palgrave macmillan

Finance, Research, Education and Growth

Edited by

Luigi Paganetto and Edmund S. Phelps

Finance, Research, Education and Growth

Also by Luigi Paganetto and Edmund S. Phelps

EQUITY, EFFICIENCY AND GROWTH: The Future of the Welfare State (*edited with Mario Baldassarri*)

INTERNATIONAL ECONOMIC INTERDEPENDENCE: Patterns of Trade Balances and Economic Policy Coordination (*edited with Mario Baldassarri*)

PRIVATIZATION PROCESSES IN EASTERN EUROPE: Theoretical Foundations and Empirical Results (*edited with Mario Baldassarri*)

THE 1990s SLUMP: Causes and Cures (*edited with Mario Baldassarri*)

WORLD SAVING, PROSPERITY AND GROWTH (*edited with Mario Baldassarri*)

INTERNATIONAL DIFFERENCES IN GROWTH RATES: Market Globalization and Economic Areas (*edited with Mario Baldassarri*)

Finance, Research, Education and Growth

Edited by

Luigi Paganetto
University of Rome 'Tor Vergata'

and

Edmund S. Phelps
Department of Economics, Columbia University
New York

palgrave
macmillan

First published 2003 by
PALGRAVE
Houndmills, Basingstoke, Hampshire RG21 6XS and
175 Fifth Avenue, New York, N. Y. 10010
Companies and representatives throughout the world

PALGRAVE is the new global academic imprint of St. Martin's Press LLC Scholarly and Reference Division and Palgrave Publishers Ltd (formerly Macmillan Press Ltd).

ISBN 0–333–73278–2

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources.

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Finance, research, education, and growth/edited by Luigi Paganetto and Edmund S. Phelps.
p. cm.
Includes bibliographical references and index.
Contents: Stock market liquidity and economic growth: theory and evidence/Ross Levine – The empirical importance of private ownership for economic growth/Darius Palia, Edmund S. Phelps – Intergenerational transfers and growth/Giancarlo Marini, Pasquale Scaramozzino – Human capital, ideas and economic growth/Charles I. Jones – A rising tide raises all ships: trade and diffusion as conduits of growth/Jonathan Eaton, Samuel Kortum – The role of education and knowledge in endogenous growth/Luigi Paganetto, Pasquale L. Scandizzo – Factors behind the Asian miracle: entrepreneurship, education and finance/Richard R. Nelson, Howard Pack – Technological globalization of national systems of innovation?/Daniele Archibugi, Jonathan Michie – Endogenizing investment in tangible assets, education and new technology/Dale W. Jorgenson – Conclusions/Edmund S. Phelps
ISBN 0–333–73278–2 (cloth)
1. Finance. 2. Economic development. 3. Capitalism. I. Paganetto, Luigi. II. Phelps, Edmund S.

HG173 .F4885 2002
338.9–dc21 2002022414

10 9 8 7 6 5 4 3 2 1
12 11 10 09 08 07 06 05 04 03

Printed and bound in Great Britain by
Antony Rowe Ltd, Chippenham, Wiltshire

Contents

List of Tables

List of Figures

Notes on the Contributors

Daniele Archibugi, University of Cambridge

Jonathan Eaton, Boston University and NBER

Charles I. Jones, Stanford University

Dale W. Jorgenson, Harvard University

Samuel Kortum, Boston University and NBER

Ross Levine, World Bank

Giancarlo Marini, University of Rome 'Tor Vergata'

Jonathan Michie, University of Cambridge

Richard R. Nelson, Columbia University

Howard Pack, University of Pennsylvania

Luigi Paganetto, University of Rome 'Tor Vergata'

Darius Palia, Columbia University

Edmund S. Phelps, McVickar Professor of Political Economy Columbia University, and Visiting Professor, University of Rome 'Tor Vergata'

Pasquale Lucio Scandizzo, University of Rome 'Tor Vergata'

Pasquale Scaramozzino, School of Oriental and African Studies, University of London

Preface

A crucial issue in the era of globalization and internationalization of real and financial markets is whether the relationship between investment and finance is beneficial to growth and development.

Two fundamental facts must be considered to illustrate the scenario in which this interaction takes place: (i) information (*ex ante* on firm's prospects or *ex post* on realized returns) cannot be gathered without costs, so that market equilibria typically occur among agents with heterogeneous information sets (Grossman and Stiglitz, 1980). As a consequence, firms' managers have informational advantages over financial investors and, in financial markets, more-informed investors have informational advantages over less-informed investors and noise traders; and (ii) capital, labour and goods are all free to circulate among countries, but the speed and cost of circulation are very different. In this scenario, the interaction between the financial and real sectors may have both positive and negative effects.

When internal finance is not sufficient, the informational gap between managers and external financiers may generate inefficiencies in terms of information-based cost differentials between external and internal finance, or in terms of constraints to the quantity of external finance available. These two outcomes may lead firms to abandon investment projects that would have been profitable in a context of perfect information.

More specifically, the three fundamental forms of financing inefficiencies analysed by the literature are bank financing inefficiencies, stock market financing inefficiencies and venture capital financing inefficiencies. Analysis of the causes and remedies for these inefficiencies shows that finance may have significant effects on investment and innovation, and that its contribution to growth may be improved if the right normative solutions to offset these inefficiencies are found.

Models explaining bank financing inefficiency show that the investor's informational advantage may cause equilibrium credit rationing (Jaffe and Russel, 1976; Stiglitz and Weiss, 1981; Williamson, 1986). In the most famous of these, Stiglitz–Weiss analyse a case with homogenous firms with size homogeneous projects (ordered according to a mean preserving spread distribution) all needing the same financing amount. In this model, the lending rate affects not only the level of lending demand, but also the level of risk assumed by investors, through an 'adverse selection' (the pool of projects selected under high lending rates has an average higher risk) and a 'moral hazard' mechanism (an increase in the lending rate shifts investors towards riskier investments). For these reasons, investors' solvency probability is inversely related to the interest rate.

The second fundamental form of financing inefficiency relates to the effect of asymmetric information on stock market financing (Myers and Majluf, 1984). Under the Modigliani–Miller hypothesis of perfectly efficient capital markets, a firm with favourable investment perspectives can always finance itself with equity issues. Equities are placed at a competitive price, and issuing costs are always zero, as the amount of external finance obtained from the market is always equal to the value of 'issued liabilities'. In this case, the firm finances all positive net present value (NPV) investments and is indifferent between internal and external financing. With imperfect information, though, firm managers may be assumed to act in the interest of existing shareholders and exploit an informational advantage on firm asset and investment perspectives. If these shareholders remain passive on the occasion of new issues, managers may find it disadvantageous to issue new equity in order to finance positive Net Present Value (NPV) investments that cannot be covered completely by internal finance or bank finance. This may occur if a firm's market value is undervalued, so that an equity issue is not in the shareholders' interest (that is, the increase in value of the shareholder stake caused by project returns is inferior to the new shareholders' stake). This effect implies that, in markets where firm managers possess an informational advantage, a new equity issue will be considered as a negative signal and this may generate additional costs of external finance in terms of excessive equity dilution.

The surveyed theoretical models show that asymmetric information causes inefficiencies, both in bank financing and stock market financing. This might lead one to think that a more advantageous external financing strategy is that in which a financing partner with some technological skills relaxes the investor/innovator cash constraint in exchange for participation in future profits from the innovation (venture capital financing).

Venture capital financing, however, also generates undesirable outcomes in the presence of asymmetric information. According to models that adopt a co-ordination failure approach, co-ordination inefficiency and excess cost of financing may occur in a simple two-agent game between a financier and an innovator. This is because, when ownership shares are bargained *ex ante* (before the innovation is achieved), an imbalance between relative bargaining strengths and relative contributions to the venture generates an inefficient division of ownership with a divergence between private and social optima. What generally happens is that the financier's excess bargaining power leads to equilibria that are nearer to his/her individual optimum than to the investor's individual optimum and to the social optimum.[1] The simple reason for this imbalance is that cash-constrained innovators possess a unique non-diversifiable asset (their talent) while financiers have the opportunity to diversify their investment over a wide range of financial assets alternative to the venture capital choice.

An interesting normative consideration in the extension of the analysis to a game with multiple agents is that if the investor/innovator has the higher relative contribution in the investment/innovation success, while the financial unit has the higher relative bargaining power, an increase in the number of financial units may reduce the costs of venture capital financing for the innovator and restore the social optimum for the given incentive structure. This result is consistent with innovation policies adopted in several countries which support, through tax relief and other instruments, the creation of venture capital specialized in financing risky ventures.[2]

What solutions may be found to improve the relationship between investment, innovation and finance? Financial intermediaries which typically find a justification for their existence in the informational economies of scale may solve this problem by improving their project monitoring and evaluating skills. A wide range of financing strategies available on the market to firms' managers may also help to solve the problem, but only if the costs and benefits of choosing different strategies are such that signalling equilibria can be realized. This occurs typically if costs are higher than gains from mimicking signalling strategies of higher-quality firms.

The identification of the crucial problem of imperfect information and of its costs leads, then, to different approaches which explain the positive role that financial intermediaries may play in this system. Financial intermediaries (FI) typically: (i) pool funds; (ii) evaluate entrepreneurs; (iii) diversify risk; and (iv) rate expected profits from innovative activities (King and Levine, 1993). In addition, the existence of strategic complementarity between financial markets and technology (both are instruments that can be used for diversification) allows entrepreneurs to spread risk through financial diversification and to choose riskier and more profitable technologies. Without financial markets, entrepreneurs can limit risk only by choosing less specialized and less productive technologies (Saint-Paul, 1992).

These theoretical conclusions lead to the formulation of other questions in a perspective of comparison among financial systems with different institutional features: How effective are different financial systems in reducing the informational problem? Which form of national interaction between FI and innovating entrepreneurs is the optimal one?

From this perspective, the typical distinction between market-orientated and bank-orientated financial systems is increasingly blurred, and the process of integration and competition among systems leading towards a new hybrid system which possesses a wide range of financial assets (and opportunities of cross-sectional risk-sharing) of market-orientated systems, requires greater intermediation.

Common currency and increased competition will play a decisive role in the process of progressive convergence of the (once called) bank-orientated and market-orientated financial systems. The old discrimination based on differences in: (i) the role of the banking system; (ii) the protection of small shareholders; (iii) the diffusion of information; and (iv) the trade-off between cross-sectional and intertemporal risk-sharing will give way to a hybrid system which will combine features of the two original ones. The new system will be market-orientated in the sense of a strong development of financial markets and of a proliferation of financial securities. The increasing flow of information will always be more difficult to select and to handle in real time, so that the role of financial intermediaries will be more important in facilitating small savers' access to financial markets.

In this changed scenario, the presumed superiority of bank-orientated financial systems does not seem so important and obvious as it appeared to be some years ago. The capacity of these systems to create long-term relationships with borrowers, and to guarantee the confidentiality of information on interim values on high-tech projects, thereby increasing incentives for long-term investment in innovation (Bhattacharya and Chiesa, 1995), were probably overvalued with respect to the costs in terms of lack of transparency of close integration between intermediaries. The recent financial crises in the Far East demonstrate that transparency is always a virtue, and that opacity may not generate serious disadvantages in terms of agency costs only if it is backed by strong ethics on the part of the most important actors in real and financial markets.

In addition, the effectiveness of 'market orientated systems' in supporting investment and innovation, on the other hand, may have been understated by earlier literature. The market for corporate control not only provides an important source of monitoring and control over managers' activities, but also represents a relevant source of internal finance for the managers themselves (Bagella and Becchetti, 1997). In addition, even though it is argued that multilateral banking may promote information sharing (and hence technological spillovers) between firms, strong theoretical and empirical support to the 'value increasing hypothesis' on mergers and acquisitions show that they represent a comparative advantage over 'market-orientated' systems.

Even though the positive role of financial markets, and of capital movements, in supporting and promoting growth cannot be neglected, it must also be recognized that short-term financial turbulence may have negative effects on the real economy.

The application of the option theory approach to investment theory recently helped to explain why investments are so sensitive to uncertainty and volatility, and not so sensitive to price effects – as postulated by previous theoretical approaches. According to these models, a high degree of volatility makes it more opportune to wait and postpone investment. This is a richer translation of the old intuitive Stiglitz story (Stiglitz, 1993) of money falling from the ceiling of a classroom and making the cost of following a lecture too high for students. As usual in economics, normative analysis is much more difficult than positive analysis, and the simple idea of 'putting a spoke in the wheels of noise traders', might have some serious drawbacks in terms of the capacity of reduced prices to play their informational role. An indirect, partial and widely acknowledged solution may be that of creating monetary unions (such as the European Monetary Union – EMU) in order to close some financial markets and eliminate some unnecessary sources of financial volatility.

Stiglitz's story then suggests another important insight: education is a public good that is crucial for development and growth in the real sector. The role of the public sector is then that of supporting education, to offset potential disturbances from increased financial-sector volatility.

The emphasis on education is the result of a long process of theoretical, empirical and applied research in economics. The passage from exogenous to endogenous growth models, clearer identification of the features of the non-decreasing return accumulated factor which originates growth, and the mistakes of past development policies which exported and installed capital plants without considering how the local human factor would be crucial in operating them, are fundamental steps in this process.

A first step was to acknowledge that the stylized facts of growth (growth in output and capital per capita, stability of the capital output ratio, and constancy of capital and labour shares of output) could not be explained by exogenous growth models.

The convergence of a restricted club of countries, and the divergence of their growth from that of many less developed countries (LDCs), in fact contradicted the hypothesis of catching-up, and could not be justified entirely by differences in savings and tax structures. The fundamental point of endogenous growth was that growth is not an exogenous process but may be affected crucially by policy decisions. A first vintage of models identified the sources of growth in: (i) the increasing variety of capital goods; (ii) research and development (R&D) activity developed inside and outside the firm; and (iii) Marshallian externalities that transformed constant returns scale (CRS) production functions at firm level into non-decreasing returns of scale, production functions of the agglomeration of productive units as a whole (Romer, 1986; Lucas, 1988; Grossman and Helpmann, 1991).

Further theoretical investigation led to the discovery that the increasing variety of capital goods was only the effect, and not the source, of growth. The idea that the human factor was crucial in creating and operating new varieties of capital goods shifted the emphasis from physical to human capital. Failures in development programmes based solely on the production and provision of capital goods and infrastructure, neglecting the education of the local human resources needed to operate them, and contributed to the development of this new growth paradigm.

At the same time, endogenous and non-orthodox theories of growth drew closer to each other, by recognizing that substitution between techniques and the choice of the preferred combination of production inputs along the path of innovation was not an easy task. Endogenous growth theorists developing their models away from the traditional neoclassical paradigm recognized that path dependence, rigidity and limited factor substitution might arise from limits in information, learning and education of the human factor (Lucas, 1988). From this perspective, recent theoretical and empirical papers analyse several aspects of the positive link between education and growth. These papers use the literacy rate as a proxy for the degree of education, and show that the positive link is much stronger for industrialized than for less developed countries. Related findings demonstrate that LDCs may catch up successfully to industrialized countries only if their human capital level is higher than the corresponding per capita growth with respect to an average world cross-country relationship between the two variables. These studies identify male secondary schooling rate and the reduction of the gender gap in schooling rates as fundamental determinants of growth. The relative share of scientific and technical education has also been proved to play an important role.

The best empirical examples of the complex interaction among the above-mentioned factors in generating growth are probably those agglomerations of Italian small-to-medium firms known as *locales* (*distretti industriali*) which increasingly are attracting economists' attention. Becattini (1991) defines a *locale* as 'a socio-territorial entity, characterised by the active presence of both a community of people and a population of firms in one naturally and historically bounded area', that generates both positive and negative spillovers. This community of people shares a homogeneous system of values and views creating a sense of belonging to the district that generates high work mobility, 'comprehensiveness of the local economic life' and easy transmission of skills.

A combination of internal co-operation and external competition, on the job and outside the job learning, and cultural and professional homogeneity, which facilitate the process of job creation and destruction, are the success factors of this experience. Recent theoretical and empirical studies, though, show that not only small-to-medium firms benefit from network externalities. In high-tech sectors, a new way of modelling the productive process focuses on the concept of 'systemic product', or a product with a complex structure which assembles different components (for example, radar, aircraft engines, but also personal computers). The systemic product is produced by a network of firms including a system company which controls the architecture of the product and several component producers. The interaction between these productive units generates positive technological externalities which in turn affect ownership of the various parts of the product.

In the light of these experiences, the most recent literature on growth therefore considers education and geographical agglomeration of industrial units as key engines of growth. Many successful examples of economic growth in different continents are now recognised as having started from well-delimited enclaves in which a high quality of human capital and other favourable conditions have fostered the productivity of local units. In this sense, the experience of Italian *locales* and of other similar areas in the rest of the world must attract the attention of researchers. What needs to be evaluated is how the geographical agglomeration of productive units and the creation of an environment which mixes elements of co-operation and competition may have helped human capital to grow by multiplying opportunities of learning both in and outside the job, and of relocating skills and jobs. An important line of research is how easier job and skill relocation might have significant positive effects on technological innovation by reducing uncertainty on intertemporal innovation profit-sharing and on the appropriateness of non-proprietary knowledge that remains part of the district's educational wealth.

In this perspective, more focus is needed on the role of intermediate entities such as public or private voluntary-based institutions aimed at increasing the quality of services and public goods needed to increase the productivity of the *locale*.

In sum, endogenous growth depends on the virtuous interaction among human capital formation, geography and service, and public good provision by intermediate entities in a way that still has to be explored thoroughly in order to develop clearer normative considerations.

This volume starts from the above-mentioned analysis on the state of the art in the relationship between finance, research education and growth. It collects contributions that attempt to shed more light on the issues outlined above. The hope is that the positive results and normative suggestions emerging from this may help to provide suggestions for an improved normative framework which promotes a growth-enhancing interaction between the real sector, financial markets, research and education.

Note

1 Further analysis, though, shows that even in a context of perfect information on relative contributions to the venture, the asymmetry between relative bargaining powers and relative contributions exists and has potential negative effects on social optimality. This is because the unit that enjoys the asymmetry has an individual convenience in maintaining this advantage and exerting all its bargaining power because 'a larger share of a smaller cake is bigger than a smaller share of a larger cake'. In this case, the bargaining outcome is individually optimal for the side with higher relative bargaining power (usually the financier), but socially suboptimal, *given the existing incentive structure*.

References

Bagella, M. and Becchetti, L. (1997) 'Real Effectiveness of National Systems of Finance, Investment and Innovation: A Review of the Literature and a Proposal for a Comparative Approach', in M. Bagella (ed.), *Finance, Investment and Innovation: Theory and Empirical Analyses*, (Aldershot: Ashgate).

Becattini, G. (1991) 'Il Distretto Industriale Marshalliano come Concetto Socioeconomico', in F. Pyke, G. Becattini and W. Sengenberger (eds), *Distretti*

Industriali e Cooperazione tra imprese in Italia, Quaderno n. 34 di Studi e Informazione della Banca Toscana.
Bhattachary, S. and Chiesa, G. (1995) 'Proprietary Information, Financial Intermediation and Research Incentives', *Journal of Financial Intermediation*, vol. 4, pp. 328–57.
Grossman, G. M. and Helpman, E. (1991) *Innovation and Growth in the Global Economy*, (Cambridge, MA: MIT Press).
Grossman, S. and Stiglitz, J. E. (1980) 'On the Impossibility of Informationally Efficient Markets', *American Economic Review*, vol. 70, pp. 393–608.
Jaffe, D. and Russel, T. (1976) 'Imperfect Information, Uncertainty, and Credit Rationing', *Quarterly Journal of Economics*, pp. 651–66.
King, R. G. and Levine, R. (1993) 'Finance and Growth: Schumpeter Might Be Right', *Quarterly Journal of Economics*, August, vol. 108, no. 3.
King, R. G. and Levine, R. (1993) 'Finance, Entrepreneurship and Growth: Theory and Evidence', *Ross Journal of Monetary Economics*, December, vol. 32, no. 3.
Lucas, R. E. (1988) 'On the Mechanics of Economic Growth', *Journal of Monetary Economics*, vol. 22, pp. 3–42.
Myers, S. C. and Majluf, N. S. (1984) 'Corporate Financing Decisions When Firms Have Investment Information That Investors Do Not', *Journal of Financial Economics*, vol. 13, pp. 187–221.
Romer, P. M. (1986) 'Increasing Returns and Long Run Growth', *Journal of Political Economy*, vol. 94, pp. 1002–37.
Rosenzweig, R. (1990) 'Population Growth, Human Capital Investments: Theory and Evidence', *Journal of Political Economy*, vol. 98, p. 5.
Saint-Paul, G. (1992) 'Technological Choice, Financial Markets and Economic Development', *European Economic Review*, vol. 36, pp. 763–81.
Stiglitz, J. E. (1993) 'The Role of the State in Financial Markets', *Proceedings of the World Bank Annual Conference on Development Economics*, pp. 19–56.
Stiglitz, J. and Weiss, A. (1981) 'Credit Rationing in Markets with Imperfect Information', *American Economic Review*, vol. 71, pp. 912–27.
Tornell, A. (1996) 'Real vs Financial Investment: Can Tobin Taxes Eliminate the Irreversibility Distortion?', *Journal of Development Economics*, vol. 32, no. 2, pp. 419–44.
Williamson, S. D. (1986) 'Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing', *Journal of Monetary Economics*, vol. 18, pp. 159–79.

Part I

Finance and Growth

1

Stock Market Liquidity and Economic Growth: Theory and Evidence

Ross Levine

Introduction

Consider the following three statements. Liquid stock markets were a pre-condition for the Industrial Revolution and a critical factor underlying long-run growth in many countries. Enhanced stock market liquidity reduces saving rates and weakens corporate control, which retard economic growth. Stock markets are basically a sideshow, a casino where players come to place bets, but where there is little feedback to the real economy. Rigorous theoretical models support each of these statements.[1]

The theoretical ambiguity can be exemplified by considering a very stylized and simplified example, the construction of a railway. While potentially very profitable, building a railway requires a long gestation period. Capital must be invested with no returns for many years. If savers are reluctant to relinquish control of their savings for long periods, this reluctance will impede railway construction. Under these conditions, an equity market where it is inexpensive to trade securities at posted prices – a liquid market – reduces this reluctance and thereby facilitates railway construction. Specifically, savers can invest in the railway, and they seek access to their wealth prior to the completion of the railway and the distribution of profits, they can sell their claim in the stock market. The greater the liquidity of the equity market, the lower will be the impediments to investing in long-run projects. By making more investment projects feasible, greater stock market liquidity boosts returns to saving.

Enhanced stock market liquidity may also impede railway construction, however. First, by increasing returns to saving, more liquid markets can lower saving rates if the income effect of higher returns dominates the substitution effect. If savings fall sufficiently, this will make it more difficult to mobilize capital for the railway. Second, more liquid securities markets may encourage ownership of the railway to become more diffuse, and for each owner to spend less time and resources overseeing the construction and operation of the railway. Put simply, if I only have a little invested in the railway and I can cheaply and confidently sell my stock in a liquid market, then I have fewer incentives to monitor the railway energetically than if I have a large portion of my wealth invested in the project and I cannot easily liquidate my holdings. If greater stock market liquidity reduces corporate control importantly, then it will have a negative influence on resource allocation and growth. Thus, the net effect of greater stock market liquidity on the ability of an economy to construct a railway efficiently is theoretically unclear.

After reviewing the theoretical literature on the relationship between stock market liquidity and growth, this chapter presents cross-country evidence using data on forty-nine countries over the period 1976–93. Conceptually, a more liquid stock market is a market where the costs of trading equities and the uncertainty concerning the price, timing and settlement of stock transactions are lower than in a less liquid market. To measure stock market liquidity for each economy, I use the total value of domestic equities traded on each country's major stock exchanges divided by gross domestic product (GDP). This indicator measures stock transactions relative to the size of the economy, and is motivated by theoretical models of stock market liquidity and growth (Levine, 1991; Bencivenga *et al.,* 1995). After controlling for many other factors associated with long-run growth, including measures of banking development and measures of stock market size, and after testing for the importance of 'outliers', I find a statistically and economically strong, positive association between growth and stock market liquidity. While much more empirical work needs to be done to dissect the causal relationship between stock market development and growth, and to identify appropriate policies towards capital markets, this chapter's analyses push one towards theories that predict a positive relationship between growth and liquidity, and away from theories that forecast a negative association between stock market liquidity and national growth rates.[2]

This contribution builds on Atje and Jovanovic's (1993) study of stock market trading and economic growth. Besides increasing importantly the sample of countries and the number of years covered, this chapter controls for initial conditions and other factors that may affect economic growth in the light of evidence that many cross-country regression results are sensitive to changes in the conditioning information set (Levine and Renelt, 1992).

A few cautionary remarks are worthwhile, to alert readers to the limitations of cross-country comparisons. Cross-country growth regressions suffer from measurement, statistical and conceptual problems. In terms of measurement problems, country officials sometimes define, collect and measure variables inconsistently across countries. Further, people with detailed country knowledge frequently find discrepancies between published data and what they know happened in fact. As I discuss below, these measurement difficulties also apply to financial transactions data. In terms of statistical problems, regression analysis assumes that the observations are drawn from the same population. Yet vastly different countries appear in cross-country regressions. Many countries may be sufficiently different that they warrant separate analyses. Conceptually, cross-country regressions do not resolve issues of causality, and they do not examine 'one piece of machinery' over time. Consequently, we should not interpret the estimated coefficients as elasticities that predict by how much growth will change following a particular policy change. Rather, the coefficient estimates and the associated t-statistics evaluate the strength of the partial correlation between stock market development and economic growth.[3]

These measurement, statistical and conceptual problems, however, should not blur the benefits that can accrue from cross-country comparisons. Elucidating cross-country empirical regularities between stock market development and economic growth will influence beliefs about this relationship, and shape future theoretical and empirical research. Put differently, beliefs about stock markets and growth that cross-country comparisons do not confirm will be viewed more sceptically than those views that are confirmed by cross-country regressions.

I organize the remainder of the chapter as follows: the second section reviews the theoretical literature on the functioning of stock markets and economic growth; the third section turns to the data and evaluates the strength of the empirical link between stock market liquidity development and long-run economic growth; while the fourth section concludes.

Theoretical overview

The theoretical literature provides ambiguous predictions regarding the influence of stock market liquidity on national economic growth rates. Liquid stock markets are markets where it is relatively inexpensive to trade equities, and where there is relatively little uncertainty concerning the price, timing and settlement of those trades. This section explains that the theoretically ambiguous relationship between growth and stock market liquidity derives from three core sources. First, stock market liquidity lowers the risk of investing in longer-run, high-return projects, and in consequence fosters a growth-accelerating reallocation of capital. The lower risk, however, affects saving and capital accumulation rates ambiguously, so that aggregate growth will slow if saving rates fall enough. Second, stock market liquidity lowers the cost of investing in longer-run, higher-return projects, and thereby induces a growth-enhancing reallocation of capital. The higher rate of return on savings, however, affects saving and capital accumulation rates ambiguously, so that growth will fall if capital accumulation rates fall enough. Finally, stock market liquidity affects incentives for investors to undertake the costly processes of researching and monitoring firms and managers ambiguously. If stock market liquidity induces agents to evaluate firms and exert corporate control more rigorously, then liquidity will affect growth positively. Alternatively, if greater stock market liquidity reduces incentives to assess firms and managers, it will influence long-run growth rates negatively.

Consider first the relationship between stock market liquidity and risk. Many high-return projects require a longer-run commitment of capital than lower-return projects. Savers, however, are generally averse to relinquishing control of their savings for long periods. In financial autarky with risk averse agents, this liquidity risk will reduce investment in longer-run, higher-return projects, Bencivenga and Smith (1991) and Levine (1991) model this liquidity risk as an agent-specific, privately observed shock to preferences.[4] They use an overlapping generations model in which agents live for three periods and have a utility function of the following form:

$$U(c_2, c_3) = -[c_2 + \varphi c_3]^{-\gamma}/\gamma,$$

where $\gamma > 0$, and where age i consumption is c_i, and where:

$$\varphi = \begin{cases} 0 & \text{with probability } 1 - \pi \\ 1 & \text{with probability } \pi \end{cases}$$

Agents make saving allocation decisions at age 1. They can invest in a high-return project that pays off in period 3, or a low-return project that pays-off in period 2. Agents care about liquidity – the ability to consume wealth at age 2 – because they may receive $\varphi = 0$ and therefore not value the payoff from the long-run project. The uncertainty associated with being a type 0 agent is 'liquidity risk'. This liquidity risk affects the period 1 allocation decision. Namely, if agents are sufficiently risk averse, liquidity risk reduces investment in high-return projects.

Liquid equity markets can reduce the negative implications of liquidity risk. If transaction costs are not too high, an equity market will arise. Agents who receive $\varphi = 0$ sell their equity claims to period 3 output from the long-term project to agents who receive $\varphi = 1$. The type 1 agents buy these equity claims with the savings they invested in the short-run liquid investment. Thus, if transaction costs are sufficiently low, equity markets reduce liquidity risk – the risk associated with being type 0. More generally, liquid equity markets make long-run investment less risky – and more attractive – because they allow savers to acquire an asset (equity) and to sell it quickly and cheaply if they need access to their savings or want to alter their portfolios. Simultaneously, projects enjoy permanent access to capital raised through equity issues. By facilitating longer-term, more profitable, investments, liquid markets improve the allocation of capital and enhance prospects for long-term growth.

Theory is unclear about the effects of lower liquidity risk on saving rates, however. As shown by Levhari and Srinivasan's (1969) classic article, lower risk may increase or decrease saving rates. Thus, an increase in stock market liquidity that lowers liquidity risk may increase or decrease saving rates in more general versions of the model sketched above, that allow for a non-trivial consumption-saving decision at age 1 (Bencivenga and Smith (1991).[5] If saving rates rise, then the reduction in liquidity risk will tend to increase growth, as both the saving rate and the efficiency of capital allocation rise. If saving rates fall, however, then growth will slow if the fall in savings dominates the improvement in capital allocation.

So far I have focused on how greater stock market liquidity can affect economic growth by altering the riskiness of longer-run, higher-return investments. Greater stock market liquidity, however, can also affect investment returns in a risk-free world (Bencivenga *et al.*, 1995, 1996). To see how, assume that (i) agents live for two periods, working and investing in period 1 and consuming their wealth at age 2; (ii) projects can extend for many periods, and longer-run projects enjoy higher technological rates of return than short-run projects ($R_j > R_{j-1}$, for all j); and (iii) there are deadweight costs associated with each stock market transactions (α), so that the net of transactions cost rate of return on a project of duration j periods is $R_j\ (1 - \alpha)^{j-1}$ because ownership must be transferred in each period throughout the gestation of the project (agents must sell their claim to projects that will produce in the future to enable them to consume their wealth before they die). Thus there will be more transactions the longer the gestation period of the project. It follows that higher transaction costs will reduce the attractiveness of longer-run projects. Thus, greater stock-market liquidity – lower transaction costs – will induce a reallocation of savings into longer-term, higher-return projects. The reallocation has a positive impact on economic growth.

Theory is unclear about the effects of higher returns on saving rates, however. Well-known income and substitution effects suggest that higher returns can increase or decrease saving rates. Thus, greater stock market liquidity will boost returns to saving, but the higher returns may increase or decrease saving rates. If saving rates fall sufficiently, then enhanced stock market liquidity reduces overall growth rates. Indeed, with capital externalities and a large fall in saving rates, enhanced stock market liquidity causes welfare to fall even as returns to investment rise.

Stock markets may also affect incentives for acquiring information about firms and managers (Grossman and Stiglitz, 1980; Kyle, 1984; Holmstrom and Tirole, 1993). Specifically, more-liquid markets may make it easier for an investor who has gained information to trade at posted prices. This will enable the investor to earn a return for expending the resources to find the information before it becomes widely available and prices change. The ability to profit from investing in information-acquisition will stimulate investors to research and monitor firms. Thus, by spurring more information-acquisition, liquid markets improve resource allocation and accelerate economic growth.

Theories differ, however, Stiglitz (1985, 1993), for example, argues that well-functioning stock markets reveal information quickly through price changes. This quick public revelation will reduce – not enhance – incentives for expending private resources to obtain information. Thus, theoretical debate still exists on the importance of stock market liquidity in enhancing incentives to acquire information.

Stock market development may also influence corporate control. More liquid stock markets ease corporate takeovers. Laffont and Tirole (1988), and Scharfstein (1988) argue that the threat of takeover induces managers to maximize the firm's equity price. Thus, by easing corporate takeovers, greater stock market liquidity can mitigate the principal–agent problem and promote efficient resource allocation and growth.

Opinion differs on this issue too. Stiglitz (1985) argues that outsiders will be reluctant to take over firms because outsiders generally have worse information about firms than do existing owners, and both insiders and outsiders recognize this information asymmetry. Thus, the threat of takeover will not be a useful mechanism for exerting corporate control; stock market liquidity, therefore, will not importantly improve corporate control. Moreover, Shleifer and Summers (1988) note that, by simplifying takeovers, stock market development can stimulate welfare-reducing changes in ownership and management. Specifically, a takeover may allow new owners and managers to transfer wealth to themselves by breaking pre-existing implicit contracts between former owners and firm workers, suppliers and other stakeholders. While new owners and managers may profit, there may be a deterioration in the efficiency of resource allocation. Finally, Shleifer and Vishny (1986) and Bhide (1993) argue that greater stock market liquidity encourages more diffuse ownership, and this impedes effective corporate governance.

Thus some theories provide a conceptual basis for believing that enhanced stock market liquidity will boost economic growth importantly. Other theoretical models, however, have a more pessimistic opinion about the importance of stock markets. Given these dissenting views, this chapter examines the empirical relationship between one measure of stock market liquidity and long-run national growth rates.

Stock market liquidity and long-run growth: cross-country evidence

This section provides cross-country evidence regarding the empirical association relationship between stock market liquidity and economic growth. This broad cross-country evidence complements important microeconomic studies of stock market liquidity. Specifically, an influential literature studies whether a security's liquidity affects its price. These studies generally find that an increase in liquidity – as measured by lower bid–ask spreads – tends to increase the security's price (for example, Amihud and Mendelson (1986, 1989)). Thus, liquidity is a positive characteristic that investors are willing to pay for. Also, Demirguc-Kunt and Maksimovic (1996b) present firm-level evidence from thirty countries consistent with the hypothesis that firms with access to liquid stock markets grow at rates faster than they could have grown without this access. This chapter supplements these microeconomic studies by addressing the question: Do countries with more liquid stock exchanges tend to grow faster, holding other factors constant?

A measure of stock market liquidity

I measure stock market liquidity as the ratio of the total value of domestic equities traded on each country's major stock exchanges to GDP. This ratio measures the value of domestic equity transactions relative to the size of the economy. This indicator of stock market liquidity does not measure directly the costs and uncertainties associated with buying and selling securities at posted prices.[6] None the less, the total value traded:GDP indicator (TVT_GDP) measures the degree of trading compared to the level of economic activity. Furthermore, theoretical models of stock market liquidity and economic growth (Levine, 1991; and Bencivenga *et al.,* 1995, 1996) motivate the TVT_GDP proxy for stock market liquidity.

It is important to recognize and avoid one potential pitfall of using TVT_GDP.[7] If investors anticipate large corporate profits, stock prices will rise. This price rise will increase the value of stock trades and therefore boost the value traded:GDP ratio. Thus, the TVT_GDP liquidity indicator would rise without a change in the number of transactions or a fall in transaction costs. It is easy to control for this price effect, however, by using the market capitalization:GDP ratio (MCAP_GDP), which equals the total value of domestic stocks divided by GDP. Note, a rise in stock prices increases MCAP_GDP in the same way that it increases TVT_GDP. Thus, one way to gauge whether the price effect is dominating the relationship between TVT_GDP and growth is to include the market capitalization ratio in the regression simultaneously. The price effect influences both indicators, but only the value traded ratio is related directly to trading. Therefore, if TVT_GDP is correlated significantly with economic growth when controlling for MCAP_GDP, then the price effect is not dominating the relationship between TVT_GDP and growth.

Cross-country regression framework

To evaluate whether stock market liquidity is strongly linked to long-run economic growth, I use cross-country growth regressions. There are data on forty-nine countries during the period 1976–93. The dependent variable, GROWTH, is the growth rate of real per capita GDP averaged over the 1976–93 period.

The structure of our regression equation is the following:

$$\text{GROWTH} = \alpha \mathbf{X} + \beta (\text{TVT\_GDP}) + u \tag{1.1}$$

where **X** is a set of control variables, α is a vector of coefficients on **X**, β is the estimated coefficient on the stock market liquidity indicator, TVT_GDP, and u is an error term.[8]

The goal of the empirical analysis is to assess the strength of the independent partial correlation between stock market liquidity and economic growth. Consequently, I select a large set of potential control variables and alter the variables included as **X** variables in regression (1.1). These variables include the logarithm of initial real per capita GDP (LRGDP), the logarithm of the initial secondary school enrolment rate (LSEC), the number of revolutions and coups (REV), the ratio of government consumption expenditures to GDP (GOVY), the inflation rate (PI), the black market exchange rate premium (BMP), the ratio of exports plus imports to GDP (TRDY), a measure of judicial efficiency (LEGAL), the market capitalization ratio (MCAP_GDP), and the ratio of bank assets to enterprises divided by GDP (BANK).[9]

Before describing the results, I first define and discuss each of the variables used as **X** variables in regression (1.1). The logarithm of initial real per capita GDP and the logarithm of the initial secondary school enrolment rate are included because recent theoretical work suggests an important link between long-run growth and the initial per capita levels of physical and human capital (see Lucas, 1988; Mankiw *et al.*, 1992). We follow Barro (1991) Barro and Sala-i-Martin (1992) and others in using LSEC and LRGDP to proxy for the initial levels of per capita human and physical capital. I include the number of revolutions and coups, since many authors find that political instability is associated negatively with economic growth (see Barro and Sala-i-Martin, 1995 for evidence and citations).

I also include a variety of macroeconomic indicators to evaluate the strength of the partial correlation between stock market liquidity and economic growth (for example, Levine and Renelt, 1992; Levine and Zervos, 1993). GOVY and PI are included because some evidence suggests a positive connection between macroeconomic stability and economic activity, as shown by Fischer (1993), Easterly and Rebelo (1993), and Bruno and Easterly (1995). Similarly, I include BMP, since international price distortions may impede economic growth, as suggested by Dollar (1992). Also, the black market premium is a general indicator of policy distortions and therefore makes a good control variable in assessing the independent relationship between growth and liquidity (Levine and Zervos, 1993). The last general macroeconomic indicator I use is the ratio of exports plus imports divided by GDP, since openness to international trade may also affect long-run growth. Thus, I include GOVY, PI, BMP and TRDY primarily to gauge the strength of the partial correlation between stock market liquidity and long-run growth.

Besides these standard initial value indicators and macroeconomic indicators, I also include a measure of judicial efficiency taken from Mauro (1995). This measure is an index ranging from 1 (lowest judicial efficiency) to 10 (highest judicial efficiency) based on subjective assessments of judicial efficiency in a broad cross-section of countries. It is important to control for judicial efficiency, since cross-country differences in stock market liquidity could primarily reflect cross-country differences in legal systems, and differences in judicial efficiency may affect growth through channels other than stock market activity. Thus, to assess whether there is an independent empirical connection between stock market liquidity and growth, I control for the level of judicial efficiency (LEGAL).

Furthermore, as discussed above, I control for the size of the stock market (MCAP_GDP). Since expectations of future corporate profits will boost TVT_GDP without implying a corresponding fall in transaction costs, I include MCAP_GDP, which is also liable to this price effect. If TVT_GDP remains correlated significantly with growth while controlling for MCAP_GDP, then readers can feel more comfortable that this relationship does not simply reflect expectations of future corporate profits.

Finally, I control for the level of banking development. A prominent line of research stresses the role of financial intermediaries in economic growth. Schumpeter (1932), Bagehot (1962), Cameron *et al.* (1967), Goldsmith (1969) and McKinnon (1973) provide conceptual descriptions of how, and empirical examples of when, financial systems affect economic growth. Building on these seminal contributions, King and Levine 1993a, 1993b show that measures of banking development are correlated strongly with economic growth in a broad cross-section of countries. Since stock market development is correlated positively with the development of banks (Demirguc-Kunt and Levine, 1996a, 1996b), I control for the level of banking development in assessing the empirical association between stock market liquidity and economic growth, using the ratio of bank loans to enterprises divided by GDP (BANK) as an indicator of banking development.[10]

Data

There are data for a maximum of forty-nine countries over the period 1976–93. The countries are Argentina, Australia, Austria, Bangladesh, Belgium, Brazil, Canada, Chile, Colombia, Costa Rica, Côte d'Ivoire, Denmark, Egypt, Finland, France, Germany, Greece, Hong Kong, India, Indonesia, Israel, Italy, Jamaica, Japan, Jordan, Korea, Luxembourg, Malaysia, Mauritius, Mexico, Nigeria, New Zealand, Norway, Pakistan, Peru, the Philippines, Portugal, Singapore, Spain, Sri Lanka, Sweden, Taiwan, Thailand, the Netherlands, Turkey, the United Kingdom, the United States of America, Venezuela and Zimbabwe. The stock market data are from the International Finance Corporation's Emerging Stock Markets act book and the International Monetary Fund's International Financial Statistics. Data on banking development are from the International Financial Statistics. Data on real per capita GDP growth, secondary school enrolment rates, government consumption spending, exports and imports are from the World Bank. The number of revolutions and coups is from Barro (1991) and initial real per capita GDP is computed from Summer and Heston 1988. Data on the black market exchange rate premium are from Picks Currency Yearbook (various issues) and International Currency Analysis (various issues).

Table 1.1 provides summary statistics on the variables. As shown, the data exhibit wide cross-country variation. Real per capita GDP growth averaged almost 10 per cent in Korea and a *negative* 2.5 per cent in Cote d'Ivoire over the eighteen-year sample period. The ratio of total value traded to GDP averaged 1.2 in Taiwan, but was very close to zero in Bangladesh, Costa Rica and Nigeria.

Table 1.2 provides correlations and P-values. Note that the TVT_GDP is correlated significantly with government spending, international trade, the efficiency of the legal system, the ratio of market capitalization to GDP, and the size of the banking system. Thus it is important to control for these variables in evaluating the strength of the partial correlation between growth and stock market liquidity. Also note that stock market size (MCAP_GDP) is very highly correlated with banking development (BANK) and the efficiency of the legal system (LEGAL).

Results

Table 1.3 presents cross-country regression results with different conditioning information sets; that is, with different sets of **X** variables. Regression (1) is a base regression that only includes a constant, the logarithm of initial income, the logarithm of secondary school enrolment, and the number of revolutions and coups, along with TVT_GDP.

Table 1.1 Summary statistics, 1976–94

	GROWTH	RGDP	SEC	REV	GOVY	PI	BMP	TRDY	LEGAL	MCAP_GDP	TVT_GDP	BANK
Mean	0.021	5959	57	0.139	0.155	0.39	16.6	0.55	7.65	0.34	0.11	0.81
Median	0.018	2762	57	0.059	0.159	0.08	2.9	0.46	7.25	0.17	0.04	0.72
Maximum	0.097	21693	92	1.588	0.347	10.93	132.6	3.04	10.00	2.45	1.16	2.68
Minimum	−0.025	0	9	0.000	0.070	0.03	−1.1	0.12	2.50	0.01	0.00	0.12
Std. Dev.	0.022	5994	26	0.254	0.056	1.53	31.8	0.47	2.04	0.44	0.19	0.56
Observations	51	51	49	51	51	51	51	51	49	50	51	51

Notes: GROWTH = real per capital GDP growth; RGDP = initial real GDP per capita, 1976; SEC = initial secondary school enrolment rate, 1976; REV = number of revolutions and coups; GOVY = government consumption spending / GDP; PI = average annual inflation rate; BMP = average black market exchange rate premium; TRDY = exports + imports divided by GDP; LEGAL = index of judicial efficiency in the 1980s (Mauro, 1995); MCAP_GDP = domestic stock market capitalization/GDP; TVT_GDP = total value of domestic equities traded/GDP

Table 1.2 Correlations

	LRGDP	LSEC	REV	GOVY	PI	BMP	TRDY	LEGAL	MCAP_GDP	BPY	TVT_GDP
LRGDP	1.00	0.59	−0.41	0.47	−0.70	−0.39	0.15	0.55	0.27	0.57	0.19
		0.00	0.00	0.00	0.53	0.01	0.00	0.00	0.00	0.00	0.00
LSEC		1.00	−0.17	0.36	−0.02	−0.62	0.09	0.57	0.32	0.58	0.22
			0.00	0.00	0.08	0.01	0.00	0.00	0.00	0.00	0.00
REV			1.00	−0.28	0.23	0.15	−0.14	−0.38	−0.18	−0.29	−0.15
				0.01	0.02	0.01	0.05	0.01	0.22	0.07	0.38
GOVY				1.00	−0.20	−0.28	−0.05	0.51	0.03	0.27	−0.01
					0.22	0.01	0.00	0.00	0.00	0.00	0.00
PI					1.00	0.08	−0.11	−0.09	−0.11	−0.26	−0.11
						0.18	0.42	0.12	0.62	0.56	0.82
BMP						1.00	−0.16	−0.26	−0.29	−0.50	−0.25
							0.07	0.01	0.58	0.29	0.86
TRDY							1.00	0.29	0.74	0.39	0.39
								0.00	0.00	0.00	0.00
LEGAL								1.00	0.54	0.51	0.23
									0.00	0.00	0.00
MCAP_GDP									1.00	0.52	0.56
										0.00	0.00
BANK										1.00	0.52
											0.00
TVT_GDP											1.00

Note: See variable definitions in Table 1.1.

Table 1.3 Economic growth and stock market liquidity

	(1)	(2)	(3)	(4)	(5)	(6)
C	0.017318152	0.054279133	0.058003928	0.052315574	0.05460077	0.052462462
	0.28343891	0.021661121	0.019982318	0.022980915	0.02766479	0.036462981
LRGDP	−0.004631255	−0.011668086	−0.012360287	−0.011509134	−0.011956503	−0.010561877
	0.025330425	0.001037283	3.93E-04	1.01E-03	0.000485461	0.001022738
LSEC	0.00919213	0.016878558	0.01620813	0.017380785	0.016753897	0.017928484
	0.117930314	0.102190485	0.127328762	0.088540133	0.12123138	0.143936243
REV	−0.012831279	−0.017763815	−0.017218416	−0.016688781	−0.016341019	−0.018840142
	0.056720694	0.086992557	0.094004164	0.147919227	0.152503445	0.103025077
GOVY		−0.053158096	−0.053994095	−0.060684361	−0.061656868	−0.030510346
		0.150835923	1.53E-01	1.17E-01	0.119208869	0.456511548
PI		−0.007950852	−0.008168934	−0.007845344	−0.007920352	−0.006610086
		0.008216261	0.006274905	0.008086469	0.007836957	0.017206639
BMP		−0.000211545	−0.000190233	−0.000185767	−0.000172885	−0.000125312
		0.024045825	0.030031674	0.043224161	0.047850811	0.247762995
TRDY		0.01023926	0.008161665	0.008373562	0.008250673	0.006589018
		3.54E−05	0.003468413	0.008910318	0.01089331	0.161138251
LEGAL						−0.002855203
						0.066902862
MCAP_GDP				0.002663578	−0.000337847	0.012028657
				0.47602357	0.950244284	0.309657112
BANK			0.007156021		0.006137066	0.007323565
			0.083471546		0.273428186	0.239794753
TVT_GDP	0.059711547	0.04018321	0.036279595	0.039522254	3.66E−02	0.032292983
	0.000806629	0.000375014	0.000579648	0.000205945	0.00049	0.009066961
Obs.	49	49	49	48	48	46
Adj. R^2	0.307540291	0.515438895	0.516631387	0.494457474	0.489309156	0.490075055

Notes: Variable definitions are given in the note to Table 1.1.
P-values are given in parentheses. Dependent variable: Real per capita GDP growth.

As shown, stock market liquidity is correlated strongly with economic growth. TVT_GDP enters with a coefficient of 0.06 and P-value of 0.001, which signifies a statistically significant relationship at any conventional significance level. The coefficient value of 0.06 suggests that the association is economically large. For illustrative purposes, assume that TVT_GDP is exogenous. Then the estimated coefficient implies that one standard deviation increase in stock market liquidity (0.2) will increase annual real per capita growth by 1.2 percent (0.2*0.06*100). This is huge, since it suggests that a one standard deviation in TVT_GDP increases growth by more than 50 per cent of the average value of GROWTH in the sample.[11]

The remainder of the regressions in Table 1.3 show that the relationship between stock market liquidity and growth remains significant statistically and economically large while altering the conditioning information set. Specifically, after controlling for government spending, inflation, the black market premium, international trade, the efficiency of the legal system, the size (as opposed to the liquidity) of the stock market, and the degree of banking development, the total value traded to GDP ratio remains associated strongly with economic growth at the 1 per cent significance level. Although the coefficient falls by almost half to 0.032, this still represents a large value in economic terms. By including so many control variables, the significance of many of these variables vanishes; it is difficult to establish an independent empirical relationship between many economic indicators and growth. For example, while the black market premium, the international trade ratio, and BANK enter the growth regression significantly and with the 'correct' signs in the more parsimonious regression (3), they all enter insignificantly in regression (6), which also includes LEGAL and MCAP_GDP. This sensitivity to changes in the conditioning information set does not affect stock market liquidity. Stock market liquidity enters all of the growth regressions significantly.

Sensitivity to outliers

In Table 1.3, I chose regression (3) because I wanted to include as many countries as possible (49), and for no better reason, regression (3) had the highest adjusted R^2. The partial scatter is computed as follows. I regress growth on all the regressors in Equation (3) except for

TVT_GDP, and collect the residuals u_g. Then I regress TVT_GDP on the same regressors and collect those residuals, u_l.

Taiwan, Korea, Jamaica, Côte d'Ivoire, Thailand and Luxembourg stand out as potential 'outliers' – as data points that may influence strongly the slope of the regression line, and the statistical strength of the relationship between growth and liquidity. More formal procedures for identifying influential observations as described by Belsley *et al.* (1980) also highlight these countries. To examine the importance of these data points I removed them from the sample systematically and re-ran regression (3). Table 1.4 presents these results. While removing different countries alters the size of the coefficient on stock market liquidity, stock market liquidity enters all of the regressions significantly at the 0.01 significance level and the coefficient remains larger than 0.03. Thus the strength of the partial correlation between growth and stock market liquidity is largely insensitive to changes in the conditioning information set, and to the removal of particularly influential observations.

Note that the strength of the empirical relationship between stock market liquidity and long-run economic growth should not be overemphasized, however. Although this chapter attempts to control for many other factors associated with growth, I may be omitting an important variable that is driving both stock market liquidity and economic growth. Similarly, while I control for outliers and heteroskedasticity, other diagnostic tests may show that the liquidity–growth relationship deteriorates under particular conditions. Also, this chapter simply looks at the broad cross-country relationship between growth and liquidity after aggregating the data over time to abstract from higher frequency interactions between liquidity and growth. Time-series procedures like those used by Neusser and Kugler (1996) to examine the relationship between financial intermediary development and manufacturing growth would provide significant value-added to this chapter's pure cross-country comparisons. Similarly, the chapter examines simple linear relationships, while there may exist non-linear relationships that distort my findings. Finally, if one selectively omits enough countries and adds enough regressors, the relationship between growth and liquidity weakens.[12] Thus much work remains in documenting the relationship between stock market liquidity and economic growth.

Table 1.4 Growth and liquidity: checking for the importance of outliers

	(1)	(2)	(3)	(4)	(5)
Excludes		TWN	KOR	JAM CIV	THA LUX
C	0.058003928	0.061368505	0.057605318	0.078610004	0.068736424
	0.019982318	0.012190745	0.023904667	3.36E-06	7.78E−08
LRGDP	−0.012360287	−0.012820956	−0.011021417	−0.010106786	−0.009702896
	0.000393235	0.000316465	0.000261072	3.17E−07	1.88E−07
LSEC	0.01620813	0.015726004	0.011973639	0.00519075	0.006585569
	0.127328762	0.121767498	0.229627929	0.209022143	0.055213249
REV	−0.017218416	−0.017983793	−0.017428258	−0.020743448	−0.022513236
	0.094004164	0.050125516	0.011545066	3.58E−06	8.48E−10
GOVY	−0.053994095	−0.042244978	−0.026447325	−0.016641318	−0.017638957
	0.153391604	0.198111991	0.250599434	0.412253423	0.372437729
PI	−0.008168934	−0.008363003	−0.006667386	−0.005921891	−0.005466838
	0.006274905	0.004798222	0.005390371	3.29E−05	2.86E−05
BMP	−0.000190233	−0.000179547	−0.000176059	−0.000243045	−0.000212139
	0.030031674	0.033794613	0.051183203	0.000158076	7.15E−05
TRDY	0.008161665	0.005895728	0.006885501	0.008600072	0.009940138
	0.003468413	0.119027129	0.018853324	1.34E−05	3.15E−07
BANK	0.007156021	0.006116379	0.006499786	0.006199264	0.006950357
	0.083471546	0.186330214	0.101697171	0.05077349	0.080148807
TVT_GDP	0.036279595	0.067138747	0.054865365	0.042963024	0.033151941
	0.000579648	0.002525116	0.00040794	5.87E−05	0.004043136
Obs.	49	48	47	45	43
Adj. R^2	0.516631387	0.493969031	0.559255843	0.750475251	0.789216747

Notes: Variable definitions are given in Table 1.1. Dependent variable is growth in real per capita GDP.
Country codes: Taiwan (TWN), Korea (KOR), Jamaica (JAM), Côte d'Ivoire (CIV), Thailand (THA), Luxembourg (LUX).

Conclusions and discussion

Theory provides ambiguous predictions about the relationship between stock market liquidity and economic growth. To shed some empirical light on this issue, this paper presents cross-country evidence on the association between one measure of stock market liquidity – the total value of stock transactions divided by GDP – and average economic growth rates over the period 1976–93 using data for forty-nine countries. Subject to various qualifications detailed above, the data suggest that there is a strong, positive relationship between long-run economic growth rates and stock market liquidity. This positive relationship is robust to various changes in the conditioning information set. Furthermore, removing outliers – particularly influential observations – does not alter the strength of partial correlation between growth and stock market liquidity. Although this chapter does not address empirically the issue of causality, Levine and Zervos (1993) show that the initial level of stock market liquidity in 1976 was a good predictor of economic growth over the next eighteen years. Thus it is not simply contemporaneous shocks to stock market activity and growth that are causing the strong positive association, and it is not simply that growth causes future increases in stock market liquidity. In sum, the data are consistent with theoretical models that predict a positive relationship between stock market liquidity and economic growth. In contrast, theories that predict a negative association between stock market liquidity and growth must reconcile this prediction with existing evidence.

Notes

1 On the potentially growth-enhancing role of stock market liquidity, see Hicks (1969), Levine (1991), Holmstrom and Tirole (1993), and Bencivenga *et al.* (1995, 1996). On the potentially growth-reducing role of stock market liquidity, see Stiglitz (1985, 1993), Shleifer and Vishny (1986), Bhide (1993), and Bencivenga *et al.* (1995, 1996). And, for arguments that the stock market is basically a sideshow, see Morck *et al.* (1990a, 1990b), and Blanchard *et al.* (1993). For a review, see Levine (1997).

2 Levine and Zervos (1996a) show that international capital control liberalizations tend to increase stock market liquidity.

3 There are also problems of aggregation. When averaging over long periods, many changes are occurring simultaneously: countries change policies; economies experience business cycles; and governments rise and fall. Thus, aggregation may blur important events and differences across countries.

4 If agent types were observable publicly, then type-contingent insurance contracts would eliminate this risk.

5 The relationship between the ease with which households can borrow and aggregate saving rates, growth rates and welfare have been studied; for example,

by Miles (1992), and Jappelli and Pagano (1994). Also, on the relationship between risk diversification through stock markets and economic growth, see Saint-Paul (1992), Devereux and Smith (1994), and Obstfeld (1994).

6 I was not able to obtain bid–ask spreads for a broad cross-section of countries.

7 While financial data are often viewed as suffering from less measurement error than other data, there are inconsistencies in the measurement of total value traded across countries. As noted by Wells (1994), some exchanges measure only those transactions that take place through the exchange (for example, Austria, Belgium, Finland, France, Greece, Luxembourg, Portugal, Spain). Other markets attempt to measure all transactions, whether they occur through the exchange or not, by having regulated traders report their trades to the regulatory agency (for example, Denmark, Germany, Ireland, Norway, Sweden, Switzerland, the Netherlands, the United Kingdom). While recognizing this problem, it is not clear how to make the data perfectly consistent. Also, for many of the countries in the sample, I have not been able to identify which type of procedure has been employed in computing the total value of transactions.

8 Throughout the analysis I use heteroskedasticity consistent standard errors as developed by White (1980).

9 I also experimented with other variables, such as the standard deviation of inflation (Levine and Renelt, 1992) and lagged GROWTH (Atje and Jovanovic, 1993). These additional conditioning variables did not alter the conclusions. See also Easterly and Levine (1997).

10 There are problems with this indicator of banking development. Bank loans to GDP is not necessarily positively correlated with how well banks research firms, exert corporate control, provide risk pooling vehicles, and mobilize resources. Also, the International Monetary Fund's *International Financial Statistics* notes that while it seeks to measure bank loans to private firms, there are inconsistencies across countries in the treatment of public enterprises. None the less, the bank loans to enterprises divided by GDP measure seems to be a better proxy for the functioning of the banking system than alternative indicators that only measure the size of bank liabilities, such as M2 divided by GDP.

11 Note that this conceptual experiment is meant to illustrate the size of the 'economic' size of the estimated coefficient on stock market liquidity. As argued above, these coefficients should not be interpreted as elasticities. Moreover, the experiment does not consider *how* to enhance liquidity.

12 For example, excluding all six 'outlier' countries and including all the regressors (which drives the sample down to 41) causes TVT_GDP to enter the growth regression insignificantly.

References

Amihud, Y. and Mendelson, H. (1986) 'Asset Pricing and the Bid–Ask Spread', *Journal of Financial Economics*, vol. 17, pp. 223–49.

Amihud, Y. and Mendelson, H. (1989) 'The Effects of Beta, Bid–Ask Spread, Residual Risk and Size on Stock Returns', *Journal of Finance*, vol. 44, pp. 479–86.

Atje, R. and Jovanovic, B. (1993) 'Stock Markets and Development', *European Economic Review*, vol. 37, no. 2/3, April, pp. 632–40.

Bagehot, W. (1962) *Lombard Street*, first published 1873 (Homewood, IL: Richard D. Irwin).

Barro, R. J. (1991) 'Economic Growth in a Cross Section of Countries', *Quarterly Journal of Economics*, vol. 56, no. 2, May, pp. 407–43.

Barro, R. J. and Sala-i-Martin, X. (1992) 'Convergence', *Journal of Political Economy*, vol. 100, pp. 223–51.

Barro, R. J. and Sala-i-Martin, X. (1995) *Economic Growth* (New York: McGraw-Hill.

Belsley, D., Kuh, E. and Welsch, R. (1980) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity* (New York: Wiley).

Bencivenga, V. R. and Smith, B. D. (1991) 'Financial Intermediation and Endogenous Growth', *Review of Economics Studies*, vol. 63, no. 2, April, pp. 152–209.

Bencivenga, V. R. Smith, B. D. and Starr, R. M. (1995) 'Transactions Costs, Technological Choice, and Endogenous Growth', *Journal of Economic Theory*, vol. 67, pp. 153–77.

Bencivenga, V. R., Smith, B. D. and Starr, R. M. (1996) 'Equity Markets, Transaction Costs, and Capital Accumulation: An Illustration', *World Bank Economic Review*, vol. 10, pp. 241–66.

Bhide, A. (1993) 'The Hidden Costs of Stock Market Liquidity', *Journal of Financial Economics*, vol. 34, pp. 31–51.

Blanchard, O., Rhee, C. and Summers, L. (1993) 'The Stock Market, Profit, and Investment', *Quarterly Journal of Economics*, vol. 108, no. 1, pp. 465–511.

Bruno, M. and Easterly, W. (1995) 'Inflation Crises and Long-Run Growth', Mines, Washington, DC: World Bank, July.

Cameron, R., Crisp, O., Patrick, H. T. and Tilly, R. (1967) *Banking in the Early Stages of Industrialization: A Study in Comparative Economic History* (New York: Oxford University Press).

De Gregorio, J. and Giudotti, P. E. (1995) 'Financial Development and Economic Growth', *World Development*, vol. 23, no. 3, pp. 433–48.

Demirguc-Kunt, A. and Levine, R. (1996a) 'Stock Market Development, Corporate Finance, and Economic Growth: An Overview', *World Bank Economic Review*, vol. 10, pp. 223–40.

Demirguc-Kunt, A. and Levine, R. (1996b) 'Stock Market Development and Financial Intermediaries: Stylized Facts', *World Bank Economic Review*, vol. 10, pp. 291–320.

Demirguc-Kunt, A. and Maksimovic, V. (1996a) 'Stock Market Development and Firm Financing Choices', *World Bank Economic Review*, vol. 10, pp. 341–70.

Demirguc-Kunt, A. and Maksimovic, V. (1996b) 'Financial Constraints, Uses of Funds, and Firm Growth: An International Comparison', Mimeo, Washington, DC: World Bank.

Devereux, M. B. and Smith, G. W. (1994) 'International Risk Sharing and Economic Growth', *International Economic Review*, vol. 35, no. 4, August, pp. 535–50.

Diamond D. W. and Verrecchia, R. E. (1982) 'Optimal Managerial Contracts and Equilibrium Security Prices', *Journal of Finance*, vol. 37, May, pp. 275–87.

Dollar, D. (1992) 'Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence From 95 LDCs, 1976–1985', *Economic Development and Cultural Change*, vol. 40, April, pp. 523–44.

Easterly, W. and Levine, R. (1997) 'Africa's Growth Tragedy: Policies and Ethnic Divisions', *Quarterly Journal of Economics*, November, vol. 112, pp. 1203–50.

Easterly, W. and Rebelo, S. T. (1993) 'Fiscal Policy and Economic Growth: An Empirical Investigation', *Journal of Monetary Economics*, vol. 32, no. 3, December, pp. 417–57.

Fischer, S. (1993) 'The Role of Macroeconomic Factors in Growth', *Journal of Monetary Economics*, vol. 32, no. 3, December, pp. 485–511.

Gelb, A. H. (1989) 'Financial Policies, Growth, and Efficiency', *World Bank PPR Working Paper WPS 202*, Washington, DC; June.

Goldsmith, R. W. (1969) *Financial Structure and Development* (New Haven, CT: Yale University Press).

Grossman, S. and Stiglitz, J. (1980) 'On the Impossibility of Informationally Efficient Markets', *American Economic Review*, vol. 70, pp. 393–408.

Hicks, J. (1969) *A Theory of Economic History* (Oxford: Clarendon Press).

Holmstrom, B. and Tirole, J. (1993) 'Market Liquidity and Performance Monitoring', *Journal of Political Economy*, vol. 101, no. 4, August, pp. 678–709.

International Currency Analysis Inc. (Various issues) *World Currency Yearbook* (Brooklyn, NY).

International Finance Corporation (IFC) (Various years) *Emerging Stock Markets factbook*.

International Monetary Fund (IMF) (Various years) *International Finance Statistics*.

Jappelli, T. and Pagano, M. (1994) 'Saving, Growth, and Liquidity Constraints', *Quarterly Journal of Economics*, vol. 109, pp. 83–110.

King, R. G. and Levine, R. (1993a) 'Finance and Growth: Schumpeter Might Be Right', *Quarterly Journal of Economics*, vol. 108, no. 3, August, pp. 717–38.

King, R. G. and Levine, R. (1993b) 'Finance, Entrepreneurship, and Growth: Theory and Evidence', *Journal of Monetary Economics*, vol. 32, no. 3, December, pp. 513–42.

Kyle, A. S. (1984) 'Market Structure, Information, Futures Markets, and Price Formation', in G. G. Storey, A. Schmitz and A. H. Sarris (eds), *International Agricultural Trade: Advanced Readings in Price Formation, Market Structure, and Price Instability* (Boulder, Col: Westview Press).

Laffont, J.-J. and Tirole, J. (1988) 'Repeated Auctions of Incentive Contracts, Investment, and Bidding Parity with an Application to Takeovers', *Rand Journal of Economics*, vol. 19, Winter, pp. 516–37.

Levhari, D. and Srinivasan, T. N. (1969) 'Optimal Savings under Uncertainty', *Review of Economic Studies*, vol. 35, pp. 153–63.

Levine, R. (1991) 'Stock Markets, Growth, and Tax Policy', *Journal of Finance*, vol. 46, no. 4, September, pp. 1445–65.

Levine, R. (1997) 'Financial Development and Economic Growth: Views and Agenda', *Journal of Economic Literature*, vol. 35, no. 2, June, pp. 688–726.

Levine, R. and Renelt, D. (1992) 'A Sensitivity Analysis of Cross-Country Growth Regressions', *American Economic Review*, vol. 46, no. 4, September, pp. 942–63.

Levine, R. and Zervos, S. (1993) 'What We Have Learned About Policy and Growth from Cross-Country Regressions?', *American Economic Review*, May, pp. 426–30.

Levine, R. and Zervos, S. (1994) 'Looking at the Facts: What We Know about Policy and Growth from Cross Sectional Analysis', in R. Solow and L. Pasinetti (eds), *Economic Growth and the Structure of Long-Term Development* (London: Macmillan), pp. 22–44.

Levine, R. and Zervos, S. (1996) 'Capital Control Liberalization and Stock Market Development', Mimeo, Washington, DC: World Bank.

Lucas, R. E. (1988) 'On the Mechanics of Economic Development', *Journal of Monetary Economics*, vol. 22, no. 1, July, pp. 3–42.

Mankiw, N. G., Romer, D. and Weil, D. (1992) 'A Contribution to the Empirics of Economic Growth', *Quarterly Journal of Economics*, vol. 107, pp. 407–37.

Mauro, P. (1995) 'Corruption and Growth', *Quarterly Journal of Economics*, vol. 110, pp. 681–712.

Mckinnon, R. I. (1973) *Money and Capital in Economic Development*, (Washington, DC: Brookings Institution).

Miles, D. (1992) 'Housing Markets, Consumption and Financial Liberalization in the Major Economies', *European Economic Review*, vol. 36, pp. 1093–127.

Morck, R., Shleifer, A. and Vishny, R. W. (1990a) 'The Stock Market and Investment: Is the Market a Sideshow?', *Brookings Papers on Economic Activity*, vol. 2, pp. 157–215.

Morck, R., Shleifer, A. and Vishny, R. W. (1990b) 'Do Managerial Objectives Drive Bad Acquisitions', *Journal of Finance*, vol. 45, no. 1, March, pp. 31–48.

Neusser, K. and Kugler, M. (1996) 'Manufacturing Growth and Financial Development: Evidence from OECD Countries', Mimeo, Switzerland: University of Berne.

Obstfeld, M. (1994) 'Risk-Taking, Global Diversification, and Growth', *American Economic Review*, vol. 84, no. 5, December, pp. 1310–29.

Picks (Various issues) *Picks Currency Yearbook* (New York: Picks).

Saint-Paul, G. (1992) 'Technological Choice, Financial Markets and Economic Growth', *European Economic Review*, vol. 36, no. 4, May, pp. 763–81.

Scharfstein, D. (1988) 'The Disciplinary Role of Takeovers', *Review of Economic Studies*, vol. 55, April, pp. 185–99.

Schumpeter, J. A. (1932) *The Theory of Economic Development*, first published 1912, transl. Redvers Opie (Cambridge, MA: Harvard University Press).

Shleifer, A. and Summers, L. (1988) 'Breach of Trust in Hostile Takeovers', in A. Auerbach (ed.), *Corporate Takeovers: Causes and Consequences* (Chicago: University of Chicago Press), pp. 33–56.

Shleifer, A. and Vishny, R. W. (1986) 'Large Shareholders and Corporate Control', *Journal of Political Economy*, pp. 461–88.

Stiglitz, J. E. (1985) 'Credit Markets and the Control of Capital', *Journal of Money, Credit and Banking*, vol. 17, no. 2, May, pp. 133–52.

Stiglitz, J. E. (1993) 'The Role of the State in Financial Markets', *Proceedings of the Annual Bank Conference on Development Economics*, pp. 19–52.

Summer, R. and Heston, A. (????) 'The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988', *Quarterly Journal of Economics*, vol. 106, no. 2, pp. 327–68.

Wells, S. (1994) 'Turnover Statstics and Trade Reporting', *London Stock Exchange Quarterly with Quality of Markets Review*, Spring.

White, H. (1980) 'A Heteroskedasticity-Consistent Covariance Matrix and Direct Tests for Heteroskedasticity', *Econometrica*, vol. 48, pp. 817–38.

2

The Empirical Importance of Private Ownership for Economic Growth

Darius Palia and Edmund S. Phelps

Another century of widening participation in the global marketplace and rising international flows of goods and capital is winding up. Yet there are still apparently durable differences in national productivity, even among the most advanced national economies. Many economies remain backward, and few of them have begun the process of catching up.

The explanation that some countries are undersupplied with capital and technology is no longer tenable. Capital is quite mobile across borders – witness the near equality of long-term national real interest rates; besides, in countries' rapid investment stage, national investors seem to be driven to step up their own saving. Technology is similarly available at a normal price. Some argue that deficiencies in skill stand in the way of equal productivity. But most training is offered by employers and, in advanced economies at any rate, they seem to finance such human capital about as easily as they fund tangible investments. The trainability of the labour force is conditioned by its education, of course, and it takes time to increase education. But in the long run the market mechanism can be expected to equalize the private rate of return to schooling, except where the state has driven it below market-sustainable levels. So little education is not the root of low productivity either.

The answer to the puzzle of permanent productivity differences among nations, such as those left after 'conditional convergence', must lie in their differing political economy. Within the neoclassical tradition of economics, starting with the work of Douglass North (1981), the emphasis is on underlying social conditions that weaken the incentives of enterprises and persons to invest. As Jakob Svensson (1994) sums up, 'Poorly enforced property rights due to, for instance, lax crime enforcement, weak court system, excessive regulations and poor patent protection create a wedge between the marginal product of capital and the rate of return that can be privately appropriated by investors.' In a related model by Aaron Tornell and Andres Velasco (1992), political polarization makes investors fear expropriation by a hostile faction. In either case, the results are a decline of investment and growth in the market sector in favour of the underground sector and capital flight.

Fear of redistribution is not the only antagonist to growth in the neoclassical perspective. Social conditions may also lower expected returns by threatening capital's *social* product, as when a plant is sacked by rioters or hit in an air raid. High tax rates are also burdensome for output per head, even for output per worker. And, as noted, illiteracy and innumeracy act as a drag on a country's ability to exploit advances in technology.

In the expropriation theory, and in the neoclassical perspective more generally, the cause of low productivity is exogenous to the *economic* system. And any profit-maximizing system will suffice for its purpose. The theory still functions when adapted to a system of *state* enterprises and persons operating in markets. Profit-maximizing state enterprises, those playing the Lange–Lerner market-socialism game, or those pursuing their own interests will shy from investing where there are risks that a new government will lower the price, tax the proceeds, cut tariff protection and so on. State enterprises also have to worry about crime and labour–union obstruction.

More broadly, the neoclassical theory sees *no* consequences for productivity growth whether the market sector is mainly private or mainly state-owned. If markets are perfectly informed and form rational expectations regarding sociopolitical risks, and as a result the resource allocation is Pareto-efficient, it does not matter for economic growth whether it is state enterprises or private ones that have to enter debt or equity markets for financing.

This chapter starts from the contrary supposition that economic institutions make a difference, and deficiencies in nations' economic systems typically account for a substantial part of their economic underperformance – the standard view in Joseph A. Schumpeter's generation, and one being re-established by modern theorists such as Mancur Olson and Oliver Williamson. In this perspective, there is some presumption that extensive private ownership and control of production tends to improve or impair, as the case may be, the performance of a country's market sector. The question is: which is it?

In the aftergloom following the interwar years it was widely believed that private enterprise was disadvantageous economically, because of its vulnerability to unemployment and underinvestment. As the conservatives, Milton Friedman and Henry Wallich, saw it, that 'cost' was worth paying to preserve personal and political freedom. Welfare-state visionaries thought the cost could be cut by public expenditure (largely on the output of the private sector) to reduce unemployment, and high budget surpluses to boost national saving.

Proponents of capitalism also did some serious theorizing on its advantages, however, such as von Mises (1963). Now, set off by the turn away from socialism in eastern Europe, there is a new round of thinking on the economic gains from private ownership and control of enterprises.

With this chapter, we set out to weigh the evidence. Our findings are highly tentative, since a great many other factors must be allowed into the statistical analysis to have confidence that any estimated influence of private ownership is not spurious, a result of omitted variables. That said, we must also say that, so far, the key statistical results are proving surprisingly robust to the introduction of additional factors.

Our provisional finding is that, in general, an increase in the share of output under private ownership and control, other things being equal, serves to raise the rate of growth of national productivity. Moreover, from our estimates, the size of this effect is impressive. If that is correct, a great many underdeveloped countries can open the door to much higher productivity by turning over to private enterprises much of the commercial activity that so far remains in the state sector. Even among the advanced economies, substantial privatization can make a significant difference to wages and output per worker.

The neoclassical growth equation

In the neoclassical set-up, a nation's gross domestic product (GDP), to be denoted Z below, is some constant-returns-to-scale function, F, of its capital, K, the labour input, L, augmented by the effective technology, and natural resources, R, similarly augmented. The concession to realism here is that, while the world technology, Λ, is posited to be freely available to entrepreneurs and their engineers, there being no imperfectly informed agents in the neoclassical set-up, only a part of it is usable as a result of limitations in the literacy and numeracy of the country's workers. ('What use is modern technology', *The Economist* asks, 'if a poor country's workers cannot read the instructions on a bag of fertilizer?') The usable technology in the ith country, A_i increases with the average schooling of its labour force, ε_i at a diminishing rate given by η (<1). Then the country's output per worker, z_{10} is:

$$z_i = \varepsilon_i^{\eta} \Lambda \mathrm{F}(k_i, 1, R_i/L_i), z \equiv Z/L, k \equiv K/(\varepsilon^{\eta}\Lambda L) \qquad (2.1)$$

where the elasticities of F(.) are positive, less than one and add up to one. The change in output per unit time is then:

$$dz_i/dt = [\eta(d\varepsilon_i/dt)/\varepsilon_i + (d\Lambda/dt)/\Lambda - \gamma(dL_i/dL_i)/L_i]z_i + \varepsilon_i^{\eta} \Lambda \mathrm{F}_k (k_i, 1, R_i/L_i)\, dki/dt \qquad (2.2)$$

where $\gamma(k, R/L)$ denotes the elasticity of output with respect to natural resources and $\mathrm{F_k}$ (.) gives the marginal productivity of capital gross of depreciation. Obviously, even if the motion of ε is treated as exogenous, completing the system requires the dynamics of k_i.

In the Solow–Swan model, the motion of k is derived by equating domestic investment to national saving, and regarding the latter as a policy variable or quasi-constant as a ratio to national income. But this describes a *closed* economy, at any rate one closed to the world capital market. The great majority of the countries in our data set were fairly open over our sample span, 1960–85; to our knowledge, large discrepancies in interest rates were few. Furthermore, interest rates were not moving with independent trends, as they would if there were disparate trends in government interference with capital flows. If this is the case, we would do better with an open-economy approach.

If capital moved instantaneously from country to country to eliminate incipient discrepancies between the domestic rate of return and the world interest rate, we would have:

$$\mathrm{F}_k (k_i, 1, R_i/L_i = r^* + \delta \qquad (2.3)$$

where r^* is the short-term world real rate and δ, the depreciation rate, is a constant. In this case,

$$dk_i/dt = (k_i \mathrm{F}_{kk}/\mathrm{F}_k)-1\ [(k_i \mathrm{F}_{kR/L}/\mathrm{F}_k)(dL_i/dt_i)/L_i + (dr^*/dt)/(r^* + \delta)]k_i \qquad (2.4)$$

In this variant of the model, dk_i/dt does not possess the convergence property that $\partial(dk_i/dt)/\partial k_i < 0$. The capital stock is not homing in on some natural path from which it is initially away since it is already on that path. Hence the growth rate equation, which is easily derivable from Equations (2.2) and (2.4), does *not* make the growth rate depend on the initial value of k_i – and hence on initial z_i. The current output growth rate depends positively on the rate of growth of ε_i and of Λ, and negatively on the growth rate of L_i – and on nothing more, apart from the political factors acknowledged in the introduction. The present variant of the model says simply that the *level* of productivity depends on the level of education and the level of natural resources, given the world interest rate and the world techno in existence. (See Equations (2.1) and (2.3)).

An alternative version of the open-economy neoclassical model has adjustment implying a capital–stock adjustment equation:

$$dk_i/dt = I(\mathrm{F}_k\,(k_i,\,1,\,R_i/\mathrm{L}_i) - r^* - \delta) \\ - (\delta + \eta\,(d\varepsilon_i/\mathrm{d}t)/\varepsilon_i + (d\Lambda/dt)/\Lambda + (dL_i/dL_\mathrm{i})/L_i)k_i \qquad (2.4')$$

where $I(.)$ denotes gross investment demand per unit of effectively augmented workers, $I/\varepsilon^\eta\,\Lambda L$, and has the usual properties, $I(0) = 0$, $1'(.) > 0$. This adjustment equation makes dk_i/dt decreasing in k_i at least locally. To be precise, we consider at each t the value of $k_i\,(t)$ such that in Equation (2.4′) $dk_i(t)/dt = 0$. Around that sequence of values, to be denoted $k_i^*\,(t)$, the effect at each t of a small increase of $k_i(t)$ is:

$$\partial(dk_i/dt)/\partial k_i = 1'(.)\mathrm{F}_{kk}(.) - (\delta + \eta(d\varepsilon_i/dt)/\varepsilon_i + (d\Lambda/dt)/\Lambda + (dL_i/dL_\mathrm{i})/L_i)) < 0 \qquad (2.5')$$

To characterize the behaviour of the growth rate of output it is helpful to approximate $dk_i(t)/dt$ around $k_i(t) = k_i^*(t)$ with:

$$dk_i(t)/dt = \partial(dk_i/\mathrm{d}t)/\partial k_{i/k\,=\,k}{}^*\,(k_i(t) - k_i^*\,(t)). \qquad (2.6')$$

Then Equations (2.6′), (2.5′), (2.2) and the further approximation:

$$z_i(t)\;z_i^*\,(t) = (k_i(t) - k_i^*(t))\varepsilon_i^\eta\,\Lambda\mathrm{F}_k\,(k_i^*,1,\,R_i/\mathrm{L}_\mathrm{i}) \qquad (2.7')$$

from Equation (2.1) give the approximate productivity–growth equation:

$$dz_i/dt = [\eta(d\varepsilon_i/dt)/\varepsilon_i + (d\Lambda/dt)/\Lambda - \gamma(dL_i/dL_\mathrm{i}/L_\mathrm{i}]z_i \\ + [I'(.)\mathrm{F}_{kk}(.) - (\delta + \eta\,(d\varepsilon_i/dt)/\varepsilon_i + (d\Lambda/dt)/\Lambda + (dL_i/\mathrm{d}L_\mathrm{i})/L_i)] \\ \times (z_i(t) - z_i^*\,(t)) \qquad (2.8')$$

or, rearranging terms to obtain the growth *rate* of productivity,

$$(dz_i/dt)/z_i = \eta\ (d\varepsilon_i/dt)/\varepsilon_i + (d\Lambda/dt)/\Lambda - \gamma(dL_i/\mathrm{d}L_i)/L_i \qquad (2.8'\text{bis})$$
$$+ [I'(.)F_{kk}(.) - (\delta + \eta\ (d\varepsilon_i/dt)/\varepsilon_i + (d\Lambda/dt)/\Lambda + (dL_i/dL_i)/L_i))]$$
$$\times (1 - (z_i^*(t)/z_i(t))).$$

A term involving the ratio $z_i^*\ (t)/z_i(t)$ appears on the right-hand side. Note that $z_i^*(t) = \varepsilon_i^\eta\ \Lambda F(k_i^*(t),\ 1,\ R_i/L_i\ (t))$ and $k_i^*(t)$ is given by Equation (2.4′) with the right-hand side set equal to zero; so $k_i^*(t)$ is a function of r^*, δ, R_i/L_i and the growth rates of ε_i^η, Λ and L_i. Hence, $z_i^*(t)$ is proportional to $\varepsilon_i^\eta\Lambda$, with the factor of proportionality a function of r^* and so on. Hence, output per worker figures in the right-hand side of Equation (2.8′bis) through the ratio $\varepsilon_i^\eta\Lambda/z_i(t)$. (The weight of this ratio in the equation is greater (in absolute value) the lower is r^*, δ and the growth rate of $\varepsilon_i^\eta\ \Lambda L_i$, since $k_i^*\ (t)$ is then greater and the subtractor in the square-bracket expression is also larger.) Thus a country's output per worker *approaches* a path indexed by the domestic education level, which may itself be exogenously changing through time.

In this alternative model, then, an explanatory variable in the growth rate equation is the *level* of domestic output per worker as a *ratio* to (a function of) the domestic education level augmented by world technology – a ratio to which the growth rate is related inversely. (The growth rate of the labour force is also an expanatory variable, besides its role in modifying the role of the output variable.) Hence, output per worker in any two countries having the similar structural equations and displaying equal growth rates of labour supply and schooling (say, zero rates) will be growing in logarithmic parallel – showing the same sequence of growth rates of output per worker – if the initially higher education level of the one is counterbalanced by a sufficiently high initial level of output and capital per worker in such a way that the marginal product of capital is equal.

In view of these results, it makes sense from the neoclassical viewpoint to regress the growth rate of output per worker in a cross-section of countries on $\varepsilon_i^\eta\ \Lambda/z_i\ (t)$. In this spirit, using the data from the set of countries we shall be using to test our private ownership hypotheses in the next section, we have estimated the regression equation:

$$\text{GR6085} = 0.0046 - 0.00843\text{GDP60} \qquad (2.9')$$
$$(t = 3.22)$$
$$+ 0.0286\text{PRIM60} + 0.0492\text{SEC60}$$
$$(2.86) \qquad (1.94)$$

where GR6085 is the annualized growth rate of output per head between 1960 and 1985, expressed in decimal form; GDP60 is gross domestic product per head in 1960; PRIM60 is the enrolment ratio of primary-school-age persons in 1960, and SEC60 is the corresponding ratio for secondary-school-age persons. The output data come from the Summers–Heston Penn world tables and the education variables are from the Barro–Lee data set.

It is encouraging that our forty-three country sample appears to generate estimates that resemble, at least qualitatively, the estimates by Robert Barro and Xavier Sala-i-Martin (1995). Increased education is effective in raising growth. And what they have dubbed 'conditional convergence' is implied, since the growth rate of output per head is lower the smaller is its initial level, and the level that it approaches is conditional on its education attainments.

Modern growth equations

Our proposition is that the real income per worker produced in a country depends not only on the world technology and the stock of schooling of its population; it also depends very much on the extent of the economy's openness to entrepreneurship, and thus innovation. There can be entrepreneurial state enterprises, of course – and there have been. But a line of economists, starting with Schumpeter and von Mises (if not Karl Marx) to, say, Roman Frydman and Andrei Shleifer in our day, have had reason to believe that state enterprises will not match the innovation of private enterprises (for surveys, see Phelps, 1992, 1993).

Private enterprise under capitalism is actuated by a vision of one or more of its owners about what product or process will be profitable – or cease to be profitable – and such visions are private, intuitive, hard to articulate and to defend; but no public justification is required. The enterprise owned by the democratic state is blocked from many actions it judges would be profitable by the need to go through bureaucratic procedures or otherwise seek public approval.

State-owned enterprises tend to slip into the service of interest groups rather than to serve the economy's productivity, while, ideally, private enterprise has only to attend to its bottom line.

Private investors have a long time horizon, as they can reap expected future benefits in the present by selling off shares in their project to one or more knowledgeable buyers from the next generation, but the politician dependent on electoral support cannot bank on persuading the majority of voters to a theoretical reward for 'jam tomorrow'.

Managers of private enterprises are further from the seat of state power than state enterprises typically are. And their applications for state aid or exemption from laws are seen as self-serving, while the claims of state enterprises may be seen as expressing a public interest. So they find lobbying relatively unproductive – and investing and innovating comparatively productive.

For these reasons, and others, it would be difficult to offer state-enterprise managers a 'carrot' of financial incentives to innovate as powerful as those presented to private enterprises. And the 'stick' of insolvency, takeover or a stoppage of financing faced by a private enterprise – under conditions such as reasonably free entry and effective corporate control by private owners, at any rate – is a better spur to innovation than the threats the state can mount against its managers. Furthermore, even where a substantial private sector exists alongside a larger public sector, the lure of contracts with monopoly enterprises in the state sector, or with the central government overseeing that sector, generates an orientation towards rent-seeking in the private sector rather than towards innovation.

The foregoing implies that enterprises do not have available to them a *known* 'world technology' – not one known by them. To innovate, enterprises must explore some of the bits of technology or other knowledge that they are capable of locating in order to make a guess about whether some product or process that is new to them would offer sufficient prospects of profit to be worth undertaking. While the investing enterprise of the neoclassical theory *knows* the possible technology and draws on that part of it that can be utilized economically by workers in view of their educational limitations, entrepreneurs of the modern model are not so up to date, but are constantly improving their knowledge. Thus the innovators operating in a country are engaged in a process of bringing the technology level realized in the economy's products and processes closer to what is possible – while technological advances are pulling in the opposite direction to leave actual practice further behind. (What is said here about scientific and engineering knowledge applies also to knowledge of legal systems, foreign cultures and much else.)

According to the hypothesis of this chapter, the pace of improvement of the actual, or realized, technology level will be faster the greater the share, π, of production under the ownership and control of private entrepreneurs, since they are the most effective innovators. Further, for a given innovative effort, the pace of improvement in the technology in use will depend on the current room for improvement, hence the size of the gap between actual and possible. Hence, denoting the *actual* technology in the *i*th country by A_i and the *possible* technology by Λ, we write:

$$dA_i(t)/dt = -\varphi(\pi)\,[A_i(t) - \Lambda(t)],\ \varphi'(\pi) > 0,\ \varphi(0) \geq 0 \qquad (2.10)$$

Equation (2.10) is implicitly behind a similar-looking equation in Nelson and Phelps (1996). Their equation expressed the distinct hypothesis that, given private ownership of production, the ability of the entrepreneurs to innovate depended on their education. (What use is the world's viniculture technology if vintners do not have at least a master's degree from the University of California at Davis in order to understand product specifications and published analyses of techniques?) We turn later to the education of entrepreneurs.

Regression tests

Testing the private ownership hypothesis requires assembling data on the extent of private ownership in a set of countries of sufficient size. Branko Milanovi (1989, p. 15) estimates for each of some forty countries the percentage of national output produced by 'state-owned enterprises' engaged in 'commercial activities', hence state production net of 'government services proper'. We have adopted the *negative* of this percentage as the variable, labelled PRIVATE1, capturing the scope of private enterprise and control in our regressions. (It is bounded by − 100 and 0 and its mean is −11.7 in the sample.)

Table 2.1 records the regression results of principal interest. The first one (column (1)) is the simple regression of GR6085, a decimal with mean value of 0.022 per annum, against PRIVATE1. According to the coefficient estimate (0.00067), a swing of the explanatory variable from −100 to 0 in a country would improve its annual growth rate by 0.067, a tripling for a country whose growth rate was average in the sample. Although the t-ratio (2.37) is merely respectable, this estimate will be seen to be surprisingly robust to the subsequent inclusion of the variables added below.

The second regression (column (2)) accompanies the private ownership variable with the education variables used in the neoclassical regression. Now the t-ratio of PRIVATE1 is considerably stronger. The t-ratio of PRIM60 is also much stronger, while that of SEC60 is weaker than before – and far weaker than that of PRIVATE1. The coefficient of PRIVATE1 has slipped only slightly.

Table 2.1 Five modern cross-section growth rate regressions

Expl. var.	(1)	(2)	(3)	(4)	(5)
Constant	0.0294	0.0114	0.0120	0.011	0.0097
	(6.83)	(1.92)	(1.97)	(1.72)	(1.39)
PRIVATE	0.00067	0.00063	0.00046	0.00048	0.00051
	(2.37)	(2.66)	(1.70)	(1.69)	(1.55)
GDP60		−0.0077	−0.0084	−0.0084	−0.0081
		(−3.17)	(−3.19)	(−3.14)	(−3.16)
PRIM60		0.031	0.030	0.031	0.033
		(3.35)	(3.10)	(3.02)	(3.31)
SEC60		0.037	0.042	0.042	0.044
		(1.54)	(1.70)	(1.66)	(1.59)
SOC			−0.013	−0.013	
			(−1.36)	(−1.26)	
MIXED		0.00014	0.00019		
			(0.025)	(0.034)	
HSGVX				−0.0000	
				(−0.32)	
INTER1					0.00065
					(0.51)
No. obs.	43	43	43	43	43
Adj. R^2	0.099	0.404	0.411	0.396	0.392

The next regression addresses the question of whether the private ownership variable is proxying for something other than private ownership *per se*. It might be thought that its significance in the regression depends on the inclusion in the country set of some countries that are socialist and therefore pursue several policies, not just a policy of extensive state enterprise, which trade off growth for other goals. (In fact, our data set excludes countries that Milanovi regarded as 'socialist' in the 1970s and early 1980s, such as China and the Soviet Union.) To this end, we add SOC, a dummy variable for social economic systems, and MIXED, a dummy variable for so-called mixed economies, both variables from the Barro–Lee data set and due to Gastil. Of these, SOC is close to being significant. Our PRIVATE1 remains significant, though at a lower level, and the coefficient falls by a quarter.

The fourth regression, which is similarly motivated, controls for government consumption (in OECD terminology). Since growth might suffer from military engagements or threats, and since education has already been allowed for, we used the share in the GDP of government purchases, excluding military and education expenditures, HSGVX. It did not prove significant.

The fifth, and last, regression to be reported here proposes that, as suggested by the Nelson–Phelps hypothesis, the effectiveness of entrepreneurial effort is potentiated by entrepreneurs' higher education, which may very well induce increased effort as well. This suggests that, besides the PRIVATE1 variable, there should be a variable capturing the interaction between the importance of private entrepreneurs and their higher education. We use SEC60 to proxy for the extent of higher education in the population. The product of SEC60 and PRIVATE1 is the interaction variable, labelled INTER1 in column (5). Unfortunately, the interaction term is not statistically significant. Some measure of higher education other than the breadth of secondary education might fare better, but we have not succeeded with the measure at hand.

We have obtained some other results that raise a question. We have found that, to paint with a broad brush, the sociopolitical variables – such as the index of political coups and the Gini coefficient, to name just two, do not contribute much once our private ownership and education variables are in place. It would be tempting to capitalize on these results by announcing the surprisingly strong finding that the choice of economic system and the provision of education transcend in importance the sociopolitical climate in which the economic system functions. But while the political variables do seem to recede in importance once the economic variables are included, the regressions giving these results all contain one of the more odd variables in the Barro–Lee data set, namely the average annual rate of improvement of the terms of trade, a variable that turns out to be strongly significant. Unfortunately, we do not understand at this point why this variable has a *negative* effect on the growth rate, and a fairly powerful one at that. Pending a resolution of this question, we prefer not to report our preliminary findings on the sociopolitical variables.

Concluding remarks

The findings here are of significance on two levels. If our findings are correct, all countries but the most dedicated adherents of capitalism have it within their power to step up markedly the pace of productivity growth in their economies by relinquishing to private interests the ownership and control of commercial enterprises still remaining in the hands of the state. For continental Western Europe, now stagnating under the weight of welfarist conceptions, and those parts of the Third World that put their faith in state-led growth, this presents an extraordinary opportunity to boost productivity and wages.

Although this practical side is justifiably the one that is bound to be of widest interest, the findings here are also important for the theoretical perspective we take on the economic growth of nations. The neoclassical approach finds itself locked into the premises that markets are perfect and expectations are rational. In such a view, there is not much in the way of an apparent role for the visions of private entrepreneurs. We suggest, however, that our understanding of the true roots of economic growth will not get very far without grasping the role of an economic philosophy, still prevalent among some nations, that is willing to entrust the ownership and management of resources to private interests.

But are our findings correct? We have to worry that private enterprise is no better than state ownership where the sociopolitical environment is not propitious – where there is great sociopolitical conflict, for example. Perhaps, in such an environment, private enterprises, though still superior to the state in picking investments, would invest far less in toto than would state enterprises. We need to test for such possible interactions before we can feel highly confident in the thrust of these, our first results.

References

Barro, R. J. and Sala-i-Martin, X. (1995) *Economic Growth* (New York, McGraw-Hill).

Economist The (1996) 'The Mystery of Growth', 25 May, pp. 15–16, 23–5.

Friedman, M. (1962) *Capitalism and Freedom* (Chicago: University of Chicago Press).

Frydman, R. and Rapaczynski, A. (1994) *The Transition in Eastern Europe: Is the State Withering Away?* (London: Francis Pinter).

Milanovi, B. (1989) *Liberalization and Entrepreneurship: Dynamics of Reform in Socialism and Capitalism* (Armonk, NY: M. E. Sharpe).

Von Mises, L. (1963) *Human Action* (2nd edn) First published 1949 (Chicago: Henry Regenery Company).

Murphy, K. M., Shleifer, A. and Vishny, R. W. (1993) 'Why Is Rent-Seeking so Costly to Growth?' *American Economic Review*, vol. 83, May, pp. 409–14.

Nelson, R. R. and Phelps, E. S. (1996) 'Investment in Humans, Technological Diffusion and Economic Growth', *American Economic Review*, vol. 56, May, pp. 69–75.

North, D. (1981) *Structure and Change in Economic History* (New York: W. W. Norton).

Olson, M. (1996) 'Big Bills on the Sidewalk: Why Some Nations are Rich and Others Poor', *Journal of Economic Literature*, March.

Phelps, E. S. (1992) 'Conclusions', Villa Mondragone International Seminar, *Rivista di Politica Economica*, December.

Phelps, E. S. (1993) 'The Argument for Private Ownership and Control', in *Annual Economic Report* (London: European Bank for Reconstruction and Development).

Schumpeter, J. (1934) *The Theory of Economic Development* (Cambridge, MA: Harvard University Press).
Svensson, J. (1994) 'Investment, Property Rights and Political Instability: Theory and Evidence', Seminar Paper 574, Institute for International Economic Studies, University of Stockholm, July.
Tornell, A. and Velasco, A. (1992) 'The Tragedy of the Commons and Economic Growth: Why Does Capital Flow from Poor to Rich Countries?', *Journal of Political Economy*, vol. 100, December, pp. 1208–31.
Wallich, H. C. (1962) *The Cost of Freedom* (New York: Harper).
Williamson, O. (1985) *The Institutions of Capitalism* (Chicago: The Free Press).

3
Intergenerational Transfers and Growth

Giancarlo Marini and Pasquale Scaramozzino

Introduction

The future of the welfare state is under threat. There is a growing consensus that current systems cannot survive, not only because of their heavy pressure on public finances but also for motives of disincentive. In particular, the alleged adverse effects on private saving are thought to undermine seriously efficiency and growth. Special attention has been paid in recent years to reforming 'unsustainable' social security schemes. The conventional view is that social security is harmful to growth, although it may have some merits in enhancing equity.

It is well known that, under the extreme hypotheses underlying Ricardian equivalence, intergenerational transfers would have neither allocative nor distributional consequences. Fully-funded schemes would be neutral even when there is no fully operational bequest motive, since private egotistic agents are indifferent between funding their pensions or investing their savings in alternative assets yielding exactly the same rate of return. Unfunded schemes, by contrast, are thought to produce negative effects on growth; in spite of this, industrialized countries have adopted mandatory pay-as-you-earn pension programmes since the end of the Second World War.

The increase in life expectancy and the decline in birth rates, combined with the pronounced productivity slowdown, have cast serious doubts on the viability of such schemes. Intergenerational transfers have become excessively generous over the years in response to the occurrence of large adverse shocks (see Marchand and Pestieau, 1991; OECD, 1995).

A central issue is to understand why compulsory transfers from the working labour force to the retired population are still so popular in the face of their presumed unsustainability. A possible answer one might expect is that the equity motive must outweigh the alleged efficiency loss brought about by unfunded intergenerational transfers. Gains may arise from several reasons. First, uncertainty could explain the need for social insurance, whereby generations hit by adverse shocks can find support from more fortunate generations (Gordon and Varian, 1988). Market forces alone could not provide intergenerational insurance, and nor could discretionary policies, when geared towards achieving short-term results.

Even in the absence of uncertainty, there are several rationales for the existence of intergenerational social security transfers. The role of partial altruism and consumption externalities has been analysed by Veall (1986). In the presence of partial altruism, the solidarity chain implied by unfunded schemes may also justify why all generations might find it optimal to maintain the previously introduced pay-as-you-earn scheme (Hansson and Stuart, 1989).

The classic motive for social security is, however, based on paternalism (Samuelson, 1975). Private agents behave on the basis of a utility function which is not the true one; in particular, the needs for old-age consumption could be underestimated systematically. Mandatory pension schemes help to solve the typical Samaritan's dilemma (Buchanan, 1975) faced by society when the old had not set aside a sufficient amount of savings when young.

Short-sighted behaviour may also be caused by unwillingness to accept the idea of growing old and grey, or 'deliberate' myopia (Diamond, 1977; Atkinson, 1987).

None of the above 'imperfections' is, however, needed to explain the existence of social security. A different rationale can be based on time consistency and intergenerational equity (Marini and Scaramozzino, 1996a): a far-sighted planner may find it optimal to introduce an unfunded scheme and maintain it over time. The benevolent government should maximize a social welfare function in such a way as to treat all generations alike. This requires discounting utilities of all generations to their birth date and not to the current date, in order to eliminate incentives to deviate from the optimal policy in the future.

Time consistency implies that future utilities will be discounted at a rate above the population growth rate. The social discount rate, however, reflects not merely impatience but also care for the currently old at any point in time. In other words, the reverse discounting procedure implies that generations must be given a lower weight when young. Unfunded schemes may be optimal even when Aaron's (1966) rule is violated (that is, when the economy is dynamically efficient) if the well-being of the elderly is sufficiently valued. Care for the old can thus justify both the introduction and the survival of intergenerational social security transfers.

Social security is, however, also bound to have allocative consequences. It is well known since the seminal work by Diamond (1965) that the introduction of unfunded social security does not result in a Pareto improvement, unless the economy is characterized by dynamic inefficiency. In other words, the transfers to the currently old need to be financed at the expense of future generations, unless the sum of productivity and population growth rates exceeds the real rate of interest. Pay-as-you-earn schemes affect the dynamic path of capital accumulation, and the usual effect is believed to be negative. In standard exogenous growth models this is certainly the case, since the production locus is shifted down by the introduction of social security. Private savings are crowded out and capital accumulation is reduced (Feldstein, 1985, 1995; Blanchard and Fischer, 1989, ch. 3).

This finding has also been extended to the endogenous growth literature, which relies even more heavily on the crucial role of savings as the engine of growth. It is well known that endogenous growth models rule out the possibility of dynamic inefficiency (Saint-Paul, 1992). The logical corollary would appear to be that social security is always harmful to growth in endogenous growth models. Hence, efficiency considerations would seem unambiguously to call against social security.

A strong warning against such a conclusion has been made by Atkinson (1995), who argues that reductions in unfunded social security might be accompanied by an increase in the private market for pension funds with unclear effects on the saving rate and the growth rate of firms.

We take up this issue in a simple learning-by-doing endogenous growth model with overlapping generations following Samuelson (1958) and demonstrate that social security may indeed contribute positively to growth. The reason is that the saving rate, while falling on impact, can increase in all subsequent periods (although at a declining rate), converging monotonically back to the rate prevailing before the introduction of social security. This result, which seems to have escaped the attention of the literature, casts serious doubt both on the conventional wisdom about the role of unfunded social security and on the appropriate policies to sustain production and growth. Properly designed unfunded intergenerational transfers, far from being harmful to growth, may indeed promote prosperity along the adjustment path.

The scheme of this chapter is as follows. The second section offers a brief review of the effects of social security on growth in the standard exogenous growth literature. The third section considers a prototypical endogenous growth model and shows the rather different implications of mandatory intergenerational transfers. The main results are summarized in the fourth, and concluding section.

Social security transfers and exogenous growth

In this section we review briefly the effects of unfunded social security in an overlapping generations (OLG) model with production and exogenous population growth. We follow Diamond (1965) in considering a simple OLG model formed of identical consumers, each of whom lives for two periods. Individual preferences are given by:

$$U(c_1^s, c_2^s) = u(c_1^s) + \frac{1}{1+\rho} u(c_2^s) \tag{3.1}$$

where the subscript denotes age, the superscript the date of birth, and where ρ is the private subjective rate of time preference. Consumers work when young, and spend when old. Population grows at the constant rate n. The individual lifetime budget constraint in the presence of a social security transfer programme is:

$$c_1^s + \frac{c_2^s}{1+r_s} = y_s - \tau + \frac{\beta}{1+r_s} \tag{3.2}$$

where y_s is income received by the young born at time s, τ are their contributions, β are the benefits received when old, and r_s is the rate of interest at time s. Following Samuelson (1969), Feldstein (1985) and Veall (1986), we assume for simplicity that the utility index in Equation (3.1) has a logarithmic form:

$$u(c) = \text{In } c \tag{3.3}$$

Firms produce a homogeneous output by using capital and labour. Technology is described by a constant-returns-to-scale production function, satisfying Inada's conditions:

$$Y_{it} = F(K_{it}, L_{it}) = L_{it} f(k_{it}) \tag{3.4}$$

where $k_{it} \equiv K_{it}/L_{it}$ is the capital–labour ratio for firm i. Total capital and labour in the economy are defined as $K_t \equiv \Sigma_i K_{it}$ and $L_t \equiv \Sigma_i L_{it}$, where L_t is the number of young consumers.

Perfect competition in the capital and in the labour markets yields the usual first-order conditions:

$$r_t = f'(k_t) \tag{3.5}$$

$$w_t = f(k_t) - k_t f'(k_t) \tag{3.6}$$

where w_t is the wage rate.

The government enacts a balanced-budget, pay-as-you-earn transfer scheme in which the benefits to the elderly are paid entirely through contributions by the young: hence, $(1 + n)\,\tau = \beta$. We consider the effects of an unanticipated transfer to the elderly at time t, financed by the young. This represents the introduction of a social security scheme. The transfer programme is then expected to remain in place in all future periods. Consumption by the different generations is given by:

$$c_2^{t-1} = \frac{1+r_{t-1}}{2+\rho} \cdot w_{t-1} + \beta \tag{3.7a}$$

$$c_1^s = \frac{1+\rho}{2+\rho}\left[w_s + \frac{\beta\,(n-r_s)}{(1+r_s)(1+n)}\right] \qquad s \geq t \tag{3.7b}$$

$$c_2^s = \frac{1}{2+\rho}\left[(1+r_s)w_s + \frac{\beta\,(n-r_s)}{(1+n)}\right] \qquad s \geq t \tag{3.7c}$$

Consumption by the elderly always increases with the benefit β. The effect on consumption by future generations will be either positive or negative, depending on whether the lifetime income of future generations increases or decreases. One critical issue is whether the population growth rate is greater or smaller than the rate of interest: if $n > r_s$, consumption by future generations will increase. This would be the case under dynamic inefficiency. By contrast, if $r_s > n$ the economy is dynamically efficient and consumption by future generations will decline. The impact on future consumption will also depend on the effect of the transfer programme on capital accumulation and, hence, on the wage rate (Equation (3.6)).

The model is closed by the asset market clearing condition, which equates the demand for capital by firms to the supply of savings by the young:

$$K_{s+1} = L_s[y_s - c_1^s] \tag{3.8}$$

The formal properties of this model have been studied by a number of authors, and are well known in the literature (see, for example, Blanchard and Fischer (1989) sect. 3.2; Azariadis (1993), ch. 18, for a formal solution of the model). The introduction of social security has the effect of reducing unambiguously both the steady-state capital stock and the speed of convergence towards equilibrium, along the dynamic adjustment path. This result is driven by the decline in the equilibrium saving rate of the economy, which leads to reduced capital accumulation.

The desirability of balanced-budget, pay-as-you-earn schemes appears to be limited to the circumstance in which the economy is over-accumulating before the introduction of the programme (Samuelson, 1975). Under dynamic efficiency, the scheme could still be justified by the requirements of symmetry across generations, and of time consistency, if the social welfare function reflects sufficient concern for the elderly (Marini and Scaramozzino, 1996a). However, if the economy is not dynamically efficient there is a trade-off between the welfare of the elderly and that of future generations. The terms of this trade-off are substantially modified when the supply-side of the economy is modified to allow for endogenous growth, as we shall see in the next section.

Endogenous growth and intergenerational transfers

The endogenous growth model we analyse is based on learning-by-doing externalities in the accumulation of capital, consistent with Sheshinski (1967) (see also Buiter (1993); Barro and Sala-i-Martin (1995), sect. 4.3). Each firm faces decreasing returns to scale with respect to its own input of capital: however, firm productivity is an increasing function of the average capital–labour ratio in the economy, which acts as a proxy for the level of technical knowledge in the economy and of the aggregate externalities from capital accumulation. Hence, the aggregate production function of the economy exhibits constant returns to an aggregate measure of capital.

The constant-returns-to-scale production function for firm i can thus be written as:

$$Y_{it} = F(K_{it}, E_{it}) = E_{it}\, f(k_{it}) \tag{3.9}$$

where $k_{it} \equiv K_{it}/E_{it}$ is capital per efficiency unit of labour, and where the efficiency variable E_{it} is defined as:

$$E_{it} = \frac{K_t}{L_t} \cdot L_{it} \tag{3.10}$$

where K_t and L_t are aggregate capital and labour, respectively. The wage of a unit of raw labour is now:

$$\tilde{w}_t = \frac{K_t}{L_t} \cdot w_t \tag{3.11}$$

The measure of the set of firms is normalized to unity. In equilibrium, $k_t = 1$ by symmetry. The economy-wide production function therefore becomes.

$$Y_t = K_t\, f(1) \equiv \alpha K_t \tag{3.12}$$

where $\alpha \equiv f(1) > 0$ is the marginal social rate of return to capital. Using $k_t = 1$, factor prices in Equations (3.5) and (3.6) become:

$$r = f'(1) \equiv \alpha' \qquad 0 < \alpha' < \alpha \tag{3.13a}$$

$$\tilde{w}_t = (\alpha - r)\frac{K_t}{L_t} \tag{3.13b}$$

The constant rate of interest r measures the marginal private product of capital, whereas the parameter α measures the social returns from capital accumulation. Following the introduction of the balanced-budget transfer scheme, consumption by the different generations is still described by equations similar to Equations (3.7a)–(3.7c), where, however, the wage w_t is replaced by $\tilde{w}_t$ as given by Equation (3.13b).

In order to assess the effects of the introduction of social security schemes, one must no longer concentrate exclusively on the implications for the long-run growth rate. Transfer programmes affect the profile of consumption and saving over time. Since such transfers modify the pattern of accumulation over time, and such modifications have *permanent* effects under endogenous growth, it becomes necessary to consider the cumulative impact of social security on both capital stock and output.

The change in capital stock is given by:

$$\Delta K_t \equiv K_t - K_{t-1} = Y_t - C_t \tag{3.14}$$

If we define the growth rate of the capital stock in period t as $g_t \equiv \Delta K_t / K_t - 1$, then g_t also denotes the rate of growth of the economy (from Equation (3.12)). In the absence of social security, $\beta = 0$ and Equation (3.14) gives (see Marini and Scaramozzino, 1996b):

$$g_t = g = \frac{r(2 + r + \rho - \alpha)}{2 + \rho(1-r) - \alpha - r} \tag{3.15}$$

The growth rate g declines with the subjective rate of time preference ρ, and increases with the rate of interest r and with the marginal social product of capital α. Interestingly, the rate of growth of output is independent of the rate of population growth: without a transfer scheme there are no externalities on the saving rate across generations. Population growth therefore exerts no influence on capital accumulation, and thus on growth. Furthermore, $g > r$: the growth rate in the absence of social security is greater than the marginal private product of capital (see also Saint-Paul, 1992).

When a social security programme is introduced, $\beta > 0$ and the level of capital stock is given by:

$$K_{t+s} = (1 + g)^s \left[K_{t-1} + (x_s - z)\, L_{t-1}\right] \qquad i = 0, 1, 2, \ldots \tag{3.16}$$

where

$$z = \beta \cdot \frac{2 + \rho + r + n + n\rho}{\left[(2 + \rho) - (1 + r)(\alpha - r)\right](1 + r)}$$

$$x_s = \beta(r - n) \cdot \frac{(2 + r + n + \rho + n\rho)[2 + \rho(1 - r) - \alpha - r]}{[2 + \rho - (1 + r)(\alpha - r)]^{s+1}(1 + r)}$$

$$\left[4 + 2\rho - 2\alpha - r(\alpha + \rho - r) + n(2 + \rho - r\rho - \alpha - r)\right]^{s-1} \quad i = 0,\ 1,\ 2, \ldots$$

From Equation (3.16), one can see that capital stock in the presence of positive social security transfers is greater than in their absence, if and only if $x > z$. The dynamics of output over time can be analysed by considering its growth rate:

$$g_t = \frac{1}{2 + \rho(1 - r) - \alpha - r}\left\{r(2 + r + \rho - \alpha) - \beta\left[\frac{2 + r + \rho + n(1 + \rho)}{1 + r}\right]\frac{L_{t-1}}{K_{t-1}}\right\} \tag{3.17a}$$

$$g_{t+1} = \frac{1}{2+\rho(1-r)-\alpha-r}\left\{r(2+r+\rho-\alpha)+\beta\,(r-n)\frac{(1+r)+(1+n)(1+\rho)}{(1+r)(1+n)}\frac{L_{t+i-1}}{K_{t+i-1}}\right\}\; i=1,\ 2,\ldots \tag{3.17b}$$

From inspection of the growth rates in Equations (3.17a) and (3.17b) one can see that the growth rate of the economy will fall, on impact, when the social security programme is introduced. The transfer will result in greater consumption by the old and a fall in the rate of capital accumulation and output. However, the long-run effects of social security on capital accumulation depend on its cumulative effects on the capital stock and output. If $r > n$, the growth rate of the economy from period $t + 1$ onwards will overshoot its long-run equilibrium value, which is given by Equation (3.15). Since $g > r > n$, the growth rate will thereafter converge monotonically towards the value g. The cumulative effect of social security could well be positive if the higher growth rate along the transitional dynamic path outweighs the impact fall in growth.

A greater social return to capital accumulation, as measured by the parameter α, makes it more likely that the introduction of the social transfer programme increases long-run capital and output. The intuition for this result is as follows. When introduced, the social security programme reduces savings. Under endogenous growth, this leads to slower capital accumulation. However, as the scheme is implemented in future periods, savings increase, provided $r > n$. This will result in accelerated accumulation and growth. The cumulative effect could well be positive, so that social security can indeed lead to greater output in the long run.

Conclusions

Unfunded social security transfers are not necessarily in conflict with growth. The conventional argument against the welfare state is based on presumptions which may not be valid. When the saving behaviour of agents is allowed to affect the rate of growth of the economy, the effects of a transfer programme can be qualitatively different from the case in which only the composition of output is affected.

As pointed out by Atkinson (1995), the verdict on the undesirability of social security has perhaps been reached too precipitously. This chapter provides additional reasons as to why 'the jury should still stay out'. In particular, when the supply-side dynamics are modelled along the lines suggested by the endogenous growth literature, it is possible that properly designed, unfunded schemes would promote growth. The reduction in savings when an intergenerational transfer programme is introduced has a negative impact on capital accumulation and output. However, in the long run, the saving rate will be greater than in the absence of social security, and this could lead to increased well-being for future generations.

The main policy implication to be derived from our theoretical analysis is that social security may indeed contribute positively to long-run capital accumulation. The terms of the trade-off between present and future generations are thus very different from what is usually assumed. The current generation always benefits from the introduction of a transfer programme. In conventional analysis, all future generations will suffer (unless the economy is dynamically inefficient). By contrast, according to the analysis in this chapter, a number of intermediate generations might still suffer in terms of lower capital and output, but in the long run future generations might be better off. Clearly, this result has important implications for the debate on equity versus efficiency of the social security system.

Rather than dismantle the welfare state altogether, the relevant policy issue might be to find an optimally designed benefit ratio which is conducive to a positive effect on growth in the long run.

References

Aaron, H. (1966) 'The Social Insurance Paradox', *Canadian Journal of Economics and Political Science*, vol. 32, no. 3, August, pp. 371–4.

Atkinson, A. (1987) 'Income Maintenance and Social Insurance', in A. Auerbach and M. S. Feldstein (eds), *Handbook of Public Economics*, vol. II (Amsterdam: North-Holland).

Atkinson, A. B. (1995) 'Is the Welfare State Necessarily an Obstacle to Economic Growth?', *European Economic Review*, vol. 39, nos 3/4, April, pp. 723–30.

Azariadis, C. (1993) *Intertemporal Macroeconomics* (Oxford: Basil Blackwell).

Barro, R. J. and Sala-i-Martin, X. (1995) *Economic Growth* (New York: McGraw-Hill).

Blanchard, O. J. and Fischer, S. (1989) *Lectures on Macroeconomics* (Cambridge, MA: MIT Press).

Buchanan, J. (1975) 'The Samaritan's Dilemma', in E. Phelps (ed.), *Altruism, Morality, and Economic Theory* (New York: Russell Sage), pp. 71–85.

Butter, W. H. (1993) 'Saving and Endogenous Growth: A Survey of Theory and Policy', in A. Heertje (ed.), *World Savings, Theory and Policy* (Oxford: Basil Blackwell), pp. 64–99.

Diamond, P. (1965) 'National Debt in a Neoclassical Growth Model', *American Economic Review*, vol. 55, pp. 1126–50.

Diamond, P. (1977) 'A Framework for Social Security Analysis', *Journal of Public Economics*, vol. 8, pp. 275–98.
Feldstein, M. (1985) 'The Optimal Level of Social Security Benefits', *Quarterly Journal of Economics*, vol. 100, no. 2, May, pp. 303–20.
Feldstein, M. (1995) 'Fiscal Policies, Capital Formation, and Capitalism' *European Economic Review*, vol. 39, nos. 3/4, April, pp. 399–420.
Gordon, R. H. and Varian, H. (1988) 'Intergenerational Risk Sharing', *Journal of Public Economics*, vol. 37, pp. 185–202.
Hansson, I. and Stuart, C. (1989) 'Social Security as Trade Among Living Generations', *American Economic Review*, vol. 79, no. 5, December, pp. 1182–95.
Marchand, M. and Pestieau, P. (1991) 'Public Pensions: Choices for the Future', *European Economic Review*, vol. 35, pp. 441–53.
Marini, G. and Scaramozzino, P. (1996a) 'Social Security and Intergenerational Equity', Centre for International Studies on Economic Growth (CEIS), University of Rome 'Tor Vergata' Working Paper No. 7.
Marini, G. and Scaramozzino, P. (1996b) 'Endogenous Growth and Social Security', Centre for International Studies on Economic Growth (CEIS), University of Rome 'Tor Vergata', Working Paper No. 8.
OECD (1995) 'Effects of Ageing Populations on Government Budgets', *OECD Economic Outlook*, vol. 57, June, pp. 33–41.
Saint-Paul, G. (1992) 'Fiscal Policy in an Endogenous Growth Model', *Quarterly Journal of Economics*, vol. 107, November, pp. 1243–59.
Samuelson, P. (1958) 'An Exact Consumption–Loan Model with or without the Social Contrivance of Money', *Journal of Political Economy*, vol. 66, December, pp. 467–82.
Samuelson, P. (1969) 'Lifetime Portfolio Selection by Dynamic Stochastic Programming', *Review of Economics and Statistics*, vol. 51, no. 3, August, pp. 239–46.
Samuelson, P. (1975) 'Optimum Social Security in a Life Cycle Growth Model', *International Economic Review*, vol. 16, pp. 539–44.
Sheshinski, E. (1967) 'Optimal Accumulation with Learning by Doing', in K. Shell (ed.), *Essays on the Theory of Optimal Economic Growth* (Cambridge, MA: MIT Press), pp. 31–52.
Veall, M. R. (1986) 'Public Pensions as Optimal Social Contracts', *Journal of Public Economics*, vol. 31, pp. 237–51.

Part II

Research and Growth

4
Human Capital, Ideas and Economic Growth

Charles I. Jones

Introduction

This chapter develops and analyses empirically a simple model of human capital, ideas and economic growth that integrates contributions from several different strands of the growth literature. These strands, and a discussion of what I try to emphasize in the paper, are outlined below:

- *Romer (1990) and the research-based new growth theory*. Recent advances in new growth theory emphasize the importance of ideas, non-rivalry and imperfect competition for understanding the engine of economic growth. Romer (1993) argues that these issues may also be important for understanding economic development. Nelson and Phelps (1966) provide a way of thinking about technology transfer that incorporates both human capital and advantages to 'backwardness'.
- *Mankiw et al. (1992)* (MRW). MRW show that a simple neoclassical model can explain up to 80 per cent of the cross-country variation in the log of per capita GDP, especially if it incorporates differences in human capital investment across countries.
- *Barro and Lee (1993) and Bils and Klenow (1996)*. Barro and Lee provide an extensive panel data set on educational attainment for a large number of countries. Bils and Klenow argue for including educational attainment in a model in a way that is consistent with Mincerian wage regressions.
- *Benhabib and Spiegel (1994), Islam (1995), Pritchett (1996), and Judson (1996)*. These papers document in various ways a puzzle involving the relationship between human capital and economic growth. The

> puzzle appears when one looks at a growth-accounting approach that involves variables, such as the Barro and Lee (1993) human capital stocks. In either simple or multivariate regressions of the growth rate output on the growth rate of the human capital stock, the human capital stock appears with a negative coefficient.

Weitzman (1996) suggests that a useful analogy for understanding the research process is a child's chemistry set: research proceeds by taking various elements (various ideas) and joining them together. Most combinations are useless, but a few combinations are extremely valuable. In this chapter, I consider the various elements of the growth literature just outlined, and combine them together in a particular – and, we hope, valuable! – way.

Several insights emerge from this combination. First, even though the model emphasizes the importance of ideas and research, one can derive an empirical specification from the model that is nearly identical to the regression estimated by Mankiw *et al.* (1992). The MRW level regressions are a very useful way to organize one's thinking about why different countries achieve different levels of income, but the specification says very little, I think, about the importance of a 'neoclassical' growth model versus a growth model based on imperfect competition and ideas.[1]

Second, many authors have interpreted the Barro and Lee (1993) data on educational attainment as measuring the stock of human capital per person in an economy. In the model presented here, the most natural interpretation of the Barro and Lee educational attainment data is as something like a rate of investment in human capital rather than as a human capital stock. More precisely, these data correspond to the fraction of an individual's time endowment that is spent accumulating skills. Unlike the physical capital stock or the capital stock per person, this variable is constant along a balanced growth path. This has implications for how these data are used in growth-accounting exercises.

Third, the empirical estimation of the level regression derived from the model verifies several of the results found by Mankiw *et al.* (1992). However, the fit of the model is far from perfect, suggesting that an important feature of the technology transfer process is not captured by the model that is presented. I discuss some avenues for future research that I am pursuing in order to address this issue.

Finally, the set-up considered here provides one possible resolution of the human capital puzzle mentioned earlier. In particular, the formulation of the model suggests that it is not the growth rate of the educational attainment variable that belongs in the specification, but rather the change in the level. Regressions that follow this approach look remarkably similar to the MRW-style level regressions in which the educational attainment variables show up strongly.

The paper is organized as follows. The second section presents the basic model, integrating several strands of the growth literature. The third section considers the empirical applications of the model, and the fourth section concludes.

The model

Production

Three kinds of goods are produced in the economy: a consumption good ('output'), a human capital good ('experience' or 'skill'), and new varieties of intermediate capital goods ('ideas').[2]

Output Y is produced by competitive firms using labour L_Y and a collection of intermediate capital goods x_i. The amount of human capital per person in the firm determines the range of intermediate capital goods that the firm can use. That is, human capital in this model is interpreted as skill or experience in using advanced intermediate goods. The production function for a firm employing workers of average skill h is:

$$Y(t) = L_y(t)^{1-\alpha} \int_0^{h(t)} x_i(t)^{\alpha} \, di \tag{4.1}$$

where $0 < \alpha < 1$, so that a firm with skill level h faces constant returns to scale in production. This kind of specification differs from that used in Romer (1990), in that the range of goods that can be used by a firm has both a non-rivalrous and a rivalrous component: the intermediate good must have been invented, and the workers in the firm must have learned to use the intermediate good. Because there are constant returns to scale, given h, and because individuals in this economy are identical, we can focus on a single, competitive representative firm.

As an alternative to producing output, individuals can spend their time acquiring skills. That is, they can learn to use more advanced intermediate capital goods. Activities such as on-the-job training, education and apprenticeships are all examples of skill acquisition. Individuals accumulate human capital according to:

$$h(t) = \mu e^{\theta u(t)} h(t) \left(\frac{A(t)}{h(t)} \right)^{\gamma} \tag{4.2}$$

In this equation, $u(t)$ is the fraction of an individual's labour endowment spent accumulating human capital, μ is an arbitrary positive constant, and $A(t)$ represents the technological frontier; that is, the total measure of intermediate goods that have been invented to date.

Equation (4.2) can be motivated in several ways. The equation is similar to the specification employed by Lucas (1988), particularly if the last term is ignored. Lucas favoured a specification that was linear in h so that the model generated endogenous growth. The least term of Equation (4.2) imposes curvature on the model, rendering it less than linear in h. With $\gamma > 0$, the equation incorporates an advantage to 'backwardness', as in Gerschenkron (1952). The curvature implies that it is easier to learn to use intermediate goods that are further from the frontier; goods close to the frontier are harder to master. More generally, the notion that time spent acquiring skills and 'backwardness' interact to affect the level of productivity in an economy dates back at least to Nelson and Phelps (1966).

Another motivation for the specification in Equation (4.2) is that it is consistent with microeconomic evidence on the relationship between wages and schooling or experience. According to Mincer (1974), an additional year of schooling or an additional year of experience should increase wages proportionally. That is, the relationship between wages and schooling or experience is a semi-log form. Equation (4.2) shares this property, as we will see shortly. Bils and Klenow (1996) emphasize this microeconomic regularity in building a model of human capital and growth.[3]

In order to enable individuals to learn to use an intermediate good, the design for the intermediate good must have been invented. In thinking about the production of 'ideas' in this economy, it is useful for the moment to interpret the model as one of a large, advanced closed economy. Later, we shall discuss the model's implications for idea flows across countries. The production function for ideas is given by:

$$\begin{aligned} A(t) &= \bar{\delta} h\,(t)^{\beta}\, L_A(y) \\ &\equiv \delta h(t)^{\beta} L_A(t) A(t) \varphi \end{aligned} \tag{4.3}$$

This production function follows the modification in Jones (1995) of the Romer (1990) specification. Units of labour L_A produce ideas based on their skill, with elasticity $\beta > 0$. The productivity of a skill-adjusted unit of labour, δ, is an increasing function of the existing stock of ideas ($\varphi > 0$). This incorporates an intertemporal knowledge spillover into the model.[4]

A feature of this equation not in Jones (1995) is that the skills of individuals augment their ability to produce ideas, apart from knowledge spillovers. One can interpret the difference as follows: h^β captures the effects of past knowledge on future production of ideas that can be 'internalized', while A^φ captures the knowledge spillovers that are external in society. The effect of education increasing an individual's abilities, either in research or in the production of output, is potentially internalized either by markets or by forward-looking individuals. On the other hand, the invention of the laser and just-in-time production presumably generate spillovers into future research that the inventors are unable to capture.

Factor accumulation

Capital K is accumulated by forgoing consumption and is measured in units of the output good:

$$K(t) = s_K(t)Y(t) - dK(t) \tag{4.4}$$

where s_k is the investment share of output (the rest going to consumption) and $d > 0$ is some constant exponential rate of depreciation.

Units of an intermediate capital good x_i are created one-for-one with units of raw capital. To simplify the set-up, we assume this transformation is effortless and can also be undone effortlessly. Thus,

$$\int_0^{h(t)} x_i(t)di = K(t) \tag{4.5}$$

Intermediate goods are treated symmetrically throughout the model, so that $x_i(t) = x(t)$ for all i. This fact, together with Equation (4.5) and the production function in Equation (4.1) implies that the aggregate production technology for this economy takes the familiar Cobb–Douglas form:

$$Y = K^\alpha(hL_Y)^{1-\alpha} \tag{4.6}$$

where we have suppressed time subscripts (which we will continue to do when the meaning is clear).

The fundamental factor of production in this model is labour, and we have already described its various uses. The total quantity of labour in the economy is given by $L(t)$, which is assumed to grow exogenously at rate $n > 0$.[5] The labour market clearing condition is:

$$I_Y + L_h + L_A \tag{4.7}$$

To summarize, the structure of the model is as follows. An individual accumulates skills h which represent the range of intermediate goods that the individual has learnt to use. The individual then spends time either producing the consumption/capital good Y, accumulating additional skills, or searching for new designs of intermediate goods. Individuals accumulate capital to smooth consumption, and the population of the economy grows exogenously at rate n.

The allocation of resources

The resource allocation decisions in this economy involve the allocation of labour over time, and the division of output into consumption and investment over time. Romer (1990) describes how the market can be used to allocate resources in this economy only in the presence of imperfect competition, and a similar approach could be taken here. Intermediate goods firms own the exclusive rights to sell their particular varieties and operate in a monopolistically competitive environment. Researchers prospect for new ideas and are rewarded for their efforts with the present discounted value of the flow of profits that can be earned in the intermediate goods sector.

An allocation decision that is not present in Romer (1990) is the decision of how much time to spend learning to produce with new varieties of intermediate capital goods. One can model this decision as being taken by forward-looking individuals who either recognize or do not recognize that learning to use a new variety has dynamic effects on future skill acquisition. In the second case, the amount of time spent accumulating skills will typically by suboptimal.

In this chapter, we choose not to spend additional time developing the market allocation of resources. Instead, we shall assume that these allocations – that is, S_K, u, L_A/L, and L_Y/L – are given exogenously. This can be justified, for example, by appealing to taxes and institutions outside the model that impinge on the forward-looking set-up to deliver allocations that are (at least asymptotically) constant. We take these allocations as given, and then ask what the steady state of the model looks like.

Steady state analysis

The steady state of the model is most easily described by considering the production function for ideas. Rewriting Equation (4.3) in terms of growth rates,

$$\frac{A}{A} = \delta \left(\frac{h}{A} \right)^{\beta} \frac{L_A}{A^{1-\beta-\phi}} \tag{4.8}$$

In a steady state, the growth rates of A and h are constant and equal.[6] Therefore, the ratio h/A is constant, and a balanced growth path requires the numerator and the denominator of the last term in Equation (4.8) to grow at the same rate. Therefore,

$$g_A = \frac{n}{1-\beta-\phi} \tag{4.9}$$

where the notation g_x will be used to denote the constant growth rate of placeholder x, and we have used the fact that the share of labour devoted to research is constant.

Equation (4.9) is the human-capital-augmented version of a result in Jones (1995). A balanced growth path for the model with a growing population exists only if $\beta + \varphi < 1$. This condition implies that the differential equation governing the production of ideas is less than linear, and leads to a 'semiendogenous' growth model. Although technological progress is endogenized, the model exhibits no long-run per capita growth unless the population is growing over time.

Analysis of the production function in Equation (4.6) and the capital accumulation equation (4.4) reveals that along the balanced growth path:

$$g_y = g_k = g_h = g_A \equiv g \tag{4.10}$$

where $y \equiv Y/L_Y$ and $k \equiv K/L_Y$.[7] Because of the labour-augmenting nature of technological change in the model, per capita (or per worker) growth rates are all equal to the rate of technological progress.

Further analysis of these equations allows us to solve for the level of output per worker in the final goods sector:

$$y^*(t) = \left(\frac{s_K}{n+g+d}\right)^{\alpha/(1-\alpha)} \left(\frac{h}{A}\right)^* A^*(t) \tag{4.11}$$

where we have used the superscript asterisk (*) to denote the balanced growth path.

Moreover, from the production function for skills in Equation (4.2), the ratio of skills to ideas along a balanced growth path is given by:[8]

$$\left(\frac{h}{A}\right)^* = \left(\frac{\mu}{g} e^{\theta u}\right)^{1/\gamma} \tag{4.12}$$

Combining these last two equations,

$$y^*(t) = \left(\frac{s_K}{n+g+d}\right)^{\alpha/(1-\alpha)} \left(\frac{\mu}{g} e^{\theta u}\right)^{1/\gamma} \tag{4.13}$$

This last equation makes clear the appeal of entering time spent accumulating skill in an exponential form. Increases in the level of u will have proportional effects on labour productivity and wages, thus matching evidence from Mincerian wage regressions.

Comparisons to previous work

Jones (1995) and Romer (1990)

The model illustrates how one can add human capital to the model of Jones (1995) without changing the basic results. In the extended model, standard policies such as investment tax credits, subsidies to research and development (R&D), or subsidies to skill acquisition – at last to the extent that we think of them as permanent increases in the rate of investment, the share of labour devoted to R&D, or the amount of time spent accumulating skills – have *level* effects but no long-run growth effects in the model. This results from fundamental lack of linearity in the production equation for ideas, and, as in Jones (1995), this lack of linearity is a necessary condition for the existence of a balanced growth path in the presence of population growth.

This result can be overturned, but only by making arbitrary assumptions about the strength of externalities in appropriate places. For example, one could set $\gamma = 0$ and have human capital be an input into the production of ideas but not into the production of output (so that designs can be used immediately after they are created). Segerstrom (1995) follows this approach, and the linearity of the human capital equation generates endogenous growth, as in Lucas (1988). However, the linearity of the human capital accumulation equation is then somewhat arbitrary, and the endogenous growth arises from human capital accumulation, not from research.

Jovanovic (1995) and 'scale effects'

Jovanovic (1995) emphasizes the importance of adoption costs relative to research costs. He argues that if there are costs proportional to the size of the population that must be paid in order to implement ideas, then these costs will swamp asymptotically any fixed cost of creating the ideas. This approach might call into question the significance of 'scale effects' and the importance of thinking about the non-rivalrous nature of ideas.

The model in this chapter incorporates both an adoption cost (skill acquisition) and the non-rivalrous nature of ideas. However, the model still contains a 'scale effect': and it is still important to recognize the non-rivalrous nature of ideas.

Consider, first, the issue of 'scale effects'. Suppose there are two nearly-identical economies of the kind described in this chapter. The economies are not allowed to interact or share ideas (for example, they are on opposite sides of the universe). The only difference between the two economies is that one has a much larger population than the other. Starting from the same initial conditions, it is obvious that the larger economy must grow more rapidly in the short run, and this transition effect leads the larger economy to be richer in the long run, when both economies are growing at the same rate. This can be seen most easily by considering the production function for ideas in Equation (4.3).

Is this kind of 'scale effect' relevant to the countries of the planet Earth? Clearly, an issue that complicates matters is the fact that countries in the world share ideas. Moreover, casual empiricism suggests that scale is at most one among many important factors. For example, China is much poorer than Hong Kong. In the next section, I will discuss interpreting this model in the context of a multi-country setting, and the particular version considered will not exhibit scale effects. However, a more detailed model in Jones (1996) still exhibits scale effects in a multi-country setting.

What about the importance of the non-rivalrous nature of ideas – does this become negligible in the presence of adoption costs that grow with the population? The answer to this question is surely in the negative, and the argument follows Romer (1990). With a non-rival input, all factors cannot be paid their marginal product, so that imperfect competition must be introduced into the model. Intermediate goods will be priced above marginal costs, with a mark-up that depends, for example, on the elasticity of substitution between intermediate goods. None of this changes as a result of adding another rivalrous factor (the training of labour) to the model.

Mankiw et al. (1992)

The result derived in Equation (4.13) is very similar to the result derived by Mankiw *et al.* (1992) for the human-capital-augmented Solow model. A country is richer along its balanced growth path the higher is its investment rate in physical capital S_K, the higher is its investment rate in human capital u, the lower is its rate of population growth n, and the higher is its level of technology A.

However, the model underlying this result is very different. The MRW approach builds on a Solow model with exogenous technological progress. There is no research, no non-rivalry, no imperfect competition, and no learning to use newly-invented technologies. This suggests that macro evidence of the kind presented in their paper cannot distinguish between a 'neoclassical' growth model and an R&D-based growth model. Additional evidence must be brought to bear in order to make this distinction.

Interpreting the model with many countries

Up to now, we have been interpreting the model primarily as one of a large, advanced, closed economy that grows by pushing out its technological frontier. In order to apply this model to a cross-section of countries, we must discuss the imprint issues of how ideas flow between economies, and which economies decide to engage in research. To push our model as far as possible, we shall make another simplifying assumption, motivated in part by what we have already developed. We assume that the world consists of a large number of relatively small economies. This is really the opposite of the assumption we have maintained so far, so it allows us to explore a different extreme. The economies will be small in the sense that the effect of an individual economy's research on the state of the world technological frontier is small, and in fact we shall ignore this effect empirically. From an individual economy's perspective, the world's technological frontier is expanding exogenously at rate $g \equiv g_A g$ $\equiv g_A$, given by Equation (4.9). We also assume that the amount of research undertaken in any single economy is small.

Under this assumption, the skill-acquisition Equation (4.2) becomes a technology transfer equation. In order to use a technology that has been invented somewhere in the world, a country must learn the skills associated with that technology. Although from the standpoint of the *invention* of ideas, knowledge is non-rivalrous but partially excludable, from the international standpoint of technology transfer, it may be a useful starting point to assume that technology is a public good. That is, a developing country sees a new technology being used and that technology is like a public good. Provided the developing country can learn to use the technology, it need not pay for its invention.[9]

In the next section, we shall discuss the empirical implications of the model. One can interpret the empirical results as describing how far the simplifying assumptions made in this model can go in terms of explaining the cross-sectional distribution of income.

Empirical applications

Three empirical applications of the model are considered. First, we address an important question of interpretation that has been overlooked by the empirical growth literature. The question is how to map data on educational attainment into our growth models. Second, we consider the empirical estimation of Equation (4.13), as in Mankiw *et al.* (1992), emphasizing our underlying model's focus on research and technology transfer. Finally, we analyse a question raised recently in empirical growth literature concerning the relationship between human capital and growth. As phrased in the title of Pritchett (1996), 'Where has all the education gone?'

Years of schooling: stocks or flows?

How to measure human capital has been one of the difficult questions faced by the empirical growth literature. Various authors have employed data on literacy rates, school enrolment rates, and public expenditure on education. In the 1990s, however, Barro and Lee (1993) assembled data on average educational attainment (that is, years of schooling) per adult in the population for a large number of countries at five-year intervals going back to 1960.[10] This data has been used in a number of later studies, including Islam (1995), Barro (1996), Pritchett (1996) and Judson (1996). In these studies, the practice has been to interpret the average educational attainment data as a measure of the stock of human capital per person in the economy.[11] This practice presumably is carried over from the labour economics literature in which individuals accumulate 'stocks' of human capital that augment their wages for a lifetime.

From the standpoint of the macroeconomic analysis of income and growth, however, we believe this interpretation is incorrect. Instead, the educational attainment data is interpreted more appropriately as a flow variable similar to an investment rate rather than a capital stock. Educational attainment per person is thought of plausibly as a constant, at least asymptotically. For example, one might think that average educational attainment in the USA will level off eventually at something like fourteen years of schooling per person. In contrast, the physical capital stock per person grows over time (for example, because of technological progress). The most natural mapping of the educational attainment data into models of economic growth is as the time an individual spends accumulating human capital. Taken as a fraction of an individual's total time endowment, this data corresponds to the variable u in the model outlined in the previous section.

It is difficult to judge how much of a problem interpreting educational attainment as a stock of human capital is in the empirical growth literature. In cross-country growth regressions such as Barro

(1996), it is plausible to reinterpret the log of average educational attainment as (the log of) an investment rate. The regression variable then proxies (perhaps with other variables) for the steady-state level of income as in Mankiw *et al.* (1992), and makes sense in terms of 'conditional convergence'. On the other hand, in growth-accounting regressions such as those employed by Benhabib and Spiegel (1994) and Pritchett (1996) the interpretation may be more difficult. In these papers, the estimation is motivated by log-differentiating the production function; that is, output growth is regressed on the growth rates of the physical capital stock, the human capital stock, and the labour force. Asymptotically, however, human capital stock should stop growing if it is measured by the average educational attainment of the labour force, and it is unclear how to interpret the results of the regression in this context.

Level regressions

With this as a background, we can now proceed to estimating Equation (4.13). First, however, consider the equation in logarithmic form:

$$\log\ y^*(\mathrm{t}) = \log A^{*(t)} + \frac{1}{\gamma}\log\frac{\mu}{g} + \frac{\alpha}{1-\alpha}\log s_K - \frac{\alpha}{1-\alpha}\log(n+g+d) + \frac{\theta}{\gamma}u \qquad (4.14)$$

As specified, the equation does not contain an error term. We shall introduce one in two ways. First, we shall assume that all countries are on their steady-state balanced growth paths. To the extent that they are not, this will be captured by an error term. Second, according to the model, all differences in labour productivity are accounted for by physical investment rates, population growth rates, and time spent learning about the technologies available in the world. To the extent that the model is misspecified, we shall find large residuals. We shall exploit this argument below as a 'test' of the model.

In general, there is no reason to suppose that these sources of the error term are uncorrelated with the variables on the right-hand side of Equation (4.14). However, we shall proceed with ordinary least squares to see what kind of relationships the data and the model together suggest.

The data used to estimate Equation (4.14) are taken primarily from the Penn World Tables Mark 5.6 of the Summers and Heston (1991) data set. For log y^* we use the log of real GDP per worker for 1990.[12] For S_K we use the average investment rate from 1980 to 1990.

To measure u, one would ideally prefer a measure that includes on-the-job training as well as time spent in the formal education sector. However, this data does not seem to be available for a large number of countries. Therefore, we measure u using the average educational attainment variable from Barro and Lee (1993), including primary, secondary and tertiary education. Data is reported at five-year intervals from 1960 to 1985, and we use the average of the 1980 and 1985 observations.

To compute u, one needs to divide educational attainment by the average time endowment of individuals in years. Instead of picking an arbitrary number, we instead simply use educational attainment as the independent variable, so that the average time endowment is included in the coefficient.[13] We shall use the notation N for the average educational attainment in years.

Table 4.1 reports the results of estimating Equation (4.4). As the regression is very similar to the one implemented by Mankiw *et al*. (1992), the basic results are familiar. In a large sample of countries, a simple specification involving physical investment rates, population growth rates and a human capital investment rate can explain a large fraction of the variation in (log) output per worker across countries. Here, the $\bar{R}^2$ is 0.713. In addition, the estimate of α, the elasticity of the production function with respect to physical capital, is 0.344, which agrees quite well with evidence from income shares and other empirical studies.

Interpreting the coefficient on u is more complicated. N terms of the parameters of the model, the coefficient is θ/γ divided by the average time endowment (lifetime) of an individual. A more direct interpretation is the econometric one: an increase in average educational attainment of one year raises output per worker by approximately 20 per

Table 4.1 Level regression, 1990

Variable	Unconstrained		Constrained	
Constant	7.402	(1.294)	7.785	(0.135)
log s_K	0.519	(0.118)	–	–
log $(n + g + d)$	−0.688	(0.567)	–	–
log $sK/(n + g + d)$	–	–	0.525	(0.118)
N	0.191	(0.031)	0.195	(0.029)
α	–	–	0.344	(0.051)
p-value	–	–	0.76 –	
$\bar{R}^2$	0.710	–	0.713	–

Note: Num Obs = 90. The p-value corresponds to the test of whether or not the coefficients on log s_K and log $(n + g + d)$ are the same. White heteroskedasticity-robust standard errors are reported in parentheses.

cent. In terms of standard deviations, a one standard deviation increase in average educational attainment is associated with an increase in log *Y/L* of 0.56 standard deviations. Finally, the coefficient suggests that, if Cameroon were to increase its educational attainment from 2.00 years per person to the US level of 11.84 years per person, its output per worker would rise from $2490 to $16,963 (compared to a US level of $36,754). Of course, these numbers are only meant to be suggestive, as the causality of the relationship is not firmly established, but clearly the educational attainment variable is economically as well as statistically significant.

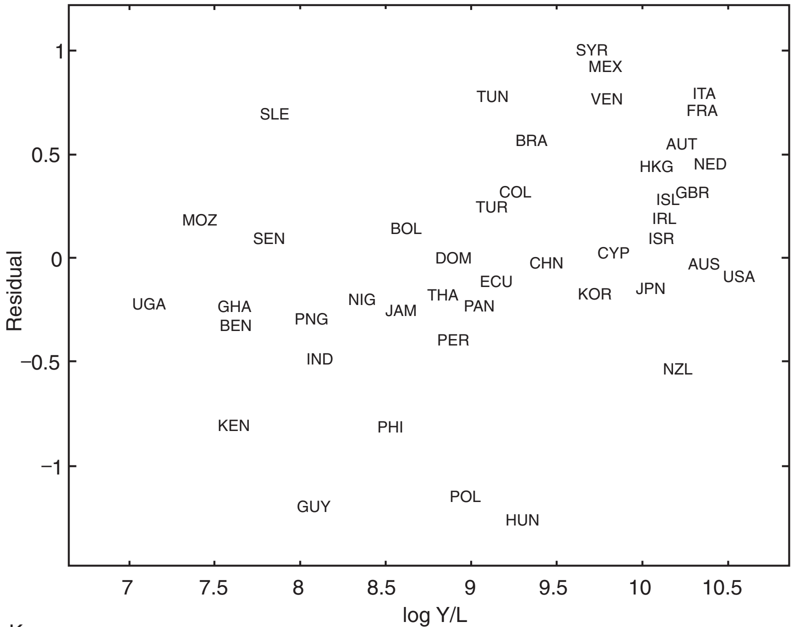
Another way to analyse these results is to think of them as a test of the model proposed in the previous section. To the extent that we have accounted successfully for the important sources of income differentials across countries, the residuals from this estimation should be small. The $\bar{R}^2$ of 0.713 is somewhat favourable, but it masks important differences in residuals across countries. These residuals are plotted in Figure 4.1.

The general upward slope in the figure suggests that rich countries are richer than the model would predict, and poor countries are poorer than the model would predict. In other words, there is a systematic difference in incomes across countries that the model does not capture. To see the magnitude of this difference, notice that the residual varies from about -1 for poor countries to about $+1$ for rich countries. That is, it is not uncommon to find countries that are either 2.7 times poorer or 2.7 times richer than the model would predict.

This suggests that, while capturing significant differences in income across countries, the model still omits important determinants.[14] An avenue explored in Jones (1996) is that circumstances beyond learning to use new technologies affect whether new ideas are implemented. In particular, ideas are likely to be put in place only when an investor expects to earn a sufficiently large profit on the idea. Even in a society in which educational attainment is fairly high, if entrepreneurs are not allowed to capture rents from their efforts, advantage may not be taken of new ideas. It remains to be seen whether a model that incorporates these additional effects can make progress in explaining the cross-sectional distribution of income.

Where has all the education gone?

Several studies in the empirical growth literature have emphasized the following finding: although levels of various measures of human capital have explanatory power in growth regressions, the *growth rate* of the stock



Key:

AUS Australia
AUT Austria
BEN Benin
BOL Bolivia
BRA Brazil
CHN China
COL Colombia
CYP Cyprus
DOM Dominican Republic
ECU Ecuador
FRA France
GHA Ghana
GUY Guyana
HKG Hong Kong
HUN Hungary
ISL Iceland
IND India
IRL Ireland
ISR Israel
ITA Italy
JAM Jamaica
KEN Kenya
KOR Korea, South
MEX Mexico
MOZ Mozambique
NED Netherlands
NZL New Zealand
NIG Niger
PAN Panama
PNG Papua New Guinea
PER Peru
PHI Philippines
POL Poland
SEN Senegal
SLE Sierra Leone
SYR Syria
THA Thailand
TUN Tunisia
TUR Turkey
UGA Uganda
GBR United Kingdom (Great Britain)
USA United States
VEN Venezuela

Figure 4.1 Residuals from Equation (4.14) versus log *Y*/*L*

of human capital has very little explanatory power and often enters regressions negatively instead of positively. These studies include Benhabib and Spiegel (1994), Islam (1995) and Pritchett (1996). This leads the authors to ask, quite naturally, why the countries that have increased their human capital more rapidly have not performed better. Why haven't these investments paid off in the aggregate? Benhabib and Spiegel (1994) and Islam (1995) answer the question by arguing for a different empirical specification, one in which the *level* of human capital enters instead of the growth rate of the human capital stocks. However, in many ways, this simply ignores the problem.

Intuitively, the problem arises because several very poor countries with very low levels of educational attainment have increased these by a large percentage amount: for example, from one year to two years, or 100 per cent. In contrast, rich countries have increased their levels by one or two years as well, but starting from a much higher base. This is shown by plotting the educational attainment data by continent in Figure 4.2.

One possible resolution of this puzzle is that it is not the percentage change in educational attainment that matters, but rather the change

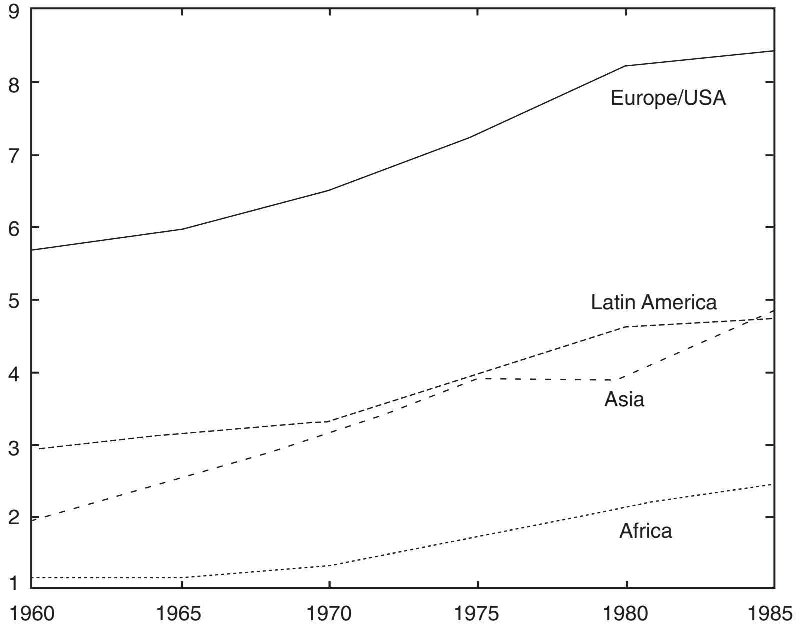


Figure 4.2 Educational attainment in years, by continent

in levels. In fact, this is exactly what a model based on the Mincerian micro-foundations suggests, as shown in the previous section.[15]

Table 4.2 illustrates the puzzle by including the logarithm of average educational attainment in the regression. The specification is first estimated in levels for 1990 and 1960, treating both years as steady-state observations.

As reported in the table, the results are very similar to those in Table 4.1. The last regression of the table is the differenced specification; all variables are the 1990 level minus the 1960 level, and the negative coefficient the change in the log of average educational attainment replicates the traditional puzzle. The partial correlation is displayed graphically in Figure 4.3.

In contrast, Table 4.3 and Figure 4.4 illustrate that there is no longer a puzzle when the *level* of average educational attainment is used, as suggested by the Mincerian approach used in the model here. Even in the specification using the 1990–1960 differenced data, the *level* of educational attainment enters positively and significantly with a coefficient that is quite close to the coefficient in the level specifications. This positive partial correlation is illustrated graphically in Figure 4.4.

Conclusion

Combining insights from Romer (1990), Mankiw *et al.* (1992), Nelson and Phelps (1996), and others to obtain a model that emphasizes the importance of technology transfer in understanding cross-country differences in income seems to be a promising avenue worthy of further research. The analysis presented here suggests that a model emphasizing research and ideas can generate the relatively successful cross-country regression pursued by MRW. But it also highlights the failings of this simple framework: it is not uncommon to find economies 2.7 times poorer or 2.7 times richer than predicted by the model.

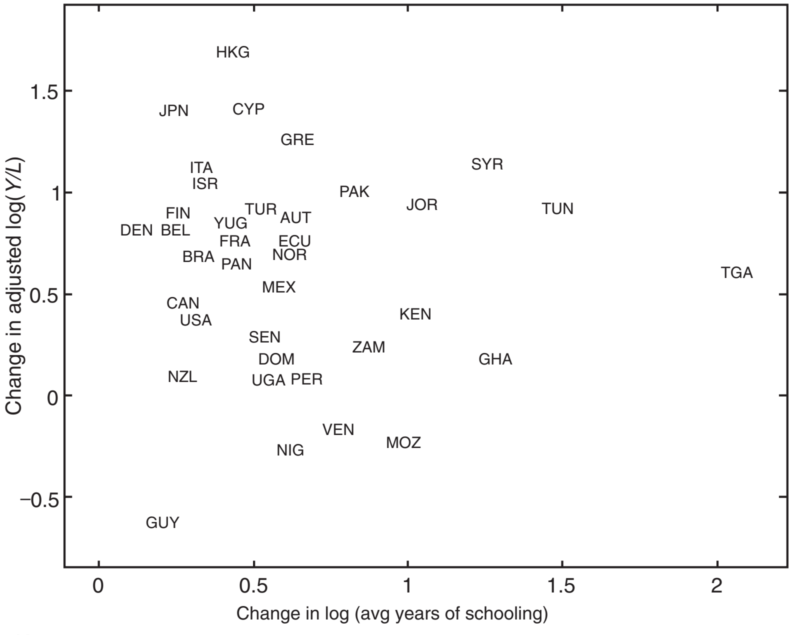
The chapter also suggests several insights related to human capital. First, the educational attainment data assembled by Barro and Lee (1993) and other similar data series are interpreted most accurately as something analogous to an investment rate rather than as a capital stock. This interpretation is consistent with the observation that educational attainment is asymptotically bounded; it does not grow without bound over time, as does the physical capital stock per worker.

Finally, the model follows the lead of Bils and Klenow (1996) by including educational attainment in the model in a way that is

Table 4.2 Regressions using the log of educational attainment

Variable	Level 1960		Level 1990		Difference 1990–1960	
Constant	5.350	(0.580)	5.814	(0.783)	0.621	(0.085)
$\log s_K/(n + g + d)$	0.425	(0.149)	0.437	(0.168)	0.394	(0.095)
$\log N$	1.032	(0.184)	0.500	(0.137)	−0.050	(0.128)
α	0.298	(0.073)	0.304	(0.081)	0.282	(0.049)
$\bar{R}^2$	0.668		0.522		0.141	

Note: Num Obs = 78. White heteroskedasticity-robust standard errors are reported in parentheses. For the 1990 regressions, s_K and n are computed as averages from 1986 to 1990, and the Barro-Lee data for 1985 is used. For the 1960 regression, s_K and n are computed as averages from 1960 to 1964, and the Barro-Lee data for 1960 and 1965 is averaged.



Key:

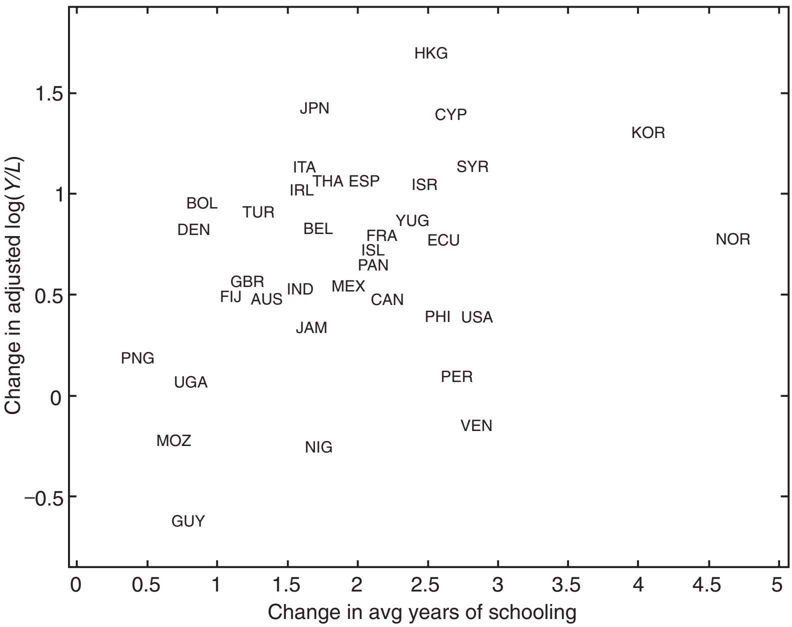
AUT	Austria	JOR	Jordan
BEL	Belgium	KEN	Kenya
BRA	Brazil	MEX	Mexico
CAN	Canada	MOZ	Mozambique
CYP	Cyprus	NZL	New Zealand
DEN	Denmark	NIG	Niger
DOM	Dominican Republic	NOR	Norway
ECU	Ecuador	PAK	Pakistan
FIN	Finland	PAN	Panama
FRA	France	PER	Peru
GHA	Ghana	SEN	Senegal
GRE	Greece	SYR	Syria
GUY	Guyana	TGA	Tonga
HKG	Hong Kong	TUN	Tunisia
ISR	Israel	TUR	Turkey
ITA	Italy	UGA	Uganda
JPN	Japan	USA	United States
		VEN	Venezuela
		YUG	Yugoslavia (Serbia and Montenegro)
		ZAM	Zambia

Figure 4.3 The puzzle using logs

Table 4.3 Regressions using the level of educational attainment

Variable	Level 1960		Level 1990		Difference 1990–1960	
Constant	5.512	(0.538)	5.950	(0.650)	0.031	(0.123)
log $s_K/(n + g + d)$	0.506	(0.128)	0.377	(0.138)	0.353	(0.095)
log N	0.191	(0.031)	0.189	(0.031)	0.159	(0.064)
α	0.336	(0.056)	0.274	(0.073)	0.261	(0.052)
$\bar{R}^2$	0.678		0.571		0.205	

Notes: See notes to Table 4.2.



Key:

AUS	Australia	KOR	Korea, South
BEL	Belgium	MEX	Mexico
BOL	Bolivia	MOZ	Mozambique
CAN	Canada	NIG	Niger
CYP	Cyprus	NOR	Norway
DEN	Denmark	PAK	Pakistan
ECU	Ecuador	PAN	Panama
FIJ	Fiji	PNG	Papua New Guinea
FRA	France	PER	Peru
GRE	Greece	PHI	Philippines
GUY	Guyana	ESP	Spain
HKG	Hong Kong	SYR	Syria
ISL	Iceland	THA	Thailand
IND	India	TUR	Turkey
IRL	Ireland	UGA	Uganda
ISR	Israel	GBR	United Kingdom (Great Britain)
ITA	Italy	USA	United States
JAM	Jamaica	VEN	Venezuela
JPN	Japan	YUG	Yugoslavia (Serbia and Montenegro)

Figure 4.4 The resolution: no logs

consistent with Mincerian wage regressions. This framework provides a natural resolution to a recently documented empirical puzzle. Benhabib and Spiegel (1994), Islam (1995) and Pritchett (1996) report negative coefficients of human capital growth rates in growth accounting regressions. The specification suggested here implies that it is not the growth rate of educational attainment that belongs in these regressions but rather the change in the level. Empirical analysis of this specification reveals a relatively stable coefficient on educational attainment, regardless of whether the specification is estimated in levels or differences.

Notes

1 Barro and Sala-i-Martin (1995) and Bernard and Jones (1996) make a similar point with respect to the convergence literature.
2 The model in this section draws heavily on Jones (1996).
3 Their specification is one in which $e^{\tau N}$ enters the production function explicitly for final output, where N is years of schooling.
4 Other effects, such as a duplication externality, can be included as well. See Jones (1995), and Jones and Williams (1996).
5 In terms of human capital, we assume that new units of labour are endowed automatically with the average skill level in the economy.
6 The steady state of the model is stable, as can be shown by examining the dynamics of the model.
7 It turns out to be convenient to think about output per worker in the final goods sector rather than output per member of the labour force. For example, much international data does not include the labour force data time spent in education.
8 We require $\mu < g\varepsilon^{-\theta}$ to guarantee that h/A is less than unity.
9 This approach implies large international knowledge spillovers of the kind explored by Coe *et al.* (1995). Eaton and Kortum (1995) provide a more detailed analysis of how ideas might be transferred across advanced economies.
10 More sophisticated approaches exist as well. For example, Judson (1996) computes a value of the human capital stock by weighting years of educational attainment by the cost of various levels of education. Jorgenson and Fraumeni (1992) and Mulligan and Sala-i-Martin (1994) use wage differentials between educated and uneducated labour to infer values of stocks of human capital.
11 This practice extends beyond the Barro and Lee data set. For example, Benhabib and Spiegel (1994) interpret alternative measures related to educational attainment as stocks of human capital.
12 Recall that, according to the model, the appropriate variable is output per worker in the final goods sector. Because time spent in school is not counted in the labour force, this is a reasonable measure. Also, we are ignoring labour employed in research. Since the measured shares of labour in research are quite small, even in advanced countries, this is probably inconsequential.

13 Notice that the levels specification cannot identify θ and γ separately in any case.
14 Given the fact that N and s_K are probably correlated with whatever it is that we are missing, the amount of variation that remains to be explained is probably even larger.
15 I do not deserve any of the credit for making this point. I first heard the suggestion from the participants of an NBER conference on human capital and economic growth in February 1996, at Stanford. As I recall, Kevin Murphy, Alwyn Young and Pete Klenow emphasized this in discussing Pritchett (1996). Independently, Julie Shaffner made a similar suggestion to me.

References

Barro, R. J. (1996) 'Democracy and Growth', *Journal of Economic Growth*, March vol. 1, pp. 1–27.

Barro, R. J. and Lee, J. (1993) 'International Comparisons of Educational Attainment', *Journal of Monetary Economics*, December, vol. 32, pp. 363–94.

Barro, R. J. and Sala-i-Martin, X. (1995) 'Technological Diffusion, Convergence, and Growth', NBER Working Paper No. 5151.

Benhabib, J. and Spiegel, M. M. (1994) 'The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data', *Journal of Monetary Economics*, vol. 34, pp. 143–73.

Bernard, A. B. and Jones, C. I. (1996) 'Technology and Convergence', *Economic Journal*, July.

Bils, M. and Klenow, P. (1996) 'Does Schooling Cause Growth or the Other Way Around?', Mimeo, University of Chicago.

Coe, D. T., Helpman, E. and Hoffmaister, A. W. (1995) 'North–South R&D Spillovers', Centre for Economic Policy Research, Discussion Paper No. 1133.

Easterly, W., King, R., Levine, R. and Rebelo, S. (1994) 'Policy, Technology Adoption and Growth', NBER Working Paper No. 4681.

Eaton, J. and Kortum, S. S. (1995) 'Trade in Ideas: Patenting and Productivity in the OECD', NBER Working Paper No. 5049.

Gershenkron, A. (1952) 'Economic Backwardness in Historical Perspective', in B. F. Hoselitz (ed.), *The Progress of Underdeveloped Areas* (Chicago, IL: University of Chicago Press).

Islam, N. (1995) 'Growth Empirics: A Panel Data Approach', *Quarterly Journal of Economics*, vol. 110, November, pp. 1127–70.

Jones, C. I. (1995) 'R&D-Based Models of Economic Growth', *Journal of Political Economy*, vol. 103, August, pp. 759–84.

Jones, C. I. (1996) 'A Simple Model of Growth and Development', Mimeo, Stanford University.

Jones, C. I. and Williams, J. C. (1996) 'Too Much of a Good Thing? The Economics of Investment in R&D', Stanford Economics Working Paper.

Jorgenson, D. W. and Fraumeni, B. M. (1992) 'The Output of the Education Sector', in Zvi Griliches (ed.), *Output Measurement in the Services Sector* (Chicago, IL: University of Chicago Press), pp. 303–38.

Jovanovic, B. (1995) 'Research, Schooling, Training, and Learning by Doing in the Theory of Growth', Mimeo, New York University.

Judson, R. (1996) 'Do Low Human Capital Coefficients Make Sense?', Board of Governors of the Federal Reserve Working Paper No. 96–13.

Lucas, R. E. (1988) 'On the Mechanics of Economic Development', *Journal of Monetary Economics*, vol. 22, pp. 3–42.
Mankiw, N. G., Romer, D. and Weil, D. (1992) 'A Contribution to the Empirics of Economic Growth', *Quarterly Journal of Economics*, May vol. 107, pp. 407–38.
Mincer, J. (1974) *Schooling, Experience, and Earnings* (New York: Columbia University Press).
Mulligan, C. B. and Sala-i-Martin, X. (1994) 'Measures of the Aggregate Value of Human Capital', Mimeo, University of Chicago.
Nelson, R. R. and Phelps, E. S. (1996) 'Investment in Humans, Technological Diffusion, and Economic Growth', *American Economic Association Papers and Proceedings*, May, vol. 56, pp. 69–75.
Pritchett, L. (1996) 'Where Has All the Education Gone?', World Bank, Policy Research Working Paper No. 1581, March.
Romer, P. (1993) 'Idea Gaps and Object Gaps in Economic Development', *Journal of Monetary Economics*, pp. 543–73.
Romer, P. M. (1990) 'Endogenous Technological Change', *Journal of Political Economy*, October, vol. 96, pp. S71–S102.
Segerstrom, P. S. (1995) 'Endogenous Growth Without Scale Effects', Mimeo, Michigan State University.
Summers, R. and Heston, A. (1991) 'The Penn World Table (Mark 5): An Expanded Set of International Comparisons: 1950–1988', *Quarterly Journal of Economics*, May, vol. 106, pp. 327–68.
Weitzman, M. L. (1996) 'Recombinant Growth', Mimeo, Harvard University.

5

A Rising Tide Raises All Ships: Trade and Diffusion as Conduits of Growth

Jonathan Eaton and Samuel Kortum

Introduction

Technology can flow from one country to another through at least two conduits. One is the diffusion of technological knowledge itself. Another is the exchange of goods embodying technological advances.[1]

Our previous work (Eaton and Kortum, 1994, 1996, 1997) focused purely on the first mechanism. Coe and Helpman (1995) take the opposite tack and relate technology flows to trade flows. Neither provides an encompassing framework in which both mechanisms are allowed to operate. In this chapter we develop a model of trade and technology to examine theoretically and empirically the connections between innovation, productivity and trade.[2]

Empirical work does suggest a connection between technology flows and trade. In Eaton and Kortum (1996) we find that flows of ideas (estimated on the basis of a model of world growth making use of data on international productivity and patenting) are in part explained by trade patterns, although the elasticity is small. Ben-David (1993) provides enticing evidence that trade leads to convergence of income levels. A model to draw out these connections is not provided in either case, however.

Our purpose here is to develop a model of trade and productivity to examine the connections between the two. A major implication of our analysis is that not only does productivity have important implications for trade but that trade can have implications for observed productivity as well. Indeed, trade can lead to convergence of countries' productivity levels even though they exchange no ideas. The reason is that trade pushes countries to concentrate on the activities they do best, and will tend to raise observed productivity everywhere. Given population, the effect is more pronounced in backward countries, whose measured productivity may then exaggerate their true command of technology.

Our model implies that trade barriers will force average productivity to be higher in industries that serve export markets relative to those that serve the home market. Higher productivity is needed to offset these barriers. This result is consistent with other evidence on the relationship between productivity and trade. For example, Bernard and Jensen (1999) find that firms that export have higher productivity than other firms in the USA. Moreover, they find that this relationship is largely explained by selection; that is only good firms can survive in export markets.

While reducing trade barriers allows less efficient firms to export, competition from abroad drives out the least efficient procedures for the domestic market. The net effect is that lower trade barriers generally enhance productivity. The idea that global exposure is good for productivity is, in fact, a major theme of the McKinsey Global Institute's (1993) monumental study of manufacturing productivity in Germany, Japan and the USA.

The next section presents the model itself, the third section calibrates the model to the five leading research economies, and provides rough calculations of the gains from trade and technological advance, while the fourth, and final, section discusses the implications of our model for the question raised in our title: Can trade by itself cause standards of living to rise in all countries even though technological advances are highly concentrated in just a few countries?

Productivity and trade

Consider a world consisting of $n = 1, \ldots, N$ countries, and a continuum of goods indexed on the unit interval, which in principle can be produced in any country. A worker in country *i* can produce z_i (*j*) units of good *j*. Hence the cost of producing a unit of that good in that country is w_i/z_i (*j*) where w_i is the wage there.

The domestic technological frontier

We first characterize what a country can do on its own. We describe a country's state of productive knowledge in terms of the distribution over goods of the parameter *z*, which we refer to as 'the domestic technological frontier'. We think of this frontier as reflecting the history of ideas about producing goods in that country, with $z(j)$ the best idea for producing good *j*. We make the specific assumption that the quality of an individual idea for producing a good is a random variable Q drawn from the cumulative Pareto distribution, $F(q) = 1 - q^{-\theta}$.[3] The good j to which the idea applies is drawn from the uniform distribution on [0, 1].[4] We use μ_n to denote the stock of ideas that have arrived in country n. As we show in Eaton and Kortum (1994), the cumulative distribution of the state of the art, given μ_n, is then:[5]

$$H(z \mid \mu_n) = e^{-\mu_n z^{-\theta}} \tag{5.1}$$

Note that the domestic technological frontier in a country depends only on the total stock of ideas there, regardless of where they came from or when they arrived. A higher value of θ implies less heterogeneity in quality across goods. A critical simplifying assumption is that the distribution of productivities is independent across countries.

International trade

We make Samuelson's standard and convenient 'iceberg' assumption about transport costs, that a fraction $1/d_{ni}$ units of what country i exports arrives in country n. We normalize $d_{nn} = 1$ for all n.[6]

Assuming perfect competition, the cost of good j in country n imported from country i is $p_{ni}(j) = w_i d_{ni}/z_i(j)$, where we measure the wage in each country in terms of final output in that country. The fraction of country i's goods whose price in country n would exceed p is $H(c_{ni}/p \mid \mu_i)$, where $c_{ni} = w_i d_{ni}$. Let $p_n(j)$ denote the lowest price for good j available in country n. The distribution of $p_n(j)$ across goods is given by:

$$1 - \prod_{i-1}^{N} H(c_{ni} / p|\mu_i) = 1 - e^{-\tilde{\mu}_n p^{\theta}} = 1 - H(p^{-1}|\tilde{\mu}_n) \tag{5.2}$$

Here:

$$\tilde{\mu}_n = \sum_{i=1}^{N} \mu_i c_{ni}^{-\theta}$$

is the sum of each source country's stock of technologies adjusted by their wage cost and the cost of transporting to country n. This term, which we can think of as the trade-augmented stock of knowledge, shows how the possibility of international trade enlarges the stock of technologies available domestically with those available from other countries. Note that if transport costs are zero, then this stock, and the consequent price distribution, is the same for all countries, while if international transport costs are prohibitive it reduces to $\mu_n w_n^{-\theta}$.

Remarkably, the distribution of $p_n(j)$ is the same as the distribution of $p_{ni}(j)$ conditional on $p_{ni}(j) = p_n(j)$. That is, the goods sold by any country i in country n have the price distribution 2, regardless of source. The probability that a good from country i will be the cheapest one available in country n is $\mu_i c_{ni}^{-\theta}/\tilde{\mu}_n$, which corresponds to the fraction of goods used by country n provided by country i.[7] Hence the fraction of goods provided by each source country corresponds to that country's contribution to the destination country's trade-augmented stock of knowledge.

To derive implications about production and trade volumes we need to specify preferences. We assume that the utility of a representative individual in any country is Cobb–Douglas, with each good having equal share;

$$u_n = \int_0^1 \text{In}[x(j)]\, dj \tag{5.3}$$

where $x(j)$ is consumption of good j. An implication is that aggregate spending on each good equals income Y. Aggregate income in country n equals wage payments, $Y_n = w_n L_n$, where L_n is labour supply in country n. Hence spending on goods from country i is just the measure of goods imported from there times income, or:

$$X_{ni} = \left(w_i d_{ni}\right)^{-\theta}\left(\frac{\mu_i}{\tilde{\mu}_n}\right) w_n L_n \tag{5.4}$$

This relationship links knowledge and export share: given wages and transport costs, countries with larger μs will have larger market shares, while, given knowledge, countries with higher factor costs will have lower market shares. Given that country n purchases a particular good j from country i, then the physical amount that it purchases will be:

$$x_{ni}(j) = Y_n/p_{ni}(j) = z_i(j) w_n L_n/c_{ni}$$

which will require employing $d_{ni} x_{ni}(j)/z_i(j) = w_n L_n/w_n$ workers.[8]

Factor-market equilibrium

Total employment is country i for production of exports to country n is just $L_{ni} = X_{ni}/w_i$. Summing across countries gives total demand for country i's labour. Equating demand to supply gives the conditions for labour-market equilibrium in each country:

$$L_i = \frac{\mu_i}{w_i^{1+\theta}} \sum_{n=1}^{N} \frac{d_{ni}^{-\theta} w_n L_n}{\tilde{\mu}_n} = \frac{\mu_i}{w_i^{1+\theta}} \sum_{n=1}^{N} \frac{d_{ni}^{-\theta} w_n L_n}{\sum_{k=1}^{N} \mu_k \left(w_k d_{nk}\right)^{-\theta}} \quad i = 1, \ldots, N \tag{5.5}$$

By Walras's Law, one equation is redundant. Hence this set of equations determines $N - 1$ relative wages as functions of the labour forces, stocks of knowledge and transport costs.

While Equation (5.5) constitutes a highly non-linear set of equations, note that they are homogeneous of degree 0 both in all countries' labour forces and in all countries' stocks of knowledge. Hence *relative* wages are not affected by proportional increases in labour forces, or in technologies across all countries.

For reasons that are familiar to students of the Ricardian model of trade, the effects of trade on relative wages depend on relative country size: smaller countries gain relatively more from trade. A useful benchmark case extracts from this effect by considering countries with identical levels of autarky GDP, which is proportional to $\mu^{1/\theta}L$. In this case, with a common cross-country transport cost d, relative real wages are proportional to $\mu_i^{1/\theta}$, which can be seen from substitution into Equation (5.5).

Prices and real wages

Since the cost of goods varies across countries, so does the cost of living. The price index appropriate to our utility function is simply the geometric mean of prices, which is proportional to:

$$P_n = \bar{\mu}_n^{-1/\theta} \tag{5.6}$$

The real wage in country n, which measures the standard of living there, is $w_n 1 P_n$. The numerator reflects productive efficiency as well as the accessibility of export markets. The denominator reflects access to cheap goods. This distinction breaks down in the special cases of autarky (all d_{ni}s infinite for $n \neq i$) and free trade (all d_{ni}s one). Under autarky, the system of Equation (5.5) is vacuous, and the real wage, determined by Equation (5.6), is given by $w_n 1 P_n = \mu_i^{1/\theta}$. Under free trade, $P_n = P$, since $\bar{\mu}_n = \bar{\mu}$ for all n. Note also that in this special case we can rearrange Equation (5.5) as:

$$Y_i = \varphi_i Y$$

where Y is world income and

$$\phi_i = \frac{\mu_i w_i^{-\theta}}{\sum_{n=1}^{N} \mu_n w_n^{-\theta}} = \mu_i \left(\frac{w_i}{P} \right)^{-\theta}$$

Under free trade, a country's share of world income depends on its stock of knowledge relative to its real wage.

Except in these limiting cases, however, the concepts of consumption per worker, as reflected by the real wage, and production per worker, diverge. We now turn to worker productivity.

In general, it is hard to infer much analytically about the role of technology and transport costs from these expressions. Hence we provide a numerical analysis below. When countries are of equal size (as measured by $\mu^{1/\theta}L$) with a common cross-country transport parameter d, the expression for the real wage reduces to:

$$w_i/P = \mu_i^{1/\theta}[1 + (N - 1)d^{-\theta}]^{1/\theta}$$

Increasing a country's own stock of knowledge and its number of trading partners raises its real wage, as does the reduction of trade barriers.

Productivity

A natural measure of productivity in a country is the average number of units of goods produced per worker. We question whether we should look at the average of what country i would produce under autarky, or at what it actually produces given its patterns of trade. The first depends on the stock of ideas available to country i, μ_i. Taking the geometric average over the state-of-the-art technologies available in country i, from Equation (5.1) productivity is proportional to $\mu_i^{1/\theta}$. This concept ignores the fact that those goods that the country is relatively poor at producing it will import instead, while it will employ more workers to produce those goods where it is relatively most advanced. The second concept, which is closer to what is in fact measured in available productivity statistics, takes into account the allocation of labour across the different goods.

To construct this second measure, we observe that the prices of goods sold by country i in country n have the distribution given by Equation (5.2) above. The productivities corresponding to these prices have distribution $H(z/c_{ni}|\tilde{\mu}_n)$. The average productivity of goods sold by country i to country n is consequently proportional to $\tilde{\mu}_n^{1/\theta} w_i \mathrm{d}_{\mathrm{ni}}$. Observe that the more expensive it is to ship goods to country n's market, the more productive country i must be in making those goods in order to compete there. Country n's weight in the sales of country i, which is also the share of country i's labour used to produce for country n, is $\varphi_{ni} = d_{ni}^{-\theta}\,\tilde{\mu}_n^{-1} Y \Big/ \left[\sum_{j=1}^{N} d_{ji}^{-\theta}\,\overline{\mu}_j^{-1} Y_j\right]$. Combining the two elements to form a geometric average we get as an index of overall productivity:

$$A_i = w_i \prod_{n=1}^{N} \left(\tilde{\mu}_n^{1/\theta} d_{ni} \right)^{\varphi_{ni}} \tag{5.7}$$

In the polar case of autarky, this measure reduces to $\mu_i^{1/\theta}$. At the other pole of free trade it becomes $w_i \tilde{\mu}^{1/\theta}$, which is also the real wage in that case. At either of these poles, relative productivities are exactly proportional to the domestic stocks of knowledge. This is not the case in general, however, since countries that are more distant from their markets will have lower wages relative to productivity in order to compensate for transport costs.

Where countries are identical in size as measured by autarky GDP (proportional to $\mu^{1/\theta} L$), and where d is common to all cross-national country pairs, this expression reduces to:

$$A_i - \mu_i^{1/\theta}[1 + (N - 1)\, d^{-\theta}]^{1/\theta}\, d^{(N-1)/(d\theta+N-1)}$$

which is decreasing in d. Hence, increasing transport costs unambiguously lowers both the real wage and measured productivity. Increased transport costs force a country to specialize in a more select set of goods in which it is on average more productive in serving export markets. But reduced exposure to foreign competition allows less productive domestic industries to serve the domestic market, which is the effect that dominates. From a situation of unimpeded trade ($d = 1$), the real wage falls faster than productivity as transport costs rise, since workers have to pay more for the goods they import. As transport costs become large enough to bring the world towards autarky, however, productivity levels converge to the autarky real wage.

Asymmetric countries

More interesting are situations in which countries differ in terms of their overall size. Unfortunately, we cannot obtain closed-form solutions for this case. We instead turn to numerical simulations to explore how international trade between unequally-sized countries influences the relationship between a country's stock of technical knowledge and its measured productivity and real wage.

To illustrate the effect of openness on relative wages and productivity, we consider a world of two countries with equal labour forces but in which one country (country 1), under autarky would have a 20 per cent productivity advantage over the other (country 2). We fix the parameter $\theta = 2$, based on estimates from our other work (Eaton and Kortum, 1994). To illustrate the effects of openness, we consider values of d ranging from

10 (at which countries buy 99 per cent of their output at home) to 1 (at which point international trade is costless). Figure 5.1 illustrates the effects of varying *d* over this range on: (i) the fraction of domestic output sold at home in each country (the two descending curves); (ii) productivity levels in each country (the top and second-from-bottom ascending curves); and (iii) the real wage in each country (the second-from-top and bottom ascending curves). Note that:

(i) As transport costs fall, the real wage and measured productivity generally rise in both countries (although productivity in country 2 falls slightly as *d* approaches unity), by a factor of about 2 for country 1, and 2.4 for country 2.

(ii) As with countries of equal size, lowering transport costs affects productivity earlier than it affects the real wage. Real wage gains are slow to emerge until *d* has reached about 2, at which point trade is about 20 per cent of output. At that point there is a dramatic effect of openness on real wages. In contrast, productivity rises more steadily over the range of transport costs that we consider. Over the whole range from essential autarky to free trade, for any given country, the real wage and productivity rise by exactly the same amount.

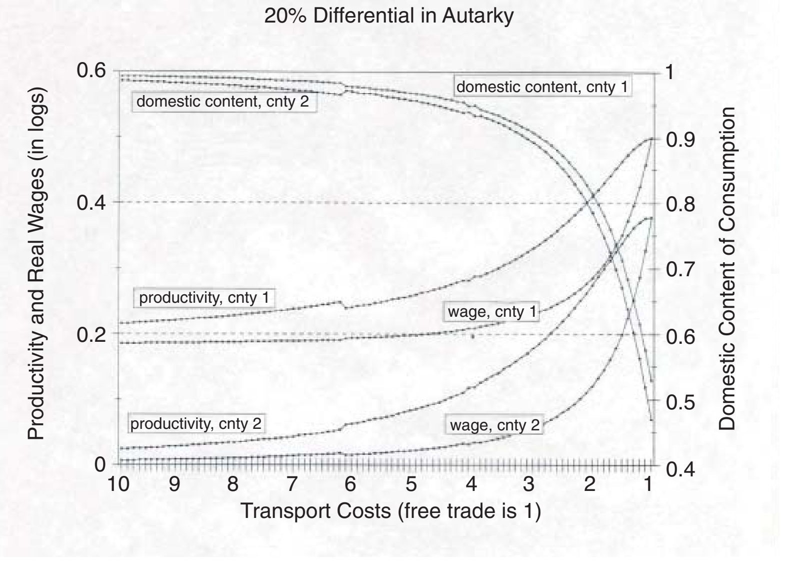


Figure 5.1 Moving towards free trade

(iii) In contrast to countries of equal size, as transport costs fall, cross-country differentials in real wages and productivity decline in percentage terms, illustrating how trade can convey the benefits of one country's more advanced technology to a less advanced country with equal physical resources. An implication is that productivity measures may give a small, open country the appearance of having more technological sophistication than it in fact possesses. Only under autarky would productivity measures fully reveal a country's state of technology as measured by μ.

Going back to Ricardo, economists have examined the implications of productivity for trade. This model illustrates the implications of trade for observed productivity. Since trade allows a smaller country to specialize in a narrower range of activities, it can be more selective in the ones that it chooses, so can appear to be more advanced than a larger country with equal know-how.

An application to five countries

Can our model explain observed trade and productivity patterns; and if it can, what does it say about the impact of trade on real wages and productivity? We examine the situation in five countries: the USA, Germany, France, the UK and Japan, in the light of our model.

To see how well trade alone can explain productivity patterns we wish to examine a world in which there is no technology diffusion between countries. Hence knowledge stocks would solely reflect ideas arising from domestic research. For this purpose we take the domestic R&D stocks calculated by Coe and Helpman (1995) to represent a country's stock of technology, along with (approximate) labour forces.[9] As a ratio to the US level, these data are as shown in Table 5.1.

Note that the distribution of R&D stocks is much more skewed than that of labour forces. As we show below, if R&D stocks correspond to

Table 5.1 Endowments of technology and labour

Country	R&D stock	Labour force
United States	1.00	1.00
Germany	0.17	0.25
France	0.11	0.25
United Kingdom	0.14	0.25
Japan	0.25	0.50

the μs of our model, with our estimate of $\theta = 2$, we would predict, under autarky, a much more unequal distribution of productivity levels than what is in fact observed. A higher value of θ would, of course, imply more equality, but gets us into trouble when we also try to predict trade volumes.

A crude calibration

We choose transport costs in order to approximate relative import shares among these five countries, with the diagonal representing 1 minus the total import share (from Coe and Helpman, 1995). We set two values of d, one applying to trade within the Europe and Pacific regions, ($d^{\text{int}\,ra} = 3$) and another to trade between these regions ($d^{\text{int}\,er} = 5$). Again we set $\theta = 2$. Table 5.2 reports the amount of trade predicted by these parameter estimates compared with actual trade (in parentheses).[10]

Overall, the model captures observed trade patterns surprisingly well, given the crudeness of the calibration. We do, however, overstate Japan's import levels substantially.[11]

The relative productivities predicted by our model, compared with productivity (value added per hour in manufacturing in 1990) (from van Ark, 1995) are as shown in Table 5.3.

In the third column we give the productivity levels predicted by the model in the case where d is infinite for all cross-country pairs.

The main thing to note is that introducing trade into a model with no international technology diffusion goes a long way towards bringing predicted productivity levels into line with actual ones. The distribution of productivity under autarky would be extremely unequal compared with actual measures, while our model implies about the right amount of dispersion. Obviously, under any assumptions about d, our model has trouble explaining why France is as productive as it is, since its R&D stock is so small. Similarly, it has trouble understanding why the UK is not more productive.

The gains from trade

With parameter estimates in hand, we can ask how different the world would be if international trade barriers were eliminated, and it was completely closed to international trade. We report in Table 5.4 the absolute productivity levels that would emerge in each case. We also report the real wage under our base case. (In the cases of autarky and free trade, the real wage and productivity measures coincide.)

Our model suggests that actual productivity measures are more than halfway towards their free trade levels from an initial autarky position.

Table 5.2 Impact shores

Destination source	USA	Germany	France	UK	Japan
USA	0.91 (0.89)	0.013 (0.018)	0.012 (0.010)	0.013 (0.013)	0.055 (0.068)
Germany	0.079 (0.57)	0.73 (0.75)	0.073 (0.092)	0.078 (0.053)	0.043 (0.048)
France	0.084 (0.046)	0.086 (0.12)	0.79 (0.77)	0.083 (0.043)	0.046 (0.025)
UK	0.081 (0.070)	0.083 (0.10)	0.075 (0.060)	0.72 (0.73)	0.044 (0.037)
Japan	0.16 (0.073)	0.021 (0.013)	0.019 (0.088)	0.020 (0.007)	0.78 (0.90)

Table 5.3 Relative productivity

Country	Predicted	Actual	Autarky
USA	1.00	(1.00)	1.00
Germany	0.79	(0.86)	0.41
France	0.69	(0.91)	0.33
UK	0.75	(0.66)	0.37
Japan	0.76	(0.78)	0.50

Table 5.4 Counterfactual trade barriers

Country	Autarky	Base productivity	Base wage	Free trade
USA	1.00	1.22	1.05	1.42
Germany	0.41	0.96	0.48	1.24
France	0.33	0.85	0.40	1.08
UK	0.37	0.92	0.45	1.18
Japan	0.50	0.93	0.56	1.12

For the real wage, the opposite is the case. Further movement towards free trade would further concentrate the distribution of productivity, with more pronounced implications for real wage inequality.

The gains from innovation

What are the benefits of technological advance, and how are they shared? We consider the effect of increasing by 20 per cent the stock of technological knowledge in the USA (the largest country where the stock is already high) and in France (one of the three smallest where the stock is low). The implications for productivity and the real wage are as shown in Table 5.5.

Starting with the USA, the 20 per cent increase in the stock of knowledge translates into an approximate 10 per cent increase in productivity and the real wage, as would be predicted in the case of autarky, given our value of $\theta = 2$. Other countries experience about a 1–2 per cent gain in productivity and only a negligible gain in real wage.

The effect of the change in France is much less pronounced. The French themselves get about a 5 per cent productivity gain and a 3 per cent gain in the real wage. The effect elsewhere is negligible.

Conclusion: trade or diffusion?

Our model's relative success in explaining observed productivity and trade levels on the basis of measured R&D inputs might lead us to con-

Table 5.5 Counterfactual rise in technological knowledge

Country	Base prod.	Case rw	Raise prod.	USA rw	Raise prod.	France rw
USA	1.22	1.05	1.33	1.15	1.22	1.05
Germany	0.96	0.48	0.98	0.48	0.97	0.48
France	0.85	0.40	0.86	0.40	0.90	0.43
UK	0.92	0.45	0.93	0.45	0.92	0.45
Japan	0.93	0.56	0.95	0.57	0.93	0.57

clude that diffusion is not needed to fit the facts. As far as explaining levels of international productivity, this may be the case. However, if one relies on trade as the sole vessel to convey technological advances across countries, one runs aground trying to explain why countries that do very little research grow, on average, as fast as the major research economies.

An implication of our expressions for the real wage and productivity is that these magnitudes will grow faster in countries whose research efforts are greater or growing at faster rates. In fact, productivity growth among our five countries has not differed that much since the mid-1980s, and there is little evidence that major research economies have been growing much faster than the others. For this reason we think that technological diffusion is central to any story that seeks to explain comparative productivity growth.

In previous work we developed a series of models of innovation and diffusion with the implication that, in a steady state, countries all grow at the same rate, with each country's relative income level determined by its ability to absorb innovations from both home and abroad. This work ignored international trade, however. The analysis in this chapter indicates that trade cannot substitute for diffusion as the vessel transporting innovations abroad. Rather, it shows how productivity and technological diffusion cannot be studied in isolation from their interactions with international trade.

Notes

1 Grossman and Helpman (1995) survey the literature on technology and trade.

2 Keller (1996) explores Coe and Helpman's results further, showing that trade patterns do not play as important a role as the original analysis would suggest.

3 Bental and Peled (1995) and Kortum (1997) also use the Pareto distribution to characterize the pool of undiscovered techniques from which researchers

draw. The Pareto distribution has the convenient feature that, if we truncate the distribution at some level z, then the random variable $Q/z(\geq 1)$ inherits the Pareto distribution. The average inventive step implied by the distribution is $\theta/(\theta - 1)$. Hence, sectors where inventions are more potent have smaller θs.

4 The continuum allows us to abstract from randomness in aggregate outcomes. To simplify further, we ignore the possibility that research could be aimed at improving the quality of a specific good.

5 That the best available technology for producing each good has this distribution across goods follows directly from our assumption that the qualities of ideas are drawn from a Pareto distribution. However, its functional form is one of only three possible limiting distributions of extrema of random sampling. See, for example, Billingsley (1986).

6 See Krugman (1995) for a discussion of this assumption.

7 The distribution of prices in country n of goods produced in countries other than i is $1 - H(p^{-1} \imath \bar{\mu}_{n,-\mathrm{i}})$ where $\bar{\mu}_{n,-\mathrm{i}} = \tilde{\mu}_n = \mu_i \mathrm{c}_{\mathrm{ni}}^{-\theta}$. The probability that a good from country i is imported by n, given that its price is p, is the probability that no other country provides the good more cheaply, or $H(p^{-1} | \tilde{\mu}_{n,} - \mathrm{i})$. From 1, goods potentially imported from country i would have a cost in country n with density given by: $\int_{n,i}(p | \mu_i) = \frac{d}{dp} [1 - H(c_{ni}/\mathrm{p} | \mu_i)] = \theta p^{\theta-1} \mu_i \mathrm{c}_{\mathrm{ni}}^{\theta}$ $\exp(- \mu_i \mathrm{p}^{\theta} \mathrm{c}_{\mathrm{ni}}^{-\theta})$. The probability that a good from i is the cheapest available in country n is therefore: $\int_0^{\infty} f_{n,i}(p | \mu_i) H(p^{-1} | \bar{\mu}_{n,}{}^{-\mathrm{i}}) dp = \mu_i \mathrm{c}_{\mathrm{ni}}^{-\theta} | \bar{\mu}_n$. The prices of goods sold by country i in country n have density $f_{n,i}(p \imath \mu_i)\ H\ (p^{-1} | \bar{\mu}_{n,} - \mathrm{i})/ (\mu_i \mathrm{c}_{\mathrm{ni}}^{-\theta} / \bar{\mu}_n)$ which, by Equation (5.2) is also the density of prices in country n.

8 Note that, even though our structure is Cobb–Douglas, countries face higher than unit elastic demand for their exports. An increase in unit cost reduces demand through two channels: the first is the unit elastic effect on the demand for each good still purchased; and the second is through the loss in the share of inputs, provided. Here our model bears a close resemblance to the analysis in Dornbusch *et al.* (1977).

9 Coe and Helpman (1995) derive this measure from domestic R&D expenditures using a perpetual inventory model. While our earlier work (Eaton and Kortum, 1994) provided alternative estimates of μ, these estimates were based on a model in which international technology diffusion occurred.

10 Obviously, we are ignoring the rest of the world. We do so by allocating each country's trade with the rest of the world to the four other countries in proportion to actual trade with the country.

11 Increasing the cost of importing into Japan to 5 improved the fit noticeably, but we preferred to stick with symmetrical transport costs.

References

Ark, B. van (1995) 'Productivity and Competitiveness in Manufacturing: A Comparison of Europe, Japan, and the United States', in K. Wagner and B. van Ark (eds), *International Productivity Differences: Measurement and Explanation* (Amsterdam: North-Holland).

Ben-David, D. (1993) 'Equalizing Exchange: Trade Liberalization and Income Convergence', *Quarterly Journal of Economics*, vol. 108, pp. 653–79.

Bental, B. and Peled, D. (1995) 'Endogenous Technical Progress and Growth: A Search Theoretic Approach', *International Economic Review* vol. 37, pp. 687–718.

Bernard, A. and Jensen, J. B. (1999) 'Exceptional Exporter Performance: Cause, Effect on Both?' *Journal of International Economics*, vol. 47, pp. 1–25.

Billingsley, P. (1986) *Probability and Measure*, 2nd edn (New York: John Wiley).

Coe, D. T. and Helpman, E. (1995) 'International R&D Spillovers', *European Economic Review*, vol. 39, pp. 859–87.

Dornbusch, R., Fischer, S. and Samuelson, P. A. (1977) 'Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods', *American Economic Review*, vol. 67, pp. 823–39.

Eaton, J. and Kortum, S. (1994) 'International Patenting and Technology Diffusion', NBER Working Paper No. 4931.

Eaton, J. and Kortum, S. (1996) 'Trade in Ideas: Productivity and Patenting in the OECD', *Journal of International Economics*, vol. 40, pp. 251–78.

Eaton, J. and Kortum, S. (1997) 'Engines of Growth: Domestic and Foreign Sources of Innovation', *Japan and the World Economy*, vol. 9, pp. 235–59.

Grossman, G. M. and Helpman, E. (1991) 'Quality Ladders and Product Cycles', *Quarterly Journal of Economics*, vol. 106, pp. 557–86.

Grossman, G. M. and Helpman, E. (1995) 'Technology and Trade', in G. M. Grossman and K. Rogoff, (eds), *The Handbook of International Economics, Vol. III* (Amsterdam: North-Holland).

Keller, W. (1996) 'Are International R&D Spillovers Trade Related? Analyzing Spillovers among Randomly Matched Trade Partners', Mimeo, University of Wisconsin.

Kortum, S. (1997) 'Research, Patenting, and Technological Change', *Econometrica*, vol. 65, pp. 1389–419.

Krugman, P. R. (1995) 'Increasing Returns, Imperfect Competition, and the Positive Theory of International Trade', in G. M. Grossman and K. Rogoff (eds), *The Handbook of International Economics, Vol. III* (Amsterdam: North-Holland).

McKinsey Global Institute (1993) *Manufacturing Productivity* (Washington, DC: McKinsey Global Institute).

Summers, R. and Heston, A. (1991) 'The Penn World Table (Mark V): An Expanded Set of International Comparisons', *Quarterly Journal of Economics*, vol. 106, pp. 327–68.

6

The Role of Education and Knowledge in Endogenous Growth

Luigi Paganetto and Pasquale Lucio Scandizzo

Introduction

In 1962, Kenneth Arrow published an article 'The Economic Implications of Learning by Doing', based on the experience of the aeronautical industry, which soon became the reference point of a vast literature on endogenous evolution of technical progress, education and growth. Although Arrow's article did not deal directly with the problems of technical progress and international trade, the link between the economies of learning and increases in productivity suggested a mechanism of accumulation and feedback that questioned the very foundation of comparative advantage in trade.

Arrow's basic idea resided in the definition of a new resource: the stock of knowledge, capable of being accumulated without sacrifice of material consumption, and linked directly to the physical stock of capital accumulated in some (but not all) sectors of the economy. Because of the peculiar properties of this resource, the sectors that can exploit the dynamic linkage between the accumulation of knowledge and ordinary production will find themselves experiencing economies of scale. These economies will occur in the form of a dynamic externality that will shift the frontier of production continually in response to the ordinary activity of capital accumulation.

After the contribution by Cass (1965) and Levhari (1966) extending Arrow's model, a young graduate of the Massachusetts Institute of Technology (MIT), Eytan Sheshinski, took the task of developing in more specific forms the idea of learning economies and its implication for comparative advantage. Sheshinski's main contributions were first presented in an unpublished essay (1966), then in his Ph.D. dissertation (1966), and finally in a famous journal article (1967). In

Sheshinki's model, technological progress reflects the experience accumulated in the production of the investment goods. However, while the growth of knowledge increases the production capacity of all sectors in the economy, at the same time it changes the terms of exchange in production in favour of investment goods.

The reason for this result is simple: only the sectors producing investment goods increase their productivity as a direct consequence of the accumulation of capital and the ensuing stock of knowledge. The other sectors benefit only indirectly, because of the release of factors of production made possible by the productivity increases discussed above. Hence, the terms of trade change in favour of the sectors producing investment goods.

The important implication of this result is that the pattern of specialization and the evolution of the economy are determined by the capital–labour ratio originally chosen, rather than by the absolute or relative endowment of any set of resources. Following this point of departure, the economy will undergo a succession of specializations, starting with either a consumption or an investment good. In approximately half of these cases, comparative advantage can be seen to result from a growth path that the economy will follow to the end only if investment, and the related accumulation of useful knowledge, is pushed to its optimal level. Since the learning process is an effect external to the firm, this event will generally not happen as a spontaneous consequence of the market mechanism, and will require a subsidy to be reached.

Although Arrow and Sheshinski's formulations have remained points of reference for economic theory since the 1970s, they focused on a rather narrow aspect of process innovation and related economies of scale. They did not have anything to say about what appeared to be a much more empirically relevant mechanism of accumulation of knowledge: the discovery of new products and the increase in the level of skill of specialized human capital.

The idea that the technical progress generated by learning economies might be related to product innovation emerged slowly in economic theory in a series of papers published in the second half of the 1980s, whose main representatives can be found in Krugman (1985), Romer (1986), Lucas (1988) and Grossmann and Helpmann (1988).

Lucas imagines an environment where economies of scale arise because new products are being introduced continuously, and this innovation in turn causes a continuous upgrading of already accumulated human capital. The skill level of human resources, inherited from the past, in other words, is expanded continuously by the discovery of new products more able to utilize it. While the learning functions for each product thus maintain their properties of decreasing returns, the aggregate learning function, because of the virtuous circle between product innovation and skill acquisition, displays increasing returns.

As in the Sheshinski case, Lucas (but see also Ethier, 1979; and Stokey, 1987) shows that the learning economies, by virtue of their nature as external to the firm, occupy a wedge between the private and the social rates of return. As a consequence, the pattern of specialization depends on the initial conditions, and private equilibrium is suboptimal. Since the private operators ignore the learning economies, it is possible to improve market performance by offering incentives to high technology, and/or to activities that are more likely to benefit from the economies of learning.

Unlike the Sheshinski case, however, specializing in high technology is not necessarily a source of long-term comparative advantage. If the elasticity of substitution between low and high technology goods is small, in fact, the expansion of the high tech sector will result in a deterioration in its terms of trade that will more than outweigh the benefits of the more dynamic increments in productivity. For the products with the highest technology and the lowest elasticity of substitution, there will thus be a tendency towards a progressive concentration of comparative advantage in the areas where resource endowment is so favourable that it is capable of 'resisting' the negative dynamics of trade. The 'high tech–low substitution' goods are thus a sort of 'residual' sector where lasting specialization occurs because of robust comparative advantage stemming from some sort of exceptional endowment of key resources.

These two somewhat extreme views on knowledge and progress expressed by Sheshinski and Lucas illustrate the range of the debate on the role of human capital formation in the mechanism of endogenous growth. Two basic ideas are opposed to one another, and sometimes combined: (i) the idea that accumulation of physical capital may have the virtue of enhancing the stock of knowledge through learning by doing; and (ii) the idea that human capital formation may be at the root of product innovation, economies of scale and technical progress. Clearly, the two ideas are not necessarily contradictory, but one may venture the conjecture that while learning by doing *per se* may generate only a weaker form of progress (via process innovation), human capital formation via formal education will not necessarily result in product innovation, learning economies and technical progress.

In this chapter we take the view that education may be represented as a somewhat ambiguous engine of growth. All forms of education will indeed promote human capital formation, but there may be ranges of development where the types of human capital produced are not what is needed to promote efficient endogenous growth. Moreover, even when it is of the 'right' type, human capital is not the same as technical progress. Innovation and growth, in fact, are both inherently stochastic phenomena, whose volatility over time may be affected by the extent and quality of human capital inputs.

The plan of the chapter is as follows. In the second and third sections we develop a simple dynamic macroeconomic model where human capital formation affects both consumption and productive capacity. In this model, technical progress is characterized as a stochastic process with a controllable drift, and education is used as an indirect control with a variable degree of effectiveness in accord with its characteristics. In the fourth section we develop some preliminary tests of the results suggested by the model against the empirical evidence, using time series data on the economic performance of South Italy over the years 1951–93, and the fifth section we summarize the results and present some conclusions.

Education and technical progress: a Harrod–Domar approach

The basic difference between education and technical progress is not the fact that the former can be considered an input and the latter an output of the same process (as in the learning by doing hypothesis), but their asymmetric effect on demand and supply. As a form of long-term-orientated, intangible and somewhat addictive consumption good, education can be expected to affect both consumption and production. On consumption, it may act as a substitute for more tangible consumption goods, as an upgrader of the whole pattern of demand or as a determinant of metapreferences (Hirschman, 1985). Its on production effects occur through the accumulation of human capital, the learning of new techniques, and the upgrading of existing skills.

Technical progress, on the other hand, has direct effects only on production. It occurs through process or product innovation in ways that may entail feedback on education and human capital formation, and it responds to investment inputs in an essentially stochastic fashion. Using Domar's Keynesian framework, we can formalize these concepts for a closed economy, as follows:

National income Y is the sum of (material) consumption C and investment I:

$$Y = C + I \tag{6.1}$$

Consumption depends both on the income level and on the level of an intangible good: education or culture, N:

$$C = c_0 + c(N)Y \tag{6.2}$$

Education N is not produced directly, but as an external effect of the rate of growth of physical capital:

$$dN = a\frac{I}{K}dt \tag{6.3}$$

where we have assumed zero depreciation ($dK = I$).

By combining Equations (6.1)–(6.3), and taking differentials, we obtain the expression for the multiplier:

$$dY = \frac{1}{1-c}\left(dI + Y_0 c_N\ a\frac{I}{K}\right)dt \tag{6.4}$$

where $c_N = dc/dN$ and Y_0 is initial level of income.

On the supply side, productive capacity Q is a function of the stock of capital K and technical progress V:

$$Q = \theta(V)K \tag{6.5}$$

where θ^{-1} is the capital/output ratio and $\theta_v = \dfrac{d\theta}{dV} > 0.$

Technical progress is a stochastic process derived from combining human and physical capital formation:

$$dV = bdN + \sigma\left(\frac{I}{K}\right)^{\frac{1}{2}}dz = ab\left(\frac{I}{K}\right)dt + \sigma\left(\frac{I}{K}\right)^{\frac{1}{2}}dz \tag{6.6}$$

where dz is the increment of a standard Wiener process, uncorrelated across time, and at any instant satisfying:

$$E(dz) = 0 \qquad E(dz^2) = dt$$

Differentiating Equation (6.5), we obtain, after taking expectations and applying Ito's lemma:

$$dEQ = \left[\left(\theta_v abI + \frac{1}{2}\theta_{vv} \ \ \sigma^2 I\right) + \theta I\right]dt \tag{6.7}$$

Equating demand in Equation (6.4) and expected supply in Equation (6.7) yields the warranted rate of growth of investment:

$$g = s\theta + s\left(\theta_v ab - \frac{\theta}{s} c_N a + \frac{1}{2}\theta_{vv}\sigma^2\right) \tag{6.8}$$

This rate can be divided into three parts: the first is identical to the warranted rate described by Domar and depends of only on the current values of the propensity to save and the capital output ratio. The second part depends also on the capital output ratio through the variation of the marginal propensity to consume caused by education. The third part is a function of the increase in the productivity of capital with technical progress, and the product of the economies or diseconomies of scale with the volatility of technical progress.

Assume now that education N is a function of two types of input:

$$N = F(E_1, E_2) \tag{6.9}$$

where E_1 represents the non-finalized educational component (that is, general education, the humanities, 'culture') and E_2 the finalized component (that is vocational and technical education). To reflect this hypothesis we can write Equation (6.9) as follows:

$$dE_i = a_i \frac{I}{K} dt \qquad i = 1,\ 2 \tag{6.10}$$

Equation (6.6), on the other hand, can be written:

$$dV = b_1 F_1 dE_1 + b_2 F_2 dE_2 + \sigma\left(\frac{I}{K}\right)^{1/2} dz \tag{6.11}$$

where $F_i = dF/dE_i; \quad i = 1, 2$

and the warranted rate will turn out to be:

$$g^* = s\theta + s\left[\theta_v \sum_{i=1}^{2} a_i b_i \ \ F_i \ - \ \frac{\theta}{s} \ \sum_{i=1}^{2} a_i \ \ c_i \ F_i + \frac{1}{2}\theta_{vv} \ \sigma^2\right] \qquad (6.12)$$

where $c_i = dC/dE$; $i = 1, 2$.

Because of its 'non-finalized' features, the E_1 component of education will presumably increase consumption ($c_1 > 0$), and it is likely to display diseconomies of scale over technical progress ($\theta_{vv} < 0$) in the range where the first education input prevails. Thus education may have negative effect on the warranted rate of growth, if its non-finalized component prevails, because it tends to depress savings and increase the volatility of technical progress.

Equilibrium growth for two regions

Consider now the problem of two regions: the North and South, specialized in the production, respectively, of manufactures (the North) and primary products (the South).[1] We assume that the North is characterized by full employment and a constant growth rate (including labour-saving technical progress) equal to n. The South, on the other hand, is characterized by structural unemployment (after Lewis), and a growth rate that clears the goods markets but is independent of labour supply.

To characterize the Southern economy fully, we use the framework in Equations (6.1)–(6.8), adding net imports M_s to Equation (6.1), according to the definition:

$$M_j = M_j(Y_j, p) \quad j = S, N \qquad (6.13)$$

where Y_j is national income (the subscripts S and N denote, respectively, South and North variables), and p indicates the terms of trade between manufactures and primary products. In Equation (6.13) $dM_j/dY_j > 0$ for $j = S, N$, while $dMj/dp < 0$ for the North.

Given Equation (6.13), with slight modifications of Equations (6.1)–(6.8), we obtain the warranted growth rate for the South:

$$g_s^* = sp\left\{\theta + \left[a\left(b\theta_v - c_N \ \frac{\theta}{s}\right) + \frac{1}{2}\theta_{vv} \ \sigma^2\right]\right\} \qquad (6.14)$$

where $s = 1 - c + m_s$, and $m_s = dM_s/dY_s$. Assuming, for simplicity, θ_v and θ_{vv} are constant (that is, $\theta(V)$ a quadratic), and the investment in education of type 2 ($c_N > 0$, $\theta_{vv} < 0$), Equation (6.14) can be interpreted as the expression of the long-run rate of growth for the South.

For the two regions to be in long-run equilibrium, the terms of trade p should adjust to guarantee equality between their growth rates:

$$p^* = \frac{n}{s\left\{\theta + \left[a\left(b\theta_v - c_N \frac{\theta}{s}\right) + \frac{1}{2}\theta_{vv}\sigma^2\right]\right\}} \tag{6.15}$$

Equation (6.15) shows that the long-run terms of trade between the two regions should be a function of full employment, once the North growth rate is given exogenously. In the South, however, they may be affected by structural parameters, including education and technical progress. If the terms of trade cannot adapt perfectly in the short run, because of institutional rigidities, not uncommon between two regions of the same country, we should expect the growth rates of the two regions to be different. In the long run we should have, however, by differentiating p^* in Equation (6.15):[2]

$$\frac{dp^*}{p^*} = -\left[\left(\theta_v + \theta_{vv}\right) ab + \frac{1}{2}\theta_{vv}\sigma^2\right] s \cdot \frac{p^*}{g^*} \tag{6.16}$$

In addition to Equation (6.15), the flow of exchange between the two regions should also respect the condition of equilibrium in the long run:

$$dp^* = 0 \tag{6.17}$$

which, by the virtue of Equation (6.16), implies:

$$\frac{\theta_v}{|\theta_{vv}|\left(ab + 0{,}5\frac{\sigma^2}{ab}\right)} \tag{6.18}$$

Because the assumption that θ_{vv} is constant is equivalent to hypothesizing that $\theta(V)$ is quadratic:

$$\theta = \theta_0 + \bar{\theta}_v \cdot V - \frac{1}{2}|\theta_{vv}|V^2 \tag{6.19}$$

Equation (6.18) can be solved to indicate the expected level of technical progress that will prevail in the South for the long run:

$$V^* = \frac{\bar{\theta}_v}{|\theta_{vv}|} - \left(ab + 0,5\frac{\sigma^2}{ab} \right) \tag{6.20}$$

The expected level of technical progress necessary to insure North–South equilibrium will be higher, the lower the ratio $\bar{\theta}_v/1\theta_{vv}1$, which indicates the strength of the decreasing (or increasing) return to scale in the relationship between the productivity of capital and technical progress.

One last requisite for the long-run equilibrium is given by the equality of the trade flows:

$$pM_N = M_s \tag{6.21}$$

By differentiating Equation (6.21), assuming $dp = 0$, we obtain:

$$m_N\text{p}^* \, dY_N = m_s dY_\text{s} \tag{6.22}$$

where $m_j = \partial M_j/\partial Y_j \quad j = N, S$.

Substituting Equation (6.15) into Equation (6.22), and recalling that $g = n$, we can solve Equation (6.22) for the equilibrium level g in the stationary state:

$$\hat{g} = \frac{m_S\bar{Y}_S}{m_n\bar{Y}_N} s\left(\theta + \left(a\left(b\theta_v - c_N \frac{\theta}{1} \right) - \frac{1}{2}|\theta_{vv}|s^2 \right) \right) \tag{6.23}$$

where Y_j ($j = N, S$) is the income level in the two regions in the stationary state. In the long run, in other words, the growth rate will converge towards a common value determined, *inter alia*, by the parameters of the less developed region.

Equation (6.23) confirms the findings of the single region model. The equilibrium growth rate is, in fact, a positive function of the effectiveness parameters of education and technical progress of the less developed region, and a negative function of the technical progress of the same region.

An empirical test: economic development in South Italy

To test the hypothesis developed above, we consider the case of development in South Italy between 1951 and 1993. Here, the investment in education, noticeable at all levels, has been characterized by the prevalence of ('classical' and juridical studies over technical training. Equation (6.24) below is a first attempt to analyze the impact of education, together with other key factors of growth, on the disappointing performance of the South in the forty-year period under examination:

$$\begin{aligned} g_t = & -\underset{(-2.03)}{0.123}\, y_{t-1} + \underset{(3.60)}{1.203}\frac{I_t}{y_t} - \underset{(2.54)}{0.032}\, CL_t + \\ & + \underset{(2.68)}{0.046}\, S_t - \underset{(-4.28)}{10.013}\, G_t - \underset{(-3.13)}{26.75}\, Le_t \\ & + \underset{(4.16)}{36.57}\, Tec_t + \underset{(3.14)}{0.102}\, D_{73-81} \\ & \qquad \bar{R}^2 = 0.52 \end{aligned} \tag{6.24}$$

where g_t = growth rate of real GNP per capita in year *t* in South Italy;
Y_t = real GNP per capita in year *t*;
I_t = real investment per capita in year *t*;
CL_t = loss of days of work because of labour conflicts, per capita, in year *t*;
S_t = km of roads per capita in year *t*;
G_t = students enrolled in the humanities school, per capita in year *t*;
Le_t = university students in humanities, per capita in year *t*;
Tec_t = university students in economics, per capita,in year *t*; and
D_{73-81} = dummy variable for the years of the oil crisis.

Equation (6.24) supports the hypothesis that an imbalance between finalized and non-finalized education may have had an adverse effect on the economic performance of South Italy. In order also to test the hypothesis that the volatility of technical progress may have played a role in the same matter, we constructed a proxy for the variance of technical progress by computing the variable:

$$V_t = \left(u_t - \sum_{j=1}^{n} \frac{u_{t-i}}{n} \right)^2 \tag{6.25}$$

where u_t is the residual of the regression:

$$q_t = a + bt + u_t \tag{6.26}$$

where q_t is the productivity of capital in manufacturing, and *t* a linear trend. Using this variable as a regressor yields:

$$g_t = -\underset{(-3.10)}{0.172}\, g_{t-1} + \underset{(3.43)}{0.82}\, \frac{I_t}{y_t} - \underset{(-3.10)}{0.035}\, CL_t$$
$$- \underset{(-3.39)}{5.45}\, G_t + \underset{(3.67)}{27.64}\, Tec_t - \underset{(-2.96)}{0.083}\, V_t \;\; \underset{(4.31)}{0.128}\, D_{73-81} \qquad (6.27)$$
$$\overline{R}^2 = 0.51$$

These results appear to corroborate the *ex ante* hypothesis, derived from the models in Equations (6.1)–(6.23), suggesting that non-finalized educational activities and the stochastic nature of technical progress may have conspired to determine the disappointing growth results of the development of South Italy between 1951 and 1993.

Conclusions

This chapter has addressed the question of human capital formation and endogenous growth by focusing on the different roles played by formal education and learning by doing on the one hand, and product innovation and technical progress on the other. We have hypothesized that, while learning is a systematic phenomenon that can be directed consciously towards consumption or production targets, innovation and progress are, while largely a direct effect of such conscious activities, are at the same time subject to a substantial random component.

This approach leads us to conclude, on purely theoretical grounds, first, that the type of education and its explicit targets do matter in determining the direction of progress and, second, that they may have a further, less evident, but not unimportant effect on its volatility. Ultimately, the growth performance of a country may be determined by the extent to which the educational inputs are used successfully to deal with these different and somewhat conflicting aspects of technical progress.

The chapter also shows that preliminary testing of these results on time series data for South Italy appears to corroborate some of these conclusions. The estimates presented support the hypothesis that growth in South Italy was affected positively by technical education, even at university level, and affected negatively by investment in more traditional, and less finalized forms of education. They also support the hypothesis that variability in the growth of productivity in manufacturing was a significantly negative factor of the regional performance over the years considered.

Appendix: some data on education in Italy, 1951–1993

The evolution of formal schooling in Italy since 1951 may be used as a test case for the role of education in regional performance. As illustrated in a recent work by Gattei (1995), Italy is characterized by modest levels of education. According to this study, in most OECD countries, more than 50 per cent of the population between the ages of 25 and 64 have completed high school, while the corresponding proportion in Italy is only 20 per cent (see Table 6.1).

In the national context, the differences between the North and the South are lower for the lower grades (primary and lower secondary; see Table 6.2). The inferior performance of formal education in the South, in comparison to the North (see Table 6.3) seems to originate from two separate causes: (i) the higher number of students who do not go beyond obligatory schooling (14 per cent in the South and 10 per cent in the North); and (ii) the higher drop-out rate (17.3 per cent in the South and 14.5 per cent in the North).

As for university education, its problems in the South are both quantitative and qualitative. The higher proportion of students enrolled in the Northern universities is matched by a higher share in technical and scientific studies (see

Table 6.1 Italy: resident population above six years of age (percentages)

	With degrees from						
Regions	University	High school	Secondary school	Primary school	No degree	Illiterates	Total
1951							
Italy	1.0	3.3	5.9	76.9	–	12.9	100
Centre–North	1.06	3.6	6.93	81.5	–	6.7	100
South	0.9	2.6	3.95	68.3	–	24.3	100
1961							
Italy	1.3	4.3	9.6	42.3	34.2	8.3	100
Centre–North	1.4	4.7	10.9	46.2	32.5	4.3	100
South	1.2	3.6	6.8	34.7	37.8	15.9	100
1971							
Italy	1.8	6.9	14.7	44.3	27.1	5.2	100
Centre–North	1.9	7.2	16.1	48.2	24.06	2.56	100
South	1.7	6.4	11.4	35.9	33.9	10.6	100
1981							
Italy	2.8	11.5	23.8	40.6	18.2	3.1	100
Centre–North	2.9	12.2	25.2	42.1	16.1	1.4	100
South	2.7	10.1	20.7	37.2	23.15	6.2	100
1991							
Italy	3.8	18.58	30.69	32.54	12.21	2.14	100
Centre–North	4.08	20.02	31.35	33.72	9.87	0.94	100
South	3.35	15.96	29.96	30.41	16.46	4.3	100

Source: Census data.

Table 6.2 Students enrolled per capita, 1951–93 (per 1000 residents between 5 and 9 years of age for primary school, 10 and 14 for secondary school, and 15 and 19 for high school)

	Primary		Secondary		High school	
Years	South	North	South	North	South	North
1951	540.4	532.5	150.4	183.6	89.6	101.6
1956	601.9	562.8	183.9	208.4	143.0	155.0
1961	556.9	504.1	283.1	315.2	192.0	186.8
1966	555.5	517.4	391.7	465.3	343.3	373.2
1971	563.1	565.3	510.4	615.4	420.5	488.3
1976	538.3	537.3	595.9	694.0	463.6	542.0
1981	494.1	487.6	609.4	637.4	486.5	540.3
1986	501.5	460.9	681.6	708.6	534.1	634.9
1991	480.6	473.7	642.4	635.1	624.5	697.1
1993	474.2	465.3	615.5	596.4	627.6	698.2

Source: ISTAT.

Table 6.3 Students enrolled in high school, 1951–93 (by type per 1000 residents between 15 and 19 years of age)

	Technical		Specialist high school		Scientific		Humanities	
Years	South	North	South	North	South	North	South	North
1951	22.0	37.3	20.8	17.7	5.9	11.5	35.4	23.6
1956	48.5	65.4	28.7	24.1	7.4	14.5	43.4	30.2
1961	78.8	85.6	34.6	22.6	11.6	18.3	46.6	31.3
1966	131.8	173.8	72.4	52.8	30.6	37.2	56.8	45.8
1971	161.7	211.5	56.8	42.3	76.2	71.4	56.1	51.3
1976	192.4	258.6	49.0	34.2	82.7	87.5	48.2	41.1
1981	205.1	249.3	55.7	36.8	73.6	75.7	50.2	40.2
1986	248.1	306.0	43.5	32.0	76.4	94.7	52.2	46.5
1991	299.2	298.5	45.4	33.8	93.4	121.7	57.8	51.4
1993	276.6	294.0	47.4	34.6	103.2	123.8	60.5	52.1

Source: ISTAT.

Table 6.4). In this respect, the gap between the North and South has widened, especially if we compare the number of students enrolled in Letters and Philosophy.

Notes

1 This problem has been treated in a world trade context by Findlay (1980) and Darity (1990). Here, we use a somewhat different approach, focusing on two regions of different development within the same

Table 6.4 Students enrolled in universities, 1951–93 (by type per 1000 residents between 20 and 29 years of age)

	Engineering		Economics		Law		Medicine		Letters and Philosophy	
Years	South	North	South	North	South	North	South	North	South	North
1951	2.2	4.1	2.1	3.8	6.1	3.9	3.5	4.6	2.9	2.3
1956	2.3	4.4	2.6	4.5	6.7	4.6	2.6	3.6	2.2	1.8
1961	2.6	5.4	4.8	7.1	7.5	5.0	2.5	3.5	2.9	1.8
1966	4.9	7.2	9.3	11.1	6.6	4.8	3.9	5.2	6.1	6.1
1971	8.4	11.6	10.2	8.9	15.1	8.2	11.0	13.9	10.8	9.9
1976	9.5	13.6	7.8	7.6	22.6	11.8	20.9	24.5	14.5	4.5
1981	8.8	12.2	12.3	11.7	23.8	14.6	19.4	21.8	12.4	4.9
1986	8.2	18.0	13.2	17.1	23.1	17.9	10.7	14.1	10.4	8.3
1991	12.9	21.2	22.0	22.9	28.4	28.4	7.8	8.4	14.9	9.2
1993	15.0	24.5	23.9	22.1	29.9	31.9	9.0	6.5	16.3	9.6

Source: ISTAT.

2 From Equation (6.15), by differentiating with respect to time, assuming $dn = 0$ and c_N and θ_{vv} constant:

$$dp^* = \frac{-s\left[\left(\theta_v + ab\theta_{vv}ab + \frac{1}{2}\theta_{vv}\sigma^2\right)\right]\frac{I}{K}}{s\left\{\theta + \left[a\left(b\theta_v - c_N\frac{\theta}{s}\right) + \frac{1}{2}\theta_{vv}\sigma^2\right]\right\}^2}$$

$I/K = g^*$, while the denominator equals $(g^*/p^*)^2$.

References

Arrow, K. (1962) 'The Economic Implications of Learning by Doing', *Review of Economic Studies*, vol. 29, no. 3, June, pp. 155–73.

Cass, D. (1965) 'Optimum Growth in an aggregative Model of Capital Accumulation', *Review of Economics Studies*, vol. 32, no. 3, pp. 233–40.

Darity, W. Jr (1990) 'The Fundamental Determinants of the Terms of Trade Reconsidered: Long Run and Long Period Equilibrium', *American Economic Review*, vol. 80, no. 4, pp. 816–27.

Ethier, W. (1979) 'Internationally Decreasing Costs and World Trade', *Journal of International Economics*, vol. 9, February, pp. 1–24.

Findlay, R. (1980) 'The Terms of Trade and Equilibrium Growth in the World Economy', *American Economic Review*, vol. 70, no. 3, pp. 291–9.

Gattei, S. (1995) 'La Scuola Secondaria Superiore Statale nel Mezzogiorno', *Rivista Economica del Mezzogiorno*, Svimez, Associazione per 6 sviluppei dell' industria nel mezzogiorno, vol. 2, pp. 395–439.

Grossman, G. M. and Helpman, E. (1988) *Comparative Advantage and Long-Run Growth*, Discussion Papers in Economics No. 142, Working Papers (Princeton, N J: Princeton University, Woodrow Wilson School).

Hirschman, A. O. (1985) 'Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse', *Economics and Philosophy*, vol. 1, pp. 7–21.
ISTAT (Central Statistics Institute) (1998) *Statische culturali*, Istituto nazionale di statistica, Rome.
Krugman, P. (1985) 'The Narrow Moving Band, the Dutch Disease and the Competitive Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Dynamic Scale Economies', Unpublished working paper, Massachusetts Institute of Technology, Cambridge, MA.
Levhari, D. (1966) 'Extension of Arrow's Learning by Doing', *Review of Economic Studies*, vol. 33, pp. 117–32.
Lucas, R. E. Jr (1988) 'On the Mechanisms of Economic Development', *Journal of Monetary Economics*, vol. 22, no. 1, July, pp. 3–42.
Romer, P. M. (1986) 'Increasing Returns and Long-run Growth', *Journal of Political Economy*, vol. 94, no. 5, October, pp. 1002–37.
Sheshinski, E. (1966a) 'An "Infant Industry" Growth Model', Unpublished working paper, Department of Economics, Massachusetts Institute of Technology, Cambridge, MA.
Sheshinski, E. (1966b) 'Some Tests of the Learning by Doing', Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA.
Sheshinski, E. (1967) 'Tests of the Learning by Doing Hypothesis', *Review of Economics and Statistics*, vol. 49, no. 4, pp. 568–78.
Stokey, N. L. (1987) 'Learning-by-doing and the Introduction of New Goods', Unpublished working paper, Northwestern University, Evanston, IL.

7

Factors Behind the Asian Miracle: Entrepreneurship, Education and Finance

Richard R. Nelson and Howard Pack

Different theories of economic development

The debate about how to explain the 'Asian miracle' puts a spotlight on a more general theoretical debate about how to explain long-run economic growth. The broader debate is between theorists who, in effect, attempt to explain economic growth in a way that is consistent with the canons of general equilibrium theory, and theorists who argue that growth must be understood as an evolutionary process driven by technological advance. The former case is, of course, familiar; see Romer (1990) for a strong statement. The latter was articulated by Nelson and Winter (1982), and is in the spirit of Schumpeter's well-known criticism of equilibrium theory as a vehicle for understanding economic growth.

The focus of this chapter is not on the general debate, but on its particular manifestation in argument about how to explain the Asian miracle. Between the 1960s and the 1990s, Korea, Taiwan, Singapore and Hong Kong transformed themselves from being technologically backward and poor, into relatively modern and affluent economies. Each has experienced a more than fourfold increase in per capita income, and each now has a significant collection of firms producing technologically-complex products competing effectively against rival firms based in the USA, Japan and Europe. The growth performance of these countries has vastly exceeded those of virtually all other economies that had comparable productivities and income levels in 1960. On these grounds alone, the question of how they did it is obviously of enormous scientific and policy importance.

It has been less well noted that their growth has been unprecedented historically. The development of Japan in the half century after the

Meiji Restoration is widely regarded as being comparable. However, Japan's growth rate over this period was less than half that of the Asian newly industralized countries (NICs) since 1960. Of course, growth rates in general were slower during this earlier period. But the rate of catch-up by the NICs is still remarkable. It certainly would seem that there is an 'Asian miracle' crying out for explanation.

Of course, economists have not been blind to or unattracted by the challenge. Over the 1980s/1990s a number of different theories have been put forward purporting to explain the phenomenon. See Westphal *et al.* (1985), Pack and Westphal (1986), Amsden (1989), Pack (1992), World Bank (1993), Young (1993), Kim and Lau (1994), Krugman (1994), Rodrik (1994). There is unanimity among the different theories regarding the identity of some of the key causal factors. All the Asian NICs have experienced rapid growth in their physical capital stock. All have been marked by very low rates of investment in human capital. Virtually all theories about how they did it place these investments centre-stage in the explanation.

However, there are significant differences in the causal mechanisms stressed. At the risk of doing some violence to the actual diversity, for our purposes we find it useful to divide up theories of the Asian miracle into two groups. One group, which we shall call 'accumulation' theories, stresses the role of these investments in moving these economies 'along their production functions'. The other group, which we call 'assimilation' theories, stresses the entrepreneurship, innovation and learning that these economies had to go through before they could master the new technologies they were adopting from the more advanced industrial nations; it sees investment in human and physical capital as a necessary, but far from sufficient, part of the assimilation process.

The 'accumulation' theory has been pushed hard by several economists, in a way clearly designed to strip away most of the 'miraculous' from the 'Asian miracle'. They say that what lies behind rapid development is, simply, very high investment rates. Economists who support this point of view do not deny that the adoption and mastery of new technology and other modern practices was an important part of the story. Rather, the position is that one should try to explain as much as one can in terms of investments that enable movement along a production function, and see if anything much is left over, thus requiring an explanation on other grounds. Several economists who have followed this path find that, according to their calculations, the lion's share of increased output per worker can be explained simply by increases in physical and human capital per worker. Thus there is little need to assign much of the credit for the growth 'miracle' to entrepreneurship, innovation or learning, except in so far as these are terms given to the shift to more capital- and education-intensive ways of production (see Young (1993), Kim and Lau (1994), Krugman (1994)).

To assimilation theorists, this point of view seems odd. The technologies that the NICs came to master progressively during the 1970s and 1980s were ones with which, in 1960, they had no experience at all. To learn to use them effectively required the development of new sets of skills, new ways of organizing economic activity, and becoming familiar with and competent in new markets. To do this was far from a routine matter, but involved risk-taking entrepreneurship as well as good management. (See Pack and Westphal (1986), Amsden (1989), Dahlman (1994).) What makes the Asian miracle miraculous is that these countries did these things so well, while other countries were much less successful. To be sure, adopting the technologies of the advanced countries required, among other things, high rates of investment in physical and human capital, and the NICs achieved these high rates. But to say that these investments simply enabled these economies to 'move along their production functions' seems a strange use of language. At the least, it poses the question of just what is meant by 'moving along a production function'.

Are we drawing a distinction without a real difference? We do not think so. The accumulation account stresses, simply, investments. The message is that other countries could have done as well as the successful NICs if they had made a similar investment effort. If the nation makes the investments, and marshals the resources, development will follow. In contrast, the assimilation account stresses learning about, risking to operate, and coming to master technologies and other practices that are new to the country, if not to the world. The 'marshalling of inputs' is part of the story, but the emphasis is on innovation and learning, rather than on marshalling. Under this view, if one marshals but does not innovate and learn, development does not follow.

A convinced accumulationist might respond by saying that, if one educates the people and provides them with modern equipment to work with, they will learn. An assimilationist might respond that the Soviet Union and the Eastern European Communist economies took exactly that point of view, made the investments, and didn't learn. There is nothing automatic about the learning business. The response of the accumulationist might be that the old Communist countries provided an economic environment where there was no incentive to learn to be efficient, either in a technological or an economic sense, much less to innovate. The assimilation theorist might agree, but then propose that it is important to understand, therefore, just how the successful NICs did it. The accumulationist would reply that they got the prices right and made the necessary public investments. Economists who stress entrepreneurship, innovation and learning would reply that it is not all that simple, and point to countries such as Spain that have had high investment rates, and have got most of the prices right, but which are developing at far lower rates than the Asian NICs.

The difference between the theories shows up strikingly in the way they treat the following four matters: what is involved in entrepreneurial decision-making; the nature of technology; the economic capabilities lent by a well-educated work force; and the role that exporting played in these countries' rapid development.

Accumulationists pay little explicit attention to firms, seeing their behavior as being determined basically by the environment – the incentives and constraints – they face, which determines the actions that are most profitable. Assimilation theorists, on the other hand, see entrepreneurial firms, and their ability to learn rapidly, as a critical factor behind the success of Korea and Taiwan, with their behaviour supported by their environments, but only partially determined by external forces (see Hobday, 1995; and Kim, 1997). For an assimilation theorist, at least our brand, when firms contemplate venturing on to ground that is new to them, the profitability of such venturing is highly uncertain, in the sense of Knight. Some firm managers will dare to venture; others will choose to stick close to the familiar. Thus what firms do is determined by the daring of their decision-makers, as well as by their environment. And whether an entrepreneurial venture will succeed or fail is also only partly determined by environmental factors. It depends, as well, on the zeal, smartness and learning abilities of firms' management and workers.

Part of the difference here resides in how the different theories see technology. Accumulationists seem to believe that the state of technological knowledge at any time is largely codified in blueprints and associated documents and that, for a firm to adopt a technology that is new to it but not to the world, primarily involves getting access to those blueprints. In contrast, assimilationists argue that only a small portion of what one needs to know to employ a technology is codified in the form of blueprints; much of it is tacit, and learning is as much by doing and using as by reading and studying (see Nelson and Winter, 1982; and Rosenberg, 1994). Further, while many economists believe that technology is defined in terms of engineering and physical science, in fact, the lines between the engineering aspects of technology and the organizational aspects are blurred, and controlling a technology often involves knowing how to manage a very complex division of labour as much as it involves knowing the relevant physics and chemistry.

Both of these differences show up in terms of how the two theories go about explaining the fact that the NICs were able to increase greatly and rapidly their capital–labour ratios (by more than fourfold over the years in question) without experiencing a significant decline in the rate of return to capital. The accumulationist would tend to invoke the concept of the elasticity of substitution, which refers to innate technological opportunities, and propose that the phenomenon in question indicates that the elasticity of substitution was high. The assimilationists, on the other hand, would argue that there is no such thing as a set of technological possibilities that can be defined independently of decision-makers' ability to search out, see and effectively take on board new technology. That is, what the accumulationist would explain in terms of the nature of the parameters of a conventionally defined production function, an assimilationist would explain in terms of skilful entrepreneurship and learning.

Along the same lines, the two theories also differ regarding how they see the effects of the rapidly rising education levels in these countries. For the accumulationist, rising human capital is treated simply as an increase in the quality or effectiveness of labour. Assimilationists, on the other hand, tend to see the effects of sharply rising educational attainments, in particular the creation by these countries of a growing cadre of reasonably well-trained engineers and applied scientists, in ways similar to that sketched out many years ago by Nelson and Phelps (1966). Well-educated in seeing new opportunities and effectively learning new things. Thus the growing human capital of the NICs was a very important support for successful entrepreneurship.

The difference between the two theories also shows up sharply in how they treat the strong export performance of the NIC manufacturing firms. The accumulationists tend to see the steep rise in manufacturing exports as just what one would expect in economies where the stocks of physical and human capital were rising rapidly, and shifting comparative advantage towards the sectors that employed these inputs intensively. From this perspective, there is nothing noteworthy about the surge of manufacturing exports, save that it is evidence that the economic policies of these countries let comparative advantage work its ways. In contrast, the assimilationists, while not denying that the NICs were building a comparative advantage in various fields of manufacturing, tend to highlight the active efforts by government to induce, almost force, firms to try to export, and the entrepreneurship, innovation and learning the firms had to do in order to compete effectively in world markets, even with government support.

Several economists of the assimilation school have argued that exporting stimulated and supported strong learning in two ways (see Pack and Westphal, 1986; Pack, 1988). First, being forced to compete in world markets made the managers and engineers in the firms pay close attention to world standards. Second, much of the exporting involved contracting with American or Japanese firms demanding high performance and providing assistance to achieve it. The story here is clearly different from one that sees the development of these new competencies as simply the more-or-less automatic result of changing factor availabilities that called them into being.

We have noted that the assimilationist's position (at least the one we espouse) sees the high rates of investment by the NICs in physical and human capital as a necessary, if not a sufficient, component of the assimilation process. These high rates themselves are remarkable, even if not miraculous. Under the argument of the assimilationists, these investments were at least partially induced by, and sustained by, the rapid innovation and learning that was going on.

Successful entrepreneurship in the NICs was certainly facilitated by the growing supply of well-trained technical people. On the other hand, it was not automatic that newly-trained engineers would find work in entrepreneurial firms. There had to be entrepreneurial firms in which to work, or the opportunity to found new ones. Thus aggressive entrepreneurship supported and encouraged rapidly rising educational attainment.

The successful manufacture of new products almost always requires that firms acquire new physical capital. There is no question that policies in these countries encouraged saving. But on the other hand, what made saving and investment profitable was the strong and effective innovative performance of the firms that were entering new lines of business.

We think it is apparent that the two broad theories differ, both in their causal structures, and in the hints they give about 'how to do it'. The emphasis of the accumulationists is on getting investment rates up and the prices right. The message of the assimilation theorists is that successful industrial development requires innovation and learning to master modern technologies; effective innovation and learning depend on investment and a market environment that presses for efficient allocations, but it also involves much more. And, indeed, to a considerable extent, the investment needed is induced by successful entrepreneurship.

Section 2 considers the argument that careful attention to the numbers and rigours calculation supports the accumulationist theory, and there is little evidence that innovation and learning played much of a role. We argue that the commonly used calculations do not do what their proponents claim. In section 3, we propose a different way to discriminate between a change in output accompanied by changes in inputs that can be considered simply 'a movement along the production function', and a change that seems to involve innovation and learning. In the light of the argument we develop there, in section 4 we consider the evidence. We propose that that evidence supports strongly the assimilationist's case. Section 5 considers in what ways the differences between the two theories matter.

Why the standard calculations in fact do not discriminate

The case put forward by its proponents for the accumulationist theory is based on calculations of two kinds. One is a growth accounting. The other involves fitting a dynamic production function. In both methods, the strategy is, basically, to try to calculate the effect of input growth on output growth, holding the 'production function' constant, and see (under growth accounting) if anything much is left over as a 'residual', or (under production function fitting) whether the passage of time itself seems to contribute to output growth over and above what is explained by input growth over time. We argue here that, contrary to widespread views in economics, neither kind of calculation can separate out growth that 'would have occurred without technical advance' from growth that involved technical advance.

Often, it is not recognized adequately that the simple logic of growth accounting is only applicable to the analysis of small changes in inputs and outputs (see Nelson, 1973). The procedure basically involves making estimates of the marginal productivities (or output partial elasticities) of the various inputs that have changed and, in effect, using these to calculate the contribution of input expansion to output growth by using a first-order Taylor series. However, in the case of the Asian tigers, the investments whose contribution to growth is being estimated have cumulatively been very large. While repressed by the format of growth accounting, which usually sets up the calculations in terms of average yearly changes, and thus makes the changes appear relatively small, in the countries in question capital per worker increased more than four times since the 1960s and years of average educational attainment also increased greatly.

The calculations in standard growth accounting take marginal productivities as estimated by factor prices (or output elasticities as estimated by factor shares) as being exogenous. However, under the assumptions of neoclassical production function theory (which lie behind the growth accounting logic), large finite changes in inputs can lead to large finite changes in marginal productivities. For this reason, the factor prices (or factor shares) that are treated as being exogenous in growth accounting need to be understood as endogenous. Thus a 'growth accounting' of the standard sort does not provide a way to calculate growth that would have occurred had there been no technical advance, if input changes are large. Sustained high marginal productivities (output elasticities) of the most rapidly growing factors, which lead a growth accountant to propose that most of the growth is explained by their expansion, could be largely the result of the fact that technical advance offset the diminishing returns that otherwise would have set in.

We know that, in the countries in question, despite the large changes in their quantities, the rates of return on physical capital and on education stayed high. We noted earlier that one explanation is that technologically determined elasticities of substitution, in the sense of standard production function theory, were quite high, and thus significant increases in these inputs relative to others had only a modest effect on marginal productivities as the economy moved along its *ex ante* production function. Under this explanation, a good share of output increase would indeed have occurred without any technical advance. This seems to be the implicit argument of the proponents of the accumulation theory. However, another explanation is that the elasticities of substitution, defined in the standard way, were quite low, and that only the rapid taking on board of new technologies prevented the sharply diminishing returns that one would have observed had these economies stayed with the production functions that existed at the start of the development traverse.

Consider the latter explanation, which we believe is the correct one. Under it, innovation and rapid learning are driving growth. However, a growth accounting of a standard sort might show a very small residual, or even a negative one. The factor shares of the more rapidly growing factors – physical and human capital – would be, and would remain, high, as a consequence of the rapid learning that made their continued expansion productive. These investments themselves would be and would remain high because rapid technical advance kept their returns high. Thus a growth accounting might 'attribute' the lion's share of output growth to input growth. There would be little left to explain in terms of innovation and learning, despite the fact that these are the basic factors driving growth.

The use by some scholars of the Tornqvist index for the weights applied to input increases represents acknowledgement that, if one is interested in the impact on output of finite changes of inputs along a production function, output elasticities can change along the way. But the use of such an index (as by Young, 1993) does not deal with the problem highlighted here. The index uses actual shares, at the end as well as the beginning of the period. But the actual shares at the end of the period can be, and in the case in question almost surely were, affected by the technological changes that occurred over the period. In general, they are not what the shares would have been at the new input quantities, had the production function stayed constant over the traverse.

We want to underline this point because many economists seem to believe that the absence of a large residual in growth accounting is strong evidence that the lion's share of growth resulted from movements along a prevailing production function. This is not so if the input changes involved are large. Growth accounting alone cannot tell whether the relevant elasticities of substitution were large or small, and thus cannot distinguish between the two stories sketched above about the sources of growth. There is an 'identification' problem.

One might think that the fitting of a dynamic production function can avoid this logical limitation of growth accounting, when input changes are large and finite. However, in practice, the identification problem cannot be resolved in this way.

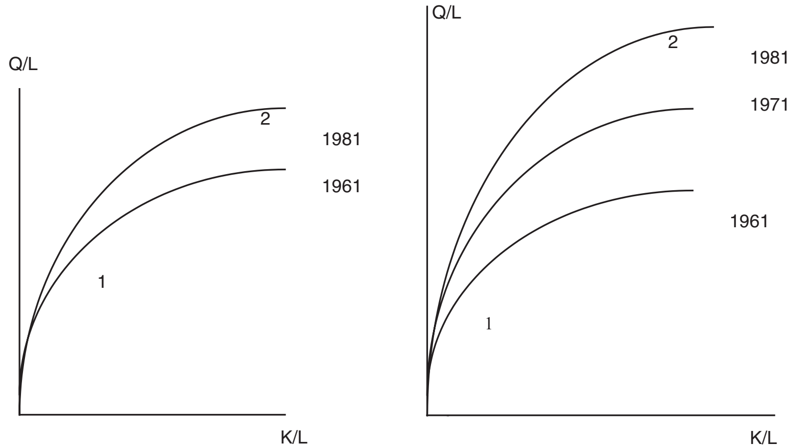
Thus, consider the two 'explanations', depicted in Figure 7.1, for a large increase in output per worker, between time one and time two, associated with a large increase in capital per worker. In the explanation on the left-hand side, much of experienced labour productivity increase would have occurred even had the economy stayed with its production function of period one (the dotted curve). The way the production function is drawn depicts only a weak diminishing return to increasing capital intensity. The firm or economy in question is presumed to know, at time one, how to operate effectively at much higher capital intensities than were employed then, but chooses not to do so because prevailing factor prices made it more profitable to operate at low capital intensity. Between time one and time two, factor availabilities changed.

In contrast, in the explanation on the right-hand side, experienced productivity growth is almost totally the result of the establishment of a new production (the solid curve), in that very little productivity growth would have occurred had the economy remained with its old production function. Under this explanatory story, at time one the firm or economy in question knew very little about how to operate effectively at significantly higher capital intensities. To have increased capital per worker without learning about and learning to use new techniques would very quickly have led to low or zero marginal returns. Thus, the economy, in order to deal productively with the changed factor price regime of period two, had to do a lot of 'learning', or 'innovating', and in fact it did so.

Both explanations fit the data at time one and two. The 'levels' and the 'slopes' of the old production functions are the same at time one, and the levels and slopes of the new production functions are the same at time two. This point was highlighted by Diamond *et al.* (1972), and Nelson (1973), almost thirty years ago. It seems to have been forgotten.

When one 'fits' a dynamic production function statistically (through many rather than just two points and slopes), how does one discriminate between these two explanations? Obviously, one needs to place some restrictions on the form fitted (for example, that the rate and direction of 'technical advance' be constants over the period, or that the underlying production function must always have a particular 'kind of general shape'. Most of the econometric exercises we are concerned with here have imposed relatively loose restrictions, although sufficient to enable a best fitting equation to be calculated. However, even if an equation that looks like the left-hand side explanation wins the 'maximum likelihood' contest (as in Kim and Lau, 1994), if the constraints on functional form are relatively loose it is a good bet that an equation that looks like the right-hand side explanation is not very far behind. Standard regression techniques of the sort that have been employed do not enable confident acceptance of one explanation and rejection of the other.

The graphs drawn in Figure 7.1 are in fact regression estimated from the actual data for Korea's manufacturing sector for the years 1962–91. The dynamic production function fitted to the data is a standard constant elasticity of substitution production function (CES), with two



Q/L = Gross product per worker   K/L = Capital per worker

Figure 7.1 Movements of the production frontier across decades

inputs – capital and labour – and constant returns to scale. To keep the analysis simple and transparent we constrained technological advance to be neutral and constant over the period in question. The key parameters to be estimated are *r*, the rate of technological progress, and *e*, the elasticity of substitution.

In the left-hand figure we forced *e* to be large, 0.9. Since growth of *K/L* then 'explains' a lot of the growth of *Q/L*, the estimated rate of technological change, *r*, comes out low, 0.16. (For regression runs in which we set *e* as greater than one, the estimated rate of technological change was even smaller, and for large values of *e* came close to zero). In the right-hand figure we constrained *e* to be low, 0.2. Since under this constraint the growth of *K/L* cannot 'explain' much of the growth of *Q/L*, the estimated rate of technological progress, *r*, came out high, 0.045. Both of these regressions, and one in which all parameters were chosen by least squares, yield values of R^2 of 0.99, leaving little to choose among the regressions on a statistical basis.

Again, we want to underline the point. The fact that the best fit of a dynamic function provides an explanation for growth in which technological advance plays a small role, and input growth accounts for the lion's share of growth, does not itself provide strong evidence against the argument that, in fact, growth would have been far less if there had not been significant technological advance. Only the imposition of particular constraints on the dynamic production function enables econometric techniques to choose between the explanation on the left-had side and the right-hand side of Figure 7.1. These constraints are basically arbitrary. And the imposition of somewhat different ones can change radically the estimated contribution of technical advance in the attribution.

The authors in question certainly have been careful with their data, and in the use of their methods. The problem is that the methods employed just do not do the job they are thought to do. Nor, at this stage of our argument, are we introducing 'new data', although we agree that the issue is an empirical one. Before considering new evidence, it would seem important to do some rethinking about the kind of data that would discriminate between growth where entrepreneurship, innovation and learning were central, and growth where they were not.

Back to basics

How is one to decide between two different explanations, each broadly compatible with the macroeconomic data, when one stresses 'movements along a production function' and the other emphasizes 'entrepreneurship, innovation and learning'? We propose that, to get an empirical answer requires that one first ask some conceptual questions. What might one mean when one says that an observed change in inputs and outputs simply reflected a move along a production function? What might one mean if one argued it was not that simple, but that entrepreneurship and innovation were in fact involved? If we agreed on answers to these conceptual questions we might be able to agree on what kind of empirical evidence would be relevant.

Regarding what we economists seem to mean by 'a move along the production function', reflects on the simple treatment in undergraduate microeconomics texts. The production function, there, is said to be the 'efficiency frontier' of the 'production set' – the set of all input–output combinations from among which a firm can choose. One way of explaining the set to students is to say that a firm 'knows' a certain set of production techniques or activities, and the production set is generated by different levels and mixes of those activities. In any case, the firm is viewed as both 'knowing about' each of the alternatives, and 'knowing how' to do whatever is associated with achieving the input–output vector associated with each.

The verbal articulation may admit that there might be modest 'set-up' costs associated with marshalling and organizing to shift operations to a point within the set that is different from what the firm is doing currently, and that some adjustments (another form of set-up cost) might be required to get the new choice operating smoothly, although these shift costs are generally repressed in the formal modelling. However, it seems inconsistent with the 'operating within the production set' idea if the set-up costs for shifting to a new point involved doing a lot of exploratory 'search and study' to identify and get a better feel for alternatives that, up to then, had been unfamiliar to the firm, and the 'adjustment' involved a lot of trial-error-try-again learning by doing and using. At least, it would seem inconsistent if the results of searching and learning were highly uncertain, both to the firm *ex ante*, and to an economist trying to predict what the results would be.

Of course, a plausible interpretation of the production set idea might admit a certain amount of statistical uncertainty regarding inputs and outputs, particularly if there were unknowable outside forces, such as the amount of rainfall, that have a bearing on the process. But if the decision-maker in question has only very rough ideas about the consequences of trying to do something, and initially about how to do it, that something does not seem to be an activity that can be regarded as within the unit's production choice set. The production set of a firm would appear to be limited at any time to those things the firm knows about and knows how to do, with confidence and skill. At least, that is how economists implicitly define the concept.

On the other hand, a move that involves a lot of study of initially hazy alternatives, or R&D where even the nature of the outcome is not clear in advance, would, according to these criteria, be regarded as a 'technological' change or 'innovation' for the firm in question. We do not see how such a move possibility can be regarded as one 'along a prevailing production function', if economists adhere to what they teach about the meaning of choice sets.

We call attention to the fact that, under the way we are proposing the distinction be drawn, a firm's production set in principle could be very extensive. Indeed, much of what some versions of the new neoclassical growth theory treat as 'technological advance' would, under the principles suggested here, be regarded as moves along a firm's prevailing production function. In these models (see Romer, 1990) investments in R&D are strictly up-front costs required to make a product or technique operational. But in these models (if not in fact) R&D is strictly a set-up cost to make an activity the firm always knew about available for use. There are no Knightian uncertainties involved.

However, once we get away from particular abstract models, most economists who have studied the processes empirically understand that the introduction into the economy of products or processes significantly different from any employed before does not look like a move 'along a prevailing production function'. It is well documented empirically that, while theoretical engineering calculations at any time encompass a wide range of techniques not yet brought to practice, the bringing to practice of new technology invariability involves 'up-front' research and development costs, with Knightian uncertainties at least early in the process (see Nelson and Winter, 1982; and Rosenberg, 1994). While R&D can resolve some of these uncertainties, there are uncertainties in the R&D process itself. Further, even after R&D, there almost always are 'bugs' at the start, and it usually takes some time before the operation is really under control. In many cases, the attempts at innovating prove to be unprofitable, and need to be abandoned, or radically revised.

Of course, in this chapter we are dealing with the adoption of technologies that, while new to the firm or country, are not new to the world. The issue, then, is whether such changes in the behaviour and performance of firms in the NICs can be explained meaningfully as changed choices within largely unchanging choice sets.

The accumulationists clearly have in mind that, if a technology is in effective use in one country, there are methods that firms in other countries can use to take on board that technology at relatively low cost, and without significant uncertainties regarding the outcome of their efforts. Quite often, detailed descriptions are available. Consultants can be hired who are familiar with the practices involved. In many cases, assistance can be gained from the firms who are operating the technology, although some licence fees may be required.

The assimilationist, in contrast, is sceptical about easy 'technology transfer'. To be sure, for many of the technologies that the firms in the NICs adopted there were available engineering texts, articles and so on. Blueprints and specific handbooks could often be obtained. There are many consultants for hire.

However, the assimilationist would stress that such a move invariably involves not only 'up-front' costs of identifying, learning about, and learning to master the technique in question, but also significant uncertainties. The range of options is hazy. Things often do not work out as expected. Consultants can seldom guarantee success. Inevitably there is a lot of learning by doing and using. The costs, and the uncertainties, are greater the further the technique being adopted is from those the firm has in fact employed. In many cases, major changes in firm organization may be required. The firm may need to learn to sense new markets. Firms attempting these changes can and often do fail. Those that succeed do so because they learn to do things successfully that they simply could not do before; that is, they succeed by expanding their production sets.

What does the evidence indicate?

We can return now to the question of what kind of evidence one would need to determine whether an observed change was within a prevailing capability or choice set, or required an expansion of the set of things the organizations in question knew how to do. The previous section argued that the standard data and techniques for deciding simply do not do the job. We propose here that the kind of evidence that is relevant involves examination of process, not simply time paths of inputs and outputs, and that the persuasive data are to be found at a quite low level of aggregation.

A major problem with highly aggregated economic data is that it masks the magnitude and even the nature of the allocational changes going on. Thus, earlier, we noted that in the 1990s Korean and Taiwanese manufacturing firms were heavily engaged in producing products that in the 1960s they were not producing at all. This is illustrated strikingly by Table 7.1 for Taiwan. In particular, note Taiwan's production of electronic goods, which by the late 1980s accounted for roughly 21 per cent of Taiwanese manufacturing exports. In 1960 virtually no electronic goods were produced in Taiwan.

Such a change in the allocation of activity within the manufacturing sector almost certainly would be associated with considerable turnover of firms, with companies going out of business in the declining sectors, and new firms entering the expanding fields. And within the expanding areas one would expect to see a certain amount of turnover as some firms try and fail, while others succeed. Unfortunately, we do not have the firm turnover data that is directly relevant to the phenomena we are characterizing.

However, there are data on the number of firms of different sizes in Korea and Taiwan for several years, and a summary of these data is presented in Table 7.2. The pattern is roughly what one would expect under the assimilationist's story. There has been a striking decline in the number of very small firms, most of which were probably locked into old technologies and producing traditional products, and a sharp rise in the number of middle-sized or larger firms; we conjecture that a

Table 7.1 Changes in physical production levels of selected industrial products Taiwan, 1960–90

Product	1960	1990
Artificial fibres (million tons)	1 762	1 785 731
Polyvinyl chloride (million tons)	3 418	920 954
Steel bars (million tons)	200 528	11 071 999
Machine tools	0	755 597
Sewing machines	61 817	2 514 727
Electric fans	203 843	15 217 438
Television sets	0	3 703 000
Motorcycles	0	1 055 297
Telephones	0	1 055 297
Radios	0	5 892 881
Tape recorders	0	8 124 253
Electronic calculators	0	44 843 192
Integrated circuits (1000s)	0	2 676 865
Electronic watches	0	5 115 695
Shipbuilding (tons)	27 051	1 211 607

Source: *Taiwan Statistical Data Book* (1992), Council for Economy Planning and Development, Republic of China, Taipei, table 5–6c.

large share of these were new firms entering the new product fields, or older firms that succeeded in taking on board modern technology. In the early 1970s, the productivity of these larger firms was strikingly higher than that of the small firms (according to the story we are proposing) they were replacing.

However, to get at the details of what was going on would seem to require the study of individual firms. Only by studying firms can one see just what was involved when they came to master new technologies and learn what was needed to operate in new product fields.

Happily, during the 1980s and 1990s, several scholars developed detailed studies of Taiwanese and Korean manufacturing firms, tracing the sources of the firms' rapidly growing range of manufacturing competencies. Thus, Alice Amsden (1989) has provided a detailed history of a Korean textile firm, which describes what was going on over a period of time when it achieved very significant productivity gains. Table 7.3 shows what happened to machine and labour productivity during the decade after 1977, when it purchased most of its capital equipment. The reduction in worker hours per unit of output was considerable, particularly in spinning. Amsden explains the productivity growth in terms of active learning. Early in the period, its

Table 7.2 Percentage distribution of employment by firm size

	Number of employees					
	4–9	10–19	20–49	50–99	100–499	500+
Taiwan						
1954	18	13	14	9	16	31
1961	18	10	14	8	17	34
1971	8	7	11	9	29	37
Index of value per worker, 1971		100	91	100	117	259
Korea						
1958	17	16	21	13	21	12
1963	15	14	16	12	21	22
1975	4	5	8	9	30	44
Index of value per worker, 1971		100	133	193	256	304

Source: Ho (1980), tables 3.1, D2, D3.

Table 7.3 Learning in a Korean textile factory

	1977	1986	1986*
Labour productivity			
Ring spinning, (kg per manhour)	52.4	78.5	156.25
Open end spinning,** (kg per manhour)	137.1	210.3	324.30
Weaving (metres per manhour)	216.2	224.1	360.36
Machine productivity			
Ring spinning, (kg per spindle)	0.20	0.23	0.21
Open end spinning,** (kg per rotor)	0.91	1.26	1.11
Weaving, metres per loom	36.1	35.4	39.8

Notes: * Relative to international best practice; ** Initial year is 1979.
Source: Cols 1 and 2 adapted from Amsden (1989), table 10.4. Col. 3 calculated from col. 2 plus coefficients from Pack (1987), tables 3.1 and 3.2, and the calculations underlying those tables.

foreign equipment-suppliers provided technical assistance. Later in the period it employed its own engineers to help it increase productivity. Note that, in 1986, while it had become much more efficient than it had been a decade earlier, its labour productivity was still lower than that in comparable plants in advanced industrial countries, a phenomenon not consistent with a move along a freely available international production function.

For our purposes, some of the most interesting sets of firm studies are those undertaken by Michael Hobday of Korean and Taiwanese electronics companies. Hobday describes in detail how these firms started out, usually producing quite simple products, and then moved on progressively to more complex ones. In most of the cases Hobday studied, these new complex products were first made to order for their foreign customers who, in the early stages, provided detailed engineering instructions and assistance. Gradually, however, many of these companies became able to do their own design work. In a number of cases they have moved on to sell under their own brand label. Throughout the history of these firms, one can see them actively working to learn to do better the things they were doing and to be able to do more sophisticated and profitable things. In the early stages, this learning involved reversed engineering. As the companies began to do their own design work, this engineering effort began to be counted as R&D.

Linsu Kim (1997) provides a set of analyses of Korean firms, in several different industries, that show much the same phenomena as does Hobday's (1995) study. The firms started out using relatively unsophisticated technologies and learned, over the years, to master progressively more sophisticated ones. By the 1990s many of these firms were approaching the technological frontier. But the paths they took were not simple, and success never was guaranteed.

The story about the development of Korean and Taiwanese firms told by Amsden, Hobday and Kim, is strikingly similar to that told by Odagiri and Goto (1997) in their study of how Japanese industry learned about and learned to master the technologies of the West in the years between the Meiji Restoration and the advent of the Second World War. They find that a major amount of searching, exploring, trying, failing and learning was required before Japanese firms acquired proficiency in the Western technologies they were adopting and adapting. The decisions of company managers to adopt the new ways involved major uncertainties. Odagiri and Goto stress their 'entrepreneuria' nature, and the innovation and learning involved. Our argument is that Korean and Taiwanese firms went through much the same process half a century later.

Table 7.4 shows the rise in accounted R&D and patenting by nationals in Taiwan. A similar progression from engineering work focused largely on mastering and adapting foreign technology, to work on designs sufficiently new that the effort could legitimately be called R&D, occurred in Korea. And, of course, the same phenomenon occurred in Japan in the early postwar period.

To return to our basic analytical argument, we do not think that the industrial development of Korea and Taiwan since the 1950s, or of Japan a half century earlier (see Saxenhouse, 1974; as well as Odagiri and Goto, 1997), can be interpreted as 'moving along production func-

Table 7.4 R&D and patenting activity in Taiwan

Year	R&D/GDP	Total patents	Taiwanese national's patents	Foreign patents
1981	0.95*	6 265	2 897	3 368
1986	0.98	10 526	5 800	4 726
1991	1.65**	27 281	13 555	13 726

Notes: * 1984; ** 1990.
Source: *Taiwan Statistical Data Book* (1992), tables 6.7, 6.8.

tions', at least if that term connotes changing choices within a largely unchanging choice set. On the other hand, if the kind of entrepreneurship, innovation and learning on the part of firms revealed in the case studies is considered to be perfectly consistent with the notion of 'moving along a production function', we do not know what that concept would exclude, and hence it becomes meaningless.

Do the differences matter, and if so, how?

The differences between the two theories would appear to matter for two different reasons. The first is, simply, regarding how one understands what happened. What lies behind the Asian miracle? The second is that the two theories might imply somewhat different things regarding appropriate economic development policy. What kinds of government policies are helpful, and what are the lessons for other countries?

It is apparent that, for many economists, one of the strongest attractions of the accumulation theory is that it is clean and simple, and its basic outlines conform with the general theory about economic activity that one finds in modern economic textbooks. It is at once delightfully iconoclastic, and comfortably conservative, to take the miraculous out of the Asian miracle by proposing that it all was a simple matter of moving along a production function. No appeal is needed to the idea of entrepreneurship or innovation, the sources of which might very well lie outside the effective province of neoclassical economics.

It also is clear that a major source of resistance to the assimilation theory is that it seems a complex theory that raises as many questions as it answers. This raises suspicions that the assimilation theory cannot be formulated cleanly. It is a comfort, therefore, that a simpler, more familiar theory seems capable of providing all the explanation needed.

And yet, what is at odds intellectually may be only a small part of the corpus of traditional economic theory. Moreover, that particular part, which proposes that production sets can be defined sharply, and that there is a clear distinction between moving along the production function and having the production function shift, came into economics a relatively short time ago. Perhaps these particular conceptions are not needed for most standard economic arguments, and maybe they have been accepted too uncritically in any case.

A strong argument can be made that the assimilationists' perspective is quite consistent with an older set of ideas in economics. The idea that economic growth can be explained by increases in the factors of production, and by improvements in their productivity, goes back at least as far as John Stuart Mill. However, a striking feature of the earlier analyses of economic growth, as contrasted with the more contemporary treatments, is that there was no compulsion to make a sharp separation between the contributions of different sources of growth. For Adam Smith, increases in the size of the market, discoveries of better ways to perform a task, growing mechanization, and changing organization of work, all go together. They would seem to also do so in Mill. The early post-Second World War growth accountants, in particular Moses Abramowitz, also stressed the interaction of technological advance, growing physical capital intensity of production, increasing exploitation of scale economies, rising educational attainments, and changes in the organization of production, as factors behind experienced economic growth. The question of which of these factors should be interpreted as moving the economy along a production function, and which should be regarded as shifting it, seems not to have been of major concern to these authors.

In the second section we argued that standard techniques do not permit a clear separation between movements along and shifts in the production function. Now we would like to argue that the very notion that one can make such clear splits, even in principle, may not be a useful theoretical premise.

In particular, we would like to argue that 'innovation' in practice is a matter of degree, not kind, and that our growth theory ought to recognize this explicitly. For any firm or organization at any time, there are some activities that are under practised control, some that are not so at present but seem easy to learn, others harder, and others at present impossible but perhaps with research and experimentation achievable over the long run. The problem with now-standard production theory is that it does not recognize these continuities, but rather presumes a sharp rift between the known and the unknown.

The case studies of firms, briefly discussed in the third section, show them moving from the known to the unknown, but cautiously, and drawing from the known as much as they can. Yesterday's unknown becomes today's known, and the firms venture further. An effective theory of what has been happening requires, we believe, abandonment of the notion that production sets at any time are sharply defined, and thus that there is a clear distinction between moving to another perceived point and innovation. Rather, there is a continuum.

If it is recognized explicitly that that distinction is in fact fuzzy, does not that mean there is a fuzzy theory? Not at all. One of the striking features of the various 'evolutionary models' of economic growth that have been built since the 1980s is that, within them, innovation is treated as a matter of degree, firms move step by step into the unknown, and in so doing seldom move very far from the known. In section 6 we develop an evolutionary model that we think fits many of the facts of the Asian miracle.

Abandoning the sharp distinction between moving along a production function and innovation is clearly a big step analytically. Such a step would involve placing learning and adaptation at centre stage of the behavioural analysis, and relinquishing analytical techniques and arguments that presume that 'profit maximization' is something that managers are in fact able to achieve, rather than something they strive for intelligently. Yet it is arguable that most of the important and useful propositions about the role of markets and competition depend on the latter rather than the former.

The notion that competition tends to force prices down towards costs, and to stimulate reform or elimination of high-cost producers, goes far back in economics. The argument does not depend on the existence of sharply-defined production sets, or the achievement by firms of policies that maximize profits, given the full set of theoretical alternatives. It is intelligent striving that does the job. Similarly, the argument that a change in factor prices will induce behaviour that economizes on the factor whose cost has risen does not require either sharply-defined production functions or maximization, but only intelligent striving.

What are the policy implications of taking an assimilationist, or evolutionary, view on what happened in the Asian miracle? Are the policy prescriptions fundamentally different under an assimilationist theory than under an accumulationist theory? In many ways, the policy prescriptions are in fact quite similar, although the reasons behind the arguments differ somewhat.

Both neoclassical and assimilationist theories put considerable weight on investments in human capital. By stressing the importance of innovation and learning, and the role of an educated workforce in these processes, the assimilationist might push even harder on the education front than would a modern neoclassical economist.

No disagreement on the importance of investment in physical capital. However, the assimilationist would highlight the role of such investments as a vehicle for taking aboard more modern technologies, and stress that, if capital formation is not linked to effective entrepreneurship, the returns to investment are almost certain to diminish greatly after a point. On the other hand, the assimilationist would point to effective entrepreneurship as a key vehicle for keeping investment rates of return high, and would put less emphasis on simply trying to lift up the saving rate.

Both theories stress the importance of exporting. However, here too the reasons for the emphasis are somewhat different. The assimilationist sees exporting as an extremely important vehicle for learning, as well as a way of exploiting evolving comparative advantage. Thus the assimilation theorist might be even stronger on the importance of exporting, and willing to bias the incentive system to induce firms to try to export.

Both theories stress the essential role of private enterprise, profit incentives, and an environment that stimulates managers to make decisions that enhance economic development. A neoclassicalist would focus on getting the prices right and making necessary public infrastructure investments. The assimilationist would take a somewhat more complex line on both of these matters. In particular, an assimilationist might stress the role of government funding and organization in building up a national scientific and technological infrastructure from which firms can draw assistance. But under both theories, it is the energy of private enterprise that is key, and under both there is deep scepticism about the value of detailed government planning.

Both neoclassical and evolutionary theorists stress the great importance of competition. However, here too the reasons differ somewhat, with the proponent of evolutionary theory pushing competition, especially in contexts where innovation is both important and risky. From this point of view, competition is valuable largely because choice sets are not clear, or not clearly defined, and it is highly valuable, therefore, to get a lot of things tried.

So, the policy differences between the theories may be significantly smaller than the conceptual or analytical differences. This should not be a surprise. Economists were stressing the importance of profit incentives, and competition, and the dangers of government planning, long before the idea of a sharply-defined production set came into fashion. Indeed, one can find these basic arguments in Adam Smith's *Wealth of Nations*.

A simple evolutionary model

The model we offer here is totally devoid of substitution possibilities within a given technology. It certainly would be strange to characterize its dynamics and the development pattern it generates as a movement 'along a prevailing production function'. Rather, all development takes place through the shifting of resources from one technology, which we shall call 'craft', to another, which we shall call 'modern'. That is, 'assimilation' is what is driving development here. Yet the growth pattern it generates could be interpreted by a growth accountant or a fitter of dynamic production functions as indicating that 'accumulation' was the basic story. Within the model, expansion of physical and educational capital per worker is an essential part of the process by which the economy takes on board modern technology. But, on the other hand, accumulation without assimilation yields no returns.

Within this model, a basic constraint on the rate of assimilation is the vigour of entrepreneurship. There are always profits to be made by expanding the modern sector. The vigour of entrepreneurship determines the rate at which this happens. We do believe that this depicts accurately an essential ingredient of the Asian miracle. The rapid expansion of human capital, another essential ingredient in our view, also plays a central role in this model, being necessary for the rapid expansion of the modern sector.

The model does not contain a third ingredient that we consider central, however: the rapid learning that took place in a firm after modern technology was first taken on board. The model assumes, in effect, that such learning took place instantly and was once-and-for-all, while in fact the firms moved progressively into increasingly complex technologies. Here we choose to keep the model simple and abstract away from the cumulative nature of learning.

Thus assume that there are two different kinds of fixed factor constant returns to scale technologies, which we shall denote as c for craft and m for modern. Capital per unit of output is the same in the two technologies. On the other hand, output per unit of labour is higher in the modern sector than craft. So also, then, is capital per unit of labour. If factor prices in the two sectors were the same, unit costs using modern techniques would be lower than costs using craft technology. However, the modern sector requires 'educated' labour, while education is not necessary or productive in craft technology.

At the start of the development traverse, almost all of capital and labour is in the craft sector. However, we assume that there is a tiny amount in the modern sector which serves, in effect, to 'seed' the development process.

At any time, output per unit of labour in the economy or industry as a whole will be the weighted average of labour productivity in the two technologies, the weights being the proportion of labour employed by each of the technologies. Let a_c be output per unit of labour in craft technology and a_m be output per unit of labour in modern technology, with $a_c < a_m$. Then:

$$Q/L = a_m L_m/L + a_c L_c/L \quad (7.1a)$$
$$Q/L = a_c + (a_m - a_c) L_m/L \quad (7.1b)$$

As L_m/L grows over the development traverse, so does Q/L. Since capital per unit of output is the same in the two sectors, an increase in L_m/L is accompanied by a rise in K/L. Indeed, within this model, Q/L and K/L grow at the same rate.

Within our model, a shift in the proportions of capital in the two sectors drives development. We assume that the price of the product is the same whether it is produced by modern or craft technology, and is constant over time. The latter can be rationalized by presuming that the product is sold on world markets and hence is insensitive to the quantity produced within the particular economy in question. We also assume that the cost of capital is the same in the two sectors. This means that differences in labour costs is the only factor affecting the relative profitabilities of the two technologies. We could modify these assumptions, but making them enables us to tell a cleaner story.

Let w be the price of labour in the craft sector, and gw its price in the modern sector, with $g > 1$. Thus g (for graduation) reflects an education premium. We assume, however, that g never is so large as to completely offset the productivity advantages of modern technology.

If one uses a prime over a symbol to denote an inverse, then the difference between the two sectors in cost, and profit, per unit of output, and capital, can be written:

$$\text{dif } C = w(a'_c - ga'_m) \quad (7.2)$$

The higher profitability of modern technology over craft provides an incentive to shift resources from the latter to the former. Within this model, the strength of the response is determined by the vigour of entrepreneurship, denoted by e:

$$d/dt(\log K_m/K_c) = ew(a'_c - ga'_m) \quad (7.3)$$
$$d/dt(\log K_m/K) = ew(a'_c - ga'_m)\,(1 - K_m/K) \quad (7.4)$$

If w and g are constants, the time path of K_m/K (and Q_m/Q) will trace out a logistic function. L_m/L will be increasing as these variables grow, but lagging behind them. Of course, at the limit they all approach one. If w increases as development proceeds, but not g, the rate of expansion of the modern sector relative to the craft will be accelerated, reflecting that, since modern technology saves on labour, an increased w increases its cost advantage. An increase in the education premium, g, over the development trajectory will diminish the cost advantage of modern technology. On the other hand, a decline in g, say, as educated labour becomes more plentiful, will enhance it.

We know from Equation (7.1) that, as capital and labour shift to the modern sector, K/L and Q/L will increase. If the amount of educated labour is responsive to demand, human capital will also be increasing. An economist looking at aggregate data would probably conclude that growth of Q/L was caused by the growth of physical and human capital per worker (and, indeed, such growth of capital was required for growth), and propose that growth arose basically from 'movements along the production function'. This 'explanation' would repress two things. First, the force driving growth was the progressive taking on board of modern technology, a technology virtually absent in the economy before development began. And second, while the profitability of employing modern technology was motivating the shift, the rate at which the modern sector replaces the craft was being determined by the strength of entrepreneurship. On the other hand, the traditional analysis would be right about the rate of growth of human capital being an enabling factor.

Thus consider two economies with exactly the same initial conditions, facing exactly the same opportunities to take on board modern technology, and having the same input supply elasticities. In one, the strength of entrepreneurship, e, is high, and in the other it is low. The expansion of the modern sector, the growth of physical capital intensity, increases in human capital, and the advance of labour productivity, would all be faster in the former than the latter. An economist, thinking in terms of production functions, would try to explain the differences in terms of different rates of 'accumulation', but the key factor behind the scenes would be differences in entrepreneurship.

Behind the scenes in the model, growth of human capital is a enabling element. Other things being equal, a high e (resulting in rapid growth of the modern sector) will cause a rapid growth of demand for educated labour. If increased supply is not forthcoming at the prevailing premium for educated labour, under various ways of modelling, the dynamics g will rise. This will slow down the rate of growth of the modern sector associated with a given e. On the other hand, a rapid expansion of the educated workforce can be absorbed productively only if e is high.

Just as, within this model, a high e tends to draw out expansion of human capital, a high e tends to generate high profits in the industry as a whole, and hence is a source of the saving to finance the investment in the modern sector. Both effects are, of course, moderated by 'supply side' variables. To keep this presentation simple we have not introduced these into this model explicitly.

Within this model, development is a process driven by a disequilibrium. The disequilibrium, and the rate at which it is eliminated, shows up in this model in the behaviour of capital's share over the development traverse. Set the constant product price as the numeraire. Then the share of capital in total income is:

$$S_K = (1 - wa'_c)\, Q_c/Q + (1 - gwa'_m)\, Q_m/Q \qquad (7.5a)$$
$$S_K = (1 - wa'_c) + w(a'_c - ga'_m)\, Q_m/Q \qquad (7.5b)$$

The first term of Equation (7.5b) is capital's share in the craft sector. The second term is the amount by which capital's share in the modern sector exceeds its share in the craft sector times the relative size of the modern sector.

Let b be the common capital output ratio in the two sectors, and r the equilibrium rate of return on capital. Assume that, at the start of the development traverse, the craft sector is in equilibrium. Then, while capital's share in the modern sector is greater than br, since Q_m/Q is very small, the share of total capital on the total industry is close to br at the start of the traverse. We also assume that, as development proceeds and the modern sector grows relative to the craft, wg grows to squeeze out excess profits in the modern sector. At the end of the development traverse, then, capital's share again is rb'. However, in between, during the course of development, capital's share will exceed rb'.

While the details depend on the exact specification, under plausible assumptions capital's share will take an inverted U-shaped path over the development trajectory. As development proceeds and the modern sector expands, capital's share will first rise, since quasi-rents per unit of capital are higher in the modern sector than in the craft sector, and a growing share of capital in that sector will more than offset the fact that rising wages will press down on rents per unit of capital in both sectors. Later, as the modern sector comes to be most of the economy, rising wages will diminish capital's share.

If one notes Equation (7.2), one can see that the expression before Q_m/Q in Equation (7.5b) is proportional to the rate at which capital is being shifted from the craft to the modern sector, and hence the rate at which output per worker and capital per worker are growing. Thus capital's share will be high when capital and output are growing most rapidly. A growth accountant would naturally assign a good share of the credit for growth of output to growth of capital. If the supply of educated labour just keeps pace with the growth of employment in the modern sector, human capital will also be growing most rapidly when output is growing fast.

The foregoing captures the spirit of our argument in the text that, in the Asian miracle, both large investments in human capital and forceful entrepreneurship were key ingredients, and they complemented each other strongly. Without the ability and inclination to expand human capital greatly, aggressive entrepreneurship would have been stymied. Without aggressive entrepreneurship, the returns to investment in human capital would have been low. And when both of these elements were present, together they made for high and rising profits in the modern sector, which provided the finance for the large investments in physical capital that were necessary for rapid assimilation.

References

Abramovitz, M. (1986) 'Catching Up, Forging Ahead, and Falling Behind', *Journal of Economic History*, vol. 46, no. 2, June, pp. 385–406.

Abramovitz, M. (1989) *Thinking About Growth* (Cambridge University Press).

Amsden, A. (1989) *Asia's Next: South Korea and Late Industrialization* (New York: Oxford University Press).

Dahlman, C. (1994) 'Technology Strategy in East Asian Developing Economies' Unpublished manuscript, July.

Dahlman, C., Pack, H. and Westphal, L. W. *Industrialization in Developing Countries* (forthcoming).

Diamond, P., MacFadden, D. and Rodriguez, M. (1972) 'Identification of the Elasticity of Substitution and the Bias of Technical Change' in D. MacFadden (ed.), *An Econometric Approach to Production Theory* (Amsterdam: North-Holland).

Ho, S. (1980) 'Small-Scale Enterprises in Korea and Taiwan', *World Bank Staff Working Paper No. 384* (Washington, DC: World Bank).

Hobday, M. (1995) *Innovation in East Asia: The Challenge to Japan* (London: Edward Elgar).

Kim, L. (1997) 'Dynamics of Korea's Technological Learning', *From Imitation to Innovation* (Boston, MA: Harvard Business School Press).

Kim, J. I. and Lau, L. J. (1994) 'The Sources of Economic Growth in the East Asian Newly Industrialized Countries', *Journal of Japanese and International Economics*, vol. 8, no. 2, pp. 235–71.

Krugman, P. (1994) 'The Myth of Asia's Miracle', *Foreign Affairs*, December, pp. 62–78.

Nelson, R. R. (1968) 'A Diffusion Model of International Productivity Differences in Manufacturing Industry', *American Economic Review*, vol. 58, pp. 1219–48.

Nelson, R. R. (1973) Recent Exercises in Growth Accounting: New Understanding or Dead End?', *American Economic Review*, vol. 73, pp. 462–8.

Nelson, R. R. (1995) 'Recent Evolutionary Theories About Economic Change', *Journal of Economic Literature*, March, pp. 48–90.

Nelson, R. R. and Phelps, E. (1966) 'Investment in Humans, Technological Diffusion, and Economic Growth', *American Economic Review*, vol. 56, May, pp. 69–75.

Nelson, R. R. and Winter, S. (1982) *An Evolutionary Theory of Economic Change* (Cambridge, MA: Harvard University Press).

Odagiri, H. and Goto, A. (1997) *Technology and Industrial Development in Japan* (New York: Oxford University Press).

Pack, H. (1987) *Productivity, Technology, and Industrial Development* (New York: Oxford University Press).

Pack, H. (1988) 'Trade and Industrialization', in H. B. Chenery and T. N. Srinivasan (eds), *Handbook of Development Economics* (Amsterdam: North-Holland).

Pack, H. and Westphal, L. E. (1986) 'Industrial Strategy and Technological Change: Theory vs Reality', *Journal of Development Economics*, vol. 22, pp. 87–128.

Pack, H. and Westphal, L. E. (1992) 'Technology Gaps between Industrial and Developing Countries: Are There Dividends for Latecomers?', *Proceedings of the World Bank Annual Conference on Development Economics*.

Rodrik, D. (1994) 'Getting Interventions Right: How South Korea and Taiwan Grew Rich', Unpublished manuscript, August.

Romer, P. (1990) 'Endogenous Technological Change', *Journal of Political Economy*, pp. 71–192.

Rosenberg, N. (1994) 'Uncertainty and Technological Advance' Unpublished manuscript.

Saxenhouse, G. (1974) 'A Tale of Japanese Technological Diffusion in the Meiji Period', *Journal of Economic History*, March, pp. 149–65.

Westphal, L., Kim, L. and Dahlman, C. (1985) 'Reflections on Korea's Acquisition of Technological Capability', in N. Rosenberg and C. Frischtak (eds), *International Technology Transfer* (New York: Praeger).

World Bank (1993) *The East Asian Miracle: Economic Growth and Public Policy* (New York: Oxford University Press).

Young, A. (1993) 'The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Enterprise', Sloan School, Massachusett Institute of Technology, July.

8

Technological Globalization of National Systems of Innovation?

Daniele Archibugi and Jonathan Michie

Knowledge and technological innovation play a crucial role in economic activities. While this has long been recognized by managers, scientists and engineers it is only really since the early 1980s that economists have devoted much effort to studying the way in which knowledge leads to the generation and diffusion of technological innovation. This attention has, however, produced a vast literature which has begun to shed some light into the 'black box' of the relationship between technology and the productive process (see, in particular, Rosenberg,). The initial hypotheses in a handful of pioneering works during the 1950s and 1960s on the economic determinants and impact of innovation have since been corroborated by a substantial amount of theoretical and empirical research.[1]

The most fruitful lesson gained by more recent research is that technological change should be explored within the social fabric in which the innovative activities are developed and used. Innovation is far more than just a series of isolated events shaped by enlightened inventors, forward-looking entrepreneurs or dynamic corporations. Certainly, individuals and firms play a crucial role in the development of specific innovations but the process which nurtures and disseminates technological change involves a complex web of interactions among a range of different subjects and institutions (see David and Foray, 1995; Smith, 1995).

To map these interactions, however, is not easy. Innovation-related information flows are of a multifarious nature:

(i) They take place through both market and non-market transactions. A substantial amount of technology and knowledge transfer takes place regardless of any economic incentives. Individuals imitate and learn; and know-how is often exchanged informally and voluntarily. See von Hippel (1987); Carter (1989); Schrader (1991); Pasinetti (1993); Lundvall and Johnson (1994).

(ii) Such flows can take the form of either tangible or intangible assets. Firms use a variety of sources to innovate: a piece of machinery and a scientific paper may both be important sources of innovation. See Pavitt (1984), von Hippel (1988), Archibugi *et al.* (1991) Evangelista (1995).

(iii) They involve not only businesses but also public institutions. Universities, research centres and other government agencies play a crucial role in fostering technological advance, as do profit-seeking business firms. See Nelson (1987), Dasgupta and David (1994), Stephan (forthcoming), Metcalfe (1995).

These various aspects of the process are unlikely to be 'captured' in their entirety simply by looking at standard economic variables such as prices and quantities alone. To understand technological change it is crucial to identify the economic, social, political and geographical context in which innovation is generated and disseminated. This space may be local, national or global. Or, more likely, it will involve a complex and evolving integration at different levels of local, national and global factors.

The relative importance of national and global forces has been the subject of a vast literature. Some authors have claimed that the current process of globalization is eroding the significance of nations as meaningful subjects of technological change (see Chesnais, 1994). Others, on the contrary, have argued that the significance of globalization has been overemphasized, since the bulk of firms' innovative activities are still carried out in their home countries (see Patel and Pavitt, 1991).

The thesis which might be dubbed 'techno-nationalism' is not necessarily contradicted by what might at first sight appear to be the alternative thesis, of 'techno-globalism'. Rather, the two concepts describe two strictly interrelated aspects of contemporary technological change. Certainly, a globalized economy is transforming the landscape for the generation and diffusion of innovation, but this does not appear to decrease the importance of national characteristics or, even less, of national institutions and their policies. On the contrary, by magnifying the potential costs and benefits that will result from any one country's competitive advantage or disadvantage – as a growing proportion of the home market risks being lost to imports, while a growing proportion of domestic output may be dependent on winning export orders – globalization will increase the impact national policy will have on domestic living standards.

Before taking this discussion further, however, some consideration is required of the two key concepts of national systems of innovation on the one hand, and the globalization of technology on the other – as well as of the main actors (broadly, private firms and public institutions) through which these systems and trends evolve.

Concepts and actors

National systems of innovation

The importance of nation-specific factors in developing technological innovation has been affirmed boldly since the mid-1980s. Chris Freeman (1995) introduced the concept of 'national systems of innovation' (NSI) to describe and interpret the performance of the economically most successful country of the post-war period, Japan. Over subsequent years the concept has experienced a remarkable diffusion and has been applied to several countries and different areas.[2] As Nelson and Rosenberg (1993) noted: There clearly is a new spirit of what might be called 'techno-nationalism' in the air, combining a strong belief that the technological capabilities of a nation's firms are a key source of their competitive prowess, with a belief that these capabilities are in a sense national and can be built by national action.

Studies in this field were pioneered by two research teams. The first team, led by Bengt-Ake Lundvall (1992) at the Aalborg University Centre, investigated the analytical content of the notion of national systems of innovation by looking at the role played by users, the public sector and financial institutions. The second team, coordinated by Richard Nelson (1993), assembled a number of case-studies to describe the main features of the innovative systems of high-, medium-, and low-income countries. More recently, the OECD has taken up the idea of national systems of innovation and is making an attempt to operationalize it through the collection and analysis of indicators. In particular, their analysis is focused on the financial dimension, the interconnections among the various institutions, and the distribution of knowledge across national agents.

Although the concept of national systems of innovation is defined and applied differently,[3] the various authors share the view that nation-specific factors play a crucial role in shaping technological change. Some of the factors are institutional, such as education, public support to industrial innovation, and defence-related technology schemes. Others are rooted in history, and concern the culture, size, language and vocation of a nation. Crucial to the definition of a national system is how the different parts, such as universities, research centres, business firms and so on interact with each other.

The globalization of technology

New technologies have always played a crucial role in the processes of economic and social globalization. Air travel, computers and satellite-based communications make possible an ever-expanding degree of information exchange, commodity trade and individual contact across the globe. Indeed, it is often argued that the current globalization would be impossible without such technologies (see Giddens, 1991). Communication and transport technologies, however, might be better described not so much as reflecting the globalization of technology as representing *the technologies of globalization*, since they service the increasingly global operation of cultural, social and economic life.

The concept of the globalization of technology is rather different in that it seeks to describe and explain how the process of economic and social globalization is not only affected by, but also itself affects, the production, distribution and transfer of technology. (See Bartlett and Ghoshal (1990), Dunning (1992), Granstrand *et al.* (1992), Howells and Wood (1993).) The strategies developed by both government and business institutions to generate technology are no longer based on a single country. Firms have to compete with a larger number of international rivals, and this often compels them to upgrade their products and processes. Inward and outward technology spillovers have also increased as a consequence of the enlarged market dimension.

The actors

The descriptions provided above indicate that these concepts of 'techno-nationalism' and 'techno-globalism' are of relevance for both public and business institutions, but also that these differing institutions will relate in their own ways to the processes under discussion. Public institutions typically operate at the scale of their own territorial state, yet are influenced heavily by the process of globalization, since the activities which take place within their own territory have effects beyond their borders and may in turn be challenged by decisions taken in other states.

National institutions at times compete to achieve leadership in science and technology (S&T), as was the case in the mid-1980s with the US Strategic Defence Initiative and the European Eureka programme (see Pianta,). In other cases, governments opt for cooperative strategies, as indicated by the large number of intergovernmental organizations in charge of specific international regimes. International property rights, international scientific exchanges, joint R&D programmes funded by international organizations such as the European Commission and so on – all illustrate S&T governmental policies that are no longer simply national in scope.

The international orientation of firms is of, course, nothing new. One of the obvious ways for firms to grow has long been to export to overseas markets. In the postwar period, however, a more demanding form of internationalization has gained importance, namely foreign direct investment (FDI), which implies the deployment of permanent facilities in host countries, and which in turn obliges firms to become familiar with more than one national institutional system. Business companies have also developed other, more sophisticated, forms of cross-border operation, such as joint ventures, non-equity collaborations and so on. The extent to which firms are still 'loyal' to their own home country is a matter of debate. Some argue that multinational corporations have lost their national identity and pursue only their global strategies, while others point out that the competitive advantage of large companies are still linked to their home country. (For a review of the different positions, see Porter (1990), Reich (1991), Dunning (1993), Chesnais (1994).)

While governments cannot be seen exclusively as national agents, neither can firms be considered stateless. And in spite of the increasing similarities of public and business actors as players in both domestic and foreign space, some basic differences persist: public institutions are, by and large, supposed to be accountable to their nation-based citizens, while business firms are allowed to be, and to some extent may be, accountable to 'stateless' shareholders. This creates at various levels a complex web of interactions between interfirm rivalry, on the one hand, and relations between nation states on the other. In order to expand their activities overseas, firms often seek the protection of foreign governments, although this in turn might jeopardize the relationship the firm has with its own home government – and such a process may also lead to a clash between the governments concerned. On the other hand, governments have to consider the pros and cons associated with inward investment into their own country: foreign direct investment might upgrade their productive capacity but might also increase their dependency on foreign capital.

These issues are explored in a growing literature on international political economy and international relations (Strange, 1988; Porter, 1990; Stopford and Strange, 1991), the implication of which is generally that governments and firms should select the capabilities to be developed in the home country and those to be acquired in the international markets when they deal with a strategic asset such as technological capability.

The origin of the concept of national systems: Friedrich List

Is there a place in economics for the study of how nation-specific factors affect the structure of production, consumption and growth? Consider the Table of Contents from Adam Smith's *The Wealth of Nations* (1976) and Paul A. Samuelson's *Economics* (1976): we find 'the division of labour', 'the commodity', 'wages', 'profit', 'the laws of supply and demand', 'the supply of money' and so on. This reflects the way that economics has developed as an analytical rather than as a historical discipline. History has been allowed to enter only when extraordinary events such as the great crash of 1929 or the post-Second World War recovery needed to be interpreted.

In 1841, Friedrich List published his book, *The National System of Political Economy*, which even from the Table of Contents looked substantially different from the main Anglo-Saxon textbooks of his age. The first part was devoted to a discussion of the history of various peoples: the Italians, the Hanseatic league, the Flemish and the Dutch, the English, the Spanish and the Portuguese, the French, the Germans, the Russians, and the North-Americans. Economic theory proper was discussed *after* history, in the second part of the treatise. It is no coincidence that List was German. At the beginning of the nineteenth century, German cultural life was dominated by the philosophy of history, which had as its main concern the explanation and prediction of the rise and fall of nations.

Influenced by the rise of American society, where List lived for several years, he tried to provide an *economic* explanation for the changing positions of nations in history. He was convinced that economic life played a crucial role in it, and therefore was highly critical of those German philosophers who ignored the material aspects of civilization. However, he also insisted that economic growth depended heavily on the social and cultural resources accumulated by a nation. Friedrich List can therefore be considered both a latecomer of the

German philosophy of history and a forerunner of the German historical school in economics.

Today, economists remember List as a fierce adversary of the theory of free trade as advocated by Adam Smith and his followers. It is certainly true that he was one of the few explicit supporters of trade protection – a doctrine that has been the focus of bitter criticism from economists, although less so from policy-makers and others. But in List's native town of Reutlingen, he is remembered as the pioneer of railways: he spent a large part of his life urging the princes who ruled 'the Germany of the one hundred homelands' to develop transportation. He understood that infrastructure, which in his day meant, above all, the railways, were a fundamental component of any strategy for economic growth, since they allowed commodities, individuals and information to circulate.

To get a balanced view of List's ideas it is perhaps necessary to combine the reminiscences of economists with those of the inhabitants of Reutlingen. List was not in favour of protection for its own sake; rather, he understood that economic growth required the creation of endogenous capabilities based on what he called 'intellectual capital' and learning.

List's main concern could be formulated in a simple basic question: which strategies should a backward nation adopt to catch up with leading countries? The free circulation of commodities was hardly the right answer. The law of comparative advantage predicts that both the leader and the follower would gain from trade. List argued, however, that in the long run the former would be likely to have preserved its advantage and the latter its underdevelopment. From a dynamic perspective, free trade was most likely to preserve and expand inequality among nations.[4]

Relatively underdeveloped countries should accept free trade policies only if the knowledge and expertise relating to the traded goods were traded equally freely. But this, of course, was not the practice followed by the then technological leader, the British empire. In spite of the free trade ideology espoused by the major English economists, the British government was keen to preserve its own technological leadership by hampering any transfer of knowledge to competing countries. Similarly, the trade of strategic machinery to other countries was heavily controlled by government policies (see Landes, 1969; Bruland, 1998). A large part of List's life was devoted to the denunciation of this covert but tenacious British protectionism.

But List was also aware that the problems involved in the circulation and assimilation of know-how go beyond the attempts of the technological leaders to defend that lead. He also pointed out the objective asymmetry that the transfer and assimilation of knowledge is much more difficult and complex than is the trading of commodities. Even if the leading nations were prepared to share their expertise with catching-up countries, the latter would still have to devote substantial amounts of energy in an attempt to assimilate it, including the development of their own endogenous scientific and technological capabilities.

List also understood that the development of endogenous capabilities had to be considered within the context of what was already, in his day, being seen as the growing globalization of economic activities. This offered an opportunity for latecomer nations to acquire best-practice techniques, although there was no guarantee that all nations would benefit to the same extent. On the question of how a latecomer could attempt to upgrade in the context of an increasingly global economy, he suggested various policy options: investing in education to promote an adequately trained workforce; creating a network of infrastructures to allow the dissemination of the most important economic resource, know-how; and creating economic ties among countries, such as customs unions. To strengthen their effectiveness, he also advocated the development of institutional systems of states; and then finally and, in fact, least, protecting infant industries to allow them to develop the expertise needed to face international competition.

National systems now

More than a century and a half after List, the concept of national systems of innovation (NSI) is once again on the academic and policy-making agenda. The country case studies published in Nelson (1993) and the thematic issues discussed in Lundvall (1992) are reminiscent of, respectively, parts one and two of List's main work. Quite rightly, Freeman (1995) starts his own historical journey on the nature of NSI from List's insights. Taken together, the resulting body of literature on NSI identifies the following crucial aspects in defining the structure and explaining the behaviour of nations.

Education and training

Education and training are vital components of economic development. In spite of the international diffusion of education and of the increasing, though still limited, number of students enrolled in foreign universities, education is still largely national in scope. Substantial differences can be found between countries in the proportion of relevant age groups participating in education, whether in primary, secondary or higher. Moreover, the distribution of students by disciplines also varies markedly across countries, as shown with reference to the East Asian countries by Mowery and Oxley (1995).

Science and technology capabilities

The level of resources devoted by each nation to formal R&D and other innovation-related activities (such as design, engineering, tooling-up, and so on) represents a basic characteristic of NSI. The bulk of the world's R&D activities is carried out in industrially advanced countries, while developing countries report a very small fraction of world R&D activities. Even within the relatively homogeneous group of OECD nations, there are significant differences in R&D intensity: a small club of countries, including the USA, Japan, Germany, Switzerland and Sweden, devote around 3 per cent of their gross domestic product (GDP) to formal R&D activities. Other countries report a much smaller R&D intensity, though they might be less disadvantaged in terms of other innovative inputs. Another difference relates to how R&D expenditure splits between the public and the business sectors; big government programmes in space, defence and nuclear technologies often shape the entire structure of the science and technology (S&T) system of a nation.

Industrial structure

Firms are the principle agents of technological innovation. The industrial structure of a nation heavily conditions the nature of its innovative activities. Large firms are more likely to undertake basic research programmes and are also more likely to be able to afford long-term investment in innovative activities whose payback may not only be spread years into the future but may also be extremely uncertain. The level of competition faced by companies in their domestic market also plays a crucial role in the R&D investment outcome.

S&T strengths and weaknesses

Each country has its own strengths and weaknesses in different S&T fields. Some nations have specializations in leading-edge technologies, while others have strengths in areas that are likely to provide only diminishing returns in the future. Moreover, some countries tend to be highly specialized in a few niches of excellence, while others have their

S&T resources distributed more uniformly across all fields (Archibugi and Pianta,). There are several determinants of national S&T specialization, including the size of a country, R&D intensity, market structure, and the international division of labour. The resulting S&T specialization may influence a nation's future economic performance, since countries with technological strengths in rising areas are likely to benefit from increasing returns, which in turn will allow them to expand their technological and production capabilities.

Interactions within the innovation system

The propensity of the different institutions to co-ordinate their activities and to interact with other actors differ widely across countries. Governments do interact heavily with large domestic firms (the so-called 'national champions'). Fransman (1995) describes the working of the Japanese Ministry for International Trade and Industry (MITI), one of the most cited successful institutions for the promotion of innovation in industry. In other countries, small firms have been keen to share their expertise and co-operate in developing a common competitive strategy, as demonstrated by the Italian industrial district. Such interactions are often able to multiply the effects of innovation undertaken at the country level, and increase its diffusion. Its absence can hamper substantially the economic effectiveness of resources devoted to S&T.

Absorption from abroad

The operation of these various aspects of national systems of innovation need to be considered within the context of increasing international integration. In the postwar period, several countries have benefited from an international regime which has deliberately encouraged the international transmission of knowledge (Nelson and Wright, 1992). Some countries, especially in the Third World, have benefited from bilateral technology transfer. A general lesson drawn from recent research, however, confirms List's original insight that no technology transfer can be effective without an endogenous effort to acquire that knowledge. See Bell and Pavitt (1997) and Mowery and Oxley (1995).

The list sketched above is far from complete. Several other aspects would need to be added to provide a complete description of a national system. But the factors singled out above do indicate that the explanatory power of the NSI notion is of a *comparative* nature. The description of a specific national system is useful when it is compared with that of other countries. These comparisons can be either qualitative or quantitative. The qualitative approach was followed by, among others, Nelson (1993), Freeman (1995) and Porter (1990), others have measured cross-country differences using indicators such as the level of resources devoted to R&D, the relative importance of the public and the business sectors, the level of international integration, and the distribution of the innovations produced across sectors. See Archibugi and Pianta (1994), Amendola *et al.* (1992), Patel and Pavitt (1994). However, we are still far from having achieved a coherent conceptual and empirical framework with which to explain the diversity between different countries' success in innovating (for a preliminary attempt, see Smith, 1995).

Implications of the national systems of innovation literature

The growing literature discussed briefly above makes clear that nations differ in their methods used to promote innovation, as well as in the quantity and significance of the innovations that have resulted from this effort. What are the implications of this for understanding the process of technological change, and for public policy?

First, while some of the key characteristics of innovative systems can be transferred from one country to another, others cannot be transplanted so easily, especially in the short term. Freeman describes the way in which the decision by a few companies based in Germany and in the US to establish internal R&D laboratories diffused gradually across several nations. Yet more than a century later, the role of industrial R&D is very far from being uniform across countries (see Archibugi and Pianta,). Only in a few advanced countries is industrial R&D at the core of the innovation system. Thus the dissemination of basic institutional innovations (such as the development of a business R&D network, or state-promoted education, or the creation of major government-led technology-intensive programmes) often requires a substantial effort as well as considerable time to be replicated successfully in other countries. But not even time and effort can eliminate the continued existence of significant cross-country differences. The route which leads each nation to build its technological competence is highly path-dependent; this would not be surprising to philosophers of history, nor to technology historians. See David (1975), Arthur (1989) and Mokyr (1990).

Second, there is no single model that alone is able to deliver successful economic performance. Over the postwar period, Japan and

Germany achieved high growth rates resulting in part from their massive investment in industrial R&D and technology. But other countries, such as Italy, managed to achieve similar goals while expending much less effort on technology. There is more than one technological avenue leading to the wealth of nations. See Abramovitz (1989), Denison (1967), Mowery and Rosenberg (1989), Maddison (1991), Fagerberg (1994).

Third, nations which fail to exploit innovation can find themselves in an underdevelopment trap. In this context, Freeman (1995) discusses why it was that the Soviet Union and East European countries, in spite of their very high investment in R&D, failed to sustain their economic development. He also compares Latin America to the countries of East Asia, pointing to a number of factors behind the industrial development of the East Asian economies that have been lacking in Latin America.

Fourth, historically, a country's innovation system has often played an important part in securing and consolidating competitive advantage and can become the driving force for economic hegemony. The change in the twentieth century from British to American economic and political leadership was associated in part with the American capability in pioneering the systematic exploitation of knowledge in the productive system The growth of East Asian countries has also been associated with their catching-up in a number of important technologies and to their acquired leadership in sectors of growing importance. The more innovative economies have also tended to be quick to adapt and imitate innovations produced elsewhere.

These implications drawn from the concept of national systems of innovation are, of course, based on historical experience. Is there any reason to believe that similar patterns will continue in the future? There are two interrelated factors that might be thought to lower the importance of nation-specific factors in the future. The first is the existence of strong technology systems which tend to be similar across countries in spite of national differences. The second is the dissemination and transfer of know-how across borders which, in principle, would allow all nations to benefit from best-practice methods and techniques.

Technology systems versus national systems?

Rosenberg (1976), Nelson and Winter (1977), Dosi (1982) and Freeman *et al.* (1982) all suggest that significant technological change is brought about generally as a result of specific regimes designed to serve specific purposes. A large number of technology and industry case studies have confirmed this.[5]

From a historical perspective, it is possible to identify technology systems which, even in the same periods, worked separately and independently. A thousand years ago, basic agricultural techniques in China were quite different from those in Europe, which in turn were different again from those in the Middle East. According to Gille (1978) this was because of the lack of circulation of information as well as to institutional rigidity. This is a far cry from the modern world system which has grown on the basis of the generation, circulation and diffusion of production techniques. The technical features of the majority of artefacts are similar across countries.

The similarities across technology systems are much broader than the narrow engineering characteristics of products (see Edquist, 1992). Technology systems are also defined by industrial concentration, barriers to entry, industrial R&D intensity, and the methods used to secure returns from innovation. Malerba and Orsenigo (1995) show that the characteristics of technological areas in terms of concentration, industrial turbulence and innovative dynamism across the four main European countries are rather similar; thus, in spite of the institutional differences of Germany, France, the UK and Italy,[6] some technology specific elements tend to be surprisingly similar.

Does this consideration reduce the significance of nation-specific factors? According to Nelson, 'if one focuses narrowly on what we have defined as *innovation systems* these tend to be sectorally specific. But if one broadens the focus the factors that make for commonality within a country come strongly into view, and these largely define the factors that make for commonality across sectors within a country'. This view is confirmed by Costello (1993), who compared the productivity growth of five major industries in six countries. Her results demonstrated stronger correlations across industries within a country than across countries within the same industry. Thus, rather than seeing the concepts of technology systems on the one hand and national systems of innovation on the other as being alternatives, only one of which at most can be applicable, it appears to be the case, rather, that both technology-specific factors shape the innovative process. The organization of industry tends to be technology-specific, while the impact of innovation is influenced heavily by the overall national economic environment. The challenge for both theory and policy is to establish these interrelations, and if possible to intervene to create positive feedbacks within this interrelationship.

What differentiates countries is not their methods of production in certain industries, but rather their relative strengths and weaknesses across industries. For example, the US innovation system is defined by strong government intervention in defence-related areas, and this is reflected in its sectoral strengths in aircraft and nuclear technology. Japan, on the other hand, has negligible industrial activity in the aircraft sector. In spite of these differences, the industrial and technological features of the aircraft sector tend to be the same in both the USA and Japan. However, it would be wrong to predict the sectoral specialization of a nation on the grounds of institutional features alone: Italy, a country with medium R&D intensity and low industrial concentration is very active in automobiles, one of the industries generally associated with both high R&D and industrial concentration.

Is the globalization of technology making the nation-state redundant?

The second factor which might be thought to diminish the importance of nation-specific factors is the increasing globalization of technological and other industrial and economic processes. Several writers have stressed that there has been a dramatic increase in the process of economic globalization. International trade and capital flows, foreign direct investment, migration – all have increased substantially since the 1970s. See Holland (1987), Dunning (1992, 1993) and Chesnais (1994). A corresponding globalization is said to have occurred in social, cultural and political life, having an impact on local communities, including nation-states, and lowering ties of national identity, citizenship and political sovereignty (Held, 1991; Robertson, 1992). On the other hand, globalization is certainly not a new phenomenon (Wallerstein, 1974).[7]

We would make a distinction (on which, see Archibugi and Michie, 1995) between three separate processes which are often subsumed within the catch-all general term 'technological globalization': (i) international exploitation of national technological capabilities: firms try to exploit their innovations on global markets either by exporting products which embody them, or by licensing the know-how; (ii) collaboration across borders among both public and business institutions to exchange and develop know-how. Firms are expanding their non-equity agreements to share the costs and risks of industrial research (see Hagedoorn and Schakenraad, 1993). Metcalfe (1995) points out that the scientific community has always been international in scope, although public research centres and academia have recently increased their proportion of cross-borders linkages substantially;[8] and (iii) the generation of innovation across more than one country, which refers particularly to the activities of multinational corporations, as discussed by Patel (1995).

On the first two of these dimensions to the globalization of technology, it is hardly controversial that they have increased in importance. Trade and patent flows, international technical agreements and scientific co-authorships have all shown a dramatic increase since about the 1970s. But it is intellectually sloppy to assume that this implies that nation-states have become less important in some way without specifying the mechanisms by which this latter conclusion follows. If, for example, increased globalization means that any loss of relative competitiveness translates into a far greater loss of markets – abroad and at home – with a concomitantly greater loss of jobs and threat to living standards than would have been the case in the days when the world economy was less 'global', then this would imply that the benefits from national action to enhance competitiveness would be that much greater. And, conversely, any inaction would risk far greater losses.

Certainly, in this case, while globalization may result in national action having greater payoffs – and national inaction greater costs – it could still be the case that while globalization makes national action more rather than less important, at the same time it makes it more difficult, or less feasible. But again, it is important not to jump to fashionable and easy conclusions unthinkingly. If national action has become more important yet more difficult, then this increased difficulty may itself call for more serious and far-reaching intervention from national governments to overcome such difficulties.

So, while for the first two of our globalization categories listed above the key controversy is about how to respond to trends that are reasonably well established (albeit exaggerated by some), on the third category of the extent to which multinational corporations have increased their technological operations in host countries, the evidence itself is less well established. Patel (1995), taking into account the patented inventions of more than five hundred of the world's largest enterprises, shows that the vast majority of inventions are developed in the firms' home nation. According to Patel, multinational corporations – the companies which by definition are globally-orientated – tend to be loyal to their own country when they have to locate a strategic asset such as technology. However, the results presented by Patel appear to be at odds with those of Cantwell (1994, 1995).[9] From a historical perspective, Cantwell shows that the share of innovations generated by firms in host countries has increased considerably.

Patented inventions, however, capture only the most formalized part of technological knowledge. Multinational corporations might be keener to decentralize forms of knowledge that do not belong to the core of their business strategy. Companies might be more willing to locate facilities abroad which are less critical to their strategy, such as software, engineering, design and so on. Less developed countries offer an adequately trained workforce, but at salaries that are much lower than in the developed countries, while information technologies make the geographical location of high-tech jobs less relevant. This justifies the widespread concern that industrial countries could lose skill-intensive jobs to the benefit of the South.[10]

On what might induce companies to centralize or decentralize their technological activities, Howells and Wood (1993) suggest that the advantages of centralization include: the benefits of economies of scale and scope which are associated with larger R&D operations; the minimum efficient size associated with indivisibilities of certain scientific instruments and facilities; the increased security over in-house research, which among other things reduces the risk of competitors copying or leap-frogging in key research fields; and the ability to create a well-established dense local innovation network with higher-education institutes, contract research companies and other support agencies. The main advantages seen by Howells and Wood as being associated with decentralization are: a more effective and applicable R&D effort focused on the real needs of the business and operational units; improved communications or coupling between R&D and other key corporate functions; fewer problems in 'programme dislocation' when a project is transferred from R&D to production; and better responsiveness to various local market needs. To this list might be added: to keep a window open on the technological developments of other countries; and to take advantage of the fields of excellence of the host country.

An extensive survey of companies' headquarters and host facilities has identified the type of work undertaken in overseas R&D laboratories (Pearce and Singh, 1992). The activities carried out most frequently in host countries are to derive new production technology and to adapt existing products to the local markets to make them accepted by local communities. Even the taste of Coca-Cola, the most typical standardized product of the global economy, is not quite the same in the USA, Japan and Italy (see Ohmae, 1990).

Multinational corporations apply a variety of strategies to capitalize on their technological advantages. Bartlett and Ghoshal (1990) have provided a useful categorization of three different, although not mutually exclusive, strategic approaches:

(i) *Centre-for-global* – the traditional 'octopus' view of the multinational corporation: a single brain located within the company's headquarters concentrates the strategic resources – top management, planning and technological expertise. The brain distributes impulses to the tentacles (that is, the subsidiaries) scattered across host countries. Even when some overseas R&D is reported, this is basically concerned with adapting products to local users' needs;

(ii) *Local-for local* – where each subsidiary of the firm develops its own technological know-how to serve local needs. The interactions among subsidiaries is, at least from the viewpoint of developing technological innovations, rather low, but subsidiaries are integrated into the local fabric. This may occur with conglomerate firms or companies that are not characterized by strong global products.

(iii) *Local-for-global* – the case of multinational corporations which, rather than concentrating their technological activities in a home country, distribute R&D and technological expertise across a variety of host countries. This allows the company to develop each part of the innovative process in the most suitable environment: semiconductors in Silicon Valley, automobile components in Turin, software in India, for example. The effectiveness of such a strategy relies on the intensity of intra-firm information flows.

Techno-nationalism versus techno-globalism?

Much of the debate about techno-nationalism and techno-globalism has direct policy implications, addressed explicitly by Fransman (1995) and Metcalfe (1995). What is the point of government policies to promote innovation in industry if the benefits can be transferred to other countries? Is there any guarantee that firms will use these benefits to the advantage of the nation which provides support? For example, Reich (1991) argues that it is not in the interests of a nation to support national champions. He advocates instead policies to foster the infrastructure of a nation:

> Rather than increase the profitability of corporations flying its flag, or enlarge the worldwide holdings of its citizens, a nation's economic role is to improve its citizens' standard of living by enhancing the value of what they contribute to the world economy. The concern over national 'competitiveness' is often misplaced. It is not what we own that counts; it is what we do.

In the USA in particular, there has been widespread concern that government policies could be benefiting foreign firms just as much as domestic ones. For example, much of the US government-funded defence and space R&D in semiconductors was exploited by Japanese companies to develop high-tech competitive products (see the debate in Lee and Reid (1991), Caldwell Harris and Moore (1992), Tyson (1992), Scherer (1992), and Nelson and Wright (1992)). The USA and other industrial countries have therefore called for a more tightly regulated international regime of intellectual and industrial property rights. In other words, the focus has shifted from the generation of technology to devices to guarantee sufficient returns from it on the international market.[11]

This has implications for industrial and technology policy. Metcalfe (1995) differentiates between two broad categories of government action, namely direct financial incentives to companies for their innovative programmes, and public supply of infrastructure to make a country attractive for the deployment of S&T activities. Globalization may be thought to have reduced the usefulness of the first kind of government policies, especially when the benefits are received by companies with subsidiaries in several countries. But policies of the second kind, which include education, effective industry–university partnerships, communications and so on, have certainly increased in importance (see also Tassey, 1991). In the global economy, nations have to upgrade their infrastructure to attract technology-intensive activities. Fransman (1995), after describing the activities of the Japanese MITI, asks: how could MITI have so much power with such a small amount of financial resources? The question itself indicates that policies aimed at creating an innovative and industrially dynamic environment can be much more important than simply handing cash to companies.

An essential factor in the postwar 'golden age of capitalism' was the existence of an international regime favourable to the diffusion of S&T (see Nelson and Wright, 1992). But at the time of writing, any such regime appears to be under constant threat from the operation of large corporations. See Holland (1987), Barnet and Cavanagh (1994), Michie and Grieve Smith (1995). From this perspective, the real opposition to *techno-nationalism* is not, as is so often suggested *techno-globalism*, but rather *techno-liberalism*. It is therefore no surprise that the literature on national systems – Porter (1990), Lundvall (1992), Nelson (1993), Fransman (1995), Freeman (1995), Metcalfe (1995), Mowery and Oxely (1995) – generally advocates a stronger role of government to foster innovation.

Notes

1. A review of the recent economic literature on technological change can be found in Dosi (1982, 1984) and Freeman (1987, 1994).
2. See, for example, the chapters by Freeman, Lundvall and Nelson in Dosi *et al.* (1988).
3. For an attempt to highlight these differences, see McKelvey (1991) and Humbert (1994).
4. See Kitson and Michie (1995) for a discussion of the political economy of trade and trade policy, where this distinction is made, between mainstream theory on the one hand, where it is asserted that all will benefit, with those lagging dragged along, and, on the other, the more likely scenario where those stuck in a vicious cycle of decline may well see their disadvantage intensified.
5. See, for example, the studies on semiconductors – (Dosi (1982, 1984), Malerba (1985)) and biotechnology (Orsenigo (1989)).
6. Highlighted by, for example, the country case-studies of Keck, Chesnais, Walker and Malerba in Nelson (1993) and by Cohendet *et al.* (1992).
7. And for a sceptical view of the above claims regarding globalisation see Michie (1995).
8. On which, see also Malerba *et al.* (1991).
9. See also Casson (1991) and Cantwell (1994, 1995).
10. See *Business Week* (December, 1994) on the argument that jobs in the North are being lost to the South, see Wood (1994) and for a more sceptical view see Eatwell (1995) and Singh and Zammit (1995); for a related argument to Wood's, see Galbraith (1996) commented on by Michie.
11. David and Foray (1995) argue that, given the nature of contemporary technological expertise, the institutional international context should favour the diffusion and imitation of innovation rather than the protection of intellectual property rights.

References

Abramovitz, M. (1989) *Thinking about Growth* (Cambridge University Press).

Amendola G., Guerrieri, P. and Padoan, P. C. (1992) 'International Patterns of Technological Accumulation and Trade', *Journal of International and Comparative Economics*, vol. 1, pp. 173–97.

Archibugi, D., Cesaratto, S. and Sirilli, G. (1991) 'Sources of Innovation and Industrial Organization in Italy', *Research Policy*, vol. 20, pp. 299–313.

Archibugi, D. and Michie, J. (1995) 'The Globalisation of technology: a new taxonomy', *Cambridge Journal of Economics*, vol. 19, no. 1 February, pp. 121–40.

Archibugi, D. and Michie, J. (1997) *Technology Globalisation and Economic Performance*, (Cambridge University Press).

Archibugi, D. and Pianta, M. (1992) *The Technological Specialization of Advanced Countries. A Report to the EC on Science and Technology Activities* (Boston: Kluwer).

Archibugi, D. and Pianta, M. (1994) 'Aggregate Convergence and Sectoral Specialisation in Innovation', *Journal of Evolutionary Economics*, vol. 4, no. 4, pp. 17–33.

Arthur, B. (1989) 'Competing Technologies, Increasing Returns, and lock-in by historical events', *Economic Journal*, vol. 99, March, pp. 116–31.

Barnet, R. J. and Cavanagh, J. (1994) *Global Dreams. Imperial Corporations and the New World Order* (New York: Simon & Schuster).

Bartlett, C. A. and Ghoshal, S. (1990) 'Managing Innovation in Transnational Corporations', in C. A. Bartlett, Y. Doz and G. Hedlung (eds), *Managing the Global Firm* (London: Routledge).

Bell, M. and Pavitt, K. (1997) 'Accumulation of Technological Capability in Developing Countries', in D. Archibugi and J. Michie (eds), *Technology, Globalisation and Economic Performance* (Cambridge University Press).

Bruland, K. (1998) *British Technology and European Industrialization* (Cambridge University Press).

Business Week (1994) 'High-tech Jobs all over the Map', December 19, pp. 42–7.

Caldwell Harris, M. and Moore, G. E. (eds) (1992) *Linking Trade and Technology Policies* (Washington, DC: National Academy Press).

Cantwell, J. A. (ed.) (1994) *Transnational Corporations and Innovator Activities* (London: Routledge).

Cantwell, J. A. (1995) 'The Globalisation of Technology: What Remains of the Product Cycle Model?', *Cambridge Journal of Economics*, vol. 19, no. 1, February, pp. 155–74.

Carter, A. (1989) 'Know-how Trading as Economic Exchange', *Research Policy*, vol. 18, no. 3.

Casson, M. (ed.) (1991) *Global Research Strategy and International Competitiveness* (Oxford: Basil Blackwell).

Chesnais, F. (1994) *La Mondialisation du Capital* (Paris: Syros).

Cohendet, P., Llerena, P. and Sorge, A. (1992) 'Technological Diversity and Coherence in Europe: an Analytical Overview', *Revue d'Economie Industrielle*, no. 59, vol. 1, pp. 9–26.

Costello, D. (1993) 'A Cross-country, Cross-industry Comparison of Productivity Growth', *Journal of Political Economy*, vol. 101, no. 2, April, pp. 207–22.

Dasgupta, P. and David, P. (1994) 'Toward a New Economics of Science', *Research Policy*, September, vol. 23, no. 5, pp. 487–521.

David, P. (1975) *Innovation, Technical Choice, and Growth* (Cambridge University Press).

David, P. and Foray, D. (1995) 'Accessing and Expanding the Science and Technology Knowledge Base', *Science Technology Industry Review*, no. 16, pp. 13–68.

Denison, E. F. (1967) *Why Growth Rates Differ* (Washington, DC: Brookings Institution).

Dosi, G. (1982) 'Technological Paradigms and Technological Trajectories', *Research Policy*, vol. 11, pp. 147–61.

Dosi, G. (1984) *Technical Change and Industrial Transformation. The Theory and Application to the Semiconductor Industry* (London: Macmillan).
Dosi, G. (1988) 'Source, Procedures and Microeconomic Effects of Innovation', *Journal of Economic Literature*, vol. 36, pp. 1126–71.
Dosi, G., Freeman, C., Nelson, D., Silverberg, G. and Soete, L. (eds) (1988) *Technical Change and Economic Theory* (London: Pinter).
Dunning, J. (1992) *The Globalisation of Business* (London: Routledge).
Dunning, J. (1993) *Multinational Enterprises and the Global Economy* (Workingham: Addison-Wesley).
Edquist, C. (1992) *Systems of Innovation. Conceptual and Theoretical Remarks*, paper presented at the Workshop Systems of Innovation, Bologna, October.
Eatwell, J. (1995) 'The International Origins of Unemployment', in J. Michie and J. Grieve Smith (eds), *Managing the Global Economy* (Oxford University Press).
Evangelista, R. (1995) 'Embodied and Disembodied Innovative Activities: Evidence from the Italian Innovation Survey', in OECD (1995).
Fagerberg, J. (1994) 'Technology and International Differences in Growth Rates', *Journal of Economic Literature*, vol. XXXII, September, pp. 1147–75.
Foray, D. and Freeman, C. (1993) *Technology and the Wealth of Nations. The Dynamics of Constructed Advantage* (London: Pinter).
Fransman, M. (1995) 'Is National Technology Policy Obsolete in a Globalised World? The Japanese Response', *Cambridge Journal of Economics*, vol. 19, no. 1, February, pp. 95–120.
Freeman, C. (1987) *Technology Policy and Economic Performance* (London: Pinter).
Freeman, C. (1994) 'The Economics of Technical Change', *Cambridge Journal of Economics*, vol. 18, no. 5, pp. 463–514.
Freeman, C. (1995) 'The National System of Innovation in Historical Perspective', *Cambridge Journal of Economics*, vol. 19, no. 1, February, pp. 5–24.
Freeman, C., Clark, J. A. and Soete, L. (1982) *Unemployment and Technological Innovation* (London: Frances Pinter).
Freeman, C. and Soete, L. (eds) (1990) *New Explorations in the Economics of Technical Change* (London: Pinter).
Freeman, C. and Soete, L. (1996) 'Conclusions', in D. Foray and C. Freeman (eds), *Technology and the Wealth of Nations. The Dynamics of Constructed Advantage*, (London: Pinter).
Galbraith, J. (1996) 'Uneven Development and the Destabilization of the North: a Keynesian View', *International Review of Applied Economics*, vol. 10, no. 1, January.
Giddens, A. (1991) *Consequences of Modernity* (Cambridge: Polity Press).
Gille, B. (1978) *Histoire des Techniques* (Paris: Editions Gallimard).
Granstarnd, O., Hakanson, L. and Sjölander, S. (eds) (1992) *Technology Management and International Business. Internalization of R&D and Technology* (Chichester: Wiley).
Hagedoorn, J. and Schakenraad, J. (1990) 'Inter-firm Partnership and Cooperative Strategies in Core Technologies', in Freeman C. and Soete L. (eds), *New Explorations in the Economics of Technical Change*, (London: Pinter).
Hagedoorn, J. and Schakenraad, J. (1993) 'Strategic Technology Partnering and International Corporate Strategies', in K. Hughes (ed.), *European Competitiveness* (Cambridge University Press).

Held, D. (1991) ‘Democracy, the Nation-State and the Global System’, *Economy and Society*, vol. 20, no. 2, May, pp. 138–72.

Hodgson, G. and Screpanti, F. (eds) (1991) *Rethinking Economics: Markets, Technology and Economic Evolution* (Aldershot: Edward Elgar).

Holland, S. (1987) *The Global Economy* (London: Weidenfeld & Nicolson).

Howells, J. and Wood, M. (1993) *The Globalisation of Production and Technology* (London: Belhaven Press).

Hughes, K. (ed.) (1993) *European Competitiveness* (Cambridge University Press).

Humbert, M. (1994) *The Globalisation of Technology as a Challenge for a National Innovation System*, EAEPE Conference, Copenhagen, 28–30 September.

Kitson, M. and Michie, J. (1995) ‘Conflict and Cooperation: the Political Economy of Trade and Trade Policy’, *Review of International Political Economy*, vol. 2, no. 4, pp. 632–57.

Landes, D. (1969) *The Unbound Prometheus. Technological and Industrial Development in Western Europe from 1750 to the Present* (Cambridge University Press).

Lee, T. H. and Reid, P. P. (eds), *National Interests in the Age of Global Technology*, (Washington, DC: National Academy Press).

List, F. (1885) [1841] *The National System of Political Economy*, English Translation (London: Longmans & Co).

Lundvall, B. Å, (ed.) (1992) *National Systems of Innovation* (London: Pinter).

Lundvall, B. Å. and Johnson, B. (1994) ‘The Learning Economy’, *Journal of Industry Studies*, vol. 1, no. 2, pp. 23–42.

Maddison, A. (1991) *Dynamic Forces in Capitalist Development. A Long-Run Comparative View* (Oxford University Press).

Malerba, F. (1985) *The Semiconductor Business. The Economics of Rapid Growth and Decline* (London: Frances Pinter).

Malerba, F., Morawets, A. and Pasqui, G. (1991) *The Nascent Globalisation of Universities and Public and Quasi-Public Research Organizations* (Brussels: European Commission, FAST).

Malerba, F. and Orsenigo, L. (1995) ‘Schumpeterian Patterns of Innovation’, *Cambridge Journal of Economics*, vol. 19, no. 1, February, pp. 47–66.

McKelvey, M. (1991) ‘How do National Systems of Innovation Differ? A Critical Analysis of Porter, Freeman, Lundvall and Nelson’, in Hodgson G. and Screpanti E. (eds), *Rethinking Economics: Markets, Technology and Economic Evolution* (Aldershot: Edward Elgar).

Metcalfe, J. S. (1995) ‘Technology Systems and Technology Policy in an Evolutionary Framework’, *Cambridge Journal of Economics*, vol. 19, no. 1, February, pp. 25–46.

Michie, J. (1995) ‘Introduction’ to J. Michie and J. Grieve Smith (eds), *Managing the Global Economy*, Oxford University Press.

Michie, J. (1996) ‘Creative Destruction or Regressive Stagnation?’, *International Review of Applied Economics*, vol. 10, no. 1, pp. 121–6.

Michie, J. and Grieve Smith, J. (eds) (1995) *Managing the Global Economy* (Oxford University Press).

Mokyr, J. (1990) *The Lever of Riches. Technological Creativity and Economic Progress*, (Oxford University Press).

Mowery, D. C. and Oxley, J. E. (1995) ‘Inward Technology Transfer and Competitiveness: the Role of National Innovation Systems’, *Cambridge Journal of Economics*, vol. 19, no. 1, February, pp. 67–94.

Mowery, D. C. and Rosenberg, N. (1989) *Technology and the Pursuit of Economic Growth* (Cambridge University Press).
Nelson, R. (1987) *Understanding Technical Change as an Evolutionary Process* (Amsterdam: North-Holland).
Nelson, R. (ed.) (1993) *National Innovation Systems* (New York: Oxford University Press).
Nelson, R. and Rosenberg, N. (1993) 'Technical Innovation and National Systems', in Nelson R. (ed.) *National Innovation Systems* (New York: Oxford University Press).
Nelson, R. and Winter, S. (1977) 'In Search of a Useful Theory of Innovation', *Research Policy*, vol. 6, pp. 36–113.
Nelson, R. and Wright, G. (1992) 'The Rise and Fall of American Technological Leadership: the Postwar Era in Historical Perspective', *Journal of Economic Literature*, vol. 30, December.
OECD, (1995) *Innovation, Patents and Technological Strategies* (Paris: OECD).
Ohmae, K. (1990) *The Borderless World: Management Lessons in the New Logic of the Global Market Place* (London: Collins).
Orsenigo, L. (1989) *The Emergence of Bio-Technologies* (London: Pinter).
Pasinetti, L. (1993) *Structural Economic Dynamics. A Theory of the Economic Consequences of Human Learning* (Cambridge University Press).
Patel, P. (1995) 'Localised Production of Technology for Global Markets', *Cambridge Journal of Economics*, vol. 19, no. 1, February, pp. 141–54.
Patel, P. and Pavitt, K. (1991) 'Large Firms in the Production of the World's Technology: an Important Case of Non-globalisation', *Journal of International Business Studies*, vol. 22, no. 1.
Patel, P. and Pavitt, K. (1994) 'National Innovation Systems: Why They are Important and How They Might Be Measured and Compared?', *Economics of Innovation and New Technology*, vol. 3, pp. 77–95.
Pavitt, K. (1984) 'Sectoral Patterns of Technical Change. Towards a Taxonomy and a Theory', *Research Policy*, vol. 13, pp. 343–73.
Pearce, R. D. and Singh, S. (1992) *Globalizing Research and Development* (London: Macmillan).
Pianta, M. (1988) *New Technologies Across the Atlantic* (Hemel Hempstead: Harvester & Wheatsheaf).
Pianta, M. (1988) 'High Technology Programmes: for the Military of for the Economy?', *Ballerin of Peace Proposals*, no. 1.
Porter, M. (1990) *The Competitive Advantage of Nations* (London: Macmillan).
Reich, R. (1991) *The Work of Nations* (London: Simon & Schuster).
Robertson, R. (1992) *Globalization, Social Theory and Global Culture* (London: Sage).
Rosenberg, N. (1976) *Perspectives on Technology* (Cambridge University Press).
Rosenberg, N. (1982) *Inside the Black Box: Technology and Economics* (Cambridge University Press).
Rosenberg, N. (1994) *Exploring the Black Box* (Cambridge University Press).
Scherer, F. M. (1992) *International High-technology Competition* (London: Harvard University Press).
Schrader, S. (1991) 'Informal Technology Transfer Between Firms: Cooperation Through Information Trading', *Research Policy*, vol. 20, pp. 153–70.

Singh, A. and Zammit, A. (1995) 'Employment and Unemployment: North and South', in J. Michie and J. Grieve Smith (eds), *Managing the Global Economy* (Oxford University Press).

Smith, K. (1995) 'Interactions in Knowledge Systems: Foundations, Policy Implications and Empirical Methods', *Science Technology Industry Review*, no. 16, pp. 69–102.

Strange, S. (1988) *States and Markets* (London: Pinter).

Stephan, P. E. (Forthcoming) 'An Essay on the Economics of Science', *Journal of Economic Literature*.

Stopford, J. and Strange, S. (1991) *Rival States, Rival Firms. Competition for World Market Shares* (Cambridge University Press).

Tassey, G. (1991) 'The Functions of Technology Infrastructure in a Competitive Economy', *Research Policy*, vol. 20, pp. 345–61.

Tyson, L. D. (1992) *Who's Bashing Whom? Trade Conflicts in High-technology Industries* (Washington, DC: Institute for International Economics).

Von Hippel, E. (1987) 'Cooperation Between Rivals: Informal Knowhow Trading', *Research Policy*, vol. 16, pp. 291–302.

Von Hippel, E. (1988) *The Sources of Innovation* (Oxford University Press).

Wallerstein, I. (1974) *The Modern World System* (New York: Academic Press).

Wallerstein, I. (1979) *The Capitalist World Economy* (Cambridge University Press).

Wallerstein, I. (1980) *The Modern World System II: Mercantilism and the Consolidation of the European World Economy* (Cambridge University Press).

Wood, A. (1994) *North-South Trade, Employment and Inequality: Changing Fortunes in a Skill-Driven World* (Oxford University Press).

Part III

Education and Growth

9

Endogenizing Investment in Tangible Assets, Education and New Technology

Dale W. Jorgenson[1]

Introduction

The early 1970s marked the emergence of a rare professional consensus on economic growth, articulated in two strikingly dissimilar books. Simon Kuznets, the greatest of twentieth-century empirical economists, summarized his decades of research in *Economic Growth of Nations* (1971). The enormous impact of this research was recognized in the same year by the Royal Swedish Academy of Sciences in awarding the third Bank Prize in Economic Sciences in Memory of Alfred Nobel to Kuznets 'for his empirically founded interpretation of economic growth which has led to new and deepened insight into the economic and social structure and process of development'.[2]

Robert Solow's book *Growth Theory* (1989) modestly subtitled 'An Exposition', contained his 1969 Radcliffe Lectures at the University of Warwick. In these lectures, Solow also summarized decades of research, initiated by the theoretical work of Roy Harrod (1939) and Evsey Domar (1946) Solow's seminal role in this research, beginning with his brilliant and path-breaking essay of 1956, 'A Contribution to the Theory of Economic Growth', was recognized, simply and elegantly, by the Royal Swedish Academy of Sciences in awarding Solow the Nobel Prize in Economics in 1987 'for his contributions to the theory of economic growth'.[3]

At the start of the twenty-first century, the consensus on economic growth of the early 1970s has collapsed under the weight of a massive accumulation of new empirical evidence, followed by a torrent of novel theoretical insights. The purpose of this chapter is to initiate the search for a new empirical and theoretical consensus. Any attempt at this thoroughly daunting task may be premature, since professional interest in growth currently appears to be waxing rather than waning. Moreover, the disparity of views among economists, always looming remarkably large for a discipline that aspires to the status of a science, is greater on growth than most other topics.

The consensus of the early 1970s emerged from a similar period of fractious contention among competing schools of economic thought, and this alone is grounds for cautious optimism. However, I believe it is critically important to understand the strengths and weaknesses of the earlier consensus, and how it was broken up by subsequent theory and evidence. It is also essential to determine whether elements have survived that could serve as a useful point of departure in the search for a new consensus.

Let me first consider the indubitable strengths of the perspective on growth that emerged victorious over its numerous competitors in the early 1970s. Solow's neoclassical theory of economic growth, especially his analysis of steady states with constant rates of growth, provided conceptual clarity and sophistication. Kuznets generated persuasive empirical support by quantifying the long sweep of historical experience of the USA and thirteen other developed countries. He combined this with quantitative comparisons among a wide range of developed and developing economies during the postwar period.

With the benefit of hindsight, the most obvious deficiency of the neoclassical framework of Kuznets and Solow was the lack of a clear connection between the theoretical and empirical components. This lacuna can be seen most starkly in the total absence of cross-references between the key works of these two great economists, yet they were working on the same topic, within the same framework, at virtually the same time, and in the very same geographical location of Cambridge, Massachusetts.

Searching for analogies to describe this remarkable coincidence of views on growth, we can think of two celestial bodies on different orbits, momentarily coinciding from our earthbound perspective at a single point in the sky and glowing with dazzling but transitory luminosity. The indelible image of this extraordinary event has been burned into the collective memory of economists, even if the details have long been forgotten. The common perspective that emerged remains the guiding star for subsequent conceptual development and empirical observation.

In the second I consider challenges to the traditional framework of Kuznets and Solow arising from new techniques for measuring economic welfare and productivity. The elaboration of production theory, and the corresponding econometric techniques, led to the successful implementation of constant quality measures of capital and labour inputs and investment goods output. However, it was not until 11 July 1994 that these measures were incorporated into a new official productivity index for the USA by the Bureau of Labour Statistics.

The recent revival of interest in economic growth by the larger community of economists can be dated from Angus Maddison's (1991) updating of Kuznets' (1971) long-term comparisons of economic growth among industrialized countries. This was by the successful exploitation of the Penn World Table (Summers and Heston, 1988) – created by Irving Kravis, Alan Heston and Robert Summers – which provided comparisons among more than a hundred developed and developing countries. Exploiting the panel data structure of these comparisons, Nasrul Islam (1995) was able to show that the Solow model is the appropriate point of departure for modelling the endogenous accumulation of tangible assets.

The new developments in economic measurement and modelling summarized in the third section of this chapter have cleared the way for undertaking the difficult, if unglamorous, task of constructing quantitative models of growth suitable for the analysis of economic policies. Models based on the neoclassical framework of Kuznets and Solow determine growth by exogenous forces, principally spillovers from technological innovations. By contrast, models based on the new framework, described in the fourth section, determine the great preponderance of economic growth endogenously through investments in tangible assets and human capital.

Endogenous models of economic growth require concepts of an aggregate production function and a representative consumer that can be implemented econometrically. These concepts imply measurements of welfare and productivity that can best be organized by means of a system of national accounts. The accounts must include production, income and expenditure, capital formation and wealth accounts, as in the United Nations (1968) A *System of National Accounts*. Alternative economic policies can then be ranked by means of equivalent variations in wealth, thus providing the basis for policy recommendations.

In the fifth section I describe quantitative models suitable for the analysis of economic policies. Econometric techniques have provided the missing link between the theoretical and empirical components of the consensus of the early 1970s. The development of these techniques was a major achievement of the 1970s, but successful applications began to emerge only in the 1980s. These techniques were unavailable when Solow (1988) first articulated the objective of constructing econometric models of growth for the analysis of economic policies.

The growth of tangible assets is endogenous within a Solow neoclassical growth model. Kun-Young Yun and I constructed a complete econometric model for postwar US economic growth with this feature in two papers published in 1986. We have used this model to analyse the economic impact of fundamental tax reforms. Subsequently this model was extended in Jorgenson and Ho (1995) to incorporate endogenous growth in human capital; we have employed the extended model to analyse the impact of alternative educational policies.

Although endogenous investment in new technology has been a major theme in growth theory since the 1960s, empirical implementation has foundered on the issues first identified by Zvi Griliches (1973), of measuring the output of research and development activities. Until this issue has been resolved successfully, a completely endogenous theory of economic growth will remain a chimera, forever tantalizing the imagination, but far removed from the practical realm of economic policy. The sixth section assesses the prospects for endogenizing investment in new technology, and concludes the chapter.

Sources and uses of growth

The objective of modelling economic growth is to explain the sources and uses of economic growth endogenously. National income is the starting point for assessments of the uses of economic growth, through consumption and saving. The concept of a measure of economic welfare, is the key to augmenting national income to broaden the concepts of consumption and saving. Similarly, gross national product (GNP) is the starting point for attributing the sources of economic growth to investments in tangible assets and human capital, but could encompass investments in new technology as well.

The allocation of the sources of economic growth between investment and productivity is critical for assessing the explanatory power of growth theory. Only substitution between capital and labour inputs resulting from investment in tangible assets is endogenous in Solow's neoclassical model of economic growth. However, substitution among different types of labour inputs is the consequence of investment in human capital, while investment in tangible assets also produces substitution among different types of capital inputs. These were not included in Solow's model of production.

Productivity growth is labour-augmenting or equivalent to an increase in population in the simplest version of the neoclassical growth model. If productivity growth predominates greatly among the sources of economic growth, as indicated by Kuznets (1971) and Solow (1989), most of growth is determined exogenously. Reliance on the Solow residual as an explanatory factor is a powerful indictment of the limitations of the neoclassical framework. This viewpoint was expressed by Moses Abramovitz (1956) who famously characterized productivity growth as 'a measure of our ignorance'.

The appropriate theoretical framework for endogenous growth is the Ramsey model of optimal growth introduced by David Cass (1965) and Tjalling Koopmans (1965). A promising start on the empirical implementation of this model was made in my 1967 paper with Griliches (1967). It appeared that 85 per cent of US economic growth could be made endogenous; determinants of the remaining 15 per cent were left for further investigation, but might be attributable to investments in new technology.[4]

The conclusions of my paper with Griliches were corroborated in two studies Christensen and Jorgenson (1969, 1970). These studies provided a much more detailed implementation of the concept of a capital as a factor of production. We utilized a model of the tax structure for corporate capital income that I had developed in a series of papers with Rober Hall 1967, 1969, 1971. Christensen and I extended this model to non-corporate and household capital incomes in order to capture the impact of additional differences in returns to capital resulting from taxation on substitutions among capital inputs.

In Christensen and Jorgenson (1973) estimates of the sources of economic growth were incorporated into a complete system of US national accounts in our paper, 'Measuring Economic Performance in the Private Sector'.[5] Our main objective was the construction of internally consistent income, product and wealth accounts. Separate product and income accounts were integral parts of both the US Income and Product Accounts[6] and the United Nations (1984) *System of National Accounts* designed by Richard Stone.[7] However, neither system included wealth accounts consistent with the income and product accounts.

Christensen and I constructed income, product and wealth accounts, paralleling the US National Income and Product Accounts for the period 1929–69. We implemented our vintage accounting system for the US on an annual basis. The complete system of vintage accounts gave stocks of assets of each vintage and their prices. The stocks were cumulated to obtain asset quantities, providing the perpetual inventory of assets accumulated at different points in time or different vintages employed by Raymond Goldsmith (1955–6, 1962).

The key innovation in our vintage system of accounts was the use of asset pricing equations to link the prices used in evaluating capital stocks and the rental prices employed in our constant quality index of capital input. In a prescient paper on the measurement of welfare, Samuelson (1961) had suggested that the link between asset and rental prices was essential for the integration of income and wealth accounting proposed by Irving Fisher.[8] Our system of accounts employed the specific form of this relationship developed in my paper, 'Capital Theory and Investment Behaviour' (Jorgenson, 1963).

Christensen and I distinguished two approaches to the analysis of economic growth. We identified the production account with a production possibility frontier describing technology. The underlying conceptual framework was an extension of the aggregate production function – introduced by Douglas (1948) and developed by Tinbergen (1942) and Solow (1957) – to include two outputs, investment and consumption goods. These two outputs were distinguished in order to incorporate constant quality indices of investment goods.

We utilized constant quality indices of capital and labour inputs in allocating the sources of economic growth between investment and productivity. Our constant quality index of labour input combined different types of hours worked into a constant quality index of labour input, using methodology that Griliches (1960) had developed for US agriculture. This broadened considerably the concept of substitution employed by Solow and altered, irrevocably, the allocation of economic growth between investment and productivity.[9]

Our constant quality index of capital input combined different types of capital inputs into a constant quality index. We identified input prices with rental rates, rather than the asset prices appropriate for the measurement of capital stock. For this purpose, we used a model of capital as a factor of production I had introduced in Jorgenson (1963), 'Capital Theory and Investment Behaviour'. This made it possible to incorporate differences in returns caused by the tax treatment of different types of capital income.[10]

Our constant quality measure of investment goods generalized Solow's () concept of embodied technical change. A Jorgenson (1966) paper, 'The Embodiment Hypothesis' showed that economic growth could be interpreted, equivalently as 'embodied' in investment or 'disembodied' in productivity growth. This indeterminacy was removed by introducing constant quality indices for investment goods.[11] Bureau of Economic Analysis (BEA) (1986) has now incorporated a constant quality price index for investment in computers into the US national accounts.[12]

Constant quality price indices for investment goods of different ages or vintages were developed by Hall (1971). This important innovation made it possible for Hulten and Wykoff (1982) to estimate relative efficiencies by age for all types of tangible assets included in the national accounts, putting the measurement of capital consumption on to a solid, empirical foundation. Estimates of capital inputs presented in Jorgenson, Gallop and Fraumeni (1987) were based on the Hulten–Wykoff relative efficiencies. BEA (1995) incorporated these relative efficiencies into measures of capital consumption in a benchmark revision of the US National Income and Product Accounts.[13]

Christensen and I identified the income and expenditure account with a social welfare function. The conceptual framework was provided by the representation of intertemporal preferences employed by Ramsey (1928), Samuelson (1961), Cass (1965), Koopmans (1965), and Nordhaus and Tobin (1972). Following Kuznets (1961), we divided the uses of economic growth between current consumption and future consumption through saving. Saving was linked to the asset side of the wealth account through capital accumulation equations for each type of asset. Prices for different vintages were linked to rental prices of capital inputs through a parallel set of capital asset-pricing equations.

The separations of production and welfare approaches to economic growth had important implications for the theory. The Ramsey model, so beautifully exposited by Solow (1989), had two separate submodels – one based on producer behaviour and the other on consumer behaviour. The production account could be linked to the submodel of production, and the income and expenditure account to the submodel of consumption. This made it possible, at least in principle, to proceed from the design stage of the theory of economic growth, emphasized by Solow, to econometric modelling, which he described accurately as 'much more difficult and less glamorous'.[14]

In summary, the dizzying progress of empirical work on economic growth had by 1973 created an impressive agenda for future research. Christensen and I had established the conceptual foundations for quantitative models of growth suitable for analysing the impact of policies affecting investment in tangible assets. However, critical tasks, such as the construction of constant quality indices of capital and labour inputs, and investment goods output, remained to be accomplished.

The final step in this lengthy process was completed only with the benchmark revision of the US National Income and Product Accounts in September 1995.

The growth revival

On 16 October 1973, the beginning of the Arab oil embargo ushered in a series of sharp increases in world petroleum prices that led to a rapidly deepening recession in industrialized countries, accompanied by a rise in inflation. Since this contradicted one of the fundamental tenets of the reigning Keynesian orthodoxy in macroeconomics, it engendered a shift in the focus of macroeconomic research from economic growth to stagflation. Debates among Keynesians, old and new, monetarists, and new classical macroeconomists took centre stage, pushing disputes among the proponents of alternative views on economic growth into the background.

In graduate courses in macroeconomics, the theory of economic growth was displaced gradually by newer topics, such as rational expectations and policy ineffectiveness. Elementary skills required for growth analysis – national income and product accounting, index number theory, the perpetual inventory method, and intertemporal asset pricing – were no longer essential for budding researchers, and fell into disuse. Even the main points of competition in the rancorous debates over growth in the early 1970s began to fade from the collective memory of economists.

Like a watercourse that encounters a mountain range, the stream of research on endogenous growth continued to flow unabated and unobserved, gathering momentum for its later re-emergence into the light of professional debate. When it did erupt in the early 1980s, the initial impulse threatened to wash away the entire agenda that had been put laboriously into place, following the canonical formulation of the neoclassical framework in the early 1970s. The renewed thrust towards endogenizing economic growth acquired startling but illusory force by channelling most of its energy into a polemical attack on the deficiencies of the 'exogenous' theories of growth of Kuznets and Solow.

The flow of new talent into research on economic growth was interrupted for a decade, sapping the high level of intellectual energy that fuelled the rapid progress of the early 1970s. The arrival of a new generation of growth economists in the early 1980s signalled a feverish period of discovery and rediscovery that is still under way. This was followed by a revival of the latent interests of many economists in economic growth after a substantial time lapse. The consequence of this time lapse has been a form of amnesia, familiar to readers who recall Washington Irving's fictional character Rip Van Winkle. To remedy this collective lapse of memory it is essential to bring our story of the dissolution of the neoclassical framework up to date.

We can fix the revival of interest in economic growth by the larger community of economists with some precision at Maddison's (1982) updating and extension of Kuznets' (1971) long-term estimates of the growth of national product for fourteen industrialized countries, including the USA. Maddison added Austria and Finland to Kuznets' list, and presented growth rates covering periods beginning as early as 1820 and extending to 1979. Maddison (1991, 1995) extended these estimates to 1992. Attempts to analyse Maddison's data led to the 'convergence debate' initiated by Abramovitz (1986) and William Baumol (1986).

Denison (1967) compared differences in growth rates for national income per capita for the period 1950–62 with differences of levels in 1960 for eight European countries and the USA. He also compared sources of these differences in both growth rates and levels. The eight European countries as a whole were characterized by much more rapid growth and a lower level of national income per capita. However, this association was not monotonic for comparisons between individual countries and the USA. None the less, Denison's conclusion was that:[15] 'Aside from short-term aberrations Europe should be able to report higher growth rates, at least in national income per person employed, for a long time. Americans should expect this and not be disturbed by it.'

Kuznets (1971) provided elaborate comparisons of growth rates for the fourteen countries included in his study. Unlike Denison (1967), he did not provide level comparisons, but Maddison (1982) filled this gap by comparing levels of national product for sixteen countries. These comparisons were based on estimates of purchasing power parities by Kravis, Heston and Summers (1978).[16] These estimates have been updated by successive versions of the Penn World Table.[17] These data have made it possible to reconsider the issue of convergence of productivity levels raised by Denison (1967).

Abramovitz (1986) was the first to take up the challenge of analysing convergence of productivity levels among Maddison's sixteen countries. He found that convergence appeared to characterize the postwar period, while the period before 1914 and the interwar period revealed no tendencies for productivity levels to converge. Baumol (1986) formalized these results by running a regressions of growth rate of gross domestic product (GDP) per hour over the period 1870–1979 on the 1870 level of GDP per hour worked.[18]

In a notable paper on 'Crazy Explanations for the Productivity Slowdown' Romer (1987) derived a version of the growth regression from Solow's (1989) growth model with a Cobb–Douglas production function. An important empirical contribution of the paper was to extend the data set for growth regressions from Maddison's (1982) group of sixteen advanced countries to the 115 countries included in the Penn World Table (Mark 3), presented by Summers and Heston (1984). Romer's key finding was that an indirect estimate of the Cobb–Douglas elasticity of output with respect to capital was close to three-quarters. The share of capital in GNP implied by Solow's model was less than half as great, on average.[19]

Mankiw, Romer and Weil (1992) provided a defence of the neoclassical framework of Kuznets (1971) and Solow (1989). The empirical portion of their study is based on data for 98 countries in the Penn World Table (Mark 4), presented by Summers and Heston (1988). Like Romer (1987), Mankiw, David Romer and Weil (1992) derived a growth equation from the Solow (1992 [1970]) model; however, they also augmented this model by allowing for investment in human capital.

The results of Mankiw, Romer and Weil produced empirical support for the augmented Solow model. There was clear evidence of the convergence predicted by the model; in addition, the estimated Cobb–Douglas elasticity of output with respect to capital was in line with the share of capital in the value of output. The rate of convergence of productivity was too slow to be consistent with the 1970 version of the Solow model, but is consistent with the augmented version.

Finally, Nasrul Islam (1995) exploited an important feature of the Summers and Heston (1988) data set overlooked in prior empirical studies. This panel data set contains benchmark comparisons of levels of the national product at five-year intervals, beginning in 1960 and ending in 1985. This made it possible for Islam to test an assumption maintained in growth regressions, such as those of Mankiw, Romer and Weil. Their study, like that of Romer (1987), was based on cross-sections of growth rates. Both studies assumed identical technologies for all countries included in the Summer–Heston data sets.

Substantial differences in overall levels of productivity among countries have been documented by Denison (1967), my paper with Christensen and Dianne Cummings (Christensen *et al.*, 1981), and Dougherty and Jorgenson (1996). By introducing econometric methods for panel data, Islam (1995) was able to allow for these differences in technology. He corroborated the finding of Mankiw, Romer and Weil (1992) that the elasticity of output with respect to capital input coincided with the share of capital in the value of output. This was further analysed in the theoretical models of Romer (1986, 1990).

In addition, Islam (1995) found that the rate of convergence of productivity among countries in the Summers and Heston (1988) data set was precisely that required to substantiate the *unaugmented* version of the Solow (1989). In short, 'crazy explanations' for the productivity slowdown, like those propounded by Romer (1987, 1994) are not required to explain the complexities of panels of data for advanced and developing countries. Moreover, the model did not require augmentation, as suggested by Mankiw, Romer and Weil (1992). However, differences in technology among these countries must be taken into account in econometric modelling of differences in growth rates.

The conclusion from Islam's (1995) research is that the Solow model is an appropriate point of departure for modelling the endogenous accumulation of tangible assets. For this purpose it is not essential to endogenize human capital accumulation as well. The rationale for this key empirical finding is that the transition path to balanced growth equilibrium requires decades after a change in policies, such as tax policies, that affect investment in tangible assets. By comparison, the transition after a change in policies affecting investment in human capital requires as much as a century.

Islam's conclusions are reinforced strongly in two important papers by Charles Jones (1995a, 1995b), testing alternative models of economic growth based on endogenous investment in new technology. Jones (1995a) tests models proposed by Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992). This model is based on an endogenous growth rate, proportional to the level of resources devoted to research and development. Jones (1995b) demonstrates that this implication of the model is contradicted by evidence from the advanced countries that conduct the great bulk of research and development. While these countries have increased steadily the resources devoted to research and development, growth rates have been stable or declining.

Jones (1995b) tests models of endogenous investment in new technology proposed by Romer (1986, 1987), Lucas (1988), the so-called 'AK' models. These models have a growth rate that is proportional to the investment rate, and Jones (1995b) shows that there are persistent changes in investment rates for advanced countries, but there are no persistent changes in growth rates. Jones concludes that 'Both AK-style models and the R&D-based models are clearly rejected by this evidence.'[20] Jones (1995a) suggests, as an alternative approach, models that make investment in new technology endogenous by preserving the feature of the Solow model that long-run growth rates are determined by exogenous forces. We consider the obstacles that remain to successful implementation of this approach in the sixth section below.

In summary, the convergence debate provided an excellent medium for the revival of interest in growth. The starting point for this debate was the revival of Kuznets' programme for research on long-term trends in the growth of industrialized countries by Maddison (1982, 1991, 1995). As the debate unfolded, the arrival of successive versions of the Penn World Table engaged the interest of new entrants into the field in cross-section variations in patterns of growth. However, a totally novel element appeared in the form of relatively sophisticated econometric techniques. In the work of Islam (1995) these were carefully designed to bring out the substantive importance of cross-section differences in technology. This proved to be decisive in resolving the debate.

Endogenous growth

Despite substantial progress in endogenizing economic growth since the 1980s, profound differences in policy implications militate against any simple resolution of the debate on the relative importance of investment and productivity. Proponents of income redistribution will not easily abandon the search for a 'silver bullet' that will generate economic growth without the necessity of providing incentives for investment in tangible assets and human capital. Advocates of growth strategies based on capital formation will not readily give credence to claims of the importance of external benefits that 'spill over' to beneficiaries that are difficult or impossible to identify.

The proposition that investment is a more important source of economic growth than productivity is just as controversial now as it was in 1973. The distinction between substitution and technical change emphasized by Solow (1957) parallels the distinction between investment and productivity as sources of economic growth. However, Solow's definition of investment, like that of Kuznets (1971), was limited to tangible assets. Both excluded investments in human capital specifically, by relying on undifferentiated hours of work as a measure of labour input.

Kuznets (1971) and Solow (1957) identified the contribution of tangible assets with increases in the stock, which does not capture adequately substitution among different types of capital inputs. Constant quality indices of both capital labour inputs and investment goods output are essential for successful implementation of the production approach to economic growth. By failing to adopt these measurement conventions, Kuznets and Solow attributed almost all US economic growth to the Solow residual.[21]

To avoid the semantic confusion that pervades popular discussions of economic growth it is essential to be precise in distinguishing between investment and productivity. Investment is the commitment of current resources in the expectation of future returns, and can take a multiplicity of forms. This is the definition introduced by Fisher (1906) and discussed by Samuelson (1961). The distinctive feature of investment as a source of economic growth is that the returns can be internalized by the investor. The most straightforward application of this definition is to investments that create property rights, including rights to transfer the resulting assets and benefit from incomes that accrue to the owners.[22]

Investment in tangible assets provides the most transparent illustration of investment as a source of economic growth. This form of investment creates transferable property rights with returns that can be internalized. However, investment in tangible assets through R&D also creates intellectual property rights that can be transferred through outright sale or royalty arrangements, and returns that can be internalized. Private returns to this forms of investment – returns that have been internalized – have been studied intensively in the literature surveyed by Griliches (1994, 1995) and Hall (1996).

The seminal contributions of Gary Becker (1993) Machlup (1962), Mincer (1974), and Schultz (1961) have given concrete meaning to the concept of 'wealth in its more general sense' employed by Fisher. This notion of wealth includes investments that do not create property rights. For example, a student enrolled in a school, or a worker participating in a training programme, can be viewed as an investor. Although these investments do not create assets that can be bought or sold, the returns to higher educational qualifications or better skills in the workplace can be internalized. The contribution of investments in education and training to economic growth can be identified in the same way as for tangible assets.

The mechanism by which tangible investments are translated into economic growth is well understood. For example, an investor in a new industrial facility adds to the supply of assets and generates a stream of rental income. The investment and the income are linked through markets for capital assets and capital services. The income stream can be divided between the increase in capital input and the marginal product of capital or rental price. The increase in capital contributes to output growth in proportion to the marginal product. This is the basis for construction of a constant quality index of capital input.

Griliches (1973, 1979, 1995) has shown how investments in new technology can be translated into economic growth. An investor in a new product design or process of production adds to the supply of intellectual assets and generates a stream of profits or royalties. The increase in intellectual capital contributes to output growth in proportion to its marginal product in the same way as the acquisition of a tangible asset. However, investments in R&D, unlike those in tangible assets, are frequently internal to the firm, so that separation of the private return between the input of intellectual capital and the marginal product or rental price of this capital is highly problematical. Bureau of Labour Statistics (BLS) (1994) and Griliches (1973) have provided estimates of the contribution of these investments to economic growth.

Finally, an individual who completes a course of education or training adds to the supply of people with higher qualifications or skills. The resulting income stream can be decomposed into a rise in labour input and the marginal product of labour or wage rate. The increase in labour contributes to output growth in proportion to the marginal product. This provides a basis for constructing a constant quality index of labour input. Although there are no asset markets for human capital, investments in human and non-human capital have the common feature, pointed out by Fisher (1906), that returns are internalized by the investor.

The defining characteristic of productivity as a source of economic growth is that the incomes generated by higher productivity are external to the economic activities that generate growth. These benefits 'spill over' to income recipients not involved in these activities, severing the connection between the creation of growth and the incomes that result. Since the benefits of policies to create externalities cannot be appropriated, these policies typically involve government programmes or activities supported through public subsidies. Griliches (1992, 1995) has provided detailed surveys of 'spillovers' from investment in R&D.[23]

Publicly supported research and development programmes are a leading illustration of policies to stimulate productivity growth. These programmes may be conducted by government laboratories or financed by public subsidies to private laboratories. The justification for public financing is most persuasive for aspects of technology that cannot be fully appropriated, such as basic science and generic technology. The benefits of the resulting innovations are external to the economic units conducting the R&D and these must be distinguished carefully from the private benefits of R&D that can be internalized through the creation of intellectual property rights.

An important obstacle to the resolution of the debate over the relative importance of investment and productivity is that it coincides with ongoing disputes about the appropriate role for the public sector. Productivity can be identified with spillovers of benefits that do not provide incentives for actors within the private sector. Advocates of a larger role for the public sector advance the view that these spillovers can be guided into appropriate channels only by an all-wise and beneficent government sector. By contrast, proponents of a smaller government search for means to privatize decisions about investments by decentralizing investment decisions among participants in the private sector of the economy.

Jorgenson and Stiroh (1995) showed that investments in tangible assets are the most important sources of postwar US economic growth. These investments appear on the balance sheets of firms, industries, and the nation as a whole, as buildings, equipment and inventories. The benefits appear on the income statements of these same economic units as profits, rents, and royalties. BLS (1983) compiled an official constant quality index of capital input for its initial estimates of total factor productivity, renamed 'multifactor productivity'.

BLS retained hours worked as a measure of labour input until 11 July 1994, when it released a new multifactor productivity measure incorporating a constant quality index of labour input as well as BEA's (1986) constant quality index for investment in computers. The final step in implementing a constant quality index of the services of tangible assets empirically was the incorporation of Hulten and Wykoff (1982) relative efficiencies into the US National Income and Product Accounts by BEA (1995). Four decades of empirical research, initiated by Goldsmith's (1955–6) monumental treatise, *A Study of Saving*, have provided a sound empirical foundation for endogenizing investment in tangible assets.

The growth of labour input is second only in importance to capital input as a source of economic growth. Increases in labour incomes have made it possible to measure investments in human capital and assess their contributions to economic growth. In Jorgenson and Fraumeni (1989), we extended the vintage accounting system developed in Christensen and Jorgenson (1973) to incorporate these investments. Our essential idea was to treat individual members of the US population as human assets with 'asset prices' given by their lifetime labour incomes. Constant quality indices of labour input are an essential first step in incorporating investments in human capital into empirical studies of economic growth. We implemented our vintage accounting system for both human and non-human capital for the USA on an annual basis for the period 1948–84.

Asset prices for tangible assets can be observed directly from market transactions in investment goods; intertemporal capital asset pricing equations are used to derive rental prices for capital services. For human capital, wage rates correspond to rental prices and can be observed directly from transactions in the labour market. Lifetime labour incomes are derived by applying asset pricing equations to these wage rates. These incomes are analogous to the asset prices in accounting for tangible assets in the system of vintage accounts developed in Christensen and Jorgenson (1973).

Fraumeni and I gave developed a measure of the output of the US education sector, presented in Jorgenson and Fraumeni (1992b). Our point of departure was that, while education is a service industry, its output is investment in human capital. We estimated investment in education from the impact of increases in educational attainment on the lifetime incomes of all individuals enrolled in school. We found that investment in education, measured in this way, is similar in magnitude to the value of working time for all individuals in the labour force. Furthermore, the growth of investment in education during the postwar period exceeded the growth of market labour activities.

Second, we have measured the inputs of the education sector, beginning with the purchase inputs recorded in the outlays of educational institutions, in Jorgenson and Fraumeni (1992a). A major part of the value of the output of educational institutions accrues to students in the form of increases in their lifetime incomes. Treating these increases as compensation for student time, we evaluated this time as an input into the educational process. Given the outlays of educational institutions and the value of student time, we allocated the growth of the education sector to its sources.

An alternative approach, employed by Schultz (1961), Machlup (1962), Nordhaus and Tobin (1972), and many others, is to apply Goldsmith's (1955–6) perpetual inventory method to private and public expenditure on educational services. Unfortunately, this approach has foundered on the absence of a satisfactory measure of the output of the educational sector and the lack of an obvious rationale for capital consumption. The approach fails to satisfy the conditions for integration of income and wealth accounts established by Fisher (1906) and Samuelson (1961).[24]

Given vintage accounts for human and non-human capital, Jorgenson and Fraumeni (1989) constructed a system of income, product and wealth accounts, paralleling the system developed earlier in Christensen and Jorgenson (1973). In these accounts, the value of human wealth was more than ten times the value of non-human wealth, while investment in human capital was five times investment in tangible assets. We defined 'full' investment in the US economy as the sum of these two types of investment. Similarly, we added the value of non-market labour activities to personal consumption expenditure to obtain 'full' consumption. Our product measure included these new measures of investment and consumption.

Since our complete accounting system included a production account with 'full' measures of capital and labour inputs, we were able to generate a new set of accounts for the sources of US economic growth. Our system also included an income and expenditure account, with income from labour services in both market and non-market activities. We combined this with income from capital services and allocated 'full' income between consumption and saving.[25] This provided the basis for a new measure of economic welfare and a set of accounts for the uses of US economic growth. Our system was completed by a wealth account containing both human wealth and tangible assets.

We aggregated the growth of education and non-education sectors of the US economy to obtain a new measure of US economic growth. Combining this with measures of input growth, we obtained a new set of accounts for the sources of growth of the US economy. Productivity contributes almost nothing to the growth of the education sector, and only a modest proportion to output growth for the economy as a whole. We also obtain a second approximation to the proportion of US economic growth that can be made endogenous. Within a Kamsey model with separate education and non-education sector we find that exogenous productivity growth accounts for only 17 per cent of growth.

The introduction of endogenous investment in education increases the explanatory power of the Ramsey model of economic growth to 83 per cent. However, it is important to emphasize that growth without endogenous investment in education is measured differently. The traditional framework for economic measurement of Kuznets (1971) and Solow (1989) excludes non-market activities, such as those that characterize the major portion of investment in education. The intuition is familiar to any teacher, including teachers of economics: what the students do is far more important than what the teachers do, even if the subject matter is the theory of economic growth.

A third approximation to the proportion of growth that could be attributed to investment within an extended Ramsey model results from incorporation of all forms of investment in human capital. This would include education, child-rearing, and the addition of new members to the population. Fertility could be made endogenous by using the approach of Barro and Becker (1989) and Becker and Barro (1988). Child-rearing could be made endogenous by modelling the household as a producing sector along the lines of the model of the educational sector I have outlined above. The results presented by Jorgenson and Fraumeni (1989) show that this would endogenize 86 per cent of US economic growth. This is a significant, but not overwhelming, gain in explanatory power for the Ramsey model.

In summary, endogenizing US economic growth at the aggregate level requires a distinction between investment and productivity as sources of growth. There are two important obstacles to empirical implementation of this distinction. First, the distinctive feature of investment as a source of growth is that the returns can be internalized. Decisions can be decentralized successfully to the level of individual investors in human capital and tangible assets. Productivity growth is generated by spillovers that cannot be captured by private investors. Activities generating these spillovers cannot be decentralized and require collective decision-making through the public sector. Successive approximations to the Ramsey model of economic growth increase the proportion of growth that can be attributed to investment, rather than productivity.

Econometric modelling

We are prepared, at last, for the most difficult and least glamorous part of the task of endogenizing economic growth-constructing models for the analysis of economic policies, the Ramsey growth model of Cass (1965) and Koopmans (1965) requires the empirical implementation of two highly problematical theoretical constructs, namely, a model of producer behaviour based on an aggregate production function and a model of a representative consumer. Each of these abstracts from important aspects of economic reality, but both have important advantages in modelling long-term trends in economic growth.

My paper, 'Accounting for Capital' (Jorgenson, 1980) presented a methodology for aggregating over sectors. The existence of an aggregate production function imposes very stringent conditions on production patterns at the industry level. In addition to value-added functions for each sector, an aggregate production function posits that these functions must be identical. Furthermore, the functions relating sectoral capital and labour inputs to their components must be identical, and each component must receive the same price in all sectors.[26]

Although the assumptions required for the existence of an aggregate production function appear to be highly restrictive, in Jorgenson and Fraumeni (1980) we estimated that errors of aggregation could account for less than 9 per cent of aggregate productivity growth.[27] In Jorgenson *et al.* (1987) Gollop, Fraumeni and I published updated data on sectoral and aggregate production accounts. We generated the data for sectoral production accounts in a way that avoids the highly restrictive assumptions of the aggregate production function. These data were then compared with those from the aggregate production account to test for the existence of an aggregate production function. We demonstrated that this hypothesis is inconsistent with empirical evidence. However, our revised and updated estimate of errors arising from aggregation over industrial sectors explained less than 3 per cent of aggregate productivity growth over the period of our study, 1948–79.[28]

In the same book, we also presented statistical tests of the much weaker hypothesis that a value-added function exists for each industrial sector, but this hypothesis was also rejected.[29] The conclusion of our research on production at the sectoral level was that specifications of technology such as the aggregate production function and sectoral valued-added functions result in substantial oversimplifications of the empirical evidence. However, these specifications are useful for particular but limited purposes. For example, sectoral value-added functions are indispensable for aggregating over sectors, while the aggregate production function is a useful simplification for modelling aggregate long-run growth, as originally proposed by Tinbergen (1942).

Sectoral value-added functions were employed by Hall (1988, 1990a) in modelling production at the sectoral level. In measuring capital and labour inputs, he adhered to the traditional framework of Kuznet (1971) and Solow (1989), by identifying labour input with hours worked and capital input with capital stock. He found large apparent increasing returns to scale in the production of value-added.[30] Basu and Fernald (1996) have pointed out that the value-added data employed by Hall are constructed on the basis of assumptions of constant returns to scale and perfect competition.

Basu and Fernald (1996) have employed the strategy for sectoral modelling of production recommended in Jorgenson *et al.* (1987), treating capital, labour and intermediate inputs symmetrically. They estimate returns to scale for the sectoral output and input data presented in Jorgenson, 1990 to be constant. These data include constant quality measures of capital, labour and intermediate input. Basu and Fernald (1996) also show that returns to scale in the production of value-added are constant, when value-added is defined in the same way as in my 1987 book with Gallop and Fraumeni, and constant quality measures of capital and labour inputs are employed.

Data for individual firms provide additional support for value-added production functions with constant, or even decreasing, returns to scale. Estimates incorporating intellectual capital have been surveyed by Griliches (1994, 1995) and Hall (1996).[31] These estimates are now available for many different time periods and several countries. Almost all existing studies employ value-added data for individual firms and provide evidence for constant or decreasing returns to scale. This evidence is corroborated further by an extensive study of plant-level data by Baily *et al.* (1992), providing evidence of constant returns at the level of individual manufacturing plants.

Turning to the task of endogenizing investment in tangible assets and education, we first review the endogenous accumulation of tangible assets. An important objective of the Christensen and Jorgenson (1973) accounting system was to provide the data for econometric modelling of aggregate producer and consumer behaviour. In collaboration with Lau, Christensen and I introduced an econometric model of producer behaviour in Christensen *et al.* (1973) We modelled joint production of consumption and investment goods from inputs of capital and labour services, utilizing data on these outputs and inputs from the aggregate production account.

In Christensen *et al.* (1975) we constructed an econometric model of representative consumer behaviour. We estimated this model on the basis of data from the aggregate income and expenditure account of the Christensen and Jorgenson (1973) accounting system. We tested and rejected the implications of a model of a representative consumer. Subsequently, in Jorgenson *et al.* (1982), we constructed a model of consumer behaviour based on exact aggregation over individual consumers that specializes in the representative consumer model for a fixed distribution of total expenditure over the population of consumers.[32]

Yun and (Jorgenson and Yun, 1986) constructed an econometric model for postwar US economic growth with endogenous accumulation of tangible assets. Our model of consumer behaviour involved endogenous labour–leisure choice, following Tinbergen's (1942) neoclassical econometric model of economic growth. Labour–leisure choice is exogenous in Solow's (1956) neoclassical model. In addition, we employed the Ramsey (1928) representation of intertemporal preferences to model saving–consumption behaviour, following Cass (1965) and Koopmans (1965). In Solow's model, the saving ratio is exogenous.

The econometric application of Ramsey's model of optimal saving was initiated by Hall (1978), removing the final remaining gap between theoretical and empirical perspectives on economic growth.[33] This occurred only eight years after Solow's classic exposition of the neoclassical theory of growth! The key to Hall's achievement in 1978 was the introduction of an econometrically tractable concept of 'rational expectations', which he combined successfully with Ramsey's theoretical model. Building on Hall's framework, Hansen and Singleton (1982, 1983) have tested and rejected the underlying model of a representative consumer.

Yun and Jorgenson and Yun (1990) have revised and updated our econometric model of US economic growth and analysed the consequences of the Tax Reform Act of 1986 for US economic growth. We also considered alternative proposals for fundamental tax reform, including proposals now under consideration by the US Congress, such as consumption-based and income-based value added taxes. We found that the 1986 Act resulted in a substantial increase in social welfare. However, we also discovered that several of the alternative proposals would have produced substantially higher gains.

The econometric model of US economic growth developed in Jorgenson and Yun (1990, 1991a, 1991b) provides the starting point for the endogenous growth model of the US economy that I constructed with Ho Jorgenson and (1995). While my model with Yun endogenized capital input, the endogenous growth model also endogenizes investment in human capital. This model includes all the elements of our Ramsey model of US economic growth. However, the new model also includes a highly schematic model of production for the US educational system.

Our production model includes a production possibility frontier for the non-education sector that is analogous to the frontier in Jorgenson and Yun (1990, 1991a, 1991b). The model also includes a production function for the education sector with investment in education as the output. The inputs include capital and labour services as well as purchases of goods and services from the non-education sector. For both submodels, we allow for exogenous growth of productivity; however, Jorgenson and Fraumeni (1992a) show that this is negligible for the education sector.

Ho and I in Jorgenson and Ho (1995) we evaluated alternative educational policies through the equivalent variation in wealth associated with each policy. As an alternative case, we consider an educational policy that would raise the participation rates and policies, keeping taxes and expenditures constant. Presumably, this would result in a lower level of 'quality'. We also consider an alternative case that would retain the base case participation rates, but raise 'quality' by increasing expenditures on consumption goods, and capital and labour services in the education sector and corresponding taxes. Hanushek (1994) has shown that the second of these alternative policies, substantial improvement in educational quality through increased expenditure, is closely comparable to the actual educational policy pursued during the 1980s.

Ho and in Jorgenson and Ho (1995) have shown that increasing participation rates without altering expenditure would produce substantial gains in social welfare. In this sense, the 'quality' level of the existing educational system is too high to be cost effective. On the other hand, increasing 'quality' with no change in participation rates would result in a sizeable loss in social welfare. These results are consistent with the literature in educational production functions surveyed by Hanushek (1986, 1989).[34]

With endogenous accumulation of tangible capital, as in the model constructed in Jorgenson and Yun (1986), almost three-quarters of growth is endogenous. By contrast, the model with endogenous investment in education constructed in Jorgenson and Ho (1995) accounts for 83 per cent of growth. By endogenizing fertility behaviour and child-rearing it would be possible, at least in principle, to add an incremental three percentage points to the explanatory power of the Ramsey model of economic growth. Modelling population growth endogenously is clearly feasible. However, the construction of an econometric model with this feature would require considerable new data development and is best left as an opportunity for future research.

In summary, the endogenous models of growth constructed Jorgenson and Yun (1986a, 1986b) and Jorgenson and Ho (1995) require the econometric implementation of concepts of an aggregate production function and a representative consumer. While each of these concepts has important limitations, both are useful in modelling long-run economic trends. Furthermore, these concepts lead naturally to a substantial increase in the level of sophistication in data generation, integrating investment and capital into a complete system of national accounts.

Conclusion

The key innovation in economic measurement required for endogenizing growth is a wealth account that can be integrated with production, and income and expenditure accounts. This encompasses the system of vintage accounts for tangible assets implemented in Christensen and Jorgenson (1973), as well as the vintage accounts for human capital developed in Jorgenson and Fraumeni (1989). These incorporate accumulation equations for tangible assets and human capital, together with asset-pricing equations. Both are essential in constructing endogenous models of growth to replace the exogenous models that emerged from the professional consensus of the early 1970s.

The framework for economic measurement developed in Christensen and Jorgenson (1973) and Jorgenson and Fraumeni (1989) incorporates the principal features of the United Nations (1993) *System of National Accounts*. This provides a production account for allocating the sources of economic growth between investment and growth in productivity. It also includes an income and expenditure account for analysing the uses of economic growth through consumption and saving. Alternative policies are ranked by means of equivalent variations in wealth for the representative consumer.

In principle, investment in new technology could be made endogenous by extending the accounting framework to incorporate investment in new technology. BEA (1994) has provided a satellite system of accounts for research and development, based on Goldsmith's (1955–6) perpetual inventory method, applied to private and public expenditure. Unfortunately, this is subject to the same limitations as the approach to human capital of Schultz (1961) and Machlup (1962). The BEA satellite system has foundered on the absence of a satisfactory measure of the output of research and development, and the lack of an appropriate rationale for capital consumption.

The standard model for investment in new technology, formulated by Griliches (1973) is based on a production function incorporating inputs of services from intellectual capital accumulated through investment in R&D. Intellectual capital is treated as a factor of production in precisely the same way as tangible assets in Christensen and Jorgenson (1973). Hall (1993) has developed the implications of this model for the pricing of the services of intellectual capital input and the evaluation of intellectual capital assets.[35]

Griliches (1973) represented the process of R&D by means of a production function which included the services of the previous R&D. This captures the notion of 'standing on the shoulders of giants', originated by Schmookler (1966) and elaborated by Caballero and Jaffe (1993), and Jones and Williams (1996). Under constant returns to scale, this representation also captures the 'congestion externality' modelled by Jones and Williams, and Stokey (1995). Research and development, leading to investment in intellectual capital, is conducted jointly with production of marketable output, and this poses a formidable obstacle to measuring the output of new intellectual capital.

The model of capital as a factor of production first proposed in Jorgenson (1963) has been applied to tangible assets and human capital. However, successful implementation of this model for intellectual capital would require a system of vintage accounts, including not only accumulation equations for stocks of accumulated research and development, but also asset-pricing equations. These equations are essential for separating the revaluation of intellectual property because of price changes over time from depreciation of this property resulting from ageing. This is required to enable measurement of the quantity of intellectual capital input and its marginal product.

Pricing of intellectual capital is the key issue remaining before investment in new technology can be endogenized in quantitative models for the analysis of alternative economic policies. Hall (1993) has constructed prices for stocks of accumulated intellectual capital from stock market valuations of the assets of individual firms. However, she points out that the high degree of persistence in expenditures on R&D at the firm level has made it virtually impossible to separate the effects of the ageing of assets from changes in the value of these assets over time. Her evaluation of intellectual capital is conditional upon a pattern of relative efficiencies imposed on past investments in new technology.

None the less, Hall's pioneering research on the pricing of intellectual assets has yielded interesting and valuable insights. For example, the gross rate of return in the computer and electronics industry, including depreciation and revaluation of these assets, greatly exceeds that in other industries. This can be rationalized by the fact that revaluation in this industry, as measured by Hall, is large and negative, mirroring the rapid decline in the price of the industry's output. This is evidence for the empirical significance of the process of creative destruction described by Schumpeter (1942) and modelled by Aghion and Howitt (1992), Stokey (1995), and Jones and Williams (1996). Since revaluation enters negatively into the gross rate of return, this rate of return exceeds that for industries with positive revaluations.

Another important result that emerges for Hall's (1996) survey of gross rates of return to R&D is the repeated finding that investment funded by the federal government contract has been unable to internalize the returns. This has the very important policy implication that public investments in new technology can be justified only by comparisons of the costs and benefits to the government. Measurement of these benefits requires careful case studies, such as those of civilian space technology by Herzfeld (1985) and commercial aircraft by Mowery (1985). Grandiose visions of spillovers from public research and development have been exposed as a rapidly fleeting mirage.

The final issue that must be resolved in order to complete the endogenization growth is modelling of spillovers. Griliches (1995) has provided a detailed survey of alternative methodologies and results, based on the model he originated in 1979. The essential idea is to include aggregate input of intellectual capital, together with the inputs of individual producers, as a determinant of output. Unfortunately, this requires precisely the same separation of marginal product and capital input for intellectual capital needed for the identification of returns that can be internalized by the individual producer.

Caballero and Lyons (1990, 1992) have attempted to circumvent the problem of measuring intellectual capital by including aggregate output as a determinant of sectoral productivity. However, Basu and Fernald (1995) have shown that Caballero and Lyons' positive results depend on the same value-added data employed by Hall (1988, 1990a). Treating capital, labour and intermediate inputs symmetrically, as in their research on economies of scale, Basu and Fernald show that the evidence for spillovers evaporates. This leaves open the question of the importance of spillovers from investment in new technology, which must await satisfactory measures of the output of R&D.

An elegant and impressive application of Griliches' (1979) framework for modelling spillovers across international boundaries has been presented by Coe and Helpman (1995). The key idea is to trace the impact of these spillovers through trade in intermediate goods. For each country, the stock of accumulated R&D of its trading partners is weighted by bilateral import shares. However, Keller (1996) has shown that the evidence of spillovers is even more impressive if the bilateral trade shares are assigned randomly, rather than matched with the countries conducting the R&D. Another vision of spillovers can be assigned to the lengthening roll of unproven theoretical hypotheses.

In summary, a great deal has been accomplished, but much remains to be done to complete the endogenization of economic growth. An important feature of recent research, for example, in the seminal papers of Romer (1986, 1987, 1990), has been the linking of theoretical and empirical investigations. This integration need no longer be left to the remarkable coincidence of empirical and theoretical perspectives that led Kuznets (1971) and Solow (1989) to the neoclassical framework. In the absence of a clear and compelling link between the theoretical model and the data generation process, the breakdown of this framework had left economists since the 1970s without a guide to long-run economic policy.

Fortunately, a new empirical and theoretical consensus on economic growth would require only a relatively modest reinterpretation of the neoclassical framework established by Solow (1956, 1970, 1988), Cass (1965), and Koopmans (1965). However, the traditional framework of economic measurement established by Kuznets (1961, 1971) and embedded in the US National Income and Product Accounts will have to be augmented considerably. The most important change is a reinterpretation of the concepts of investment and capital to encompass Fisher's notion of 'wealth in its more general sense'.

In closing, I must emphasize that my goal has been to provide a new starting point in the search for a consensus on economic growth, rather than to arrive at final conclusions. The new framework I have outlined is intended to be open-ended, permitting a variety of different approaches to investment – in tangible assets, human capital and new technology. There is also ample, if carefully delimited, space within this framework for endogenizing spillovers; for example, by using the Lindahl–Samuelson theory of public goods. New entrants to the field will continue to find a plethora of opportunities for modelling economic growth.

Notes

1 I have benefited greatly from the help of my colleague, Zvi Griliches, in navigating the recent empirical literature on investment in new technology. Susanto Basu and Charles Jones, as well as Griliches, have kindly provided me with access to unpublished material. Financial support was provided by the Program on Technology and Economic Policy of Harvard University. Responsibility for any remaining deficiencies rests solely with the author.

2 Assar Lindbeck (1992) (ed.) *Nobel Lectures in Economic Sciences, 1969–1980* (River Edge, NJ: World Scientific Publishing), p. 79.

3 Karl-Goran Maler (1992) (ed.) *Nobel Lectures in Economic Sciences, 1981–1990*, (River Edge, NJ: World Scientific Publishing), p. 191.

4 See Jorgenson and Griliches (1967), table X, p. 272. We also attributed 13 per cent of growth to the relative utilization of capital, measured by energy consumption as a proportion of capacity; however, this is inappropriate at the aggregate level, as Denison (1974), p. 56, pointed out. For additional details, see Jorgenson *et al.* (1987), esp. pp. 179–81.

5 This paper was presented at the thirty-seventh meeting of the Conference on Research in Income and Wealth, held at Princeton, New Jersey in 1971.

6 See, for example, Office of Business Economics (1966).

7 The United Nations System of National Accounts (SNA) was summarized by Stone (1984) in his Nobel Prize address. The SNA has been revised in United Nations (1993).

8 See Samuelson (1961), esp. p. 309.

9 Constant quality indices of labour input are discussed in detail by Jorgenson, *et al.* (1987), chs 3 and 8, pp. 69–108 and 261–300; and Jorgenson *et al.* (1994).

10 A detailed survey of empirical research on the measurement of capital input is given in my 1996 paper, 'Empirical Studies of Depreciation', and Jack Triplett's (1996) paper, 'Measuring the Capital Stock: A Review of Concepts and Data Needs', both presented at a meeting of the Conference on Research in Income and Wealth, held in Washington, DC, in May 1992.

11 A detailed history of constant quality price indices is given by Berndt (1991), Triplett's (1990) contribution to the Jubilee of the Conference on Research in Income and Wealth discusses obstacles to the introduction of these indices into government statistical programmes. Gordon (1990) constructed constant quality indices for all types of producers' durable equipment in the national accounts, and Pieper (1989, 1990) gave constant quality indices for all types of structures.

12 Cole *et al.* (1986) reported the results of a joint project conducted by BEA and IBM to construct a constant quality index for computers. Triplett (1986) discussed the economic interpretation of constant quality price indices in an accompanying article. Dulberger (1988) presented a more detailed report, while Triplett (1989) gave an extensive survey of empirical research on constant quality price indices for computers. Young (1989) answered Denison's (1989) objections and reiterated BEA's rationale for introducing a constant quality price index for computers.

13 The methodology is described by Fraumeni (1997).

14 See Solow (1992 [1970]), p. 105. He went on to remark, 'But it may be what God made graduate students for. Presumably He had something in mind.'

15 See Denison (1967), esp. ch. 21, 'The Sources of Growth and the Contrast between Europe and the United States', pp. 296–348.
16 For details, see Maddison (1982), pp. 159–68. Purchasing power parities were first measured for industrialized countries by Gilbert and Kravis (1954) and Gilbert (1958).
17 A complete list up to Mark 5 is given by Summers and Heston (1991), while the results of Mark 6 are summarized by the World Bank in the *World Development Report 1993*.
18 This 'growth regression' has spawned a vast literature, summarized by Levine and Renelt (1992), Baumol (1994) and Barro and Sala-i-Martin (1994). Much of this literature has been based on successive versions of the Penn World Table.
19 Unfortunately, this Mark 3 data set did not include capital input. Romer's empirical finding has spawned a substantial theoretical literature, summarized at an early stage by Lucas (1988) and, more recently, by Grossman and Helpman (1991, 1994), Romer (1994) and Barro and Sala-i-Martin (1994). Romer's own important contributions to this literature have focused on increasing returns to scale, as in Romer (1986), and spillovers from technological change, as in Romer (1990).
20 Jones (1995b), p. 519.
21 The measurement conventions of Kuznets and Solow remain in common use. See, for example, the references given in Jorgenson (1990), 'Productivity and Economic Growth', based on a paper presented at the Jubilee of the Conference on Research in Income and Wealth, held in Washington, DC in 1988. For recent examples, see Baily and Gordon (1988), Englander and Mittlestadt (1988), Blanchard and Fischer (1989), pp. 2–5, Baily and Schultze (1990), Gordon (1990), Englander and Gurney (1994) and Lau (1994).
22 Fischer ( ) discusses property rights in ch. 2, pp. 18–40.
23 Griliches (1992) also gives a list of survey papers on spillovers. Griliches (1979, 1995) has shown how to incorporate spillovers into growth accounting.
24 For a more detailed discussion, see Jorgenson and Fraumeni (1989).
25 Our terminology follows that of Becker's (1965, 1993) theory of time allocation.
26 A detailed survey of econometric modelling of production is included in Jorgenson (1986): 'Econometric Methods for Modeling Producer Behavior'. This is also the focus of Solow's (1967) survey article, 'Some Recent Developments in the Theory of Production'. The conceptual basis for the existence of an aggregate production function was provided by Hall (1973).
27 Fraumeni and Jorgenson (1980), table 2.38, lines 4 and 11.
28 Jorgenson *et al.* (1987), table 9.5, lines 6 and 11.
29 Jorgenson *et al.* (1987), table 7.2, pp. 239–241. The existence of an aggregate production function requires identical value added functions for all sectors.
30 Hall (1990a) reports the median degree of returns to scale in value added from two-digit US manufacturing industries of 2.2!
31 Hall (1996) gives a list of survey papers.
32 A survey of empirical approaches to aggregation is given by Stoker (1993).
33 Hall's (1978) paper and his subsequent papers on this topic have been reprinted in Hall (1990b), *The Rational Consumer*. Hall and Deaton (1992)

have presented surveys of the literature on econometric modelling of consumer behaviour within the Ramsey framework.

34 Note that the meaning of 'production function' in this context is different from the meaning of this term in our model of the education sector. In Hanushek's terminology, the output of the education sector is measured in terms of measures of educational performance, such as graduation rates or test scores. Our terminology is closer to the Hanushek's (1994) concept of 'value added' by the educational system. The output of the education system is the addition to the lifetime incomes of all individuals enrolled in school.

35 These implications of the model are also discussed by Jones and Williams (1996).

References

Abramovitz, M. (1956) 'Resources and Output Trends in the United States since 1870', *American Economic Review*, vol. 46, no. 3, May, pp. 5–23.

Abramovitz, M. (1986) 'Catching Up, Forging Ahead, and Falling Behind', *Journal of Economic History*, vol. 46, no. 2, June, pp. 385–406.

Aghion, P. and Howitt, P. (1992) 'A Model of Growth through Creative Destruction', *Econometrica*, vol. 60, no. 2, March, pp. 323–51.

Baily, M. N. and Gordon, R. J. (1988) 'Measurement Issues, the Productivity Slow-down, and the Explosion of Computer Power', *Brookings Papers on Economic Activity*, no. 2, pp. 1–45.

Baily, M. N. and Schultze, C. L. (1990) 'The Productivity of Capital in a Period of Slower Growth', *Brookings Papers on Economic Activity: Microeconomics*, vol. 1, pp. 369–406.

Baily, M. N., Hulten, C. R. and Campbell, D. (1992) 'Productivity Dynamics in Manufacturing Plants', *Brookings Papers on Economic Activity: Microeconomics*, vol. 1, pp. 187–267.

Barro, R. J. and Becker, G. S. (1989) 'Fertility Choice in a Model of Economic Growth', *Econometrica*, vol. 7, no. 2, March, pp. 481–502.

Barro, R. J. and Sala-i-Martin, X. (1994) *Economic Growth* (New York: McGraw-Hill).

Basu, S. and Fernald, J. G. (1995) 'Are Apparent Productive Spillovers a Figment of Specification Error?', *Journal of Monetary Economics*, vol. 36, no. 1, August, pp. 165–88.

Basu, S. and Fernald, J. G. (1996) 'Returns to Scale in U.S. Production: Estimates and Implications', International Finance Discussion Papers, No. 546, Washington, DC, Board of Governors of the Federal Reserve System, March.

Baumol, W. J. (1986) 'Productivity Growth, Convergence, and Welfare', *American Economic Review*, vol. 76, no. 5, December, pp. 1072–85.

Baumol, W. J. (1994) 'Multivariate Growth Patterns: Contagion and Common Forces as Possible Sources of Convergence', in W. J. Baumol, R. R. Nelson and E. N. Wolff (eds), *Convergence of Productivity* (New York: Oxford University Press), pp. 62–85.

Becker, G. S. (1965) 'A Theory of the Allocation of Time', *Economic Journal*, vol. 75, no. 296, September, pp. 493–517.

Becker, G. S. (1967) 'Human Capital and the Personal Distribution of Income: An Analytical Approach', Woytinsky Lecture No. 1, Institute of Public Administration, University of Michigan, Ann Arbor.

Becker, G. S. (1993) *Human Capital*, 3rd edn, (1st edn 1964; 2nd edn 1975) (Chicago, University of Chicago Press).

Becker, G. S. and Barro, R. J. (1988) 'A Reformulation of the Economic Theory of Fertility', *Quarterly Journal of Economics*, vol. 103, no. 1, February, pp. 1–25.

Berndt, E. R. (1991) *The Practice of Econometrics: Classic and Contemporary* (Reading, MA: Addison-Wesley).

Blanchard, O. J. and Fischer, S. (1989) *Lectures on Macroeconomics* (Cambridge, MA: MIT Press).

Bureau of Economic Analysis (BEA) (1982) *Measuring Nonmarket Economic Activity: BEA Working Papers*, Bureau of Economic Analysis Working Paper No. 2, US Department of Commerce, Washington DC, December.

Bureau of Economic Analysis (BEA) (1986) 'Improved Deflation of Purchases of Computers', *Survey of Current Business*, vol. 66, no. 3, March, pp. 7–9.

Bureau of Economic Analysis (BEA) (1994) 'A Satellite Account for Research and Development', *Survey of Current Business*, vol. 74, no. 11, November, pp. 37–71.

Bureau of Economic Analysis (BEA) (1995) 'Preview of the Comprehensive Revision of the National Income and Product Accounts: Recognition of Government Investment and Incorporation of a New Methodology for Calculating Depreciation', *Survey of Current Business*, vol. 75, no. 9, September, pp. 33–41.

Bureau of Labour Statistics (BLS) (1983) *Trends in Multifactor Productivity*, Bulletin No. 2178, US Department of Labour, Washington, DC.

Bureau of Labour Statistics (1993) *Labour Composition and U.S. Productivity Growth 1948–90*, Bulletin 2426, US Department of Labour, Washington, DC.

Bureau of Labour Statistics (1994) 'Multifactor Productivity Measures 1991 and 1992', News release USDL 94–327, 11 July.

Caballero, R. J. and Jaffe, A. B. (1993) 'How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth', in O. J. Blanchard and S. Fischer (eds), *NBER Macroeconomics Annual 1993* (Cambridge, MA: MIT Press), pp. 15–74.

Caballero, R. J. and Lyons, R. K. (1990) 'Internal and External Economics in European Industries', *European Economic Review*, vol. 34, no. 4, April, pp. 805–30.

Caballero, R. J. and Lyons, R. K. (1992) 'External Effect in U.S. Procyclical Productivity', *Journal of Monetary Economics*, vol. 29, no. 2, April, pp. 209–26.

Carson, C. and Landefeld, S. (1994) 'Integrated Economic and Environmental Satellite Accounts', *Survey of Current Business*, vol. 74, no. 4, April.

Cass, D. (1965) 'Optimum Growth in an Aggregative Model of Capital Accumulation', *Review of Economic Studies*, vol. 32(3), no. 91, July, pp. 233–40.

Christensen, L. R. and Jorgenson, D. W. (1969) 'The Measurement of U.S. Real Capital Input, 1929–1967', *Review of Income and Wealth*, series 15, no. 4, December, pp. 293–320.

Christensen, L. R. and Jorgenson, D. W. (1970) 'U.S. Real Product and Real Factor Input, 1929–1967', *Review of Income and Wealth*, series 16, no. 1, March, pp. 19–50.

Christensen, L. R. and Jorgenson, D. W. (1973) 'Measuring Economic Performance in the Private Sector', in M. Moss (ed.), *The Measurement of Economic and Social Performance*, pp. 233–338.

Christensen L. R., Cummings, D. and Jorgenson, D. W. (1981) 'Relative Productivity Levels 1947–1973', vol. 16, no. 1, May, pp. 61–94.

Christensen, L. R., Jorgenson, D. W. and Lau, L. J. (1973) 'Transcendental Logarithmic Production Frontiers', *Review of Economics and Statistics*, vol. 55, no. 1, February, pp. 28–45.

Christensen, L. R., Jorgenson, D. W. and Lau, L. J. (1975) 'Transcendental Logarithmic Utility Functions', *American Economic Review*, vol. 65, no. 3, June, pp. 367–83.

Coe, D. T. and Helpman, E. (1995) 'International R&D Spillovers', *European Economic Review*, vol. 39, no. 5, May, pp. 859–87.

Cole, R., Chen, Y. C., Barquin-Stolleman, J. A., Dulberger, E., Helvacian, N. and Hodge, J. H. (1986) 'Quality-adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment', *Survey of Current Business*, vol. 66, no. 1, January, pp. 41–50.

Deaton, A. (1992) *Understanding Consumption* (Oxford University Press), pp. 76–135.

Denison, E. F. (1967) *Why Growth Rates Differ* (Washington, DC: The Brookings Institution).

Denison, E. F. (1974) *Accounting for United States Economic Growth* (Washington, DC: The Brookings Institution).

Denison, E. F. (1989) *Estimates of Productivity Change by Industry* (Washington, DC: The Brookings Institution).

Domar, E. (1946) 'Capital Expansion, Rate of Growth and Employment', *Econometrica*, vol. 14, no. 2, April, pp. 137–47.

Dougherty, C. and Jorgenson, D. W. (1996) 'International Comaprisons of the Sources of Economic Growth', *American Economic Review*, vol. 86, no. 2, May, pp. 25–9.

Douglas, P. H. (1948) 'Are There Laws of Production?', *American Economic Review*, vol. 38, no. 1, March, pp. 1–41.

Dulberger, E. (1988) 'The Application of a Hedonic Model to a Quality-Adjusted Index for Computer Processors', in D. W. Jorgenson and R. Landau (eds), *Technology and Capital Formation* (Cambridge, MA: MIT Press), pp. 37–76.

Englander, A. S. and Mittelstadt, A. (1988) 'Total Factor Productivity: Macroeconomic and Structural Aspects of the Slowdown', *OECD Economic Studies*, vol. 10, Spring, pp. 7–56.

Englander, A. S. and Gurney, A. (1994) 'OECD Productivity Growth: Medium Term Trends', *OECD Economic Studies*, vol. 22, Spring, pp. 111–30.

Fisher, I. (1906) *The Nature of Capital and Income* (New York: Macmillan).

Fraumeni, B. M. (1997) 'The Measurement of Depreciation in the U.S. National Income and Wealth Accounts', *Survey of Current Business*, July, vol. 77, no. 7, pp. 7–23.

Gilbert, M. and Kravis, I. B. (1954) *An International Comparison of National Products and the Purchasing Power of Currencies* (Paris: OEEC).

Gilbert, M., Beckerman, W., Edelman, J., Marris, S., Stuvel, G. and Teichert, M. (1958) *Comparative National Products and Price Levels* (Paris: OEEC).

Goldsmith, R. (1955–6) *A Study of Saving in the United States*, 3 vols (Princeton, NJ: Princeton University Press).

Goldsmith, R. (1962) *The National Wealth of the United States in the Postwar Period* (New York, National Bureau of Economic Research).

Gordon, R. J. (1990) *The Measurement of Durable Goods Prices* (University of Chicago Press).

Griliches, Z. (1960) 'Measuring Inputs in Agriculture: A Critical Survey', *Journal of farm Economics*, vol. 40, no. 5, December, pp. 1398–1427.

Griliches, Z. (1973) 'Research Expenditures and Growth Accounting', in B. R. Williams (ed.), *Science and Technology in Economic Growth* (London: Macmillan), pp. 59–95.

Griliches, Z. (1979) 'Issues in Assessing the Contribution of Research and Development to Productivity Growth', *Bell Journal of Economics*, vol. 10, no. 1, Spring, pp. 92–116.

Griliches, Z. (1992) 'The Search for R&D Spillovers', *Scandinavian Journal of Economics*, vol. 94, Supplement, pp. 29–47.

Griliches, Z. (1994) 'Productivity, R&D, and the Data Constraint', *American Economic Review*, vol. 84, no. 1, March, pp. 1–23.

Griliches, Z. (1995) 'R&D and Productivity Econometric Results and Measurement Issues', in P. Stoneman (ed.), *Handbook of the Economics of Innovation and Technological Change* (Oxford: Basil Blackwell), pp. 52–89.

Grossman, G. M. and Helpman, E. (1991) *Innovation and Growth* (Cambridge, MA: MIT Press).

Grossman, G. M. and Helpman, E. (1994) 'Endogenous Innovation in the Theory of Growth', *Journal of Economic Perspectives*, vol. 8, no. 1, Winter, pp. 23–44.

Hall, B. (1993) 'Industrial Research in the 1980s: Did the Rate of Return Fall?', *Brookings Papers on Economic Activity: Microeconomics*, vol. 2, pp. 289–331.

Hall, B. (1996) 'The Private and Social Returns to Research and Development', in B. L. R. Smith and C. E. Barfield (eds), *Technology, R&D, and the Economy* (Washington, DC: The Brookings Institution and the American Enterprise Institute), pp. 140–62.

Hall, R. E. (1971) 'The Measurement of Quality Change from Vintage Price Data', in Z. Griliches (ed.), *Price Indexes and Quality Change* (Cambridge, MA: Harvard University Press), pp. 240–71.

Hall, R. E. (1973) 'The Specification of Technology with Several Kinds of Output', *Journal of Political Economy* vol. 81, no. 4, July/August, pp. 878–92.

Hall, R. E. (1978) 'Stochastic Implications of the Life-Cycle–Permanent Income Hypothesis: Theory and Evidence', *Journal of Political Economy*, vol. 86, no. 6, December, pp. 971–87.

Hall, R. E. (1988) 'The Relation between Price and Marginal Cost in U.S. Industry', *Journal of Political Economy*, vol. 96, no. 5, October, pp. 921–47.

Hall, R. E. (1990a) 'Invariance Properties of Solow's Productivity Residual', in P. Diamond (ed.), *Growth/Productivity/Employment* (Cambridge, MA: MIT Press), pp. 71–112.

Hall, R. E. (1990b) *The Rational Consumer* (Cambridge, MA: MIT Press).

Hall, R. E. and Jorgenson, D. W. (1967) 'Tax Policy and Investment Behavior', *American Economic Review*, vol. 57, no. 3, June, pp. 391–414.

Hall, R. E. and Jorgenson, D. W. (1969) 'Tax Policy and Investment Behavior: Reply and Further Results', *American Economic Review*, vol. 59, no. 3, June, pp. 388–401.

Hall, R. E. and Jorgenson, D. W. (1971) 'Applications of the Theory of Optimal Capital Accumulation', in G. Fromm (ed.), *Tax Incentives and Capital Spending* (Amsterdam: North-Holland), pp. 9–60.

Hansen, L. P. and Singleton, K. J. (1982) 'Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models', *Econometrica*, vol. 50, no. 6, November, pp. 1269–86.

Hansen, L. P. and Singleton, K. J. (1983) 'Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Stock Market Returns', *Journal of Political Economy*, vol. 91, no. 2, April, pp. 249–65.

Hanushek, E. A. (1986) 'The Economics of Schooling', *Journal of Economic Literature*, vol. 24, no. 4, December, pp. 1141–78.

Hanushek, E. A. (1989) 'The Impact of Differential Expenditures on School Performance', *Educational Researcher*, vol. 18, no. 4, May, pp. 45–51.

Hanushek, E. A. (1994) *Making Schools Work* (Washington, DC: The Brookings Institution).

Harrod, R. (1939) 'An Essay in Dynamic Theory', *Economic Journal*, vol. 49, no. 194, March, pp. 14–33.

Hertzfeld, H. R. (1985) 'Measuring the Economic Impact of Federal Research and Development Activities', Workshop on the Federal Role in Research and Development, National Academies of Science and Engineering, November 21–22.

Hulten, C. R. and Wykoff, F. C. (1982) 'The Measurement of Economic Depreciation', in C. R. Hulten (ed.), *Depreciation, Inflation and the Taxation of Income from Capital* (Washington, DC, Urban Institute Press), pp. 81–125.

Inter-Secretariat Working Group on National Accounts (1993) *System of National Accounts 1993* (New York: United Nations), pp. 379–406.

Islam, N. (1995) 'Growth Empirics', *Quarterly Journal of Economics*, vol. 110, no. 4, November, pp. 1127–70.

Jones, C. I. (1995a) 'R&D-Based Models of Economic Growth', *Journal of Political Economy*, vol. 103, no. 4, August, pp. 759–84.

Jones, C. I. (1995b) 'Time Series Tests of Endogenous Growth Models', *Quarterly Journal of Economics*, vol. 110, issue 2, May, pp. 495–526.

Jones, C. I. and Williams, J. C. (1996) 'Too Much of a Good Thing? The Economics of Investment in R&D', Department of Economics, Stanford University, 26 February.

Jorgenson, D. W. (1963) 'Capital Theory and Investment Behaviour', *American Economic Review*, vol. 53, no. 2, May, pp. 247–59.

Jorgenson, D. W. (1966) 'The Embodiment Hypothesis', *Journal of Political Economy*, vol. 74, no. 1, February, pp. 1–17.

Jorgenson, D. W. (1973) 'The Economic Theory of Replacement and Depreciation', in W. Sellekaarts (ed.), *Econometrics and Economic Theory* (New York: Macmillan), pp. 189–221.

Jorgenson, D. W. (1980) 'Accounting for Capital', in G. M. von Furstenberg (ed.), *Capital, Efficiency, and Growth* (Cambridge, MA: Ballinger), pp. 251–319.

Jorgenson, D. W. (1986) 'Econometric Modeling of Producer Behaviour', in Z. Griliches and M. D. Intriligator (eds), *Handbook of Econometrics*, vol. 3, (Amsterdam, North-Holland), pp. 1841–915.

Jorgenson, D. W. (1990) 'Productivity and Economic Growth', in E. R. Bernd and J. Triplett (eds), *Fifty Years of Economic Measurement* (University of Chicago Press), pp. 19–118.

Jorgenson, D. W. (1996) 'Empirical Studies of Depreciation', *Economic Inquiry*, vol. 34, no. 1, January, pp. 24–42.

Jorgenson, D. W. and Fraumeni, B. M. (1989) 'The Accumulation of Human and Nonhuman Capital, 1948–1984', in R. E. Lipsey and H. S. Tice (eds), *The Measurement of Saving, Investment, and Wealth* (University of Chicago Press), pp. 227–82.

Jorgenson, D. W. and Fraumeni, B. M. (1992a) 'Investment in Education and U.S. Economic Growth' *Scandinavian Journal of Economics*, vol. 94, Supplement, pp. 51–70.

Jorgenson, D. W. and Fraumeni, B. M. (1992b) 'The Output of the Education Sector', in Z. Griliches (ed.), *Output Measurement in the Services Sector* (University of Chicago Press), pp. 303–38.

Jorgenson, D. W., Gallop, F. M. and Fraumeni, B. M. (1987) *Productivity and U.S. Economic Growth* (Cambridge: Harvard University Press).

Jorgenson, D. W., and Ho, M. S. (1995) 'Modeling Economic Growth', Harvard University, Department of Economics.

Jorgenson, D. W., Ho, M. S. and Fraumeni, B. M. (1994) 'The Quality of the U.S. Work Force, 1948–90', Harvard University, Department of Economics.

Jorgenson, D. W., Lau, L. J. and Stoker, T. M. (1982) 'The Transcendental Logarithmic Model of Aggregate Consumer Behavior', in R. L. Basmann and G. Rhodes (eds), *Advances in Econometrics*, vol. 1 (Greenwich: JAI Press) pp. 97–238.

Jorgenson, D. W. and Stiroh, K. (1995) 'Computers and Growth', *Economics of Innovation and New Technology*, vol. 3, December, pp. 295–316.

Jorgenson, D. W. and Yun, K.-Y. (1986a) 'The Efficiency of Capital Allocation', *Scandinavian Journal of Economics*, vol. 88, no. 1, pp. 85–107.

Jorgenson, D. W. and Yun, K.-Y. (1986b) 'Tax Policy and Capital Allocation', *Scandinavian Journal of Economics*, vol. 88, no. 2, pp. 355–77.

Jorgenson, D. W. and Yun, K.-Y. (1990) 'Tax Reform and U.S. Economic Growth', *Journal of political Economy*, vol. 98, no. 5, Part 2, October, pp. 151–93.

Jorgenson, D. W. and Kun, K.-Y. (1991a) 'The Excess Burden of U.S. Taxation', *Journal of Accounting, Auditing, and Finance*, vol. 6, no. 4, Fall, pp. 487–509.

Jorgenson, D. W. and Kun, K.-Y. (1991b) *Tax Reform and the Cost of Capital* (Oxford University Press), pp. 1–38.

Keller, W. (1996) 'Are International R&D Spillovers Trade-related? An Analysis Using Monte-Carlo Techniques', Department of Economics, University of Wisconsin, Madison, March 16.

Koopmans, T. C. (1965), 'On the Concept of Optimum Growth', in *The Econometric Approach to Development Planning* (Chicago: Rand-McNally) pp. 225–300.

Kravis, I. B., Heston, A. and Summers, R. (1978) *International Comparisons of Real Product and Purchasing Power* (Baltimore: Johns Hopkins University Press).

Kuznets, S. (1961), *Capital in the American Economy* (Princeton University Press).

Kuznets, S. (1971) Economic Growth of Nations (Cambridge: Harvard University Press).

Kuznets, S. (1992) 'Modern Economic Growth: Findings and Reflections', in A. Lindbeck (ed.), *Nobel Lectures: Economic Sciences, 1969–1980*, pp. 87–102.

Lau, L. J. (1996) 'The Sources of Long-Term Economic Growth: Observations from the Experience of Developed and Developing Countries', in R. Landau,

T. Taylor, and G. Wright, (eds), *The Mosaic of Economic Growth* (Stanford University Press) pp. 63–91.

Levine, R., and Renelt, D. (1992) 'A Sensitivity Analysis of Cross-Country Regressions', *American Economic Review*, vol. 82, no. 4, September, pp. 942–63.

Lindbeck, A. (1992) (ed.) *Nobel Lectures in Economic Sciences, 1969–1980* (River Edge: World Scientific Publishing Company).

Lucas, R. E. (1988) 'On The Mechanics of Economic Development', *Journal of Monetary Economics*, vol. 22, no. 1, July, pp. 2–42.

Machlup, F. (1962) *The Production and Distribution of Knowledge in the United States* (Princeton University Press).

Maddison, A. (1982), *Phases of Capitalist Development* (Oxford University Press).

Maddison, A. (1991), *Dynamic Forces in Capitalist Development* (Oxford University Press).

Maddison, A. (1995), *Monitoring the World Economy* (Paris: OECD).

Maler, K.-G. (1992) *Nobel Lectures in Economic Sciences* (River Edge: World Scientific Publishing Company).

Mankiw, N. G., Romer, D. and Weil, D. (1992) 'A Contribution to the Empirics of Economic Growth', *Quarterly Journal of Economics*, vol. 107, no. 2, May, pp. 407–37.

Mincer, J. (1974) *Schooling, Experience, and Earnings* (New York: Columbia University Press).

Mowery, D. C. (1985) 'Federal Funding of R&D in Transportation: The Case of Aviation', Workshop on the Federal Role in Research and Development, National Academies of Science and Engineering, November 21–22.

Nordhaus, W. D. and Tobin, J. (1972) *Is Growth Obsolete?*, New York, National Bureau of Economic Research, reprinted in M. Moss (ed.), The Measurement of Economic and Social Performance (New York: Columbia University Press) 1973, pp. 509–32.

Pieper, P. E. (1989) 'Construction Price Statistics Revisited', in D. W. Jorgenson and R. Landau, (eds), *Technology and Capital Formation* (Cambridge: MIT Press) pp. 293–330.

Pieper, P. E. (1990) 'The Measurement of Construction Prices: Retrospect and Prospect', in E. R. Berndt and J. Triplett (eds), *Fifty Years of Economic Measurement* (University of Chicago Press) pp. 239–68.

Ramsey, F. (1928) 'A Mathematical Theory of Saving', *Economic Journal*, vol. 28, no. 112, December, pp. 543–59.

Romer, P. (1986) 'Increasing Returns and Long-Run Growth', *Journal of Political Economy*, vol. 94, no. 5, October, pp. 1002–37.

Romer, P. (1987) 'Crazy Explanations for the Productivity Slowdown', in Stanley Fischer (ed.), *NBER Macroeconomics Annual* (Cambridge: MIT Press) pp. 163–201.

Romer, P. (1990) 'Endogenous Technological Change', *Journal of Political Economy*, vol. 98, no. 5, Part 2, October, pp. S71–S102.

Romer, P. (1994) 'The Origins of Endogenous Growth', *Journal of Economic Perspectives*, vol. 8, no. 1, Winter, pp. 3–20.

Samuelson, P. A. (1961) 'The Evaluation of 'Social Income': Capital Formation and Wealth', in F. A. Lutz and D. C. Hague, (eds), *The Theory of Capital* (London: Macmillan) pp. 32–57.

Schmookler, J. (1966) *Invention and Economic Growth* (Cambridge: Harvard University Press).
Schumpeter, J. (1942) *Capitalism, Socialism, and Democracy* (New York: Harper).
Schultz, T. W. (1961) 'Investment in Human Capital', *American Economic Review*, vol. 51, no. 1, March, pp. 1–17.
Solow, R. M. (1956) 'A Contribution to the Theory of Economic Growth', *Quarterly Journal of Economics*, vol. 70, no. 1, February, pp. 65–94.
Solow, R. M. (1957) 'Technical Change and the Aggregate Production Function' *Review of Economics and Statistics*, vol. 39, no. 3, August, pp. 312–20.
Solow, R. M. (1960) 'Investment and Technical Progress', in K. J. Arrow, S. Karlin, and P. Suppes (eds), *Mathematical Methods in the Social Sciences, 1959* (Stanford University Press) pp. 89–104.
Solow, R. M. (1988) 'Growth Theory and After', *American Economic Review*, vol. 78, no. 3, June, pp. 307–17.
Solow, R. M. (1989) *Growth Theory*: An *Exposition*, (New York: Oxford University Press).
Solow, R. M. (1994) 'Perspectives on Growth Theory', *Journal of Economic Perspectives*, vol. 8, no. 1, Winter, pp. 45–54.
Stoker, T. M. (1993) 'Empirical Approaches to the Problem of Aggregation over Individuals', *Journal of Economic Literature*, vol. 31, no. 4, December, pp. 1827–74.
Stokey, N. L. (1995) 'R&D and Economic Growth', *Review of Economic Studies*, vol. 62(3), no. 212, July 1995, pp. 469–90.
Stone, R. (1992) 'The Accounts of Society', in K.-G. Maler (ed.), *Nobel Lectures: Economic Sciences, 1981–1990*, pp. 115–39.
Summers, R. and Heston, A. (1984) 'Improved International Comparisons of Real Product and Its Composition: 1950–1980', *Review of Income and Wealth*, series 30, no. 2, June, pp. 207–62.
Summers, R. and Heston, A. (1988) 'A New Set of International Comparisons of Real Product and Price Levels: Estimates for 130 Countries, 1950–1985', *Review of Income and Wealth*, series 34, no. 1, March, pp. 1–25.
Summers, R. and Heston, A. (1991) 'The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988', *Quarterly Journal of Economics*, vol. 106, issue 2, May, pp. 327–68.
Summers, R., Kravis, I. B. and Heston, A. (1980) 'International Comparisons of Real Product and Its Composition, 1950–77', *Review of Income and Wealth*, series 26, no. 1, March, pp. 19–66.
Tinbergen, J. (1942) 'On the Theory of Trend Movements', in J. Tinbergen, *Selected Papers*, Amsterdam, North-Holland, 1959, pp. 182–221 (translated from 'Zur Theorie der Langfristigen Wirtschaftsentwicklung', *Weltwirtschaftliches Archiv*, band 55, nu. 1, 1942, pp. 511–49).
Triplett, J. (1986) 'The Economic Interpretation of Hedonic Methods', *Survey of Current Business*, vol. 66, no. 1, January, pp. 36–40.
Triplett, J. (1989) 'Price and Technological Change in a Capital Good: Survey of Research on Computers', in D. W. Jorgenson and R. Landau (eds), *Technology and Capital Formation* (Cambridge: MIT Press) pp. 127–213.
Triplett, J. (1990), 'Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy', in E. R. Berndt and J. Triplett (eds), *Fifty Years of Economic Measurement*, pp. 207–38.

Triplett, J. (1996) 'Measuring the Capital Stock: A Review of Concepts and Data Needs', *Economic Inquiry*, vol. 34, no. 1, January, pp. 93–115.
United Nations (1968) *A System of National Accounts* (New York: United Nations).
World Bank (1994) *World Development Report 1993* (Washington, DC: World Bank).
Young, A. (1989) 'BEA's Measurement of Computer Output', *Survey of Current Business*, vol. 69, no. 7, July, pp. 108–15.

10
Conclusions

Edmund S. Phelps

This volume provides an authoritative survey of present-day economics of growth. Little would be accomplished in reviewing this survey. I propose instead to contrast the early growth economics with the new insights and re-emphases now emerging and, in light of recent work and recent policy developments, to assess the prospects that countries have for stepping up their economic growth through better economic policy.

The processes of economic growth were the subject of my first sustained economic research – over the first half of the 1960s – and I drew a few tentative conclusions at that time. I could not understand learning by *doing* as a *source* of economic growth. Such learning seemed to me a sort of friction that operates as a drag on progress, since it would be better if every invention were instantaneously and costlessly known. I came to the conclusion that the prime mover of economic growth was the innovator. Richard Nelson helped me to understand the process of technological diffusion. The capacity of a small producer (such as my farming ancestors in downstate Illinois) to adopt a technological improvement – some recently created process or product, or even a fairly old one not yet in wide use – is facilitated by a higher education, which enables or eases the grasp of technological and other needed information. (This is learning by *reading* technical evaluations, manufacturers' specifications, and so on.) Ultimately, though, it is the entrepreneur – the heroic figure of Joseph Schumpeter's theory – who, if he or she can come up with the necessary resources, *creates* innovations, generally with scientists, engineers, designers, lawyers, and other specialists as employees or partners. I thought of these innovating producers as private entrepreneurs rather than agents of the state.

This emphasis on innovation of the private entrepreneur was common ground among several growth theorists – for example, Edwin Mansfield, Karl Shell, Hirofumi Uzawa and Kenneth Arrow, Nelson, and myself – by the mid-1960s. If readers wonder why we left this important field about that time, the answer was not, I think, that we did not know how to reconcile research and innovation with market competition, perfect or imperfect. Schumpeter's basic answer was familiar enough to us, and Arrow's work on the atomistic innovator versus the monopolist innovator had extended the analysis to free entry. The answer, in my opinion, is that it would take a great deal of time and thought on our part to understand the role of private ownership, and thus to understand why state enterprises could not generally be expected to be innovative.

The postwar history of growth economics generally did not emphasize private entrepreneurship, though – not until the shock of the collapse of the alternative systems in Eastern Europe. In the 1950s, the standard thesis was that the low-wage countries are suffering from an undersupply of capital. However, Wassily Leontieff came along in the middle of the decade with the suggestion that human labour in the high-wage countries is augmented by superior skills. For that reason, the world's capital stock was allocated predominantly to the high-wage countries. The less developed countries were therefore seen as receiving the right amount of capital.

That step forward raised a question, though: why do workers in low-wage countries suffer from low levels of skills? The neoclassical answer given was that their education is poor. As a result, their employers find that they are less trainable, and more costly to train. To this day, education is a factor heavily emphasized by neoclassical investigators – see, for example, the growth equations of Robert Barro, Gregory Mankiw and co-authors.

This answer raises in turn the question: why is educational attainment so low in those countries where wages are found to be low? The explanation cannot be that those countries are poor, since the answer seeks to explain their poverty by their low education, not the other way around. A plausible answer, if only a partial one, is that a country's education improves, and approaches the world average, only in small steps from generation to generation. In countries where few people have a lengthy education, most children are handicapped in their own education by their parents' low educational attainment; and, in such conditions, both private and public support for education is apt to be low. This line of argument was originated by Yoram Ben-Porath and Gary Becker.

The answer on which the neoclassical brand of growth studies has put more stress, however, is that, generally speaking, enterprises face serious investment risks in a large number of countries, with the result that the productivity of education in those countries is lower than in the more fortunate nations. If, as argued by Douglass North and more recently by Aaron Tornell and Andreas Velasco, property rights are endangered because of social instability and unresolved political conflict, less tangible capital is invested in the affected countries; too much human activity is diverted to crime, rent-seeking and political power-seeking, and too little to building enterprises. As a result, the same amount of human capital invested in these unfortunate nations would show lower returns than in the countries without these sociopolitical handicaps.

This answer, however, raises yet another question: why are these countries so occupied with the activity of redistribution, through crime and rent-seeking, while the high-wage countries are not? Some commentators like to theorize in terms of the purported benefits of a Confucian–Tao culture or the purported burdens of a 'Latin' culture. However, most analysts are reluctant, and reasonably so, to accept answers that assert the force of a culture, since that is hard to measure, and that assume cultures to be more cause than effect, when that is hard to test.

The answer that many of us would propose is that in the countries where too much activity revolves around redistribution, the prevailing economic institutions have not given the populace anything better to do. Private enterprise has not spread and strengthened enough to compete for the attention of the working-age population. Of course, where there is a dearth of private enterprise – an excess of state enterprise – it is likely that there is an entire syndrome of impediments to economic growth. This answer derives from a methodological perspective according to which it is poverty of the economic institutions inherited in some countries, rather than a set of sociopolitical hazards inherent in some parts of the world, that is the obstacle to the achievement of high wages. Today, we associate the emphasis on private enterprise with models stimulated by events in Eastern Europe; for example, some of the work by Roman Frydman and by Andrei Shleifer, and their collaborators. But we should remember that the role of private entrepreneurship was common currency among growth theorists in the 1960s, as I noted above, and we should also acknowledge that the importance of capitalism has been insisted on in a series of works by Mancur Olson and Oliver Williamson, to name just two of a select group, since the 1980s. Since the early 1990s, particularly in a series of reports for the European Bank for Reconstruction and Development, I have been navigating my own re-entry into this hugely important area.

Of course, the answer that a shortage of capitalism is the problem raises, in turn, the question: why have the poor countries not liberated their populations from redistribution conflicts by fostering private enterprise? A tenable answer, it seems to me, is the operation of so-called 'path dependence'. I am thinking here of the work by Raquel Fernandez and Dani Roderick in which they demonstrate a bias in favour of the status quo ante. All the individuals in the society would prefer the alternative state, say, capitalism, to the current state, say, predominantly socialist enterprises, *if* they did not know their particular position in the current state, and hence thought they all had the same chance of a gain as everyone else, as in John Rawls's hypothetical 'veil of ignorance'; but in fact the vested interests know perfectly well that they are winners in the existing system, and these 'insiders' do not want to take the risk of being 'outsiders' in the alternative system. We see this problem all the time, for example, in Mexico, where the government's attempts to privatize some downstream operations of Pemex, the state-owned petrochemical giant, met the opposition of the entrenched old-boy network who wanted to keep their existing advantages.

Does this mean that the movement towards private enterprise has gone about as far as possible? I think not. As the populace of a country sees much of the world forging ahead under private enterprise, the existing system with its large state sector and high level of rent-seeking will remain attractive for fewer and fewer people. At some point, there will emerge a majority of the population who see it as preferable to take the risk of being an outsider in the alternative system for a chance of a large gain that would come from turning out a winner in that system.

What I suggest we have witnessed since the 1980s is just such a momentous swing towards capitalism – a swing in one country after another away from huge state employment and rent-seeking and towards a broader role for private enterprise. But, if the above sketch of the dynamic process is right, it will take another decade or so for this swing to take hold in all or nearly all the countries of the world.

As this process unfolds, we shall undoubtedly see that capitalism is not a panacea. It will not solve the problem of poverty among the least advantaged. But the quickened pace of growth that this process promises to bring will help to make possible an earlier and stronger attack on these problems.