining are never subsidized. The within-province distributions are different, with 27% of all

Table 9. Characteristics of Always Subsidized and Never Subsidized Firms, Annual Average.

	All Firms	Always Subsidized Firms	Never Subsidized Firms
1. Distribution of Employees by Firm Size			
Less than 19	11.0	11.3	3.2
20–99	16.8	27.4	5.0
100–499	18.4	28.4	12.7
500 or more	53.8	32.9	79.7
2. Distribution of UI Claims by Reason for Separation			
Temporary	0.478	0.715	0.432
Permanent	0.370	0.211	0.404
Unknown	0.152	0.074	0.164
3. Distribution Across Industries (top three)			
	Services (36.5)	Construction (30.7)	Services (41.4)
	Trade (23.2)	Services (23.8)	Trade (19.1)
	Construction (10.8)	Trade (10.7)	Finance (14.1)
4. Distribution Within Industries (top three)			
		Forestry (34.7)	Finance (45.4)
		Fishing (29.0)	Services (26.1)
		Construction (17.6)	Mining (24.1)
5. Distribution Across Provinces (top three)			
	Ontario (33.1)	Quebec (37.8)	Ontario (38.5)
	Quebec (23.5)	Ontario (15.0)	Quebec (14.7)
	British Columbia (13.2)	New Brunswick (9.7)	Alberta (14.6)
6. Distribution Within Provinces (top three)			
		Newfoundland (27.3)	Saskatchewan (31.7)
		Prince Edward Island (21.4)	Alberta (30.6)
		New Brunswick (19.5)	Manitoba (27.4)

Note: Derivations are based on the subset of firms in operation in all 11 years between 1986 and 1996.

firms in Newfoundland being always subsidized, and one-fifth in Prince Edward Island and New Brunswick.

More detail on the industrial distribution of always- and never-subsidized firms is presented in Tables 10 and 11. The 20 three-digit SIC industries accounting for the highest proportions of always-subsidized firms are presented in Table 10. These 20 industries account for over 71% of always-subsidized firms. Most of the always-subsidized firms belong to the always-subsidized industries with fully one-third in the construction industries (SIC 420, 401, 412, 456, and 402). However, almost six per cent of

Table 10. Distribution of Always Subsidized Firms by Three-Digit Industry (the Highest 20 Industries).

Sic-80	Industry	RBT Ratio	Number of Firms	Percent of Always Subsidized Firms
420	Trade contracting industries	3.21	3,910	19.7
010	Agricultural industries	4.21	1,384	7.0
401	Residential building & development	3.82	944	4.8
041	Logging industry	5.10	816	4.1
412	Highway and heavy construction	3.94	777	3.9
965	Sports and recreation clubs service	2.54	719	3.6
921	Food services	1.39	692	3.5
456	Truck transport industries	1.38	649	3.3
911	Hotels motels and tourist courts	1.37	626	3.2
031	Fishing industries	17.08	498	2.5
601	Food stores	0.92	418	2.1
690	Other retail store and non-store retail industries	1.20	399	2.0
830	Local government services	0.52	400	2.0
910	Accommodation service excluding motels, hotels	4.74	390	2.0
457	Public passenger transit system industries	0.63	321	1.6
402	Non-residential building & development	3.21	284	1.4
990	M&E rental ,other repair, other service	1.81	249	1.3
960	Commercial spectator, sport & recreation	1.03	234	1.2
995	Services to buildings and dwellings	1.92	215	1.1
102	Fish products industry	13.23	213	1.1
Total			14,138	71.1

always-subsidized firms operate in surcharged industries (SIC 601 Food Stores, but notably also SIC 830 local government and SIC 457 public transit). Table 11 presents the 20 industries with the highest proportions of never-subsidized firms. These 20 industries account for 62% of never-subsidized firms. A large fraction of never-subsidized firms (31%) belong to the service industries, while there are no industries in this table associated with the manufacturing and public sectors. Fully half of these industries have an RBT greater than one. Further, six out of the twenty also appear in the first panel of the table among industries with a large fraction of always-subsidized firms. Cross-subsidization, in other words, exists not only between industries but also within them.

This point is made more clearly in Table 12, albeit at a more aggregated industrial classification. Between-industry cross-subsidization is clearly illustrated in these data. Over 70% of firms in the Forestry and Fishing sectors are either frequent or always subsidized, but only 5% in the Financial sector

Table 11. Distribution of Never Subsidized Firms by Three-Digit Industry (the Highest 20 Industries).

Sic-80	Industry	RBT Ratio	Number of Firms	Percent of Never Subsidized Firms
865	Office of physicians, surgeons and dentists	0.77	6,255	8.90
981	Religious organizations	0.53	5,797	8.25
010	Agricultural industries	4.21	5,097	7.25
720	Investment intermediary industries	0.94	3,468	4.93
750	Real estate operator, insurance industries	0.96	3,162	4.50
761	Insurance and real estate agencies	0.56	2,188	3.11
690	Other retail store and non-store retail	1.20	1,910	2.72
420	Trade contracting industries	3.10	1,888	2.69
777	Management consulting services	1.01	1,493	2.12
456	Truck transport industries	1.38	1,461	2.08
980	Membership org industries, excl religious	1.10	1,336	1.90
775	Architectural, engineering and other scientific	0.90	1,324	1.88
974	Private households	1.86	1,203	1.71
776	Offices of lawyers and notaries	0.68	1,166	1.66
601	Food stores	0.92	1,082	1.54
990	M&E rental, other repair, other service	1.81	1,044	1.49
773	Accounting and bookkeeping services	1.02	1,030	1.47
590	Other products industries, Wholesale	0.97	1,014	1.44
779	Other business services	0.95	969	1.38
635	Motor vehicle repair shops	1.42	919	1.31
Total			43,806	62.33

belong to that class. At the same time, however, within-industry cross-subsidization is also apparent. In both Mining and Transportation, 49% of firms never or only occasionally receive positive transfers from UI, while a large percentage (34% and 32%, respectively) always or frequently received transfers. Even in the Public sector (a sector with the lowest RBT ratio) almost one-third of enterprises always or frequently account for more benefits than contributions made. This within-industry cross-subsidization is sometimes more important than between-industry cross-subsidization. For instance, Agriculture is a subsidized industry with an RBT ratio of 3.2, but a third of firms in this industry never received a subsidy and a further one-quarter received a subsidy for only one, two, or three years out of the 11 under study. It is the minority of firms (27%) that lead benefits to be persistently greater than contributions for the industry as a whole. The same story holds, though perhaps not to the same degree, in other cross-subsidized industries. In Construction nine per cent of firms never receive a positive net

Table 12. Within Industry Distribution of Firms by UI Status: For firms in Operation in Each Year from 1986 to 1996.

Industry (One-Digit SIC 80)	Never Subsidized	Occasionally Subsidized	Sometimes Subsidized	Frequently Subsidized	Always Subsidized	Total
Agriculture	6,798	5,662	3,112	3,822	1,995	21,389
	(32.0)	(26.0)	(15.0)	(18.0)	(9.0)	
Forestry	155	262	336	1,009	913	2,675
	(6.0)	(10.0)	(13.0)	(38.0)	(34.0)	
Fishing and trapping	138	116	194	822	516	1,786
	(8.0)	(6.0)	(11.0)	(46.0)	(29.0)	
Mining	447	493	312	432	239	1,923
	(23.0)	(26.0)	(16.0)	(22.0)	(12.0)	
Manufacturing	3,050	8,915	6,616	4,653	1,484	24,718
	(12.0)	(36.0)	(27.0)	(19.0)	(6.0)	
Construction	3,304	6,140	7,695	12,734	6,035	35,908
	(9.0)	(17.0)	(21.0)	(35.0)	(17.0)	
Transportation	2,512	3,447	2,340	2,699	1,220	12,218
	(21.0)	(28.0)	(19.0)	(22.0)	(10.0)	
Trade	12,498	30,574	17,365	9,359	2,159	71,955
	(17.0)	(42.0)	(24.0)	(13.0)	(3.0)	
Finance	9,966	8,654	2,728	944	153	22,445
	(44.0)	(39.0)	(12.0)	(4.0)	(1.0)	
Business & per. service	30,311	45,844	23,336	14,338	4,750	118,579
	(26.0)	(39.0)	(20.0)	(12.0)	(4.0)	
Public administration	964	1,238	814	957	413	4,386
	(22.0)	(28.0)	(19.0)	(22.0)	(9.0)	
Total	70,275	111,391	64,868	51,794	19,889	318,217

Note: Never Subsidized is based on RBT never >1; Occasionally Subsidized is defined as RBT >1 for 1–3 years; Sometimes Subsidized is defined as RBT >1 for 4–6 years; Frequently Subsidized is defined as RBT >1 for 7–10 years; and Always Subsidized is defined as RBT >1 for all 11 years. Numbers in ( ) are row percentages.

transfer and a further 17 per cent receive one for just one to three years. In a similar fashion a significant fraction of firms operating in surcharged industries frequently or always receive a subsidy. In Mining as many firms receive a net transfer in seven or more years out of eleven as do those in three or fewer. A substantial one-quarter to one-third of firms in Manufacturing, Transportation, and the Public Sector also fall into the former category.

In sum these data suggest that the behaviour and characteristics of individual firms may play a significant role in determining both between- and within-industry patterns in the flows of UI funds. It is therefore informative to explore what fraction of the variance in the RBT ratios is industry-specific, firm-specific, or due to other factors. We adopt the approach used in Anderson and Meyer (1993) by estimating the following equation.

$$\mathrm{RBT}_{jpt} = \alpha_t + \beta_p + \delta_i + \gamma_j + \varepsilon_{jpt} \tag{2}$$

The dependent variable is RBT ratio for firm j in province p in year t. This is modelled as a function of a number of fixed effects: α_t captures changes from year to year; β_p and δ_i are province and industry effects respectively; γ_j captures differences between firms; and ε_{jpt} serves as an error term. Note that the subscript for dependent variable is *jpt* because each firm may have more than one plant located in different provinces in a given year. Province fixed effects are included in the model because assessments of the nature of cross-subsidies through UI are often cast in regional terms. Using Least Squares we estimate a series of models of this sort by successively adding each block of fixed effects, with the change in the adjusted R^2 from the most restrictive to the least restrictive versions providing a measure of the relative contribution of province, industry, firm and other factors to the total variance in the RBT ratio. The data cover firms located in the 10 provinces. Self-employed firms, those located in one of the territories, as well as those with an unknown industry are excluded from sample.

Table 13 shows the changes in adjusted R^2 by five different specifications from the most restrictive to the least restrictive, respectively, for the firms operating in every year between 1986 and 1996. Column 1 includes only year dummies in the regression and shows no year effect. The impact of business cycle or any other year effect is likely to be removed by the standardization on the countrywide RBT ratio. In column 2, the province effect significantly increases adjusted R^2 by 10 percentage points showing substantial cross-subsidies between provinces. The next two columns include, respectively, one-digit and three-digit industry indicators. Adding the one-digit industry dummies (Column 3) further increases the adjusted R^2 another 10 percentage points, and an additional 3.6 percentage points when the finer industry categories are used (Column 4). The most significant gain in adjusted R^2, however, is found when firm dummies are introduced. The final column shows that adding firm dummies results in a large increase in the adjusted R^2: an additional 35 percentage points to the explained variance, leaving 41% of total variance unexplained. The effect of province and industry may be influenced by the order in which we have introduced the blocks of fixed effects. To assess this we reverse the order by adding industry dummies first then the province dummies. The results are in the second row of the top panel. The between industry effect now has a larger impact with a 13 percentage points increase in adjusted R^2. The size of inter-industry effect is about the same as before, but the inclusion of province effect only adds about 6.7 percentage

Table 13. Analysis of Variance in Relative Benefit-Tax Ratios: Firms Operating in Every Year between 1986 and 1996.

Dependent Variable: RBT Ratio for Firm *j* in Year *t* and Province *p*

Specifications	Adjusted R^2				
	(1)	(2)	(3)	(4)	(5)
	Year	(1) + Province	(2) + One-digit SIC	(3) + three-digit SIC	(4) + Firms
All	0.0005	0.1027	0.2071	0.2435	0.5888
All*	0.0005	0.1316	0.1766	0.2435	0.5888
By One-Digit Industry					
Agriculture	0.0042	0.1516	–	0.1737	0.6024
Forestry	0.0049	0.2457	–	0.2466	0.5910
Fishing/Trapping	0.0790	0.1777	–	0.1810	0.4472
Mining	0.0072	0.2055	–	0.2490	0.5836
Manufacturing	0.0004	0.1043	–	0.2709	0.5888
Construction	0.0080	0.1612	–	0.1657	0.4785
Transportation	0.0009	0.1247	–	0.1488	0.5696
Trade	0.0006	0.0872	–	0.1036	0.5190
Finance	0.0003	0.0257	–	0.0388	0.3161
Service	0.0010	0.0678	–	0.1462	0.5140
Public Administration	0.0023	0.1476	–	0.1479	0.5239
By Province					
Newfoundland	0.0046	–	0.0817	0.2191	0.6388
PEI	0.0082	–	0.2463	0.3186	0.5813
Nova Scotia	0.0026	–	0.2038	0.2974	0.6545
New Brunswick	0.0082	–	0.3455	0.4014	0.6959
Quebec	0.0014	–	0.1022	0.1580	0.5494
Ontario	0.0052	–	0.0843	0.1237	0.3855
Manitoba	0.0024	–	0.1217	0.1630	0.4333
Saskatchewan	0.0039	–	0.1249	0.1670	0.4573
Alberta	0.0056	–	0.0555	0.0809	0.2801
BC	0.0067	–	0.0853	0.1374	0.4149

Note: 1,058 firms are dropped because of a location outside 10 provinces, and further 4,920 excluded because the industry is unknown. The resulting sample for long-lived firm is 2,907,757. All plants operated by a firm within a province are aggregated, but plants operated in more than one province contribute separate observations.

*The adjusted-R^2 with reverse regression order for SIC and province variables. One-digit SIC is added after year effect in Column (2), then three-digit SIC in Column (3) and province effect in Column (4).

points. Both results suggest that variations in the RBT ratio across firms are much greater than that across industry and province. Among the explained variation in the RBT, 59% can be attributed to firms, about 11–17% to province-specific factors, and the remaining 24–30% to industry-specific factors.

We also extend the estimation by examining each one-digit industry as well as each of 10 provinces separately, offering the lower panel of Table 13. Once again there is no year effect, but adding province fixed effects produces quite distinct results across industries. For example, provincial controls increase the explained variation by as much as 25 percentage points in Forestry, but only 2.6 percentage points in Finance. The inter-industry variation (at the three-digit level) is generally unimportant except in manufacturing, registering a gain of 17 percentage points in the adjusted R^2. Firm effects are still dominant but the impacts are quite different across industries. Adding firm dummies results in an additional 42 percentage-point gain in the explained variance for Agriculture, Transportation and Trade, but only 27 points for Fishing and Finance. These results echo findings from Table 12. Industries that have a high proportion of both subsidized and surcharged firms tend to have more important firm effects.

Similarly, the effects of industry are also different across provinces. Adding one-digit industry dummies increase the adjusted R^2 by nearly 35 percentage points in New Brunswick, but less than six percentage points in Alberta. The within-industry variation is the largest in the Atlantic Provinces (especially in Newfoundland), least important in Alberta. Adding firm dummies again results in a significant increase in explained variation for most provinces. It is, however, surprising that industry-specific variation is more important than firm-specific variation in provinces such as PEI and New Brunswick, suggesting heterogeneity among industries rather than firms is significant factor in determining cross-subsidization in these provinces.

This analysis is repeated in Table 14 using the entire population of firms, those operating for at least one year and highlighted in the bottom panel of Table 8. The general conclusions continue to hold when this larger population is examined, though the overall explanatory power falls. The much larger sample size causes the analysis to bump up against some computing limitations, and as a result some of the estimations are based upon a Monte Carlo analysis, the details of which are in the note to the table. With respect to the country wide results the full set of fixed effects accounts for about 30% of the total variation, about half of that when the sample of long-lived firms is used. However, there is a sense in which the firm effects are even more important as they raise the adjusted R^2 more than seven-fold from just 4% to 30%. At the same time the explained variance in some of the industry specific analyses rival that obtained for the sample of long-lived firms. In the case of finance it is actually higher. But in all cases the firm fixed effects play the dominant role. If the causes of business failure are firm specific as opposed to industry specific then it may well be expected that firm effects

Table 14. Analysis of Variance in Relative Benefit-Tax Ratios: All Firms Operating for at least One Year between 1986 and 1996.

Dependent Variable: RBT Ratio for Firm *j* in Year *t* and Province *p*

Specifications	Adjusted R^2				
	(1)	(2)	(3)	(4)	(5)
	Year	(1) + Province	(2) + 1 digit SIC	(3) + 3 digit SIC	(4) + Firms
All	0.0003**	0.0201**	0.0359**	0.0429**	0.3064**
All*	0.0002**	0.0161**	0.0282**	0.0411**	0.2975**
By One-Digit Industry					
Agriculture	0.0048	0.0970	–	0.1044	0.4839
Forestry	0.0064	0.0981	–	0.0982	0.3920
Fishing/Trapping	0.0250	0.0511	–	0.0542	0.4506
Mining	0.0114	0.0845	–	0.0973	0.3873
Manufacturing	0.0013	0.0638	–	0.1292	0.4313
Construction	0.0029	0.0782	–	0.0803	0.3174
Transportation	0.0019	0.0822	–	0.0902	0.4169
Trade	0.0012	0.0544	–	0.0632	0.4003
Finance	0.0005	0.0314	–	0.0383	0.3776
Service	0.0014	0.0826	–	0.1075	0.3404**
Public Administration	0.0017	0.1019	–	0.1053	0.5051
By Province					
Newfoundland	0.0071	–	0.0214	0.0722	0.4064
PEI	0.0078	–	0.1017	0.1459	0.4876
Nova Scotia	0.0026	–	0.0852	0.1231	0.4865
New Brunswick	0.0074	–	0.1157	0.1635	0.5143
Quebec	0.0015	–	0.0383	0.0609	0.2691**
Ontario	0.0038	–	0.0368	0.0513	0.2306**
Manitoba	0.0041	–	0.0554	0.0692	0.3632
Saskatchewan	0.0042	–	0.0460	0.0566	0.3775
Alberta	0.0049	–	0.0180	0.0276	0.2969**
BC	0.0061	–	0.0372	0.0541	0.3562**

Note: The sample includes all firms ever operating during 1986–1996. Self-employed firms, as well as firms located outside the 10 provinces, and those with unknown industry are dropped. The resulting sample consists of 8,685,112 firm-year observations. All plants operated by a firm within a province are aggregated, but plants operated in more than one province contribute separate observations.

*The adjusted-R^2 with reverse regression order for SIC and province variables. Here one-digit SIC is added after year effect in Column (2), then three-digit SIC in Column (3) and province effect in Column (4).

**To ease the computing burden we run the fixed-effect models by randomly selecting a sample of 5,000 firms from the underlying population for the country-wide analysis in rows one and two of the table, and a sample of 1,000 firms at industry and province level. Once a firm was selected, all year records associated with this particular firm were also selected for fixed-effect analysis. A Monte-Carlo-type simulation was performed with 30 replications, and the mean results are reported. All simulation results are significant at five per cent level.

will continue to dominate. Further, if the causes of business failure are firm specific in the context of competition within an industry then it may also be expected that the explanatory power within industries would be even greater. This explanation meshes with the analysis in Baldwin (1995), which emphasizes the role of intra-industry competition and managerial practices in explaining firm failures. In this sense the firm effects are capturing much more than just an implicit-contract between firms and workers. As firms fail and lay off workers this could lead to large increases in the RBT, while still being a firm specific result.

In sum, the conventional view of high UI cross-subsidization in Canada is often interpreted as the result of geography and an unavoidably large proportion of seasonal employment. However, estimates from these fixed effect models suggest that a substantial proportion of explained variance in RBT ratios is firm-specific. For long-lived firms geography and industry are not as important in determining cross-subsidization once across-firm variations are considered. When combined with findings for all firms, these results suggest that within-industry cross-subsidization may be a more important source of persistent cross-subsidization. There are a considerable number of firms predictably and persistently receiving subsidies year after year regardless of the geographical and industrial conditions they face.

5. ESTIMATING EFFICIENCY COSTS

The economic framework for an analysis of the efficiency costs associated with cross-industry/firm subsidies is well known. The simple static model assumes that there are only two firms (or sectors of identical firms) and that workers are completely mobile between them. One sector has a stable demand for labour and the other doesn't. If a perfectly experience rated unemployment insurance programme is in operation, one in which expected benefits paid are equal to contributions, a competitive labour market leads to an equilibrium allocation of labour at a common wage rate. If UI is not perfectly experience rated the less stable sector receives a subsidy from the stable sector, which reduces its labour costs and shifts its demand for labour so that it increases its size at the expense of stable sector. This transfer results in a misallocation of labour characterized by a welfare or deadweight loss (DWL). Under the assumption of linearity this can be calculated for a given sector as $1/2\,\Delta W\,\Delta N$. Our data can be used to estimate dollar values of the efficiency loss associated with the Canadian UI programme for every year from 1986 to 1996. The DWL can be expressed as a fraction of total payroll

$$\frac{\text{DWL}}{WN} = \frac{1}{2}\frac{\Delta W \Delta N}{WN} = \frac{1}{2}\left(\frac{\Delta N}{N}\frac{W}{\Delta W}\right)\left(\frac{\Delta W}{W}\right)^2 = \frac{1}{2}\eta_{LL}S^2 \tag{3}$$

where η_{LL} is the wage elasticity of demand for labour and S represents the dollar subsidy to an industry over its total payroll. Our analysis is not an attempt to make a definitive estimate of these costs in large part because of the uncertainty in the literature over the true value of the elasticity of labour demand and the appropriate level of aggregation. Further, it is not clear over what time span is appropriate for the measurement of the DWL. Rather our calculations are intended to illustrate the sensitivity of the results to these issues. For this reason we calculate DWL for an elasticity of one and invite readers to scale the results according to their reading of this literature.[7] The DWL is estimated at four different levels by deriving the subsidy at the one-digit SIC, the one-digit SIC and province, the three-digit SIC and province, and the firm level. As mentioned by Anderson and Meyer (1993), the first three estimates are likely understatements because they all assume firms in a given industry (and by extension province) have the same subsidy rate. The use of industry aggregates disguises the across-firm variation and would result in a lower estimate of efficiency loss. We are able to assess the significance of this by also calculating the subsidy at the firm level.

The sample used in these derivations includes all firms operating in any year between 1986 and 1996 in the 10 provinces. Owner operated firms without paid employees and firms in unknown industries are excluded. An example of the calculation is provided in Table 15 using information on the subsidy at the one-digit SIC level. Estimates of DWL for other levels are calculated in a similar way with more cells (provinces, three-digit SIC, or firms) involved. Columns (1) through (3) represent total industry employment, annual payroll, and annual subsidy, respectively. Column (4) offers the per cent of subsidy over payroll which is labelled S in Eq. (3). The dollar value of the subsidy per employee is given in Column (5). Finally, the DWL is presented in Column (6) assuming $\eta_{LL} = 1$. This example is based on data for 1986. The primary industries (agriculture, forestry, and fishing) have fairly high subsidies over their payroll. For example, in the fishing industry the subsidy amounts to nearly 66% of total payroll, and annual subsidies per worker are as high as \$5,321. On the other hand, every worker in transportation as well as in the public sector was paying a \$440 surcharge. The total DWL in this example is about \$126 million, with almost one-third

Table 15. Accounting for Subsidy on One-Digit SIC Level (Excluding Self-Employment Firms).

Industry (One-Digit SIC 80)	Employment	Annual Payroll	Annual Subsidy	% Subsidy over Payroll	$ Value of Subsidy per Employee	$ Value of Deadweight loss
	(1) ('000s)	(2) ($'000s)	(3) ($'000s)	(4)	(5)	(6) ($'000s)
Agriculture	294	2,124,633	250,583	11.79	852.69	14,777
Forestry	125	1,615,523	267,584	16.56	2,149.01	22,160
Fishing	18	142,500	93,518	65.63	5,320.80	30,687
Mining	277	9,224,763	89,826	0.97	324.42	437
Manufacturing	3,077	70,985,163	−200,478	−0.28	−65.15	283
Construction	1,115	16,726,608	1,173,121	7.01	1,051.97	41,138
Transportation	1,201	33,106,311	−527,324	−1.59	−439.13	4,200
Trade	3,356	47,163,421	−149,713	−0.32	−44.61	238
Finance	1,167	25,157,384	−402,524	−1.60	−344.94	3,220
Service	6,609	87,912,063	−630,242	−0.72	−95.36	2,259
Public admin.	1,654	40,905,769	−727,505	−1.78	−439.81	6,469
Total	18,892	335,064,138	−763,155			125,869

Note: All dollar values expressed in1997 dollars. Column (1–3) are derived directly from BNOP files. Column (1) represents the number of T4s issued; Column (4) = $[(3)/(2)] \times 100$; Column (5) = (3)/(1); Column (6) = $1/2 \times (\text{Column } 4)^2 \times \eta_{LL} \times (2)$ assuming $\eta_{LL} = 1$.

($41 million) coming from construction alone, and 24% and 18% from fishing and forestry, respectively. In this example the DWL from manufacturing comprises only 2% of the total loss. As mentioned, these calculations likely underestimate the true value because of the assumption that all firms in a given industry in all provinces have the same subsidy rate. Total UI benefits paid in this year accounted for just over $10 billion, orders of magnitude greater that the estimated DWL.

Estimates of the total dollar value of the DWL for the years 1986–1996 inclusive is offered in Table 16 for each one-digit industry and using different levels of aggregation in the calculation. The DWL is calculated for each cell and then summed across all cells in each broad industry. The estimates are very sensitive to the level of aggregation used in deriving the subsidy level. When the calculation is based on one-digit industries the total DWL is about $1.75 billion, about one per cent of total benefits paid in this period. However, the estimated DWL increases rapidly as finer industry and across-province variations are considered. When across-firm variations are taken into account it reaches a $27.6 billion, about 16.5% of total UI benefits. This is nearly 16 times larger than the estimate based on the one-digit SIC,

Table 16. Variations in Estimates of Deadweight Loss by Level of Aggregation (1986–1996).

Industry (One-Digit SIC-80)	Level of Aggregation Upon which Calculation of Subsidies is Based			
	One-Digit SIC	One-Digit SIC Within Province	Three-Digit SIC Within Province	Firm Level
Agriculture	116,276	220,473	296,288	1,239,740
Forestry	228,773	547,112	564,925	1,260,322
Fishing	361,982	578,376	596,200	2,301,065
Mining	1,602	12,191	64,439	222,002
Manufacturing	6,761	454,164	1,744,463	5,982,893
Construction	786,193	1,158,460	1,228,539	3,917,573
Transportation	49,382	65,934	173,516	2,025,111
Trade	2,817	98,473	272,808	4,306,576
Finance	44,107	49,364	70,112	397,981
Service	50,593	166,092	685,773	4,937,575
Public administration	100,827	132,320	146,862	976,109
Total	1,749,313	3,482,958	5,843,925	27,566,949
Percentage of total UI benefits	1.05	2.08	3.50	16.5

Note: Expressed in thousands of 1997 dollars.

and five times larger than the estimate using the three-digit SIC/province variations.

Table 16 also shows that the increases of DWL are not distributed proportionally across industry when a finer level of aggregation is used. The most significant change concerns the role of the manufacturing sector. Manufacturing's share of the total DWL rises from 0.4% ($6.8 million) with one-digit SIC information to 21.7% ($6 billion) with firm-level information, indicating a good deal of heterogeneity among firms in this sector. Surprisingly, Services and Trade surpass Construction and are the second and third largest contributors ($4.9 billion and $4.3 billion, respectively) to the total DWL when firm-level information is used. Construction's share of the total drops from as high as 45% based on one-digit SIC to only 14% when across-firm sources of variations are recognized.

Once again it should be stressed that all of these estimates are based on the assumption of the unit labour demand elasticity. As such they are not meant to represent estimates of the actual DWL. If we apply the lower ($\eta_{LL} = 0.5$) and upper ($\eta_{LL} = 2.6$) bounds of elasticity suggested by a survey of the existing literature the total DWL could be as low as \$13.8 billion or as high as \$71.8 billion. Furthermore, Anderson and Meyer (1993) also note that the true DWL would be larger if a distinction could be made between the average and marginal subsidy.

6. CONCLUSION

The research summarized in this paper uses administrative data on the universe of Canadian firms, workers, and UI claimants to paint a picture of patterns in the use of UI. Firms and industries are the units of analysis. We document patterns in the flow of UI benefits and contributions, and examine their nature.

There are at least four major findings. First, the Canadian UI programme – in spite of significant changes in eligibility rules and benefit entitlements and rates since the early 1970s – entails a relatively stable and long-lasting pattern of transfers across industries and provinces. Second, when examined at a finer level these patterns reflect subsidies and surcharges that are concentrated among particular industries. Some industries never receive a net transfer from the programme; others always do. To some important degree these patterns reflect greater than average separation rates (particularly temporary separations) and lower than average wages (and hence contributions). In contrast to the other determinants of cross-subsidization – benefit durations and weekly benefit rates – both of these dimensions can be significantly influenced by the firm or reflect the implicit or explicit contract between employers and employees. The third major conclusion deals with the finding that individual firm effects are very important in understanding the variations in Benefit/Tax ratios across and within industries. Our analysis of firm effects focuses in part on long-lived firms, those operating in all 11 years under study, for two reasons: they represent a significant proportion of economic activity, accounting for over 70% of all jobs; and credible long-term contracts (either implicit or explicit) between employers and employees are most likely to have evolved among this sector. We find that cross-subsidies occur not only between industries but also within them. Most "always-subsidized" firms belong to "always-subsidized" industries, but many "never subsidized" firms are also part of these same industries. Our fourth major finding refines this point and suggests that while industry and province effects represent 20–25% of the total variation in Benefit/Tax ratios, firm effects account for as much as 35%. In addition, the impact of firm effects is very different across industries, accounting for over 40% of explained variation in some industries but as less than 30% in others.

Our work raises two major implications for the economic analysis of the labour market consequences of UI. First, we point out that estimates of the DWL associated with no experience rating of UI contributions are very sensitive to the level of aggregation. Incorporating firm-level information in the calculation of efficiency losses leads to higher estimates than those based just on industry information. More generally, our findings also suggest that it is important to use perspectives on the interaction between UI and the labour market that recognize the role of the demand side of the market in future analysis and policy making. Implicit contract models might in this sense prove particularly valuable.

NOTES

1. The benefit rate was reduced to 57% from 60% in 1993, and to 55% (60% for low-income claimants) in 1994. In addition, those quitting without just cause were no longer eligible for benefits beginning in 1993.

2. Our definition of a temporary separation may be more liberal than often used. Individuals are considered to have experienced a temporary separation if they are found to have employment income from the same firm in the tax year after the year of separation. In the extreme this would classify an individual who experienced a separation of almost two years from the same firm as temporary if the separation occurred early in the year and the rehire late in the next year. See Appendix A for more details.

3. The entries for Tables 2 and 3 are calculated using the formula $B_i - T_i(B/T)$, where B_i represents benefits received and T_i taxes paid by a particular industry/province (B and T represent benefits and taxes for Canada as a whole). The industry/province contributions are multiplied by the country wide Benefit/Tax Ratio (B/T) because the UI account was not exactly in balance over the period. In essence, the \$1.95 billion annual surplus is allocated to each industry/provinces in proportion to the contributions made. The result represents the excess of benefits over taxes for each industry/province that would prevail if the overall program were in balance. In a similar manner the entries for Table 4 are derived as $(B_i/T_i)/(B/T)$.

4. A careful reading of these three studies will reveal notable variations in RBT ratios in certain industries (especially in primary sectors), but no change in status between subsidized and surcharged status. Further, some important part of the explanation for these variations has to do with differences in the industry coding (SIC 1970 versus SIC 80). We produced information similar to that presented in Tables 1 through 3 for 1997 and this general conclusion would continue to hold using this additional year of data, the first full year in which substantial changes associated with legislation that renamed the program "Employment Insurance" came into effect.

5. This is available under the same title at www.statcan.ca.

6. Potential claimants had to accumulate between 10 and 14 insured weeks of employment in order to qualify for UI benefits. The exact number of weeks depended upon the unemployment rate in the applicant's region of residence. This eligibility rule was known as the VER. It was introduced in December 1977, but with the stipulation that it would expire after three years. Each year successive governments passed enabling legislation to prevent it from sun-setting. This was done until 1990 when the government of the day bundled the enabling legislation with a broader legislative package associated with the introduction of the Goods and Services Tax. Passage of this package was delayed in third reading with the result that the VER was suspended and reverted to14 weeks in all regions regardless of economic conditions. This had a disproportionate impact in high unemployment regions, notably many parts of the Atlantic provinces where the entrance requirement had historically been 10 weeks. This was the case from mid-February to mid-November.

7. Hamermesh (1993) reviews various studies of the estimates of constant-output labour demand elasticity among developed countries from both aggregate and micro economic data. In his summary, the mean estimate of $-\eta_{LL}$ is 0.39 for studies using aggregate data, while the mean value is 0.45 for those using micro-economic data. He suggests that a reasonable range for $-\eta_{LL}$ is probably between 0.15 and 0.75 for the typical firm. However, several studies suggest that Canada has a relatively higher elasticity of labour demand. Appendix C summarizes estimates of labour demand elasticity for Canada. In general, nearly all-Canadian studies produce estimates greater than 0.5. The magnitudes could go as high as 2.6 in Symons and Layard (1984) or 2.24 in Lawerence (1989). Lawrence shows the own price elasticity of labour demand increases from 0.21 in 1962 to 2.24 in 1980. He suggests that the Canadian economy has become more price responsive in recent decades owing to increasing openness in international trade, deregulation, and associated improvements in flexibility.

ACKNOWLEDGEMENTS

An earlier draft of this paper was presented to the first IZA/SOLE Transatlantic Meeting of Labor Economists in Buch/Ammersee, Germany and to the 2002 meetings of the Canadian Economics Association at the University of Calgary. The authors thank participants at these presentations, an anonymous referee, and the editor for their comments. At the same time the responsibility for the content of this paper rests solely with the authors and in particular should not be attributed to Statistics Canada.

REFERENCES

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, *67*(2), 251–333.

Anderson, P. M., & Meyer, B. D. (1993). The unemployment insurance payroll tax and interindustry and interfirm subsidies. In: J. L. Poterba (Ed.), *Tax policy and the economy* (Vol. 7). Cambridge, MA: NEBR.

Atkinson, A. B., & Micklewright, J. (1991). Unemployment compensation and labor market transitions: A critical review. *Journal of Economic Literature*, *29*(December).

Baldwin, J. R. (1995). *The dynamics of industrial competition: A North American perspective*. New York: Cambridge University Press.

Corak, M. (1993a). Unemployment insurance once again: The incidence of repeat participation in the Canadian UI program. *Canadian Public Policy*, *19*(June).

Corak, M. (1993b). Is unemployment insurance addictive? Evidence from the benefit durations of repeat users. *Industrial and Labor Relations Review*, *47*(October).

Corak, M. (1994). Unemployment insurance, work disincentives, and the Canadian labour market: An overview. In: C. Green, F. Lazar, M. Corak & D. Gross (Eds), *Unemployment insurance: How to make it work*. Toronto: C.D. Howe Institute.

Corak, M., & Pyper, W. (1995a). *Firms, industries and cross-subsidies: Patterns in the distribution of UI benefits and taxes*. UI Evaluation Technical Report. Human Resource Development Canada.

Corak, M., & Pyper, W. (1995b). *Workers, firms and unemployment insurance* (Cat. No. 73–505). Ottawa: Statistics Canada.

Feldstein, M. (1976). Temporary layoffs in the theory of unemployment. *Journal of Political Economy*, *84*(October).

Gray, D., & Sweetman, A. (2001). A typology analysis of the users of Canada's unemployment insurance system. In: S. Schwartz & A. Aydemir (Eds), *Essays on the repeat use of UI: The earnings supplement project*. Ottawa: Social Research and Demonstration Corporation.

Haltiwanger, J. C., Lane, J. I., Spletzer, J. R., Theeuwes, J. J. M., & Troske, K. R. (Eds) (1999). *The creation and analysis of employer–employee matched data*. Amsterdam: Elsevier.

Hamermesh, D. S. (1990). Unemployment insurance, short-time compensation and labor demand. *Research in Labor Economics*, *13*.

Hamermesh, D. S. (1993). *Labor demand*. Princeton: Princeton University Press.

Karagiannis, E. (1986). *Experience rating UI premiums: An assessment*. Unpublished mimeo, 120 pp. Employment and Immigration Canada.

Kesselman, J. R. (1983). *Financing Canadian unemployment insurance*. Canadian Tax Paper No. 73. Toronto: Canadian Tax Foundation.

Lawerence, D. (1989). An aggregator model of Canadian export supply and import demand responsiveness. *Canadian Journal of Economics*, *22*(3), 503–521.

Lemieux, T., & MacLeod, W. B. (1995). *State dependence and unemployment insurance*. UI Evaluation Technical Report. Human Resource Development Canada.

Lemieux, T., & MacLeod, W. B. (2000). Supply side hysteresis: The case of the Canadian unemployment insurance system. *Journal of Public Economics*, *78*, 139–170.

Lin, Z. (1998). Employment insurance in Canada: Recent trends and policy changes. *Canadian Tax Journal*, *46*(1).

Statistics Canada. (1988). *Developing a longitudinal database on businesses in the Canadian economy* (Cat. No. 18–501E). Ottawa: Statistics Canada.

Statistics Canada. (1992). *Worker turnover in the Canadian economy: separations and hiring, 1978–1989* (Cat. No. 71–539). Ottawa: Statistics Canada.

Symons, J., & Layard, R. (1984). Neoclassical demand for labour functions for six major economies. *Economic Journal*, *94*, 788–799.

APPENDIX A. DATA DEVELOPMENT

The analysis is based upon a number of administrative data sets. These include the Benefits and Overpayments (BNOP) file, T4 information, and data from the Longitudinal Employment Analysis Program (LEAP). The BNOP contains information on all UI claims initiated in a given year. Data from 1986 through 1996 is used to derive the total number of claims, the total amount of benefits paid, and the average duration of benefit receipt for the workers of each firm. Each BNOP record contains a Payroll Deduction Account Number associated with a particular firm. These account numbers are established and used by Revenue Canada for tax remittance purposes. A firm may have several account numbers. These are all aggregated up to the firm level using the information in LEAP, a longitudinally consistent catalogue of all firms operating in Canada. (See Statistics Canada (1988) for a detailed description of this file.) A firm is defined according to the Longitudinal Business Register Identifier as used in LEAP. The categorization of a claim as being due to a temporary or a permanent separation is also done in the manner of Statistics Canada (1992). A temporary separation is said to have occurred if the individual had any employment earnings from the same firm in the year following the separation. This is determined by whether or not the firm has issued a T4 indicating some earnings for that individual. If an individual initiates more than one UI claim in a given year the firm information on each record in the BNOP is used to determine if the claims were supported with employment from the same firm and the first claim is categorized directly as resulting from a temporary or permanent separation.

The T4 is also the source of information on the amount of UI contributions made. T4s are issued by firms to all paid employees, and used for tax purposes. They also have a payroll deduction account number and these are aggregated to the firm level using the LEAP in the same manner as the BNOP information. Total contributions by the workers of a firm are summed from the T4 file, and employer contributions are derived by marking these up by 1.4, reflecting the legislated employer contribution rate. No adjustments are made for contribution reductions to those firms participating in a wage-loss reduction plan. The error introduced by this is small. UI contributions of self-employed fishermen are not available in the T4. As such this group is not included in any of the tabulations. The number of T4s issued is used as an indication of the number of jobs in each firm or industry over the course of a given year. While there are a small number of cases in which employers issue more than one T4 per job to their paid employees, equating a T4 with a job does not entail too much of an error.

(The exception to this is the fishing industry, which is dominated by self-employed fishermen. It is not uncommon for these individuals to receive two or three T4Fs in a single calendar year).

The structure the Payroll Deduction Account Numbers changed in 1997 with the result that a longitudinally consistent labelling of firms beyond this year is not possible.

APPENDIX B. UI CONTRIBUTION RATES AND MAXIMUM INSURABLE EARNINGS, 1986–2001

Year	Contribution Rate		Maximum Annual Insurable Earnings	Maximum Annual Contribution
	Employer	Employee		
1986	\$3.29	\$2.35	\$25,740	\$1,452
1987	\$3.29	\$2.35	\$27,560	\$1,555
1988	\$3.29	\$2.35	\$29,380	\$1,657
1989	\$2.73	\$1.95	\$31,460	\$1,473
1990	\$3.15	\$2.25	\$33,280	\$1,797
1991	\$3.15	\$2.25	\$35,360	\$1,910
	(\$3.92)	(\$2.80)		(\$2,377)
1992	\$4.20	\$3.00	\$36,920	\$2,659
1993	\$4.20	\$3.00	\$38,740	\$2,790
1994	\$4.30	\$3.07	\$40,560	\$2,990
1995	\$4.20	\$3.00	\$42,380	\$3,052
1996	\$4.13	\$2.95	\$39,000	\$2,762
1997	\$4.06	\$2.90	\$39,000	\$2,714
1998	\$3.78	\$2.70	\$39,000	\$2,527
1990	\$3.57	\$2.55	\$39,000	\$2,387
2000	\$3.36	\$2.40	\$39,000	\$2,246
2001	\$3.15	\$2.25	\$39,000	\$2,107

Note: The rates indicated by ( ) became effective part-way through 1991.

APPENDIX C. SELECTED STUDIES ON THE ESTIMATES OF LABOUR DEMAND ELASTICITY FOR CANADA

Study	Category	Description	Elasticity ($-\eta_{LL}$)
Homogeneous labor			
Pindyck (1979)	Constant-output demand elasticity	Aggregate on large industries, annual 1963–1973, translog cost function	0.66
Symons and Layard (1984)	Varying-output demand elasticity	Manufacturing employment, quarterly 1956–1980	2.6
Halvorsen and Smith (1986)	Constant-output demand elasticity	Aggregate on small industry (Metal mining), annual 1954–1974, translog cost function	0.51
Lawerence (1989)		Aggregate import and export industries, 1962–1980, flexible functional form	0.21–2.24
Wylie (1990)	Constant-output demand elasticity	Aggregate on small industry (four two-digit manufacturing), annual 1900–1929, translog cost function	0.51
Card (1990c)	Constant-output demand elasticity	Aggregate on firm level (union contracts), 1968–1983	0.62
Currie (1991)	Constant-output demand elasticity	Aggregate on firm level (Ontario's teachers' contracts), 1975–1983,	0.53–0.68
Christofides and Oswald (1991)	Constant-output demand elasticity	Aggregate on firm level (union contracts), 1978–1984	$<$ 0–0.22
Heterogeneous labor			
Merrilees (1982)		Aggregate, annual 1957–1978, 4 labour types	
		Young men	−0.56
		Young women	0.44
		Adult men	0.07
		Adult women	−0.11
Ferguson (1986)		Atlantic provinces, 1966–1979, 7 labour types	0.33–1.00

Source: Hamermesh (1993) (Chapter 3).

THE IMPACT OF DEUNIONISATION ON EARNINGS DISPERSION REVISITED

John T. Addison, Ralph W. Bailey and
W. Stanley Siebert

ABSTRACT

This paper examines the effects of union change in Britain on changes in earnings dispersion 1983–1995. We investigate not only the decline in union density but also the greater wage compression among unionised workers, as well as changes in union density across skill groups. For the private sector, we find that deunionisation accounts for little of the increase in earnings dispersion. What unions have lost on the swings (lower density), they have gained on the roundabouts (greater wage compression). But for the public sector we find strong effects, because unions are increasingly organising the more skilled. This change in the character of public sector unions means that they no longer reduce earnings variation nearly as much as they once did.

1. INTRODUCTION

The British earnings distribution has widened considerably since Mrs Thatcher's sustained attack on the unions. The possibility of there being a

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 337–363

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26009-6

connection between the two developments has been the subject of a fairly large literature. In this paper, we revisit the subject, using the general variance decomposition technique first put forward by Freeman (1980) and Metcalf (1982). We follow Card's (2001) modification of this approach to allow for changes in the "structure" of unionisation across the workforce; specifically, the greater decline in union density among the lower paid than the higher paid. Using this method, and allowing for changes in union wage and variance gaps, as well as union density, we show that the effect of deunionisation on earnings dispersion has on the whole been more modest than generally believed. We concentrate on the period up to 1995, because most of the changes in unionisation and earnings dispersion had occurred by this point (see Fig. 1 below).

Certainly, casual inspection shows a striking association between movements in union density over time and changes in the earnings dispersion (see Leslie & Pu, 1996, Fig. 4d). Emphasising this link, Schmitt (1995, p. 201) has calculated that the decline in union density could account for 21 per cent of the rise in the pay premium for a university degree and for 13 per cent of the increase in the non-manual differential, 1978–1988. Machin (1997, p. 653) obtains more dramatic results: comparing 1983 with 1991, he calculates that the male earnings variance would have been 40 per cent less had the 1983 levels of union coverage prevailed in 1991. Bell and Pitt (1998, pp. 520–523)

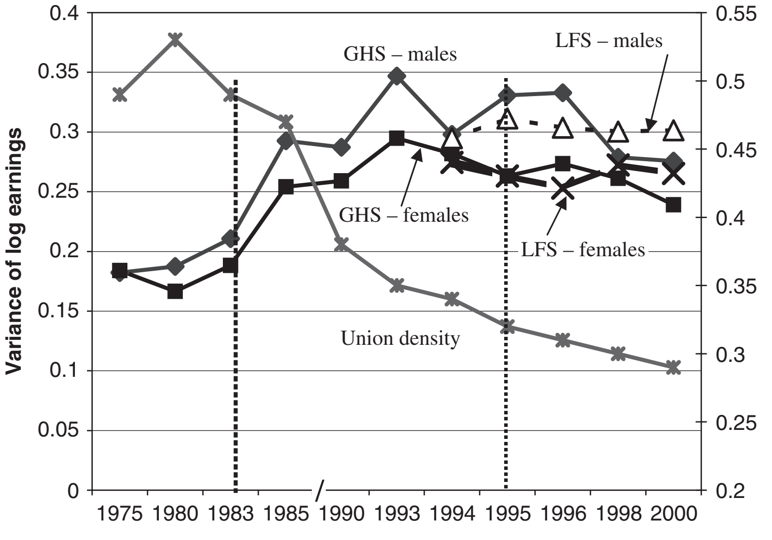


Fig. 1. Earnings Variance in the GHS and LFS Compared.

also conclude that deunionisation between the early 1980s and 1990s widened the male earnings distribution – in this case by about 20 per cent.

That said, not all research points the same way. Notably, in their thorough analysis of the wage distribution of U.K. males, Gosling, Machin, and Meghir (2000, p. 661) emphasise education rather than deunionisation: the way recent cohorts have improved their acquisition of education, as well as changes over time in the returns to education. Moreover, Card (2001) has pointed out that the equalising effects of unionism can be exaggerated if we do not allow for the fact that unionisation effects vary across the wage distribution. He shows that if the *structure* of unionisation changes, so that union density falls less over time for the higher paid – as has happened both in the U.S. and the U.K. (see below) – then estimates of the equalising tendency of unionisation can be reduced.

The plan of the paper is as follows. In the next two sections, we first describe the datasets used before reviewing the variance decomposition approach; here, we also derive some descriptive results on changes in unionisation over time. In the fourth section, we give the results of the variance decomposition analysis. (Because variance decomposition is central to our approach, our measure of wage dispersion is naturally the variance, rather than other commonly used measures such as the Gini coefficient or the ratio of the top to the bottom deciles.) Then, in the fifth section, given the diverging trends of unionisation in the public and private sectors, we present some results for the two sectors separately. The final section provides a summary and conclusion.

2. THE DATA

We require data on earnings, unionisation, and individual characteristics over the last two decades. Just about the earliest dataset available with good union and earnings information is the 1983 General Household Survey (GHS) (OPCS, 1986). 1983 is the only year in which the GHS included a union membership question, but this year is early enough to represent the "golden age" of unionism. The Family Expenditure Survey (Bell & Pitt, 1998), or British Social Attitudes Survey (Blanchflower & Bryson, 2003) also offer possibilities. The Family Expenditure Survey asks a question on whether union dues are paid, from which it would be possible to infer union membership beginning in 1982. However, as Bell and Pitt (1998, p. 515), acknowledge this method is likely to omit union members who do not pay dues regularly. For its part, the British Social Attitudes Survey has union membership and earnings data available from 1984 (SCPR, 1985). However, the earnings data provided are in categories rather than continuous in form, and there is a small sample (867 employees in 1984), which would raise problems for our study of earnings dispersion. The GHS, by contrast, has the advantage of a large sample of employees (over 8,000), which is important since we aim to split the sample into private and public sectors, and analyse males and females separately. Accordingly, we use the 1983 GHS for our early period, as have Machin (1997), Gosling and Lemieux (2001), and Bell and Pitt (1998).

For the later period, we use the Labour Force Survey (LFS), which provides detailed earnings data from 1993 onward. The LFS also provides a large sample of over 8,000 employees. We choose the 1995 LFS (OPCS, 2000), because 1995 represents the nadir of the union movement's fortunes, and well precedes Labour's 1997 election victory. Most of the changes in unionisation and earnings dispersion had occurred by 1995, as shown in Fig. 1 (see also Card, Lemieux, & Riddell, 2003), and we therefore concentrate on this period.

Fig. 1 further indicates that the two datasets are comparable. It can be seen that earnings inequality in the GHS steadily increased from the late 1970s to the early 1990s, with the two surveys yielding similar measures of inequality in 1995. While the measures are more divergent in 2000, both sources agree that the rise in inequality plateaued in the 1990s. Moreover, union status is measured by the same question in both surveys: 'Are you a member of a trade union or staff association?' As regards union coverage, however, which would arguably better address the issue of union impact on wages, the survey questions differ. In the GHS the question is: 'Is there a trade union or staff association where you work, which people in your type of job can join if they want to?' In the LFS the question is simply: 'At your place of work, are there unions, staff associations, or groups of unions?' Hence, as with most of the literature, we restrict the analysis to union membership alone.

As regards the wage variable, we take several steps to ensure comparability. For both datasets, we restrict the sample to individuals aged 16–66 years, and not self-employed. For both, we use the same hourly wage variable computed by dividing weekly earnings by usual hours. In addition, we convert the 1983 wage data to 1995 values using the retail price index. Finally, for both years we trim off observations with implausibly low or high wage rates, excluding hourly wages outside the £1–£45 range. These adjustments have a minor effect. Our 1995 figure for aggregate union density (the percentage of employees who are union members) is 33.1 per cent,

Table 1. Hourly Wage Distributions, Union and Non-Union Workers, 1983 and 1995.

	Men		Women	
	Non-Union	Union	Non-Union	Union
1983				
Union density (%)	56.7		42.1	
Overall variance log wages	0.223		0.192	
Variance log wage	0.289	0.151	0.197	0.147
Mean log wage	1.639	1.854	1.280	1.534
Adjusted union wage gap (t-value)	0.149 (12.9)		0.195 (15.5)	
1995				
Union density (%)	37.4		30.7	
Overall variance log wages	0.309		0.262	
Variance log wage	0.358	0.205	0.241	0.226
Mean log wage	1.876	2.066	1.55	1.89
Adjusted union wage gap (t-value)	0.091 (6.41)		0.195 (13.7)	

Notes: Samples are taken from the 1983 General Household Survey and the 1995 third quarter Labour Force Survey (LFS) with Northern Ireland excluded. Samples comprise respondents aged 16–66 years who were not self-employed and whose hourly wage was between £1 and £45 in 1995 pounds (1983 wages valued in 1995 pounds according to the retail price index). For the LFS, the income weights supplied with the data are used. The adjusted union wage gap is the union coefficient from a regression controlling for years of education, years of experience (plus experience squared and cubed), and dummies for non-white, marital status, and 5 regions.

comparable with Brook's (2002), (Table 1) figure of 32.3 per cent for employees in Great Britain.

3. ACCOUNTING FOR THE IMPACT OF DEUNIONISATION

There are different ways to account for the impact of deunionisation on earnings dispersion. First, various counterfactuals are possible. It is natural to compute the impact of deunionisation by asking what earnings dispersion would be if union density had not declined. However, there are two other important dimensions of unionism: the union wage gap, and the variance gap (the difference in the variance of wages for union and non-union workers). It is worth considering counterfactual changes in these dimensions as well. Second, as noted above, we can allow for differences in union density across skill groups. Let us look at these points in turn.

Beginning with the basic two-sector formulation, average wages $\overline{w}$ are

$$\overline{w} = U\overline{w}^u + U'\overline{w}^n \tag{1}$$

where U is union density, $U' = 1-U$ and the superscripts u and n refer to union and non-union, respectively. This equation can be rewritten in terms of union "power", namely, union density multiplied by the union/non-union wage gap

$$\overline{w} - \overline{w}^n = U\Delta_w \tag{2}$$

where $\Delta_w = \overline{w}^u = \overline{w}^n$ is the wage gap. This equation shows that the term $U\Delta_w$ determines the extent to which average wages are pushed above non-union wages; hence, the conventional use of the term "union power." It is important to consider how union power differs across the skill groups, which we do below.

The impact of unionism on the variance of average wages is what we wish to assess. Eq. (1) provides a framework for estimating this effect. According to this equation, the variance of wages can be expressed in terms of union density, and the union–non-union wage and variance gaps. Using Freeman's formula (1980, p. 19) the variance (V) is

$$V = V^n + U\Delta_v + UU'\Delta_w{}^2$$

or,

$$D = V - V^n = U\Delta_v + UU'\Delta_w{}^2 \tag{3}$$

where $\Delta_v = V^u - V^n$ is the union–non-union variance gap, V^u and V^n being the variance of wages in the union and non-union sectors, respectively. The impact of unionism on the overall wage variance is then D, namely, the overall wage variance minus the (larger) wage variance that would prevail without unionism. As can be seen, the impact can be decomposed into a term involving the union variance gap, $U\Delta_V$, the so-called *within-sector* effect, which is generally negative since Δ_v is generally negative. The impact will also depend on the term $UU'\Delta_w$, the *between-sector* effect, which is positive since unions widen wage dispersion due to the union wage gap. Note that the impact of unionism depends not only upon U but also upon the wage and variance gaps, Δ_w and Δ_v.

In assessing the impact of unionism on changes in wage variance over time – our focus here – we need to hypothesise what would have happened if unionism had taken a different path, that is, develop a counterfactual.

Various approaches are possible. First, let us write an equation for the change in union impact, ΔD, between time periods 0 and 1

$$\Delta D = \Delta V - \Delta V^n = U_1\Delta_{1v} - U_0\Delta_{0v} + U_1 U_1' \Delta_{1w}{}^2 - U_0 U_0' \Delta_{0w}{}^2 \qquad (4)$$

where $\Delta D = D_1 – D_0$, $\Delta V = V_1 – V_0$, and $\Delta V^n = V_1^n – V_0^n$. The counterfactual here is then the change in the non-union wage variance, ΔV^n For example, if deunionisation is causing a decline in union impact on the wage variance, the (negative) impact of unionisation will be smaller absolutely in period 1 than period 0; that is, $\Delta D > 0$. This condition requires the change in the overall wage variance to be greater than the change in the non-union wage variance, or $\Delta V > \Delta V^n$. Thus, changes in the non-union wage variance are meant to control for changes in the "other factors" which determine the overall wage variance.

We can also develop a counterfactual by writing

$$\Delta D = \Delta_{1v}\Delta U + \Delta_{1w}{}^2\Delta(UU') + U_0\Delta\Delta_v + U_0 U_0' \Delta\Delta_w{}^2 \qquad (5)$$

or,

$$\Delta D = \Delta D' + \Delta D'' \qquad (5')$$

where $\Delta D' = \Delta_{1v}\Delta U + \Delta_{1w}{}^2\Delta(UU')$; $\Delta D'' = U_0\Delta\Delta_v + U_0 U_0' \Delta\Delta_w{}^2$; $\Delta U = U_1 - U_0$; $\Delta(UU') = U_1(1-U_1) - U_0(1-U_0)$; $\Delta\Delta_v = \Delta_{1v} - \Delta_{0v}$; and $\Delta\Delta_w{}^2 = \Delta_{1w}{}^2 - \Delta_{0w}{}^2$.

In other words, the change in union impact can be decomposed into two parts: $\Delta D'$, the change in impact due to movements in U alone, weighted by period 1's wage and variance gaps, and $\Delta D''$, the change in impact due to movements in wage and variance gaps, weighted by period 0's U level. $\Delta D'$ is sometimes reported (e.g., Machin, 1997, p. 653), since it builds a natural counterfactual based on changes in union density *alone*. However, variance gaps are also important as a measure of union power. These gaps have in fact increased over time in Britain, as we will see. Therefore, while we will report $\Delta D'$ values for comparative purposes, we will generally rely on the ΔD measure.

Let us now turn to the point that unionisation varies across skill groups. A way of showing this variation, following Card (2001), is to define skill groups using predicted earnings percentiles based on the non-union wage structure. We can then compare union densities across these skill groups.[1] We can also consider how union "power" (viz. density multiplied by the wage gap, noted earlier) varies across skill groups. The picture for males (females) is given in Figs. 2a and 2b (Figs. 3a and 3b).

Fig. 2a shows that, for males in 1983, union density was lowest among the least skilled (lowest decile), highest at the third decile and then somewhat lower for the more skilled. Corresponding data for 1995 show density falling most among the least skilled, leaving the highest density at the top decile. The male union density measure thus suggests that unions help a labour

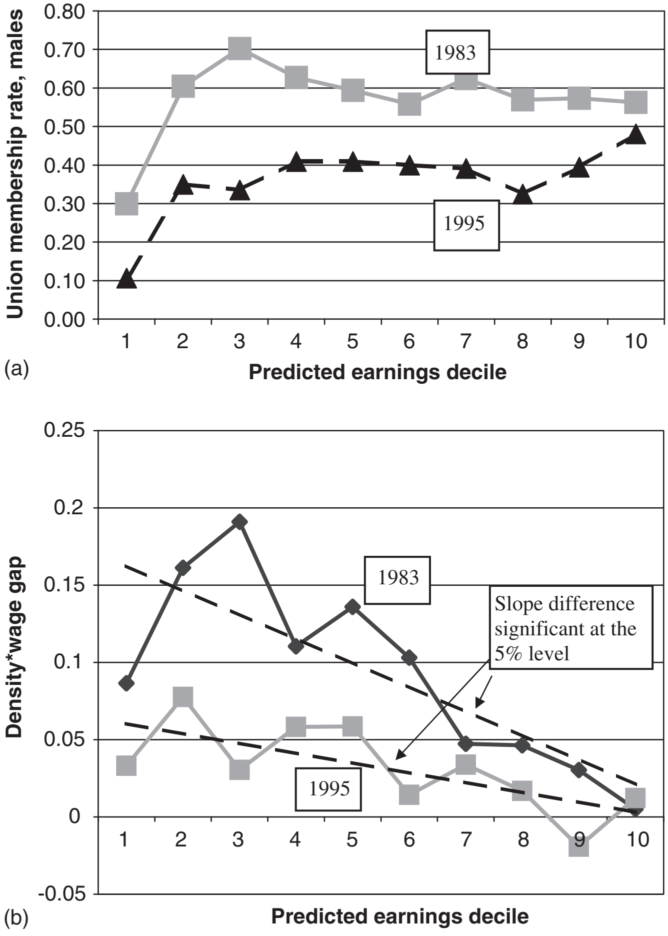
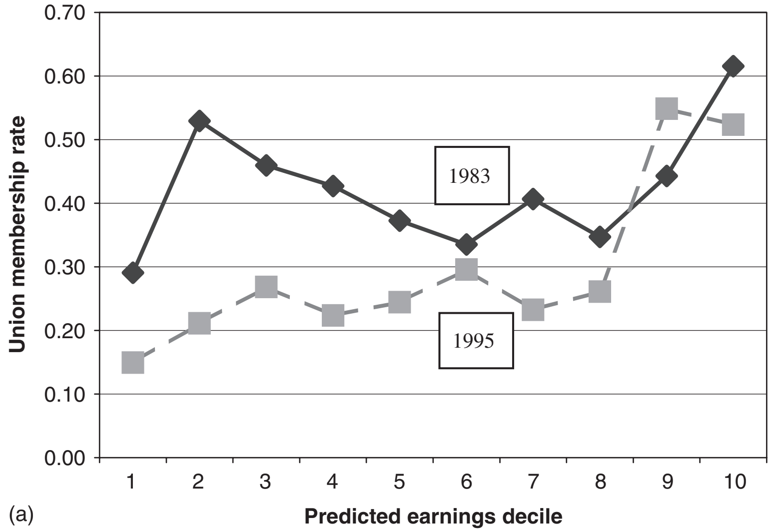


Fig. 2. (a) Union Density by Skill, (b) Union Power by Skill, Males in 1983 and 1995.



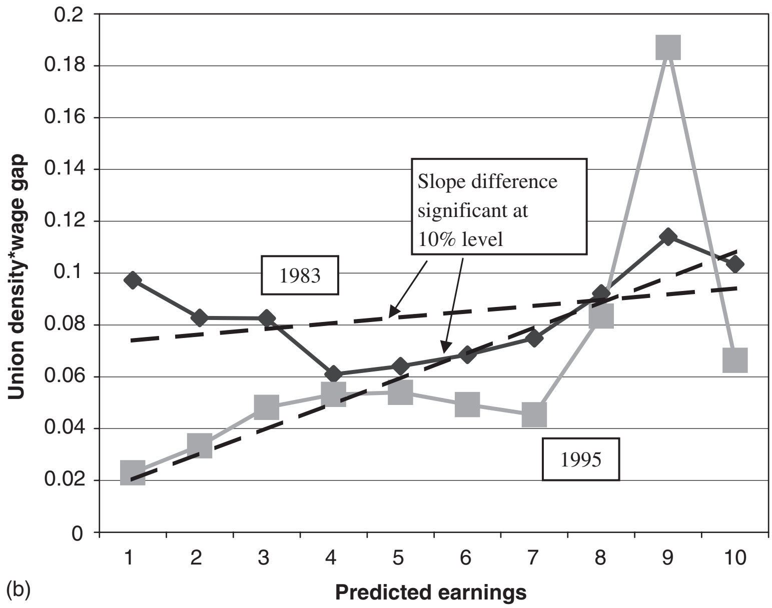


Fig. 3. (a) Union Density by Skill, (b) Union Power by Skill, Females, in 1983 and 1995.

"elite". However, the picture is different for union power. Fig. 2b shows that union power was definitely greater both in 1983 and in 1995 for the least skilled. Nevertheless, it is evident that there has been a significant fall in union power among this group by 1995. For females, the union density and union power graphs are more similar. Over time, both density and power have fallen among low-skill groups, but have remained quite steady in the top three deciles. Thus, Fig. 3b shows that, particularly in 1995, there is a positive covariance between union power and skill for women, so that unionisation appears to benefit an elite.[2]

Allowing for different union effects by skill category requires modification of Eq. (3). Card (2001, p. 298) shows that the formula becomes

$$V = V^{n*} + \overline{U\Delta_V} + \overline{U(1-U)\Delta_w{}^2} + \mathrm{Var}[U(c)\Delta_w(c)] + 2\mathrm{Cov}[w^n(c),\ U(c)\Delta_w(c)] \tag{6}$$

V^{n*} is the non-union wage variance, namely, the variance that would result if all workers were paid according to the non-union wage structure.[3] $U(c)$ is union density in the c groups, $\Delta_v(c)$ are the variance gaps, $\Delta_w(c)$ are the wage gaps, and $w^n(c)$ are the non-union wage rates. The over-bar terms indicate averages over the c skill categories, and are analogous to the terms in Eq. (3). But the terms for variance and covariance between categories are new, and in practise we will find the covariance to be the most important. This covariance is precisely that between skill, w^n, and union power, $U\Delta_w$, which we have been discussing above in connection with Figs. 2b and 3b. A negative covariance term will indicate that unions are more helpful to the least skilled, and this will pull the overall wage variance, $V(c)$, below the variance prevailing without unions, V^{n*}. A positive covariance term indicates the opposite.

Over time, as Figs. 2b and 3b have made clear, union power in the cases of both men and women has been shifting towards more skilled workers (the covariance term in Eq. (6) is becoming less negative). This factor will have offset the equalising tendency of unions brought about, in particular, by the variance gap. We now consider the size of these effects.

4. FINDINGS

4.1. The Economy as a Whole

Table 1 contains panels for 1983 and 1995 that show how the overall variance in log wages has increased over the period. For men the increase has been 86 log points (from 0.223 to 0.309), and for women it has been 70 log points (from 0.192 to 0.262). These large increases are what we are concerned to explain. Notice that the increase in wage variance for non-union workers has been smaller: 69 points for men (0.289 to 0.358), and 44 points for women (0.197 to 0.241). Thus, forces operating on the non-union sector alone cannot explain the increase in overall wage variance, suggesting a role for deunionisation. The table also shows that the union wage variance is lower than the non-union variance, thereby pointing to the equalising effect of greater unionisation.[4] Interestingly, it can also be seen that while both the union and non-union wage variances have risen over time, the union variance for men remains much smaller than the corresponding non-union variance: the variance gap has even increased. In other words, even though they are less extensive than heretofore, male unions can still strongly "standardise" their members' wages.

Table 1 also contains information on the wage gap, both unadjusted and adjusted for a set of conventional human capital variables.[5] The unadjusted wage gaps are always larger than the adjusted gaps because union workers have higher skills than their non-union counterparts. However, the difference between adjusted and unadjusted wage gaps grows between 1983 and 1995, reflecting the increased unionisation of high skill groups in 1995. For men, the adjusted wage gap falls over time as well, reflecting reduced union power on this dimension (but we must remember that male unions can still standardise members' wages). By contrast, female unions seem to have increased their power to bring about a wide wage gap (0.205 in 1995, compared to 0.197 in 1983), but not so strongly to standardise their members' wages.

We now estimate basic union effects on wage dispersion, using Eq. (3). The results are given in Table 2. Taking males in 1983, for example, the within-sector effect is $U\Delta_V = -0.078$, which is negative because the variance gap is negative. The between-sector effect is $U(1-U)\Delta_W{}^2 = 0.006$ which is positive, following the wage gap, but bound to be small since the wage-gap term is squared. The total effect is −0.072. This figure represents a sizeable contribution – about one-third – to reducing male wage variance in 1983 (0.223 from Table 1). In 1995, the impact is smaller, −0.055, or about one-sixth of the male wage variance (0.309 from Table 1). Taking changes over time, as in Eq. (4), male deunionisation contributes to a rise in wage variance of 0.017, which is 19.8 per cent of the overall increase. Turning to women, we see that in 1983 unionism is weakly egalitarian, reducing wage variance by −0.012. By 1995, however, women's unionism actually *widens* the wage variance by 0.004. Over time, then, the impact of deunionisation for women is similar – namely, 22.9 per cent – but is achieved by a different route.

Table 2. Basic Estimates of the Contribution of Declining Unionisation to Wage Inequality, 1983–1995.

	Men	Women	Remarks
1983			
Union density, U	0.567	0.421	From Table 1
Union wage gap, Δ_W	0.151	0.197	Adjusted difference between union and non-union wages (Table 1)
Union variance gap, Δ_V	−0.138	−0.050	Difference in union and non-union wage variances (Table 1)
Union effect, *between* sectors, $U(1-U)\Delta_W^2$	0.006	0.009	Small effect of unions in raising wage inequality by widening mean pay as between union and non-union sectors
Union effect, *within* sectors, $U\Delta_V$	−0.078	−0.021	Larger effect of unions is to reduce wage dispersion within union sectors
Total union effect	−0.072	−0.012	Estimated total effect of unions is to reduce wage variance; for example, for men the reduction is −0.072
1995			
Union density, U	0.374	0.307	From Table 1
Union wage gap, Δ_W	0.086	0.205	
Union variance gap, Δ_V	−0.153	−0.015	
Union effect, *between* sectors, $U(1-U)\Delta_W^2$	0.002	0.009	See explanations for 1983 above
Union effect, *within* sectors, $U\Delta_V$	−0.057	−0.005	
Total union effect	−0.055	0.004	Variance-reducing effect of unions is smaller for men in 1995 than 1983, and unions even increase dispersion for women in 1995
Changes: 1983–1995			
Change in variance of wages	0.086	0.070	See Table 1; for example, for men 0.086 = 0.309–0.223
Change in effect of unions	0.017	0.016	Change in total union effect derived above; for example, for men 0.017 = −0.055–(−0.072)
Contribution of unions (%)	19.8	22.9	For example, for men 0.198 = 0.017/0.086

Table 2. (*Continued*)

	Men	Women	Remarks
Memo item			
Amount 1995 V would be lowered given 1983 U (%)[a]	0.030 (34.3)	0.002 (2.4)	This figure depends mainly on $(U_1-U_0)\ \Delta_{v1}$, the change in U weighted by the 1995 variance gap. This gap is small for women; hence the 2.4% figure

[a]This number gives $\Delta D'$, the deunionisation effect assuming changes in union density alone; see Eq. (5′).

The last row of Table 2 shows the different estimates for deunionisation that are obtained when we use the counterfactual, $\Delta D'$, from Eq. (5′). It will be recalled that here we are estimating what the 1995 wage variance would have been had the 1983 level of union density prevailed, taking as given the 1995 union wage and variance gaps. Using this method, deunionisation contributes 34.3 per cent to the widening in the male wage variance, but only 2.4 per cent in the case of females. However, as we have also noted, this method ignores changes in wage and variance gaps.[6]

The next step is to allow for differences in union structure (i.e. in coverage and in wage and variance gaps) across skill groups, where the latter are defined using Card's (2001) predicted earnings deciles. We have already seen (from Fig. 2b) how union power, for men, although tending to be pro-poor, has become less so with the passage of time. And the trend is the same for women (Fig. 3b). Table 3 now quantifies the impact of these trends.

The estimates in Table 3 indicate a reduced impact of deunionisation on wage dispersion for men, although not for women. Looking first at men, unions reduce overall wage variance in both years: by −0.041 in 1983 and by −0.042 in 1995. However, as can be seen, the reduction is as great in 1995, which implies that deunionisation cannot be a factor in the widening male wage variance. To put this finding another way: the counterfactual variance of wages if all were paid according to the non-union wage structure, V^{n^*}, has increased by 0.087, which is as much as the increase in the overall wage variance, 0.086. Since the male non-union wage variance has increased so much, there is little room for a deunionisation effect.

The main factor behind the strong variance-reducing effect of unions for men in 1995 is the larger variance gap term: $\overline{U\Delta_V} = -0.033$ in 1995 compared with −0.024 in 1983 (see the lower panel of the table). In other words, unions standardise their members' pay more in 1995 than 1983. This factor

Table 3. Adjusted Estimates of the Contribution of Declining Unionisation to Wage Inequality, Allowing for Different Union Effects across Pay Deciles.

	Men	Women	Remarks
1983			
Variance of wages, V	0.223	0.192	From Table 1
Adjusted variance of non-union wages, $V^{n}*$	0.264	0.207	Allowing for different union impacts across pay deciles (see Notes below)
Adjusted union effect	−0.041	−0.015	Example for men −0.041 = $V - Vn*$; (see text Eq. (6))
1995			
Variance in wages	0.309	0.262	From Table 1
Adjusted variance in non-union wages, $V^{n}*$	0.351	0.261	Allowing for different union impacts across pay deciles (see Notes below)
Adjusted union effect	−0.042	0.001	E.g. for men −0.041 = $V - Vn*$; (see text Eq. (6))
Changes: 1995 – 1983			
Variance of wages ΔV	0.086	0.070	For men, unionism reduces wage dispersion about as much in 1995 as 1983. So decline of unions cannot have increased dispersion. But for women, unionism has a role
Adjusted variance of non-union wages $\Delta V^{n}*$	0.087	0.054	
Adjusted union effect	−0.001	0.016	
Union effect is % of ΔV	0	23	

	Men		Women	
	1983	1995	1983	1995
$\overline{U\Delta_V}$	−.024	−0.033	−0.030	−0.022
$\overline{U(1-U)\Delta_w{}^2}$	0.009	0.003	0.011	0.009
$\mathrm{Var}[U(c)\Delta_w(c)]$	0.003	0.001	0.001	0.002
$2\mathrm{Cov}[w^n(c), U(c)\Delta_w(c)]$	−0.028	−0.013	0.004	0.013
Total	−0.041	−0.042	−0.016	0.001
Memo: Average variance gap $\overline{\Delta_V}$	−0.04	−0.09	−0.06	−0.06

Notes: The adjusted formula, allowing for different union effects on wage variance by skill category, is given in Eq. (6) in the text. Values for the terms in the equation (taken from the $c = 10$ decile groups in Appendix B) are given above.

counteracts the tendency for union power to become less pro-poor, as shown by the diminution of the covariance term (see also the significant flattening of the union power line in Fig. 2b). On the other hand, the adjusted and simple estimates are similar for women. The variance-reducing effect of unions is estimated to be much larger in 1983 (at −0.015) than in 1995 (0.001). For women, union power has tended over time to become less egalitarian (see also Fig. 3b).[7] Consequently, the change in the *character* of women's unionisation appears to play a considerable role in the widening of women's wage variance.

These results differ from the received wisdom. In particular, it seems that the increase in wage dispersion for men can hardly be attributed to deunionisation. What unions have lost on the swings (less power among the unskilled) they have gained on the roundabouts (more wage compression for their members). It is true that deunionisation still seems to have a role to play in explaining increased wage dispersion among women. Nevertheless, we conclude that the equalising effects of unions are less than might be thought. Let us now consider whether distinguishing between the public and private sectors upsets this conclusion.

4.2. Public-Private Sector Comparisons of Unionism

It is interesting to assess the impact of deunionisation on wage inequality in the public and private sectors separately, since union trends have been so different. As can be seen from Table 4, public sector union density in 1995 is 78–86 per cent of its 1983 value. Indeed, some public sector groups such as women with further or higher education, have even maintained or increased their union density reflecting the rise in unionism among teachers and nurses. However, private sector density has declined considerably. In particular, the 1995 value for women (men) is now only 57 (69) per cent of the 1983 value.

At the same time, the private and public sectors are similar in that the more educated categories have maintained their union density better than less educated groups. The picture is best appreciated with the aid of Figs. A1 through A4 in Appendix A, which graph the union power variable – union density multiplied by the wage gap – against predicted earnings (the covariance term in Eq. (6)). Men and women are shown separately by sector. As can be seen, the 1995 relationship is significantly less negatively sloped than that for 1983 in all cases (though marginally for private sector males),

Table 4. Trade Union Membership Rates, 1983 and 1995.

	Men			Women		
	1983	1995	Ratio 1995/1983	1983	1995	Ratio 1995/1983
(a) Private sector						
Overall	41.4	27.5	68.8	26.0	14.9	57.3
By education						
Degree or equivalent	13.4	18.4	94.8	30.2	14.2	47.0
Further education	40.6	24.3	59.9	27.3	21.8	79.8
'A' level or equivalent	39.3	32.2	81.9	20.2	17.7	87.6
'O' level or equivalent	30.0	17.2	57.3	21.4	14.4	67.3
Other	47.8	38.3	80.1	21.4	11.8	55.1
None	49.8	25.5	51.2	30.9	13.4	43.4
Observations	2,851	3,199		2,149	2,875	
(b) Public sector						
Overall	85.1	66.3	77.9	68.9	59.4	86.2
By education						
Degree or equivalent	81.2	71.9	88.5	76.1	73.3	96.3
Further education	85.2	78.1	91.7	73.9	79.0	1.07
'A' level or equivalent	83.6	56.2	67.2	68.2	45.2	66.3
'O' level or equivalent	79.5	60.8	76.5	64.3	46.8	72.8
Other	85.5	55.4	64.8	65.0	49.0	75.4
None	88.9	75.8	85.3	67.9	46.5	68.5
Observations	1,535	979		1,334	1,582	

Note: Public sector employment is defined to include nationalised industries, public corporations, or central or local government.

indicating that the more educated have maintained their union power better than the less educated.

We now calculate the basic union effects on wage dispersion. The necessary data on union density, and the wage and variance gaps are assembled in Table 5. It is interesting to note how, in the public sector, even though union density has been maintained, there have been considerable changes in wage gaps and variance gaps. For males, the wage gap – both the raw and the adjusted gap – has fallen almost to zero. However, the *variance gap* has been maintained, indicating that public sector unions have retained the power to compress male wages. However, for females in the public sector, there have been opposite tendencies, with the wage gap in particular rising.

Basic estimates of the impact of deunionisation, following Eq. (3), are given in Table 6. This table is analogous to Table 2 for the whole economy. For example, for private-sector men in 1983, −0.062 is an estimate of the amount by which unionisation reduces the wage variance. As can be seen,

Table 5. Hourly Wage Distributions, 1983 and 1995.

	Men		Women	
	Non-Union	Union	Non-Union	Union
(a) Private sector				
1983				
Union density (%)	41.4		26.0	
Overall variance log wages	0.231		0.168	
Variance log wage by group	0.291	0.131	0.179	0.112
Mean log hourly wage	1.62	1.78	1.22	1.41
Adjusted union wage gap (t-value)	0.128 (8.60)		0.202 (10.9)	
1995				
Union density (%)	28.5		15.1	
Overall variance log wages	0.314		0.239	
Variance log wage by group	0.359	0.187	0.242	0.198
Mean log hourly wage	1.85	1.97	1.52	1.68
Adjusted union wage gap (t-value)	0.113 (6.3)		0.133 (5.96)	
(b) Public sector				
1983				
Union density (%)	85.1		68.9	
Overall variance log wages	0.178		0.172	
Variance log wage by group	0.250	0.162	0.201	0.154
Mean log hourly wage	1.77	1.89	1.49	1.61
Adjusted union wage gap (t-value)	0.112 (4.46)		0.095 (4.99)	
1995				
Union density (%)	66.5		59.5	
Overall variance log wages	0.238		0.235	
Variance log wage, by group	0.294	0.206	0.216	0.206
Mean log hourly wage	2.09	2.18	1.67	2.00
Adjusted union wage gap	0.016 (0.57)		0.191 (9.49)	

Note: See Table 2.

the impact of unions has fallen over time in both public and private sectors, just as for the economy as a whole. However, the fall has been particularly marked for women in the public sector, implying a greater role for de-unionisation (in terms of Eq. (4), the inequality $\Delta V > \Delta V^n$ holds strongly for this group). This is a surprising result given the fact that their union density has fallen least. The penultimate row gives the basic estimates for the contribution of deunionisation to the increased wage variance: 18.1 per cent for private-sector men, 5.6 per cent for private-sector women, 23.3 per cent for public-sector men, and 54.0 per cent for public-sector women. The final row shows, as a matter of interest, the very different estimate we would obtain using the counterfactual $\Delta D'$ of Eq. (5′).

Table 6. Basic Estimates of the Contribution of Declining Unionisation to Wage Inequality in the Private and Public Sectors, 1983–1995.

	Private Sector		Public Sector	
	Men	Women	Men	Women
1983				
Union effect, *between* sectors, $U(1-U)\ \Delta_W{}^2$	0.004	0.008	0.002	0.002
Union effect, *within* sectors, $U\Delta_V$	−0.066	−0.017	−0.075	−0.032
Total effect	−0.062	−0.009	−0.073	−0.030
1995				
Union effect, *between* sectors, $U(1-U)\ \Delta_W{}^2$	0.002	0.002	0.000	0.010
Union effect, *within* sectors, $U\Delta_V$	−0.049	−0.007	−0.059	−0.006
Total effect	−0.047	−0.005	−0.059	0.004
Changes: 1983–1995				
Change in variance of wages	0.083	0.071	0.060	0.063
Change in effect of unions	0.015	0.004	0.014	0.034
Contribution of unions (%)	18.1	5.6	23.3	54.0
Memo item				
Amount 1995 V would be lowered given 1983 U (%)	0.022 (26.7)	0.005 (6.8)	0.016 (27.3)	0.001 (1.5)

Note: See Table 2.

We now turn to estimates that allow for different union effects by skill category. The results are given in Table 7, which is analogous to Table 3 for the whole economy. For men in the private sector, as for the economy as a whole, the adjusted estimate is smaller than the basic estimate. This outcome is primarily because the variance gaps within skill categories are smaller than the variance gap for the sector. An indication of this fact is provided in the memo item in the last row of the table, which gives the average variance gap across skill categories, $\overline{\Delta_V}$. For private sector men in 1983 this gap averages −0.05, whereas for the private sector as a whole it is −0.160 (= 0.131 – 0.291, Table 5).[8] At the same time, notice how the variance gap for men in this group has increased over time, from −0.05 to −0.07, as the memo item in the bottom panel indicates. On this measure, then, unions have increased their power over male wages in the private sector, even as union density has declined.

Pushing against this equalising effect of unions for private sector men has been the shift in union membership towards the labour elite. The shift is given by the decline (in absolute value) in the covariance term given in the lower panel of Table 7. The shift is also illustrated by the flatter union power graphs for 1995 (see Fig. A1 in Appendix A). For private-sector men, the net

Table 7. Adjusted Estimates of the Contribution of Declining Unionisation to Wage Inequality, Allowing for Different Union Effects across Pay Deciles.

	Private Sector		Public Sector	
	Men	Women	Men	Women
1983				
Variance in wages	0.231	0.169	0.178	0.172
Adjusted variance of non-union wages, V^{n*}	0.263	0.177	247	0.215
Adjusted union effect	−0.032	−0.008	−0.069	−0.044
1995				
Variance in log wages	0.314	0.239	0.238	0.235
Adjusted variance of non-union wages, V^{n*}	0.343	0.242	287	0.251
Adjusted union effect	−0.029	−0.003	−0.049	−0.016
Changes: 1983–1995				
Variance of wages, ΔV	0.083	0.070	0.060	0.063
Adjusted variance of non-union wages, ΔV^{n*}	0.080	0.065	040	0.036
Adjusted union effect	0.003	0.005	020	0.028
Union effect as % of ΔV	3.6%	7.1	33	44

	Private Sector				Public Sector			
	Men		Women		Men		Women	
	1983	1995	1983	1995	1983	1995	1983	1995
$\overline{U\Delta_V}$	−0.020	−0.021	−0.012	−0.007	010	−0.021	−0.018	−0.041
$\overline{U(1-U)\Delta_w^2}$	0.007	004	0.009	0.003	0.007	0.002	0.004	008
$\mathrm{Var}[U(c)\Delta_w(c)]$	0.003	001	000	000	023	004	006	0.002
$2\mathrm{Cov}[w^n(c), U(c)\Delta_w(c)]$	−0.021	−0.013	−0.006	0.001	−0.109	−0.034	−0.035	0.016
Total	−0.032	−0.029	−0.008	−0.003	−0.069	−0.049	−0.044	−0.016
Memo: average variance gap $\overline{\Delta_V}$	−0.05	−0.07	−0.05	−0.04	01	−0.03	−0.03	−0.06

Notes: See Table 4. The adjusted formula (allowing for different union effects by skill category) for the effect of unions on the variance of wages is given in Eq. (6) in the text. Values for the terms in the equation are given above.

result is that unions reduce earnings variance by about the same amount (around −0.03) in both 1983 and 1995. Therefore, deunionisation has apparently not contributed to the rise in male private sector wage variance.

For the other groups, the adjusted estimates are similar to the basic estimates, though the new method reveals interesting consequences of the change in the nature of unionism, particularly in the public sector. As can be seen, for public-sector men and women, the deunionisation effect remains large, 0.020 and 0.028 respectively (33–44 per cent of the increase in variance). The large effect in the public sector does not result from a fall in union density, as might be thought, but rather from the shift towards elite workers in union organising. This effect is shown by the decline in the covariance term in the lower panel of Table 7, which we have already noted for private sector men, and is also shown in Appendix A's Figs. A3 and A4. In fact, for public-sector women in 1995, the usual negative, pro-poor covariance between skill and union power turned positive, 0.016, as Table 7 shows. This change in the covariance term overwhelms the dispersion-reducing effect of a tendency towards larger variance gaps (for example, for public sector women, the last row of Table 7 shows the average variance gap to have increased in absolute value from −0.03 to −0.06). In short, there has been a change in the character of public-sector women's unionism, which the union density figures alone do not capture.

5. CONCLUSIONS

In this paper, we have analysed the impact of deunionisation on earnings dispersion over the period 1983–1995, taking men and women separately and also distinguishing between the private and public sectors. We have seen that unionism is a many-dimensioned entity. Union density is by no means the most important dimension. The variance and wage gaps attributable to unions are also important. So, too, is the "pro-poor" – or otherwise – distribution of union density. In fact, we show (following Card, 2001) that the distribution of union density has become less pro-poor over time, shifting for example from the less educated to the better educated. Accordingly, the "sword of justice" effect of unions (see Metcalf, 2005, p. 102) has become weaker.

Our headline finding is that the large decline in union density accounts for little of the increase in earnings variation in the private sector, either for men or women. This finding can be explained by allowing for unionism's other dimensions. We show that the variance gap has widened sufficiently over time to offset both the decline in density and the adverse shift in density towards the more skilled. In the private sector, therefore, unions appear to have maintained their power – at least as regards standardising their members' wages – notwithstanding all Mrs Thatcher's reforms.

In the public sector there has been less of a decline in union density. Yet, paradoxically, it is here that unionism has had more of a role to play. In the public sector, as in the private sector, variance gaps – and thus the power to standardise – have been maintained. The difference lies in the shift towards organising the more skilled in the public sector, particularly amongst women. This means that unions no longer reduce earnings variation as much as they once did. Changes in the character of public sector unionism – not so much deunionisation as "re-unionisation" – can thus account for a large percentage (30–40 per cent) of the increased earnings dispersion in the public sector. But, to repeat, of the private sector no such statement can be made.

NOTES

1. The prediction equation is based on Card's (2001, p. 303) specification, and includes years of education, dummies for race, marital status and (5) regions, linear, quadratic and cubed experience, and interactions of five levels of education with linear and quadratic experience. It is fitted to non-union workers only, and then used to assign union and non-union workers into 10 equally sized groups.

2. It is likely that union power is overstated for low-skilled workers, and understated for the high skilled. Card (2001, p. 300) finds that low-skilled union workers have higher unobserved skills than their non-union counterparts, and the opposite for high-skilled union workers. Hence, the true wage (in efficiency units) for the low-skilled union worker will be lower than the observed wage, leading to an overstatement of union power here, with precisely the opposite result for the high skilled. We do not make an adjustment for this factor, but it should be kept in mind when assessing the extent to which union power is "pro-poor".

3. V^{n*} will differ from V^n in Eq. (3). $V^{n*} = \overline{V_i^n} + \mathrm{var}(\overline{X_i^n})$, where $\overline{V_i^n}$ is the weighted average of wage variances of the c groups, and $\mathrm{var}(\overline{X_i^n})$ is the variance of wage averages of the c groups.

4. Union wage variance remains much lower than the non-union variance when we standardise the differences in the characteristics of union and non-union workers. The variance of residuals from a wage regression for union workers is also lower than that for non-union workers.

5. The adjusted union wage gap is the union coefficient from a regression controlling for years of education, years of experience (plus experience squared and cubed), and dummies for non-white, marital status, and 5 regions. As will be seen, this two-sector wage gap does not play a major role in later calculations, and so we do not refine it.

6. The position here would be assessed by computing $\Delta D''$. For men, Δ_v has increased, indicating that 1995 is *superior* for this dimension of union power. Hence, male V in 1995 would be reduced given 1983 Δ_v.

7. We have the counterintuitive result for women that their average variance gap within skill groups, $\overline{\Delta_V}$, is larger than the variance gap for the labour force as a whole, Δ_v. In 1995, for example, $\overline{\Delta_V} = -0.06$ (bottom panel, Table 3), yet $\Delta_v = -0.015$ (Table 2). The reason is that Δ_v depends upon the distribution of union density across skill groups, as well as variance gaps within groups. The fact that most female union members are in the high skill groups, coupled with the fact that variance gaps are small for some of these groups, drives Δ_v down to -0.015.

8. For public sector males in 1983 we have the extreme result that the average within skill group gap $\overline{\Delta_V} = 0$ (Table 7, bottom panel), while the overall gap $\Delta_v = -0.088$ (= 0.162−0.250, Table 5b). This result arises because males in public sector unions in 1983 tended to be found in skill groups with high variance gaps, although variance gaps were zero averaged across skill groups (going the "wrong" way for several groups, with higher variance for union than non-union workers).

ACKNOWLEDGMENTS

The authors acknowledge the Office of Population Censuses and Surveys, Social Survey Division (OPCS), as the originator of the 1983 General Household Survey; the OPCS, the Central Statistical Office, the Departments of Finance and Personnel and of Economic Development (Northern Ireland) and the Commission of the European Communities as originators and sponsors of the 1995 Quarterly Labour Force Survey; and the Data Archive at the University of Essex as the distributor of the data. None of these organisations bears any responsibility for the authors' analysis and interpretation of the data. The authors also thank participants at the July 2004 IZA *Workshop on Wage Inequality* and an anonymous reviewer for their comments.

REFERENCES

Bell, B., & Pitt, M. (1998). Trade union decline and distribution of wages in the UK: Evidence from kernel density estimation. *Oxford Bulletin of Economics and Statistics*, *60*(4), 509–528.

Blanchflower, D., & Bryson, A. (2003). Changes over time in union relative wage effects in the UK and the US revisited. In: J. T. Addison & C. Schnabel (Eds), *International handbook of trade unions* (pp. 197–245). Cheltenham: Edward Elgar.

Brook, K. (2002). Trade union membership: An analysis of data from the autumn 2001 LFS. *Labour Market Trends*, *110*(7), 343–354.

Card, D. (2001). The effect of unions on wage inequality in the U.S. labor market. *Industrial and Labor Relations Review*, *54*(2), 296–315.

Card, D., Lemieux, T., & Riddell, W. C. (2003). Unions and the wage structure. In: J. T. Addison & C. Schnabel (Eds), *International handbook of trade unions* (pp. 246–292). Cheltenham: Edward Elgar.

Freeman, R. (1980). Unionism and the dispersion of wages. *Industrial and Labor Relations Review*, *34*(1), 3–23.

Gosling, A., Machin, S., & Meghir, C. (2000). The changing distribution of male wages in the U.K. *Review of Economic Studies*, *67*(4), 635–666.

Gosling, A., & Lemieux, T. (2001). *Labor market reforms and changes in wage inequality in the United Kingdom and the United States.* NBER Working Paper No. 8413, Cambridge, MA: National Bureau of Economic Research.

Leslie, D., & Pu, Y. (1996). What caused rising earnings inequality in Britain? Evidence from time series, 1970–1993. *British Journal of Industrial Relations*, *34*(1), 111–130.

Machin, S. (1997). The decline of labour market institutions and the rise in wage inequality in Britain. *European Economic Review*, *41*(3-5), 647–657.

Metcalf, D. (1982). Unions and the distribution of earnings. *British Journal of Industrial Relations*, *20*(2), 163–169.

Metcalf, D. (2005). Trade unions: Resurgence or perdition? In: S. Fernie & D. Metcalf (Eds), *Trade unions: Resurgence or demise?* London: Routledge.

OPCS. (1986). General Household Survey, 1983 [computer file]. London: Office of Population Censuses and Surveys Social Survey Division; Colchester, U.K.: The Data Archive [distributor], SN: 2099.

OPCS. (2000). Quarterly Labour Force Survey, June–August, 1995 [computer file]. 3rd ed., Office of Population Censuses and Surveys Social Survey Division, Department of Finance and Personnel (Northern Ireland) Central Survey Unit, Colchester, Essex: UK Data Archive [distributor], SN: 3490.

Schmitt, J. (1995). The changing structure of male earnings in Britain, 1974–1988'. In: R. B. Freeman & L. F. Katz (Eds), *Differences and changes in wage structures* (pp. 177–204). Chicago, Il.: University of Chicago Press.

SCPR. (1985). Social and Community Planning Research, British Social Attitudes Survey, 1984 [computer file]. Colchester, Essex: UK Data Archive [distributor], 1985. SN: 2035.

APPENDIX A

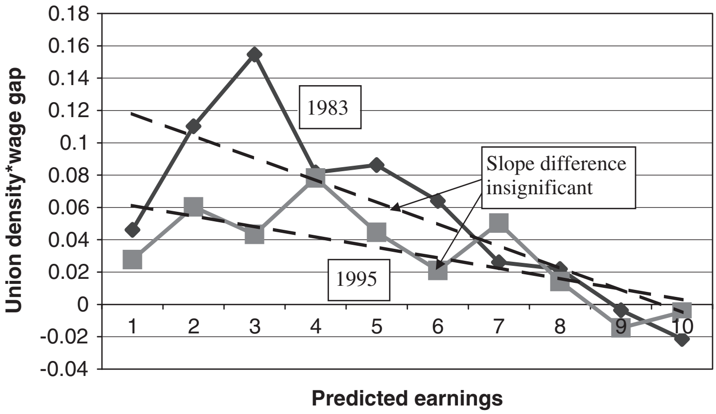


Fig. A1. Union Power by Skill, Private-Sector Males 1983 and 1995.

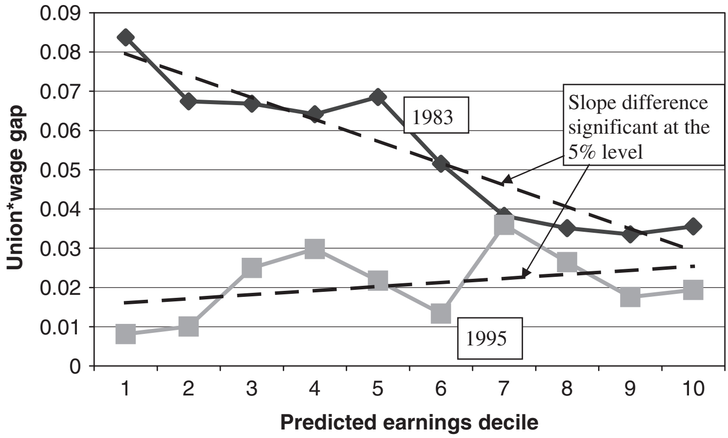


Fig. A2. Union Power by Skill, Private-Sector Females 1983 and 1995.

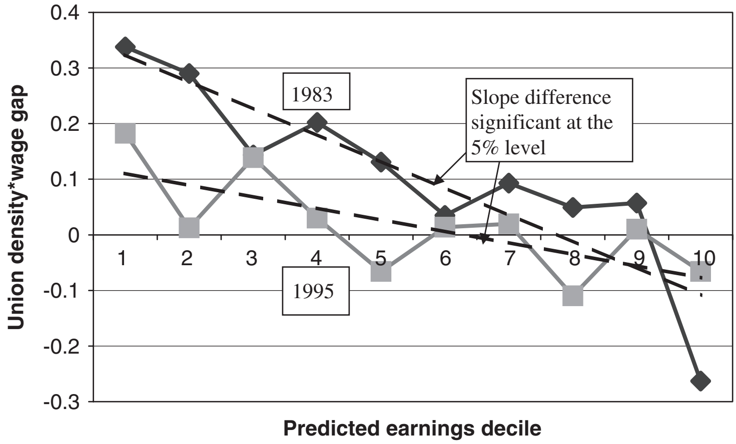


Fig. A3. Union Power by Skill, Public-Sector Males 1983 and 1995.

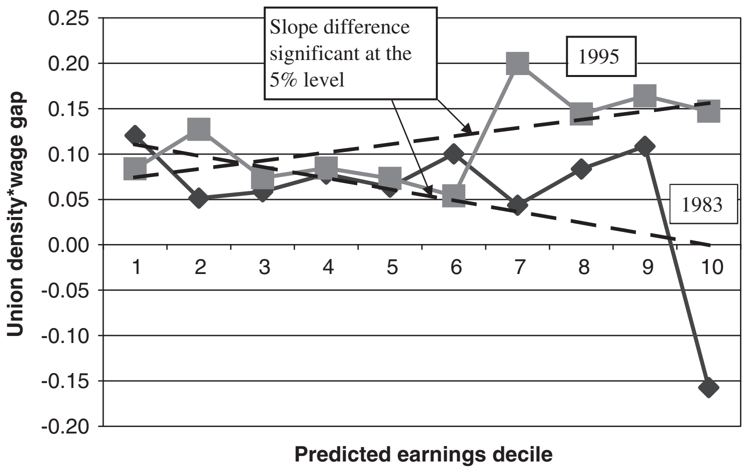


Fig. A4. Union Power by Skill, Public-Sector Females 1983 and 1995.

APPENDIX B

Table B1. Union Membership Rates and Union Wage Effects by Pay Decile.

Predicted Earnings Decile	Men				Women			
	Proportion Union	Log W_N	Wage Gap	Variance Gap	Proportion Union	Log W_N	Wage Gap	Variance Gap
1983								
1	0.22	0.99	0.38	−0.03	0.30	0.93	0.32	−0.11
2	0.52	1.35	0.31	−0.01	0.49	1.12	0.17	−0.05
3	0.68	1.45	0.28	−0.04	0.42	1.13	0.20	−0.04
4	0.66	1.57	0.17	−0.02	0.46	1.19	0.13	−0.03
5	0.64	1.54	0.21	−0.02	0.37	1.24	0.17	−0.05
6	0.55	1.61	0.19	0.06	0.36	1.27	0.19	−0.07
7	0.60	1.71	0.08	−0.02	0.37	1.31	0.20	−0.11
8	0.57	1.83	0.07	−0.05	0.39	1.39	0.24	−0.07
9	0.57	2.00	0.05	−0.05	0.38	1.46	0.30	0.01
10	0.57	2.28	0.01	−0.12	0.59	1.79	0.18	−0.17

1995								
1	0.12	1.24	0.28	−0.06	0.16	1.22	0.15	−0.03
2	0.36	1.53	0.21	−0.06	0.23	1.31	0.14	−0.06
3	0.34	1.72	0.09	−0.09	0.28	1.37	0.18	−0.04
4	0.42	1.73	0.14	−0.10	0.23	1.46	0.24	−0.02
5	0.44	1.83	0.13	−0.02	0.28	1.54	0.17	−0.05
6	0.43	1.98	0.03	−0.12	0.27	1.57	0.18	−0.04
7	0.41	2.02	0.08	−0.02	0.26	1.66	0.18	−0.07
8	0.36	2.14	0.05	−0.12	0.28	1.73	0.30	−0.04
9	0.39	2.42	−0.05	−0.05	0.57	1.85	0.33	−0.11
10	0.48	2.54	0.03	−0.18	0.51	2.21	0.13	−0.14

Notes: Predicted earnings decile is based on a prediction equation for the non-union sector, using an equation with years of education, experience, experience squared and cubed, dummies for marital status, non-white and 5 regions, and interaction of five levels of education with education and linear and quadratic experience. The wage gap is the difference between the log of hourly earnings between union and non-union workers for the given decile. The variance gap is the difference in the variance of log earnings between union and non-union workers for the given decile.

MATERNAL EDUCATION AND CHILD SCHOOLING OUTCOMES IN NEPAL

Diane Dancer and Anu Rammohan

ABSTRACT

This paper uses a sample of school age children from the Nepal Demographic Health Survey (NDHS) to examine the relationship between maternal education and child schooling in Nepal. Taking advantage of the two-stage stratified sample design, we estimate a sample selection model controlling for cluster fixed effects. These results are then compared to OLS and Tobit models. Our analysis shows that being male significantly increases the likelihood of attending school and for those children attending school, it also affects the years of schooling. Parental education has a similarly positive effect on child school, but interestingly we find maternal education having a relatively greater effect on the schooling of girls. Our results also point to household wealth as having a positive effect on both the probability of schooling and the years of schooling in all our models, with the magnitude of these effects being similar for male and female children. Finally, a comparison of our results with a model ignoring cluster fixed effects produces results that are statistically different both in signs and in the levels of significance.

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 365–396

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26010-2

1. INTRODUCTION

Improvements in education and health are crucial to improving labour market productivity. In particular, the positive link between education and labour market returns is well-established in the literature. However, in developing countries where the informal sector dominates, education levels continue to be low, partly because the relationship between education and labour market returns is not so clear-cut. In particular, due to the existence of imperfections in labour and education markets, a complex interaction between economic and social factors leads to a greater role for non-economic factors in household decision making on children's schooling. Hence, an important developmental objective is to identify the factors influencing schooling decisions at the household level, particularly in poor countries where there is no market for educational loans. An additional consideration is that the gender gap in education is particularly severe in South Asia, where adult literacy rates in Nepal (for ages 15 and above) are low, at approximately 48.6% in 2003. These low levels of literacy are likely to have negative effects on labour productivity, and are also likely to perpetuate intergenerational poverty by restricting the ability of future generations to engage productively in the labour market.

Previous research has found maternal education and labour market status to be influential in improving schooling outcomes for children, and girls in particular.[1] However, others studies such as Handa (1996), Rosenzweig and Wolpin (1994), Lillard and Willis (1994), and Unni (1998), have identified a gender dimension where father's schooling has a greater impact on boys relative to girls whereas mother's schooling has the opposite effect. Indeed there is a large literature from developing countries, which finds that resources in the hands of the mother are more likely to be spent on improvements in child health and education.[2] Under such circumstances, maternal education and labour market status are crucial predictors of child schooling.

There are at least three channels through which maternal education levels can improve child-schooling outcomes. The first is the direct effect of a better-educated mother having better employment prospects, higher income levels and higher labour productivity in general. However, since employment in the formal sector is limited in Nepal, particularly for females, it is difficult to estimate the labour market returns accruing to female education. Behrman, Foster, Rosenzweig, and Vahsishtha (1999) however show that, even if the mother did not participate in the labour market, having a better educated mother increases child schooling benefits, through being able to spend more time in home teaching. They argue that this effect is independent of the mother's labour force participation. A third avenue, through which maternal education improves child schooling, is through an improvement in the mother's bargaining power in the household, which in turn is assumed to improve her command over household resources.

The aim of this paper is to examine the role of maternal education and employment status on the schooling participation and schooling levels of children in Nepal. Specifically, we analyse if parental education levels and, in particular, mother's education and labour market status have any gender-specific influences. This paper contributes to the labour econometrics literature in a number of ways. It is the first multivariate study of child education and gender differentials in education using the nationally representative *Nepal Demographic and Health Survey* (*NDHS-2001*). A key contribution of our paper is that our econometric methodology takes advantage of the two-stage sample design of the dataset, where sample households are not randomly distributed over space but are geographically grouped. These geographical groups or clusters could refer to villages in a rural sample. Unlike previous studies that only use cluster effects for correcting standard errors, our econometric strategy recognises the possibility of neighbourhood effects and uses a unique cluster fixed effects technique where households belonging to a cluster (or village) are recognised as having similar characteristics. As Deaton (2000) points out, the cluster fixed effects methodology also allows us to control for unobservable characteristics of the villages (for example, supply side characteristics such as distance to school) and make the analysis robust to the lack of data. Our analysis shows that ignoring these cluster fixed effects in studies that use datasets with cluster-based sampling provides potentially biased results. The cluster fixed effects methodology is discussed in detail in Section 2 below.

The fixed effects methodology could potentially be applied to any cross-sectional dataset where the samples are clustered such as the *Demographic Household Surveys* (DHS) and the *Living Standard Measurement Survey* (LSMS). Datasets such as DHS and LSMS are increasingly collected in a large number of countries and using clustered sampling designs. The *Demographic Household Surveys*, for example, have been conducted in over 60 countries. Fixed effects' modelling is also frequently used with panel datasets such as the *Panel Study of Income Dynamics* (PSID) from the US, which is a longitudinal panel, which has been ongoing since 1968. The use of fixed effects modelling with panel data allows for arbitrary correlation between the unobserved effect and the observed explanatory variables. Here the unobserved effect in panel data analysis is a random variable which is unobserved and time-constant. This methodology could also be used in large panel datasets, such as the PSID, where either fixed or random effects could be used to control for possible unobserved heterogeneity, which is another name for the unobserved effect.

The analysis in this paper shows that ignoring these cluster fixed effects will produce very different estimates, which are clearly biased. For example, a comparison of our results with a model ignoring cluster fixed effects produces results that are statistically different both in signs and in the levels of significance. Our main results can be summarised as follows. First, we find that being male significantly increases the likelihood of school attendance and, once in school, it also increases the number of years of schooling. Second, we find that maternal education (both primary and secondary) has a relatively greater effect on the schooling of girls. In terms of resource constraints, we find that household wealth increases both the probability of schooling, and the years of schooling in all our models and the magnitude of these effects are similar for male and female children.

The rest of the paper is organised as follows. In Section 2, we introduce the model and estimation strategy, which is followed by Section 3 with a discussion of the data and summary statistics on the variables used in our analysis. In Section 4 we present our main empirical results, and finally our conclusions follow in Section 5.

2. MODEL AND ESTIMATION STRATEGY

The econometric approach used in this paper is based on the collective household framework due to McElroy and Horney (1981) and Chiappori (1988). We assume that parents have differing preferences regarding the consumption and health of their children (see Hoddinott and Haddad (1995) and Alderman, Chiappori, Haddad, Hoddinott, and Kanbur (1995) for comparisons of the unitary and collective household models). Assume that the household consists of two parents, mother (*m*) and father (*f*), and children of both genders, daughters (*d*) and sons (*s*) respectively. Parents care about their children's education and therefore invest in the schooling of their children to the extent that the marginal benefits of the schooling investment exceed or equal investment costs.

Parents derive utility from both market and non-market goods, and the utility of the mother and father are denoted by U^m and U^f respectively, so that, parental preferences can be represented by a utility function:

$$U^i = u(c, h) \quad i = m, f \tag{1}$$

where c denotes parental consumption and h the schooling investments in children (both male and female).

Further assume that the father's and mother's reservation utility levels or threat point are given by $\tilde{U}_m$ and $\tilde{U}_f$, respectively. The threat point refers to the outside option of each parent, which could refer to their re-marriage options, support networks and resources that they can take away should the marriage break down. An improvement in autonomy could also improve one's reservation utility. It is reasonable to assume that mothers with better earning abilities, higher education or decision making power have better outside options should their union fail. Hence, they presumably have greater power to allocate resources towards child schooling. Assume that the two parents (m and f) choose c and h to maximise

$$V = [U^m(c,h) - \tilde{U}^m(p, A_m; \gamma_m)] \times [U^f(c,h) - \tilde{U}^f(p, A_f; \gamma_f)] \tag{2}$$

where p is the price inputs. Children's schooling is financed through parental income, both from wage income (w) and from non-income household assets (A). Hence the full income constraint is given by:

$$w_m T_{wm} + w_f T_{wf} + A_m + A_f = p_c c + p_h h \tag{3}$$

where T refers to parental time spent in wage labour and p_c and p_h are the prices of consumption and schooling inputs respectively.

The schooling investments of a child, h_i, are determined by a reduced form demand function:

$$h_i = f(I, H, Z, \varepsilon_i) \qquad (j = d, s) \tag{4}$$

where I is a vector of parental characteristics (such as demographics, education levels, labour market status), H a vector of household characteristics (such as wealth, household size, number of pre-school age siblings), Z a vector of individual child characteristics (such as age, age-squared, gender), and ε represents unobservable individual, household and community characteristics that affect the child's schooling. We do not observe schooling preferences. However, we do observe the actual number of years of schooling that the child has currently attained. Therefore, assuming that these choices reflect the household's preference for children's schooling, we can use the number of years of schooling as our dependent variable.

There are however a number of limitations with this measure, as the data has a large number of zeros, indicating that there are a large number of children that have no schooling. In particular, 39.8% of all children have no schooling, of which 63.7% are female children. The variable, years of schooling, however, is a continuous variable but its range is constrained.

This implies that optimizing behaviour in this model leads to a corner solution for some households where it is optimal to choose zero years of schooling. If a least squares (LS) model is used with a large number of zeros in the dependent variable, the predictions may well be negative. This problem does not manifest itself if the Tobit model is used instead.

Another consideration to take into account is the sequential nature of schooling decisions. In the first stage, parents make a decision on whether or not the child should go to school, and if the child is currently in school, the number of years of schooling that they should get. The question then arises whether there is a likelihood of sample selection bias in estimating schooling demand, if children with no schooling are ignored. In particular, if we only use data from those children who are in school, then the resulting LS estimates are potentially inconsistent due to an omitted variable bias. Hence, we first estimate a two-stage sample selection model where, in stage 1, we estimate a probit equation examining the probability of going to school. This is a binary variable taking on a value of 1 if the child is currently attending school and 0 otherwise. In the next stage, for those children that are currently in school, we estimate the years of schooling using a maximum likelihood estimator. For completeness we also estimate LS and Tobit models with years of schooling as our dependent variable, and a sample selection model.

Cluster Fixed Effects

Like most surveys in developing countries, the *Nepal Demographic Household Survey* (*NDHS 2001)* is a two-stage, stratified sample of households. The stratification process breaks down a single survey into multiple independent surveys, one for each stratum. The survey data for the NDHS 2001 was collected in two stages, first sampling clusters, or primary sampling units (PSUs), and then selecting households from within each cluster. In other words, households are not randomly distributed over space but are geographically grouped. Two hundred and fifty seven PSUs were randomly selected to be included in the dataset using a systematic sampling process, with the probability of being in the sample proportional to the size of the population. At the second stage of the sampling process, on average, systematic samples of 34 households per PSU were selected in all the regions. Each cluster represents a PSU. In our dataset, there are 247 clusters in the full sample and 246 and 245 in the male and female samples, respectively.

In using cross-section data therefore, we have an additional econometric problem in that households within the cluster may have similar characteristics, and ignoring these cluster fixed effects is likely to give us biased estimates. This is because clusters are typically villages in rural samples, so that households within a cluster not only live near one another but are also interviewed around the same time.

Deaton (2000) points out that, if the cluster design of the data is not specified in the model (as is done with the fixed effects models), standard formulas for variances of estimated means are too small. Hence, the intra-cluster co-relationships lead to heteroscedasticity, which will bias the estimated standard errors.[3] Additionally, there is an efficiency issue in that the error terms in the regressions are correlated across observations if the cluster design of the data is ignored. This implies that the OLS regression is not efficient even within the class of linear estimators.

Moreover, as we do not have information on community characteristics such as the proximity and availability of schools, quality of schools, schooling costs and job opportunities, ignoring community/cluster fixed effects is likely to give us biased estimates. These supply side variables that influence schooling decisions also have the disadvantage of being potentially endogenous. This could happen because individuals might choose their residence based on the availability of public schooling services (see Rosenzweig and Wolpin, 1988). Second, local infrastructure itself might be placed selectively by public policy, perhaps in response to local schooling conditions (see Rosenzweig and Wolpin, 1986). While selective migration in response to local infrastructure variables is unlikely to be particularly common in a developing country such as Nepal, selective placement of education services is potentially an important issue. The inclusion of cluster fixed effects has the advantage of controlling for the possibility that people living in the cluster may share the same schooling characteristics and be able to access common information.

The full model for the cluster fixed effects selection model is as follows where (5) and (6) refer to the probit model and (7) refers to the selection model.

$$y_{it}^* = \alpha_i + \delta' w_{it} + u_{it} \tag{5}$$

$$y_{it} = \begin{cases} 1 & \textit{if} \;\; y_{it}^* > 0 \text{ (child attends school)} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$\begin{gathered} h_{it} = \theta_i + \beta' x_{it} + \varepsilon_{it} \quad [u_{it}, \varepsilon_{it}] \sim BVN[0, 0, \sigma, \rho] \\ y_{it},\ h_{it} \text{ only observed when } y_{it} = 1 \end{gathered} \tag{7}$$

The probit estimation assumes that there is a latent variable y_{it}^* which can be written as a linear function of variables that affect the probability of a child attending school (see (5)) where w_{it} is a matrix of explanatory variables, δ' is the vector of coefficients that will be estimated and u_{it} is a random error term. The latent variable is unobservable and instead we observe the dummy variable $y_{it} = 1$ if a child is currently attending school, and zero otherwise. We assume that the probability of a child's schooling attendance is contingent on a range of child characteristics (c), maternal characteristics (m) and household characteristics (H). In the selection equation, the variable h_{it} only pertains to children currently in school. We assume that this variable is a linear function of variables, x_{it}, that affect the number of years of schooling that a child has.

Since neither the LS nor the Tobit model take into account the possibility that the standard errors maybe incorrect due to the cluster effects, we use a cluster fixed effects model to correct the standard errors for clustering effects. This technique implies that each cluster has a different intercept. However, although we include cluster fixed effects estimates to reduce the potential for omitted variable bias, in Table 5 we also present results which do not control for the cluster fixed effects, since it is possible that the cluster fixed effect will absorb some of the variation in child schooling.

In addition to the sample selection model, we use a LS model to analyse years of schooling. The LS model (without cluster fixed effects) is:

$$h_i = \alpha + \gamma' \mathbf{x}_i + \varepsilon_i \tag{8}$$

and the LS model using cluster fixed effects is:

$$h_{it} = \alpha_i + \beta' \mathbf{x}_{it} + \varepsilon_{it} \tag{9}$$

where h_i (h_{it}) is the dependent variable, years of schooling, and $\mathbf{x}_i$ ($\mathbf{x}_{it}$) is a matrix of parental, household and individual child characteristics. Here, α_i is the separate constant term for each cluster and ε_{it} is the cluster-varying error because it represents the unobserved factors that change over the clusters and affect h_{it}.

Although (8) and (9) may in fact be good approximations of schooling choices, it is likely that some of the predictions for the years of schooling will be negative. To account for this possibility of predicted negative years of schooling, a Tobit model is also estimated, since the model will have non-negative predicted values for years of schooling and reasonable partial effects over the explanatory variables. In the Tobit model without cluster fixed effects, there is an underlying latent variable, h_i^* which is the desired level of schooling. However, since we only observe the actual years of schooling for each child, the Tobit model is therefore given by the following set of equations.

$$\begin{aligned} h_i^* &= \beta' \mathbf{x}_i + \varepsilon_i \\ h_i &= \max\,\{h_i^*, 0\} \end{aligned} \tag{10}$$

The Tobit model using cluster fixed effects is characterised as follows:

$$f(h_{it}) = f(\beta' \mathbf{x}_{it} + \alpha_i) + \varepsilon_{it} \tag{11}$$

where h_i (h_{it}) and $\mathbf{x}_i$ ($\mathbf{x}_{it}$) are as defined in (8). The latent variable, h_i^*, satisfies the classical linear model assumptions. Note that the observed variable, h_i, equals h_i^* when $h_i^* \geq 0$ and equals 0 when $h_i^* < 0$. All we observe is the actual number of years of schooling, h_i, not the desired years of schooling. The cluster fixed effects method allows the relaxation of the restriction in the Tobit model which requires the same variables to affect the probability of a non-zero observation, as well as a zero observation, and with the same sign. This is a restrictive structure that the Tobit model imposes. The sample selection model on the other hand allows for a correlation coefficient between the disturbances of the two equations. If the disturbances are uncorrelated (that is, if the estimated ρ is not significant), then the y_{it} equation could be estimated by the Least Squares method.

Wooldridge (2003) suggests informally evaluating the Tobit model by using a probit model with $y = 1$ if $h > 0$ and 0 otherwise. He notes that the relationship between $\hat{\gamma}_i$, the estimated coefficients of the probit model, should be "close to" $\hat{\beta}_i/\hat{\sigma}$ where $\hat{\beta}_i$ and $\hat{\sigma}$ are the estimates from the Tobit model. In our models, the estimated probit coefficients suggest that the Tobit model is an appropriate model as $\hat{\gamma}_i$ are very similar to $\hat{\beta}_i/\hat{\sigma}$ with the same variables being significant in both models and with the same level of significance.

In the section below we describe the data and the variables used in the estimation process.

3. DATA AND SUMMARY STATISTICS

The data for our analysis come from the 2001 *Nepal Demographic and Health Survey* (NDHS). The survey was conducted by Macro International Inc. with funding from USAID, and was administered to ever-married females aged 18–49 years. It contains detailed information on household structure, labour market participation, asset ownership, health and educational characteristics for all the household members.

Descriptive statistics of the variables used in the empirical analysis are presented in Tables 1 and 2, disaggregated by gender and schooling outcomes, respectively. Our analysis is based on data for 7464 children in the 6–17 age group for whom complete information is available on schooling and household characteristics. Thus, we exclude those households where there are no children in the school-going age, or where data are missing. Since the descriptive statistics are roughly similar across male and female children, we restrict our discussion to the combined sample (all children).

According to Table 1, a little over half of the children in the sample are girls. In general we observe low levels of schooling among mothers in our sample, with approximately 80% of the mothers in the sample having no schooling, and only around 13% of the mothers with a primary education. Despite this, maternal employment levels are high in this sample, with 88% being employed. However, a large proportion of these mothers (65%) are in unpaid employment, which is likely to be of a self-employed nature or to be working for the family. The education levels for fathers, although low, are substantially higher than that for mothers in the sample. For example, from Table 1, we observe that 42% of the children have fathers with no schooling, with over a quarter having primary education and over 33% being secondary or higher graduates.

In Table 2 we disaggregate the sample by schooling outcomes and compare the sample of children that have some schooling with those children that have no schooling. Here we observe that children with no schooling also have a greater proportion of mothers with no schooling (nearly 88%) relative to school-going children (73%). Moreover, children attending school are also three times more likely to have mothers with at least a secondary education.

Explanatory Variables

Since the primary focus of this paper is to examine for the role of maternal education and labour market status on schooling outcomes and to see if there are gender differences if any, we include a range of child, sibling, parental and household characteristics (including wealth) among our explanatory variables.

Child Characteristics

The child characteristics considered here are the child's gender, age and age squared. The child's gender is a dichotomous variable that takes on a value

Table 1. Descriptive Statistics.

Variable	Mean (Standard Error)		
	All ($n = 7464$)	Male ($n = 3545$)	Female ($n = 3919$)
Child characteristics			
Education (in years, corrected for age 6 and in school = 1)	1.912	2.312	1.551
	(2.452)	(2.585)	(2.266)
Age (in years)	10.844	10.798	10.886
	(3.688)	(3.626)	(3.743)
Male (= 1 if male, 0 otherwise)	0.475		
Household characteristics			
Number of siblings under 5	1.692	1.658	1.722
	(0.985)	(0.981)	(0.987)
Proportion of daughters	0.434	0.360	0.502
	(0.302)	(0.290)	(0.298)
Father's age	34.806	35.019	34.613
	(9.142)	(9.164)	(9.119)
Mother's age	30.255	30.449	30.080
	(7.392)	(7.394)	(7.386)
Household size	8.320	8.330	8.310
	(3.609)	(3.604)	(3.613)
Religion – Hindu	0.821	0.816	0.825
Mother owns land (dummy variable)	0.079	0.076	0.081
Mother owns livestock and can sell without permission	0.101	0.103	0.100
Wealth 1 (= 1 if household belongs to 1st wealth quintile)	0.252	0.250	0.253
Wealth 2 (= 1 if household is in 2nd wealth quintile)	0.202	0.206	0.199
Wealth 3 (= 1 if household is in 3rd wealth quintile)	0.208	0.217	0.199
Wealth 4 (= 1 if household is in 4th wealth)	0.199	0.195	0.202
Wealth 5 (= 1 if household is in 5th wealth quintile)	0.139	0.132	0.146
Father's education – no schooling	0.418	0.427	0.411
Father's educational attainment – primary	0.251	0.254	0.248
Father's educational attainment – secondary school or higher	0.331	0.319	0.341
Father's occupation – clerical, sales or professional	0.214	0.209	0.219
Father's occupation – agriculture	0.554	0.228	0.235
Father's occupation – manual worker	0.232	0.563	0.546
Father's all-year employment status = 1 if employed all year	0.712	0.711	0.712

Table 1. (*Continued*)

Variable	Mean (Standard Error)		
	All ($n = 7464$)	Male ($n = 3545$)	Female ($n = 3919$)
Mother's education – no schooling	0.795	0.806	0.784
Mother's educational attainment- primary	0.125	0.123	0.128
Mother's educational attainment-secondary	0.080	0.071	0.088
Mother is in unpaid, informal employment	0.650	0.657	0.644
Mother is in paid employment	0.230	0.223	0.236

of 1 for males and 0 for females. To take into account the possibility that older children are more likely to be working, and to also capture possible non-linearities, we include both the child's age and age squared as explanatory variables.

Maternal Characteristics

Control over economic assets is an important means of gaining autonomy for women in traditional societies. Maternal economic status is enhanced by the mother being better educated, participating in the labour market, and if she has direct ownership of income generating assets such as land and livestock. However, labour market participation has the potential of being endogenous and further, the link between maternal labour market participation and schooling outcomes is ambiguous. For example, while an increase in labour market participation increases household income, so more schooling inputs can be purchased, it also restricts the amount of time that the mother can spend with the child. Furthermore, having a working mother may also mean a greater demand for child labour, leading to a substitution away from schooling. Moreover, the researcher is unaware as to whether the greater labour force participation was motivated by greater female autonomy or due to poverty. Therefore, since a majority of the mothers in our sample is employed, we consider whether or not the mother works in paid employment. Three dummy variables are constructed to capture the impact of maternal education on children's schooling: no education (the base category), primary school, and secondary school and above.

Table 2. Descriptive Statistics Disaggregated by Schooling Outcomes.

Variable	Mean (SE)	
	No Schooling ($n = 2974$)	Some Schooling ($n = 4490$)
Child characteristics		
Education (in years, corrected for age 6 and in school = 1)	N/A	3.179 (2.443)
Age (in years)	11.671 (4.180)	10.296 (3.207)
Male (= 1 if male, 0 otherwise)	0.363	0.549
Household characteristics		
Number of siblings under 5	1.720 (1.020)	1.674 (0.960)
Proportion of daughters	0.428 (0.308)	0.438 (0.299)
Father's age	34.752 (9.657)	34.842 (8.785)
Mother's age	30.241 (8.028)	30.265 (6.939)
Household size	8.282 (3.468)	8.344 (3.699)
Religion – Hindu	0.785	0.844
Mother owns land (dummy variable)	0.057	0.093
Mother owns livestock and can sell without permission	0.078	0.116
Wealth 1 (= 1 if household is in 1st wealth quintile)	0.285	0.229
Wealth 2 (= 1 if household is in 2nd wealth quintile)	0.226	0.187
Wealth 3 (= 1 if household is in 3rd wealth quintile)	0.236	0.189
Wealth 4 (= 1 if household is in 4th wealth quintile)	0.169	0.218
Wealth 5 (= 1 if household is in 5th wealth quintile)	0.084	0.176
Father's education – no schooling	0.535	0.307
Father's educational attainment – primary	0.225	0.268
Father's educational attainment – secondary school or higher	0.214	0.407
Father's occupation – clerical, sales or professional	0.137	0.245
Father's occupation – agriculture	0.582	0.535
Father's occupation – manual worker	0.250	0.219
Father's employment status = 1 if employed all year	0.793	0.858

Table 2. *(Continued)*

Variable	Mean (SE)	
	No Schooling ($n = 2974$)	Some Schooling ($n = 4490$)
Mother's education – no schooling	0.877	0.732
Mother's educational attainment – primary	0.085	0.152
Mother's educational attainment – secondary	0.036	0.109
Mother is in unpaid, informal employment	0.601	0.683
Mother is in paid employment	0.264	0.207

Household Characteristics

We control for household financial resource constraints by including variables relating to household wealth, father's occupation and household size. While low incomes impose resource constraints and increase direct schooling costs, they also make the monetary benefits from child labour more attractive. Hence, in areas where there are possibilities for child labour, the opportunity cost of schooling increases for poor households. Per capita household expenditure is typically used as a proxy for per capita household income since it is relatively easy to obtain and it is also measured with less error. A shortcoming of this dataset, however, is that it contains no information on wages, household expenditure patterns and community characteristics. There is however a household wealth index that divides households into 5 wealth quintiles (Wealth 1 – Wealth 5), with Wealth 1 representing the poorest quintile. Since the wealth index is calculated using household asset ownership, it has the advantage of providing a reasonably reliable measure of the household's economic status, and furthermore it is not affected by the endogeneity and transitory nature of labour income. Using this index, we observe that poorer households are disproportionately represented in our sample, with only 13% of households belonging to the highest wealth quintile whereas the lowest wealth index has a much larger proportion (25%).[4] To examine if increases in wealth affect schooling investments differentially for male and female children, we additionally interact each of the wealth dummies with the dummy variable for child's gender in the combined sample.

Father's education and employment status are also included among our explanatory variables. Three dummy variables are constructed for father's education: no education (the base category), primary school, and secondary school and above. Restricting our analysis to only those children whose fathers are employed (since very few fathers in the sample are unemployed), we categorise father's employment status into manual labourer, office worker and agricultural employment (our base category). Given the rural nature of our sample, we also include the possibility of seasonal employment by constructing a dummy variable if the father is employed all year.

We control for household demographic characteristics by including household size, the number of children, the presence of pre-school age siblings and the proportion of female children in the household. All have been found to be influential in affecting schooling outcomes. The 'quantity-quality trade-off theory' of Becker and Lewis (1973) posits an inverse relationship between the numbers of children and schooling investments. However, this relationship is not so clear cut in the developing country literature where there are possibilities for combining work and school, selective schooling for some children and economies of scale in schooling costs.

The presence of younger siblings in the household is typically found to reduce school attendance, both due to resource constraints and also because it opens up possibilities for combining work with child-minding, particularly in the rural informal sector (see Patrinos and Psacharopoulos, 1997; Lloyd and Gage-Brandon, 1994). However, the effect of this variable is unclear and may be different for male and female children and may also depend on the gender of the younger sibling(s). For example, parents may prefer to educate their male children, in which case resources may be directed to male children and an older female child may not be sent to school in anticipation of future schooling resource constraints. Both Lloyd and Gage-Brandon (1994) and Patrinos and Psacharopoulos (1997) find that, particularly for girls, schooling outcomes in sub-Saharan Africa and Peru are more likely to be adversely affected by the presence of younger siblings. For example, older siblings may act as carers for younger pre-school age siblings.

Finally, we include the proportion of school-age female children in the household among our explanatory variables. This variable is included because there is strong evidence in the literature that, in traditional societies, female children are particularly disadvantaged in the intrahousehold allocation of resources (see Rosenzweig and Schultz, 1982; Behrman, 1988; Harriss, 1990; Haddad *et al.,* 1994; Strauss and Thomas, 1995). If indeed there is a gender bias operating at the household level, then, for a given family size, it must be the case that a male child, growing up in a household with brothers only, may have fewer resources than if he were to grow up with sisters only. This is likely to be true for females as well. This would imply that the educational attainment of children depends not only on their own gender but also differs depending on whether their siblings are male or female. Hence, siblings become rivals in a competition for greater access to household resources. This view is supported in studies by Parish and Willis (1993), Garg and Morduch (1998) and Morduch (2000) where education and health outcomes are better for children growing up with more sisters rather than more brothers.

4. RESULTS

Initially, the entire sample was analyzed to examine if there were any sample selection issues. In the first stage we estimated a Cluster fixed effects Probit model where the dependent variable = 1 if the child is currently in school and 0 otherwise. Then, for those children that are currently in school, we use a maximum likelihood (MLE) cluster fixed effects model to estimate the number of years of schooling. The results for the sample selection model are presented in Tables 3 and 4, with Table 3 presenting the first-stage probit cluster fixed effects estimates and Table 4 presenting the ML estimates for children in school. We report the coefficients and standard errors (in parentheses). For comparison, we present the results for a sample selection model using Heckman's Two Step estimator (without cluster fixed effects) in Table 5. Finally, in Table 6 we present the Least Squares cluster fixed effects results (columns 2, 3 and 4) and the Tobit cluster fixed effects results (columns 5, 6 and 7) for the combined sample, male and female samples, respectively.

To take into account the gender-specific impact of some of the explanatory variables, we estimate the combined sample and then the male and female samples separately. A likelihood ratio test confirms that there is a significant difference between male and female children.

Our results show that the male child has a higher probability of going to school and once in school also has more years of schooling. Second, having better educated parents increases the probability of attending schooling for both boys and girls, with maternal education having a relatively greater effect on the probability of girls being in school. Third, having a mother in unpaid, informal employment reduces the probability of a male child going to school. In terms of resource constraints, we find that household wealth increases both the probability of schooling and the years of schooling in all our models, and the magnitude of these effects are approximately the same for male and female children. Finally, a comparison of the model using cluster fixed effects with the model that does not control for cluster fixed effects shows that there are important differences in the signs and in the levels of significance between these models.

Table 3. First-Stage Estimates – Cluster Fixed Effects Probit Model.

Variables	Coefficient (SE)		
	Full Sample	Males	Females
Child characteristics			
Age	0.604***	0.761***	0.549***
	(0.034)	(0.052)	(0.050)
Age square	−0.030***	−0.035***	−0.030***
	(0.001)	(0.002)	(0.002)
Male	0.882***		
	(0.069)		
Household characteristics			
Religion	0.164**	0.334***	−0.001
	(0.076)	(0.118)	(0.110)
Number of siblings < 5 years	−0.020	0.025	−0.063*
	(0.022)	(0.034)	(0.033)
Household size	0.010	−0.009	0.026***
	(0.007)	(0.011)	(0.010)
Wealth 2	0.262***	0.113	0.254***
	(0.072)	(0.086)	(0.080)
Wealth 3	0.363***	0.271***	0.323***
	(0.080)	(0.096)	(0.093)
Wealth 4	0.594***	0.463***	0.502***
	(0.081)	(0.103)	(0.094)
Wealth 5	0.950***	0.792***	0.726***
	(0.109)	(0.148)	(0.135)
Wealth 2*male	−0.202**		
	(0.103)		
Wealth3*male	−0.177*		
	(0.103)		
Wealth 4*male	−0.302***		
	(0.106)		
Wealth 5*male	−0.572***		
	(0.124)		
Mother owns land	0.211***	0.194	0.322***
	(0.076)	(0.118)	(0.109)
Mother owns livestock and can sell without permission	0.038	0.117	0.002
	(0.065)	(0.104)	(0.091)
Mother's age	0.001	0.002	0.000
	(0.004)	(0.007)	(0.006)
Mother's educational attainment – primary	0.220***	0.111	0.326***
	(0.064)	(0.102)	(0.089)
Mother's educational attainment – secondary	0.439***	0.330**	0.584***
	(0.090)	(0.150)	(0.124)
Father's age	0.003	−0.001	0.005
	(0.004)	(0.006)	(0.005)

Table 3. (*Continued*)

Variables	Coefficient (SE)		
	Full Sample	Males	Females
Father's educational attainment – primary	0.290***	0.264***	0.312***
	(0.047)	(0.072)	(0.069)
Father's educational attainment – secondary school or higher	0.543***	0.556***	0.629***
	(0.052)	(0.082)	(0.076)
Mother is in unpaid, informal employment	−0.181**	−0.028***	−0.040
	(0.083)	(0.124)	(0.125)
Mother is in paid employment	0.007	−0.093	0.177
	(0.084)	(0.126)	(0.127)
Father's employment status = 1 if employed all year	0.066	−0.118	−0.019
	(0.060)	(0.093)	(0.087)
N	7464	3545	3919

Standard errors are in parentheses.
*Significant at 10% level.
**Significant at 5% level.
***Significant at 1% level.

Below we discuss these results in some detail focusing first on the sample selection model estimates, a comparison of sample selection models with and without controlling for fixed effects, and then on the Least Squares and Tobit results.

Sample Selection Results

As indicated earlier, the sample selection estimates allow us to examine the probability of a child going to school and, contingent on this, analyze factors influencing schooling levels for those children that are currently in school. One important result is the significance or otherwise of the estimated coefficient for *rho*. This estimated coefficient is significant for the full sample and for females, but not for males. This indicates that, for female children, we should not ignore those that have no schooling, and a failure to take their characteristics into account will result in biased estimates. According to Table 3, our first-stage probit cluster fixed effects results indicate that, relative to a female child, a male child has a greater probability of attending school. Furthermore, relative to the base category of no schooling, those children in the combined and female samples with primary and secondary level educated mothers are more likely to attend school. Interestingly, mother's primary education has no effect on the probability of a male child

Table 4. Second-Stage Estimates – Selection Model Cluster Fixed Effects Least Squares.

	Coefficient (Standard Error)		
	Full sample	Males	Females
Child characteristics			
Age	−0.001	0.037	−0.022
	(0.159)	(0.156)	(0.161)
Age squared	0.029	0.028	0.029
	(0.034)	(0.050)	(0.050)
Male	0.222***		
	(0.060)		
Household characteristics			
Religion	−0.044***	0.065***	−0.155***
	(0.006)	(0.022)	(0.028)
Number of siblings < 5	−0.051***	−0.045	−0.062
	(0.000)	(0.050)	(0.050)
Household size	0.017	0.020***	0.010***
	(0.016)	(0.000)	(0.001)
Proportion of daughters	0.180**	0.217***	0.148***
	(0.073)	(0.027)	(0.034)
Wealth 2	0.129*	0.128	0.098
	(0.069)	(0.079)	(0.075)
Wealth 3	0.157**	0.104	0.129
	(0.078)	(0.102)	(0.081)
Wealth 4	0.355***	0.475***	0.361***
	(0.089)	(0.100)	(0.096)
Wealth 5	0.581***	0.553***	0.548***
	(0.123)	(0.141)	(0.148)
Wealth 2*male	−0.027		
	(0.099)		
Wealth 3*male	−0.064		
	(0.110)		
Wealth 4*male	0.082		
	(0.087)		
Wealth 5*male	−0.115		
	(0.109)		
Mother owns land	0.217***	0.214**	0.252**
	(0.066)	(0.089)	(0.101)
Mother's age	0.003	0.004	−0.002
	(2.121)	(0.007)	(0.008)
Mother's education – primary	0.031	0.029	−0.006
	(1.448)	(0.086)	(0.085)
Mother's education – secondary	0.146*	0.002	0.188*
	(0.087)	(0.117)	(0.110)

Table 4. (*Continued*)

	Coefficient (Standard Error)		
	Full sample	Males	Females
Mother is in unpaid, informal employment	−0.055	0.062	−0.193
	(0.106)	(0.127)	(0.150)
Mother is in paid employment	−0.053	0.088	−0.259
	(0.115)	(0.127)	(0.196)
Father's age	0.003	0.004	0.003
	(0.004)	(0.654)	(0.713)
Father's education – primary	0.081	0.150	0.028
	(0.054)	(1.096)	(1.159)
Father's education – secondary	0.366***	0.431***	0.308***
	(0.057)	(0.075)	(0.096)
Father's occupation – clerical, sales or professional	0.088***	0.113	0.074
	(0.000)	(0.078)	(0.081)
Father's occupation – manual	−0.048***	−0.125*	0.062
	(0.002)	(0.074)	(0.079)
Father's employment status = 1 if employed all year	−0.003	−0.062	0.102
	(0.003)	(0.089)	(0.116)
σ	1.080***	1.060***	1.037***
	(0.019)	(0.021)	(0.028)
ρ	−0.101**	−0.015	−0.167*
	(0.048)	(0.074)	(0.094)
N	7464	3545	3919

Standard errors are in parentheses.
*Significant at 10% level.
**Significant at 5% level.
***Significant at 1% level.

attending school. However, while having a secondary educated mother significantly increases the probability of schooling across all three samples, the size of the effect is much larger for females. Similar significant results hold for the effect of father's education, with both primary and secondary education increasing the probability of schooling across all three samples. Again the size of the effect is largest for female children.

Table 3 also indicates that an increase in household wealth significantly increases the probability of a child being in school, across all three samples. Not surprisingly, being male significantly increases the probability of attending school and once in school, it also increases the number of years of schooling. However, the greatest increase in the probability of attending school comes from belonging to the wealthiest quintile, as opposed to the lowest wealth quintile.

Table 5. Estimates for the Sample Selection Model for the Full Sample.

	Coefficient (SE)	
	Probit Estimates	OLS Estimates
Constant	−3.703***	2.802***
	(0.199)	(0.830)
Child characteristics		
Age	0.529***	−0.401***
	(0.031)	(0.092)
Age squared	−0.026***	0.049***
	(0.001)	(0.004)
Male	0.752***	−0.382***
	(0.063)	(0.145)
Household characteristics		
Religion	0.187***	−0.048
	(0.042)	(0.064)
Number of siblings < 5	−0.041**	−0.038
	(0.020)	(0.027)
Household size	0.017***	−0.003
	(0.006)	(0.007)
Proportion of daughters		0.164***
		(0.062)
Wealth 2	0.136**	0.011
	(0.064)	(0.097)
Wealth 3	0.177***	0.031
	(0.065)	(0.099)
Wealth 4	0.441***	−0.067
	(0.066)	(0.117)
Wealth 5	0.629***	0.067
	(0.081)	(0.144)
Wealth 2*male	−0.213**	0.164
	(0.094)	(0.132)
Wealth 3*male	−0.199**	0.087
	(0.093)	(0.131)
Wealth 4*male	−0.255***	0.325**
	(0.097)	(0.135)
Wealth 5*male	−0.517***	0.317*
	(0.113)	(0.162)
Mother owns land	0.282***	0.002
	(0.064)	(0.085)
Mother owns livestock and can sell without permission	0.187***	
	(0.056)	
Mother's age	0.000	0.001
	(0.004)	(0.005)
Mother's education – primary	0.340***	−0.103
	(0.055)	(0.078)

Table 5. (*Continued*)

	Coefficient (SE)	
	Probit Estimates	OLS Estimates
Mother's education – secondary	0.549***	−0.038
	(0.078)	(0.106)
Father's age	0.006**	0.003
	(0.003)	(0.004)
Father's education – primary	0.431***	−0.232***
	(0.041)	(0.088)
Father's education – secondary	0.694***	−0.068
	(0.045)	(0.116)
Mother is in unpaid, informal employment	0.209***	−0.115
	(0.066)	(0.094)
Mother is in paid employment	−0.022	−0.004
	(0.066)	(0.091)
Father's occupation – clerical, sales or professional		0.155***
		(0.048)
Father's occupation – manual		0.031
		(0.044)
Father's employment status = 1 if employed all year	0.273***	−0.293***
	(0.044)	(0.073)
λ		−1.507***
		(0.282)
N	7464	7464

Standard errors are in parentheses.
*Significant at 10% level.
**Significant at 5% level.
***Significant at 1% level.

To get a better perspective of the influence of gender in the combined sample, we interact the wealth variable with the male dummy variable. The estimated coefficients are higher in the full sample than those in the male sample. This is particularly noticeable for the wealthiest quintile where the coefficient in the full sample is 1.26 (Male + Wealth 5 + Wealth 5 * Male = 0.882 + 0.950–0.572) compared with 0.792 in the male sample. The full sample estimates indicate that the effect for males is always greater than the effect for females because the coefficient for being male is positive and much larger than the coefficients for the interaction variables, which are always negative.

In terms of maternal education and labour market variables, we observe that maternal land ownership has a positive and significant effect on the probability of a female child attending school, with no effect on the schooling

Table 6. Results from LS and Tobit Fixed Effects Models for the Full Sample, Males and Females.

	Coefficient (Standard Error)					
	Least Squares			Tobit		
	All	Males	Females	All	Males	Females
Child characteristics						
Age	1.089***	1.204***	0.984***	1.695***	1.726***	1.693***
	(0.045)	(0.069)	(0.058)	(0.071)	(0.093)	(0.105)
Age square	−0.038***	−0.038***	−0.037***	−0.065***	−0.060***	−0.071***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)	(0.005)
Male	1.080***			2.078***		
	(0.094)			(0.148)		
Household characteristics						
Religion	0.077	0.332**	−0.149	0.214	0.600***	−0.192
	(0.101)	(0.158)	(0.130)	(0.156)	(0.211)	(0.226)
Number of siblings < 5	−0.071**	−0.008	−0.131***	−0.097**	0.010	−0.226***
	(0.030)	(0.046)	(0.039)	(0.047)	(0.062)	(0.069)
Household size	0.036***	0.008	0.060***	0.046***	−0.001	0.097***
	(0.009)	(0.014)	(0.012)	(0.014)	(0.019)	(0.021)
Proportion of daughters	0.280***	0.160	0.182*	0.479***	0.300*	0.370*
	(0.083)	(0.131)	(0.108)	(0.129)	(0.173)	(0.190)
Wealth 2	0.322***	0.056	0.258***	0.607***	0.093	0.544***
	(0.100)	(0.115)	(0.097)	(0.163)	(0.153)	(0.175)
Wealth 3	0.429***	0.246*	0.346***	0.805***	0.414**	0.737***
	(0.108)	(0.128)	(0.111)	(0.176)	(0.172)	(0.200)
Wealth 4	0.800***	0.754***	0.675***	1.492***	1.037***	1.271***
	(0.109)	(0.134)	(0.113)	(0.173)	(0.178)	(0.198)
Wealth 5	1.304***	1.036***	1.025***	2.373***	1.504***	1.887***
	(0.144)	(0.190)	(0.162)	(0.221)	(0.249)	(0.278)
Wealth 2*male	−0.264*			−0.497**		
	(0.140)			(0.219)		
Wealth 3*male	−0.204			−0.400*		
	(0.139)			(0.219)		
Wealth 4*male	−0.170			−0.692***		
	(0.141)			(0.216)		
Wealth 5*male	−0.621***			−1.456***		
	(0.158)			(0.236)		
Mother owns land	0.204**	0.136	0.309**	0.359**	0.274	0.570***
	(0.097)	(0.150)	(0.127)	(0.146)	(0.195)	(0.213)
Mother's age	0.019***	0.013	0.020**	0.020**	0.015	0.019
	(0.006)	(0.010)	(0.008)	(0.010)	(0.013)	(0.014)
Mother's education – primary	0.191**	0.082	0.304***	0.334***	0.158	0.540***
	(0.082)	(0.127)	(0.106)	(0.121)	(0.164)	(0.175)

Table 6. (*Continued*)

	Coefficient (Standard Error)					
	Least Squares			Tobit		
	All	Males	Females	All	Males	Females
Mother's education – secondary	0.687*** (0.108)	0.429** (0.173)	0.893*** (0.137)	0.863*** (0.158)	0.486** (0.222)	1.169*** (0.221)
Mother is in unpaid, informal employment	0.025 (0.112)	−0.133 (0.171)	0.160 (0.148)	0.009 (0.179)	−0.235 (0.232)	0.352 (0.276)
Mother is in paid employment	−0.102 (0.110)	−0.256 (0.168)	0.069 (0.143)	−0.323* (0.177)	−0.544** (0.231)	0.034 (0.268)
Father's age	−0.001 (0.005)	−0.002 (0.008)	−0.002 (0.006)	0.005 (0.008)	−0.001 (0.010)	0.008 (0.011)
Father's education – primary	0.236*** (0.065)	0.280*** (0.099)	0.193** (0.086)	0.583*** (0.101)	0.518*** (0.131)	0.616*** (0.153)
Father's education – secondary	0.726*** (0.071)	0.838*** (0.109)	0.684*** (0.092)	1.283*** (0.109)	1.212*** (0.144)	1.465*** (0.162)
Father's occupation – clerical, sales or professional	0.094 (0.072)	0.072 (0.112)	0.077 (0.092)	0.153 (0.109)	0.117 (0.147)	0.177 (0.159)
Father's occupation – manual	−0.178*** (0.066)	−0.240** (0.101)	−0.092 (0.086)	−0.294*** (0.103)	−0.352*** (0.135)	−0.150 (0.152)
Father's employment status = 1 if employed all year	0.015 (0.081)	0.142 (0.125)	−0.080 (0.106)	0.110 (0.128)	0.274 (0.170)	−0.063 (0.190)
N	7464	3545	3919	7464	3545	3919

Standard errors are in parentheses.
*Significant at 10% level.
**Significant at 5% level.
***Significant at 1% level.

of a male child. However, for those children currently in school, maternal land ownership significantly increases schooling by 0.2 of a year for male children and by 0.25 for females. Maternal employment status does not appear to have any significant effect on schooling outcomes.

Considering the Maximum Likelihood estimates from the selection equation in the second stage, we observe that wealth has a significant and positive effect, particularly for children from the 4th and 5th quintiles for all three samples. The results indicate that school-going children belonging to households in the two highest wealth quintiles get approximately half a year's extra schooling compared with those from the lowest wealth quintile. While maternal secondary education increases the likelihood of both male and female children attending school, it has no influence on the number of years of schooling that a male child gets. For female children on the other hand, mother's secondary schooling increases the likelihood of having an extra 0.19 years of schooling. Father's secondary education on the other hand increases the years of schooling of boys and girls, where the estimated coefficients show that schooling increases by 0.43 years for a male and 0.30 years for a female child.

In terms of parental occupation, with the exception of manual labour (which is mildly significant for male children), father's occupation in general does not appear to influence the number of years of schooling that a child has. Relative to a male child whose father is employed in agriculture, a male child whose father is a manual labourer has approximately 0.12 years less schooling.

Household demographic control variables, such as household size and the proportion of daughters in the household, also significantly increase the number of years of schooling for both girls and boys. Religion (being Hindu) on the other hand significantly increases the number of years of schooling for boys but has the opposite effect on girls. Consistent with the findings of Garg and Morduch (1998), our results also show that the proportion of daughters in the household significantly increases schooling levels for children in the full sample and for both male and female children. The size of these effects is large, with male children getting an extra 0.22 of a year of schooling and female children an additional 0.15.

Comparison of Results from the Sample Selection Models with and Without Cluster Fixed Effects

Estimates from the original sample selection model without cluster fixed effects are presented in Table 5 for comparison purposes with the cluster fixed effects selection model presented in Tables 3 and 4. There are some significant differences between the two models.

Consider first the probit equation results (Table 3 – with cluster fixed effects and Table 5 (column 2) – without cluster fixed effects). First we observe that variables that are significant in the sample selection model without cluster fixed effects lose their significance when we incorporate cluster fixed effects. These include several household characteristics variables such as: the number of siblings under 5, household size, whether mother owns livestock and can sell without permission, father's age and whether the father is employed all year. By ignoring the cluster effects, the results are biased and, for some variables, the significance is over-stated.

Further, we observe that, for two of the variables (mother owns livestock and can sell without permission, and father is employed all year), there have been large decreases in the estimated coefficients from incorporating cluster fixed effects, implying that their effect on the probability of a child attending school has decreased.

Having a mother in unpaid, informal employment is positive and very significant in a sample selection model without cluster fixed effects. However, when we incorporate cluster fixed effects the variable is still very significant but is now negative, implying that the child has a lower probability of going to school if the mother is in unpaid employment. This is an important result and would not have been observed if cluster fixed effects were not taken into account.

Finally, we observe that the estimated coefficients have increased in magnitude in the cluster fixed effects model for the variables, child's age, males, and wealth quintiles 2–5. These changes imply that the probability of going to school has increased when cluster fixed effects are taken into account. The wealth variables in particular exhibit the largest change. For example, for a female child belonging to the 4th wealth quintile for wealth, the probability of going to school is 0.223 in the cluster fixed effects model but 0.159 in the original sample selection model.

As in the probit equation, the second-stage estimation results (Tables 4 and 5 (column 3)) show that there are a number of changes both in the levels of significance and in the magnitude of the effects. A number of variables (such as religion, number of siblings under 5, wealth quintiles 2–5, parental education variables, mother owning land and the father being a manual worker), that were originally insignificant (without cluster fixed effects) are now significant. In particular, the incorporation of cluster fixed effects implies that a child belonging to the highest wealth quintile has over half a year more schooling than a child belong to the poorest wealth quintile. These effects are substantial and ignoring cluster fixed effects would underestimate the magnitude of these wealth effects. Having a father with secondary education similarly increases the level of schooling of a school-going child by nearly 0.37 of a year. Importantly, household-specific variables such as religion, the number of siblings under 5, and having a father working in manual labour, all reduce child schooling, when cluster fixed effects are taken into account.

The incorporation of cluster fixed effects has also made some variables that were originally significant become more insignificant. Variables such as father's employment all year that were significant in the original model (without fixed effects) are now insignificant. Further the variable, male, was significant and negative in the original model indicating that a male child would have less schooling than a female. This is contrary to expectations in Nepal where previous studies have identified a son preference and aggregate data shows female education levels to be substantially below male. However, the cluster fixed effects results show that being male in fact increases the number of years of schooling and the estimated coefficient is significant at the 1% level. The variable, father's employment all year, is now insignificant in the cluster fixed effects model having been very significant in the original model.

In both models, there is strong evidence that a sample selection model for the full sample is essential. The estimated coefficient, λ, is negative (-1.507) and very significant. In the cluster selection equation, the relevant estimated coefficient ρ is also negative (0.101) and significant at the 5% level. Therefore, using the results from either model would result in biased conclusions if the children who have no schooling were ignored and only children in school were used in an Ordinary Least Squares model. However, the important point to note is that our results indicate that ignoring cluster fixed effects would produce results that are biased and would overestimate the significance of some variables and underestimate others for females but not for males.

Least Squares and Tobit Results

To keep the discussion manageable we focus on the more important results in the combined sample, with a discussion of gender effects next. Note that approximately the same set of variables are significant in the Least Squares and Tobit cluster fixed effects results, with the Tobit results having greater effects on the desired level of schooling. Are there any perceptible gender differences in schooling outcomes? Our results show that a male child has a significant and positive schooling outcome relative to a female child, in both the Least Squares and Tobit models. From Table 6, we further observe that the size of these effects is largest in the Tobit fixed effects model, where, relative to a female child, a male child is likely to have 2.1 years more schooling.

Further we observe that the explanatory variables have a differential effect on the male and female sample. Interestingly, and in keeping with previous studies, we find that an improvement in maternal education has a far greater effect on the education of girls than on boys. For example, in the cluster fixed effects Tobit model, a female child with a mother who has secondary schooling or above has, on average, 1.2 years more schooling relative to a girl whose mother has no education. For boys on the other hand, both in the LS and Tobit fixed effects models, maternal secondary education has a much smaller, but still significant, effect (0.49 years more in Tobit FE) relative to the base category. Studies by Behrman *et al.* (1999) have found that the positive effect of mother's education on child schooling is independent of labour market returns. For example, a primary educated mother may be spending more time in the home schooling of children.

Father's education has a similarly positive and highly significant effect in all our models. Having a father with higher education has a much larger effect than having a primary educated father, possibly capturing the effect of higher household income. For example, in the Tobit model, a male child, whose father has a higher education compared with one whose father has a primary education, will have an extra 0.7 years of schooling on average.

Household wealth emerges as being an important factor in improving schooling outcomes for children in all our models. Relative to the base category (lowest wealth quintile), a child who is born in the each of the higher wealth quintiles has a significantly larger probability of better schooling outcomes in both LS and Tobit models. Interestingly, the size of these effects increases with each wealth category with a child born in the highest wealth quintile having nearly 2.4 years more schooling in the Tobit fixed effects model relative to the base category for the full sample. In the LS model on the other hand, although highly significant, the size of the effect is much smaller (1.3 years).

What is interesting however is that, in the Tobit model, household wealth has a larger positive effect on girls than it does on boys when we estimate the male and female samples separately. For example, in the sample for females, we observe that relative to the base category (poorest quintile), an increase in wealth in each of the other four categories is associated with a positive and significant increase in schooling outcomes. The size of these effects is largest for girls in the 4th and 5th wealth quintiles, where we see improvements in schooling of nearly 1.9 years in the 5th quintile in the Tobit model and 1.0 years in the corresponding least squares fixed effects model. Although the size of these effects is smaller for boys, they are nevertheless of the magnitude of 1.5 years for boys in the highest wealth category relative to the base case in the Tobit cluster fixed effects model.

Interacting the wealth variable with the male dummy variable, we observe that the size of the wealth effect on schooling is enhanced for male children. For example, a male child who is born in the highest wealth category has 3 extra years of schooling compared to a female child also in the highest wealth category in the Tobit fixed effects model.

As discussed above there is strong support for the hypothesis that, relative to a female child, being male is both statistically and economically significant at increasing years of schooling in all our models. When we disaggregate by gender and run the models separately, we still observe similarly large wealth effects on both boys and girls.

Maternal labour force participation is generally high in this sample, with a majority of mothers working in the agricultural informal sector. It is however of interest to know how the different categories of employment affect child schooling outcomes. Hence, maternal employment is divided into two categories – those mothers working for free and those who are in paid employment. The only estimated coefficient that is significant in the Tobit cluster fixed effects model is that for a male child whose mother is in paid employment. The male child is likely to have 0.5 years less schooling compared with a male child whose mother is not working.

The other household characteristics have similar results as in the sample selection model.

5. CONCLUSIONS

In this paper we estimate the effects of maternal education on child schooling outcomes in Nepal. Our estimation strategy uses a unique cluster fixed effects methodology to control for unobserved village-level heterogeneity. Our analysis shows that it is important to control for the cluster fixed effects, given that individuals living in the same village share common community characteristics and there is also likely to be information sharing. The incorporation of cluster fixed effects methodology also helps in addressing the problem of unobserved heterogeneity and omitted variable bias. A comparison of the estimates from the model using cluster fixed effects with estimates from the model that does not control for cluster fixed effects demonstrates some of the reasons for not ignoring these cluster effects. Our results show that there are important differences in the signs and in the levels of significance between these models.

Increasing the schooling levels of children and, in particular girls, is an important development objective in many developing countries where schooling attendance is low and literacy rates among adults are low. This study examines factors affecting schooling demand in Nepal. Our analysis based on Nepalese children aged 6–17 years shows that, in all our three models (sample selection, LS and Tobit), the male child has a higher probability of going to school and once in school also has more years of schooling. Further, parental education and household wealth are crucial in improving child schooling outcomes. In particular, having a secondary educated mother both increases the probability of schooling and once in school, also increases the number of years of schooling that a child gets. It is interesting however, that maternal education has a stronger effect on the schooling of girls. Gender differentials persist with male children more likely to be educated and having a lower likelihood of having schooling disrupted due to the presence of pre-school age siblings. The only significant result for mother's employment is for a male child having a mother in unpaid, informal employment, which has a negative effect of the probability of a male attending school. Given the high levels of maternal labour force participation and the low levels of schooling, these results indicate that there may be some substitution of household chores when the mother increases her labour market participation.

NOTES

1. Studies by Psacharopoulos and Arrigada (1989), Grootaert (1998); Dreze and Kingdon (2001), all show that parental schooling affects the probability of whether or not the child will go to school.
2. See, for example, Strauss (1990); Glewwe (2000); and Pitt and Khandker (1998).
3. Using US data, Moulton (1986, 1990) shows that ignoring cluster effects can lead to a very large underestimation of standard errors.
4. It would have been expected that as the household wealth was in quintiles that there would have been 20% in each group. This was true of the whole sample but households without children and those without school age children were deleted from the analysis. Therefore this has led to an over representation of poorer households and an under representation of wealthy households in this analysis.

REFERENCES

Alderman, H., Chiappori, P. A., Haddad, L., Hoddinott, J., & Kanbur, R. (1995). Unitary versus collective models of the household: Time to shift the burden of proof? *World Bank Research Observer*, *10*(1), 1–20.

Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, *81*(2), s279–s288.

Behrman, J. R. (1988). Intrahousehold allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion? *Oxford Economic Papers*, *45*, 32–54.

Behrman, J., Foster, A., Rosenzweig, M., & Vahsishtha, P. (1999). Women's schooling, home teaching, and economic growth. *Journal of Political Economy*, *107*(4), 682–715.

Chiappori, P. (1988). Nash bargained household decisions. *International Economic Review*, *29*(4), 791–796.

Deaton, A. (2000). *The analysis of household surveys*. Baltimore, Maryland: The Johns Hopkins University Press.

Dreze, J., & Kingdon, G. G. (2001). School participation in rural India. *Review of Development Economics*, *5*(1), 1–24.

Garg, A., & Morduch, J. (1998). Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana. *Journal of Population Economics*, *11*, 471–493.

Glewwe, P. (2000). Why does mother's schooling raise child health in developing countries? Evidence from Morocco. *Journal of Human Resources*, *34*, 124–159.

Grootaert, C. (1998). *Child labor in Cote D'Ivoire: Incidence and determinants"*. Washington, DC: The World Bank Social Development Department, World Bank.

Haddad, L., Hoddinott, J., & Alderman, H. (1994). *Intrahousehold resource allocation: An overview*. Policy Research Working Paper no. 1255, World Bank, Washington, DC.

Handa, S. (1996). The determinants of teenage schooling in Jamaica: Rich vs. poor, females vs. males. *Journal of Development Studies*, *32*, 554–580.

Harriss, B. (1990). The intrafamily distribution of hunger in South Asia. In: J. Dreze & A. K. Sen (Eds), *The political economy of hunger, entitlement and well-being*. New York, NY: Oxford University Press.

Hoddinott, J., & Haddad, L. (1995). Does female income share influence household expenditures? Evidence from Cote d'Iviore. *Oxford Bulletin of Economics and Statistics*, *57*(1), 77–96.

Lillard, L. A., & Willis, R. J. (1994). Intergenerational educational mobility. *Journal of Human Resources*, *29*(4), 1126–1166.

Lloyd, C., & Gage-Brandon, A. (1994). High fertility and children's schooling in ghana: Sex differences in parental contributions and educational outcomes. *Population Studies*, *48*(2), 293–306.

McElroy, M. B., & Horney, M. J. (1981). Nash bargained household decisions. *International Economic Review*, *22*(2), 333–350.

Morduch, J. (2000). Sibling rivalry in Africa. *American Economic Review*, *90*(2), 405–409.

Moulton, B. R. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics*, *32*, 385–397.

Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics*, *72*, 334–338.

Parish, W. L., & Willis, R. J. (1993). Daughters, education and family budgets: Taiwan experiences. *Journal of Human Resources*, *28*(4), 863–898.

Patrinos, H. A., & Psacharopoulos, G. (1997). Family size, schooling and child labor in Peru: An empirical analysis. *Journal of Population Economics*, *10*(4), 387–405.

Pitt, M., & Khandker, R. (1998). The impact of group-based credit programs on poor households in Bangladesh: Does the gender of participants matter? *Journal of Political Economy*, *106*(5), 958–996.

Psacharopoulos, G., & Arrigada, A. M. (1989). The determinants of early age human capital formation: Evidence from Brazil. *Economic Development and Cultural Change*, *37*, 683–708.

Rosenzweig, M. R., & Schultz, T. P. (1982). Market opportunities, genetic endowments and intrafamily resource distribution: Child survival in rural India. *American Economic Review*, *72*, 803–815.

Rosenzweig, M. R., & Wolpin, K. R. (1986). Evaluating the effects of optimally distributed public programs. *American Economic Review*, *76*, 460–482.

Rosenzweig, M. R., & Wolpin, K. R. (1988). Migration selectivity and the effects of public programs. *Journal of Public Economics*, *37*(3), 265–289.

Rosenzweig, M. R., & Wolpin, K. R. (1994). Are there increasing returns to the intergenerational production of human capital-maternal schooling and child intellectual achievement? *Journal of Human Resources*, *29*(2), 670–693.

Strauss, J., & Thomas, D. (1995). Human resources: Empirical modelling of household and family decisions. In: T. N. Srinivasan & J. Behrman (Eds), *Handbook of development economics* (Vol. 3). New York: North Holland Press.

Strauss, J. (1990). Household, communities and preschool child nutrition outcomes: Evidence from Cote d'Ivoire. *Economic Development and Cultural Change*, *38*(2), 231–261.

Unni, J. (1998). Gender differentials in schooling. In: M. Krishnaraj, R. M. Sudarshan & A. Shariff (Eds), *Gender, population and development*. Delhi: Oxford University Press.

Wooldridge, J. M. (2003). *Introductory econometrics*. Australia: Thomson.

MINIMUM WAGE EFFECTS ON WAGES, EMPLOYMENT AND PRICES IN BRAZIL

Sara Lemos

ABSTRACT

This paper presents new evidence on the effects of the minimum wage using Brazilian monthly household and firm panel data between 1982 and 2000. By examining the effects on wages, employment and prices together we are able to provide an explanation for the small employment effects prevalent in the literature. Our principal finding is that increasing the minimum wage raises wages and prices with small adverse employment effects. This suggests a general wage-price inflationary spiral, where persistent inflation offsets some of the wage gains. The main policy implication deriving from these results is that the potential of the minimum wage for the policy maker as a tool to help the poor is bigger under low inflation. Under high inflation, the resulting wage-price spiral makes the minimum wage increase – as well as its antipoverty policy potential – short lived. In this case, the wage effects are volatile and the permanent scars are lower employment and higher inflation in Brazil.

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 397–413

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26011-4

1. INTRODUCTION

The minimum wage helps the poor if it increases wages and does not destroy jobs or cause inflation. It is well established in the literature that minimum wage increases compress the wage distribution (Brown, 1999). As a result, the policy debate hinges on whether employers respond to the associated higher labor costs by reducing profits, reducing employment, or raising prices. Firstly, the empirical evidence on the profit effects is very limited, but standard theory suggests that low wage firms operate in competitive markets with zero profits (Card & Krueger, 1995). Therefore, changes in profits are hard to detect. Secondly, evidence of negative employment effects, predicted by the standard theoretical model, conflicts with evidence of non-negative effects in the literature. Although there is yet no consensus, small employment effects have been frequently reported (Freeman, 1996; Brown, 1999; Dickens, Machin, & Manning, 1999). Thirdly, with employment and profits not significantly affected, higher prices are the obvious alternative response to minimum wage increases. This is consistent with the standard theory prediction that an industry wide cost shock is passed on to prices. Nonetheless, there is very little empirical evidence on price effects in the literature (Brown, 1999; Lemos, 2004a).

The main contribution of this paper is to present new evidence on all three of these minimum wage effects together. By examining wages, employment and price effects together, we are able to provide an explanation for the small employment effects prevalent in the literature. This has potentially important policy implications, and yet empirical analysis has been unable to shed sufficient light at it. The price effect evidence we provided is, in turn, another contribution of this paper to a very under researched area.

A further contribution of this paper is to provide evidence on what Brown (1999, p. 2157), in his recent comprehensive survey, reckons is "the largest and most important gap in the minimum wage literature". We estimate anticipated and lagged wages, employment and price responses to minimum wage increases. This is another aspect of minimum wage effects that has important policy implications, as we demonstrate in this paper.

The data used is monthly Brazilian household and firm panel data from 1982 to 2000. As the non-US literature is relatively scarce, an additional contribution of this paper is to extend the current understanding on the effects of the minimum wage in developing countries. The limited available empirical evidence for Brazil suggests that the minimum wage compresses the wage distribution and has a small adverse employment effect (Fajnzylber, 2001; Carneiro, 2002; Neumark, Cunningham, & Siga, 2006).

Our principal finding is that increasing the minimum wage raises wages and prices with small adverse employment effects in Brazil. This suggests a general wage-price inflationary spiral, where persistent inflation offsets some of the wage gains. Minimum wage indexation and reinforced inflationary expectations were a phenomenon first noticed by Gramlich (1976) and Cox and Oaxaca (1981), and more recently discussed by Card and Krueger (1995) and Freeman (1996). If this is the context, it is perhaps not so surprising that adverse employment effects are small. The main policy implication deriving from these results is that the potential of the minimum wage to help the poor is bigger under low inflation. Under high inflation, the resulting wage-price spiral makes the minimum wage increase – as well as its antipoverty policy potential – short lived. In this case, the wage effects are volatile and the permanent scars are lower employment and higher inflation.

Another important finding is that the poorest only benefit from higher wages in the month of the minimum wage increase. However, they start suffering from higher unemployment and inflation one month before. Furthermore, they are faced with higher inflation for the following three months, by which time some of their wages gains are offset. Under this scenario, a better antipoverty policy is perhaps to lower inflation. A stable growing economy will aid the poor perhaps more than quickly eroded minimum wage increases. Other options include structural reforms and direct cash transfers (Harrison, Rutherford, Tarr, & Gurgel, 2004; Jayaraman & Lanjouw, 2004; Bourguignon, Ferreira, & Leite, 2003). The remainder of this paper is organized as follows. Section 2 describes the institutional background of the minimum wage in Brazil. Section 3 describes the data. Section 4 discusses the empirical equations and identification issues. Section 5 presents the results and Section 6 concludes.

2. MINIMUM WAGE INSTITUTIONAL BACKGROUND

The minimum wage was introduced as a social policy in Brazil under the 1940s populist government. After a steep decline during the 1940s, the real minimum wage was adjusted and reached its peak during the boom of the 1950s. It then decreased as a result of the subsequent recession. With the installation of the dictatorship in the mid 1960s, the real minimum wage was systematically devalued because the government associated the then high inflation with wage adjustments. Even after the end of the military regime in the mid 1980s, the minimum wage continued to be used as an anti-inflationary policy throughout the 1980s and most of the 1990s. During this time, minimum wage increases were subject to the rules of five different stabilization plans. The increases were large and frequent, but were quickly eroded by the subsequent inflation.

In early 1986, the nominal minimum wage was increased by 15% and adjusted biannually initially. It was then adjusted whenever inflation was higher than 20%. Despite this, the real minimum wage was 25% lower in mid 1987 than it was in early 1986. The nominal minimum wage was then initially frozen for three months before it was indexed monthly by past inflation. In early 1989, it was again frozen, and in mid 1989 it was again indexed monthly. In early 1990, the real minimum wage was 45% lower than it was in early 1989. In late 1991, the nominal minimum wage was again indexed monthly. In 1993, adjustments were bimonthly and then monthly. In early 1994, adjustments were made daily, which did not prevent the real minimum wage from falling to 40% lower in mid 1994. In mid 1995 the nominal minimum wage was increased by 42%, and since then it has been annually adjusted. Since the mid 1990s, under reasonably stable inflation, the minimum wage has again been used as a social policy.

Since 1984, the minimum wage in Brazil has been the same for all individuals. There have been no differentiated minimum wage rates for different regions, specific demographic groups or labor market categories. Coverage is full, although accommodation and food costs can be deducted from the wage.

3. DATA

The data we use is the PME (Monthly Employment Survey), the PIM (Monthly Industrial Survey), the Consumers Price index, and the minimum wage. All data is available from the IBGE (Instituto Brasileiro de Geografia e Estatistica).

The PME is a rotating household panel, similar to the US Current Population Survey, which has been collected since 1982. The IBGE interviews on average 30.000 households per month in the six main Brazilian metropolitan regions (Salvador, Recife, Belo Horizonte, Rio de Janeiro, Sao Paulo and Porto Alegre). Households are interviewed for four consecutive months, not interviewed for eight months, and then interviewed again for four additional months, before being dropped from the sample. In the PME the panels are refreshed every two years, rather than every year, as is the case in the CPS. The PIM is a rotating firm panel, similar to the US Production Index, which has been collected since 1968. The IBGE interviews on average 6.000 firms per month in most of the Brazilian metropolitan regions including the six regions above. Firms are assigned a random number when they are first selected for the sample. They are then interviewed monthly for a maximum of four years, but they may be dropped from the sample before then, depending on the initial random number assigned. The sample is refreshed once a year.

We aggregate the PME and PIM across regions and months; the average number of observations per region-month cell is respectively 13,000 and 600. The cross-region variation in the data is considerable and we exploit this in order to identify the minimum wage effect in the econometric models below. In Table 1 we show statistics for the poorest region (Recife) and the richest region (Sao Paulo) in the sample. Wages, prices and employment are lower in Recife, where the fraction of workers earning the minimum wage is larger. In Fig. 1 we show that the patterns of the log nominal minimum wage and average log wages in differences are remarkably synchronized in the aggregate over time, with a correlation of 0.77. In Fig. 2 we show that the correlation between the log nominal minimum wage and the employment rate in differences is much weaker, 0.09. Finally, in Fig. 3 we show that the patterns of the log nominal minimum wage and log prices in differences are also fairly synchronized, with a correlation of 0.55.

4. EMPIRICAL EQUATION SPECIFICATIONS

4.1. Wage Effects

A standard empirical wage equation in the literature is delivered by a labor market equilibrium reduced form equation (Brown, 1999; Card & Krueger, 1995):

$$\begin{aligned}\Delta \ln W_{rt} = \alpha^{w} + \sum_{l=-k}^{L} \beta_{l}^{w} \Delta \ln MW_{t-l} + \gamma^{w}\pi_{rt-1} + \delta^{w}\Delta u_{rt-1} \\ + \lambda^{w} X_{rt} + f_{r}^{w} + f_{t}^{w} + \epsilon_{rt}^{w}\end{aligned} \tag{1}$$

where W_{rt} is nominal average wages in region r and month t, $r = 1,\ldots,6$, and $t = 1,\ldots,214$; MW_t is nominal minimum wage; π_{rt-1} is past inflation; u_{rt-1} is the past unemployment rate; f_r^w and f_t^w are region and time fixed effects; X_{rt} are labor supply shifters; and ϵ_{rt}^w is the error term. The supply shifters we include are the proportion of the total population who are younger than 10 years old, between 10 and 24 years of age, women, illiterates, retirees,

Table 1. Descriptive Statistics Across Regions.

Variables	Recife (Poor Region)	Sao Paulo (Rich Region)
Average hours worked in the labor force	18.56	34.26
Hours worked per worker	38.61	41.31
Employment rate	44.9%	46.3%
"Fraction (of workers) at" the minimum wage	15.1%	4.0%
Log price index	−9.01	−9.13
Log real minimum wage	4.95	5.09
Log 25th percentile real earnings distribution	5.12	5.70
Log 50th percentile real earnings distribution	5.61	6.18
Log 75th percentile real earnings distribution	6.23	6.76
Log average real earnings distribution	5.72	6.26
Log standard deviation real earnings distribution	0.87	0.85
Log price of industrial power consumption	7.93	9.30
Log of average productivity in the metallurgic industry	0.14	0.21
Percentage of Population which is:		
Aged 0 to 14 years old	0.18	0.15
Aged 15 to 24 years old	0.27	0.25
Aged 25 to 64 years old	0.47	0.53
Aged over 65 years old	0.07	0.07
Women	0.45	0.43
Students	0.31	0.22
Enrolled in schooling	0.38	0.31
Literates	0.86	0.95
Elementary education (8 years of schooling)	0.43	0.38
Secondary education (11 years of schooling)	0.14	0.14
Graduates	0.08	0.11
Retired	0.13	0.11
Urban	0.93	0.97
Percentage of Workers in the:		
Metallurgic industry	0.07	0.19
Building construction	0.03	0.04
Commerce	0.09	0.09
Services	0.26	0.29
Public sector	0.07	0.05
Informal sector	0.23	0.36
Sample size	1507171	3292027

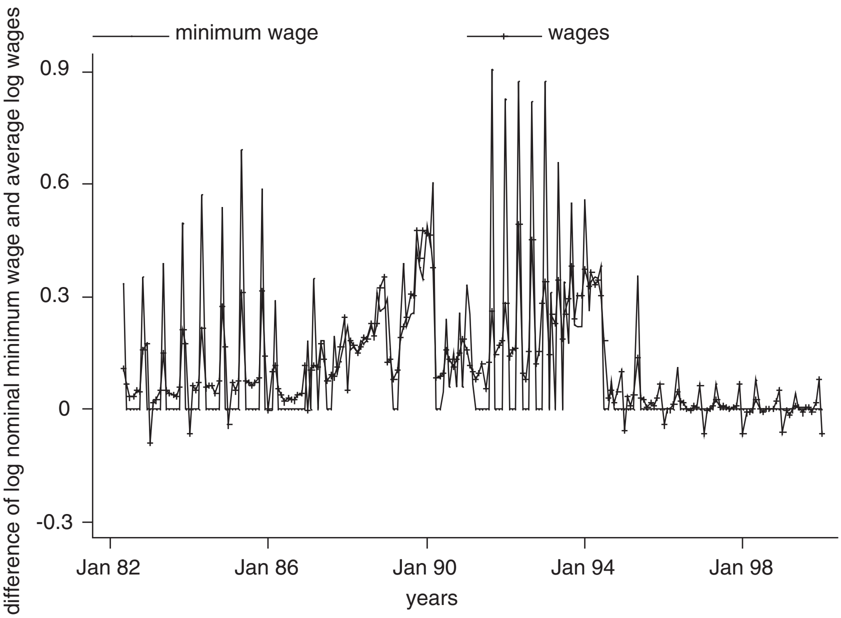


Fig. 1. Minimum Wage and Wages, Brazil 1982–2000.

students, in urban areas, with completed basic (8 years) education and high school (11 years) education; the average years of schooling in the total population; the proportion of the working population holding two jobs, in the informal, public, construction and metallurgy sectors. We include lags and leads of the minimum wage (indexed by $l = -k,\ldots,L$) to allow the effect of the minimum wage on average wages to be complete. The number of lags and leads is an empirical matter and is discussed in Section 5. A GLS correction is performed in all models in the paper to correct for heteroskedasticity arising from aggregation and to account for the relative importance of each region. Also, standard errors are corrected for serial correlation across and within regions.[1]

We re-estimate Eq. (1) taking W_{rt} to mean, in turn, the 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th and 90th percentiles of the wage distribution. This gives an overall picture of the effect of the minimum wage in the entire wage distribution (Dickens et al., 1999). Because the nominal minimum wage is constant across regions in Brazil, we cannot use it as our shock variable. This also prevents us to use other common minimum wage variables such as the real minimum wage and the "Kaitz index", which is

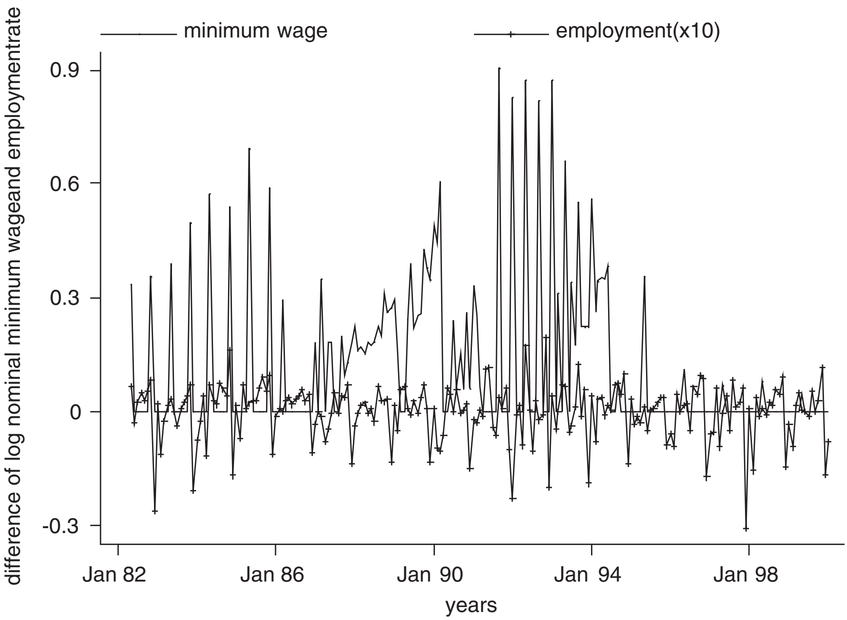


Fig. 2. Minimum Wage and Employment, Brazil 1982–2000.

defined as the ratio of the minimum wage to average wage adjusted for coverage of the legislation (Kaitz, 1970). This is because the variation in these "relative minimum wage" measures is driven by the variation in their denominator, which does not ensure full identification of the minimum wage effect (Welch & Cunningham, 1978). Consequently, "degree of impact" measures are becoming more common in the literature (Brown, 1999). Examples are "fraction affected" (Card, 1992), which is defined as the proportion of workers earning a wage between the old and the new minimum wage, and "fraction at" (Dolado et al., 1996), which is defined as the proportion of workers earning one minimum wage. The rationale is that an increase in the nominal minimum wage affects a different proportion of people across regions depending on the initial level and shape of the wage distribution in each region. Although these two variables are closely related, "fraction affected" does not capture the erosion of the minimum wage in relation to other wages, while "fraction at" does. This is because "fraction affected" is constant at zero when the minimum wage is constant (Brown, 1999). Thus, in a similar fashion to Card and Krueger (1995), we use

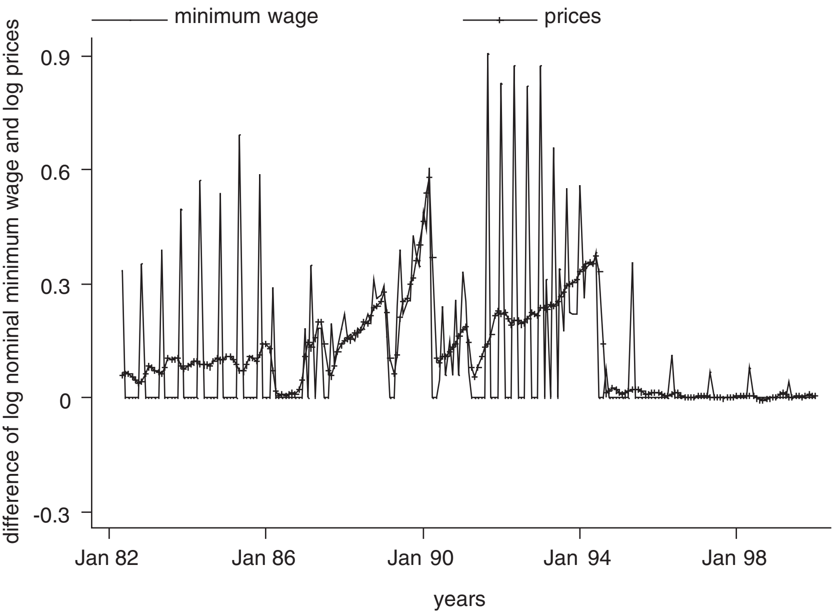


Fig. 3. Minimum Wage and Prices, Brazil 1982–2000.

"fraction at" (plus or minus 0.02%)[2] in place of the log nominal minimum wage in Eqs. (1) (2) and (3) to ensure identification of the effect of the minimum wage.

4.2. Employment Effects

The counterpart empirical employment equation (Brown, 1999) is:

$$\Delta \ln N_{rt} = \alpha^n + \sum_{l=-k}^{L} \beta_l^n \Delta \ln MW_{t-l} + \gamma^n \pi_{rt-1} + \lambda^n X_{rt} + f_r^n + f_t^n + \epsilon_{rt}^n \tag{2}$$

where N_{rt} is taken in turn to mean total average hours worked in the labor force (includes zero hours worked for those unemployed) T, average hours worked for those working (hours worked per worker) H, and the employment rate E. As Eq. (2) is separately estimated using each of these three dependent variables, the estimates in the T equation equal the sum of the estimates in the H and E equations, i.e. $\beta^n_T = \beta^n_H + \beta^n_E$. This makes it possible to decompose the total effect of a minimum wage increase on employment into hours effect and jobs effect. This decomposition is important because employment can be adjusted in two margins following a minimum wage increase: the number of jobs and the number of hours per worker. Some authors have argued that the non-negative jobs effects often found in the literature might be a sub-product of adjustments in hours (Card & Krueger, 2000; Neumark & Wascher, 2000).

4.3. Price Effects

A standard empirical price equation – largely used in the literature on the price response to industry wide shocks (Poterba, 1996; Goldberg & Knetter, 1997) – is the inverse of the profit maximizing condition under imperfect competition. This equation expresses prices as a markup over costs:

$$\Delta \ln P_{rt} = \alpha^p + \sum_{l=-k}^{L} \beta^p_l \Delta \ln MW_{t-1} + \xi^p \Delta \ln E_{rt} + \delta^p \Delta \ln A_{rt} + f^p_r + f^p_t + \epsilon^p_{rt} \tag{3}$$

where P_{rt} is prices; E_{it} is the cost of industrial power consumption, and A_{it} is productivity. We define productivity as the total industrial output divided by total number of workers directly employed in production in the metallurgy industry.[3] The cost of industrial power consumption is a proxy for costs of inputs other than labor.[4]

5. RESULTS

In Table 2 we show generalized least squares β estimates. Row 1 shows evidence of anticipated effects of the minimum wage on average wages, but no evidence of lagged effects. The coefficient of the first lead of the shock variable, one month before the increase, is positive and significant. The contemporaneous coefficient is also positive and robust. The coefficients of further leads and lags are not statistically different from zero. This suggests that on average, wages adjustment in response to minimum wage increases happens in the month of the increase and in the month before, and that no lagged adjustment follows the increase. However, the estimate of the minimum wage effect on average wages is a summary measure of wage effects

Table 2. Effect of the Minimum Wage on Wages, Employment and Prices.

Dependent Variable	2 Months Before		1 Month Before		Month of the Increase		One Month After		2 Months After		Total	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
Average wage	0.02	0.17	0.39	0.18	0.51	0.19	−0.07	0.18	−0.24	0.17	0.59	0.60
10th wage distribution decile	−0.42	0.25	0.09	0.25	1.60	0.26	−0.02	0.25	−0.82	0.25	0.43	0.74
20th wage distribution decile	−0.38	0.27	0.33	0.27	1.28	0.27	−0.59	0.27	−0.84	0.27	−0.20	0.80
30th wage distribution decile	−0.23	0.22	0.47	0.23	0.49	0.23	−0.38	0.23	−0.51	0.23	−0.16	0.69
40th wage distribution decile	−0.06	0.19	0.40	0.19	0.26	0.20	−0.13	0.20	−0.48	0.19	−0.02	0.61
50th wage distribution decile	−0.04	0.17	0.31	0.18	0.20	0.18	−0.12	0.18	−0.30	0.17	0.04	0.56
60th wage distribution decile	0.15	0.17	0.34	0.18	−0.03	0.18	−0.02	0.18	−0.30	0.17	0.14	0.55
70th wage distribution decile	0.11	0.16	0.39	0.18	0.16	0.18	−0.14	0.18	−0.11	0.17	0.42	0.57
80th wage distribution decile	0.12	0.17	0.42	0.18	−0.10	0.18	−0.02	0.18	−0.21	0.17	0.21	0.56
90th wage distribution decile	0.28	0.17	0.40	0.18	−0.02	0.19	−0.04	0.19	−0.21	0.18	0.42	0.59
Total hours worked	−0.04	0.05	−0.10	0.05	−0.03	0.05	−0.02	0.05	−0.07	0.05	−0.26	0.13
Hours worked per worker	−0.04	0.05	−0.09	0.05	−0.01	0.05	0.00	0.05	−0.08	0.05	−0.22	0.13
Employment rate	0.00	0.01	−0.01	0.01	−0.02	0.01	−0.02	0.01	0.01	0.01	−0.05	0.03
prices	0.10	0.09	0.28	0.12	0.38	0.13	0.32	0.12	0.17	0.09	1.26	0.45

(a) The dependent variable is, in turn, the log of various deciles of the wage distribution, (average) total hours worked for the labour force, hours worked per worker, the employment rate, and logs of prices. The hours worked per worker estimate plus the employment rate estimate add to the total hours worked estimate.

(b) These are the GLS estimates of the shock variable "fraction at" in Eqs. (1) to (3). The weights are the square root of the inverse of the sample size. Standard errors are White-corrected and serial correlation corrected across and within regions.

(c) Labour supply shifters are included as controls in the wages and employment equations, namely, the proportion of the total population younger than 10 years old, between 10 and 24 years of age, women, illiterates, retirees, students, in urban areas, with completed basic and high school education; the average years of schooling in the total population; the proportion of the working population corresponding to workers holding two jobs, workers in the informal, public, construction and metallurgy sectors. A measure of productivity and a measure of other inputs' prices is included in the price equation.

(d) To reflect a 10% increase in the minimum wage, the estimates and standard errors need to be multiplied by 0.6, which is the approximate elasticity of the minimum wage with respect to "fraction at".

throughout the wage distribution. A closer look at the estimates of the minimum wage effect on each decile of the wage distribution reveals a more intricate picture. For example, while there is evidence of lagged, but not anticipated effects at the very bottom of the distribution; conversely, there is evidence of anticipated but not lagged effects at the top half of the distribution. This suggests that the higher paid workers have greater bargaining power and revise their labor contracts in anticipation of the minimum wage increase.

Row 2 shows that the contemporaneous coefficient is positive and robust at the 10th percentile of the distribution. It is three times larger than the coefficient for the average wages (row 1). This suggests that the wages of the poorest increase three times more than average wages do. However, the coefficient of the second lag of the shock variable, two months after the increase, is negative and significant. It is half the size of the contemporaneous coefficient. This suggests that after two months, the poorest lose half of the wage gains they had in the month of the increase. Neumark, Schweitzer, and Wascher (2004) also find evidence of strongly negative lagged minimum wage effects for the US. They argue that employers take advantage of inflation in the following periods to partly undo the wage gains resulting from minimum wage increases. Row 3 shows a similar picture for the 20th percentile.

The results for the 30th percentile in row 4 show that the contemporaneous coefficient is positive and significant. It is about as large as the coefficient for the average wages (row 1). The coefficient of the first lead is positive and significant and the coefficient of the second lag is negative and significant. As both are roughly of the same magnitude, whatever those at the 30th percentile gain one month before the increase, they loose two months after the increase. In the remainder of the distribution, anticipated gains are roughly about the same magnitude as the effect on average wages (row 1). Further leads and lags are not statistically different from zero. This suggests that most labor contracts – especially those of higher paid workers – are revised in anticipation of the minimum wage increase.

Concurrently, there is a decrease in total hours worked in the labor force. Row 11 shows a negative and significant effect on total hours worked one month before the increase. Further leads and lags are not statistically different from zero. Row 13 shows that the coefficient of the first lead of the employment rate is not statistically different from zero. This suggests that while revising labor contracts in anticipation of the minimum wage increase, employers and employees negotiate not only wage increases, but also the number of hours worked. It also suggests that employers do not fire employees at this stage. Instead they first increase prices to offset some of the higher labor costs, as shown by the significant and positive coefficient of prices in row 14. However, in the month of the increase and in the subsequent month, not only do employers continue to increase prices, but they also start adjusting employment through firing employees. The coefficient of the employment rate is negative and significant in the month of the increase and the following month, while the coefficient of prices is significant and positive for four consecutive months. The price coefficient is about three quarters of the average wage coefficient (row 1) in the month before and in the month of the increase. The prices coefficient remains positive and significant in the two following months, even though wage effects become negative and often insignificant. This suggests more stickiness in price than in wages following a minimum wage increase. These results are consistent with those of Aaronson (2001), who included lags and leads of the minimum wage in his price equation specifications. He found that in the US most of the price response occurs in the two months period immediately after a minimum wage increase.

The last column of Table 2 shows long run effects. The wage effects are not statistically different from zero, suggesting no wage gains associated with the minimum wage increase in the long run. The long run total hours worked effect is significant, although month-by-month this effect is mostly insignificantly different from zero. The price effect is positive and significant, consistent with month-by-month persistent increases. This suggests that firms' responses to higher labor costs resulting from minimum wage increases is a mix of lower employment and higher prices.

In summary, the anticipated wage gains are roughly about the same magnitude throughout the wage distribution (except for the very poor) one month before the increase. The price effects are about half the size and there is no evidence of disemployment effects in that month (although there is some evidence of reduction in hours worked). This suggests a general wage-price spiral, where nominal variables are affected but not real ones. In the month of the increase, the poorest benefit relatively more than other workers, as there is no spillover effects above the 30th percentile. However, the inflation effects are now larger and persistent, and some small disemployment effects start to take place. One month after the increase, inflation persists and some of the wage gains are undone for the poor, with some further small disemployment effects. Finally, two months after the increase inflation starts to ease, employment effects disappear and those at the bottom half of the distribution have wage losses. In the long run, the wage effects are volatile and the permanent scars are lower employment and higher inflation.

We calibrate the estimates above to ensure comparability with those in the literature (Brown, 1999; Card & Krueger, 1995). Following Card and Krueger (1995), the ''fraction at'' estimates are multiplied by 0.6, which is the approximate elasticity of the ''fraction at'' with respect to the nominal minimum wage (Lemos, 2004b). A 10% increase in the minimum wage decreases employment by 0.2% and increases prices by 0.8% in the long run. These results are in line with previous evidence for Brazil, where wage effects of minimum wage increases are large and employment effects are small (Fajnzylber, 2001; Carneiro, 2002; Neumark, Cunningham, & Siga, 2006). Our results compare with, respectively, 1% (mainly in the food industry) employment decrease and 0.2% to 0.4% economy wide price increases for the US (Brown, 1999; Sellekaerts, 1981; MaCurdy & McIntyre, 2001). Thus, a smaller employment effect in Brazil is consistent with a larger price effect. However, these are economy wide estimates that might have diluted more negative employment effects in low wage industries.

6. CONCLUSIONS

This paper fills a gap in the literature by providing an overall picture on the effects of the minimum wage on wages, employment and prices using monthly Brazilian monthly household and firm panel between 1982 and 2000. The evidence we provide indicates that increasing the minimum wage raises wages throughout the wage distribution in the month before the increase, although it only raises the wages of the poorest in the month of the increase. However, persistent inflation effects offset some of the wage gains in the following months. This suggests a general wage-price spiral, where nominal variables are affected but not real ones. It is then perhaps not so surprising that adverse employment effects are small. Small employment effects – frequently reported in the recent literature – are sensible when relatively large price effects are uncovered. In the long run, the wage effects are volatile and the permanent scars are lower employment and higher inflation.

A 10% increase in the minimum wage decreases employment by 0.2% and increases prices by 0.8% after five months of adjustment, when wage gains have already vanished. These results compare with, respectively, 1% (mainly in the food industry) employment decrease and 0.2% to 0.4% economy wide price increases for the US. One potential criticism here is that aggregate estimates might have diluted more negative employment effects in low wage industries. Estimates for such industries are not available for Brazil. Thus, a fruitful avenue for future research is to estimate wages, employment and price effects for industries overpopulated by minimum wage workers in Brazil and other developing countries.

The main policy implication deriving from these results is that the potential of the minimum wage for the policy maker as a tool to help the poor is bigger under low inflation. Under high inflation, the resulting wage-price spiral makes the minimum wage increase – as well as its antipoverty policy potential – short lived. In this case, the wage effects are volatile and the permanent scars are lower employment and higher inflation in Brazil. The poorest only benefit from higher wages in the month of the minimum wage increase. However, they start suffering from higher unemployment and inflation one month before the increase. Furthermore, they are faced with higher inflation for the following three months, by which time some of their wages gains are offset. Under this scenario, a better antipoverty policy is perhaps to lower inflation. A stable growing economy will aid the poor perhaps more than quickly eroded minimum wage increases.

NOTES

1. The GLS estimates were robust to GMM estimation using lags of the minimum wage variable as well as a number of political variables as instruments (Lemos, 2005). This suggests that any endogeneity bias arising from the simultaneous determination of "fraction at" and employment is not too severe.
2. The bounds account for measurement error introduced by rounding approximations. All estimates in the paper were robust to defining "fraction at" with and without bounds (the correlation between the two is 0.91).
3. Data for all industries was not available, and thus the productivity in the metallurgic industry is taken as a proxy to overall productivity.
4. We also used other measures of labor costs above and below the minimum wage, but this did not alter our main findings. See, for example, Lemos (2004c, 2006).

ACKNOWLEDGMENTS

Special thanks to Alan Manning, Arthur Van Soest, Christian Dustman, David Neumark, Gianni De Fraja, Steve Machin, and Steve Wheatley Price. Thanks to Solomon Polacheck as the editor and to an anonymous referee. Also thanks to Alan Krueger, Charles Brown, Dave Wittenburg, and Kenneth Couch. And thanks to comments of various discussants and participants in the following conferences: EEA-Italy, LACEA-Spain, RES-England, SBE-Brazil, LACEA-Mexico, and IZA-EBRD-Italy.

REFERENCES

Aaronson, D. (2001). Price pass-through and the minimum wage. *Review of Economics and Statistics*, *83*, 158–169.

Bourguignon, F., Ferreira, F. H. G., & Leite, P. G. (2003). Conditional cash transfers, schooling, and child labor: Micro-simulating Brazil's Bolsa Escola program. *The World Bank Economic Review*, *18*, 289–317.

Brown, C. (1999). Minimum wages, employment, and the distribution of income. In: O. Ashenfelter & D. Card (Eds), *Handbook of labor economics* (pp. 2101–2163). Amsterdam; New York and Oxford: Elsevier Science, North-Holland.

Card, D. (1992). Do minimum wages reduce employment? A case study of California, 1987–1989. *Industrial and Labor Relations Review*, *46*, 38–54.

Card, D., & Krueger, A. B. (2000). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania: Reply. *American Economic Review*, *90*, 1397–1420.

Card, D. E., & Krueger, A. B. (1995). *Myth and measurement: The new economics of the minimum wage*. Princeton: Princeton University Press.

Carneiro, F. G. (2002). *Uma Resenha Empirica Sobre Os Efeitos Do Salario Minimo No Mercado De Trabalho Brasileiro*. The Jobs Report, The World Bank, Washington, D.C.

Cox, J., & Oaxaca, R. (1981). Effects of minimum wage policy on inflation and on output, prices, employment and real wage rates by industry. *Report of the Minimum Wage Paper Commission*, *6*, 171–210.

Dickens, R., Machin, S., & Manning, A. (1999). The effects of minimum wages on employment: Theory and evidence from Britain. *Journal of Labor Economics*, *17*, 1–22.

Dolado, J., Kramarz, F., Machin, S., Manning, A., Margolis, D., & Teulings, C. (1996). The economic impact of minimum wages in Europe. *Economic Policy: A European Forum*, *23*, 317–357.

Fajnzylber, P. (2001). Minimum wage effects throughout the wage distribution: Evidence from Brazil's formal and informal sectors. *Anais do XXIX Encontro Nacional de Economia*.

Freeman, R. (1996). The minimum wage as a redistributive tool. *Economic Journal*, *106*, 639–649.

Goldberg, P. K., & Knetter, M. M. (1997). Goods prices and exchange rates: What have we learned? *Journal of Economic Literature*, *35*, 1243–1272.

Gramlich, E. M. (1976). Impact of minimum wages on other wages, employment, and family incomes. *Brookings Papers on Economic Activity*, *2*, 409–451.

Harrison, G. W., Rutherford, T. F., Tarr, D. G., & Gurgel, A. (2004). Trade policy and poverty reduction in Brazil. *The World Bank Economic Review*, *18*, 289–317.

Jayaraman, R., & Lanjouw, P. F. (2004). Small-scale industry, environmental regulation, and poverty: The case of Brazil. *The World Bank Economic Review*, *18*, 443–464.

Kaitz, H. (1970). Experience of the past: The national minimum, youth unemployment and minimum wages. *US Bureau of Labor Statistics Bulletin*, *1657*, 30–54.

Lemos, S. (2004a). *A survey of the effect of the minimum wage on prices*. IZA Discussion Paper 1072.

Lemos, S. (2004b). *Are wage and employment effects robust to alternative minimum wage variables?* IZA Discussion Paper 1070.

Lemos, S. (2004c). Minimum wage policy and employment effects: Evidence form Brazil. *Economia*, *5*, 219–266.

Lemos, S. (2005). Political variables as instruments for the minimum wage. *Contributions to Economic Analysis & Policy*, *4*(1), 1–31.

Lemos, S. (2006). Anticipated effects of the minimum wage on prices. *Applied Economics, 2006*, *38*(3), 325–337.

Macurdy, T., & Mcintyre, F. (2001). *Winners and losers of federal and state minimum wages*, Employment Policies Institute Foundation Paper, 6/0.

Neumark, D., Cunningham, W., & Siga, L. (2006). The effects of the minimum wage in Brazil on the distribution of income 1996–2001. *Journal of Development Economics*, *86*(1), 136–159.

Neumark, D., Schweitzer, M., & Wascher, W. (2005). The minimum wage effects throughout the wage distribution. *The Journal of Human Resources*, *39*, 425–450.

Neumark, D., & Wascher, W. (2000). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania: Comment. *American Economic Review*, *90*, 1362–1396.

Poterba, J. M. (1996). Retail price reactions to changes in state and local sales taxes. *National Tax Journal*, *49*, 165–176.

Sellekaerts, B. (1981). Impact of minimum wage legislation on wage and price inflation. *Report of the minimum wage paper commission*, *6*, 1–17.

Welch, F. R., & Cunningham, J. (1978). Effects of minimum wages on the level and age composition of youth employment. *Review of Economics and Statistics*, *60*, 140–145.

BARGAINING AND ARBITRATION WITH ASYMMETRIC UNCERTAINTY☆

Cary Deck and Amy Farmer

ABSTRACT

Arbitration is often used to settle bargaining disputes. Frequently in such disagreements, one party has better information with respect to the surplus to be allocated. This paper considers the impact that the choice of dispute resolution mechanism, conventional or final offer arbitration, has on settlement. This paper shows that theoretically final offer arbitration can systematically favor the informed party by shifting the contract zone towards more profitable allocations while conventional arbitration is theoretically less likely to generate a mutually agreeable settlement. Laboratory results find that the surplus shares are consistent with the predicted favoritism. However, settlement is positively correlated with the width of the contract zone and the data suggest that the location of the contract zone in final offer arbitration generates more disputes.

☆This research was supported by the Decision Risk and Management Science Program of the National Science Foundation (SES-0350709).

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 415–445

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26012-6

When bargaining disagreements occur, arbitration is increasingly being employed to resolve the dispute. The U.S. Supreme Court has ruled that employers can force employees to use arbitration to settle labor disputes (Circuit City Stores Inc. vs. Saint Clair Adams). Many companies have begun including clauses in contracts that stipulate the use of arbitration in the event of a conflict. Arbitration offers several advantages over traditional litigation including shorter decision times, lower costs (Bernstein, 1993) and the ability to not disclose proprietary information (Fuller, 1993). The widespread use of arbitration has lead to the development and implementation of several alternative arbitration methods. Perhaps the simplest form is conventional arbitration (CA) in which an arbiter determines the allocation. In contrast, final offer arbitration (FOA) allows each party to submit a proposal to the arbiter who is bound to select one of the proposals. The prominence of final offer has been enhanced by Major Leagues Baseball's decision to use it in resolving salary disputes. While these are the most commonly used forms of arbitration, practitioners and researchers have developed variants such as tri-offer (Ashenfelter, Currie, Farber, & Spiegel, 1992), combined offer (Brams & Merril, 1986), and amended final offer (Zeng, 2003).

From a theoretical standpoint, the existence of arbitration essentially provides both parties with a threat point during the initial bargaining period. Given what each side expects to receive in the event that arbitration is employed, the two parties know the range of outcomes that would be mutually preferred to arbitration. The set of mutually beneficial outcomes is referred to as the contract zone. Regardless of the width of the contract zone, as long as it exists, players with symmetric information should always reach a settlement. Of course, in naturally occurring settings, disputes that reach arbitration are frequently observed. The same is true of controlled laboratory investigations of bargaining and arbitration; see Ashenfelter et al. (1992); Dickinson (2004), and Deck and Farmer (2003). From a behavioral standpoint, it is not clear how the size of the contract zone impacts settlement. One view is that the larger the contract zone the more room there is for compromise and agreement. Currie and McConnell (1991) and Babcock, Lowenstein, Issacharoff, and Camerer (1995) Babcock, et al. (1995) have found empirical support for this view using field data. The opposing view is that a smaller contract zone offers fewer alternatives and thus reduces conflict about how to divide the gains from settlement.

In the traditional model, two parties divide a known fixed surplus. To explain why arbitration is regularly used, researchers have developed a richer set of models. One explanation that has received considerable attention is asymmetric information typically with respect to the settlement the arbiter is likely to choose.[1] Different arbitration mechanisms generate different incentives with regard to bidding strategies and making and accepting offers, and consequently the choice of mechanism can affect both settlement rates and the ultimate allocation.[2] An alternative explanation is disputant optimism in which agents with similar information still fail to agree.[3] Both of these alternative explanations have received support from recent experimental analysis.[4]

Deck and Farmer (2003) present an alternative model of arbitration in which parties are unsure of the surplus to be divided. Rather than modeling the dispute as a zero sum game as would be expected in a civil suit regarding damages, for example, they examine bargaining and arbitration when neither party the true value to de divided. This is more representative of many situations such as labor markets where employers and employees do not know what surplus will be created if the job applicant is hired.[5] The uncertainty could be due to uncertainty about the applicant's ability or fit with the organization's needs or it might represent uncertain future conditions in the output market. In this framework, the employee's wage equals the payout of the employer, but the employer faces uncertainty even after settlement.

The goals of the employee and employer are to maximize and minimize that payment, respectively.[6] This is not unlike the strategies of a plaintiff or defendant in a model of civil litigation, and previous models of arbitration follow the setup of those litigation models with arbitration as the alternative rather than trial. However, that literature emphasizes theoretical analyses of settlement rates, and is largely unconcerned with how arbitration mechanisms affect the distribution of the allocation or the size of the contract zone available for settlement.[7] Deck and Farmer (2003) find that the chosen arbitration method can systematically favor one party over the other, even when the arbiter is unbiased. Specifically, FOA is found to shift the contract zone towards lower wages relative to CA.[8]

Our paper introduces asymmetric information into the model of Deck and Farmer (2003). Specifically, we consider the case where the employer has better information regarding the value of the worker. This paper provides a two pronged approach for exploring this type of asymmetry in arbitration. First, we provide an initial theoretical treatment of this situation in the next section. One important theoretical finding is that with the informational asymmetry the contract zone may fail to exist in certain cases even with risk neutrality and consistent beliefs regarding the arbitrator's preferences; however, we find that if it exists in CA it will also exist in FOA.

Given that this may clearly influence settlement rates, this provides an additional dimension upon which these mechanisms might be compared. The theoretical findings are complimented by a separate section detailing a series of laboratory experiments that allow us to examine behavior in this new setting. With these experiments we can identify behavioral regularities, evaluate the predictive power of the model, and provide insight into areas where the theory is silent such as how contract zone width impacts agreement rates. Taken together, the theory and the experiments provide a more complete picture of arbitration with asymmetric uncertainty.

THEORETICAL MODEL

The framework of the model follows that in Deck and Farmer (2003), but the information structure differs. Consider a risk neutral worker and a risk neutral firm bargaining over the worker's wage. The employment of the worker generates some total value to the employer who retains the residual surplus after having paid the worker. Define w to be the wage payment. Now suppose that the total value available to be allocated between the two players is uncertain and that only the firm is informed about the true level of surplus that the worker's employment will generate. Note that Deck and Farmer (2003) assume that neither party knows the true value, and as a result their baseline model does not involve an informational asymmetry. The asymmetry in this model could be the result of information the firm has concerning its needs and the ability of the employee to fulfill them.[9] If negotiations fail, the dispute may be settled by an arbiter who is informed about the true value and whose preferences for the allocation are dependent upon this true value.[10] While neither the firm nor the worker knows the arbiter's preferences concerning how the surplus should be allocated, the firm is informed about the distribution from which that decision will be made. The worker, on the other hand, does not know the true value and hence does not know the exact distribution from which the arbiter's preferences will be drawn.[11]

The basic structure of the bargaining game is as follows:

(0) The firm and the worker observe 2 possible values (high and low) generated by the worker's employment. Denote these two amounts as v_H and v_L where $v_H > v_L$, and v_H has a probability p of being the true value.
(1) Nature determines the true amount of money available to be either v_H or v_L. The firm is then informed of this result, but the worker is not. The

arbiter's preference is drawn from f_H or from f_L depending on the true value. The firm knows the true distribution describing the arbiter's preferences while the worker only knows that the distribution will be f_H with probability p and f_L with probability $1-p$.

(2) In the first round of bargaining, the worker and firm submit an offer; the worker asks for a wage a while the firm submits a wage offer, or bid for the labor, denoted b.

(3) If the players' bids are compatible, i.e., if $b > a$, then the worker receives $w = (a+b)/2$ while the firm receives a residual surplus equal to either v_H-w or v_L-w depending on the result in stage 1. Note that it is possible for the firm to receive a negative payment. If the bids are incompatible, the game proceeds to stage 4.

(4) The allocation, w, is determined via conventional or FOA.

The following subsections explore each alternative dispute resolution technique and what is the optimal behavior in each case. Specifically, the contract zone, optimal bidding behavior and expected profits for each player are considered. Fig. 1 presents the complete game tree for each alternative using the bargaining process described above. We then compare the theoretical properties of the mechanisms and conclude with a summary of our theoretical findings.

Conventional Arbitration

If bids in round 1 are incompatible, an arbiter will decide the wage based on the known value of the surplus. Define Y to be the arbiter's preference for how this surplus should be allocated. Thus $Y \sim f_H$ when the true value is v_H and $Y \sim f_L$ when the true value is v_L. From the worker's perspective $Y \sim \mathrm{f_H}$ with probability p and $Y \sim f_L$ with probability $1-p$. Further, arbitration is assumed to be costly. Let c_w and c_f denote the costs of proceeding to arbitration for the worker and the firm respectively. Now stage 4 becomes 4′.

(4′) The arbiter determines the final allocation based on his preferences, $w = Y$. The worker receives w-c_w and the firm receives either $v_H - w - c_f$ or $v_L - w - c_f$ depending on the realization of total value.

What will be the players' optimal bidding strategy in round 1 given that CA is the dispute resolution mechanism? If workers end up in arbitration, the expected value of w will be $p\mu_H + (1-p)\mu_L$ where μ_H and μ_L are the means of the distribution f_H and f_L respectively; thus, the worker will settle for $p\mu_H + (1-p)\mu_L - c_w$ to avoid arbitration. A firm who knows the true value to

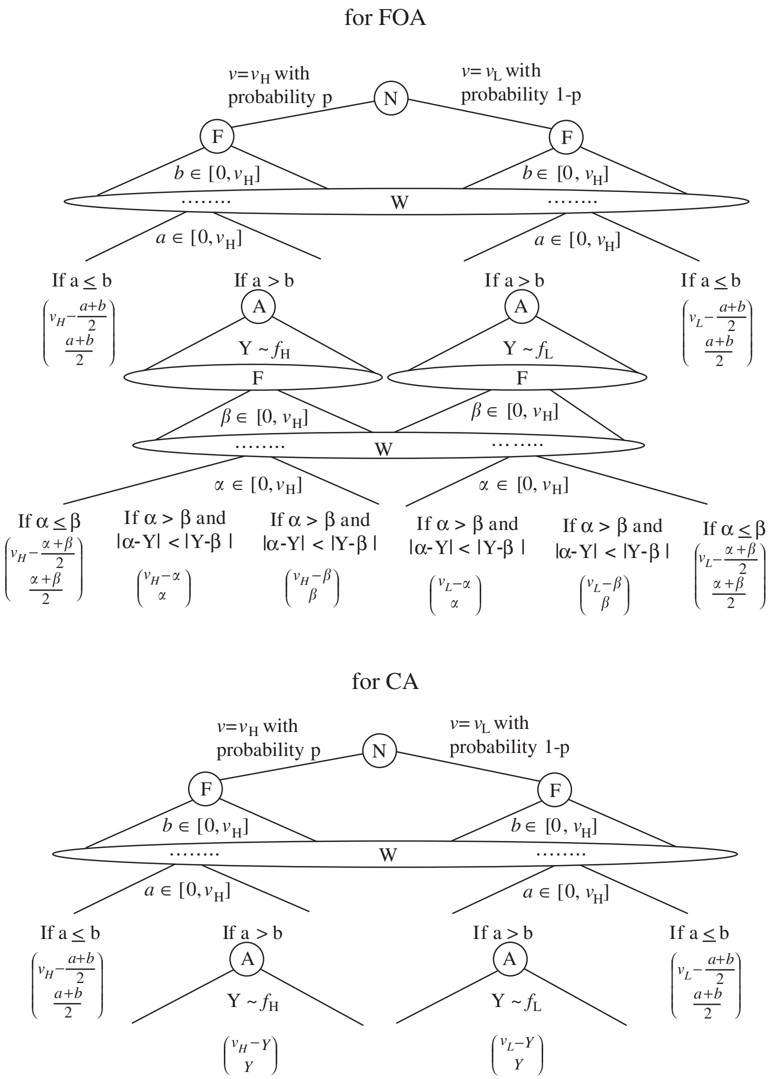


Fig. 1. Complete Game Tree.

be v_H will pay $\mu_H + c_f$ to avoid arbitration while a firm who knows the true value to be v_L will offer no more than $\mu_L + c_f$. As a result, the firm knows the true contract zone, but the worker only knows that the contract is one of two possibilities with probability p and $1-p$. This generates Result 1.

Result 1. When CA is the dispute resolution mechanism, the contract zone in round 1 is either $[p\mu_H + (1-p)\mu_L - c_w, \mu_H + c_f]$ *or* $[p\mu_H + (1-p)\mu_L - c_w, \mu_L + c_f]$ .The firm knows the true contract zone and the worker knows only the probability that it is one or the other.

Note that the expected size of the contract zone is the sum of the arbitration costs $c_w + c_f$. However, examination of the two possible contract zones reveals that for some parameter values the low value contract zone may fail to exist; i.e., the upper bound is below the lower bound. By the upper bound of the contract zone we mean the maximum wage a firm would offer and by the lower bound we mean the minimum wage a worker would accept. This result is similar in spirit to the traditional litigation models in which asymmetric information can produce settlement failure. Intuitively, if the likelihood of a high value is large but the true value observed by the firm is low, the worker will be unwilling to settle for a wage that is compatible with what the employer is willing to pay. Similarly, if the difference between a high and a low value outcome is great, it is possible that when the firm knows the truth to be low but the worker is making decisions based on the expected value, the divide between their positions is too large to overcome despite the costs of arbitration. These possibilities are explored further in section ‘‘Behavioral Examination.’’

Final Offer Arbitration

In FOA, the arbiter is forced to choose between two final offer bids submitted by each player; there is no freedom to choose some middle ground. Thus step 4 is replaced with 4″.

(4") The worker and firm submit final proposals α and β respectively. The arbiter chooses the proposal that lies closer to his preferred valuation, denoted Y; recall that from the worker's perspectives the arbiter's valuation is drawn from distribution f_H with probability p and from f_L with probability $1-p$ while the firm knows the correct distribution. Thus, $w=\alpha$ or $w=\beta$ and the worker receives $w-c_w$ while the firm receives either v_H-w-c_f or v_L-w-c_f.

The arbiter's decision rule is to choose

$$w = \alpha \text{ if } \alpha - \mathrm{Y}| < |\mathrm{Y} - \beta|$$
$$w = \beta \text{ otherwise}$$

If $Y \sim f_H$, then the firm's bid is chosen with probability $F_H((\alpha+\beta)/2)$. If instead $Y \sim f_L$, the firm wins with probability $F_L((\alpha+\beta)/2)$.[12] Given this decision rule, the worker's optimal bids in stage 4″ can be found by maximization of Eq. 1.

$$E\pi_w = p\left[\beta_H F_H\left(\frac{\alpha+\beta_H}{2}\right) + \alpha\left(1 - F_H\left(\frac{\alpha+\beta_H}{2}\right)\right)\right] + (1-p)\left[\beta_L F_L\left(\frac{\alpha+\beta_L}{2}\right) + \alpha\left(1 - F_L\left(\frac{\alpha+\beta_L}{2}\right)\right)\right] \quad (1)$$

Note that the worker will only make one bid, α, but the firm's bid will be specific to the true surplus since the firm will place its bid with that information in mind. The firm's optimization can be found in Eq. (2) where (2a) represents expected profit when the surplus is high and (2b) represents the expected profit when the surplus is low.

$$E\pi_{fH} = v_H - \left[\beta_H F_H\left(\frac{\alpha+\beta_H}{2}\right) + \alpha\left(1 - F_H\left(\frac{\alpha+\beta_H}{2}\right)\right)\right] \quad (2a)$$

$$E\pi_{fL} = v_L - \left[\beta_L F_L\left(\frac{\alpha+\beta_L}{2}\right) + \alpha\left(1 - F_L\left(\frac{\alpha+\beta_L}{2}\right)\right)\right] \quad (2b)$$

Both players choose their bids to maximize expected profits in this round. Given the bid/ask pair, substitution back into (1) and (2) and subtracting their respective arbitration costs provides the expected profit each player can expect if they should reach FOA. Plug in these optimized values to find each player's expected payout in arbitration, denoted π_w^* and π_f^*H, and π_f^*L. As was true with CA, the worker does not know at what amount the firm is willing to settle. This leads us to Result 2.

Result 2. Depending upon the realization of v, the contract zone is either $[\pi_w^* - c_w, \pi_f^*L + c_{fL}]$ or $[\pi_w^* - c_w, \pi_f^*H + c_{fH}]$. The firm knows the truth and the worker knows only the probability of each contract zone.

As was true with CA, for some parameter values the low value contract zone may fail to exist.

Arbitration Mechanisms Using a Uniform Distribution

To generate explicit predictions regarding bidding behavior and expected profits, some specificity in terms of the distributions f_H and f_L must be assumed. For this theoretical treatment as well as the experimental design that follows, these distributions are assumed to be uniform.[13] Let $f_H \sim U[0,v_H]$ and $f_L \sim U[0,v_L]$; thus, $F_H(w) = w/v_H$ and $F_L(w) = w/v_L$. Using these distributions, reconsider the 2 results above. Result 1 can be rewritten as follows.

Result 1′. When CA is the dispute resolution mechanism and the arbiter's value is drawn from a uniform distribution for each of the true values, the contract zone in round 1 is $[p\ v_H/2+(1-p)\ v_L/2-c_w,\ v_H/2+c_f]$ *or* $[p\ v_H/2+(1-p)\ v_L/2-c_w,\ v_L/2+c_f]$. The firm knows the truth and the worker knows that the contract zone is the former with probability p and the latter with probability $1-p$.

Recall from the discussion in section on "conventional arbitration that the low contract zone may fail to exist. Given the assumption of a uniform distribution, the contract zone will not exist iff $c_f + c_w < p(v_H - v_L)/2$. This leads to Result 3:

Result 3. In *CA*, as the level of uncertainty regarding the worker's value to the firm rises, *i.e.*, the difference between the high and low values rises, the greater the possibility that the contract zone fails to exist when the firm knows the true value to be low. Furthermore, as the probability that the worker is of a high value rises, the likelihood that there is no contract zone in the low value realization rises.

Intuitively, as the difference between the uninformed worker's expectation of his worth and the firm's knowledge of his worth when the realized state is low rises, agreement becomes increasingly impossible. It is worth noting that the contract zone always exists when the worker has a high value to the firm assuming that the costs of going to arbitration are nonnegative. The condition for the existence of the contract zone in this case is $c_f + c_w > (1-p)(v_L - v_H)/2$. The right hand side is weakly negative as $v_H > v_L$ and $0 < p < 1$.

Now consider FOA. Given uniform distributions over f_H and f_L, Eqs. (1), (2a) and (2b) are rewritten below.

$$E\pi_w = p\left[\beta\frac{\alpha+\beta}{2v_H} + \alpha\left(1 - \frac{\alpha+\beta}{2v_H}\right)\right] + (1-p)\left[\beta\frac{\alpha+\beta}{2v_L} + \alpha\left(1 - \frac{\alpha+\beta}{2v_L}\right)\right] \tag{1′}$$

$$E\pi_{fH} = v_H - \beta\frac{\alpha+\beta}{2v_H} - \alpha\left(1 - \frac{\alpha+\beta}{2v_H}\right) \tag{2a′}$$

$$E\pi_{fL} = v_L - \beta\frac{\alpha+\beta}{2v_L} - \alpha\left(1 - \frac{\alpha+\beta}{2v_L}\right) \quad (2b')$$

Maximization of (1′), (2a′) and (2b′) yields optimal bids in (3).

$$\alpha = \frac{v_H v_L}{(1-p)v_H + p v_L} \quad (3a)$$

$$\beta_H = \beta_L = 0 \quad (3b)$$

Note that if $v_H = v_L$, then there is no uncertainty regarding the amount of money available and the worker should ask for the upper bound of the distribution. Since the lower bound is known, the firm will always bid zero regardless of where the upper bound lies. This is consistent with bidding behavior found in Farmer and Pecorino (1998) and Deck and Farmer (2003).

From the values found in (3), both players can substitute these expected bids into equations (1′), (2a′) and (2b′) in order to determine what they can both expect in FOA. Given those values, the contract zones in Result 2′ can be determined.

Result 2′. Depending upon the realization of v, the contract zone is either $[\alpha/2-c_w,\ \alpha(1-\alpha/2v_H)+c_f]$ or $[\alpha/2-c_w,\ \alpha(1-\alpha/(2v_L))+c_f]$ where α is found in Eq. (3a). The firm knows the truth and the worker knows only the probability of each contract zone.

As in CA, under FOA the contract zone may fail to exist if the worker's value to the firm is low, but the contract zone will exist if the worker has a high value. Also, the expected size of the contract zone is the sum of the arbitration costs $c_w + c_f$, as in CA.

Comparison of Conventional Arbitration and Final Offer Arbitration

This section compares the location and existence of the contract zone under the two dispute resolution techniques.

Result 4. When the arbiter's draw is uniformly distributed over $[0,\ v_H]$ or $[0,\ v_L]$, $v_H \neq v_L$, with probabilities p and $1-p$ respectively, the upper and lower endpoints of the contract zone under FOA are lower than when CA is used for all $p \in (0,\ 1)$.

Result 5. The magnitude of the shift in the endpoints of the contract zone differs for each endpoint.

These results follow directly from a comparison of the contract zone boundaries given in Results 1′ and 2′. One implication of Results 4 and 5 is that the width of the contract zone can differ across arbitration mechanisms for a given realization of the worker's value. This is not the case when neither party knows the true value of the worker (see Deck & Farmer, 2003) or when both parties know the true value of the worker. As stated previously, depending on the variance, the upper bound of the contract zone (determined by what the firm will pay) can fall below the lower bound of the contract zone (determined by what the worker will accept). Thus it is feasible that a contract zone might not exist. However, Results 4 and 5 indicate that when the contract zone exists under both methods of arbitration, final offer favors the firm by generating lower wages. As variance increases, there is a greater downward shift under FOA relative to CA.

In the event that a contract zone does not exist one would expect the parties to always employ arbitration thus creating a social loss as the costs c_w and c_f are incurred. Interestingly, FOA is weakly more efficient than CA as is made explicit in the following result.

Result 6. If the contract zone exists under CA it also exists under FOA.

Proof. The contract zone for CA fails iff $c_f+c_w<p(v_H-v_L)/2$. The contract zone for FOA fails iff $c_f+c_w<\alpha/2-\alpha(1-\alpha/2v_L)$. Substitution for α from Eq. (3a) and rearranging reveals that $\alpha/2-\alpha(1-\alpha/2v_L)<p(v_H-v_L)/2$.

Beyond some level of variance, the contract zone under CA disappears if nature draws v_L. However, in FOA it is possible that the contract zone can actually go out of existence at one variance level and then come back into existence as variance continues to increase. This non-linearity is evidenced by the condition under which the contract zone disappears in FOA, which can be written as $\frac{\alpha^2}{2v_L}-\frac{\alpha}{2}>c_w+c_f$.

Result 3′. In FOA, as with *CA*, as the probability that the worker is of a high value rises, the likelihood that there is no contract zone in the low value realization rises. But, unlike in *CA*, as the level of uncertainty regarding the worker's value to the firm rises, *i.e.*, the difference between the high and low values rises as the variance between the high and low increases, there exists a non-monotonic relationship between this variance and the existence of the contract zone.

Summary of Theoretical Results

(1) In both CA and in FOA, the worker knows the expected position and width the of the contract zone while the firm knows the true contract zone. In both cases the expected width is the sum of the arbitration costs.
(2) FOA produces a contract zone whose endpoints are shifted to the left relative to CA.
(3) In both CA and FOA, it is possible for the contract zone to fail to exist when the worker's value to the firm is low.
(4) In CA as the variance between high and low value workers rises, the contract zone is less likely to exist. However, in FOA, when the worker's value is low the contract zone may not exist at one level of variance, but still exist at a higher level of variance.
(5) If the contract zone exists in under CA, it also exists under FOA. This would suggest theoretically that at least in these cases, FOA should produce greater settlement.
(6) In both CA and FOA, the likelihood that the contract zone does not exist when the worker's value is low increases in the ex ante probability that a worker's value is high.

Note that these theoretical results do not serve the sole purpose of providing a set of testable hypotheses that we will systematically examine. Rather, they identify the relationship between placement and existence of the contract zones as the variance and arbitration method vary.

BEHAVIORAL EXAMINATION

Having identified the theoretical prosperities of CA and FOA with asymmetric uncertainty, we now present a complimentary analysis of the behavioral properties of these mechanisms. First we note that arbitration (as well as mediation and litigation) is a special case of the more general problem of bargaining with outside options. From previous work in arbitration specifically (see Ashenfelter et al., 1992; Deck & Farmer, 2003; Dickinson, 2004, 2006; and Pecorino & Van Boening, 2001, 2004) as well as bargaining experiments more generally (see, for example, Binmore, Morgan, Shaked, & Sutton, 1991 and Kahn & Murnighan, 1993; Binmore, McCarthy, Ponti, Samuelson & Shaked, 2002) we know that behavior in the lab often differs from the theoretical predictions.[14] Depending upon the specific situation, these predictive failures could be due to fairness concerns, risk attitudes, or disputant optimism.[15] However, there is as yet no definitive agreement on how to incorporate such concerns into a theoretical model. This is not to say that the theoretical models lack predictive power, but rather that they are not complete. Hence, our experiments help to provide a more complete picture of arbitration with asymmetric information.[16] The summary results of the previous section guide our experimental investigation. They allow us to identify parameters that will isolate specific changes in order to determine which factors (contract zone, arbitration method or variance independent of its effect on the contract zone) influence bidding behavior and ultimately settlement.

Experimental Design

While the model developed in the previous section identifies conditions under which complete agreement should be reached, the model cannot identify exactly how the surplus will be divided within the contract zone when there is agreement. To address this issue and to establish behavioral regularities in such situations, we turn to the laboratory. Unlike naturally occurring field data, in the laboratory the potential value of the worker, the realized value of the worker, the arbiter's preferences, the arbitration method, the cost of going to arbitration and each party's knowledge of these items can be systematically manipulated.

By altering parameter values such that the contract zone changes position and size allows us to observe bidding behavior as these changes take place, thereby evaluating the predictive power of the summary results. To explore bargaining with this type of informational asymmetry a $2 \times 2 \times 2$ experimental design is used. The first dimension is the arbitration method manipulated across subjects; CA vs. FOA. The second dimension is the variance in the potential value of the worker manipulated within subjects; high variance versus low variance. The third dimension is the realization of the surplus manipulated within subjects; high value or low value. In all cases the worker's expected value from employment is 100, the probability of the worker being of high value is $p = 0.5$, and the costs to each party for going to arbitration is $c_w = c_f = 15$. In the high variance condition $v_H = 180$ and $v_L = 20$ while in the low variance condition $v_H = 130$ and $v_L = 70$.

The motivation for these specific parameter choices are as follows. First, as shown in summary Result 3, Results 4 and Results5 we know that with CA, as the variance increases the contract zone becomes smaller and can even cease to exist when the value of a worker is low; the parameter values we have chosen cause the contract zone to disappear in the low realization of the high variance case. However, as summary Result 4 indicates, the contract zone under FOA exhibits non-monotonicity with regard to variance changes. The parameter values we have chosen generate contract zones with the same width in FOA in both the low and high variance condition.[17] Further, these parameters generate a contract zone with this same width in CA under the low variance condition. Of course, the increase in variance pushes the contract zone in FOA down to lower wage levels (see summary Result 2). Thus, in addition to allowing us to examine behavior with respect to the various positions and potential lack of existence of the contract zone, these parameter choices allow for comparisons between arbitration methods when the contract zones are similar and between variances under FOA when the width of the contract zone is unchanged. In other words, we are able to control for contract zone width and therefore determine whether behavioral responses are due to specifically to the arbitration method or perceptions regarding variance or whether behavior is driven by movements of the contract zone itself. Ultimately, it is not the contract zone itself that is of interest but how this translates into final agreement rates; it is this criteria that is used in order to compare the performance of CA vs. FOA. By manipulating in a controlled fashion that method of arbitration and contract zone width, we are able to ascertain which factors influence the ability to reach agreement. Table 1 gives the contract zone for each cell in the $2 \times 2 \times 2$ design.

Another primary rationale for the parameter choices is due to previous experimental studies of bargaining. Many researchers have found that subjects tend to focus on the equal split outcome. This has lead researchers conducting arbitration experiments to institute procedures designed to avoid

Table 1. Endpoints of the Contract Zones.

	High Variance Environment		Low Variance Environment	
	$v_L = \$20$ and $v_H = \$180$		$v_L = \$70$ and $v_H = \$130$	
	Realization is v_L	Realization is v_H	Realization is v_L	Realization is v_H
Conventional arbitration	[35, 25]	[35, 105]	[35, 50]	[35, 80]
Final offer arbitration	[3, 18]	[3, 47]	[31, 46]	[31, 74]

Note: Only integer wages were permitted. These contract zones give the integer wages, which both parties would prefer to arbitration. The upper endpoint is the maximum wage a firm would be willing to pay and the lower endpoint is the minimum wage a worker would be willing to accept. The discreteness of the wage proposals explains why the expected contract zone width is not always exactly $30 = c_f + c_w$.

this outcome, see Dickinson (2004) and Ashenfelter et al. (1992). Under both environments the wage that splits the expected surplus is $w = 50$. While such a wage results in positive profits to both parties in the low variance condition, this wage would generate a loss to firms in the high variance condition who are bargaining with a low valued worker. As previous experimental research suggests that people exhibit loss aversion, the equal split outcome is unattractive to firms in the high variance environment.

The last advantage of choosing these parameters is that they provide similarity to previous work. Specifically, Deck and Farmer (2003) explore the impact of arbitration when neither party knows the true surplus to be divided. Thus, these parameters afford insight on the impact of moving from two sided uncertainty to one sided uncertainty in arbitration.

Each laboratory session involved 8 subjects, four were assigned the role of a firm and four were assigned the role of a worker. In the experiments no references were made to firms, workers, wages, or arbitration. Subjects retained their roles throughout the experiment. Upon entering the laboratory each subject read the computerized directions and then completed a comprehension quiz.[18] Once every subjected had completed the quiz and had an opportunity to ask questions, the experiment began. Subjects bargained for a total of 30 periods. Each period the subject was randomly matched with someone in the opposite role. The first half of the periods were in one variance condition and the second half were in the other variance condition. To control for sequence effects, the order was blocked so that in half of the sessions subjects first experienced the high variance condition and in the other sessions subject's first experienced the low variance condition. For consistency, one sequence of 30 v_H and v_L realizations were drawn and used in all sessions. In each session subjects resolved disputes with one arbitration method. There were four CA and four FOA sessions. Table 2 provides the details for each experimental session.

The structure of the experiment was that specified in steps 0–4 of section seen in "theoretical model" and in Fig. 1. For initial wage proposals, each party privately submitted an amount that that was simultaneously an offer to the other party and a minimum acceptance threshold, thus eliminating bargaining power. The subject screens showed two parallel lines from 0 to v_H and v_L. On the firm's screen the realized value of employment was identified, but not for the workers. Subjects submitted a proposal by first moving a slider to the desired division of the surplus and then clicking a button. Wage proposals were restricted to be between 0 and v_H. The decision screen for a worker and firm is shown in Fig. 2.[19] If the proposals were compatible, $a<b$, then both parties were informed of the wage and the

Table 2. Details of Experimental Sessions.

Session		1	2	3	4	5	6	7	8
Arbitration Method		CA	CA	CA	FOA	FOA	FOA	FOA	FOA
Periods 1–15	$v_H =$	130	130	180	180	130	130	180	180
	$v_L =$	70	70	20	20	70	70	20	20
Periods 16–30	$v_H =$	180	180	130	130	180	180	130	130
	$v_H =$	20	20	70	70	20	20	70	70
Cost of arbitration, c =		15	15	15	15	15	15	15	15
# workers		4	4	4	4	4	4	4	4
# firms		4	4	4	4	4	4	4	4
Exchange Rate 100 $EXP =		$US 1	$US 1	$US 1	$US 1	$US 1	$US 1	$US 1	$US 1

Note: The session numbers are for expositional purposes and do not reflect the order in which the sessions were run.

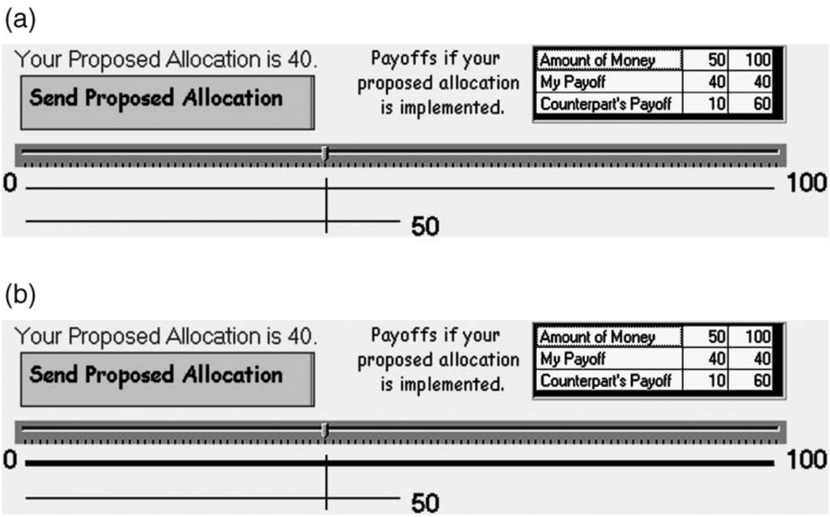


Fig. 2. Wage Proposal Screen. Wage Proposal Decision from (a) the Perspective of the Worker and (b) the Prospective of the Firm.

profits to each party. If the proposals were incompatible, the appropriate form of arbitration was used to determine the wage. In CA, subjects were informed of the randomly determined wage and the resulting profits. Under FOA, after observing the first round proposals each party privately typed a proposal for the arbiter to consider. This was on a separate part of the screen from the initial wage proposal. Once both parties had submitted a second round proposal the resulting wage and profits were announced.

The parameters are in terms of experimental dollars. At the conclusion of the experiment subjects' payoffs were calculated as 100 Exp$ = 1 US$. The exchange rate was announced prior to the start of the experiment. Each session lasted approximately one hour and the average salient payoff was US$ 11.98. Subjects, who are undergraduates were recruited from classes at the University of Arkansas, also received a $5.00 show up fee.

EXPERIMENTAL RESULTS

We present the results as a series of findings on agreement rates, the level of wages, and finally profits. To control for learning effects, we only report behavior from the last 10 periods of each environment for each session.[20] For ease of exposition we introduce the following notation. A treatment is referred to as either CA or FOA depending on the arbitration method. High variance conditions are indicated by a V superscript while low variance conditions are indicated by a V subscript. Similarly, a Π superscript indicates that there is a relatively large surplus to be divided due, and a relatively small surplus is indicated by a Π subscript.

Observed agreement rates for the 80 bargaining pairs in each condition and arbitration method are presented in Table 3. The fact that agreement is not 100% when the contract zone exists has been observed in all previous experimental studies. In examining summary Result 3, we have now found the not too surprising result that agreement is not 0% when the contract zone does not exist.[21] In fact we observe just over 25% agreement in the situation where the contract zone does not exist (CA^{V}_{Π}). This is nominally more agreement than in FOA^{V}_{Π} (17.5%) where the contract zone does exist.

Table 3. Aggregate Agreement Rates.

Treatment	$CA^{V\Pi}$	$FOA^{V\Pi}$	CA^{V}_{Π}	FOA^{V}_{Π}	CA^{Π}_{V}	FOA^{Π}_{V}	$CA_{V\Pi}$	$FOA_{V\Pi}$
Agreement Rate	0.4750	0.2875	0.2625	0.1750	0.4750	0.4500	0.4125	0.2625
Contract Zone	[35,105]	[3,47]	—	[3,18]	[35,80]	[31,74]	[35,50]	[31,46]

Note: The contract zones reflect the fact that subjects could only enter integer wage proposals. A V or Π superscripts denotes a high level of variance or surplus, respectively. A subscript denotes a low level.

This ordering in agreement rates between CA_{Π}^{V} and FOA_{Π}^{V} is the same as has been observed by previous experimental studies. The first finding shows that our data are consistent with these previous studies in this respect.

Finding 1. Aggregate agreement rates are lower under FOA than under CA.

Support. Fig. 3 provides the qualitative support for this finding. This figure shows agreement rates by session for each of the four environments. Within each of the environments the solid diamond FOA markers tend to be below the square outline CA markers. The quantitative support is given by a Wilcoxon Rank Sum test of Ho: no arbitration mechanism effect vs. Ha: FOA leads to lower agreement rates than CA. As the agreement rates are not independent across conditions within a session, the observational units for this test are the eight session agreement rates. With a test statistic of 24, the null hypothesis is rejected at the 10% significance level.

Note that theory does not offer a prediction regarding this ranking except that settlement is less likely when the contract zone fails to exist at all. However, this finding holds true even when the contract zone fails to exist in CA. In other words, even when CA offers no contract zone, it still generates greater settlement. This provides stronger evidence in support of previous literature that CA as a mechanism produces more agreement than does FOA. On the basis of the data graphed in Fig. 3, when considering CA only, one can see that agreement rates are lower in the case where the contract

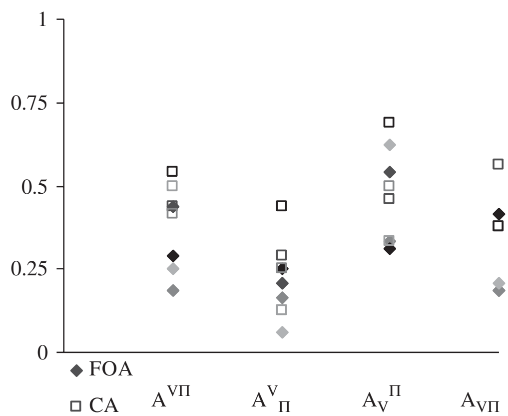


Fig. 3. Agreement Rates by Session and Condition.

zone does not exist than where it does. This begs the question of how agreement rates are related to the size of the contract zone independent of the arbitration mechanism. The parameter values we chose allow us to test this effect. Consistent with previous empirical work by Currie and McConnell (1991) and Babcock et al. (1995), we find that a larger contract zone leads to greater agreement.

Finding 2. Agreement rates are positively correlated with the width of the contract zone for both CA and FOA, conditional upon the arbitration mechanism being held constant.

Support: Given the number of sessions and the number of conditions, standard non-parametric tests are not available to test Ho: contract zone width and agreement rates are uncorrelated vs. Ha: there is a positive relation. Thus we rely on the following non-parametric procedure. First, for each session the conditions are ordered based upon observed agreement rates. The correlation between the observed rankings and the theoretical rankings based upon the width of the contract zones is then calculated. For CA this correlation is 0.6 and for FOA this correlation is 0.623. On the basis of 40,000 randomly generated rankings[22] of four conditions in four sessions, the probability of observing a correlation at least as large as 0.6 is 0.013 for CA and the probability of observing a correlation at least as large as 0.623 is 0.015 for FOA.

Note again that theory does not predict how settlement rates vary with contract zone width. A greater contract zone width permits more settlement possibilities, but a smaller zone may provide a focal point at which to settle. However, as with previous work, settlement does not always take place and as mentioned following summary results, one goal is to determine how the contract zone's position and size contributes to settlement independent of the arbitration mechanism. Under the parameters chosen for the low variance environment, the width of the contract zone does not depend on the arbitration mechanism.[23] As Finding 2 establishes that agreement rates and contract zones are correlated, we next ask if agreement rates are similar across mechanisms when the width of the contract zone is unchanged. The answer is affirmative.

Finding 3. When changing the arbitration mechanism does not change the width of the contract zone, agreement rates are similar in FOA and CA.

Support: Only in the low variance environment does the arbitration mechanism not change the size of the contract zone. However, the width of the contract zone is dependent on the realized pie. To be succinct, we conducted similar analysis to that reported in Finding 1, using agreement rates by session in the appropriate condition as the unit of observation. Based on the two Wilcoxon Rank Sum tests the null hypotheses of no treatment effect cannot be rejected at the 90% confidence level for either the large pie or the small pie condition.

Note that Finding 1 indicates that CA performs better in terms of settlement rates, but Finding 3 suggests that it is not the result of contract zone width. Further experiments use a controlled environment to test for the impact of the variance and the resulting positioning of the contract as a possible explanation for the differences between CA and FOA. Thus, the analysis now turns to an examination of how movements in the contract zone and changes in arbitration mechanism influence the variance affect subjects' choices of proposals. Fig. 4 shows the distribution of wage proposals by workers and firms. As the workers do not know the surplus to be divided, their responses are aggregated along this dimension.

The most striking feature in this figure is the bimodal distribution of proposals by workers (the solid line) under CA in the high variance condition, the situation in which the contract zone may not exist. In each of the other three cases, worker proposals are well behaved. The central tendencies are similar between FOA and CA in the low variance condition. Interestingly, switching from the low variance to the high variance conditions leads to higher wage proposals by workers in FOA even though the contract zones shift to the left.

To quantitatively determine the effect that the variance and arbitration mechanisms have on worker wage proposals we employ linear mixed effects models.[24] This repeated measures model allows each treatment to have a fixed effect while allowing a random effect for each session and each subject within the session.

Finding 4. Changes in the variance generate the following bidding behavior. These directly test summary results 1 and 2.

(1) In the low variance condition, the arbitration mechanism does not affect wage proposals.
(2) The variance in potential values does not effect worker proposals under CA. However, under FOA workers ask for higher wages in the high variance condition even though the contract zone shifts downward.

Support: This finding is based upon the mixed effects estimation results presented in Table 4. The baseline condition is FOA in the low variance

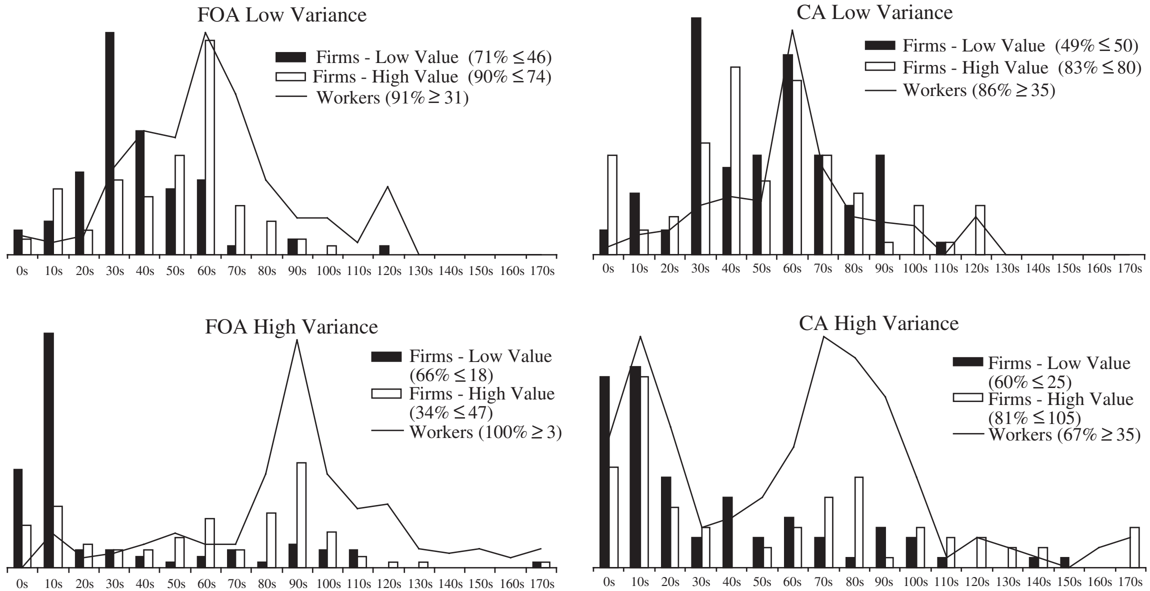


Fig. 4. Distribution of Wage Proposals. The parenthetical text gives the percentage of proposals that were consistent with the relevant contract zone endpoint.

Table 4. Mixed Effects Model of Worker Proposals.

$\text{WProposal}_{ijt} = \beta_0 + \beta_1 \text{ CA}_j + \beta_2 \text{ HiVar}_{jt} + \beta_3 \text{ CA}_j \times \text{HiVar}_{jt} + \beta_4 \text{ Order}_j + \varepsilon_j + \varepsilon_{ij} + \varepsilon_{ijt}$				
Parameter	Estimate	df	*t*-statistic	*p*-value
β_0	62.44	606	10.818	<0.001***
β_1	0.53	5	0.077	0.9413
β_2	29.11	606	10.951	<0.001***
β_3	−33.19	606	−7.977	<0.001***
β_4	2.97	5	0.456	0.668

Note: WProposal_{ijt} is the wage proposal by worker *i* in session *j* during period *t*. CA and HiVar are indicator functions that take on the value one if the observation is from the conventional arbitration or high variance treatment, respectively. To control for an order effect, the indicator function Order is included. Order takes a value of one if the observation is from a session that first experienced the high variance condition and is 0 otherwise.
***Significance at 1% level.

environment. The similarity between wage proposals between CA and FOA in this environment is based on the failure to reject Ho: $\beta_2 = 0$. The impact of the variance on proposals in FOA is demonstrated by the rejection of Ho: $\beta_2 = 0$ at any standard level. The lack of a variance effect for workers under CA is evidenced by the failure to reject Ho: $\beta_2 + \beta_3 = 0$. The failure to reject Ho: $\beta_4 = 0$ indicates the order in which the variance conditions were introduced did not effect wage proposals.

Unlike the workers who only knew the variance of potential values, firms made proposals under complete information as to what was to be divided. From Fig. 4 it appears that firms make lower proposals under FOA than under CA when the value of a worker is low (the dark bars are shifted left). However, this shift is not significant as reported in Finding 5.

Finding 5. Firms propose similar wages under FOA and CA when there is a small value to be divided. This is despite the fact that the contract zone under FOA favors the firm. (summary result 2) However, firms offer a lower (but still positive) percentage of each additional dollar to workers under CA than under FOA. In other words, firms are not responding to fact that the contract zone has shifted in their favor.

Support: This finding is based upon the mixed effects estimation results presented in Table 5. The finding that at low worker values the mechanisms generate the same results is based on the failure to reject Ho: $\lambda_1 = 0$. On the basis of the estimation results, firms under FOA offer a worker \$0.20 for every \$1 the worker is worth. This percentage is significantly different from

Table 5. Mixed Effects Model of Firm Proposals.

$$\text{FProposal}_{ijt} = \lambda_0 + \lambda_1\ \text{CA}_j + \lambda_2\ \text{Money}_{jt} + \lambda_3\ \text{CA}_j \times \text{Money}_{jt} + \lambda_4\ \text{Order}_j + \varepsilon_j + \varepsilon_{ij} + \varepsilon_{ijt}$$

Parameter	Estimate	df	*t*-statistic	*p*-value
λ_0	23.42	606	2.982	0.003***
λ_1	13.23	5	1.412	0.217
λ_2	0.20	606	7.418	<0.001***
λ_3	−0.09	606	−2.270	0.024
λ_4	4.24	5	0.497	0.640

Note: FProposal_{ijt} is the wage proposal by firm *i* in session *j* during period *t*. CA is an indicator function that takes on the value one if the observation is from the conventional arbitration. Money is the actual value of employment which is known to the firm. To control for an order effect, the indicator function Order is included. Order takes a value of one if the observation is from a session that first experienced the high variance condition and is 0 otherwise.
***Significance at 1% level.

zero as λ_2 is significant at standard levels. The difference in offer rates due to the arbitration method is evidenced by the significance of λ_3. However, firms under CA still offer a positive percentage to workers, i.e., one rejects Ho: $\lambda_2 + \lambda_3 = 0$. Again, no environment order effect was observed ($\lambda_4 = 0$).

While FOA generates more disputes resulting in a loss of efficiency, FOA leads to higher wage proposals by firms and in the high variance case for workers as well. So even though self negotiated wages maybe higher conditional on settlement in FOA, settlement is less likely and hence firms are more likely to pay arbitration costs and have the wage determined through arbitration. Thus, the impact on profits of these mechanisms is unclear. The next finding examines the impact that arbitration has on realized profits for firms.

Finding 6. Firms profits are statistically the same in FOA and CA when the value of the worker is low. However, firms receive a statistically lower share of each additional dollar under CA.

Support: Similar analysis as given in support of Finding 5 was conducted with firm profit as the dependent variable. This analysis is presented in Table 6. The findings are identical to those of Finding 5. Specifically, the coefficient on CA was not significant, the coefficient on Money was significantly positive, the coefficient on the interaction term was significantly negative, and the sum of the coefficients on Money and the interaction term is statistically positive.

Table 6. Mixed Effects Model of Firm Profits.

$\text{FProfit}_{ijt} = \gamma_0 + \gamma_1\ \text{CA}_j + \gamma_2\ \text{Money}_{jt} + \gamma_3\ \text{CA}_j \times \text{Money}_{jt} + \gamma_4\ \text{Order}_j + \varepsilon_j + \varepsilon_{ij} + \varepsilon_{ijt}$				
Parameter	Estimate	df	*t*-statistic	*p*-value
γ_0	−38.48	606	−5.877	<0.001***
γ_1	13.00	5	1.644	0.161
γ_2	0.74	606	27.205	<0.001***
γ_3	−0.10	606	−2.529	0.012**
γ_4	−1.82	5	−0.263	0.803

Note: FProfit_{ijt} is the profit to firm *i* in session *j* during period *t*. CA is an indicator function that takes on the value one if the observation is from the conventional arbitration. Money is the actual value of employment, which is known to the firm. To control for an order effect, the indicator function Order is included. Order takes a value of one if the observation is from a session that first experienced the high variance condition and is 0 otherwise.
**Significance at 5% level.
***Significance at 1% level.

Worker profits are not simply the amount of money to be divided minus the firm's profits due to the social savings that occur when agreements are reached. For example, given that CA leads to more agreement, there is more profit under that mechanism. However, when analysis similar to that presented in support of Finding 6 was conducted for worker profits the nominal results and statistical conclusions were consistent with those reported in Table 6 and are thus omitted here.

To this point our analysis has focused exclusively on pre-arbitration bargaining. Here we briefly discuss the behavior of subject offers during arbitration. Given that CA requires no decisions once in arbitration, this section pertains only to FOA. Given Eqs. (3a) and (3b) the firm should always offer $\beta = 0$ regardless of the environment, while the worker should offer $\alpha = 36$ in the high variance environment and $\alpha = 91$ in the low variance environment. Fig. 5 plots the distribution of offers by environment, conditional on a failure to settle. Even though workers are not informed of the actual amount to be allocated, to allow for the possibility of first round firm offers serving as a signal, the distribution of workers is conditioned on the realization. From this figure several things are apparent. First, workers' offers are well above 36 in the high variance environment and well below 91 in the low variance environment. Second, it does not appear that workers are learning from the firm proposals as the solid and dashed lines are similar. It is also interesting to note that Fig. 5 is distinct from the appropriate panels of Fig. 4, indicating that subjects are not simply selecting the

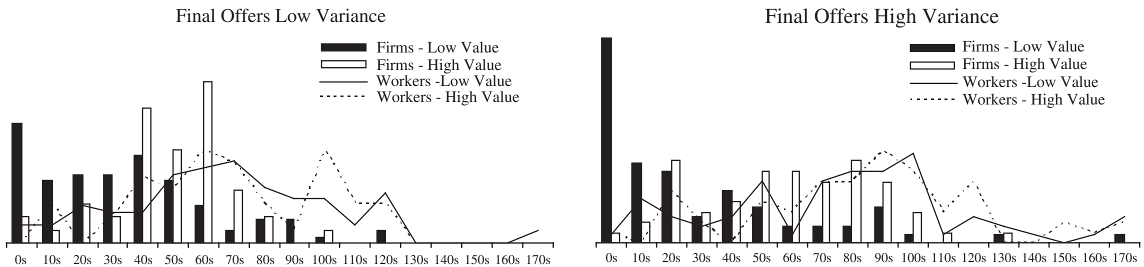


Fig. 5. Offers in FOA.

Table 7. Mixed Effects Estimation of Offers in Final Offer Arbitration.

Dependent Variable		Constant	Order	HiVar	Money	Wask	Fbid
Firm offer	Estimate	17.60	−4.64	–	0.06	0.09	0.41
	Standard error	6.41	7.21	–	0.02	0.04	0.06
	p-value	0.006***	0.586	–	0.003***	0.018**	<0.001***
Worker offer	Estimate	27.05	−5.11	4.53	–	0.41	0.27
	Standard error	8.40	10.47	2.59	–	0.05	0.05
	p-value	0.001***	0.674	0.082*	–	<0.001***	<0.001***
Worker offer	Estimate	29.81	−6.61	3.15	0.03	0.47	–
	Standard error	8.48	10.36	2.74	0.02	0.05	–
	p-value	<0.001***	0.589	0.250	0.140	<0.001	–

*Significance at 10% level.
**Significance at 5% level.
***Significance at 1% level.

same number for both a proposal and an offer. With respect to firms, in all cases firms offer more than the predicted level of 0.[25] However, when the realized value is low, the firms frequently place very low offers. These results are borne out econometrically as well. Again we rely upon linear mixed effects models to control for repeated measures. The variables are as above with the addition of *Wask* and *Fbid* to denote the initial proposals of the worker and the firm respectively. Table 7 provides the estimation results.

That workers and firms do not simply reuse their proposals is evident from the fact that the coefficients on *Wask* and *Fbid*, respectively, are not 1. As in Finding 4*ii* with initial proposals, going from a low variance to a high variance environment actually increases offers even though theory suggests they should fall, in this case the effect is a marginally significant increase of 4.53. Interestingly, both parties use their opponent's proposals when making a final offer. The more generous the proposal of one's counterpart, the more aggressive is one's offer. For every additional dollar of a firm's initial proposal a worker's offers increases by additional 27 cents; whereas for each dollar reduction in a worker's initial proposes a firm's offer decreases by 9 cents. This seems to suggest that proposals convey information about how generous or aggressive a person is. While Finding 5 determines that firms do base initial proposals on the actually amount of money at stake ($\lambda_2 > 0$), workers are not able to identify the realization or at least do not incorporate it into their final offers. As support for the lack of information revelation, we point to the bottom rows of Table 7. In this estimation, *Money* replaces *Fbid* in estimating worker offers. The coefficient on *Money* turns out to be a small and insignificant 0.03. Perhaps this lack of information revelation is due to the money effect being overpowered by the perceived generosity information.

CONCLUSIONS

This paper addresses the situation in which a firm has information concerning the true value of an employee and the employee knows only the distribution of his or her value to the firm. Faced with alternative dispute resolution mechanisms, how does this asymmetry affect the contract zone and settlement patterns? From a theoretical perspective, the contract zone is obviously larger when the firm knows that the worker's value is high, regardless of the form of arbitration to be used. However, we find that in all cases, FOA causes the contract zone to shift to the left relative to CA; i.e., possible settlement values are those with lower wages and, as a result, FOA favors the firm. Moreover, as the variance rises, this shift increases. This finding is consistent with that found by Deck and Farmer (2003) in a model with no asymmetric information.

Where our results deviate from that work is that asymmetric information produces not only a shift in the contract zone, but the size changes as well depending upon the realization of the information. As a result, asymmetric information can result in the disappearance of the contract zone. When the firm knows the workers true value to be low, the worker may have unreasonable expectations given the possibility from their viewpoint that the firm may benefit greatly from their employment. As the variance in the information increases or the probability that a worker is of a high values rises, the gap between the worker's expectations and the true value when the value is low, will increase; as a result, the chance that the contract zone disappears rises. Finally, we find that this disappearance of the contract zone is more likely under CA than under FOA. Moreover, the theoretical findings indicated that the relationship between contract zone size and variance is monotonic for CA but is non-monotonic for FOA. Importantly, therefore, we were able to carefully identify parameters that allowed us to independently analyze the impact of contract zone size and position and the variance independent of its impact on the contract zone and arbitration mechanism.

We find that not only does CA outperform FOA in this structure, a result that is consistent with previous arbitration experiments including Deck and Farmer (2003), but that this result holds true even when the contract zone ceases to exist in CA. We also find that agreement rates are positively correlated with the width of the contract zone. Given that theory suggests either result may occur, this finding sheds light on that debate. Moreover, when we conducted a controlled comparison and held the width of the contract zone constant as the arbitration mechanism changed, we found no difference in settlement rates between CA and FOA. This suggests that the aggregate difference in settlement rates is being driven by the changing location of the contract zone as opposed to its width.

Investigations into the specific wage proposals by both parties shed further insight into the reasons driving settlement failure and the poorer performance of FOA. When the variance between the low and high values is small, the arbitration mechanism does not impact wage proposals by either party. However, as this variance rises, workers raise their wage demands in FOA despite the fact that the contract zone shifts downward. This is the same behavior reported in Deck and Farmer (2003) in the case where neither party knew the distribution of potential values. When workers see the possibility of a higher value of their work (conditional on a constant expected value) they expect greater wages. Interestingly, CA does not have the same effect on workers. Simultaneously, when firms know the true value of the worker to be high, they raise their wage offer, but this increase is less significant in FOA, a reaction that is consistent with the location of the contract zone in FOA relative to CA. In other words, the greater aggregate settlement failure in FOA is likely caused by the behavior of the workers who raise their wage demand when there is an increase in the upper bound of what they are worth. Finally, in this high variance case we find that firm profits are greater under FOA than CA, a result that is consistent with both our theoretical predictions.

NOTES

1. See Farmer and Pecorino (1996) for a survey of the role of asymmetric information on bargaining failure.
2. See, for example, Farmer and Pecorino (2003).
3. See Shavell (1982), Neale and Bazerman (1985), Farber and Bazerman (1989), Babcock and Lowenstein (1997) and Farmer, Pecorino and Stango (2004).
4. See Pecorino and Van Boening (2004) for an analysis of asymmetric information when there is renegotiation and Dickinson (2006) for an analysis of optimism.
5. This model is applicable to a wide variety of situations such as one partner buying the other partner out of a firm or negotiations for the construction of a new development. For simplicity, we follow Deck and Farmer (2003) and couch the discussion as a conflict between a worker and a firm.

6. As is standard in this literature, we assume the parties are monetarily self interested and do not exhibit fairness concerns or other regarding preferences.

7. The models of civil litigation generally assign one player with the power to make an offer, and as a result, the existence of the contract zone merely ensures settlement will occur. An analysis of its location or width is irrelevant when one player is given the power of the offer. In this model, we do not provide this structure, and instead consider the possible ranges for settlement and compare how the arbitration mechanism influences that range.

8. In controlled laboratory experiments, Deck and Farmer (2003) conclude that the location of the contract zone roughly describes aggregate behavior.

9. Note that it is possible for the worker to have private information concerning his or her true ability, work effort or other attributes, but these are not the focus of this paper. While two-sided asymmetric information is realistic, our analysis isolates one aspect of this asymmetry, an approach that is consistent with the law and economics literature concerning asymmetric information and bargaining.

10. It is a standard assumption in the law and economics literature that the arbiter (or judge or jury in a litigation setting) is informed. Usually these models involve asymmetric information on the part of the players, and the arbiter has the information, which is possibly gained through the course of the hearings. While the context of this paper differs, we choose to frame our problem in a consistent fashion in order to draw parallels.

11. Previous work also models the arbiter's decision as a draw from a distribution which may be known by both players (see Farber, 1980) or an asymmetry may exist in which one player knows the correct distribution and the other does not (see Farmer and Pecorino, 1998). Similarly, there exists a large literature on traditional litigation in which it is the judge or jury's decision, which is drawn from a distribution, and it is the distribution which one party may know with greater accuracy than another. (see Bebchuk, 1984 for the seminal model in this literature).

12. The arbiter's decision is modeled similar to Farber (1980). F_i denotes the cumulative density function for f_i.

13. It is the experimental investigation that drives our choice of distributional forms. A uniform distribution can be explained more simply to a layperson than can a normal distribution. The general theoretical findings are robust to the distributional form, see Appendix 1.

14. For a more general discussion of bargaining experiments, the reader is referred to Davis and Holt (1993) and Kagel and Roth (1995).

15. In a simple alternating offer experiment, Binmore et al. (2002) find that that the one's outside offer does influence one's perception of fairness.

16. Kagel, Kim, and Moser (1996) consider simple ultimatum games with asymmetric information. There are several key distinction between our experiments and those of Kagel et al. (1996) Kagel, Kim and Moser (1996). In arbitration, the outside options are unknown during bargaining even for the more informed player and in the case of FOA depended upon subsequent strategic moves. Also, the bargaining structure we employ does not make unique predictions as both players are equally powerful.

17. This is true for continuous proposals but given the integer restriction, the contract zone widths vary slightly.

18. A copy of the directions as well as the comprehension handout is available from the authors upon request.
19. These screen shots are from the directions. So as not to bias the subjects, values for v_H and v_L differed form those used in the experiment.
20. It is not uncommon for subjects to make mistakes in the first periods as they become familiar with the computer interface. The choice of 5 periods is somewhat arbitrary, but it does not significantly impact the conclusions.
21. Risk loving preferences could cause a contract zone to not exists when it would for risk neutral parties while risk averse preferences may cause the zone to exist when it would not for risk neutral preferences. Similarly fairness concerns may cause a contract zone to exist when it would not for self interested parties, while a preference to win in a relative payoff sense may cause a contract zone not to exist when it would for self interested parties. Disputant optimism may also impact the existence of a contract zone.
22. One could also have found the exact distribution of the correlation value. However, there are $(4!)^4$ possible orderings that could be observed.
23. There are slight (1 or 2 unit) differences due to the discrete nature of the experiment.
24. See Longford (1993) for a discussion of this model that is commonly used in experimental sciences.
25. While this could be due to risk aversion, risk aversion cannot explain the observed pattern of worker offers.

ACKNOWLEDGMENT

The authors thank Nathan Espinosa and Chenin Vadalma for their research assistance. The data collected for this study are available upon request. Please direct correspondence to Cary Deck at cdeck@walton.uark.edu

REFERENCES

Ashenfelter, O., Currie, J., Farber, H. S., & Spiegel, M. (1992). An experimental comparison of dispute rates in alternative arbitration systems. *Econometrica*, *60*(6), 1407–1433.

Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives*, *11*(1), 109–126.

Babcock, L., Lowenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased judgments of fairness in bargaining. *American Economic Review*, *85*(5), 1337–1343.

Bebchuk, L. A. (1984). Litigation and settlement under imperfect information. *RAND Journal of Economics*, *15*(Autumn), 404–415.

Bernstein, L. (1993). Understanding the limits of court-connected ADR: A critique of federal court-annexed arbitration programs. *University of Pennsylvania Law Review*, *141*(6), 2169–2259.

Binmore, K., McCarthy, J., Ponti, G., Samuelson, L., & Shaked, A. (2002). A backward induction experiment. *Journal of Economic Theory*, *104*, 48–88.

Binmore, K., Morgan, P., Shaked, A., & Sutton, J. (1991). Do people exploit their bargaining power? An experimental study? *Games and Economic Behavior*, *3*(3), 295–322.

Brams, S., & Merril, S. (1986). Binding versus final offer arbitration: A combination is best. *Management Science*, *32*(10), 1346–1355.

Currie, J., & McConnell, S. (1991). Collective bargaining in the public sector: The effect of legal structure on dispute costs and wages. *American Economic Review*, *81*(4), 693–718.

Davis, D., & Holt, C. (1993). *Experimental economics*. Princeton: Princeton University Press.

Deck, C., & Farmer, A. (2003). Bargaining over an uncertain value: Arbitration mechanisms compared. *Journal of Law, Economics, and Organization*, forthcoming.

Dickinson, D. (2004). A comparison of conventional, final-offer and combined arbitration for dispute resolution. *Industrial Labor Relations Review*, *57*, 288–301.

Dickinson, D. (2006). The chilling effect of optimism: The case of final-offer arbitration. *Journal of Socio-economics*, *35*(1), 17–30.

Farber, H. (1980). An analysis of final offer arbitration. *Journal of Conflict Resolution*, *24*(4), 683–705.

Farber, H. S., & Bazerman, M. H. (1989). Divergent expectations as a cause of disagreement in bargaining. *Quarterly Journal of Economics*, *104*(Feb.), 99–120.

Farmer, A., & Pecorino, P. (1996). Issues of informational asymmetry in legal bargaining. In: D. A. Anderson (Ed.), *Dispute resolution: Bridging the settlement gap* (pp. 79–105). Greenwich, Connecticut: JAI Press, Inc.

Farmer, A., & Pecorino, P. (1998). Bargaining with informative offers: An analysis of final offer arbitration. *Journal of Legal Studies*, *27*(2), 415–432.

Farmer, A., & Pecorino, P. (2003). Does the use of final offer arbitration impede settlement? *The Journal of Law, Economics and Organization*, *19*(1), 64–82.

Farmer, A., Pecorino, P., & Stango, V. (2004). *The causes of bargaining failure: Evidence from major league baseball, Journal of Law and Economics*, *57*(2), 543–568.

Fuller, J. P. (1993). U.S. tax review. *Tax Notes International*, *7*(Oct.), 1041–1042.

Kagel, J., Kim, C., & Moser, D. (1996). Fairness in ultimatum games with asymmetric information and asymmetric power. *Games and Economic Behavior*, *13*, 100–110.

Kagel, J., & Roth, A. (1995). *The handbook of experimental economics*. Princeton: Princeton University Press.

Kahn, L., & Murnighan, J. K. (1993). A general experiment on bargaining in demand games with outside ptions. *American Economic Review*, *83*(5), 1260–1280.

Longford, N. T. (1993). *Random coefficient models*. New York: Oxford University Press.

Neale, M. A., & Bazerman, M. H. (1985). Perspectives for understanding negotiation: Viewing negotiation as a judgmental process. *Journal of Conflict Resolution*, *29*(1), 33–55.

Pecorino, P., & Van Boening, M. (2001). Bargaining and information: An empirical analysis of a multistage arbitration game. *Journal of Labor Economics*, *19*(4), 922–948.

Pecorino, P., & Van Boening, M. (2004). An empirical analysis of bargaining with voluntary transmission of private information. *The Journal of Legal Studies*, *33*, 131–156.

Shavell, S. (1982). Suit, settlement and trial: A theoretical analysis under alternative methods for the allocation of legal costs. *Journal of Legal Studies*, *11*(1), 55–81.

Zeng, D. (2003). An amendment to final offer arbitration. *Mathematical Social Sciences*, *46*(1), 9–19.

THE EMPLOYMENT–PRODUCTIVITY RELATIONSHIP WITH EMPLOYMENT CRITERIA

Sumati Srinivas and Michael Sattinger

ABSTRACT

This paper analyzes labor market responses to productivity shocks when firms set employment criteria on the basis of the likelihood of hiring high or low productivity workers. In response to a positive productivity shock, firms do not raise the criterion as much as the shock, increasing the proportion of low productivity workers among the employed. The observed average productivity may respond negligibly even if employment changes substantially. Interest rate fluctuations can yield an opposite relation between productivity and employment, explaining the weak empirical relationship.

1. INTRODUCTION

This paper proposes an alternative view of the labor market to explain weak and contradictory relations between employment and productivity over the business cycle. Instead of choosing an amount of labor to employ at a particular real wage, firms set an employment criterion to determine which applicants to hire. In contrast to the implications of the standard model with

Aspects of Worker Well-Being
Research in Labor Economics, Volume 26, 447–466

ISSN: 0147-9121/doi:10.1016/S0147-9121(06)26013-8

homogeneous labor, the employment criterion approach yields the following major results.

- In response to a productivity shock, observed average productivity varies less than proportionately with the initial productivity shock, so that large employment fluctuations can be associated with small observed productivity changes.
- Average productivity can be positively or negatively related to employment, depending on whether the fluctuation in the economy is from a productivity shock or a monetary disturbance (affecting the interest rate).

In the Real Business Cycle approach initially developed by Kydland and Prescott (1982), productivity shocks are the source of aggregate fluctuations in employment and economic activity (see the review by Hartley, Hoover, & Salyer, 1998). One question considered in the literature is whether adjustments to productivity shocks would generate employment fluctuations of the magnitudes observed empirically. Hansen (1985) shows that large fluctuations in employment can be generated by the model if labor is assumed to be indivisible (see also Hansen & Wright, 1992). A related question is why the posited relation between productivity shocks and aggregate fluctuations is not reflected in empirical estimates of the correlation between productivity and employment (measured in hours worked). Low correlation between productivity and employment has been explained within the literature by introducing other sources of shocks. Benhabib, Rogerson, and Wright (1991) introduce household production and home production shocks, and Christiano and Eichenbaum (1992) introduce government spending shocks (see also Hansen & Wright, 1992).

This paper proposes an alternative resolution of these issues. Instead of choosing an amount of labor, firms choose an employment criterion. Workers are heterogeneous and have either high or low productivity. Firms observe worker productivity imperfectly and set an employment criterion for the observed worker productivity to determine which workers to employ. In response to a productivity shock that increases worker productivities proportionately, firms raise the employment criterion less than proportionately to the shock and may even lower it. Employment expands along with a decrease in the proportion of high productivity workers among the employed, which counters the increase in average productivity from the productivity shock. The result can be a substantial increase in employment associated with a negligible change in observed average productivity. As a simple, paradigmatic model of the labor market, the analysis developed here abstracts from dynamics and expectations that are important in fully delineated macroeconomic models.

Heterogeneous labor has been introduced previously in the study of business cycle fluctuations (Kydland, 1984; King, Plosser, & Rebelo, 1988; Hansen & Sargent, 1988; Cho & Rogerson, 1988; and Cho, 1995). Prasad (1996) investigates aggregation bias in measurements of productivity and the real wage from ignoring skill heterogeneity in workers. Employment criteria or hiring standards have been considered in labor economics as an alternative margin of adjustment over the business cycle. Gaston (1972) shows that hiring standards vary depending on labor market conditions. Thurow (1975) develops the concept of an employment criterion in developing a theory of job competition to explain statistical discrimination (see also Schlicht, 1981). Lockwood (1991) analyzes a matching model in which firms observe worker productivity imperfectly and use unemployment duration as an employment criterion. Schlicht (2005) develops a model in which firms simultaneously set wage offers and hiring standards, generating results that differ from efficiency wage models. Firm behavior in setting employment criteria in the model developed here derives most directly from Sattinger's analysis of statistical discrimination (Sattinger, 1998), and Srinivas's analysis of labor market compositional effects of productivity shocks (Srinivas, 2001).

Pissarides (1985, 2000, Chapter 6) has considered hiring standards in the context of the equilibrium search and matching model. Following Jovanovic (1979), the productivity outcome of a match between a worker and a job is determined by a random drawing from an urn. This ex post productivity is unrelated to any ex ante characteristics of the worker or job, and all workers are ex ante identical. With a zero profit condition on firm entry, Pissarides derives conditions that determine the reservation productivity for a match to form and the ratio of vacancies to unemployed (2000, p. 156). Pissarides uses the model to explain determinants of shifts in the Beveridge curve. Without ex ante heterogeneous labor, Pissarides' model cannot consider the changes in the composition of labor that generate the weak relationship between productivity and unemployment in this paper. Also, unlike this paper, Pissarides' development does not compare actual productivity shocks with observed productivity shocks, nor does it consider other disturbances that would generate alternative empirical relations between observed productivity and unemployment.

Furthermore, Shimer (2004, 2005; see also Hall, 2005) has argued that the Mortensen–Pissarides search and matching model cannot explain the magnitude of business cycle fluctuations. In the alternative model considered by

Shimer, rigid wages lead firms to adjust to productivity shocks through employment variation rather than wage changes. This paper also assumes rigid wages in the short run. Although it is consistent with large employment fluctuations, the model developed here is more concerned with the relation between observed productivity and unemployment.

Section 2 develops the model, beginning with firm behavior. While the proportion of unemployed workers is exogenous to the individual firm, it is endogenously determined in the market. Incorporating the determination of this proportion into the firm first order condition yields a market condition for equilibrium, permitting comparative statics. Section 3 introduces productivity shocks, which change worker productivities proportionately without changing firm abilities to distinguish between high and low productivity workers. It is shown that the employment criterion increases less than proportionately to a productivity shock, and that average productivity increases by a smaller proportion than the productivity shock. Section 4 considers fluctuations in the interest or discount rate as a source of shock to the labor market, presumably generated by monetary disturbances. Since the benefits of employing workers are distributed over time, a change in the discount rate leads firms to alter their employment criteria. It is shown that an interest rate increase reduces employment while raising average productivity, since the employment criterion becomes stricter. This is the opposite relationship from a productivity shock. Section 5 discusses the conclusions.

2. MODEL

2.1. Firm Determination of Employment Criterion

The labor market is characterized as follows. Identical firms interview workers and decide on the basis of a test or interview whether a worker is likely to have high or low productivity. Firms decide to hire a worker if the expected profit from the worker is positive (or non-zero). It will be shown that firms set an employment criterion and hire any worker with a test score greater than or equal to that employment criterion. Workers have a common quit rate, but different hiring rates depending on whether they have high or low productivity. The two types of workers then have unequal unemployment rates, which determine endogenously the proportion of high productivity workers among the unemployed. This section shows how the employment criterion is determined and examines its existence and uniqueness.

First consider the problem facing the firm. Suppose workers either have high productivity, p_1, or low productivity, p_2, with $p_1 > p_2$. Suppose the firm observes the productivity imperfectly as in the statistical discrimination literature:

$$y_i = p_i + \varepsilon_i \tag{1}$$

where p_i is the productivity of worker i, y_i is the observed productivity for worker i, and ε_i is an independently and identically distributed random error term. Suppose ε_i is distributed normally with mean 0 and variance σ^2, and let $f(\varepsilon)$ and $F(\varepsilon)$ be the probability density function and cumulative distribution function, respectively.

Let w be the wage rate, taken as given in this section (Section 4 will consider the determination of the wage rate). The wage is the same for all workers. Let q be the quit rate, the same for all workers, and let r be the discount rate, the same for all firms. Suppose firms incur a cost of c for all workers hired. By integration, the present discounted value from hiring a worker with productivity p_i is

$$\pi_i = \frac{p_i - w}{q + r} - c, \quad i = 1, 2 \tag{2}$$

It is assumed that $\pi_1 > 0 > \pi_2$. If instead $\pi_1 > \pi_2 > 0$, the firm would hire all workers, and if $0 > \pi_1 > \pi_2$, the firm would hire no workers.

Now consider the likelihood that a worker with observed productivity (or score) y_i has high productivity, p_1. If the worker actually has productivity p_1, then $\varepsilon_i = y_i – p_1$, while if the worker actually has productivity p_2, then $\varepsilon_i = y_i – p_2$. Let μ be the proportion of high productivity workers among the unemployed. (This will be determined endogenously by firm hiring decisions, but a single firm's decision will not affect μ so the firm will take μ as given.) Applying Bayes Rule, the likelihood that a worker with score y_i is high productivity is

$$\frac{\mu f(y_i - p_1)}{\mu f(y_i - p_1) + (1 - \mu) f(y_i - p_2)} \tag{3}$$

The probability that the worker is low productivity is one minus the amount in (3). The expected added profit from hiring worker i is then

$$E(\pi_i) = \frac{\mu f(y_i - p_1)\pi_1 + (1 - \mu) f(y_i - p_2)\pi_2}{\mu f(y_i - p_1) + (1 - \mu) f(y_i - p_2)} \tag{4}$$

The firm should hire the worker whenever $E(\pi_i) \geq 0$.[1]

The following argument shows that there will be a threshold value of y_i, the employment criterion y_0 such that it will be profitable for the firm to hire all workers with $y_i \geq y_0$. A consequence of the normality assumption is that $d \log f(\varepsilon)/d\varepsilon = -2\varepsilon/2\sigma^2$, a decreasing function of ε. Then

$$\frac{d \log\left(f\left(y_i - p_1\right)/f\left(y_i - p_2\right)\right)}{dy_i} = \frac{f'\left(y_i - p_1\right)}{f\left(y_i - p_1\right)} - \frac{f'\left(y_i - p_2\right)}{f\left(y_i - p_2\right)} > 0 \tag{5}$$

so that $f(y_i - p_1)/f(y_i - p_2)$ is an increasing function of y_i. Let y_0 be the value of y_i such that the numerator in (4) is zero:

$$\mu f(y_0 - p_1)\pi_1 + (1-\mu) f(y_0 - p_2)\pi_2 = 0 \tag{6}$$

The value y_0 can be shown to exist.[2] Then if $y_i > y_0$,

$$\frac{f\left(y_i - p_1\right)}{f\left(y_i - p_2\right)} > \frac{f\left(y_0 - p_1\right)}{f\left(y_0 - p_1\right)} \tag{7}$$

so that $E(\pi_i) > 0$. Therefore the firm will hire all workers with $y_i \geq y_0$. The firm's strategy concerning which workers to hire is analogous to the reservation wage property in search theory.

Because of the foregoing result, it is possible to reformulate the firm's problem as follows. Let β be the number of interviews per period for the firm (assumed to be exogenous to the firm) and let C_I be the cost per interview. Then the firm's expected profit per period is

$$E(\pi) = \beta\left(\mu\left[1 - F\left(y_0 - p_1\right)\right]\pi_1 + (1-\mu)\left[1 - F\left(y_0 - p_2\right)\right]\pi_2 - C_I\right) \tag{8}$$

The firm maximizes $E(\pi)$ with respect to y_0, yielding the first order condition in (6). The second order condition, after applying the first order condition and rearranging, is given by (5) and is satisfied because of the normality assumption.

The firm's problem in choosing the employment criterion y_0 can be understood as follows. Any unemployed worker appearing for an interview could be either a high productivity or a low productivity worker. The firm is willing to risk hiring a low productivity worker (and losing money on that worker) if there is a sufficient chance of getting a high productivity worker and making money. If $y_i = y_0$, the addition to profits from the chance of hiring the high productivity worker just balances the loss from the risk of hiring a low productivity worker, and the expected gain in profit from hiring the worker is zero. At any higher value of y_i, the expected profit from hiring a high productivity worker outweighs the risk of loss, and the expected added profit is positive.

2.2. Proportions of the Employed and Unemployed That Are High Productivity

Firms, in choosing the employment criterion y_0, take μ, π_1, and π_2 as given. However, μ, the proportion of the unemployed that are high productivity, depends on the hiring decisions of firms. This section examines how μ is determined.

Suppose workers receive interviews at the rate of θ per period. A worker with productivity p_i will have $y_i < y_0$ in a proportion $F(y_0 - p_i)$ of interviews. Then the proportion of interviews that yield job offers is $1 - F(y_0 - p_i)$. The rate at which an unemployed worker with productivity p_i gets a job is therefore $\theta\,[1 - F(y_0 - p_i)]$. Let u_i be the unemployment rate for workers with productivity p_i. The long run equilibrium level of unemployment will be achieved when the flow of workers from employment to unemployment, $(1 - u_i)q$, equals the flow from unemployment to employment

$$(1 - u_i)q = u_i\theta\left[1 - F\big(y_0 - p_i\big)\right], \quad i = 1, 2 \tag{9}$$

Then

$$u_i = \frac{q}{q + \theta\big[1 - F\big(y_0 - p_i\big)\big]}, \quad i = 1, 2 \tag{10}$$

Let ρ be the proportion of high productivity workers among the population. The proportion of workers unemployed in equilibrium is $\rho\, u_1 + (1 - \rho)u_2$. The proportion of unemployed workers that are high productivity is then

$$\mu = \frac{\rho\, u_1}{\rho\, u_1 + (1 - \rho)u_2} \tag{11}$$

2.3. Market Determination of Employment Criterion

The previous derivations can now be combined to yield the following theorem on the market determination of the employment criterion.

Theorem 1. *In the Employment Criterion Model, the criterion y_0 exists and satisfies the following condition*:

$$\frac{\rho}{1-\rho}\frac{q+\theta\left[1-F\left(y_0-p_2\right)\right]}{q+\theta\left[1-F\left(y_0-p_1\right)\right]}\frac{f\left(y_0-p_1\right)}{f\left(y_0-p_2\right)}=\frac{-\pi_2}{\pi_1} \tag{12}$$

Proof:. Using (11) to substitute for μ in the firm first order condition (6) yields

$$\frac{\rho u_1 f\left(y_0-p_1\right)\pi_1+(1-\rho)u_2 f\left(y_0-p_2\right)\pi_2}{\rho u_1+(1-\rho)u_2}=0 \tag{13}$$

Multiplying by $\rho u_1+(1-\rho)u_2$, substituting for u_1 and u_2 from (10) and rearranging yields (12). When this condition is satisfied, firms' first order conditions are satisfied, and μ is consistent with firm choices of y_0. The value of y_0 satisfying (12) is such that the labor market is in equilibrium.

Next, consider existence and uniqueness of the employment criterion satisfying (12). The strategy of the proof can be demonstrated using Fig. 1. The upward-sloping curve in the figure shows the left-hand side of (12) as a function of y_0.[3] The horizontal line is at the ratio of profits $-\pi_2/\pi_1$, where π_2 is negative. The value of y_0 where the curve reaches $-\pi_2/\pi_1$ is the equilibrium value. If the left-hand side of (12) is monotonically increasing, starting below $-\pi_2/\pi_1$ and going above that value, then existence and

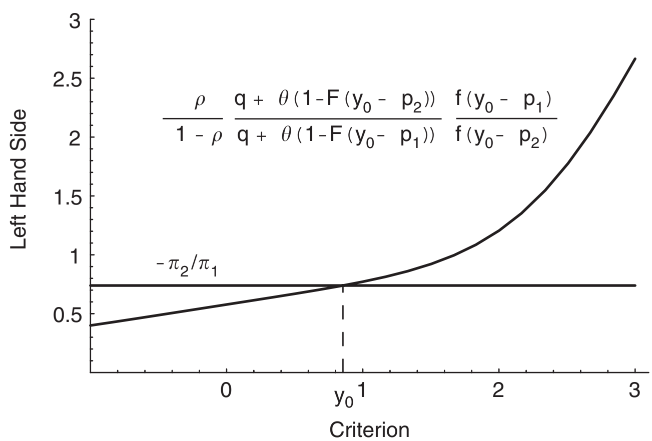


Fig. 1. Market Equilibrium Condition.

uniqueness would follow immediately. While $f(y_0 - p_1)/f(y_0 - p_2)$ increases monotonically because of the assumption of normality,

$$\frac{u_1}{u_2} = \frac{q + \theta\left[1 - F\left(y_0 - p_2\right)\right]}{q + \theta\left[1 - F\left(y_0 - p_1\right)\right]} \tag{14}$$

does not. It is therefore necessary to investigate the individual components of (12) that depend on y_0.

From the normal distribution,

$$\begin{aligned}\frac{f\left(y_0 - p_1\right)}{f\left(y_0 - p_2\right)} &= \frac{\left(1/\sqrt{2\pi}\sigma \mathrm{e}^{-\left(y_0 - p_1\right)^2/\left(2\sigma^2\right)}\right)}{\left(1/\sqrt{2\pi}\sigma \mathrm{e}^{-\left(y_0 - p_2\right)^2/\left(2\sigma^2\right)}\right)} \\ &= \mathrm{Exp}\left(\left(p_1 - p_2\right)\left(2y_0 - p_1 - p_2\right)/\left(2\sigma^2\right)\right)\end{aligned} \tag{15}$$

where $\mathrm{Exp}(x) = \mathrm{e}^x$. Since $p_1 > p_2$, this is an increasing exponential function of y_0, starting at zero for indefinitely small y_0 and increasing indefinitely as y_0 increases indefinitely.

Now consider u_1/u_2. Fig. 2 shows the behavior of the unemployment rates individually and Fig. 3 shows the ratio u_1/u_2. From the functional forms, this ratio first decreases over an interval and then increases. Since the unemployment rates approach each other at arbitrarily low and high values of y_0, the ratio u_1/u_2 starts at one and ends at one as y_0 increases over its range. Then the left-hand side of (12) starts at zero for sufficiently

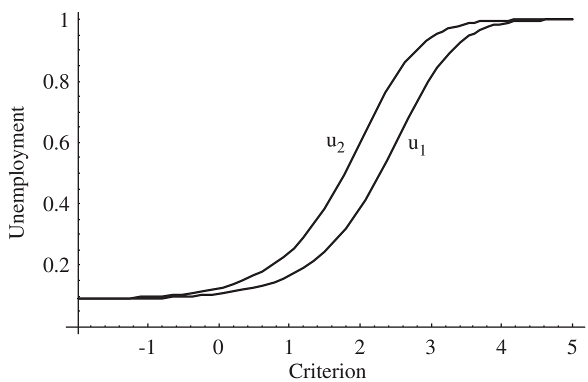


Fig. 2. Unemployment Rates.

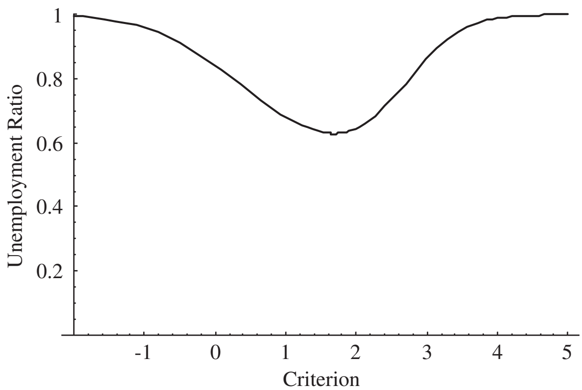


Fig. 3. Ratio of Unemployment Rates, u_1/u_2.

low y_0 (and is less than $-\pi_2/\pi_1$) and eventually goes above $-\pi_2/\pi_1$ for sufficiently large y_0. By continuity, the left-hand side of (12) must equal $-\pi_2/\pi_1$ at some value of y_0, establishing existence. This completes the proof of Theorem 1.

The theorem does not include a statement of uniqueness because of complications. The ratio $f(y_0-p_1)/f(y_0-p_2)$ is monotonically increasing, but u_1/u_2 is not. However, if the solution for y_0 yields sufficiently low unemployment rates, the slope of u_1/u_2 approaches zero so that the product of the two ratios is monotonically increasing in a lower range for y_0. Also, it can be shown that changes in the ratio $f(y_0-p_1)/f(y_0-p_2)$ dominate changes in u_1/u_2, so the left-hand side of (12) is a monotonically increasing function of the employment criterion, y_0. The market equilibrium value of the employment criterion would then be unique.

2.4. Comparative Statics

The condition for the equilibrium employment criterion in (12) yields the following comparative statics results:

Theorem 2. *The employment criterion y_0 will be greater when c, w or θ are greater or when r, ρ, p_1 or p_2 are less.*

Proof. The equilibrium employment criterion occurs when the left-hand side of (12) crosses the horizontal line at the level $-\pi_2/\pi_1$. When the left-hand side of the condition shifts down or $-\pi_2/\pi_1$ goes up, the employment criterion increases. As c or w increase, π_1 decreases and π_2 decreases, so that $(-\pi_2)$, a positive amount, increases. Then the ratio $-\pi_2/\pi_1$ increases, leading to an increase in y_0. Similarly, an increase in the interest rate r raises $-\pi_2/\pi_1$ resulting in an increase in y_0. Declines in either p_1 or p_2 raise $-\pi_2/\pi_1$, leading to a greater value of y_0. The parameters ρ and θ affect the left-hand side of the condition but not $-\pi_2/\pi_1$. If ρ declines, the left-hand side shifts down. If θ increases, the ratio u_1/u_2 declines since $1-F(y_0-p_1) > 1-F(y_0-p_2)$, shifting the left-hand side of the condition down and increasing y_0. This completes the proof.

In general, parameter changes that raise the profitability of a marginal applicant (with criterion equal to y_0) lead firms to risk hiring more low productivity workers (by lowering the criterion y_0) in order to hire more high productivity workers. Comparative static effects of parameter changes are relevant to the analysis of labor market responses to cyclical conditions, which will be examined in Section 4.

3. THE EFFECTS OF PRODUCTIVITY SHOCKS

3.1. Assumptions

At this point it is possible to introduce productivity shocks into the model. In the RBC models, positive productivity shocks raise the demand for labor. Then at a constant wage rate, employment increases along with an increase in the average productivity of labor. In the model developed here (with heterogeneous labor, imperfect observation of worker productivity and employment criteria), the net response depends on the adjustment in the employment criterion. It will be shown that the employment criterion adjusts less than the productivity shock, raising the proportion of low productivity workers among the employed and moderating the observed productivity change. The change in average productivity may be negligible in comparison to the productivity shock.

Productivity shocks are assumed to affect all workers' productivities by the same proportion. Then in (1), the productivity of a worker of type i is

$$p_i s, \quad i = 1, 2 \tag{16}$$

where s is the productivity shock and p_1 and p_2 are the productivities of the high and low productivity workers when s equals one. The standard deviation of the error term in (1) also changes in proportion to the shock, so that the shock has no effect on the ability of firms to distinguish between high and low productivity workers. This assumption is fully consistent with the type of productivity shock assumed in RBC models. To incorporate this assumption, write the probability density function for the normally distributed error term in (1) as $f(\varepsilon;\sigma)$, where σ is the standard deviation. In response to a productivity shock s, the probability density function and cumulative distribution function do not change values if ε increases by the same proportion as s. This is achieved if the probability density function and cumulative distribution function are $f(\varepsilon;s\sigma)$ and $F(\varepsilon;s\sigma)$. Then the assumption concerning the productivity shocks yields

$$f(s\varepsilon; s\sigma) = f(\varepsilon; \sigma), \quad F(s\varepsilon; s\sigma) = F(\varepsilon; \sigma) \tag{17}$$

3.2. Effects of Productivity Shocks on the Employment Criterion

With productivity shocks given by (16), the test scores in (1) will be raised by a positive productivity shock. More of both types of workers will have scores that exceed the existing employment criterion, y_0. However, the employment criterion will also adjust, as described in the following theorem.

Theorem 3. *In response to a positive productivity shock $s>1$ (starting from $s = 1$), the ratio of the equilibrium employment criterion to the parameter s declines. Unemployment rates of both types of workers decline but if the unemployment rates initially are sufficiently low, the ratio u_1/u_2 increases, and the proportion of high productivity workers among the unemployed, μ, increases.*

Proof. The proof proceeds by considering whether an increase in the employment criterion proportional to s satisfies the market condition (12). Let $\hat{y}_0$ be the initial value of the employment criterion. By construction, $f(\hat{y}_0 s - p_i s;\ s\sigma) = f(\hat{y}_0 - p_i;\ \sigma)$ and $F(\hat{y}_0 s - p_i s;\ s\sigma) = F(\hat{y}_0 - p_i;\ \sigma)$. Then the left-hand side of 12 will have the same value at $\hat{y}_0 s$ after the shock that it had at $\hat{y}_0$ before the shock. However, $-\pi_2/\pi_1$ will be lower. The profit ratio is given by

$$\frac{-\pi_2}{\pi_1} = \frac{w + c(q + r) - p_2 s}{p_1 s - w - c(q + r)} \tag{18}$$

An increase in s reduces the numerator and raises the denominator, reducing the ratio on the right side of (12). After the positive productivity shock, the new equilibrium value of the employment criterion will thus be less than $\hat{y}_0 s$. The effect of a productivity shock can be understood using Fig. 1. If the horizontal axis is now y_0/s, the curve representing the left-hand side of (12) does not move in response to a productivity shock. Only the horizontal line at $-\pi_2/\pi_1$ is affected by s, and it moves down when s goes up. The new equilibrium market criterion will be less than $\hat{y}_0 s$. If y_s is the equilibrium employment criterion after the shock, then $y_s/s < y_0$ and $F(y_s/s - p_i; \sigma) < F(y_0 - p_i; \sigma)$. The effect on the unemployment rates is therefore the same as a reduction in the employment criterion, holding the productivity shock fixed. The unemployment rates from both types of workers decline, and from Fig. 3, when the unemployment rates are sufficiently low, the reduction in the employment criterion raises u_1/u_2. From 11, it follows that μ also increases. This completes the proof.

3.3. Average Productivity

The most important consequence of Theorem 3 is that the change in the mix of employed will have opposite effects from the productivity shock itself. Although productivity shocks will substantially affect the aggregate levels of employment and unemployment, the observed effect on average productivity will be substantially moderated. While a positive productivity shock will by itself raise average productivity, the increase in the proportion of employed who are low productivity will reduce it. Then substantial fluctuations in employment could be associated with negligible or undetectable productivity changes. This section examines the effects of productivity shocks on average productivity.

The proportion of employed workers who are high productivity is given by

$$v = \frac{\rho(1 - u_1)}{\rho(1 - u_1) + (1 - \rho)(1 - u_2)} \tag{19}$$

Then average productivity, ϕ, is

$$\phi = v p_1 s + (1 - v) p_2 s \tag{20}$$

The average productivity depends both on the productivity shock, s, and on v. In turn, v depends on the ratio of employment rates

$$v = \frac{\rho(1 - u_1)/(1 - u_2)}{\rho(1 - u_1)/(1 - u_2) + (1 - \rho)} \tag{21}$$

The ratio of employment rates can be rewritten as

$$\frac{1-u_1}{1-u_2} = 1 + \frac{1-(u_1/u_2)}{(1/u_2)-1} \tag{22}$$

When $1/u_2$ is substantially greater than one, the ratio of employment rates decreases when the ratio of unemployment rates, u_1/u_2, rises. The relations between the employment criterion y_0 and μ and v are shown in Fig. 4, holding the productivity shock s fixed. As shown, a reduction in the employment criterion raises the proportion of unemployed that are high productivity and reduces the proportion of employed that are high productivity, if the unemployment rates are sufficiently low.

From Theorem 3, a positive productivity shock reduces y_0/s. From (17), the effect of a productivity shock that reduces y_0/s is equivalent to a reduction in y_0, holding s fixed. As a result, v declines in response to a positive productivity shock. The net effect on ϕ in 20 is summarized in the following theorem:

Theorem 4. *In response to a productivity shock $s>1$ (starting from $s = 1$), observed average productivity increases less than proportionately to the productivity shock when unemployment rates are sufficiently small.*

Fig. 5 shows the relationship between average productivity and the productivity shock using the same assumptions as in Fig. 1. As shown, starting at low levels of s, a positive productivity shock reduces average productivity instead of raising it.

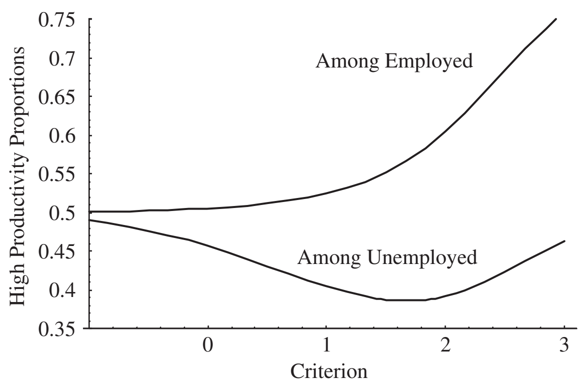


Fig. 4. Proportions of Workers that Are High Productivity.

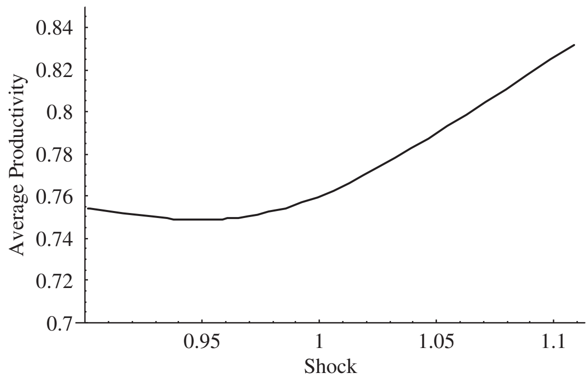


Fig. 5. Response of Average Productivity to Shock.

4. COMPARISONS OF LABOR MARKET DISTURBANCES

This section demonstrates that the relation between employment and average productivity that depends on the source of disturbances to the economy. The previous section developed the consequences of productivity shocks in the employment criterion model, showing that a positive productivity shock generates an increase in employment (or decrease in unemployment) in combination with an observed average productivity change that is smaller than the productivity shock. An alternative disturbance to the economy in the employment criterion model is a fluctuation in the interest or discount rate, presumably caused by a monetary disturbance. Fluctuations in the interest rate will be shown to generate a negative relation between average productivity and employment, just the opposite of the relation generated by productivity shocks.

Effects of interest rate fluctuations are simpler to analyze than productivity shocks. A change in the interest rate has no effects on the left-hand side of the market equilibrium condition in (12). From (18), an increase in the interest rate raises the ratio $-\pi_2/\pi_1$. In Fig. 1, the upward sloping line stays fixed while the horizontal line at the level $-\pi_2/\pi_1$ goes up, so that the market equilibrium employment criterion is higher. The higher employment criterion then raises the unemployment rates of both types of workers and, if the unemployment rates are sufficiently low, lowers the ratio u_1/u_2, lowers the proportion of high productivity workers among the unemployed, μ, and raises the proportion of high productivity workers among the employed, v. With no change in the productivities of high and low productivity workers (i.e., there is no productivity shock), the average productivity ϕ goes up. The decline in employment is then associated with an increase in average productivity. These results are summarized in the following theorem.

Theorem 5. *In the Employment Criterion Model, a higher interest rate yields a higher employment criterion, higher unemployment rates and lower employment. If the unemployment rates are sufficiently small, a higher interest rate yields a higher proportion of high productivity workers among the employed, v,* and higher average productivity.

The possibility of opposite relationships between employment and average productivity, depending on the source of the disturbance, explains weak or contradictory evidence of the cyclicality of productivity. The possibility of opposite relationships is summarized in the following theorem.

Theorem 6. *In the Employment Criterion Model (with the wage rate fixed), an increase in employment can occur with an increase or a decrease in average productivity depending on whether the source of the disturbance is a productivity shock or a fluctuation in the interest rate.*

Fig. 6 shows the two different relations between average productivity and the aggregate employment rate, given by $\rho(1-u_1) + (1-\rho)(1-u_2)$. Consistent with Theorem 6, productivity shocks generate a negative and positive

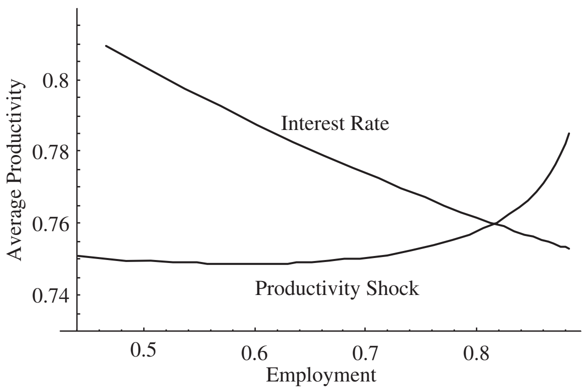


Fig. 6. Alternative Relations Between Employment and Average Productivity.

relation between employment and average productivity, while interest rate fluctuations generate a negative relation.

Another disturbance often discussed in the macroeconomic literature is a change in the real wage. With homogeneous labor, an exogenous increase in the real wage reduces employment by moving firms back up their derived demand curves. The lower employment levels result in higher average labor productivity (since the capital to labor ratio is increased). In the model developed here (with heterogeneous labor and an employment criterion), an increase in the wage has the same effect as an increase in the interest rate. The employment criterion must be higher to satisfy (12), employment decreases, and the average productivity increases. The consequences of a wage increase are therefore the same as in the standard macroeconomic analysis, even though the productivity of a given worker does not decline as more workers are hired. Wage rate fluctuations (holding the productivity shock and interest rate fixed) generate the same relationship between employment and average productivity as interest rate fluctuations.

In the analysis of Sections 2 and 3, the wage has been taken to be exogenously determined. If productivity shocks have only short-term effects, they can be expected to have negligible effects on the wage rate and this assumption is reasonable. If on the other hand a productivity shock has a lasting effect, then eventually the wage would adjust. Wage adjustment to a continuing productivity shock can be determined from the aggregate condition that in the long run, the wage must be such that firm profits are zero. Using this approach, the firm profit expression in (8) can be set equal to zero and solved for the wage w as a function of the employment condition, y_0. The wage rate generated by the resulting function is such that firm profits are zero. This relation can then be combined with the relation generated by the market condition (12) to yield the long run determination of the wage. This is demonstrated in Fig. 7. The downward sloping curve shows combinations of w and y_0 that yield zero firm profits. From (8), the firm faces a cost for every worker interviewed, C_I. If the firm hires fewer workers (because of a higher employment criterion), then the profit on each hired worker must be greater. This in turn requires that the wage be lower, generating the downward sloping zero profit curve in Fig. 7. The upward sloping curve in the figure arises from the market condition (12). At higher wage rates, the profits from the high and low productivity workers are lower, leading firms to choose a higher employment criterion. The intersection of the two curves yields the wage rate and employment criterion consistent with long run equilibrium.

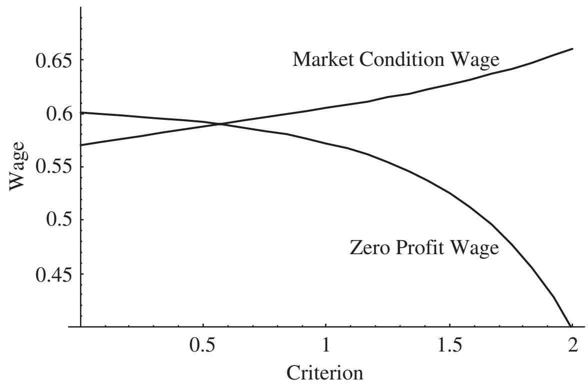


Fig. 7. Determination of Wage Rate.

5. CONCLUSIONS

The phenomenon that drives the conclusions of this paper is that the mix of workers changes in response to productivity shocks. In response to a positive productivity shock, the proportion of employed workers that are low productivity increases. This change in the mix of workers has an effect on the average productivity that is opposite to the productivity shock itself. As a result, there can be a substantial and positive employment response to a positive productivity shock without a large observed increase in an average productivity.

The change in the mix of workers arises because firms in the paradigmatic model face a problem of choosing an employment criterion rather than an amount of labor to hire at a given wage rate. With the employment criterion as the variable subject to firm control, labor market reactions to productivity shocks take the form of adjustments in the employment criterion rather than direct changes in employment and, indirectly, in wage rates. The fluctuations in the employment criterion yield the changes in mix of workers and employment. Wage changes take place through a process that may take longer than the adjustments in firm employment criteria.

The weak relationship between employment and observed average productivity is demonstrated in the case worked out in the paper, in which the positive productivity shock results in a decline in the employment criterion (instead of just an increase that is smaller than the productivity shock, the outcome proven in Theorem 4). Then employment increases both because of the positive productivity shock and because of the reduction in the criterion.

Opposite relations between employment and observed average productivity can also arise because of different sources of disturbances in the model. Productivity shocks can yield a positive but weak relationship between employment and observed average productivity, while disturbances that generate a fluctuation in the interest rate could yield a negative relationship. Estimates of correlations would then be sensitive to time periods included.

The employment criterion model developed here provides a simple means to explain observed relationships among major macroeconomic variables – including employment, wage rates and productivity – that are inconsistent with a simple homogeneous worker view of the labor market. When firms use employment criteria as the margin of adjustment during business cycles, productivity shocks can generate large fluctuations in employment with no strong correlation between observed productivity and employment.

NOTES

1. A worker with $y_i = y_0$ will yield zero expected profit to the firm. Such workers will have measure zero. As a convention, it will be assumed that firms hire them too.
2. Existence can be demonstrated as follows using features of the normal distribution. At low values of y_0, $f(y_0 - p_1)/f(y_0 - p_2)$ will approach zero, so that the left side will be negative (since $\pi_2 > 0$). As y_0 increases indefinitely, the ratio $f(y_0 - p_1)/f(y_0 - p_2)$ will increase indefinitely and the left hand side will be positive since $\pi_1 > 0$. By continuity, there will be a value y_0 such that the left hand side of the equation is zero.
3. The parameters for this figure are $\sigma = 1$, $q = 1$, $w = 0.6$, $r = 0.05$, $p_1 = 1$, $p_2 = 0.5$, $\rho = 0.5$, c $= 0.75$, $C_I = 0.2$ and $\mu = 0.4$. The equilibrium value of y_0 is 0.854 and the value of μ is 0.412.

REFERENCES

Benhabib, J., Rogerson, R., & Wright, R. (1991). Homework in macroeconomics: Household production and aggregate fluctuations. *Journal of Political Economy*, *99*, 1166–1187.

Cho, J.-O. (1995). Ex post heterogeneity and the business cycle. *Journal of Economic Dynamics and Control*, *19*, 533–551.

Cho, F., & Rogerson, R. D. (1988). Family labor supply and aggregate fluctuations. *Journal of Monetary Economics*, *21*, 233–245.

Christiano, L. J., & Eichenbaum, M. (1992). Current real business cycle theories and aggregate labor market fluctuations. *American Economic Review*, *82*, 430–450.

Gaston, R. (1972). Labor market conditions and employer hiring standards. *Industrial Relations*, *11*, 272–278.

Hall, R. E. (2005). Employment fluctuations with equilibrium wage stickiness. *American Economic Review*, *95*(1), 50–65.

Hansen, G. D. (1985). Indivisible labor and the business cycle. *Journal of Monetary Economics*, *16*, 309–327.

Hansen, G. D., & Sargent, T. (1988). Straight time and overtime in equilibrium. *Journal of Monetary Economics*, *21*, 281–308.

Hansen, G. D., & Wright, R. (1992). The labor market in real business cycle theory. *Federal Reserve Bank of Minneapolis Quarterly Review*, *16*, 2–12.

Hartley, J. E., Hoover, K. D., & Salyer, K. D. (Eds) (1998). *Real business cycles: A reader*. New York: Routledge.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, *87*, 972–990.

King, R. G., Plosser, C., & Rebelo, S. (1988). Production, growth and business cycles II; new directions. *Journal of Monetary Economics*, *21*, 309–341.

Kydland, F. E. (1984). Labor force heterogeneity and the business cycle. *Carnegie-Rochester Conference Series on Public Policy*, *21*, 173–208.

Kydland, F. E., & Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica*, *50*, 1345–1369.

Lockwood, B. (1991). Information externalities in the labor market and the duration of unemployment. *Review of Economic Studies*, *58*, 733–753.

Pissarides, C. (1985). Short-run equilibrium dynamics of unemployment, vacancies, and real wages. *American Economic Review*, *75*, 676–690.

Pissarides, C. (2000). *Equilibrium unemployment theory* (Second ed.). Cambridge: MIT Press.

Prasad, E. (1996). Skill heterogeneity and the business cycle. *Canadian Journal of Economics*, *XXIX*, 910–929.

Sattinger, M. (1998). Statistical discrimination with employment criteria. *International Economic Review*, *39*, 205–237.

Schlicht, E. (1981). Training costs and wage differentials in the theory of job competition. *Zeitschrift fur die Gesamte Staatswissenschaft*, *137*, 212–221.

Schlicht, E. (2005). Hiring standards and market clearing. *Metroeconomica*, *56*(2), 263–279.

Shimer, R. (2004). The consequences of rigid wages in search models. *Journal of the European Economic Association*, *2*(2-3), 469–479.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, *95*(1), 25–49.

Srinivas, S. (2001). *Labor market compositional effects of productivity shocks.* Doctoral Dissertation, University at Albany, State University of New York.

Thurow, L. (1975). *Generating inequality: Mechanisms of distribution in the US economy*. New York: Basic Books.