Ramaprasad Bhar
and
Shigeyuki Hamori

Hidden Markov Models

Applications to Financial Economics

Advanced Studies in Theoretical and Applied Econometrics

KLUWER ACADEMIC PUBLISHERS

Hidden Markov Models

Advanced Studies in Theoretical and Applied Econometrics

Volume 40

The titles published in this series are listed at the end of this volume.

Hidden Markov Models

Applications to Financial Economics

by

Ramaprasad Bhar

School of Banking and Finance,
The University of New South Wales, Sydney, Australia

and

Shigeyuki Hamori

Graduate School of Economics,
Kobe University, Japan

KLUWER ACADEMIC PUBLISHERS

NEW YORK, BOSTON, DORDRECHT, LONDON, MOSCOW

eBook ISBN: 1-4020-7940-0
Print ISBN: 1-4020-7899-4

Created in the United States of America

**To Rajiv, Mitra,
Hitoshi, Makoto, and
Naoko**

This page intentionally left blank

Contents

Dedication v

Acknowledgments xi

List of Figures xiii

List of Tables xvii

1. INTRODUCTION 1

  1 Introduction 1

  2 Markov Chains 1

  3 Passage Time 5

  4 Markov Chains and the Term Structure of Interest Rates 6

  5 State Space Methods and Kalman Filter 11

  6 Hidden Markov Models and Hidden Markov Experts 13

  7 HMM Estimation Algorithm 16

  8 HMM Parameter Estimation 18

  9 HMM Most Probable State Sequence: Viterbi Algorithm 22

  10 HMM Illustrative Examples 24

2. VOLATILITY IN GROWTH RATE OF REAL GDP 29

  1 Introduction 29

  2 Models 31

    2.1 GARCH Model 31

    2.2 Markov Switching Variance Model 32

  3 Data 33

  4 Empirical Results 33

5 Conclusion 38

3. LINKAGES AMONG G7 STOCK MARKETS 41

1 Introduction 41

2 Empirical Technique 44

2.1 Markov Switching Stock Return Model 44

2.2 Concordance Measure 45

3 Data 46

4 Empirical Results 46

5 Conclusion 51

4. INTERPLAY BETWEEN INDUSTRIAL PRODUCTION AND STOCK MARKET 55

1 Introduction 55

2 Markov Switching Heteroscedasticity Model of Output and Equity 58

3 Data 62

4 Empirical Results 63

5 Conclusion 76

5. LINKING INFLATION AND INFLATION UNCERTAINTY 81

1 Introduction 81

1.1 Inflation and Inflation Uncertainty 81

1.2 Inflation Uncertainty and Markov Switching Model 83

2 Empirical Technique 85

2.1 Markov Switching Heteroscedasticity Model of the Inflation Rate 85

2.2 Non-Nested Model Selection using Vuong Statistic 86

3 Data 87

4 Empirical Results 91

5 Conclusion 107

6. EXPLORING PERMANENT AND TRANSITORY COMPONENTS OF STOCK RETURN 117

1 Introduction 117

2 Markov Switching Heteroscedasticity Model of Stock Return 119
3 Data 120
4 Empirical Results 121
5 Conclusion 125

7. EXPLORING THE RELATIONSHIP BETWEEN COINCIDENT FINANCIAL MARKET INDICATORS 127
1 Introduction 127
2 Markov Switching Coincidence Index Model 129
3 Data 131
4 Empirical Results 131
5 Conclusion 139

References 145

Index 153

This page intentionally left blank

Acknowledgments

We dedicate this monograph to express our thanks to our families, friends, and colleagues for their invaluable support in the writing of this volume. Without the excellent editorial guidance of Mrs. Cathelijne van Herwaarden and Mrs. Herma Drees, the kind and helpful comments of Mr. Akira Tokihisa, and the love of our family members Rajiv, Mitra, Hitoshi, Makoto, and Naoko, we never would have succeeded in completing this work. This research was partly supported by a grant-in-aid from the Japan Society for the Promotion of Science.

This page intentionally left blank

List of Figures

1.1	Filtering Algorithm	13
2.1	Estimated variance from the Markov switching heteroskedasticity model: Japan	36
2.2	Estimated variance from the Markov switching heteroskedasticity model: UK	37
2.3	Estimated variance from the Markov switching heteroskedasticity model: USA	37
4.1	Industrial production growth: Canada	67
4.2	Excess return: Canada	67
4.3	Estimated filtered probability, $\Pr(S_t = 2)$: Canada	67
4.4	Industrial production growth: France	68
4.5	Excess return: France	68
4.6	Estimated filtered probability, $\Pr(S_t = 2)$: France	68
4.7	Industrial production growth: Germany	69
4.8	Excess return: Germany	69
4.9	Estimated filtered probability, $\Pr(S_t = 2)$: Germany	69
4.10	Industrial production growth: Italy	70
4.11	Excess return: Italy	70
4.12	Estimated filtered probability, $\Pr(S_t = 2)$: Italy	70
4.13	Industrial production growth: Japan	71
4.14	Excess return: Japan	71
4.15	Estimated filtered probability, $\Pr(S_t = 2)$: Japan	71
4.16	Industrial production growth: UK	72
4.17	Excess return: UK	72
4.18	Estimated filtered probability, $\Pr(S_t = 2)$: UK	72

4.19 Industrial production growth: USA 73
4.20 Excess return: USA 73
4.21 Estimated filtered probability, $\Pr(S_t = 2)$: USA 73
5.1 Inflation rate: Germany 88
5.2 Inflation rate: Japan 88
5.3 Inflation rate: UK 89
5.4 Inflation rate: USA 89
5.5 Relationship among the probability of high variance state for permanent shocks, $\Pr(S_t = 1)$, the probability of high variance state for transitory shocks, $\Pr(S_t = 2)$, and inflation rates: Germany 92
5.6 Relationship among the probability of high variance state for permanent shocks, $\Pr(S_t = 1)$, the probability of high variance state for transitory shocks, $\Pr(S_t = 2)$, and inflation rates: Japan 92
5.7 Relationship among the probability of high variance state for permanent shocks, $\Pr(S_t = 1)$, the probability of high variance state for transitory shocks, $\Pr(S_t = 2)$, and inflation rates: UK 93
5.8 Relationship among the probability of high variance state for permanent shocks, $\Pr(S_t = 1)$, the probability of high variance state for transitory shocks, $\Pr(S_t = 2)$, and inflation rates: USA 93
5.9 Probability of high variance state for permanent shocks: Germany 96
5.10 Probability of high variance state for permanent shocks: Japan 96
5.11 Probability of high variance state for permanent shocks: UK 97
5.12 Probability of high variance state for permanent shocks: USA 97
5.13 Probability of high variance state for transitory shocks: Germany 98
5.14 Probability of high variance state for transitory shocks: Japan 98
5.15 Probability of high variance state for transitory shocks: UK 99
5.16 Probability of high variance state for transitory shocks: USA 99

5.17 Components of forecast variance at $k = 2$: Germany 103
5.18 Components of forecast variance at $k = 4$: Germany 103
5.19 Components of forecast variance at $k = 2$: Japan 104
5.20 Components of forecast variance at $k = 4$: Japan 104
5.21 Components of forecast variance at $k = 2$: UK 105
5.22 Components of forecast variance at $k = 4$: UK 105
5.23 Components of forecast variance at $k = 2$: USA 106
5.24 Components of forecast variance at $k = 4$: USA 106
7.1 Estimated coincident indicator and excess return: Japan 136
7.2 Estimated coincident indicator and excess return: UK 136
7.3 Estimated coincident indicator and excess return: USA 137
7.4 Inferred probability of low variance state: Japan 137
7.5 Inferred probability of low variance state: UK 138
7.6 Inferred probability of low variance state: USA 138

This page intentionally left blank

List of Tables

2.1 Growth rate in real GDP: GARCH(1,1) estimation 34
2.2 Diagnostics using standardized residuals from GARCH(1,1) model 34
2.3 Growth rate in real GDP: Markov switching heteroskedasticity estimation 35
2.4 Diagnostics using standardized residuals from Markov switching heteroskedasticity model 35
3.1 Summary statistics on stock return 47
3.2 Correlation of stock returns 47
3.3 Parameter estimates: Markov switching heteroscedasticity model of stock return 48
3.4 Diagnostics using standardized returns from Markov switching heteroscedasticity model 49
3.5 Estimated Markov probabilities of staying in the same state for the G7 countries 49
3.6 Correlation statistics between probability of high variance state 50
3.7 Concordance statistics between probability of high variance state 50
4.1 Summary statistics: monthly excess return 62
4.2 Summary statistics: monthly industrial production growth 63
4.3 Parameter estimates: bivariate Markov switching heteroscedasticity model of output and equity 64
4.4 Diagnostics using standardized residuals from excess return equation 66

4.5 Correlation between probabilities of expansion states 74

4.6 Concordance statistics, probabilities of expansion states or low volatility states 75

4.7 Concordance statistics, expected probabilities in three months for expansion states 75

5.1 Summary statistics 87

5.2 Unit root test 90

5.3 Parameter estimates: Markov switching heteroscedasticity model of inflation rate 94

5.4 Residual diagnostics and model adequacy tests 100

5.5 Regime classification measure (RCM) 101

5.6 Vuong statistics for non-nested model selection: Markov switching model against two different GARCH models 101

6.1 Summary statistics 121

6.2 Permanent and transitory components of equity return: Markov switching heteroscedasticity framework 122

6.3 Residual diagnostics and model adequacy tests 123

6.4 Correlations: return, permanent and temporary components 123

7.1 Parameter estimates for the Markov switching coincidence index model 132

7.2 Diagnostic statistics for the residuals of measurement equations 133

7.3 Correlation between the coincident financial indicator and its components 133

7.4 Granger causality tests 134

Chapter 1

INTRODUCTION

1. Introduction

Markov chains have been growing more useful as a method for capturing the stochastic nature of many economic and financial variables. While hidden Markov processes have been widely employed for some time in engineering applications such as speech recognition, its effectiveness is now recognized in areas of social science research as well. The chief aim of this chapter is to summarize the basic properties of Markov chains and illustrate these properties with the help of useful examples. We also discuss the algorithmic structures of the processes used to estimate such models. Wherever possible we draw the reader's attention to the features of hidden Markov models (HMM) that distinguish them from similar modeling approaches, for example, the regime switching models familiar to most financial economists. The contents of later chapters of this book also require some understanding of state space methodology (SSM) and the related filtering techniques. To help readers gain a basic understanding of SSM, we include a section on state space methods and the application of Kalman filters.

2. Markov Chains

To model uncertainty, we usually adopt probability distributions to quantitatively describe the set of possible outcomes. In doing so, it is important to base the specification of these distributions on an understanding of the processes. A stochastic process is a collection of random variables indexed to time, t and the state, x. For example, we can write $\{x, t \geq 0\}$, $t \in T$. When T is finite we refer it to as *a countable stochastic process*. The indices can assume discrete or continuous values. Various stochastic processes such as random walk, Markov chains, Wiener processes, stochastic differential equations are applied to different applications. In this section we present only a brief introduction of Markov chains. [1]

The Markov chain, a stochastic process originally proposed by the Russian mathematician Markov in 1907, has been extensively applied to problems in social science, economics, finance, computer science, computer-generated music, and many other fields.

Consider a communications system that transmits the digits 0 and 1. Each digit transmitted must pass through several stages, and at each stage there is a probability p that the digit will leave unchanged. Where X_n denote the digit entering the $n-$th stage, $\{X_n, n = 0, 1, \cdots\}$ is a two-state Markov chain with the following transition probability matrix:

$$P = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}. \qquad (1.1)$$

If there are three states in this Markov chain, then the transition probability matrix will have the following form:

$$P = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{bmatrix}. \qquad (1.2)$$

In this matrix we note that $\Pr(X_n = j | X_{n-1} = i) = p_{ij} \geq 0$, $i = 1, 2, 3$; $j = 1, 2, 3$; $\sum_{j=1}^{3} p_{ij} = 1$. The properties of the Markov chain are then defined by the mathematical properties of the probability matrix.

Take the case of a new airline about to commence operation in a deregulated market that formerly protected the national airline. The new entrant will offer various inducements to attract passengers from the national airline, and in fact it has been estimated that (1/6) of the clients of the national company will remain loyal and refrain from switching in a given month. The clients of the new company have a (2/3) probability of remaining loyal. Our aim (given the number of passengers in the beginning) is to find the expected number of passengers flying the two airlines after a month, after two months, and after a long time i.e. in the long run. Tapeiro (1998) provides several other interesting illustrations.

Define state 0: A customer flies with the national airline, and state 1: A customer flies with the new airline.

The transition probability matrix is given by

$$P = \begin{bmatrix} 1/6 & 1/3 \\ 5/6 & 2/3 \end{bmatrix} \equiv \begin{bmatrix} p_{00} & p_{10} \\ p_{01} & p_{11} \end{bmatrix}. \qquad (1.3)$$

Let $N_0(0)$ be the number of clients in the national airline at the beginning, and let $N_1(0)$ be the number of clients in the new airline at the beginning. We assume $N_0(0) = 600$ and $N_1(0) = 0$, i.e. the number of clients at month 0 in the new airline is 0.

After the first month, the distribution of clients among the companies is given by:

$$N_0(1) = p_{00}N_0(0) + p_{10}N_1(0), \tag{1.4}$$

$$N_1(1) = p_{01}N_0(0) + p_{11}N_1(0). \tag{1.5}$$

In numbers this becomes $N_0(1) = 100$, and $N_1(1) = 500$. In matrix notation this can be written as:

$$\begin{bmatrix} N_0(1) \\ N_1(1) \end{bmatrix} = \begin{bmatrix} p_{00} & p_{10} \\ p_{01} & p_{11} \end{bmatrix} \begin{bmatrix} N_0(0) \\ N_1(0) \end{bmatrix}. \tag{1.6}$$

Thus, when we consider two consecutive months, the matrix notation becomes,

$$\begin{bmatrix} N_0(t+1) \\ N_1(t+1) \end{bmatrix} = \begin{bmatrix} p_{00} & p_{10} \\ p_{01} & p_{11} \end{bmatrix} \begin{bmatrix} N_0(t) \\ N_1(t) \end{bmatrix}, \tag{1.7}$$

or

$$N(t+1) = PN(t). \tag{1.8}$$

For the second month we have $N(2) = PN(1)$, which gives us (in term of numbers) $N_0(2) = 183$ and $N_1(2) = 417$.

The matrix P^T denotes the transition probability from one state to another in T steps. If the number of steps is large, the transition probabilities are then called ergodic transition probabilities and are given by the equilibrium probabilities (assuming they exist):

$$\pi_{ij} = \lim_{t\to\infty} p_{ij}(t), \tag{1.9}$$

or

$$\pi = \lim_{t\to\infty} P^t. \tag{1.10}$$

An application of the Chapman-Kolmogorov matrix multiplication formula yields:

$$\pi = \lim_{t\to\infty} P^t = \lim_{t\to\infty} PP^{t-1} = P \lim_{t\to\infty} P^{t-1} = P\pi. \tag{1.11}$$

This provides a system of linear equations that can be used to calculate the ergodic probabilities along with the fact that $\sum_{i=1}^{n} \pi_i = 1, \pi_i \geq 0$.

For our airline problem this implies:

$$\pi_0 p_{00} + \pi_1 p_{10} = \pi_0, \tag{1.12}$$

$$\pi_0 p_{01} + \pi_1 p_{11} = \pi_1, \tag{1.13}$$

$$\pi_0 + \pi_1 = 1, \tag{1.14}$$

$$\pi_0 \geq 0, \pi_1 \geq 0. \tag{1.15}$$

Equations (1.12) to (1.15) can be solved by any of several well-known methods, and the solution provides the answer to our question on distribution of passengers in the long run.

The implications of these equations are very important in practice, as they reveal the state the probabilities towards which a process will incline. Thus, if we compare two approaches that lead to two different Markov chains, we can study the long run effects of these methods.

We next consider a more efficient approach to solving a system of equations such as that in (1.12) to (1.15). From equation (1.11) we obtain:

$$(I_M - P)\pi = 0_M \tag{1.16}$$

where I_M is the identity matrix of order M (i.e. the order of the probability transition matrix P) and 0_M is the $M \times 1$ vector of zeros. The condition where the steady state probabilities sum to one is captured by:

$$i'_M \pi = 1 \tag{1.17}$$

where $i_M = [1, 1, \cdots, 1]'$.

Combining equations (1.16) and (1.17) we get:

$$\begin{bmatrix} I_M - P \\ i'_M \end{bmatrix} \pi = \begin{bmatrix} 0_M \\ 1 \end{bmatrix}, \tag{1.18}$$

or

$$A\pi = \begin{bmatrix} 0_M \\ 1 \end{bmatrix}. \tag{1.19}$$

And by multiplying both sides of (1.19) by $(A'A)^{-1}A'$, we get:

$$\pi = (A'A)^{-1}A' \begin{bmatrix} 0_M \\ 1 \end{bmatrix}. \tag{1.20}$$

The steady state probabilities may thus be obtained from the last column of the product matrix, $(A'A)^{-1}A'$.

3. Passage Time

Oftentimes we must ascertain the time required to attain a particular state. In a wealth process, for example, we may want to determine how long it will take to reach a bankrupt state, i.e. the state without any wealth. Assume that we are in state i and let $f_{ij}(n)$ be the probability of a first transition from state i to state j in n steps. This is the probability of having not gone through the $j-$th state in prior transitions. For a transition in one step, the transition matrix gives this probability. For a transition in two steps it equals the probability in two steps conditional on not having transited in one step. This implies two step transition probability less the one step transition probability times the probability that if it reaches such a state, it does not stay there. This can be represented by:

$$f_{ij}(1) = p_{ij}(1) = p_{ij}, \tag{1.21}$$

$$f_{ij}(2) = p_{ij}(2) - f_{ij}(1)p_{ij}, \tag{1.22}$$

and in general,

$$\begin{aligned} f_{ij}(n) &= p_{ij}(n) - [f_{ij}(1)p_{ij}(n-1) \\ &\quad + f_{ij}(2)p_{ij}(n-2) + \cdots + f_{ij}(n-1)p_{ij}] \end{aligned} \tag{1.23}$$

When the states i and j communicate (i.e. it is possible to switch from state i to state j in a finite number of transitions) we can compute the expectation of this passage time. This expectation is defined by:

$$\mu_{ij} = \sum_{n=0}^{\infty} n f_{ij}(n). \tag{1.24}$$

With further analysis (see Tapeiro, 1998) we find that the mean first passage time can be obtained by solving the set of equations given by:

$$\mu_{ij} = 1 + \sum_{k \neq j} p_{ik}\mu_{kj}. \tag{1.25}$$

Note here that if $k = j$, then $\mu_{kj} = \mu_{kk} = 0$.

Now consider the example of hedge fund market share. The current market position of hedge fund and its two main competing funds are 12%, 40% and 48%. Clients switch from one fund to another based on industry data, for example, the underperformance of a fund, and so on. The switching fund matrix is estimated as:

$$P = \begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.1 & 0.5 & 0.4 \\ 0.1 & 0.3 & 0.6 \end{bmatrix}. \tag{1.26}$$

Our aim is to find the mean first passage time for clients of funds 2 and 3 to fund 1.

This is given by the system of equations from (1.25):

$$\mu_{21} = 1 + \mu_{21} \times 0.5 + \mu_{31} \times 0.4, \tag{1.27}$$

$$\mu_{31} = 1 + \mu_{21} \times 0.4 + \mu_{31} \times 0.6. \tag{1.28}$$

Solving these we get $\mu_{21} = \mu_{31} = 10$.

Other interesting applications for the first passage time would be to calculate the time to bankruptcy, the first time cash reaches a given level, and so on. Continuing with the same example of hedge funds, it would also be interesting to ascertain the long-term market share of each of the three funds.

4. Markov Chains and the Term Structure of Interest Rates

In this section we briefly discuss an approach suggested by Das (1996), namely, the use of Markov chain techniques to model the term structure of interest rates and the pricing of interest-rate-sensitive securities. Two fundamental approaches are used for the modeling of term structures for interest rates: a) equilibrium approach and b) no-arbitrage approach. The equilibrium approach is based on a representative agent scenario that requires assumptions about the production and consumption behavior. The process for the interest rate is derived from this setting, and the security prices dependent on interest rates are obtained as solutions to a fundamental partial differential equation. The no-arbitrage approach, on the other hand, starts with the assumption of an interest rate process calibrated against a set of observable prices, and its evolution is designed to reflect absence of arbitrage possibilities over time.

More recently, the kernel approach has been adopted as an alternative for dealing with the issues encountered in modeling the term structure of interest rates. Pricing kernels or state price densities imply that asset prices follow martingales. The functional specification of the kernel depends on the preferences of the representative agent, and it also ensures the existence of no-arbitrage. In a sense, this approach combines the two methods mentioned earlier. This section discusses the implementation of the kernel approach by exploiting the properties of Markov chains.

Markov chains are characterized by the transition probability matrix and the choice of the state space. This probability matrix offers great flexibility in the modeling of the various stochastic behaviors of the term structure of interest rates. Das (1996) also shows that it is relatively simple to estimate this transition matrix. Thus, the approach offers a theoretically solid basis with empirical ease of implementation. One point to note here is that the Markov chain operates in the historical probability measure and not in the risk-neutral probability measure. For this reason, the risk premia demanded by the investor has to be built into the model.

Cox, Ingersoll, and Ross (1985) (CIR) can claim credit for the widespread use of the term structure of the interest rate model both among practitioners and academic researchers. The driving state variable in this setup is the short rate process, hence we refer to it as a *one-factor model.* Our discussion on the chain below will also rely on a one-factor situation, although this can be extended to more than one factor. Even if the state variable is not strictly Markov, it can be converted to a Markov process by expansion of states. Examples of such state expansion can be found in Bhar and Chiarella (1997) in the framework of Heath, Jarrow, and Morton (1992).

We assume that the state variable, the short rate $r(t)$ at time t can take on a finite number of values, $[r_1, r_2, \cdots, r_N]$, and we observe its evolution over the time scale $t = 1, 2, \cdots, T$, with interval Δt. Within this interval the short rate can move from state i to state j with probability $p_{ij}(t)$. Due to the assumption of Markov property (i.e. the history does not matter) we can write the transition probabilities as follows:

$$Pr[r_j(t+1)|r_i(t)] = p_{i,j}(t). \tag{1.29}$$

If we assume that the transition probabilities remain unchanged over time, then we get the time-homogeneous version of the transition probability matrix. The overall transition probability matrix appears thus:

$$P = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{N1} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1N} & p_{2N} & \cdots & p_{NN} \end{bmatrix}, \tag{1.30}$$

with $\sum_{j=1}^{N} p_{ij} = 1$ for all i.

To account for the time-varying risk premium we stipulate that the required rate of return over the small interval Δt shall be $(r_i \Delta t + \lambda(t))$. Here $\lambda(t)$ captures the additional return demanded by the investor in the historical probability measure. The one period discount rate for a given state is:

$$\begin{aligned} d_i &= exp(-r_i \Delta t - \lambda(t)) \\ &= exp(-r_i \Delta t) exp(-\lambda(t)) \\ &= exp(-r_i \Delta t)\pi(t). \end{aligned} \tag{1.31}$$

The variable $\pi(t)$ here is the time-dependent risk premium, and we can define

$$\pi = [\pi(1), \pi(2), \cdots, \pi(T)] \tag{1.32}$$

as the complete set of premiums applicable for the entire observation period.

For later assistance in defining the pricing operator, let us now define a new matrix combining the probabilities and the risk premia as:

$$Q = \begin{bmatrix} q_{11} & q_{21} & \cdots & q_{N1} \\ \vdots & \vdots & \ddots & \vdots \\ q_{1N} & q_{2N} & \cdots & q_{NN} \end{bmatrix}, \tag{1.33}$$

where $q_{ij} = p_{ij}\pi(t)$. The usual constraints need not be satisfied since the elements of the Q matrix are clearly not probabilities. We can operate in the historical measure by incorporating the applicable risk premium in this model, hence the measure need not be changed in order to move to a risk neutral measure. As the definitions show that the elements of the Q matrix are clearly positive, we can apply the Markov model to the observed interest rates.

Let us define the cash flow from a security at any time t as $X(t) = [X_1(t), X_2(t), \cdots, X_N(t)]'$, and the discount factor for each of the states as $R = exp(-r_i \Delta t)$, $i = 1, 2, \cdots, N$. With these definitions we can express the discounted state price matrix as:

$$\tilde{M}(t) = Q(t) \otimes R, \quad \forall\ t, \tag{1.34}$$

where the operator $\otimes$ denotes a Kronecker product, and not in sense of matrix multiplication. The elements of the discount state price matrix are of the form, $m_{ij} = q_{ij} exp(-r_i \Delta t)$. In this situation, the no-arbitrage pricing of securities implies that:

$$X(t) = \tilde{M}(t+1) X(t+1), \tag{1.35}$$

or, in other words the current price of the security in term of its cash flow at time T is given by:

$$X(0) = \left(\prod_{t=1}^{T} \tilde{M}(t) \right) X(T) \equiv h(T)X(T). \tag{1.36}$$

In this equation $h(T)$ is the pricing kernel that transforms T period cash flow to today's price. It also ensures no-arbitrage. In other words, the kernel-based pricing of securities and the Markov chain representation of interest rate movement are captured by equation (1.36). We can also write this in term a discount bond price of maturity $\tau = T\Delta t$, as:

$$B(0, \tau) = h(T)i, \tag{1.37}$$

where i is an N-dimensional vector of ones representing the dollar payoff at time T. Similarly, for coupon bonds of periodic coupon amount c, the price today is:[2]

$$B(0, \tau) = h(T)i + c \sum_{t=1}^{T} h(t)i. \tag{1.38}$$

Next we will illustrate this methodology with the help of a simple example. There are two main inputs to the process, the transitions probability matrix, P, and the vector of risk premia, π. In practice the matrix P can be approximated from historical data with frequency histogram. Das (1996) also indicates another way of populating the matrix that relies on prior understanding of a stochastic process with known density function driving the transitions. With a known P matrix, the risk premia vector can be estimated by calibrating the model with observable security prices covering the entire time scale.

For the purposes of illustration, let $\Delta t = 1$ year, let the short rate state space equal $[0.05, 0.10, 0.15]$, and let the P matrix equal:

$$P = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.2 & 0.7 \end{bmatrix}. \tag{1.39}$$

We also observe the three discount bond prices of maturities 1, 2 and 3 years as, $[0.900, 0.805, 0.712]$, respectively. The discount vector covering the state space for the next period can be expressed as:

$$R = \begin{bmatrix} e^{-0.05} \\ e^{-0.10} \\ e^{-0.15} \end{bmatrix} = \begin{bmatrix} 0.951229 \\ 0.904837 \\ 0.860708 \end{bmatrix}. \tag{1.40}$$

Knowing the current one period interest rate as 0.10, we can set up the equation for determining $\pi(1)$ as:

$$0.900 = \pi(1)(P \otimes R) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{1.41}$$

In this case, we are interested in the second equation. Thus:

$$0.900 = \pi(1) \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.570737 & 0.285368 & 0.095122 \\ 0.090483 & 0.723869 & 0.090483 \\ 0.086070 & 0.172141 & 0.602495 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{1.42}$$

From the middle equation we get $\pi(1) = 0.994654$. Knowing $\pi(1)$, we can ascertain $\pi(2)$ from the second equation as

$$0.805 = \pi(2)\pi(1)(P \otimes R)^2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \tag{1.43}$$

and thereby obtain:

$$0.805 = \pi(2)\pi(1) \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.359749 & 0.385814 & 0.137522 \\ 0.124928 & 0.565384 & 0.128621 \\ 0.116557 & 0.252884 & 0.386764 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{1.44}$$

This solves $\pi(2)$ as 0.988267, and we can use the same approach to solve $\pi(3)$ as 0.977335.

Once these risk premia are computed, the model is completely specified and can be applied to determine security prices that depend on the interest rate movements over the applicable time scale. Coupon bonds as well as European style options on bonds could be priced from the model just developed. Additional illustrations are available in Das (1996). In a more recent application, Lyn, Allen, and Morkel-Kingsbury (2002) adopt the lattice Markov chain model of Pliska (1997, Chapter 6) to determine the bond prices subject to default risk. They generalize it by inclusion of a hidden Markov chain, describing the economic condition that drives the evolution of the interest rate as well as the corporate credit rating. We encourage interested readers to examine Pliska's treatment of the two probability measures, the historical and risk-neutral, in comparison to the treatment described above.

5. State Space Methods and Kalman Filter

The state space model (SSM) was originally intended for aerospace-related research, but it has found immense application in economics and finance. This approach is used to analyze typically dynamic time series models that involve unobserved components. A great many potential applications in econometrics involve unobserved variables such as permanent income, expectations, the ex ante real interest rate, and so on.

The SSM in its basic form retains a VAR (1) (vector autoregressive) structure for the state equation

$$y_t = \Gamma y_{t-1} + w_t, \tag{1.45}$$

where the state equation determines the rule for generation of the states $y_{t,i}$ from the past states $y_{t-1,j}$, $j = 1, 2, \cdots, p$ for $i = 1, 2, \cdots, p$ and time points $t = 1, 2, \cdots, n$. For completeness we assume that w_t are $p \times 1$ Gaussian white noise vectors with covariance Q. The state process is assumed to have started with the initial value given by the vector, y_0, taken from normally distributed variables with mean vector μ_0 and the $p \times p$ covariance matrix, Σ_0 .

Although the state vector itself is not observed, some transformation of this vector is observed, albeit in a linearly added noisy environment. Thus, we can express the measurement equation as:

$$z_t = A_t y_t + v_t. \tag{1.46}$$

In this sense, the $q \times 1$ vector z_t is observed through the $q \times p$ measurement matrix A_t together with the $q \times 1$ Gaussian white noise v_t, with the covariance matrix, R. We also assume that the two noise sources in the state and the measurement equations are uncorrelated. In the original space tracking area, the state equation defines the motion equations for the space position of a spacecraft with location y_t and z_t reflects information that can be observed from a tracking device such as velocity and height.

The next step is to make use of the Gaussian assumptions and produce estimates of the underlying unobserved state vector given the measurements up to a particular point in time. In other words, we would like to ascertain $E[y_t|(z_{t-1}, z_{t-2}, \cdots, z_1)]$ and the covariance matrix $P_{t|t-1} = E[(y_t - y_{t|t-1})(y_t - y_{t|t-1})']$. This is achieved with the use of a Kalman filter and the basic system of equations described below.

Given the initial conditions $y_{0|0} = \mu_0$, and $P_{0|0} = \Sigma_0$ for observations made at time $1, 2, 3, \cdots, T$,

$$y_{t|t-1} = \Gamma y_{t-1|t-1}, \tag{1.47}$$

$$P_{t|t-1} = \Gamma P_{t-1|t-1}\Gamma' + Q, \tag{1.48}$$

$$y_{t|t} = y_{t|t-1} + K_t(z_t - A_t y_{t|t-1}). \tag{1.49}$$

where the Kalman gain matrix is

$$K_t = P_{t|t-1}A_t'(A_t P_{t|t-1}A' + R)^{-1} \tag{1.50}$$

and the covariance matrix $P_{t|t}$ after the t-th measurement has been made is

$$P_{t|t} = (I - K_t A_t)P_{t|t-1}. \tag{1.51}$$

Equation (1.47) forecasts the state vector for the next period given the current state vector. The use of this one-step-ahead forecast of the state vector lets us define the innovation vector as

$$v_t = z_t - A_t y_{t|t-1}, \tag{1.52}$$

and its covariance as

$$\Sigma_t = A_t P_{t|t-1}A_t' + R. \tag{1.53}$$

Since all the observations are already available in most applications in finance and economics, we can improve the estimates of the state vector based on the whole sample. This is referred to as a *Kalman smoother*, and it starts with initial conditions at the last measurement points, i.e. $y_{T|T}$ and $P_{T|T}$. The following set of equations describes the smoother algorithm:

$$y_{t-1|T} = y_{t-1|t-1} + J_{t-1}(y_{t|T} - y_{t|t-1}), \tag{1.54}$$

$$P_{t-1|T} = P_{t-1|t-1} + J_{t-1}(P_{t|T} - P_{t|t-1})J_{t-1}', \tag{1.55}$$

$$J_{t-1} = P_{t-1|t-1}'\Gamma'(P_{t|t-1})^{-1}. \tag{1.56}$$

On the basis of the above, we clearly need to store the quantities $y_{t|t}$ and $P_{t|t}$ generated during the filter pass in order to implement the smoothing algorithm.

The description of the above filtering and the smoothing algorithms assumes that these parameters are known. In actuality we need to determine these parameters, and this can be achieved by maximizing the innovation form of the likelihood function. The one-step-ahead innovation and its covariance matrix are defined by equations (1.52) and (1.53).

t		$t+1$		$t+2$
$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$
$\widehat{y}_{t\|t}$	System	$\widehat{y}_{t+1\|t}$	Updating	$\widehat{y}_{t+t\|t+1}$
$\widehat{P}_{t\|t}$	Dynamics	$\widehat{P}_{t+1\|t}$	Equations	$\widehat{P}_{t+1\|t+1}$
		$\Downarrow$		
		$\widehat{z}_{t+1\|t}$		
		Prediction Error		
		$v_{t+1} = z_{t+1} - \widehat{z}_{t+1\|t}$		
		Build Prediction Error Form of Likelihood Function		
		Using v_{t+1}, $\widehat{P}_{t+1\|t}$ and Accumulate Through all t		$\Longrightarrow$

Figure 1.1. Filtering Algorithm

As these are assumed to be independent and conditionally Gaussian, the log likelihood function (without the constant term) is expressed as:

$$\log(L) = -\frac{1}{2}\sum_{t=1}^{T}\log|\Sigma_t(\Theta)| - \frac{1}{2}\sum_{t=1}^{T} v_t'(\Theta)\Sigma_t^{-1}(\Theta)v_t(\Theta). \tag{1.57}$$

The term Θ (the vector of unknown parameters) in this expression is specifically used to emphasize the dependence of the log likelihood function on the parameters of the model. Once the function is maximized with respect to the model parameters, those parameters can be used to start the next step of smoothing. Several numerical approaches can be applied to maximize the log likelihood function, and some of these will be discussed in later chapters.[3]

At this point, it would be useful to review the entire process with the help of a diagram that depicts the adaptive filtering algorithm (Figure 1.1).

6. Hidden Markov Models and Hidden Markov Experts

This section will begin with a simple example to help readers understand the basic concepts in Hidden Markov Models (HMMs) and how they relate to the state space models (SSMs). Suppose we observe a series of counts of a specific medical episode, for example, epileptic seizures, in one patient on successive days, or the counts of home burglaries in a particular district. The simplest model we can use to generate these counts would be a homogeneous Poisson process. The counts so generated would be independently and identically distributed Poisson random variables, hence the variance would be equal to the mean. In practice, however, such counts usually display a dispersion that deviates from that suggested above, and they may also display a serial dependence.[4]

An alternative way to describe the data would be to generate the counts by two Poisson processes with different means and then choose one of the means by another random process known as an emission process. Let us take an example where the first mean is selected with a particular probability (p) and the second mean is selected with the probability $(1 - p)$. If the emission process were a Markov chain, then the observed counts process would display serial dependence in addition to over-dispersion. Such probabilistic functions of Markov chains have been in wide use in the engineering literature.[5]

Thus, the basic idea here is to determine the observed sequence by the underlying unobservable process, i.e. the state sequence of the HMM with an emission probability. The word *hidden* emphasizes that these states cannot be directly estimated from the observed data. The hidden process is also Markov in nature, as the next state will depend solely on the current state and the transition probability between the states. Both the states and the observed process can be either discrete or continuous. For time series analysis, we use discrete states relating to different regimes and continuous observations referring to the time series. In an HMM, an underlying and unobservable sequence of states follows a Markov chain with finite state space, and the probability distribution of the observation at any time is dependent only on the current state of the Markov chain.

We can enrich our understanding of the basic issues in HMM in works such as Weigend and Shi (1998) and Weigend, Mangeas, and Srivastava (1995). These authors suggest that the basic linear time series models rest on two assumptions: a) the time series has weak stationarity, and b) complete characterization is possible within a finite embedding space. Many financial and economic time series are non-stationary, and display time varying means and/or variances. ARCH (Autoregressive Conditional Heteroscedasticity) processes address some of these issues, provided that the variance of the time series depends on past variances. One way to deal with the non-stationarity has been to partition the time series in such a way that it traverses different regimes or regions. Different architectures have also been proposed to decompose a global model into modular local models. The main point here is how to split the data space. If the different regions are independent, then simply reshuffling the data does not lead to a different final model.

The concept of Hidden Markov Experts (HMEs) explicitly deals with the time dependency between adjacent regions. A context-dependent HMM is an important speech recognition application that can also handle time-dependent regime-switching, as demonstrated by Hamilton (1990).[6] In Hamilton's work, the regions or states can be estimated from the current observation. In an HME, the states are hidden from the observation and depend on the entire history of observations. While Hamilton uses all linear predictors, an HME allows non-linear predictors such as neural networks as well.

An HMM may be viewed as an alternative to the Kalman filter in modeling time structure. Some authors liken the Kalman filter to a factor analysis over time, whereas an HMM is a mixture of densities over time. A Kalman filter normally represents the linear evolution of states and linear measurements, whereas an HMM can represent highly non-linear evolution and measurements, and only one state at any time. The Kalman filter may also be viewed as an *on-line* algorithm as it utilizes recursions. In the case of the HMM, however, the linearity of the model makes this impossible. To estimate the unknown parameters in the Kalman filter, we can apply the prediction error form of the likelihood function maximization, as explained in the previous section. In the HMM we would have to apply the Viterbi algorithm to estimate the most likely state sequence and the Baum-Welch forward-backward algorithm to estimate the parameters following Expectation Maximization (EM). We will explore these topics in the next section.

Before concluding this section, we will re-state the HMM approach and cite some related works as well. An HMM is a parameterized stochastic probability model used to analyze time series. It consists of two interrelated processes: a) a finite state Markov chain that cannot be observed, and b) an emission model associated with each state. The Markov chain is characterized by its transition probability matrix, and the probability densities given by the emission model can have two components: 1) either a parametric or non-parametric specification, and 2) dependence on either additional input (a conditional HME) or an unconditional HME. As Weigend and Shi (1998) describe, a Markov chain generates a sequence of discrete states as a path, and the emission model generates the probability density for each time step. A particular observation is then generated from this probability density at each time step.

Based upon transition among states, Hamilton (1990) introduced regime-switching models in economics. Hamilton and Susmel (1994) extended the approach to capture time-varying conditional variance where the parameters of the variance process are taken from one of two states that evolve as a Markov chain. Similarly, Gray (1996) proposed an approach where a GARCH model is nested in a regime-switching model. Smith (2002) further enhanced this area for modeling stochastic volatility in interest rates by incorporating a regime-dependent variance parameter. All these attempts are restricted to the first two moments of the distribution, rather than the complete density function. Weigend and Shi (1998) demonstrate the application of an HMM to generate a complete density function using both linear and non-linear HMEs.

7. HMM Estimation Algorithm

In this section we first define the mathematical structure of the HMM and then go on to describe the estimation algorithms for the different components.[7] In some processes, an observed sequence is probabilistically related to an underlying Markov process. When this is so, the number of observable states may differ from the number of hidden states. To handle such cases, an HMM must comprise two sets of states and three sets of probabilities. These include the following:

- Hidden states: The states of a system that can be described by a Markov process,
- Observable states: The states of the process that are visible, i.e. measurable,
- Pi-Vector: Contains the probability that the model is in one of the hidden states at the initial time,
- State transition matrix: Contains the probability that a hidden state will evolve to another state, given the previous state,
- Emission probability matrix: Contains the probability that a particular measurable state can be observed, provided that the model is in one of the hidden states.

Thus, a hidden Markov model is a standard Markov process augmented by a set of measurable states and several probabilistic relations between those states and the hidden states. In developing the algorithms for the model, we write the joint probability over hidden (y_t) and observed (x_t) states, and then use the Markov property to simplify it. This property allows us to assume that all information about the history of the states is summarized by the value of the state at the previous time step. The system does not have a long memory, and the observations (x_t) do not depend on the previous states or the observations given the states at that time (y_t). The total probability is,

$$p(y^T, x^T) = p(y_1) \prod_{t=2}^{T} p(y_t|y_{t-1}) \prod_{t=2}^{T} p(x_t|y_t). \tag{1.58}$$

The superscript is used to denote $y^T = (y_1, y_2, \cdots, y_T)$. As a general term, this expression is similar to the *would be* expression used in a Kalman filter. A particular HMM can be characterized by the following three matrices,

$$A_{i,j} = p(y_t = i|y_{t-1} = j), \tag{1.59}$$

$$B_{t,i} = p(x_t|y_t = i), \tag{1.60}$$

$$\pi_{t,i} = p(y_t = i). \tag{1.61}$$

In the literature, an HMM is thus referred to as a set, $\lambda = \{A, B, \pi\}$, i.e. three components that uniquely define a particular HMM for a problem. In what follows we also need the following definitions:

$$\gamma_{t,i} = p(y_t = i|x^T), \tag{1.62}$$

$$\Omega_{t,ij} = p(y_t = i, y_{t-1} = j|x^T), \tag{1.63}$$

$$\alpha_{t,i} = p(y_t = i|x^t), \tag{1.64}$$

$$\kappa_t = p(x_t|x^{t-1}), \tag{1.65}$$

$$\beta_{t,i} = \frac{p(x_{t+1}, \cdots, x_T|y_t = i)}{p(x_{t+1}, \cdots, x_T|x^t)}. \tag{1.66}$$

While the definitions above may still appear somewhat out of context; their usefulness becomes clearer as we go on to develop the recursive algorithm. The expressions help to reduce the computational complexity of the procedure and facilitate the computer implementation. The label of the state appearing in the above expressions explicitly shows its own dependence on the state, and the use of the subscript t shows the dependence of the label on observation x_t. Additionally, and quite importantly, we also find from equation (1.59) that the elements of the transition matrix $A_{i,j}$ are independent of the observations and time invariant.

The two main problems we need to address are to estimate or identify a) the unknown parameters in model $\lambda = \{A, B, \pi\}$, and b) the sequence of states most likely to generate those observations for a given model and set of observations.

8. HMM Parameter Estimation

In order to estimate the model parameters, we start by writing the complete likelihood function of the complete dataset over N iterations. This is given by,

$$
\begin{aligned}
Q &= \sum_{n=1}^{N} \int dy^T p(y^T|x^{T,n}) \log \left[p(y_1) \prod_{t=2}^{T} p(y_t|y_{t-1}) \prod_{t=1}^{T} p(x_t^n|y_t) \right] \\
&= \sum_n \int dy_1 p(y_1|x^{T,n}) \log p(y_1) \\
&\qquad + \sum_n \sum_{t=2}^{T} \int dy_t dy_{t-1} p(y_t, y_{t-1}|x_t^n) \log p(y_t|y_{t-1}) \\
&\qquad + \sum_n \sum_{t=1}^{T} \int dy_t p(y_t|x_t^n) \log p(x_t^n|y_t) \\
&= \sum_n \sum_t \gamma_{1,i}^n \log \pi_{1,i} + \sum_n \sum_{t=2}^{T} \sum_{i,j} \Omega_{t,ij}^n \log A_{i,j} + \sum_n \sum_{t=1}^{T} \sum_i \gamma_{t,i}^n \log B_{t,i}.
\end{aligned} \tag{1.67}
$$

The above likelihood function cannot be directly maximized since the hidden states are known. The solution to this problem is due to Baum et al (1970) and is referred to as the Baum-Welch algorithm. It would be seen that it turns out to be equivalent to the EM (Expectation Maximization) algorithm of Dempster, Laird, and Rubin (1977). In the above objective function we also need to include the constraints due to the probability terms in the parameter set. In other words, the following two constraints are needed:

$$\sum_i \pi_{1,i} = 1, \tag{1.68}$$

$$\sum_i A_{i,j} = 1, \quad \forall j. \tag{1.69}$$

The augmented objective function using Lagrange multiplier becomes,

$$L = Q - \lambda_\pi (\sum_i \pi_{1,i} - 1) - \sum_j \lambda_j (\sum_i A_{i,j} - 1). \tag{1.70}$$

In the maximization step, we need to take the derivatives of the above augmented objective function and equate to zero. Thus,

$$\frac{\partial L}{\partial \pi_{1,i}} = \sum_n \frac{\gamma_{1,i}^n}{\pi_{1,i}} - \lambda_\pi = 0 \ \Rightarrow \ \pi_{1,i}^{new} = \frac{\lambda_\pi}{N}\sum_{n=1}^{N}\gamma_{1,i}^n, \tag{1.71}$$

where the Lagrange coefficient is determined from the constraint relation. Thus,

$$\pi_{1,i}^{new} = \frac{1}{N}\sum_{n=1}^{N}\gamma_{1,i}^n. \tag{1.72}$$

This equation implies the expected number of times the system is found in state i at the beginning time. Similarly, for the matrix A, we find,

$$\frac{\partial L}{\partial A_{i,j}} = \sum_n \sum_{t=2}^{T} \frac{\Omega_{t,ij}^n}{A_{i,j}} - \lambda_j = 0, \tag{1.73}$$

$$\begin{aligned} A_{i,j}^{new} &= \frac{\lambda_j}{N(T-1)}\sum_{n=1}^{N}\sum_{t=2}^{T}\Omega_{t,ij}^n \\ &= \frac{\sum_{n=1}^{N}\sum_{t=2}^{T}\Omega_{t,ij}^n}{\sum_{n=1}^{N}\sum_{t=2}^{T}\gamma_{t-1,j}^n}, \end{aligned} \tag{1.74}$$

where $\sum_i \Omega_{t,ij} = \gamma_{t-1,j}$. This represents the ratio of the expected number of transitions from state j to state i divided by the expected number of moves from state j.

We now focus on the emission probabilities and assume that it could be parameterized as,

$$p(x_t|y_t = i;\theta) = B_{t,i}(\theta), \tag{1.75}$$

and the parameter θ is updated as,

$$\theta^{new} = \arg\max_\theta \sum_{n=1}^{N}\sum_{t=1}^{T}\sum_i \gamma_{t,i}^n \log B_{t,i}(\theta). \tag{1.76}$$

Although, different structures could be associated with the output generation for various states, here we concentrate on discrete output only. The parameters are, therefore, the probability masses, i.e. for k states of the output, x_t,

$$B_{k,i} = p(x_t = k|y_t = i). \tag{1.77}$$

We also assume these are time independent and the constraint that must be satisfied is,

$$\sum_k B_{k,i} = 1, \quad \forall i. \tag{1.78}$$

Again with the help of Lagrange multiplier it can be shown that,

$$B^{new}_{k,i} = \frac{\sum_{n,t\ s.t.x^n_t=k} \gamma^n_{t,i}}{\sum_{n=1}^N \sum_{t=1}^T \gamma^n_{t,i}}. \tag{1.79}$$

This update of the B matrix indicates the expected number of times the system is in state i and we observe k, divided by the expected number times the system is in state i. The maximization or the M step of the algorithm is now completed.

We now discuss the strategy of efficiently computing the quantities, $\Omega_{t,ij}$, $\gamma_{t,i}$. This is done through intermediate computation of the quantities $\alpha_{t,i}$, κ_t, $\beta_{t,i}$ recursively.

$$\alpha_t = p(y_t|x^t) = \frac{\sum_{y_{t-1}} p(x_t|y_t)p(y_t|y_{t-1})p(y_{t-1}|x^{t-1})}{p(x_t|x^{t-1})}, \tag{1.80}$$

$$\alpha_{t,i} = \frac{\sum_j B_{t,i}A_{i,j}\alpha_{t-1,j}}{\kappa_t}. \tag{1.81}$$

The initialization of the forward variable α is achieved as follows:

$$\alpha_1 = p(y_1|x_1) = \frac{p(x_1|y_1)p(y_1)}{\sum_{y_1} p(x_1|y_1)p(y_1)}, \tag{1.82}$$

$$\alpha_{1,i} = \frac{B_{1,i}\pi_{1,i}}{\sum_j B_{1,j}\pi_{1,j}}. \tag{1.83}$$

For the κ variable, we note that,

$$\kappa_t = p(x_t|x^{t-1}) = \sum_{y_t}\sum_{y_{t-1}} p(x_t|y_t)p(y_t|y_{t-1})p(y_{t-1}|x^{t-1}), \tag{1.84}$$

$$\kappa_t = \sum_{i,j} B_{t,i}A_{i,j}\alpha_{t-1,j}, \tag{1.85}$$

and the initial value is given by,

$$\begin{aligned}\kappa_1 &= p(x_1)\\ &= \sum_{y_1} p(x_1|y_1)p(y_1)\\ &= \sum_i B_{1,i}\pi_{1,i}.\end{aligned} \tag{1.86}$$

Let us now focus on the backward recursions. We start with,

$$\begin{aligned}\beta_{t-1} &= \frac{p(x_t,\cdots,x_T|y_{t-1})}{p(x_t,\cdots,x_T|x^{t-1})}\\ &= \frac{\sum_{y_t} p(x_t,\cdots,x_T,y_t|y_{t-1})}{p(x_t|x^{t-1})p(x_{t+1},\cdots,x_T|x^t)}\\ &= \frac{\sum_{y_t} p(x_t,\cdots,x_T|y_t)p(y_t|y_{t-1})}{p(x_t|x^{t-1})p(x_{t+1},\cdots,x_T|x^t)}\\ &= \frac{\sum_{y_t} p(x_t|y_t)p(x_{t+1},\cdots,x_T|y_t)p(y_t|y_{t-1})}{p(x_t|x^{t-1})p(x_{t+1},\cdots,x_T|x^t)}. \end{aligned} \tag{1.87}$$

$$\tag{1.88}$$

This leads to

$$\beta_{t-1,j} = \frac{\sum_i B_{t,i}\beta_{t,i}A_{i,j}}{\kappa_t}. \tag{1.89}$$

We initialize this recursion as follows:

$$\begin{aligned}\beta_{T-1} &= \frac{p(x_T|y_{T-1})}{p(x_T|x^{T-1})}\\ &= \frac{\sum_{y_T} p(x_T|y_T)p(y_T|y_{T-1})}{\kappa_T}. \end{aligned} \tag{1.90}$$

Here again for initialization,

$$\beta_{T-1,j} = \frac{\sum_i B_{T,j}A_{i,j}}{\kappa_T}, \tag{1.91}$$

$$\beta_{T,j} = 1. \tag{1.92}$$

With the help of the quantities, α_t, κ_t and β_t, we would be able to complete the recursions of γ_t and Ω_t. Thus,

$$\begin{aligned}\gamma_t &= p(y_t|x_t)\\ &= \frac{p(y_t,x^T)}{p(x^T)}\\ &= \frac{p(x_{t+1},\cdots,x_T,y_t|x^t)p(x^t)}{p(x_{t+1},\cdots,x_T|x^t)p(x^t)}\\ &= \frac{p(y_t|x^t)p(x_{t+1},\cdots,x_T|y^t)}{p(x_{t+1},\cdots,x_T|x^t)}. \end{aligned} \tag{1.93}$$

This leads to

$$\gamma_{t,i} = \alpha_{t,i}\beta_{t,i}. \tag{1.94}$$

The final recursion we need to formalize is,

$$\begin{aligned}
\Omega_t &= p(y_t, y_{t-1}|x^T) \\
&= \frac{p(x_t, \cdots, x_T|y_t)p(y_t|y_{t-1})p(y_{t-1}|x^{t-1})}{p(x_t, \cdots, x_T|x^{t-1})} \\
&= \frac{p(x_t|y_t)p(x_{t+1}, \cdots, x_T|y_t)p(y_t|y_{t-1})p(y_{t-1}|x^{t-1})}{p(x_t|x^{t-1})p(x_{t+1}, \cdots, x_T|x^{t-1})}.
\end{aligned} \tag{1.95}$$

Thus the final relation in the E-step (Expectation step) is,

$$\Omega_{t,ij} = \frac{B_{t,i}\beta_{t,i}A_{i,j}\alpha_{t-1,j}}{\kappa_t}. \tag{1.96}$$

The essence of the EM-algorithm is to alternate between the M-step and E-step until the desired convergence is reached.[8]

9. HMM Most Probable State Sequence: Viterbi Algorithm

Next we address the question of how to infer the hidden states given the observations. The task is comparable to estimating the smoothed states in the Kalman filter application, where the mean state is estimated given all the observations. In discrete situations it would be more meaningful to estimate the most likely state at each time instant, t. In other words,

$$\widehat{y}_t^T = \arg\max_{y_t}\ p(y_t|x^T) = \arg\max_{i}\ \gamma_{t.i} \quad \forall t. \tag{1.97}$$

Note that this does not imply the single most likely state sequence. If all transitions are not permissible, the above sequence given by equation (1.97) may have no probability of occurring. At first glance, however, the task appears daunting. If there are M hidden states, then there are M^T sequences to consider. Fortunately, we can solve this dilemma by applying a Viterbi algorithm, a special formulation of dynamic programming utilizing the Markov structure. We are attempting to find,

$$\widehat{y}^T = \arg\max_{y^T}\ p(y^T|x^T)$$

$$= \arg\max_{y^T} \frac{p(y^T, x^T)}{p(x^T)}$$

$$= \arg\max_{y^T} \; p(y^T, x^T). \tag{1.98}$$

As before, the superscript indicates the whole sequence.

We now explore ways to compute this recursively. In this context, we define,

$$\delta_t(y_t) = \max_{y^{t-1}} \; p(y^t, x^t), \tag{1.99}$$

which implies,

$$\delta_{t,i} = \max_{y^{t-1}} \; p(y^{t-1}, y_t = i, x^t). \tag{1.100}$$

As we are clearly maximizing over the entire sequence y^{t-1}, we initialize with,

$$\delta_1 = p(y_1, x_1) = p(x_1|y_1)p(y_1). \tag{1.101}$$

In other words,

$$\delta_{1,i} = B_{1,i}\pi_{1,i}, \tag{1.102}$$

and

$$\max_{y^T} \; p(y^T, x^T) = \max_{y^T} \; \delta_T = \max_{i} \; \delta_{T,i}. \tag{1.103}$$

Now, the Markov structure of the model gives the recursion as follows:

$$\begin{aligned}
\delta_{t+1} &= \max_{y^t} \; p(y^{t+1}, x^{t+1}) \\
&= \max_{y^t} \; [p(x_{t+1}|y_{t+1})p(y_{t+1}|y_t)p(y^t, x^t)] \\
&= p(x_{t+1}|y_{t+1}) \max_{y^t} \; \{p(y_{t+1}|y_t) \max_{y^{t-1}} \; [p(y^t, x^t)]\} \\
&= p(x_{t+1}|y_{t+1}) \max_{y^t} \; [p(y_{t+1}|y_t)\delta_t].
\end{aligned} \tag{1.104}$$

The recursion thus turns out to be,

$$\delta_{t+1,i} = B_{t+1,i} \max_j [A_{i,j}\delta_{t,j}]. \tag{1.105}$$

We start the recursion by initializing with equation (1.102), then compute $\delta_2, \cdots, \delta_T$ with the above recursion, and finally use equation (1.103) to get the overall maximum. Regarding the computation, we should understand that since δ_t is generated by multiplying probability terms, its value decreases as the value of t rises. As a consequence, we may have to re-normalize δ_t at each iteration in order to avoid computing exceptions due to underflow. This step in no way compromises our objective of inferring the state sequence that maximizes the overall probability, and we are not interested in the actual value of that probability.

In order to extract the sequence that maximizes the probability, we need to store the relevant values. The following expressions achieve this task by storing the y_t values that maximize $[p(y_{t+1}|y_t)\delta_t(y_t)]$ for all y_{t+1}. Accordingly,

In order to extract the sequence that maximizes the probability we need to store the relevant values. The following expressions achieve this task by storing the y_t values that maximizes $[p(y_{t+1}|y_t)\delta_t(y_t)]$ for all y_{t+1}. Accordingly,

$$\Psi_{t+1}(y_{t+1}) = \arg\max_{y_t} \; [p(y_{t+1}|y_t)\delta_t(y_t)], \tag{1.106}$$

$$\Psi_{t+1,i} = \arg\max_{j} \; (A_{i,j}\delta_{t,j}), \quad t = 1, 2, \cdots, T-1. \tag{1.107}$$

We thus find the optimal state sequence that maximizes the probability $p(y^T|x^T)$.

10. HMM Illustrative Examples

In this section we describe two relatively recent research findings that make use of the algorithms discussed above. The first, from Manton, Muscatelli, Krishnamurthy, and Hurn (1998), is reported in an analysis of excess stock market returns in a number of OECD countries recorded at a monthly frequency from the period of 1960 to 1988.

In concordance with related works on excess stock market returns, these authors also find negative skewness and excess kurtosis in the data. Instead of following the widely accepted ARCH (Auto-Regressive Conditional Heteroskedasticity) type models, they model the non-linearity in the data using a Hidden Markov structure with two or three states at each observation point driving the data-generating process. The emission mechanism is a straightforward set of mean-variance pairs corresponding to each state. After estimating the model parameters using the EM algorithm, the most probable state sequence is determined by applying the Viterbi algorithm. In this example, the optimal state estimates indicate whether the excess return in a particular month is in a high, medium, or low state. The accuracy of the model is established by the results for the residuals, which turn out to be white noise.

A number of interesting observations were obtained through this approach. To begin with, the well-known leverage effect seen by other researchers using ARCH type models are also seen here. In addition, the HMM approach shows an absence of the leverage effect over prolonged periods. This provides an observation that would be impossible to make using ARCH type models, namely, that the leverage effect may only become observable after sudden sharp movements in equity values. ARCH type models also allow volatility to persist far too long, while failing to detect sharp changes. In contrast, the HMM approach accurately detects sharp changes in the market, and increases in volatility does not persist beyond a few months. While the HMM approach seems to capture the major features of the data quite well, it may be more suitable for analyzing minor changes in the excess return or volatility. When working with financial applications, the HMM approach may thus be more suitable for high-frequency data. This is a prime candidate for future research.

In the second example, Guha and Banerji (1998/1999) develop the interesting approach of comparing two HMMs applied to economic data. They explore whether the underlying drivers of the business cycle data are the same in different levels of aggregation within an economy. To be more precise, they investigate non-farm employment data in several regions of the USA and in the US economy as a whole. Their results show significant differences between the US employment cycle at the national and regional levels. Guha and Banerji (1998/1999) claim that this is consistent with economic intuition.

The approach is to fit a two state HMM to this data, i.e. data at national and regional levels, and to estimate the parameters by applying the EM algorithm discussed earlier. The observation variable is the log difference of the quarterly non-farm employment data. The two hidden states describe either the low-growth or the high-growth phases of the economy. The emission process is simply two different rates of growths corresponding to the underlying states, and the residual is assumed to be Gaussian with the same variance in both states. Once the parameters are estimated using the EM algorithm, the most probable state sequence is estimated for both datasets using Viterbi algorithm. The innovation in the paper is to construct two residual series: one based on the regional growth rate and the means of the regional model based on the regional state sequences, and one based on the same observation data based on the means of the national model and national state sequence. If the estimated state sequences are the same, then intuition holds that the computed residual series will also be the same, on average.

These authors compare the square differences of the above two residual series by computing the probability that this difference is greater than zero. If the two residual series are independent normal, then this comparison can be accomplished by matching only the first two moments. Since this is not the case in this study, the authors resort to the moving block bootstrap method suggested by Kunsch (1989) to compute the probability estimate mentioned above. The entire statistical approach developed in this article appears to significantly enhance the procedure to compare pairs of HMMs applied to business cycle research. We are very encouraged by this development and expect it to be applicable in other areas of financial economics as well.

Notes

1 The illustrations in this section follow Ross (2000) and Tapeiro (1998).

2 See Das (1996) for further information on extending this methodology to contingent claims, e.g. for options or for time invariant risk premia.

3 Additional information on the state space models and their wide array of applications can be found in Shumway and Stoffer (2000) and Bar-Shalom and Li (1993).

4 See MacDonald and Zucchini(1997).

5 See Rabiner (1989).

6 For example, see Rabiner (1989).

7 The discussion of this section relies primarily on the exposition in Rabiner (1989).

8 The analysis presented here is due to Sahani (1999) and this has similarity to the Kalman filter recursions.

This page intentionally left blank

Chapter 2

VOLATILITY IN GROWTH RATE OF REAL GDP

1. Introduction

As many know, instability in a structural model describing a data-generating process can be modeled as a switching regression. Hamilton (1989) shows that a state-dependent switching model might capture some form of non-linearity where the states are driven by an unobserved Markov process. This has been successful in capturing the characteristics of the mean growth rate of the US GDP. In the original formulation of the Hamilton's Markov switching model, the volatility is assumed to remain constant irrespective of the changes from state to state. On the other hand, recent studies by McConnell and Quiros (2000) suggest that volatility among different components of the GDP was lower during the 1990-91 recovery in the United States. As a matter of course, this merits testing for a break in the GDP volatility during expansions and recessions. Kim, Nelson and Piger (2001) also adopt a Bayesian test for the structural break in variance to document some stylized facts regarding volatility reduction.

This chapter attempts to empirically characterize the volatility in the growth rate of real GDP for Japan, the UK and the USA using quarterly data spanning the period from 1960 through to 1996. A recent article by Hamori (2000) presents evidence that the symmetric GARCH (1,1) structure applies reasonably well to these three countries when quarterly observations are applied.[1] Although it is meaningful to expect some asymmetric effect, Hamori (2000) does not find support for this view in the data analyzed. Hamori (2000) explains that economic growth requires expansion of production capacity, increases in labor supply, technological progress, and other prerequisites, whereas economic downturn is somewhat easier to achieve. A fall in demand may be the only necessary trigger to achieve contraction. Based on this premise, Hamori (2000) models the growth rate of real GDP in the GARCH model, a framework capable of factoring in asymmetry.

While GARCH effects are highly significant with daily and weekly financial data, they tend to be much less pronounced in less frequently sampled data, for example, quarterly data on GDP (Bollerslev, Chou and Kroner, 1992). With some types of financial time series, it may be possible to use estimates obtained from more frequently sampled data to make inferences about parameters for less frequently sampled data. Though Drost and Nijman (1993) have demonstrated this for foreign currencies, Bollerslev, Chou and Kroner point out that this temporal aggregation has yet to be confirmed workable for other economic series.

More recently, Markov switching heteroskedasticity has been adopted as an alternative method for dealing with ARCH effects in economic data.[2] The behavior of the unconditional variance constitutes the main difference between the ARCH type conditional heteroskedasticity and the Markov switching variance model. Specifically, the unconditional variance remains constant in the case of the former, whereas it changes with the state of the economy in the latter. In the case of GDP, a more intuitive approach is to think in terms of the different regimes through which the economy may have passed. Hamilton and Susmel (1994) suggest that the long-run variance dynamics may be subjected to regime shifts, and follow an ARCH type process within a given regime. Using weekly data, Hamilton and Susmel (1994) show that the ARCH effect completely dies out after a month. This gives us sufficient reason to examine infrequently sampled data, e.g. quarterly or yearly GDP data, in a Markov switching heteroskedasticity framework. In view of the structural break in volatility foreshadowed in Kim, Nelson and Piger (2001), it becomes important in the modeling of the GDP volatility to reflect its differences during expansion and recession and to drive it by an unobserved Markov process.

Kim, Nelson and Startz (1998) demonstrate good fitting of the Markov switching variance model to monthly stock return data, particularly in terms of normality of the standardized return. Although, Hamori (2000) obtains significant coefficient statistics in the GARCH framework, the model residual still exhibits non-normality. In this chapter we examine the growth rates of real GDP in three countries (Japan, the UK and the USA) within the Markov switching variance framework and compare the results with those derived from the GARCH model using quarterly data. This chapter also seeks to document whether the structural break in the GDP variance in the USA pointed out by Kim, Nelson and Piger (2001) can also be found in the GDP data on the other two OECD members, Japan and the UK.

2. Models

2.1 GARCH Model

This chapter uses two kinds of volatility model. One is a GARCH model and the other is the Markov switching variance model. The GARCH(1,1) model is specified as follows:

$$r_t = \phi_0 + \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + \varepsilon_t, \quad \varepsilon_t | I_{t-1} \sim N(0, h_t), \tag{2.1}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}, \tag{2.2}$$

where $I_{t-1} = (r_{t-1}, r_{t-2}, \cdots, r_{t-p})$.

Equation (2.1) is the mean equation and is specified as an AR(p) process. Equation (2.2) is the conditional variance equation and is specified as the GARCH(1, 1) process. By successively substituting for the lagged conditional variance into equation (2.2), the following expression is found:

$$h_t = \frac{\alpha_0}{1-\beta} + \alpha_1 \sum\nolimits_{i=1}^{\infty} \beta_{i-1} \varepsilon_{t-i}^2. \tag{2.3}$$

An ordinary sample variance would give each of the past squares an equal weight rather than declining weights. Thus the GARCH variance is like a sample variance but it emphasizes the most recent observations. Since h_t is the one period ahead forecast variance based on past information, it is called the conditional variance. The surprise in squared residuals is given by,

$$v_t = \varepsilon_t^2 - h_t. \tag{2.4}$$

Equation (2.4) is by definition unpredictable based on the past. Substituting Equation (2.4) into equation (2.2) yields an alternative expression as follows:

$$\varepsilon_t^2 = \omega + (\alpha_1 + \beta)\varepsilon_{t-1}^2 + v_t - \beta v_{t-1}. \tag{2.5}$$

It can immediately be seen that the squared errors follow an ARMA(1, l) process. The autoregressive root is the sum of α_1 and β, and this is the root which governs the persistence of volatility shocks.

2.2 Markov Switching Variance Model

The second approach, the Markov switching variance model is specified as follows:

$$y_t \sim N(0, \sigma_t^2), \tag{2.6}$$

$$\sigma_t^2 = \sigma_1^2 S_{1t} + \sigma_2^2 S_{2t}, \tag{2.7}$$

$$S_{kt} = \begin{cases} 1 & \text{if } S_t = k \\ 0 & \text{otherwise} \end{cases} \quad (k = 1, 2), \tag{2.8}$$

$$\Pr(S_t = j | S_{t-1} = i) = p_{ij}, \;\; i, j = 1, 2, \tag{2.9}$$

$$\sum_{j=1}^{2} p_{ij} = 1, \tag{2.10}$$

$$\sigma_1^2 < \sigma_2^2, \tag{2.11}$$

where y_t is the demeaned real growth rate, S_t is an unobserved state variable which evolves according to a first order Markov process with transition probabilities in equation (2.9). Equation (2.9) denotes the conditional probability that the date t state is the value $S_t = j$ and the date $t-1$ is the value $S_{t-1} = i$. Equation (2.11) shows that state 1 corresponds to the low volatility state and state 2 corresponds to the high volatility. This would help us establish any break in variance in the GDP process as suggested in Kim, Nelson and Piger (2001).

In this model, the long-run dynamics are governed by regime shifts in unconditional variance according to a first order Markov switching process and as discussed earlier with quarterly data no ARCH terms may be necessary. The model is estimated using the maximum likelihood method and the likelihood function is constructed following the discussion in chapter 4 of Kim and Nelson (1999). The smoothed probabilities i.e. the transition probabilities given the complete data are also obtained following Kim and Nelson. The variance of the growth rate in GDP may then be computed as,

$$E(\sigma_t^2 | \Psi_T) = \hat{\sigma}_1^2 E(S_t = 1 | \Psi_T) + \hat{\sigma}_2^2 E(S_t = 2 | \Psi_T), \tag{2.12}$$

where, $\hat{\sigma}_1^2$ is the estimate of volatility in state 1, $\hat{\sigma}_2^2$ is the estimate of the volatility in state 2, and Ψ_T is the complete observation set, i.e. $\Psi_T = [y_1, y_2, \cdots, y_T]'$.

3. Data

This chapter uses the data on the quarterly real GDP in the United States, the United Kingdom, and Japan. The source is the OECD Main Economic Indicators. The sample period is the first quarter of 1960 through the fourth quarter of 1996. Each variable is seasonally adjusted. The quarterly real growth rate is calculated as $(Y_t - Y_{t-1}) \times 100/Y_{t-1}$, where Y_t is the original data series (real GDP for the United States, the United Kingdom and Japan) at time t.

4. Empirical Results

Table 2.1 shows the results of GARCH model estimation. The AR order for the mean equation is selected by the AIC criterion and is found to be three for all three countries. As the table clearly indicates, the ARCH term and GARCH term are both significant in Japan, with coefficients of 0.0911 and 0.8811, respectively. Similar parameter estimates are obtained for both coefficients in the USA, whereas the estimates are quite discrepant in the UK, with a relatively large estimate for ARCH (0.3682) and small estimate for GARCH (0.4539). Moreover, the sum of α_1 and β_1, a parameter that shows the persistence of volatility, is 0.8221 for the UK, versus 0.9722 and 0.9876 for Japan and the USA, respectively. Thus, the persistence of volatility is relatively high in Japan and the USA but relatively low in the UK.

Table 2.2 gives the residual diagnostics corresponding to the estimation in Table 2.1. The Ljung-Box test is used to check the autocorrelation of the residuals (Ljung and Box, 1979) and the Jarque-Bera test is used to check the normality of residuals (Jarque and Bera, 1987). The entries in this table are $P-$values. $\text{LB}^2(12)$ is the Ljung-Box test of order 12 using squared standardized residuals. As the table indicates, the null hypothesis of no autocorrelation is not rejected for any of the three countries, whereas the null hypothesis of normality is rejected for all three countries at the 1% significance level. The result of non-normality in residuals shows that the GARCH effect is insufficient to capture the characteristics of the distribution.

The empirical results of the Markov switching heteroskedasticity estimation are shown in Table 2.3. The probability p_{11} (the tendency to remain in the low variance state once in that state) is estimated at 0.9764 for Japan, 0.9412 for the UK, and 0.9346 for the USA, and all three values are significant at the 5% significance level. Similarly, the probability p_{22} (the tendency to remain in the high variance state once in that state) is estimated at 0.0128 for Japan, 0.1566 for the UK, and 0.0379 for the USA, but only that for the UK is significant, and only at the 10% level.

Table 2.1. Growth rate in real GDP: GARCH(1,1) estimation

	Japan	UK	USA
ϕ_0	0.2517*	0.5186*	0.4413*
	(0.1024)	(0.1190)	(0.0958)
ϕ_1	0.1418**	0.0368	0.2660*
	(0.0837)	(0.1006)	(0.0899)
ϕ_2	0.3082*	0.0454	0.1678*
	(0.0888)	(0.0857)	(0.0863)
ϕ_3	0.2695*	0.2205*	0.0132
	(0.0794)	(0.0951)	(0.0976)
α_0	0.0227	0.2420	0.0148
	(0.0279)	(0.1554)	(0.0269)
α_1	0.0911*	0.3682**	0.1478*
	(0.0394)	(0.2170)	(0.0578)
β_1	0.8811*	0.4539*	0.8398*
	(0.0440)	(0.2281)	(0.0777)

Note: AR order for the mean equation selected by the AIC criterion is found to be three for all three datasets. The numbers in parentheses below the parameter estimates are standard errors obtained from the heteroskedasticity consistent covariance matrix of the parameters. Significance at 5% level is indicated by * and at 10% level is indicated by **.

Table 2.2. Diagnostics using standardized residuals from GARCH(1,1) model

	Japan	UK	USA
$LB^2(12)$	0.962	0.498	0.307
Normality Test	0.000	0.000	0.029

Note: $LB^2(12)$ is the Ljung-Box test of order 12 using squared standardized residuals and the normality test is obtained from Jarque-Bera statistic. Entries represent corresponding $P-$values. $P-$value less than 0.05 implies the hypothesis of remaining no ARCH effect is rejected and the hypothesis of normal distribution is rejected at the 5% level of significance.

The variance in the low volatility state is estimated at 0.4968 for Japan, 0.4719 for the UK, and 0.1517 for the USA, and all values are significant at the 5% significance level. The variance in the high volatility state is estimated at 2.1081 for Japan, 2.9757 for the UK, and 1.3118 for the USA, and all values are all significant at the 5% significance level. Note that the variances in the high volatility state ($S_t = 2$) in Japan, the

Table 2.3. Growth rate in real GDP: Markov switching heteroskedasticity estimation

	Japan	UK	USA
p_{11}	0.9764*	0.9412*	0.9346*
	(0.0203)	(0.0344)	(0.0475)
p_{22}	0.0128	0.1566**	0.0379
	(0.0148)	(0.0865)	(0.0311)
σ_1^2	0.4968*	0.4719*	0.1517*
	(0.1019)	(0.0927)	(0.0353)
σ_2^2	2.1081*	2.9757*	1.3118*
	(0.3502)	(0.9215)	(0.2165)

Note: Growth rate in real GDP is modeled as $y_t \sim N(0, \sigma_t^2)$, where y_t is the demeaned variable. $p_{11} = \Pr(S_t = 1|S_{t-1} = 1)$, $p_{22} = \Pr(S_t = 2|S_{t-1} = 2)$, S_t is the unobserved state variable that evolves according a first-order Markov process with probabilities defined by p_{11}, p_{22}. σ_1^2 and σ_2^2 are the variances in the two states and $\sigma_1^2 < \sigma_2^2$. The numbers in parentheses below the parameter estimates are standard errors. Significance at 5% level is indicated by * and at 10% level is indicated by **.

Table 2.4. Diagnostics using standardized residuals from Markov switching heteroskedasticity model

	Japan	UK	USA
$\mathrm{LB}^2(12)$	0.510	0.547	0.141
Normality Test	0.473	0.245	0.619

Note: $\mathrm{LB}^2(12)$ is the Ljung-Box test of order 12 using squared standardized residuals and the normality test is obtained from Jarque-Bera statistic. Entries represent corresponding $P-$values. $P-$value less than 0.05 implies the hypothesis of remaining no ARCH effect is rejected and the hypothesis of normal distribution is rejected at the 5% level of significance.

UK, and USA are more than four times, six times, and eight times as great as the variances in the low volatility state, respectively. We are thus able to quantify the differences or the breaks in the variance of the GDP process reported in studies such as that of Kim, Nelson, and Piger (2001).

Table 2.4 shows the diagnostics of the estimation results given in Table 2.3. The table clearly indicates that the null hypotheses of normality and of no autocorrelation are not rejected for any of three countries at the 5% significance level. Note that while the normality of residuals is not supported for the GARCH model, it is supported for the Markov switching heteroskedasticity estimation.

How can we explain these empirical results? One possibility would be a structural change in volatility during the sample period. This accords with the explanation proposed by Kim, Nelson and Piger (2001). As Perron (1989) pointed out, researchers tend to find high persistence in economic variables when they ignore the effects of structural change in their empirical analyses. Diebold (1986) suggested that the high persistence of volatility might be due to the regime shifts in conditional variance. Lamoureux and Lastrapes (1990) conclusively demonstrated that ignoring simple structural shifts in unconditional volatility can lead to the spurious appearance of extremely strong persistence in variance. This spurious persistence may correspond to stationary GARCH movements within regimes with unconditional jumps occurring between regimes. If this hypothesis is true, then the Markov switching variance model should give us good empirical results. In our analysis, the standardized residual of the Markov switching variance model actually supported the normal distribution, whereas the standardized residual of the GARCH model did not.

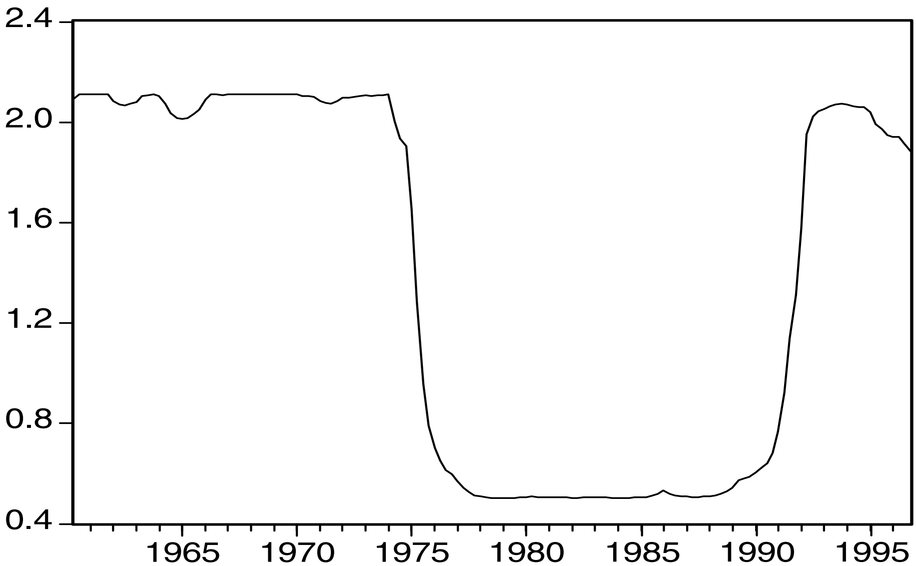

Figure 2.1. Estimated variance from the Markov switching heteroskedasticity model: Japan

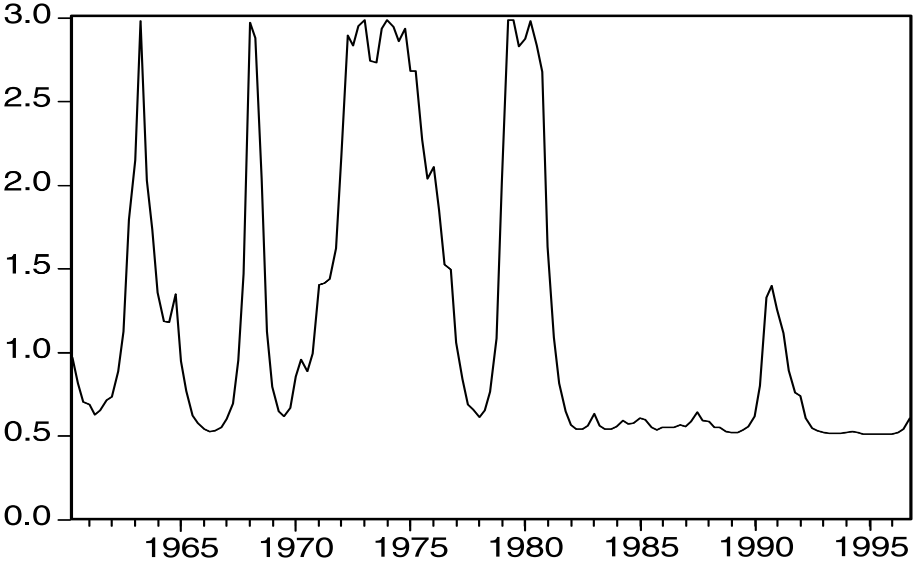

Figure 2.2. Estimated variance from the Markov switching heteroskedasticity model: UK

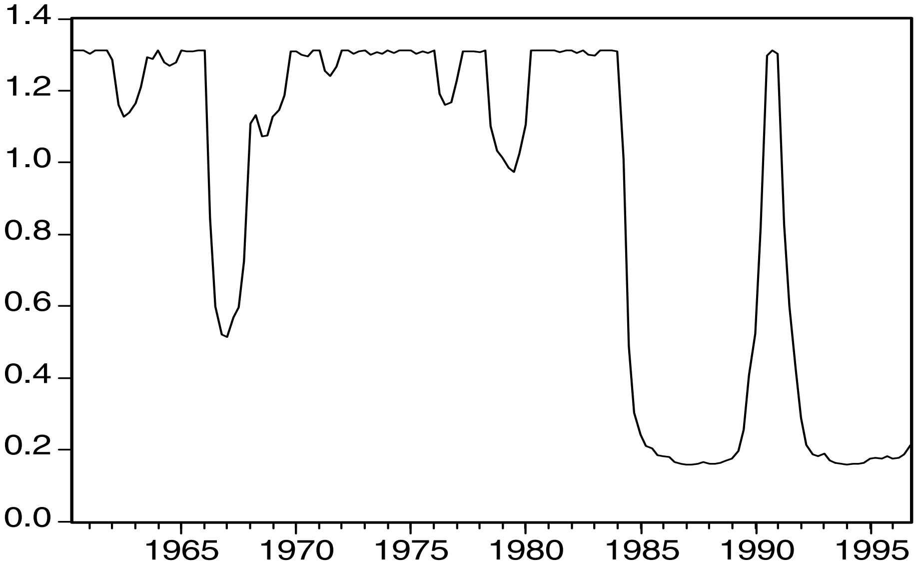

Figure 2.3. Estimated variance from the Markov switching heteroskedasticity model: USA

Figure 2.1, Figure 2.2, and Figure 2.3 show the estimated variance from the Markov switching heteroskedasticity model for each country. These figures depict clear evidence of the structural break in the variance of the GDP process for Japan and the USA. This corresponds to the high persistence of the GARCH estimation reported in Table 2.1. In the case of Japan, we obtain a high variance for the period between 1960 and the middle of the 1970s, whereas we observe a low variance between the middle of the 1970s and the beginning of the 1990s. These years correspond to periods of high and low economic growth in Japan. The apparent similarity of the estimates of variance for all three countries after the middle of the 1980s may be due to the synchronization of business cycles among these countries induced by the liberalization of international capital markets. Upon entering the 1990s, the high variance continues only in Japan, however. When the collapsing bubbles of the Japanese stock markets triggered depression in the Japanese economy in the early 1990s, the resulting economic uncertainty might have been the cause for the increase in the variance.

5. Conclusion

Recent studies have uncovered evidence of a structural break in the variance of the GDP process in the USA. For further clarification, this chapter has investigated an alternative method to characterize the volatility in the growth rate of real GDP for three countries, i.e. Japan, the UK, and the USA. Previous work has documented the usefulness of a GARCH (1,1) model without asymmetry in the innovation for this data. However, the standardized residuals did not support model diagnostics such as normality in the distribution. Besides, the ARCH effect may be largely absent or weak in infrequently sampled data. In stock market applications, the effects of ARCH have been shown to filter out in monthly data.

This chapter has modeled a state-dependent process in order to model the volatility in the growth rate of real GDP. An unobserved Markov process drives the states. Using a Markov switching specification with two states, we find that the model does a credible job in capturing the different episodes. The residual diagnostics all support the proposed specification. In addition to inferring the probability of remaining in any one of these two states, the estimation result quantifies the levels of variance in these two states.[3]

Notes

1 The GARCH (generalized autoregressive conditional heteroskedasticity) model was originally developed by Bollerslev (1986) as an extension of the ARCH (autoregressive conditional heteroskedasticity) model developed by Engle (1982). Bollerslev, Chou and Kroner (1992) published a compelling article surveying the application of ARCH modeling in financial economics.

2 For an example, see Kim, Nelson, and Startz (1998).

3 This chapter is an edited version of Bhar and Hamori (2003a) with permission from Elsevier.

This page intentionally left blank

Chapter 3

LINKAGES AMONG G7 STOCK MARKETS

1. Introduction

Financial deregulation, technological advancements, and other influences over the last three decades have resulted in a closer integration of world stock markets. Economic researchers have been studying the various aspects of this market integration and the way it evolves. Some studies have focused on the interrelationships among national stock markets immediately following significant world events. Malliaris and Urrutia (1992) confirm that almost all stock markets fell together during the crash of October 1987, in spite of the considerable differences in national economies at that time. Similarly, Malliaris and Urrutia (1997) confirm the simultaneous fall of national stock market returns following the Kuwaiti war in July 1990. In the absence of global events, however, national stock markets are dominated by domestic fundamentals. These analyses are at the heart of the question of the benefits of international diversification.

McCarthy and Najand (1995) have reviewed the literature on linkages among international stock markets. Using the state-space methodology, these authors infer the linkage relationship between the stock markets in Canada, Germany, Japan, the UK, and the USA. The authors hold that this methodology delivers the causal relationship in the Granger sense with the minimum number of parameters needed. Their data suggest that the US market exerts a strong influence on all of the other markets studied, while none of the other markets exert a strong influence on the US market. McCarthy and Najand (1995) use daily data and attempt to appropriately handle the overlapping of trading in these markets in the interpretation of their results. Their results concord well with those obtained by Eun and Shim (1989) in their examination of nine stock markets over the period from 1980 to 1985 in the VAR framework.

According to the results obtained by Kasa (1992) and Corhay, Rad and Urbin (1993), stock prices are cointegrated in such a way that world stock markets are driven by one or more common stochastic trends. The presence of a common trend can be interpreted as a natural consequence of well functioning, integrated capital markets that are freely accessible to both domestic and foreign investors. On the other hand, a re-examination by Ahlgren and Antell (2002) using small-sample corrections finds no evidence of cointegration between international stock prices. To explain their finding, they point out that the previous empirical results may have been subject to small-sample bias and size distortion of cointegration tests.

Stock market volatility has been a widely researched topic in financial economics. Shiller (1993) presents a number of articles documenting the sources of market volatility. Some studies have suggested that investors must understand these sources in order to take effective actions to reduce their impact. Based on the premise that excess volatility is the portion of volatility beyond that justifiable by the efficient market hypothesis, Shiller (1993) explores several popular models in order to understand the events surrounding the stock market crash in October 1987. It may also be noteworthy that high levels of volatility have not been recorded exclusively during such isolated events. Jochum (1999), for example, reports regular occurrences of such events in the Swiss market.

Much research effort has also been directed toward modeling market volatility. This has important implications on the derivatives markets as well. Volatility models based on autoregressive conditional heteroskedasticity (ARCH) have been particularly successful in capturing some of the stylized facts. Various researchers have used GARCH (generalized ARCH) models to account for the leptokurtosis, skewness, and volatility clustering often characterizing stock returns. Nelson (1991) and Glosten, Jagannathan and Runkle (1993) extend the standard GARCH model to account for the difference in the effects of negative and positive shocks of the past period on the conditional volatility.

Bollerslev, Chou and Kroner (1992) suggest that although GARCH effects may be highly significant with many daily and weekly financial data, its effect tends to be much milder in less frequently sampled data e.g. quarterly data. The presence of sequential structural shifts due to the nature of news releases in the market, as a cause of conditional heteroscedasticity, has also been proposed by other researchers e.g. Kim and Kon (1996, 1999). In this context, the stock market returns may be viewed as drawn from a mixture of normal distributions. As a rationale for different regimes in the stock market, Cecchetti, Lam and Mark (1990) propose that this may be due to the switching of the economy's endowment between high and low growth phases.

More recently, Kim, Nelson and Startz (1998) adopt Markov switching heteroskedasticity as an alternative method for dealing with the ARCH effects in economic data. The behavior of the unconditional variance constitutes the main difference between the ARCH type conditional heteroskedasticity and the Markov switching variance model. Specifically, the unconditional variance remains constant in the case of the former, whereas it changes with the state of the economy in the latter. If there are sequential changes in regime, as suggested by some authors, a more intuitive approach would be to think in terms of the different regimes contributing to the return-generating process in the stock market. Hamilton and Susmel (1994) also suggest that the long-run variance dynamics may be subjected to regime shift, while those within a single regime are more likely to follow an ARCH type process. Using weekly data, Hamilton and Susmel show that the ARCH effect completely dies out after a month. This gives us sufficient reason to examine less frequently sampled data (e.g. monthly) in a Markov switching heteroskedasticity framework.

In the Markov switching framework Chu, Santoni and Liu (1996) adopt a two-stage process to describe the return behavior in the stock market. In the first stage they model the stock return as a Markov switching process, and in the second stage they estimate a volatility equation using different return regimes derived from the first stage. Returns either above or below some *normal* level are construed as evidence of high volatility. According to their results, the increase in volatility is larger for negative deviations in returns than for positive deviations. Thus, they conclude that the return and volatility relate to each asymmetrically rather than linearly.

In this chapter we attempt to analyze the stock return characteristics of the G7 countries using monthly returns, and to capture the changes in mean-variance in a two-state framework, where an unobserved Markov process drives the states. Our work complements the study by Chu, Santoni and Liu (1996). Whereas Chu, Santoni and Liu (1996) adopt a two-stage approach, we estimate the model in one stage, allowing both the mean and variance to depend on the state. The estimation of this model for the G7 countries can be used to infer the differences in the volatility regimes and the likelihood that an economy will remain in one of the two states.

The contributions of this chapter are therefore twofold. First, we characterize the national stock markets of the G7 countries using a Markov switching specification to the mean and variance, and thus document the probability that these markets will stay in low or high volatility states. Compared to the GARCH process, the Markov switching variance structure is a more intuitive method for modeling changes in variance in less frequently sampled data. In this context we extend the analysis of Schaller and Van Norden (1997) to other members of the G7. Secondly, we use concordance, a measure of co-movement of stock prices proposed by Harding and Pagan (1999), to analyze the informational dependence between these markets, as well as to describe the proportion of time during which the stock prices of two countries remain concurrently in the same phase. Information on the clustering of turning points is also summarized.

2. Empirical Technique

2.1 Markov Switching Stock Return Model

We model the series as an AR process with mean and variance depending on an unobserved state. Denoting the return at time t by y_t, the two-state Markov mean-variance model is written as,

$$y_t - \mu_{S_t} = \phi_1(y_{t-1} - \mu_{S_{t-1}}) + \epsilon_t, \quad \epsilon_t \sim i.i.d.N(0, \sigma^2_{S_t}), \tag{3.1}$$

where

$$\Pr(S_t = j | S_{t-1} = i) = p_{ij}, \quad i, j = 1, 2, \quad \sum_{j=1}^{2} p_{ij} = 1, \tag{3.2}$$

$$\mu_{S_t} = \mu_1 S_{1t} + \mu_2 S_{2t}, \tag{3.3}$$

$$\sigma^2_{S_t} = \sigma^2_1 S_{1t} + \sigma^2_2 S_{2t}, \tag{3.4}$$

$$S_{mt} = \begin{cases} 1 & \text{if } S_t = m, \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

State 1 corresponds to the low volatility state and state 2 corresponds to the high volatility ($\sigma^2_1 < \sigma^2_2$). The estimation of the model is carried out by constructing the probability weighted likelihood function and maximizing this with respect to the model parameters. In this context we follow the procedure discussed in Kim and Nelson (1999, p. 65) to deal with the unobserved state variable. The estimation process also yields the filtered probabilities i.e. the probabilities about the state S_t conditional on the information, Ψ_t, up to time. Conditional on the information in the whole sample period, Ψ_T, and using the algorithm described in Kim and Nelson (1999, p. 68) we develop the estimates of the smoothed probabilities of the states.

Given the estimates of the parameters and the smoothed probabilities, it is then straightforward to compute the variance of the stock returns as follows:

$$E(\sigma_t^2|\Psi_T) = \hat{\sigma}_1^2 E(S_t = 1|\Psi_T) + \hat{\sigma}_2^2 E(S_t = 2|\Psi_T), \tag{3.6}$$

where, $\hat{\sigma}_1^2$ is the estimate of volatility is state 1, $\hat{\sigma}_2^2$ is the estimate of the volatility in state 2, Ψ_T is the complete observation set. Besides, we perform the model diagnostics using the residual generated during the construction of the likelihood function. This may be described as,

$$y_t - E\left(y_t|\Psi_{t-1}\right) = y_t - E\left(\mu_{S_t}|\Psi_t\right) - \phi_1 E\left(y_{t-1} - \mu_{S_{t-1}}|\Psi_t\right). \tag{3.7}$$

The above expression may be interpreted as the forecast error.

2.2 Concordance Measure

The recursive estimation process (Hamilton, 1994) generates the probability that a particular month is in high volatility or low volatility state. Using these probability state estimates we form the concordance statistics. Here is the brief description of this new statistic. Concordance measure is a non-parametric statistic, first proposed by Harding and Pagan (1999) and later extended by McDermott and Scott (1999) for its distributional properties. This statistic has been successfully applied in studies of co-movement of prices in seemingly unrelated commodities, e.g. Cashin, McDermott, and Scott (1999). The concordance statistic between the two series, x_i and x_j is defined by,

$$C_{i,j} = T^{-1}\left\{\sum_{t=1}^{T}\left(S_{i,t}S_{j,t}\right) + \left(1 - S_{i,t}\right)\left(1 - S_{j,t}\right)\right\}, \tag{3.8}$$

where T is the number of observations in each series, $S_{i,t}$ is a binary variable taking on the value 0 when the corresponding value of x_i is below a certain reference level, otherwise it is 1. Similarly, $S_{j,t}$ is defined. In this chapter we are dealing with the probability series, we choose 0.25 as a reference value. This implies that when the estimated state probability is less than 25% we consider it low and assign 0 to the corresponding S variable. In order to make statistical significance test of the computed concordance statistic between two series, Cashin, McDermott, and Scott (1999) propose and carry out simulation experiment to establish validity of their approach. We follow that guideline and compute the critical values of the concordance statistic for 10%, 5%, and 1% level of significances under the assumption that a Brownian motion without a drift has generated the probability state realizations. The application of Markov switching model fitted to the stock return process makes this a valid assumption. The relevant critical values are included in the table describing the concordance matrix of the seven series under investigation.

3. Data

This study uses data on the monthly stock prices of the G7 countries, i.e. Canada, France, Germany, Italy, Japan, the UK, and the USA. The data were taken from the International Financial Statistics of the International Monetary Fund. The sample period spans the approximately 30-year period from January 1970 through March 1999. The rate of return is calculated as $R_t = 100 \times (P_t - P_{t-1})/P_{t-1}$, where P_t is the stock price index at time t. Thus, the stock returns are obtained for the period between February 1970 and March 1999. Table 3.1 gives statistics summarizing the national stock market return in each country. The table includes descriptive statistics such as the mean, standard deviation (Std. Dev.), skewness, kurtosis, and the $P-$value of the Jarque-Bera test (JB test). The hypothesis of normal distribution is rejected at the 5% (1%) level of significance if the $P-$value for the JB-test is less than 0.05 (0.01). In Table 3.2 we include the correlation among these national stock market returns. Table 3.1 clearly shows that France, the UK, and Italy are relatively high risk and high return countries, whereas Canada, Germany, and Japan are relatively low risk and low return countries. The hypothesis of normal distribution is rejected at the 1% significance level for all of the countries.

4. Empirical Results

The parameter estimates of the Markov switching heteroskedasticity models are given in Table 3.3. The estimates of the transition probability of both p_{11} (low variance state) and p_{22} (high variance state) are significant at the 1% level for all of the data sets. The p_{11} estimates are higher than the p_{22} estimates for all of the G7 countries except Japan.

The variance estimates in the low volatility state are much lower than the corresponding estimates in the high volatility state. Though all variance estimates are statistically significant, those in the high volatility state for France and the UK are much higher than those for the other G7 members. This confirms the conclusion of Schaller and Van Norden

Table 3.1. Summary statistics on stock return

	Canada	France	Germany	Italy	Japan	UK	USA
Mean	0.6636	0.9640	0.6302	0.9556	0.6382	0.9571	0.8517
Std. Dev.	4.8020	6.9157	4.8397	6.3748	4.2285	5.1632	3.6323
Skewness	-0.6382	0.9685	-0.5094	0.2674	-0.1959	1.2558	-0.4279
Kurtosis	5.9324	12.3093	5.1450	3.6035	3.7579	16.9842	4.8675
JB test	0.0000	0.0000	0.0000	0.0087	0.0050	0.0000	0.0000

Note: The hypothesis of normal distribution is rejected at the 5% (1%) level of significance if the $P-$value for the JB test is less than 0.05 (0.01).

Table 3.2. Correlation of stock returns

	Canada	France	Germany	Italy	Japan	UK	USA
Canada	1.0000	0.4221	0.3877	0.2149	0.2730	0.3378	0.5897
France		1.0000	0.4294	0.1829	0.1847	0.2278	0.3173
Germany			1.0000	0.2914	0.2515	0.3355	0.4242
Italy				1.0000	0.2423	0.3629	0.3225
Japan					1.0000	0.3545	0.3788
UK						1.0000	0.5642
USA							1.0000

(1997) on the US stock market and applies to other members of the G7 as well. The means in the low volatility state range from a low of -0.4504 for Italy to a high of 1.1520 for UK, and these values are significant at the 1% level for all countries except Italy. The mean in the high volatility state is only significant at the 5% level for Italy. Notice that all of the countries except Italy have higher return in the low volatility state. This is consistent with the findings of Chu et al. (1996).

Table 3.4 provides some diagnostics of the empirical results. As the table shows, neither the null hypothesis of no autocorrelation (in both the standardized residual and squared standardized residual) nor the null hypothesis of normality is rejected for any of the seven countries at the 1% significance level. These results are very encouraging and

Table 3.3. Parameter estimates: Markov switching heteroscedasticity model of stock return

	Canada	France	Germany	Italy
p_{11}	0.9593**	0.9735**	0.9632**	0.8427**
	(0.0203)	(0.0521)	(0.0252)	(0.1199)
p_{22}	0.8402**	0.6493**	0.9080**	0.8182**
	(0.0709)	(0.1503)	(0.0620)	(0.1252)
σ_1^2	12.7513**	27.6203**	12.7855**	15.0558**
	(1.4493)	(3.6934)	(1.9127)	(4.2326)
σ_2^2	62.0948**	317.8201*	49.7709**	59.9137**
	(13.8951)	(151.7166)	(11.1484)	(10.0798)
μ_1	1.0239**	1.0248**	0.7720**	-0.4504
	(0.2389)	(0.3305)	(0.2699)	(0.5863)
μ_2	-0.8325	0.1402	0.3004	2.6457**
	(1.1027)	(0.7428)	(0.9785)	(1.0183)
ϕ_1	0.0041	0.0521	0.0704	0.2519**
	(0.0437)	(0.0613)	(0.0557)	(0.0564)

	Japan	UK	USA
p_{11}	0.9809**	0.9893**	0.8555**
	(0.0142)	(0.0084)	(0.0599)
p_{22}	0.9906**	0.8509**	0.5372**
	(0.0089)	(0.1165)	(0.1480)
σ_1^2	5.5655**	14.7712**	5.6962**
	(0.8346)	(1.3439)	(0.7759)
σ_2^2	22.5227**	143.7775*	31.9665**
	(2.5471)	(57.2016)	(7.3916)
μ_1	1.0259**	1.1520**	1.1024**
	(0.3134)	(0.2842)	(0.2329)
μ_2	0.4167	-2.2098	0.0017
	(0.4793)	(3.4521)	(0.0630)
ϕ_1	0.3043**	0.2258**	0.2502**
	(0.0519)	(0.0494)	(0.0468)

Note: The parameters are described in text. Standard errors are given in parentheses below the parameter estimates. Significance at the 1% level is indicated by ** and at the 5% level is indicated by *.

strongly support the empirical results obtained in Table 3.3, as well as the modeling approach adopted in this chapter.

Table 3.4. Diagnostics using standardized returns from Markov switching heteroscedasticity model

	Canada	France	Germany	Italy	Japan	UK	USA
LB (16)	0.789	0.107	0.416	0.045	0.673	0.030	0.615
LB^2 (16)	0.048	0.984	0.485	0.381	0.512	0.924	0.187
JB test	0.344	0.064	0.219	0.318	0.623	0.412	0.033

Note: LB (16) and LB^2 (16) are the Ljung-Box tests of order 16 using standardized residuals and squared standardized residuals respectively. The JB test is for normality obtained from Jarque-Bera statistic. Entries represent corresponding $P-$values. For LB (16) test, $P-$value less than 0.05 (0.01) implies that the hypothesis of white noise is rejected at the 5% (1%) level of significance. Similarly, for LB^2 (16) remaining no ARCH effect is rejected for $P-$value less than 0.05 (0.01) at the 5% (1%) level of significance. The hypothesis of normal distribution is rejected at the 5% (1%) level of significance if the $P-$value for the JB-test is less than 0.05 (0.01).

Table 3.5. Estimated Markov probabilities of staying in the same state for the G7 countries

	Canada	France	Germany	Italy	Japan	UK	USA
Regime 1	0.9593	0.9735	0.9632	0.8427	0.9809	0.9893	0.8555
Regime 2	0.8402	0.6493	0.9080	0.8182	0.9906	0.8509	0.5372

Note: Regime 1 represents the low variance state and the regime2 represents the high variance state.

Table 3.5 highlights the estimated Markov probabilities that the markets in the G7 countries will stay in the same states. The probability of staying in regime 1, i.e. the low volatility state, is very high for all G7 members. The probability of staying in regime 2, the high volatility state, appears to be relatively low for six of the G7 countries, that is, all of them but Japan. Japan has the highest propensity to remain in the high volatility state once it enters that state. The probability of staying in regime 2 is the lowest for the US market, followed by France. If the current month falls inside the low variance state, the probability of remaining in the same state the next month is very high for all of the countries but Japan. Again, to the best of our knowledge, this obser-

Table 3.6. Correlation statistics between probability of high variance state

	Canada	France	Germany	Japan	Italy	UK	USA
Canada	1.0000	0.3208*	0.0288	0.1981*	-0.0857	0.4819*	0.6012*
France		1.0000	0.1836*	0.2736*	0.0143	0.4115*	0.3287*
Germany			1.0000	0.2662*	0.4352*	0.1236*	0.1318*
Japan				1.0000	-0.0279	0.2306*	0.1765*
Italy					1.0000	0.1800*	0.0651
UK						1.0000	0.5301*
USA							1.0000

Note: Critical value for individual correlation at 5% level significance is 0.1049 (indicated with *). This is calculated as $1.96/\sqrt{N}$, where N is the number of observations. Test statistic for significance as a group is 549.82. This is computed as $-2\ln(|R|)^{0.5N}$, where R is the correlation matrix. The critical value for the group statistic is obtained from a Chi-squared distribution with degrees of freedom $0.5p(p-1)$, where p is the number of series, and at 5% level this is 32.67.

Table 3.7. Concordance statistics between probability of high variance state

	Canada	France	Germany	Japan	Italy	UK	USA
Canada	1.0000	0.7794*	0.6132*	0.5158	0.3696	0.8138*	0.7994*
France		1.0000	0.6619*	0.4499	0.3553	0.8968*	0.7507*
Germany			1.0000	0.5072	0.6132*	0.6676*	0.6189*
Japan				1.0000	0.4527	0.4499	0.5158
Italy					1.0000	0.3897	0.4269
UK						1.0000	0.7794*
USA							1.0000

Note: The critical values are computed following McDermott, and Scott (1999) under the assumption that a Brownian motion without a drift has generated the realizations of the state probabilities. These values are 0.5456, 0.5636 and 0.5949 for 10% (***), 5% (**) and 1% (*) level of significances.

vation of the behavior of national stock markets has not been reported earlier.

Table 3.6 shows the correlations between the probabilities of high variance states. The analysis using probabilities of low volatility states will produce the same result. This occurs because these two probabilities for a particular month add up to one. Individual cross-correlations exceeding 0.1049 in magnitude are significant at the 5 percent level. As the table clearly shows, 16 out of 21 correlations satisfy this criterion: the USA is correlated with Canada, France, Germany, Japan, and the UK; the UK is correlated with Canada, France, Germany, Japan, Italy, and the USA; Italy is correlated with Germany; Japan is correlated with Canada, France, and Germany; Germany is correlated with France; and France is correlated with Canada.

Table 3.7 shows the concordance statistics. The $(i, j)-$th cell represents concordance between the $i-$th and $j-$th countries. Thus, the numbers along with the diagonal are unity. For all the combinations, 11 statistics are significant at the 1 percent level. The highest concordance, 0.8968, is recorded between the UK and France, and this is significant at the 1 percent level. This high concordance indicates that the stock prices of the two countries are in the same phase most of the time. By comparing the results in Table 3.6 and Table 3.7, we see that the two measures support quite discrepant conclusions on co-movement. For example, it would appear that some of the significant correlations in the case of the UK and USA are not supported by the concordance statistics. Similarly, the significant correlations in the case of Japan do not constitute evidence of phase synchronization among Japan, Canada, France, and Germany. The underlying reason could be the different economic forces at play in these markets. The overall evidence suggests that international co-movement in equity markets, at least in these G7 economies, is doubtful and requires further exploration.

5. Conclusion

This chapter enhances the investigation of international linkages in stock markets by focusing on the information dependence between the markets. The characteristics of the monthly stock market return from G7 countries are first captured in a Markov switching framework. Although the traditional notion of volatility clustering is less predominant in the monthly frequency, volatility clustering is likely to represent a mixture of distribution. Accordingly, this study adopts an approach that allows both the mean and variance to depend on an unobserved state driven by a Markov process. The efficacy of this model to capture the monthly stock return characteristic is illustrated by the diagnostics of the standardized residuals.

The results show that the low variance states are quite stable, whereas the high variance states are relatively short-lived. Thus, if the current month falls inside the low variance state, the probability of remaining in the same state the next month is very high. Among the seven countries, it is also interesting to note that the USA has the lowest propensity to stay in the high volatility state once already in that state, whereas the Japanese market has the highest propensity to stay in the high volatility state. Moreover, the magnitude of the variance in the high variance state is several times higher than that in the low variance state for all the markets.

When we defined co-movements in terms of concordance, i.e. the proportion of time that stock prices are in the same state, we do not find clear evidence that international stock prices move together. This is consistent with Ahlgren and Antell (2002). The underlying reason could be the different economic forces at play in these markets. The overall evidence suggests that international co-movement in equity markets, at least in these G7 economies, is doubtful, and requires further exploration.[1]

Notes

1 This chapter is an edited version of Bhar and Hamori (2003b) with permission from Finance Letters.

APPENDIX 3.A
Data

The stock price indices are obtained from the International Financial Statistics of the International Monetary Fund. The series code of stock price index in each country is shown as follows:

Canada: 15662...ZF
France: 13262...ZF
Germany: 13462...ZF
Italy: 13662...ZF
Japan: 15862...ZF
UK: 11262...ZF
USA: 11162...ZF

This page intentionally left blank

Chapter 4

INTERPLAY BETWEEN INDUSTRIAL PRODUCTION AND STOCK MARKET

1. Introduction

Extensive studies have focused on the empirical regularity in the degree of correlations and related issues between different equity markets. These studies have provided insight into the nature of comovements across markets and what might be driving this phenomenon. Baca et al. (2000) ascribe the phenomenon to the global industry factor. Brooks and Del Negro (2002) find that there may be only one industry factor (high technology) that might explain this notion. They also demonstrate that this might be a temporary phenomenon, as country-based diversification appears to have remained effective for portfolio risk since the collapse of the technology bubble.

In studies adopting a slightly different focus, researchers such as Kasa (1992) and Engsted and Lund (1997) rely on the results from cointegration tests in international equity prices. They also examine whether the underlying dividend process can explain this cointegrating behavior. Although the return patterns from these markets can differ in the short term, cointegrating behavior would suggest that they are closely linked over the long term. In a paper stressing the unlikelihood of cointegrating behavior from the perspective of economic theory, Richard (1995) proves that Kasa's (1992) finding may result from the inappropriate use of critical values in their statistical tests.

In this chapter we attempt to explore the phenomenon of comovement among the G7 equity markets from a different perspective. We make use of a concordance measure to document whether any of these markets tend to be in phase with the equivalent volatility state of the economy. Specifically, our analysis focuses on the G7 markets over a

30-year period at a monthly frequency. The first step in our approach is to model the interaction between the state of the economy and the stock market. This stems from the work of Hamilton and Lin (1996) and Chauvet (1998/1999), who demonstrate important relationships between stock markets and business cycle variables such as industrial production. Once equipped with our model, we can develop a probabilistic picture of whether a particular month is in the state of expansion or contraction. We then utilize this information and apply the concordance measure (Harding and Pagan, 1999) to capture the likelihood that the stock markets are in the same phase. Our results show that not all the G7 economies are commoving in this respect.

In modeling stock return and return volatility, the return volatility can be forecasted though the return itself cannot. Another important point to recognize is the dynamic relationship between the stock market and business cycle observed by Chauvet (1998/1999) in his attempts to anticipate turning points in the business cycle using stock market factors in a dynamic factor framework at a monthly frequency.

As large changes in volatility may alarm investors and adversely influence their investment behavior, many investigators have sought to capture the patterns of stock return volatility. Several studies have applied time-varying conditional moments to capture such stylized facts. ARCH and GARCH models have been popular for this purpose, and Nelson (1991) and Glosten, Jagannathan and Runkle (1993) have proposed a variation of the GARCH model to capture asymmetries in stock return and stock index return.

Schwert (1989, 1990) focuses on two major aspects of stock market volatility, namely changes in stock price volatility over time and the economic factors related to significant changes in stock price volatility over time. Based on an analysis of 105 years of stock price history, Schwert (1989) reports that the greatest single-day decline took place on October 19, 1987, while almost all other large declines took place from 1929 to 1939. Schwert also finds weak evidence that financial and operating leverage measures are positively correlated with stock market volatility. His findings additionally demonstrate that stock return volatility rises during periods of recessions, and a similar analysis by Jones and Wilson (1989) demonstrates that the modeled volatility during the 1980s depends on the measures of volatility employed and interval used.

Others have proposed an alternative characterization of volatility episodes focusing on different regimes in the return generating process. Cecchetti, Lam and Mark (1990) consider a Lucas asset pricing model where the economy's endowment switches between high- and low-growth phases. In doing so, they demonstrate that such switching in fundamentals can explain several features of stock market return. Chu, Santoni and Liu (1996) adopt a two-stage process to describe the return behavior. In the first stage the stock return is modeled as a Markov switching process, and in the second a volatility equation is estimated using different return regimes derived from the first stage. They find evidence of higher volatilities when the returns are either above or below some "normal'" level. In a multivariate extension of the regime switching approach, Schaller and Van Norden (1997) show that the response of returns to the past price-dividend ratio in the US is strongly asymmetric. More specifically, they showed that the impact of the past price-dividend ratio is about four times larger in the low-return state than in the high-return state.

Thus, regime switching behavior is now well established as a viable method to analyze stock market behavior. As a natural extension of the most commonly applied stock return model (GARCH), Hamilton and Susmel (1994) propose a switching ARCH model in which the parameters of the ARCH process can be drawn from one of several different regimes. Although the ARCH process controls the short-run dynamics, the long-run dynamics are governed by regime shifts in unconditional variance, and an unobserved Markov switching process drives the regime changes. When these authors apply the model to weekly return data, the ARCH effects almost completely fade away after a month. This suggests that no ARCH term is necessary in modeling the monthly return.

Hamilton and Lin (1996) adopt the switching ARCH structure of Hamilton and Susmel (1994) to characterize the stock return process. They enhance the model to a bivariate system where a relevant macroeconomic variable is used as a determinant of stock market volatility. They propose that an unobserved Markov switching variable drives the learning process of the market participants and thereby affects the industrial output and stock market with a time lag. This joint specification does a credible job in capturing the behavior of the stock market volatility when applied to US data at the monthly frequency.

In this chapter we adopt Hamilton and Lin's (1996) basic precept that the anticipation of an economic downturn affects the stock market before the industrial output actually starts to decline. At the same time, we presume that most of the ARCH effect would fade away at monthly frequency, as suggested by Hamilton and Susmel (1994), and therefore model the conditional heteroskedasticity directly in the Markov switching framework. Kim, Nelson and Startz (1998) also demonstrate good fitting of Markov switching variance model to monthly stock return data.

Moving forward with the above ideas described by Hamilton and Lin (1996) and Kim, Nelson and Startz (1998), we set up a bivariate framework for the growth in industrial production and stock return and extend the results of Hamilton and Lin to all the G7 countries using more recent data. As a result, we find a similar yet varying interaction between the two variables. Our work then extends to analyzing comovement in these stock markets using a new statistical measure of concordance. Specifically, we measure whether the state probabilities of being in the expansionary phase coincide in these markets. Within the same experimental framework, we can use the estimated model parameters to check whether the three-month ahead expectations of the state probabilities are in phase.

2. Markov Switching Heteroscedasticity Model of Output and Equity

We discuss the joint modeling of the industrial output and the equity market return in this section following Hamilton and Lin (1996). We present a simplified exposition of the generalized version of Hamilton and Lin (1996) along with our changes to the excess equity return process. We have endeavored to present the essential concepts in such a way that it can be easily applied to other similar situations as well as help develop the algorithm using a suitable computer language.

The main hypothesis in this chapter concerns the delayed impact of upswing and downswing of industrial output on the volatility of equity return. It is assumed that an unobserved two-state Markov switching process drives these variables. The impact of this Markov variable, however, is not contemporaneous and there is a lag of one time period. In other words, the impact of increasing industrial output this period, for example, during economic expansion will be transmitted to the equity market volatility next period. Let y_t represent the growth rate of industrial output and its dynamics consist of two components - an autoregressive component and an economic state dependent mean term. We express this as,

$$y_t = a_1 z_{t-1} + \mu_{S,t-1} + e_t, \tag{4.1}$$

where, $z_{t-1} = y_{t-1} - \mu_{S,t-2}$, $e_t \sim N(0, \sigma_y^2)$ is the error term with time invariant variance. The economic state dependent mean component is captured by $\mu_{S,t}$, where,

$$\mu_{S,t} = \begin{cases} \mu_1, & S_t = 1, \\ \mu_2, & S_t = 2, \end{cases} \tag{4.2}$$

and we may associate $S_t = 1$ with the state of economic slowdown and $S_t = 2$ with the state of economic expansion.

We next express the excess equity market return as an autoregressive process with Markov switching variance, as,

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma^2_{S,t}), \tag{4.3}$$

where,

$$\sigma^2_{S,t} = \begin{cases} \sigma^2_1, & S_t = 1, \\ \sigma^2_2, & S_t = 2. \end{cases} \tag{4.4}$$

Our model of the equity excess return is different from that employed by Hamilton and Lin (1996) and this is explained below. Since we propose to carry out the empirical investigation using monthly data it is more appropriate to model the volatility process as consisting of different regimes rather than an ARCH process. Hamilton and Lin (1996) also reinforce our observation and state that ARCH terms could not be estimated from the monthly data for their model.[1]

The bivariate model of the industrial production and the excess return given by equations (4.1) and (4.3) respectively may be estimated by constructing the probability weighted conditional density function. To achieve this we need to specify the probability law governing the evolution the Markov switching variable. We define the transition probability of the Markov chain defined by S_t as below:

$$\begin{bmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{bmatrix}. \tag{4.5}$$

It may be observed from equation (4.1) and (4.3) that we need to carry information of the Markov switching variable S_t, for three time periods at any one time i.e. for the time periods, $t-2$, $t-1$, and t. This implies that there are eight relevant states that are of importance to this bivariate system. We can define them as a set of eight combinations where the ordering of the state occurrences of 1 or 2 reflects the time steps $t-2$, $t-1$, and t respectively:

$$\begin{aligned} &\{1,1,1\}, \{1,1,2\}, \{1,2,1\}, \{1,2,2\}, \\ &\{2,1,1\}, \{2,1,2\}, \{2,2,1\}, \{2,2,2\}. \end{aligned} \tag{4.6}$$

With the help of (4.5) it is now possible to represent the transition probability matrix of the eight- state Markov chain in (4.6) as,

$$\boldsymbol{P} = \begin{bmatrix} p_{11} & 0 & 0 & 0 & p_{11} & 0 & 0 & 0 \\ 1-p_{11} & 0 & 0 & 0 & 1-p_{11} & 0 & 0 & 0 \\ 0 & 1-p_{22} & 0 & 0 & 0 & 1-p_{22} & 0 & 0 \\ 0 & p_{22} & 0 & 0 & 0 & p_{22} & 0 & 0 \\ 0 & 0 & p_{11} & 0 & 0 & 0 & p_{11} & 0 \\ 0 & 0 & 1-p_{11} & 0 & 0 & 0 & 1-p_{11} & 0 \\ 0 & 0 & 0 & 1-p_{22} & 0 & 0 & 0 & 1-p_{22} \\ 0 & 0 & 0 & p_{22} & 0 & 0 & 0 & p_{22} \end{bmatrix}. \tag{4.7}$$

Before we can apply the probability weighting to the conditional density function we describe the structure of the bivariate density of the model. We can rewrite equation (4.1) as,

$$e_t = (y_t - a_1 y_{t-1}) - (\mu_{S,t-1} - a_1 \mu_{S,t-2}), \tag{4.8}$$

and equation (4.3) as,

$$\nu_t = x_t - \alpha_0 - \alpha_1 x_{t-1}. \tag{4.9}$$

The variance of equation (4.8) is time invariant and that of equation (4.9) is state dependent as described earlier. Both these are conditionally normal and uncorrelated. It is, therefore, straightforward to write the conditional density function given the parameters and the states as the product of the two normal densities. We then multiply the conditional densities for different states by the corresponding probabilities of the states and sum them to obtain the likelihood function. It is this likelihood function we maximize with respect to the parameters of the model. The algorithm is described in detail in Hamilton (1994, chapter 22) and Kim and Nelson (1999, chapter 4). The unknown parameters of the model include, $\Theta \equiv \left(a_1, \mu_1, \mu_2, \sigma_y^2, \alpha_0, \alpha_1, \sigma_1^2, \sigma_2^2, p_{11}, p_{22}\right)$.[2]

In order to understand the effectiveness of the model we further carry out diagnostics using standardized residual from the excess return equation. This is generated as follows. From the estimated filtered probabilities we can define,

$$\begin{aligned} E\left(\nu_t^2 \left| x_{t-1}, x_{t-2}, \ldots\right.\right) = \; & \sigma_1^2 \times \Pr\left(S_t = 1 \left| x_{t-1}, x_{t-2}, \ldots\right.\right) \\ & + \sigma_2^2 \times \Pr\left(S_t = 2 \left| x_{t-1}, x_{t-2}, \ldots\right.\right). \end{aligned} \tag{4.10}$$

Using equation (4.10) and the excess return residual defined by equation (4.9), we construct the standardized residual series and test for correlation in the level as well as in squared series. These tests are discussed in the empirical results section.

As is explained in Chapter 3, The recursive estimation process generates the probability that a particular month is in high volatility or low volatility state. Using these probability state estimates we form the concordance statistics. The concordance statistic between the two series, $x_i$and x_j is defined by,

$$C_{i,j} = T^{-1} \left\{ \sum_{t=1}^{T} (S_{i,t} S_{j,t}) + (1 - S_{i,t})(1 - S_{j,t}) \right\}, \tag{4.11}$$

where T is the number of observations in each series, $S_{i,t}$ is a binary variable taking on the value 0 when the corresponding value of x_i is below a certain reference level, otherwise it is 1. Similarly, $S_{j,t}$ is defined. This measure is different from correlations since that is influenced by the magnitude of the variable, whereas here the emphasis is on being in the same phase in some meaningful way. In this chapter since we are dealing with the probability series, we choose 0.50 as a reference value. This implies that when the estimated state probability is less than 50% we consider it low and assign 0 to the corresponding S variable. In order to make statistical significance test of the computed concordance statistic between two series, Cashin, McDermott, and Scott (1999) propose and carry out simulation experiment to establish validity of their approach. We follow that guideline and compute the critical values of the concordance statistic for 10%, 5%, and 1% level of significances under the assumption that a Brownian motion without a drift has generated the probability state realizations. The application of Markov switching model fitted to the stock return process makes this a valid assumption in our model.

We are not only interested in finding whether the probability of being in a particular volatility state is in phase, we would also like to explore whether the expectation of such volatility states are in phase a few months ahead. The Markov model structure allows us to investigate this relatively easily. Given the current state probabilities we could form the expectation for k period ahead using the estimated transition probabilities. The following equation captures this relationship:

$$\Pr(S_{t+k} = 1 | \Im_t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{bmatrix}^k \begin{bmatrix} \Pr(S_t = 1 | \Im_t) \\ \Pr(S_t = 2 | \Im_t) \end{bmatrix}, \tag{4.12}$$

Table 4.1. Summary statistics: monthly excess return

	Canada	France	Germany	Italy	Japan	UK	USA
Mean (%)	0.4437	0.6583	0.5922	0.3368	0.5370	0.6339	0.5926
Std. Dev.	5.0066	6.1178	5.3756	7.3216	5.4592	6.2581	4.4140
Skewness	-0.4206	-0.0941	-0.3232	0.4601	-0.0431	1.3837	-0.3168
Kurtosis	5.3067	4.0583	4.8181	3.8929	4.1826	18.2241	5.3182
JB test	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000

Note: The hypothesis of normal distribution is rejected at the 5% (1%) level of significance if the $P-$value for the JB test is less than 0.05 (0.01).

where, $\Im_t$ is the information available at time t.

3. Data

We use monthly data on industrial production indices and stock returns in the G7 countries, i.e. Canada, France, Germany, Italy, Japan, the UK, and the USA. The sample period runs from January 1971 through June 2000, and thus includes 354 observations for each sample. Our data set covers more of recent years compared to that in Hamilton and Lin (1996), though it also starts a few years later. The industrial production indexes are seasonally adjusted data from the International Financial Statistics provided by the International Monetary Fund. The growth rate of the industrial production index is used for empirical research. The stock returns are obtained from the Morgan Stanley Capital International (MSCI) Index. The MSCI takes into consideration both capital gain and dividend income, and is thus appropriate for our analysis. Stock returns are calculated as the growth rate of the MSCI index. Nominal short-term interest rates are also obtained from the International Financial Statistics of the International Monetary Fund. The excess stock return is calculated from the growth rate of the MSCI index and the nominal short-term interest rate. Further details on our data sources are given in the appendix.

Tables 4.1 and 4.2 summarize the monthly excess returns and the monthly growth rates of industrial production indexes in the G7 countries. The figures given are means expressed percentages, standard deviations (Std. Dev.), skewness, kurtosis, and the $P-$value of Jarque-Bera test statistics (JB test) for testing normality of the series. Under the null hypothesis, the Jarque-Bera statistic has a chi-square distribution with

Table 4.2. Summary statistics: monthly industrial production growth

	Canada	France	Germany	Italy	Japan	UK	USA
Mean (%)	0.2444	0.1670	0.1383	0.1993	0.2307	0.1287	0.2585
Std. Dev.	1.3008	1.4602	1.7668	2.5404	1.5603	1.5487	0.7831
Skewness	0.5135	0.0580	0.2522	0.0060	-0.0555	0.1147	-0.8284
Kurtosis	6.7899	4.4421	10.2027	5.8543	3.2803	12.6923	7.3010
JB test	0.0000	0.0000	0.0000	0.0000	0.5115	0.0000	0.0000

Note: The hypothesis of normal distribution is rejected at the 5% (1%) level of significance if the $P-$value for the JB test is less than 0.05 (0.01).

two degrees of freedom. When the reported probability for the Jarque-Bera statistic is small, the null hypothesis of a normal distribution is rejected.

As seen in Table 4.1, the lowest mean excess return is found Italy (0.3368) and the highest is found in France (0.6583). However, the highest standard deviation in excess return is observed in Italy (7.3216) whereas the lowest one is found in the USA (4.4140). The null hypothesis of normal distribution in excess equity return is rejected for every country.

Turning to the descriptive statistics of the growth of industrial production in Table 4.2, the highest mean is recorded in the USA (0.2585) and the lowest is recorded in the UK (0.1287). The standard deviation in the growth of industrial production is the lowest in the USA (0.7831) and the highest in Italy (2.5404). Interestingly, the null hypothesis of normal distribution in the growth of industrial production is rejected for all countries except Japan.

4. Empirical Results

In table 4.3 we show the maximum likelihood estimates of the model given by equations (4.1) through (4.4) along with the transition probabilities defined by equation (4.5).[3] During the periods of contraction ($S_t = 1$), the mean growth of industrial production falls by 1.01% each month for the USA. The corresponding figures for Canada and France are 0.69% and 1.56%, respectively. The parameter μ_1 is not significant for Germany, Italy, Japan, or the UK at the 5% significance level. This indicates zero growth in industrial production during the downturns in these four countries.

Table 4.3. Parameter estimates: bivariate Markov switching heteroscedasticity model of output and equity

	Canada	France	Germany	Italy
a_1	-0.2258**	-0.4532**	-0.4022**	-0.4209**
	(0.0548)	(0.0518)	(0.0490)	(0.0484)
μ_1	-0.6939**	-1.5652**	0.2086	0.0711
	(0.1817)	(0.4272)	(0.1275)	(0.1759)
μ_2	0.4065**	0.2805**	0.1025	0.3245*
	(0.0627)	(0.0578)	(0.0823)	(0.1585)
σ_y^2	1.4603**	1.5176**	2.6177**	5.2768**
	(0.1170)	(0.1439)	(0.1972)	(0.3997)
α_0	0.7544**	0.6477*	0.4621	0.0714
	(0.2324)	(0.3105)	(0.2545)	(0.6830)
α_1	-0.0267	0.0530	0.0469	0.0783
	(0.0492)	(0.0503)	(0.0530)	(0.0562)
σ_1^2	82.0369**	150.6109**	56.2699**	83.2761**
	(18.5777)	(56.1026)	(9.9937)	(16.7178)
σ_2^2	15.2857**	29.4379**	14.8588**	26.4668**
	(1.4262)	(2.7362)	(2.1981)	(6.7655)
p_{11}	0.8303**	0.3555*	0.9486**	0.8691**
	(0.0638)	(0.1810)	(0.0321)	(0.0840)
p_{22}	0.9706**	0.9563**	0.9747**	0.8834**
	(0.0131)	(0.0207)	(0.0180)	(0.0612)
	Japan	UK	USA	
a_1	-0.2861**	-0.1747**	0.2210**	
	(0.0512)	(0.0528)	(0.0604)	
μ_1	0.1387	-0.2114	-1.0194**	
	(0.0808)	(0.3262)	(0.1727)	
μ_2	0.4314**	0.1708*	0.4068**	
	(0.1289)	(0.0763)	(0.0483)	
σ_y^2	2.2174**	2.3003**	0.4009**	
	(0.1677)	(0.1746)	(0.0329)	
α_0	0.5813*	0.7110**	0.8167**	
	(0.2365)	(0.2630)	(0.2194)	
α_1	0.0222	0.0275	-0.0643	
	(0.0577)	(0.0606)	(0.0540)	
σ_1^2	40.4185**	153.0646**	66.5685**	
	(4.9706)	(44.8695)	(16.8037)	
σ_2^2	7.1741**	20.2164**	14.1243**	
	(1.9494)	(1.9383)	(1.2891)	
p_{11}	0.9656**	0.8372**	0.7657**	
	(0.0201)	(0.0858)	(0.0797)	
p_{22}	0.9261**	0.9827**	0.9736**	
	(0.0350)	(0.0110)	(0.0107)	

Note: The parameters are described in the text. Standard errors are given in parentheses below the parameter estimates. Significance at the 1% level and 5% level is respectively indicated by ** and *.

The parameter μ_2 implies the mean growth rate during periods of expansion ($S_t = 2$). Japan has the highest value, at 0.43% monthly, but those of the USA and Canada trail behind by only a very small margin, at 0.41% each. The mean growth is also significantly positive for France (0.28%), Italy (0.32%) and the UK (0.17%). In the case of the last country, Germany, the model is unable to differentiate the mean growth of industrial production during the expansionary phase. Focusing on the variances in the excess stock return during the two phases, we note that the variance during the contraction of industrial production is much higher than that during expansionary phases in all of the G7 countries.. The ratio of σ_1^2 to σ_2^2 ranges from a low of 3.79 for Germany to a high of 7.57 for the UK. The ratios for Canada and the USA are 4.42 and 4.71, respectively. Thus, the unpredictable component of the excess stock returns in regime 1 can be concluded to have a variance threefold higher than that in regime 2 for all the G7 countries.

Next, we analyze the persistence of each regime based on the estimated parameters. Each regime appears highly persistent. The parameter p_{11}, the probability that a month of depression will be followed by another month of contraction, is 83% for Canada, 35% for France, 95% for Germany, 87% for Italy, 97% for Japan, 84% for the UK, and 77% for the USA. This regime will persist on average for $1/(1 - p_{11})$ months, i.e. for about 5.89 months for Canada, 1.55 for France, 39.53 for Germany, 7.64 for Italy, 29.07 for Japan, 6.14 for UK, and 4.27 for the USA. Interestingly, France had the lowest the average duration per economic depression, followed by the USA. Germany had the highest average duration.

Following the same argument as above, the parameter p_{22} expresses the probability that a month of expansion will be followed by another month of expansion. The values here are 97% for Canada, 96% for France, 97% for Germany, 88% for Italy, 93% for Japan, 98% for the UK, and 97% for the USA. An economic upturn will typically persist for $1/(1 - p_{22})$ months, i.e. about 34.01 for Canada, 22.88 for France, 39.53 for Germany, 8.58 for Italy, 13.53 for Japan, 57.80 for the UK, and 37.88 for the USA. When the G7 countries enter expansionary phases, apparently the duration of expansion will be briefest in Italy. The UK, in contrast, will continue expanding for the longest period.

The unconditional probability that an economy will be in state 1 at any time can be calculated as $\Pr(S_t = 1) = (1-p_{22})/(2-p_{11}-p_{22})$.[4] The maximum likelihood estimates in table 4.3 show this probability to be 14.8% for Canada, 6.3% for France, 67% for Germany, 47.1% for Italy, 68.2% for Japan, 9.6% for the UK, and 10.1% for the USA. Thus, Germany, Italy, and Japan are the mostly likely to be in the contractionary

Table 4.4. Diagnostics using standardized residuals from excess return equation

	Canada	France	Germany	Italy	Japan	UK	USA
LB (36)	0.444	0.862	0.606	0.570	0.352	0.604	0.702
LB^2 (36)	0.989	0.762	0.265	0.207	0.920	0.314	0.331

Note: LB (36) and LB^2 (36) are the Ljung-Box tests of order 36 using standardized residuals and squared standardized residuals respectively. For LB (36) test, $P-$value less than 0.05 (0.01) implies that the hypothesis of white noise is rejected at the 5% (1%) level of significance. Similarly, for LB^2 (36) remaining no ARCH effect is rejected for $P-$value less than 0.05 (0.01) at the 5% (1%) level of significance.

phase at any given time among the G7 countries. The results also show, however, that Italy will continue contracting for an average of only 7.64 months, a far shorter duration than the contractions in Germany and Japan. These peculiarities in the economies analyzed cannot be captured if only the excess stock return is modeled in a univariate setup. This becomes useful when we later analyze the concordance measure to describe the comovement of the stock markets.

We now analyze the model fitting statistics using the data in table 4.4. This table shows the diagnostics using standardized residuals from the excess return equation. The entries in the table are the $P-$values of the Ljung-Box $Q-$statistics (Ljung and Box, 1979). The $Q-$statistic at lag k is the statistic for the null hypothesis that there is no autocorrelation up to order k for residuals. LB (36) and LB^2 (36) are the $P-$values of the Ljung-Box tests of order 36 using standardized residuals and squared standardized residuals, respectively. We can infer from this table that the null hypothesis of no autocorrelation up to order 36 is not rejected for all cases. In other words, the hypothesis of whiteness of the residuals in both the level and the squared form cannot be rejected. This test supports our modeling approach in this chapter.

Figures 4.1 through 4.21 show each country's industrial production growth, each country's excess return, and the estimated filtered probability that the economy of each country is in a state of higher economic growth or a relatively lower variance state, i.e. $Pr(s_{t+1} = 2|w_t, w_{t-1}, \cdots, w_1)$ for $w_t = (y_t, x_t)'$. This corresponds to the probability that the industrial production would be in the higher output growth state at month $t+1$. This estimate of the volatility state probability is inferred from the joint distribution of the data under the assumption that an ergodic Markov chain drives its transition from month to month.[5]

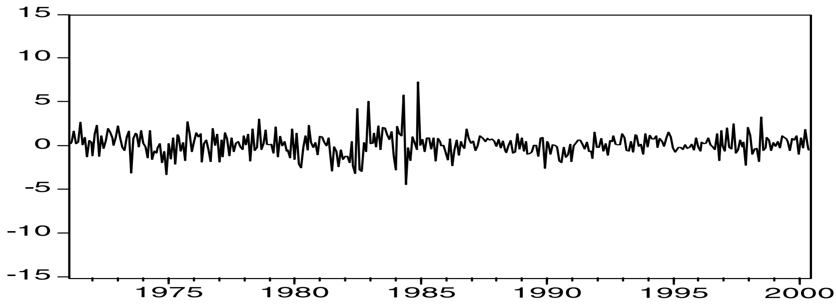

Figure 4.1. Industrial production growth: Canada

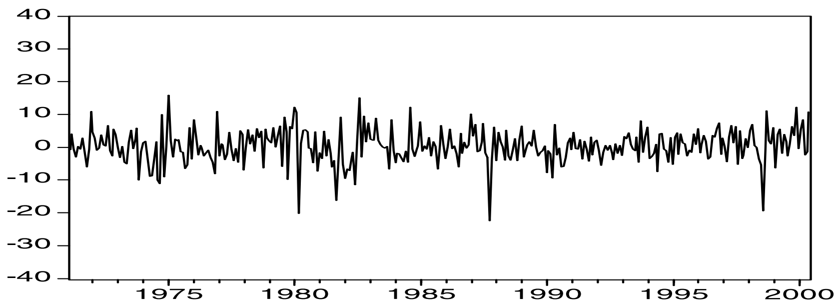

Figure 4.2. Excess return: Canada

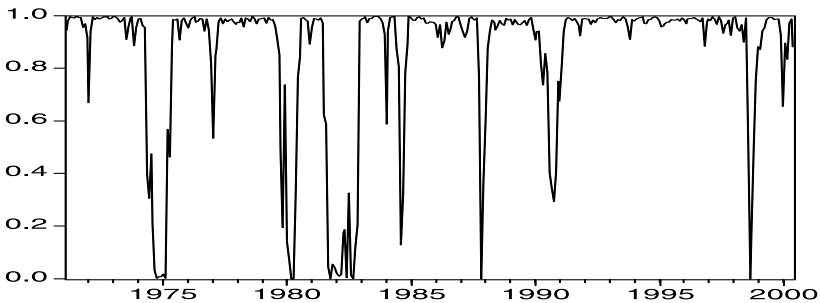

Figure 4.3. Estimated filtered probability, $\Pr(S_t = 2)$: Canada

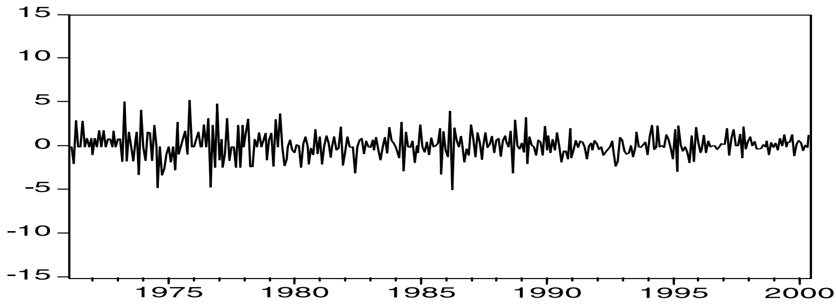

Figure 4.4. Industrial production growth: France

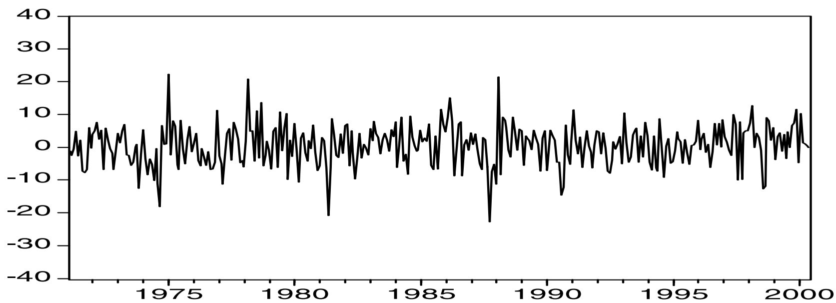

Figure 4.5. Excess return: France

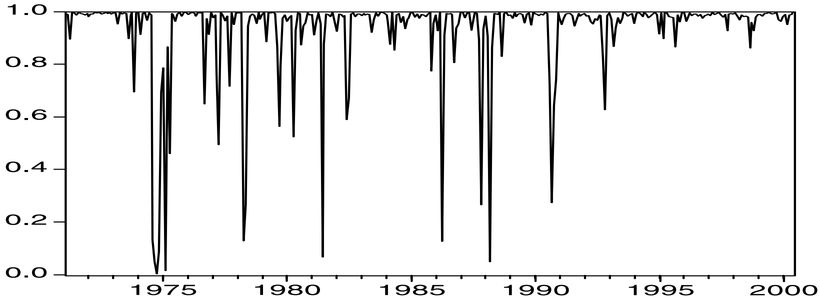

Figure 4.6. Estimated filtered probability, $\Pr(S_t = 2)$: France

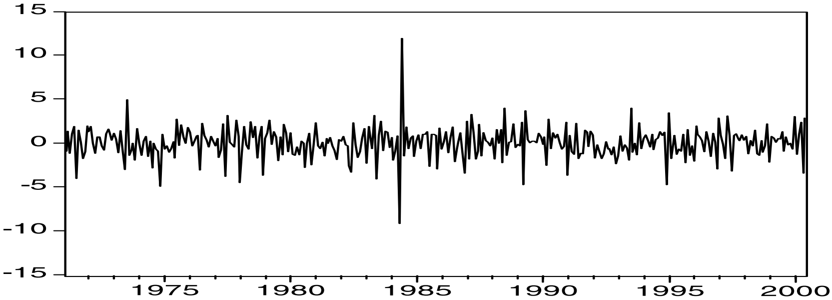

Figure 4.7. Industrial production growth: Germany

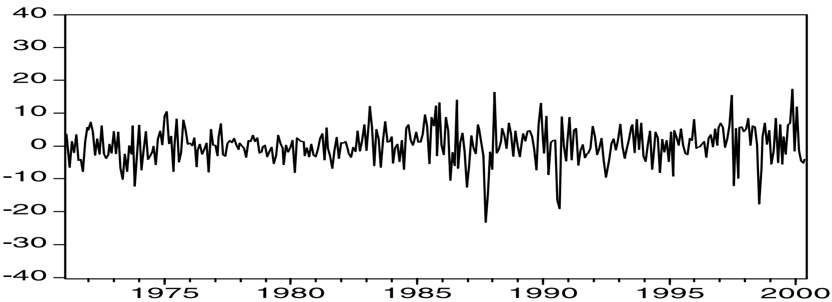

Figure 4.8. Excess return: Germany

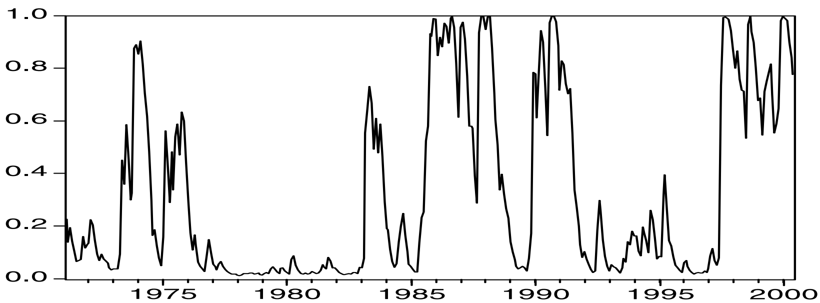

Figure 4.9. Estimated filtered probability, $\Pr(S_t = 2)$: Germany

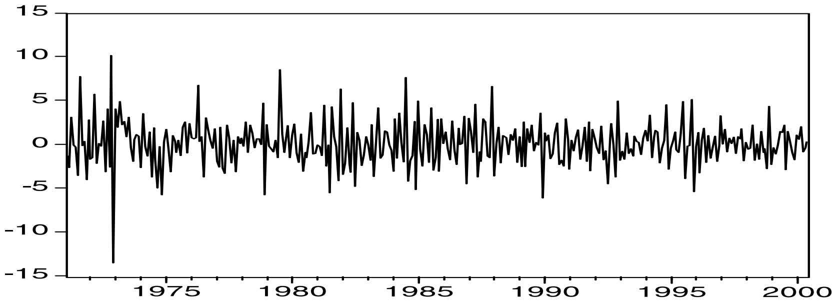

Figure 4.10. Industrial production growth: Italy

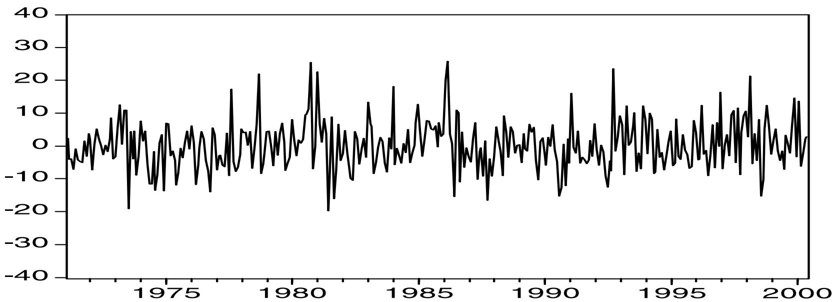

Figure 4.11. Excess return: Italy

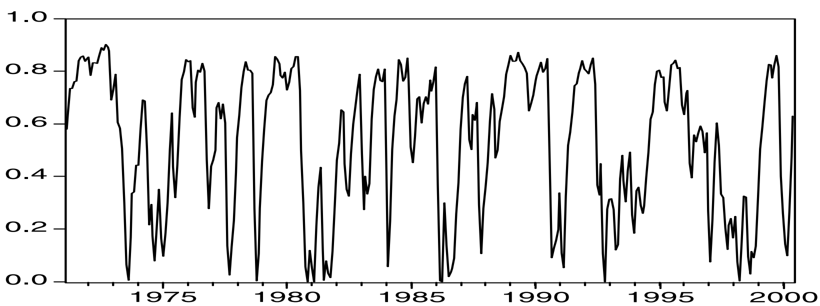

Figure 4.12. Estimated filtered probability, $\Pr(S_t = 2)$: Italy

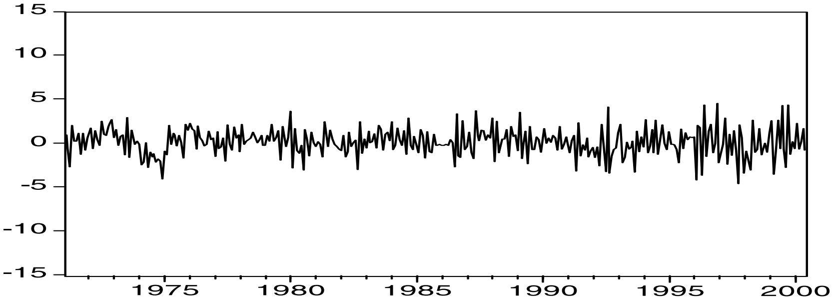

Figure 4.13. Industrial production growth: Japan

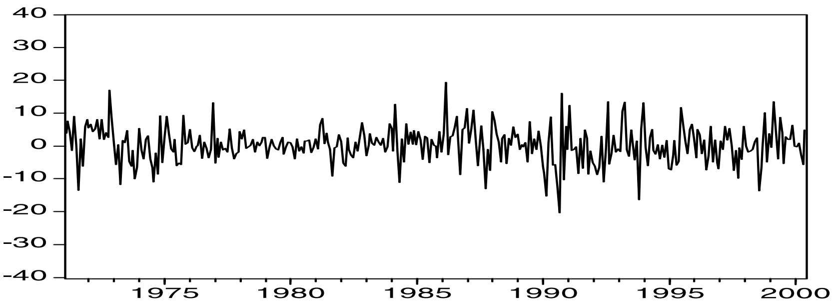

Figure 4.14. Excess return: Japan

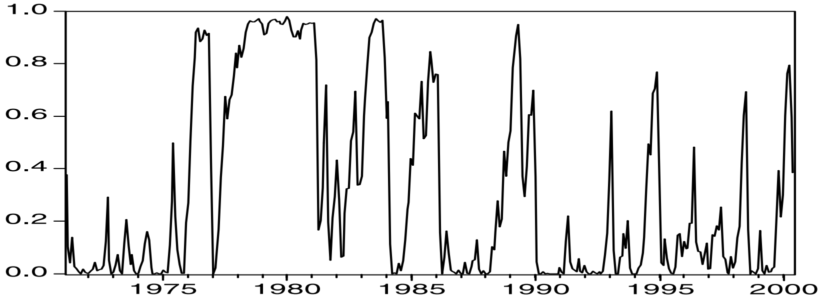

Figure 4.15. Estimated filtered probability, $\Pr(S_t = 2)$: Japan

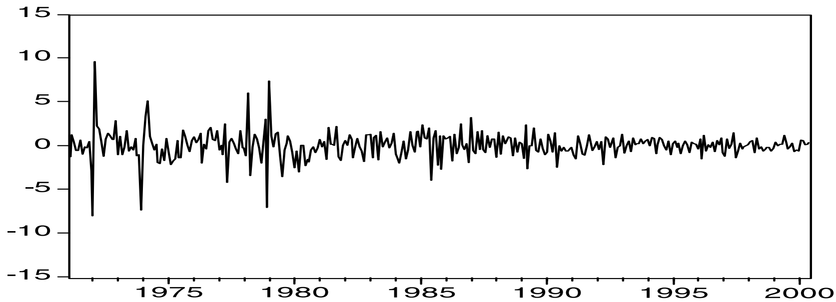

Figure 4.16. Industrial production growth: UK

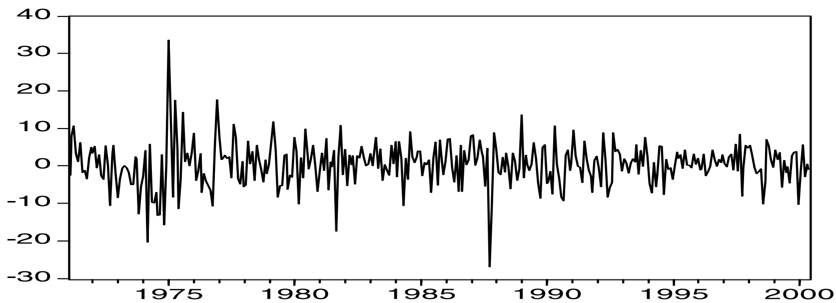

Figure 4.17. Excess return: UK

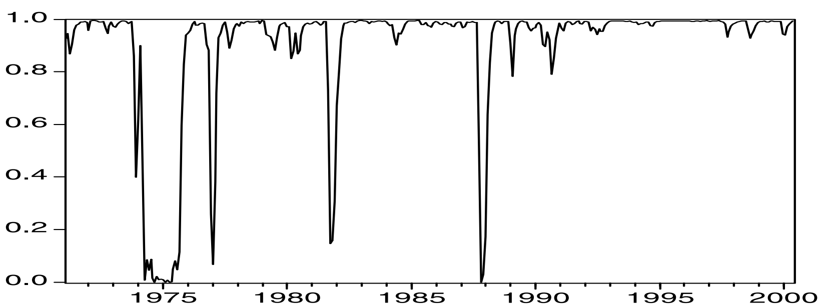

Figure 4.18. Estimated filtered probability, $\Pr(S_t = 2)$: UK

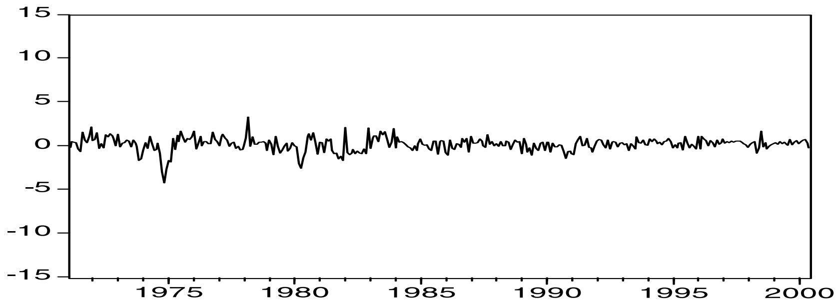

Figure 4.19. Industrial production growth: USA

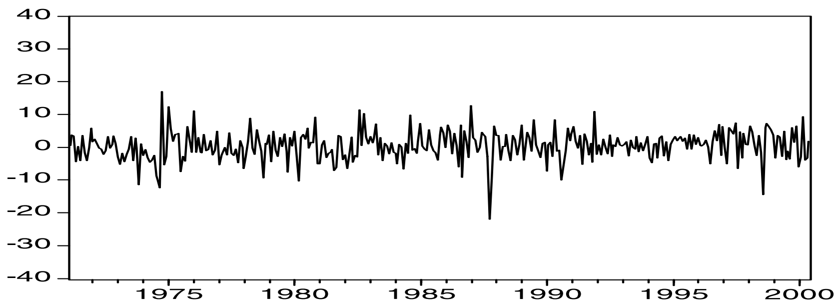

Figure 4.20. Excess return: USA

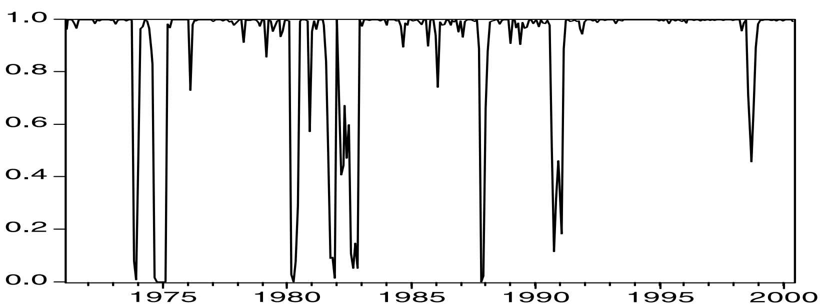

Figure 4.21. Estimated filtered probability, $\Pr(S_t = 2)$: USA

Table 4.5. Correlation between probabilities of expansion states

	Canada	France	Germany	Italy	Japan	UK	USA
Canada	1.0000						
France	0.3701*	1.0000					
Germany	0.0077	-0.0437*	1.0000				
Italy	0.2085*	0.2049*	-0.2543*	1.0000			
Japan	0.0627	0.0669	-0.2841*	0.1277*	1.0000		
UK	0.4802*	0.3584*	-0.0944	0.2128*	0.1956*	1.0000	
USA	0.7568*	0.3690*	-0.0593*	0.2291*	0.0742	0.4369*	1.0000

Note: Critical value for individual correlation at 5% level significance is 0.1043 (indicated with a *). This is calculated as $1.96/\sqrt{N}$, where N is the number of observations. Test statistic for significance as a group is 573.10. This is computed as $-2\ln(|R|)^{0.5N}$, where R is the correlation matrix. The critical value for the group statistic is obtained from a Chi-squared distribution with degrees of freedom $0.5p(p-1)$, where p is the number of series, and at 5% level this is 32.67.

These figures indicate that there is a close correspondence between the econometric inference and the actual data process.[6]

As the figures clearly show, the probabilities that the volatility states will move, i.e. the probabilities that the economies are in expansionary phases, differ from country to country. However, these probabilities appear to move together internationally during three periods, namely, the periods corresponding to the 1^{st} oil crisis of 1973, the oil crisis of 1979, and Black Monday in 1987. These are typically observed in the UK and USA. In contrast to Hamilton and Lin (1996), this model successfully indicates the high probability that the US economy will enter a recession just after Black Monday. Only Japan shows a high probability of economic downturn continuing into the 1990s. This may be due to the collapse of the stock market bubbles, which increased the likelihood that the Japanese economy would enter a recession in the 1990s. The probability of entering a recession is also high in Germany in 1990, the period of German reunification, and in the late 1990s, a period of recession.

Following from the above graphical analysis, we would like to quantify the possible comovements in terms of the concordance statistic discussed earlier. Before doing so, we check the sample correlations between these probability series given in the Table 4.5. Most of these correlations are statistically significant, indicating that most of these markets are in the same volatility state during the sample period. As mentioned before, the correlation statistics are influenced by the magnitude of the

Table 4.6. Concordance statistics, probabilities of expansion states or low volatility states

	Canada	France	Germany	Italy	Japan	UK	USA
Canada	1.0000						
France	0.8814*	1.0000					
Germany	0.3729	0.3220	1.0000				
Italy	0.6102*	0.5876**	0.3842	1.0000			
Japan	0.3672	0.3277	0.5085	0.5198	1.0000		
UK	0.8842*	0.9181*	0.3362	0.6243*	0.3814	1.0000	
USA	0.9266*	0.8983*	0.3333	0.6158*	0.3503	0.8955*	1.0000

Note: The critical values are computed following McDermott, and Scott (1999) under the assumption that a Brownian motion without a drift has generated the realizations of the state probabilities. These values are 0.5454, 0.5633 and 0.5944 for 10% (***), 5% (**) and 1% (*) level of significances, respectively. Reference probability 0.50 is used in concordance statistic.

Table 4.7. Concordance statistics, expected probabilities in three months for expansion states

	Canada	France	Germany	Italy	Japan	UK	USA
Canada	1.0000						
France	0.1225	1.0000					
Germany	0.3675	0.6866*	1.0000				
Italy	0.6125*	0.4131	0.3846	1.0000			
Japan	0.3647	0.6781*	0.5071	0.5242	1.0000		
UK	0.8832*	0.0855	0.3305	0.6268*	0.3789	1.0000	
USA	0.9259*	0.1054	0.3276	0.6182*	0.3476	0.8946*	1.0000

Note: The critical values are computed following McDermott, and Scott (1999) under the assumption that a Brownian motion without a drift has generated the realizations of the state probabilities. These values are 0.5454, 0.5633 and 0.5944 for 10% (***), 5% (**) and 1% (*) level of significances, respectively. Reference probability 0.50 is used in concordance statistic.

observation, whereas the concordance measure better captures whether the observations are in the same phase in some meaningful way. Both the Tables 4.6 and 4.7 use 50% probability as the reference point for the concordance measure, and Germany and Japan are clearly never in phase, as neither ever show a more than 50% probability of being in the same volatility state. This observation is strikingly different from the correlations measure in table 4.5, which shows that Germany and Japan are correlated with some of the other markets. Table 4.7 refers to the expected volatility state, as given by the equation (4.12), and the message is essentially the same as that of Table 4.6.

5. Conclusion

In this chapter we analyze the question of comovement in stock prices between the G7 economies with the help of a new statistical measure of concordance. This allows us to quantify which of these markets are in phase, in terms of sharing the same volatility state. This is done not only in the context of the volatility state as inferred from the data, but also from the expectations of the volatility state predicted by the model parameters.

The relevant volatility state of the economy is obtained from a bivariate model describing the joint distribution of the stock price and the industrial production data. Some researchers have previously reported a causal link between the growth in industrial production and the stock market volatility for US data. We extend that analysis to all the G7 countries using monthly data spanning the period from January 1971 through June 2000. The basic idea of the model originates from the work of Hamilton and Lin (1996), who show that the anticipation of an economic downturn affects stock market returns before the industrial output actually starts to decline. In the present study we modify their approach and take the view that most of the ARCH effect fades away at the monthly frequency. This was suggested by Hamilton and Susmel (1994), and we model the conditional heteroskedasticity directly in the Markov switching framework.

This framework does a credible job in capturing the behavior of the G7 economies over the sample period. The model diagnostics support our approach and allow us to quantify the differences in stock market volatility during the contractionary and expansionary phases of the economies. We are also able to quantify the likelihood that an economy will be in a contractionary or expansionary phase, as well as the average duration of these phases. We find that each phase is highly persistent among the G7 countries, although there are some variations. We also identify a relatively high unconditional probability that the economies of Germany, Italy and Japan will be in recession.

The concordance measure of the comovement suggests that with the exception of Germany and Japan, the economies in the G7 countries display statistically significant comovement over the sample period. The same is true in terms of the expectation of the volatility state as well.

This result differs from that obtained from simple sample correlations of the probability series.

Notes

1 Glosten, Jagannathan, and Runkle (1993) propose a variation of standard GARCH model to capture the asymmetries in stock return and in particular in stock index return. Hamilton and Lin (1996) also incorporate such asymmetries in their model of the equity excess return. But since their results indicate that such a parameter is statistically insignificant we do not include this effect in our model of excess return, thereby reducing the number of parameters to be estimated.

2 In order to make the algorithm operational we need starting values of the state probabilities and a mechanism to update it from observation to observation. Starting values of the state probabilities are derived following the procedure in Hamilton (1994, p. 693), Kim and Nelson (1999, p. 70). The probability updating process is described in Hamilton (1994, p. 692). The maximization of the likelihood function can be carried out by any suitable numerical method. As a by-product of this algorithm we also get the filtered probabilities of the states i.e. probability of the state occurring given the information up to that point in time. However, once the parameters are estimated, it is possible to utilize the whole sample information to infer the probabilities of the states and this referred to as the smoothed probabilities. The algorithm is described in Hamilton (1994, p. 694). We have implemented this algorithm in Gauss. Although the code for Hamilton and Lin (1996) is available from the journal, we chose to write the program in the most recent version of the Gauss package. Besides, our model is somewhat different from that in Hamilton and Lin (1996).

3 The estimation results for the USA data closely match those of Model C of Hamilton and Lin (1996). The apparent differences can be attributed to the different sample coverage, as well as the differences in the excess equity return equation.

4 See Hamilton and Lin (1996).

5 See Hamilton (1994, p. 681)

6 We could not obtain the business cycle peak and trough dates for all G7 countries.

APPENDIX 4.A
Data

Industrial Production Index:

Industrial production index is obtained from the International Financial Statistics of International Monetary Fund.

Canada: 15666..CZF...
France: 13266..CZF..
Germany: 13466..CZF..
Italy: 13666..CZF..
Japan: 15866..CZF.
UK: 11266..CZF...
USA: 11166..CZF...

Stock Returns:

Nominal stock returns are obtained from the Morgan Stanley Capital International Index.

Nominal interest rate:

Nominal interest rate is obtained from International Financial Statistics of International Monetary Fund.

Canada: 15660C..ZF..
France: 13260C..ZF...
Germany: 13460B..ZF...
Italy: 13660B..ZF...
Japan: 15860B..ZF...
UK: 11260C..ZF...
USA: 11160C..ZF...

Excess return

The excess return is calculated as follows:

$$\text{Excess return (\%)} = \text{stock return (\%)} - \frac{\text{interest rate (percent per annum)}}{12}$$

This page intentionally left blank

Chapter 5

LINKING INFLATION AND INFLATION UNCERTAINTY

1. Introduction

1.1 Inflation and Inflation Uncertainty

Macroeconomists have long been studying the impacts of inflation on various economic activities. Unanticipated inflation is generally thought to lead to a redistribution of wealth, although moderate inflation is not necessarily considered harmful (Cooley and Hansen, 1991). The exact impacts of inflation are often difficult to document. Some researchers examine the relationship between inflation, inflation uncertainty, and growth. In one study, Judson and Orphanides (1996) investigate this link using a cross-country panel data approach over a long period of time. According to their results, inflation volatility is negatively correlated with income growth across the inflation level, time, and type of country. They also find that the level and volatility of inflation are independently significant in influencing growth. In other words, high inflation is detrimental to growth above a certain level of inflation, while volatile inflation is correlated to lower growth at all levels of inflation.

Hess and Morris (1996) take a slightly different perspective. They attempt to establish the long-run costs of moderate inflation. Their hypothesis holds that anti-inflationary policies are costly in the short-run. In response, we question whether it is costly to allow low inflation to rise? Their article identifies three potential consequences of inflation, namely, inflation uncertainty, real growth variability, and relative price volatility. Using long-run data from countries with low to moderate inflation, they show that rising inflation is associated with higher values for all three of these consequences. In this context, there seem to be long-run benefits in not allowing inflation to rise even above low levels.

Holland (1995), on the other hand, attempts to clarify whether an increase in inflation increases inflation uncertainty. The author stresses that a positive association between the inflation rate and inflation uncertainty implies a particular temporal ordering. Based on the results of three different tests of temporal ordering, Holland concludes that an increase in the rate of inflation Granger-causes an increase in inflation uncertainty.

The preceding discussions indicate the importance of understanding the relationship between inflation and inflation uncertainty. Grier and Perry (1998) explore this aspect further for G7 economies in a two-step procedure. After fitting the GARCH (and some variation of it) model to generate a measure of inflation uncertainty, they use the Granger causality approach to determine the relationship between average inflation and inflation uncertainty. They use the autoregressive-GARCH(1,1) model for inflation at time t (π_t) as follows:

$$\pi_t = \beta_0 + \sum_{i=1}^{N} \beta_i \pi_{t-i} + \epsilon_t, \tag{5.1}$$

$$\sigma_{\epsilon t}^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \sigma_{\epsilon t-1}^2, \tag{5.2}$$

where $E_{t-1}(\epsilon_t) = 0$ and $E_{t-1}(\epsilon_t^2) = \sigma_{\epsilon t}^2$. Equation (5.1) is an autoregressive representation of the conditional mean of inflation. Equation (5.2) is a GARCH(1,1) representation of the conditional variance. In all G7 countries they found that inflation significantly raises inflation uncertainty, as predicted by Friedman (1977). On the other hand, there was less evidence to suggest that inflation uncertainty Granger-causes inflation. In the same paper, Grier and Perry also suggest that GARCH estimation is much better than other survey-based measures of volatility. GARCH estimation directly tests the statistical significance of time variation of conditional variance, whereas survey-based measures do not.[1]

Some researchers have also emphasized the impact of inflation on the distribution of relative prices in an economy. Relative price dispersion at time t (RPD_t) is typically measured as:

$$RPD_t = \frac{1}{n} \sum_{i=1}^{n} (\pi_{it} - \pi_t)^2, \tag{5.3}$$

where π_t is the aggregate inflation rate and π_{it} is rate of price change in the $i-$th individual commodity group.[2] For example, Grier and Perry (1996) report that menu cost models imply that higher trend inflation tend to increase price dispersion, whereas signal extraction models usually predict that increased inflation uncertainty raises relative price dispersion. These authors construct a bi-variate GARCH-M model of inflation and price dispersion to test these differing hypotheses. They use the following simplified bi-variate GARCH(1,1)-M model for inflation and relative price dispersion:

$$\pi_t = \beta_0 + \beta_1 \pi_{t-1} + \beta_2 \pi_{t-2} + \beta_3 \pi_{t-6} + \beta_4 \epsilon_{t-12} + \epsilon_t, \tag{5.4}$$

$$\sigma^2_{\epsilon t} = \alpha_0 + \alpha_1 \epsilon^2_{t-1} + \alpha_2 \sigma^2_{\epsilon t-1}, \tag{5.5}$$

$$RPD_t = \gamma_0 + \gamma_1 RPD_{t-1} + \gamma_2 \nu_{t-1} + \gamma_3 \pi^2_{t-1} + \gamma_4 \sigma^2_{\epsilon t} + \nu_t, \tag{5.6}$$

$$\sigma^2_{\nu_t} = \bar{\sigma}^2_{\nu}, \quad \forall t, \tag{5.7}$$

$$Cov_t = \rho \sigma_{\epsilon t} \bar{\sigma}_{\nu}. \tag{5.8}$$

where $E_{t-1}(\epsilon_t) = E_{t-1}(\nu_t) = 0$, $E_{t-1}(\epsilon^2_t) = \sigma^2_{\epsilon t}$, and $E_{t-1}(\nu^2_t) = \sigma^2_{\nu t}$. Equation (5.4) describes the inflation rate as a function of the first, second, and sixth lags of inflation and a 12-th order moving average term. Equation (5.5) is a GARCH(1,1) model of the conditional variance of inflation. Equation (5.6) shows that the relative price dispersion depends on the first-order moving average term, the first lag of the squared inflation rate, and the conditional variance of inflation. Equation (5.7) indicates the homoskedasticity of relative price dispersion, where $\bar{\sigma}^2_{\nu}$ is constant. Equation (5.8) is a simple constant correlation model of covariance between ϵ_t and ν_t, where ρ is the coefficient of correlation between ϵ_t and ν_t. Grier and Perry (1996) conclude that inflation uncertainty dominates trend inflation as a predictor of relative price dispersion.

1.2 Inflation Uncertainty and Markov Switching Model

Although GARCH models measure the conditional variation of volatility, the unconditional volatility remains constant. Evans and Wachtel (1993) remark on the paucity of studies investigating how uncertainty in the inflation regime may be the underlying source of the positive relation observed between inflation rates and inflation uncertainty. Conditional volatility models like GARCH ignore the structural instability due to changes in regimes. Evans and Wachtel develop a Markov switching model of inflation that decomposes uncertainty about future inflation into two components: a certainty-equivalent component and a regime uncertainty component. This decomposition allows them to examine Friedman's conjecture that uncertainty concerning regime changes depresses real economic activity. According to the authors, the model will seriously underestimate both the degree and impact of uncertainty unless the consequences of regime switches are much more closely considered in dealing with inflation uncertainty.

Kim (1993) and Kim and Nelson (1999) take this analysis further in four respects. First, since volatility or some measure of uncertainty remains essentially an unobserved component (UC), they extend the standard UC model to capture the observed heteroskedasticity in the data.[3] This is achieved by allowing an unobserved Markov process to change the volatility regime over the sample period. Unlike the case in GARCH models, the unconditional volatility in this scenario does not remain constant. This conforms with the observations of inflation regimes in Evans and Wachtel (1993), as well as the papers by Raymond and Rich (1992), who conclude that regime changes may constitute a single source of persistence in the conditional variance of the inflation series.

Second, the inflation series modeled by Kim (1993) and Kim and Nelson (1999) consists of both a temporary and permanent component. A paper by Ball and Cecchetti (1990) is the first to suggest a separation of this kind for inflation series. By establishing this separation, investigators can study the impact of uncertainty on each component.

Third, Kim (1993) and Kim and Nelson (1999) let another unobserved Markov process drive the mean inflation rate, thereby allowing them to test the link between inflation and its uncertainty over both short and long horizons.

Finally, since the resulting model is no longer conditionally Gaussian, as it is normally assumed in estimating UC models using Kalman filter, Kim (1993) develops an algorithm that results in a quasi-optimal filter. Kim also demonstrates a numerically efficient way to implement such a filter by an application to the US GDP deflator series. This holds an advantage over alternative approaches such as that of Grier and Perry (1998), as the entire estimation is performed in one step. For the US data, Kim reports a positive association between a higher inflation rate and long-run uncertainty.

This chapter takes advantage of the abovementioned advances in dealing with unobserved component models (e.g. Kim, 1993) and extends the inflation model of Evans and Wachtel (1993) by considering two different Markov processes, one driving the permanent component and one driving the transitory component. This approach allows more insight into the process of inflation forecast variance decomposition and its impact on the level of inflation at different horizons. We also extend the analysis to more recent data and cover four major countries, i.e. Germany, Japan, the UK, and the USA.

2. Empirical Technique

2.1 Markov Switching Heteroscedasticity Model of the Inflation Rate

As in Kim (1993) and Ball and Cecchetti (1990), we model the inflation series for a particular country as consisting of a stochastic trend component and a stationary component around the trend. Ball and Cecchetti adopt this approach of distinguishing between the long-term and the short-term components in order to resolve the empirical issue of the link between the level of inflation and its uncertainty. If we assume that the inflation series consists of a temporary component (white noise) and a permanent component (random-walk) then the investigation of this link is possible on both short and long horizons. We may think of the trend in money growth is responsible for determining the trend component and the other fiscal adjustments contributing to the stationary deviation from this trend.

The next step of the model development has to deal with the measure of uncertainty. As outlined earlier, GARCH type of model is not suitable for capturing the changes in regime as foreshadowed in Evans and Wachtel (1993). Although in GARCH type model conditional volatility changes, the unconditional volatility is, however, constant. When the inflation series is decomposed into two parts, it is possible to think of different uncertainties associated with these two components. In this respect the idea in Kim (1993) is an elegant way to approach this problem. Kim assumes that two different volatility regimes characterize the two components and these are driven by two different (unrelated) Markov stochastic processes. This concept brings in the required richness in the model, although it becomes more complex from the estimation point of view. In the section on results we compare two different GARCH models with the Markov switching model adopted in this paper. These models are clearly non-nested, and we compare these using a recently developed statistical procedure by Vuong (1989).

We focus in this article on individual inflation rate (π_t) modeling for each of the four countries and following the argument in Kim (1993) and Kim and Nelson (1999, p. 154), we assume the specification as follows:

$$\pi_t = T_t + \mu_2 S_{1,t} + \mu_3 S_{2,t} + \mu_4 S_{1,t} S_{2,t} + (h_0 + h_1 S_{2t}) e_t, \quad (5.9)$$

$$T_t = T_{t-1} + (Q_0 + Q_1 S_{1,t}) v_t, \quad (5.10)$$

where $e_t \sim N(0,1)$, and $v_t \sim N(0,1)$. The model specified in equation (5.9) and (5.10) is justified by Ball and Cecchetti (1990, p. 225). Inflation rate consists of a stochastic trend (random walk) component and a stationary component. Trend inflation for example is determined by trend money growth, and thus the permanent shock (v_t) captures events that change trend inflation. A negative shock occurs if the central bank creates a recession to disinflation. A positive shock occurs, when the central bank allows trend inflation to rise in accommodating a supply shock. The deviation of inflation from its trend is caused by monetary and other shocks. Thus, the transitory shock (e_t) captures events that affect inflation temporarily but do not affect the trend. Some examples are supply shocks that are not accommodated, fluctuations in velocity, and so on. The two unobserved Markov processes, $S_{1,t}, S_{2,t}$, determine the regime the economy is in at any point in time and these evolve independently of each other according to their own probability transition matrices (defined fully in the appendix B). The Markov process $S_{1,t}$ is associated with the trend component and $S_{2,t}$ is associated with the temporary component. The two Markov switching variables can take on values of 0 or 1. The value of 1 represents the high variance states of the corresponding component and 0 represent the low variance state.

The parameters μ_2, μ_3, μ_4 measure the shift in the mean depending on the state. From the formulation of the problem above it can be observed that μ_2 is associated with the high variance state of the trend (or the permanent) component. Similarly, μ_3 is associated with the high variance state of the stationary (or the transitory) component. We include μ_4 accounting for the shift in the mean inflation rate when both components are in a high variance state. Q_1 determines the extent of shift in variance during the high variance state of the trend component. Similarly, h_1 determines the increase in variance of the temporary component during its high variance state.

The estimation results will help us comment on the time series behavior of the inflation series for a particular country. The nature of the model given by the equations (5.9) and (5.10) suggest unobserved component modeling in a state space framework. The details of the estimation algorithm are quite involved and we have presented a simplified version in the appendix A. For a general understanding of the issues in unobserved component modeling Harvey (1991) is an excellent reference. The modification to the standard estimation of state space models needed due to the presence of the Markov switching variables are developed in Kim (1993).

2.2 Non-Nested Model Selection using Vuong Statistic

Vuong's (1989) test for selection of non-nested models is related to the classical likelihood ratio based test and uses Kullback-Leibler Information criterion to measure the closeness of the model to the true data

Table 5.1. Summary statistics

	Germany	Japan	UK	USA
Mean	0.7799	0.9650	1.7144	1.0065
Std. Dev.	0.6892	1.1855	1.4888	0.6327
Skewness	0.2672	1.4840	1.5726	0.9566
Kurtosis	3.0138	6.5674	6.3268	3.3451
JB Test	0.3973	0.0000	0.0000	0.0000

Note: The hypothesis of normal distribution is rejected at the 5% (1%) level of significance if the $P-$value for the JB test is less than 0.05 (0.01).

generating process. The test is directional and may be used to decide which model is performing better than the other one in explaining the data under consideration.

Consider two competing models with the conditional densities for the observations, y, with the explanatory variables, z, are given by, $f(y|z;\Theta)$, and $g(y|z;\Psi)$, where Θ and Ψ are the parameter vectors for the first and second model respectively. For non-nested models, under certain regulatory conditions, Vuong showed that the following test statistic,

$$\frac{\sum_{i=1}^{n}\{\ln f(y_i) - \ln g(y_i)\}}{\sqrt{n}\widehat{\omega}_n} \xrightarrow{D} N(0,1), \tag{5.11}$$

where n is the number of observations and $\widehat{\omega}_n$ given by,

$$\hat{\omega}_n^2 = \frac{1}{n}\sum_{i=1}^{n}[\ln f(y_i) - \ln g(y_i)]^2 - \{\frac{1}{n}\sum_{i=1}^{n}\{\ln f(y_i) - \ln g(y_i)]\}^2 \tag{5.12}$$

is the variance of the test statistic.

This statistic is easy to compute once the maximum likelihood estimation of the parameters has been carried out. The procedure needs to store all the conditional densities computed at the point of convergence. If the computed test statistic is higher than the chosen critical value we reject the hypothesis that the models are equivalent in favor of the model represented by the conditional densities, $f(y|z;\Theta)$.

3. Data

This paper uses the data on the quarterly price level in four major countries, i.e. Germany, Japan, the UK, and USA. The sample period

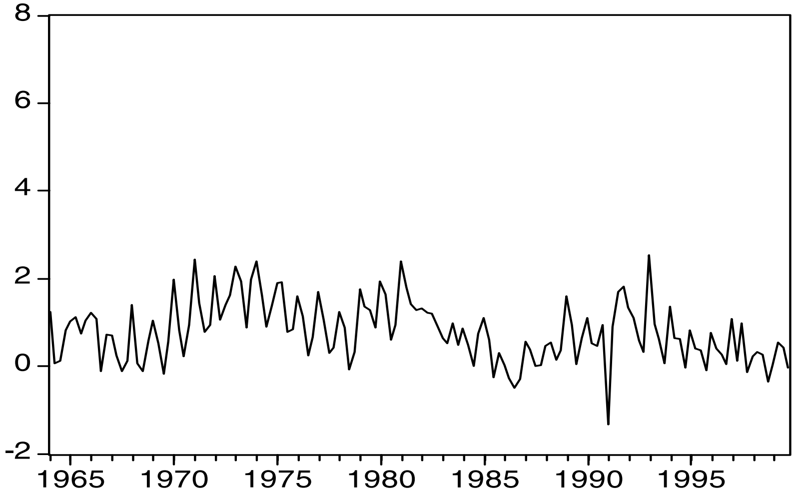

Figure 5.1. Inflation rate: Germany

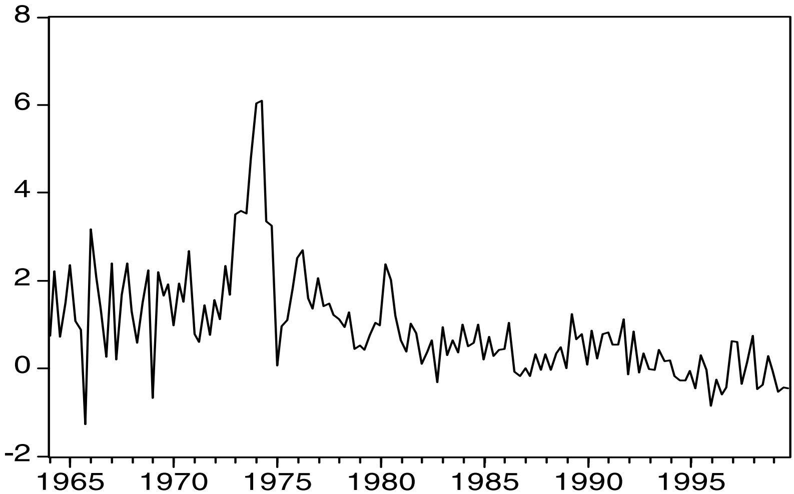

Figure 5.2. Inflation rate: Japan

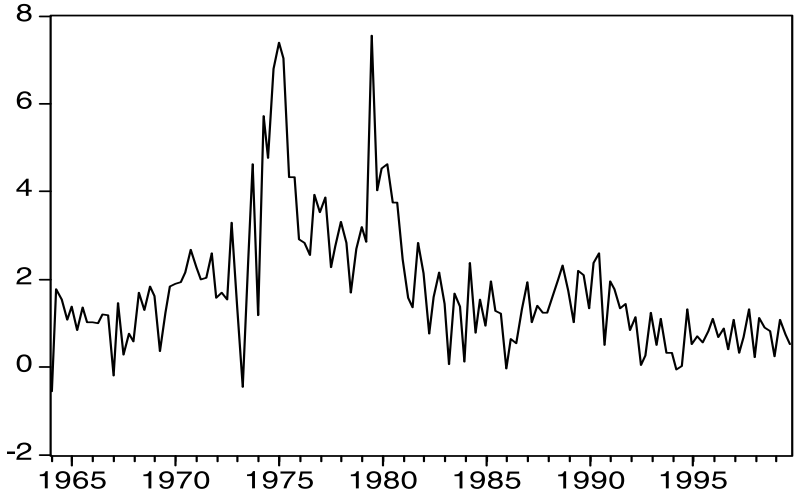

Figure 5.3. Inflation rate: UK

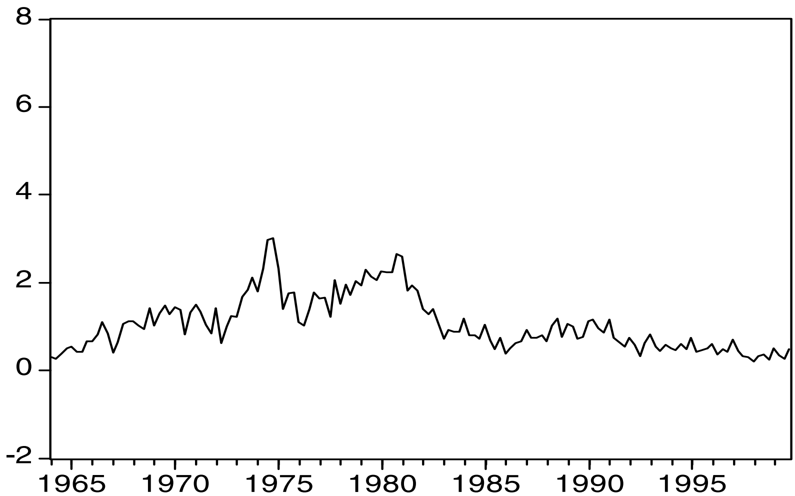

Figure 5.4. Inflation rate: USA

Table 5.2. Unit root test

	Test Statistic					
	CT		C		None	
Germany	-2.541	(0.308)	-2.089	(0.249)	-1.199	(0.210)
Japan	-2.769	(0.211)	-1.030	(0.741)	-1.163	(0.222)
UK	-2.749	(0.219)	-2.453	(0.129)	-1.512	(0.122)
USA	-2.633	(0.267)	-1.920	(0.322)	-0.888	(0.330)

Note: Numbers in parentheses are $P-$values. The hypothesis of a unit root is rejected at the 5% (1%) level of significance if the $P-$value for the test is less than 0.05 (0.01). CT shows that the test regression includes a time trend and a constant term. C shows that the test regression includes a constant. None shows that the test regression does not include deterministic term.

runs from the first quarter of 1961 through the fourth quarter of 1999. GDP deflators are used for Japan, the UK, and USA. The consumer price levels are used for Germany since GDP deflators were unavailable for this country during the corresponding period. The data sources include the Main OECD Economic Indicators for Japan, the UK, and USA, and the International Financial Statistics of the International Monetary Fund for Germany. The rate of inflation (π_t) is calculated as $\pi_t = (P_t - P_{t-1}) \times 100/P_{t-1}$, where P_t is the price level at time t. Thus, inflation rates are obtained for the period between the second quarter of 1961 and the fourth quarter of 1999. Figures 5.1 through 5.4 show the movements of inflation rate for each country.

Table 5.1 summarizes the inflation rate statistics in each country. This table shows the mean, standard deviation (Std. Dev.), skewness, kurtosis, and $P-$value of the Jarque-Bera test (JB test). The hypothesis of normal distribution is rejected at the 5% (1%) level of significance if the $P-$value for the JB test is less than 0.05 (0.01). As clearly shown in the table, the mean and standard deviation of inflation rates are relatively high for the UK, and relatively low for Germany. The hypothesis of normal distribution is rejected at the 1% significance level for all countries except Germany.

Table 5.2 shows the results of the augmented Dickey and Fuller (ADF) test.[4] This test is used to analyze whether inflation rate in each country has a unit root.[5] The hypothesis of a unit root is rejected at the 5% (1%) level of significance if the $P-$value for the ADF-test is less than 0.05 (0.01). Importantly, ADF test statistics support the assumption of the integrated inflation rate series for all countries examined in this chapter. This is important, since the model of decomposition adopted here to measure the impact of the high variance state on the mean inflation rate at both short and long horizons depends on this feature of the data.

4. Empirical Results

Table 5.3 shows the empirical results of Markov switching heteroscedasticity estimation for the sample period of the first quarter of 1964 through the fourth quarter of 1999.[6] The observations between the second quarter of 1961 and the fourth quarter of 1963 are used to obtain initial values for the filter. The elements of the transition probability matrix of the switching variable, $S_{1,t}$, given by p_{11} and p_{00}, are all statistically significant. Similarly, q_{11} and $q_{00}$are all statistically significant for the switching variable, $S_{2,t}$. These quantities may be interpreted in the context of the conditional variance of the inflation rate given by the equation (5.B.30) and (5.B.18) in the appendix. The first part of the equation (5.B.18) is contributed by the trend component and its persistence may be measured by $(p_{11} + p_{00} - 1)$.[7] Similarly, the second part of (5.B.18) is contributed by the temporary component and its persistence may be measured by $(q_{11} + q_{00} - 1)$.

Based upon the estimated parameters the persistence of the trend component contribution to the conditional variance is 0.9613, 0.7238, 0.9565, and 0.9671 for Germany, Japan, the UK and the USA respectively. Similarly, the persistence of the temporary component contribution to the conditional variance is 0.9205, 0.9817, 0.8461, and 0.9224 for Germany, Japan, the UK and the USA respectively. Therefore, the inflation series for and Japan behave differently from the other countries in the sample, with respect to the contribution from the two components to the persistence in conditional variance. A GARCH based variance model would not allow such insight into the variance process. Figures 5.5, 5.6, 5.7, and 5.8 show the relationship among the probability of high variance state for permanent shocks ($\Pr(S_t = 1)$), the probability of high variance state for transitory shocks ($\Pr(S_t = 2)$), and inflation rates for four countries.

The parameter μ_2 is the contribution to the mean inflation rate associated with the high variance state of the trend (or the permanent) component. Judging from the estimates of μ_2, it is found that high uncertainty about long-run inflation is associated with a significant positive shift in inflation for Germany and Japan. The positive association between the long-run uncertainty and level of inflation is consistent with Kim (1993)'s finding for US data for 1958 through 1990 although we do not obtain a significant estimate for USA in our dataset. As suggested

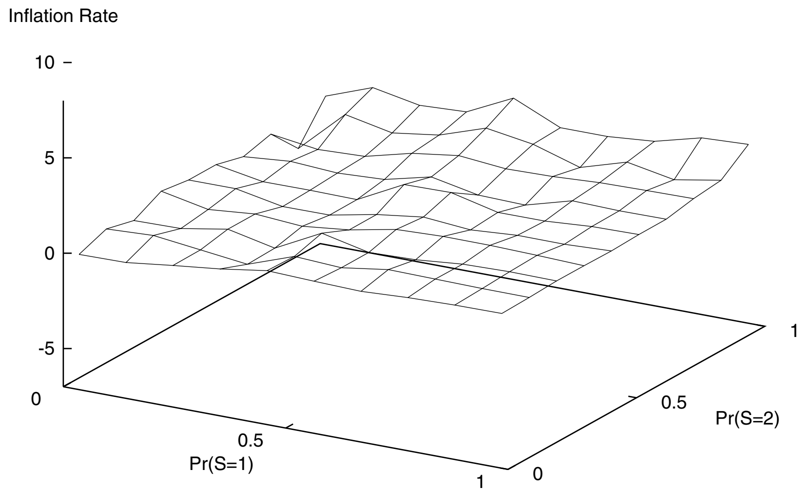

Figure 5.5. Relationship among the probability of high variance state for permanent shocks, $\Pr(S_t = 1)$, the probability of high variance state for transitory shocks, $\Pr(S_t = 2)$, and inflation rates: Germany

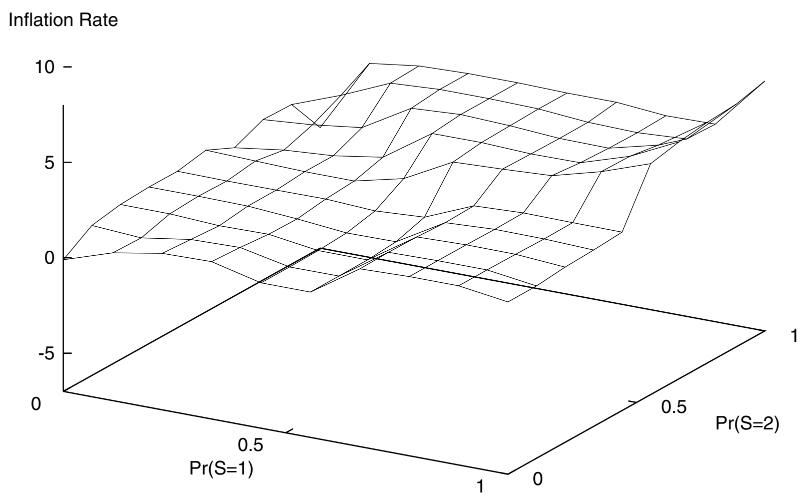

Figure 5.6. Relationship among the probability of high variance state for permanent shocks, $\Pr(S_t = 1)$, the probability of high variance state for transitory shocks, $\Pr(S_t = 2)$, and inflation rates: Japan

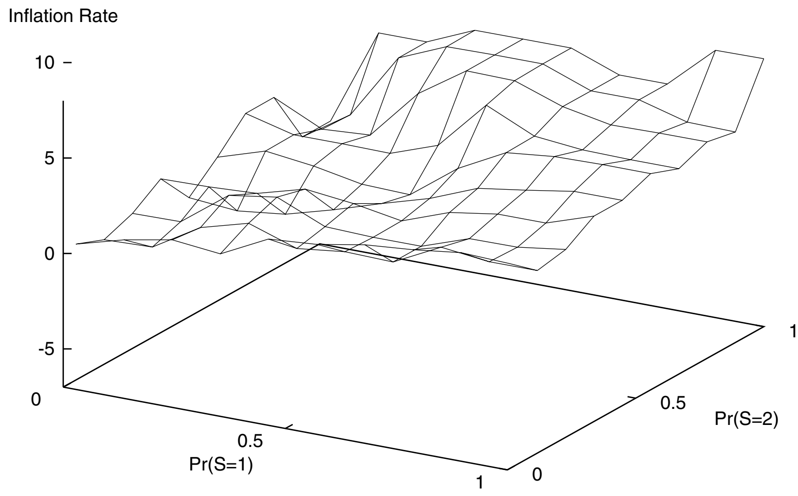

Figure 5.7. Relationship among the probability of high variance state for permanent shocks, $\Pr(S_t = 1)$, the probability of high variance state for transitory shocks, $\Pr(S_t = 2)$, and inflation rates: UK

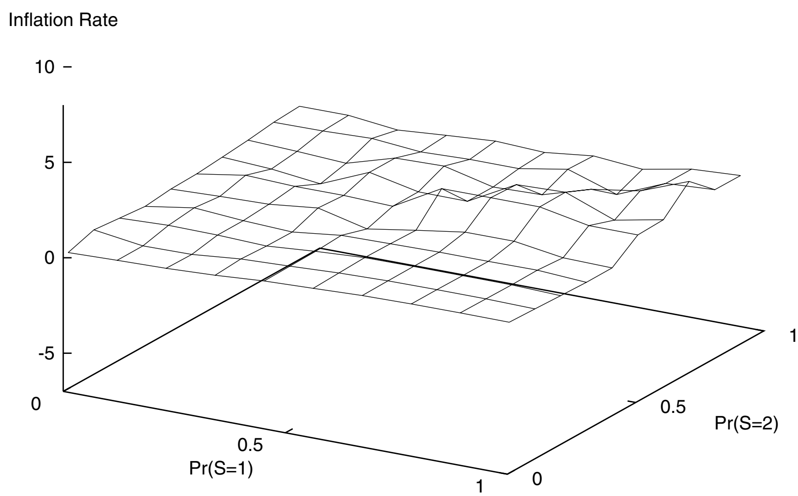

Figure 5.8. Relationship among the probability of high variance state for permanent shocks, $\Pr(S_t = 1)$, the probability of high variance state for transitory shocks, $\Pr(S_t = 2)$, and inflation rates: USA

Table 5.3. Parameter estimates: Markov switching heteroscedasticity model of inflation rate

	Germany	Japan	UK	USA
p_{11}	0.9698**	0.7397**	0.9653**	0.9808**
	(0.0276)	(0.1425)	(0.0436)	(0.0210)
p_{00}	0.9915**	0.9841**	0.9912**	0.9863**
	(0.0091)	(0.0120)	(0.0098)	(0.0131)
q_{11}	0.9636**	0.9819**	0.8674**	0.9559**
	(0.0276)	(0.0176)	(0.0821)	(0.0540)
q_{00}	0.9569**	0.9998**	0.9787**	0.9665**
	(0.0377)	(0.0009)	(0.0157)	(0.0365)
Q_0	0.0001	0.0958**	0.1419**	0.0382
	(0.0449)	(0.0313)	(0.0408)	(0.0235)
h_0	0.3891**	0.3592**	0.5383**	0.1241**
	(0.0615)	(0.0338)	(0.0445)	(0.0186)
Q_1	0.0068	0.0000	0.0125	0.2467**
	(0.2708)	(0.0010)	(0.1391)	(0.0530)
h_1	0.2044*	0.6054**	0.9266**	0.0562
	(0.0825)	(0.1091)	(0.2355)	(0.0398)
μ_2	0.6946*	1.1988**	0.9841	0.3663
	(0.3087)	(0.3405)	(0.6408)	(0.2258)
μ_3	0.5421**	0.4794	-0.2644	0.3049**
	(0.1785)	(0.4673)	(0.6343)	(0.1008)
μ_4	0.1594	1.6493**	2.5369**	-0.2236
	(0.3802)	(0.5798)	(0.7803)	(0.5485)

Note: The parameters are described in the text. Standard errors are given in parentheses below the parameter estimates. Significance at the 1% level is indicated by ** and at the 5% level is indicated by *.

in Kim (1993), if inflation increases above normal, it increases long-run uncertainty by making monetary policy less stable for Germany and Japan. This positive association between long-run uncertainty and level of inflation is also consistent with Ball and Cecchetti (1990) and Evans (1991).

Figures 5.9 through 5.12 show the probability of high variance state for permanent shocks for four countries. These figures reveal at least three facts about the relationship between the inflation rate and the probability of high variance state for permanent shocks. Firstly, the movements of probability of high variance state for permanent shocks differ from country to country. It is clear from the figures, however, there are two periods that variance moves together internationally. One corresponds to the 1^{st} oil crisis occurred in 1973 and the other corresponds to second oil crises occurred in 1979. These are typically observed in Japan and the UK. There are two peaks in the Japanese probability of high variance state for permanent shocks, which correspond to these two incidents. Secondly, there is clear positive correspondence between the inflation rate and the probability of high variance state for permanent shocks for Germany, and Japan. This leads to the significant positive estimates of μ_2 for these two countries. Thirdly, there is a possibility of structural change for each country. For example, the probability of high variance state of the permanent component is close to one between 1973 and 1982 for USA. An important motivation for considering the Markov-switching heteroscedasticity model is this possibility of structural change. As pointed out by Lastrapes (1989) and Lamoureux and Lastrapes (1990), failure to allow for regime shifts leads to an overstatement of the persistence of the variance of a series. This paper successfully takes into consideration this possibility of regime shift in the framework of Markov-switching model.

The parameter μ_3 is the contribution to the mean inflation rate associated with the high variance state of the stationary (or the transitory) component. Judging from the estimates of μ_3, it is also found that high uncertainty about short-run inflation is associated with a significant positive shift in inflation for Germany and the USA. Kim (1993) found the negative association between the short-run uncertainty and level of inflation for US data. Our findings support his results for Germany and the USA, but not for Japan and the UK. When inflation increases above normal, it increases short-run uncertainty by making monetary policy less stable for Germany and USA. This may result in more stable short-run monetary policy for these two countries, but not for Japan and the UK.

The parameter μ_4 is statistically significant only for Japan and the UK. This implies that uncertainty in both the trend component and the temporary component contributes to a significant shift in the inflation rate for these two countries. Kim (1993) also reports insignificant μ_4 for the US data analyzed there.

Figures 5.13 through 5.16 show the probability of high variance state for transitory shocks for four countries. At least two observations can be made about the relationship between the inflation rate and the probability of high variance state for the transitory shocks. Firstly, contrary to the case of permanent shocks, there are no clear international comovements about the probability of high variance state for transitory shocks. This is probably due to the fact that the source of the temporary shocks is within the national economy. Second, there is a possibility

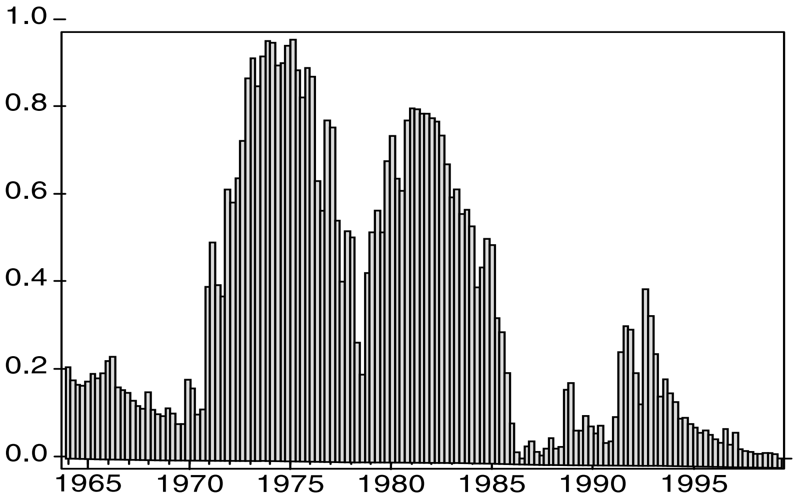

Figure 5.9. Probability of high variance state for permanent shocks: Germany

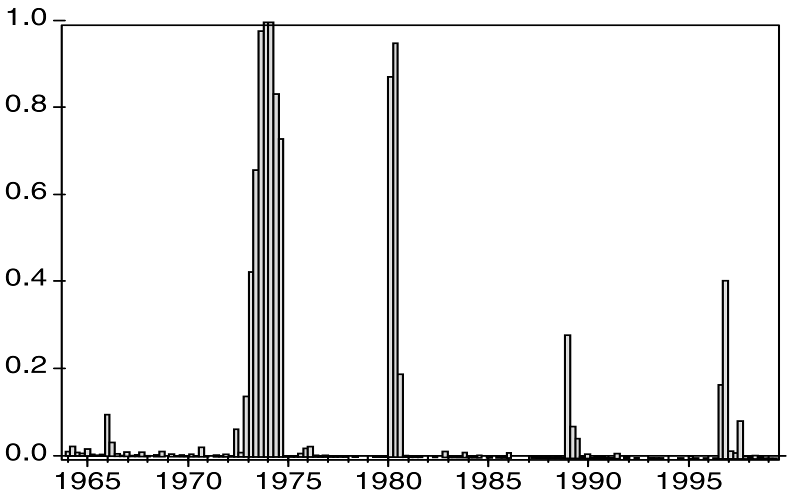

Figure 5.10. Probability of high variance state for permanent shocks: Japan

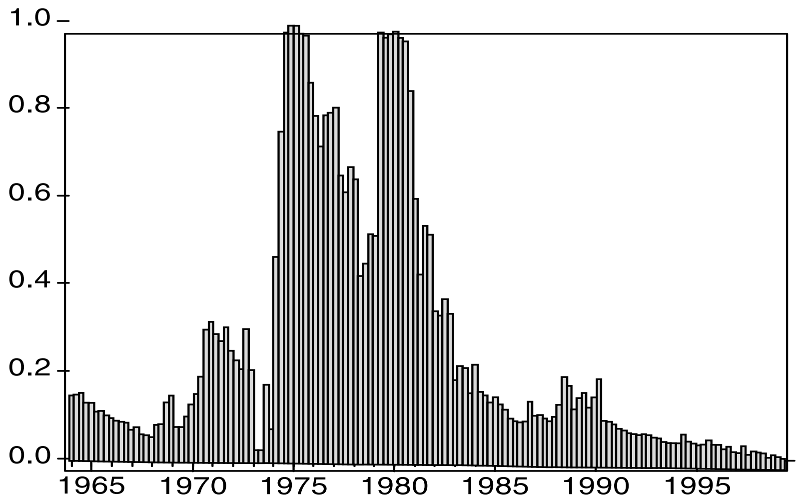

Figure 5.11. Probability of high variance state for permanent shocks: UK

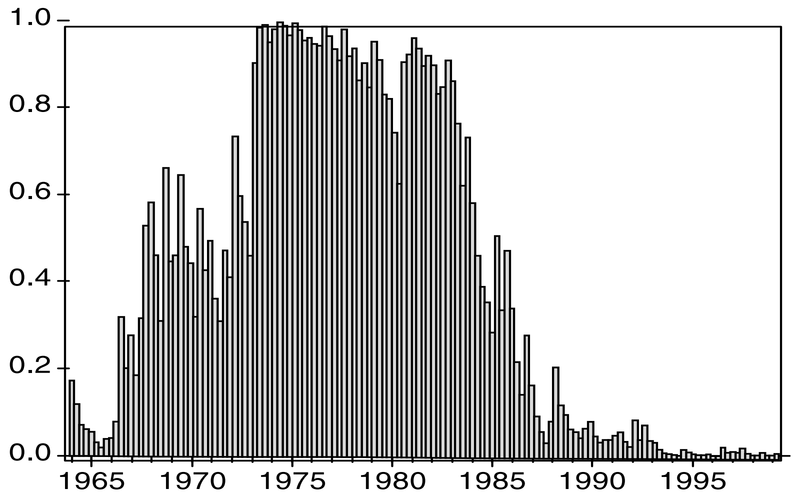

Figure 5.12. Probability of high variance state for permanent shocks: USA

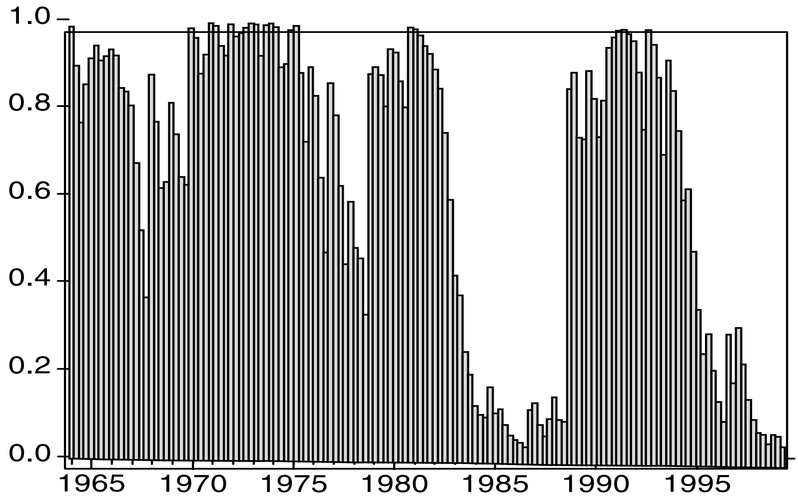

Figure 5.13. Probability of high variance state for transitory shocks: Germany

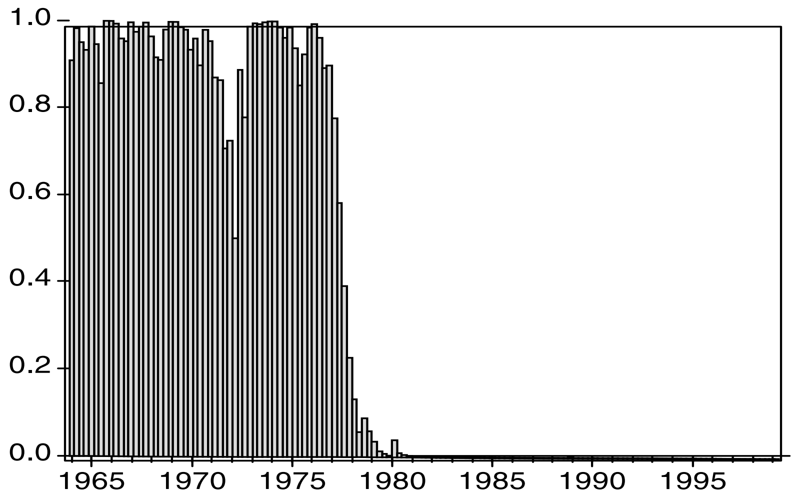

Figure 5.14. Probability of high variance state for transitory shocks: Japan

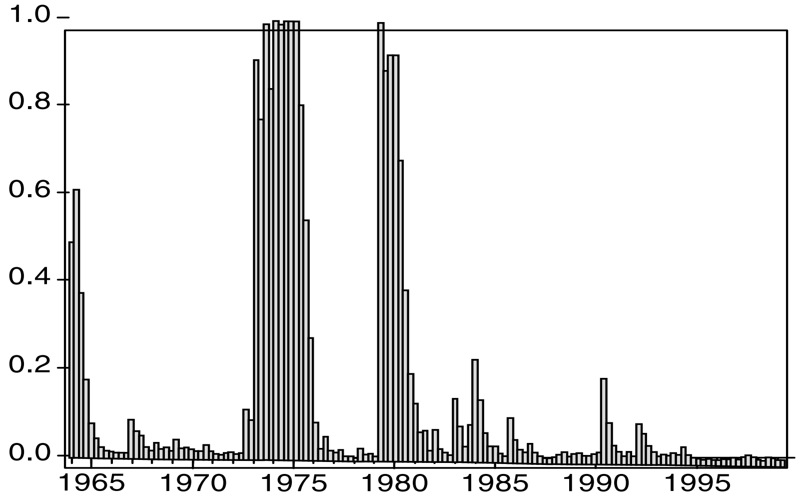

Figure 5.15. Probability of high variance state for transitory shocks: UK

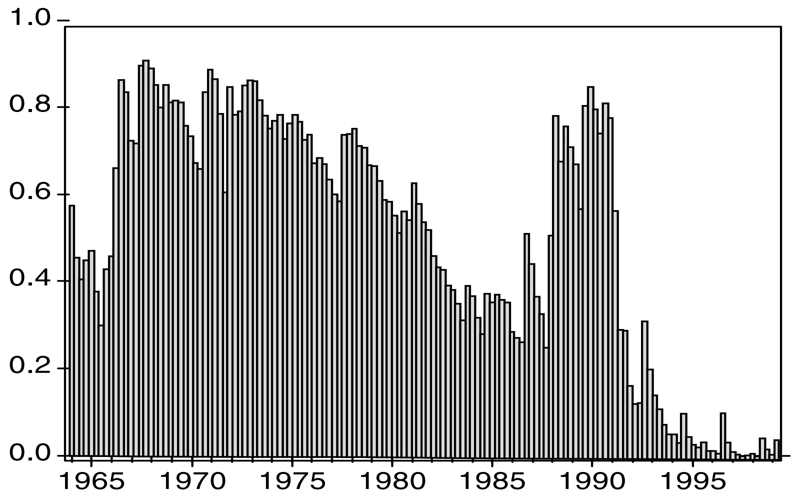

Figure 5.16. Probability of high variance state for transitory shocks: USA

Table 5.4. Residual diagnostics and model adequacy tests

	BDS	ARCH	KS Test	MNR	Recursive T
Germany	0.023	0.802	0.063	0.835	0.901
Japan	0.495	0.251	0.058	0.616	0.673
UK	0.492	0.374	0.065	0.574	0.824
USA	0.020	0.707	0.063	0.481	0.656

Note: Entries are $P-$values for the respective statistics except for the KS statistic. These diagnostics are computed from the recursive residual (standardized) of the measurement equation. BDS test checks for the i.i.d. assumption in the residual from the non-linear model. This test statistic has a N(0,1) distribution under the null i.i.d. hypothesis. The details of this statistic may be found in Brock et al (1996). The two parameters needed for this test are: embedding dimension = 2, and epsilon = 0.1*range. The ARCH test checks for no serial correlations in the squared residual up to lag 26. This test is applicable to recursive residuals as explained in Wells (1996, p. 27). MNR is the modified Von Neumann ratio test using recursive residual for model adequacy (Harvey, 1990, chapter 5). Similarly, if the model is correctly specified then Recursive T has a Student's t-distribution (Harvey, 1990, p. 157). KS statistic represents the Kolmogorov-Smirnov test statistic for normality. 95% significance level in this test is 0.113. When KS statistic is less than 0.113 the null hypothesis of normality cannot be rejected at the indicated level of significance.

of structural change for each country. For example, the probability of high variance is close to one between 1964 and 1977 for Japan. The Markov-switching model can take this regime shift into consideration.

Kim (1993) found that the ratio of high to low variances of permanent shocks is larger than that of transitory shocks for US, which means $Q_1 > h_1$. The entries in Table 5.3 suggest that Kim's finding is true for the USA but it is the reverse for Japan, and the UK. Kim also points out that the variance of permanent-shocks when low, denoted by Q_0, is close to 0. However, it is clear from Table 5.3, Q_0 is significantly different from 0 for all countries except for Germany and the USA. This suggests that infrequent permanent shocks to the price level does not necessarily account for most of the persistence in the price level.

Before proceeding further we would like to ascertain the performance of the model. In this respect, we analyze the residual from the model using a variety of diagnostics tests. We present these test results in Table 5.4. The BDS test for non-linear models of Brock et al (1996) checks for the i.i.d. assumption for the disturbances. The two parameters required for this test are to be supplied. The first one is the embedding dimension and for our model we set this to 2, given the size of our sample. The second one is the distance measure, epsilon, and we set this

Table 5.5. Regime classification measure (RCM)

	Germany	Japan	UK	USA
Trend	54.07	7.06	44.62	40.89
Temporary	45.13	11.12	18.79	67.07

Note: Ang and Bekaert (2002) introduce this RCM measure. Good regime classification is associated with a low value of the measure. A value of 0 implies perfect classification and a value of 100 implies no information about regimes is found in the data.

Table 5.6. Vuong statistics for non-nested model selection: Markov switching model against two different GARCH models

	AR(1)-GARCH(1,1)	MA(1)-GARCH(1,1)-in Mean
Germany	1.64	1.81
Japan	4.31	2.85
UK	2.89	1.92
USA	1.28	1.97

Note: The limiting distribution of Vuong statistic is N(0,1). The critical value at 5% level is 1.64. If the entry in the table is greater than the critical value then the Markov switching model is performing better in explaining the data generating process than the other model.

to 0.10 times the range of the data being examined.[8] The tests support the whiteness of the residuals for all the countries. The ARCH tests indicate no remaining heteroscedasticity in the residuals. Besides, the Kolmogorov-Smirnov tests support the normality of the residuals. These three tests overwhelmingly support the modeling approach adopted here and, therefore, the conclusions drawn are statistically meaningful.

In addition to the three tests just outlined above, we also include two additional tests particularly designed for recursive residuals produced by the state space system adopted in this study. The modified von-Neumann ratio tests against serial correlations in the residuals where as the recursive t-test is used to check for correct model specification. As the entries in Table 5.4 indicate the model applied to different country dataset perform extremely well in respect of these two tests. There is overwhelming support for the adequacy of the model in describing the inflation process involving the permanent and the transitory components.

In Table 5.5 we show the result regime classification measure recently introduced by Ang and Bekaert (2002). In addition to other diagnostic tests it provides a measure of the information of regime switches available in the data. It relies upon the estimated filtered probability of the states from the model that indicate the likelihood of a particular regime from which a particular data point is drawn. This is what underlies any regime-switching model as inferred by the econometrician. If there is insufficient information in the series then the regime classification will be weak. Based upon the notion that the true regime classification is a Bernoulli random variable, Ang and Bekaert develop the following measure:

$$RCM = 400 \times \frac{1}{T} \sum_{t=1}^{T} p_{S1,t} \left(1 - p_{S1,t}\right). \tag{5.13}$$

Similarly, the measure based on the switching variable S_2 can be defined. In the case of a perfect regime classification the inferred state probability for a particular data point would be either 0 or 1. This leads to the conclusion that the RCM measure is simply the sample estimate of variance of the Bernoulli variable. A value of 0 indicates perfect regime classification and a value of 100 indicates no regime switching information available in the data. The entries in Table 5.5 support to a varying degree our modeling approach based upon the Markov switching heteroscedasticity. The sharpest classification for both the components are obtained or Japan.

In Table 5.6 we present the results of Vuong test statistics to compare the Markov switching model with an AR(1)-GARCH(1,1) and also with a MA(1)-GARCH(1,1)-in-mean model. We include in the appendix sufficient details of this test and the associated references. This is a fairly recent development and has been successfully adopted by, for example, by Danielsson (1998), Ball and Torous (1999), and Smith (2002). This test statistic has a well-defined limiting distribution i.e. N(0,1). If the statistic is greater than the critical value at, say, 5% level then the Markov switching model captures the data generating process better than the other model. Thus, the entries in Table 5.6 demonstrate the dominance of the Markov switching model for all cases except for the USA compared with the AR(1)-GARCH(1,1) model. However, compared to MA(1)-GARCH(1,1)-in-mean model for the USA the Markov switching model performs better. This test result along with that of the RCM measure discussed above, we are confident that the modeling approach adopted here is a sound one.

We also derive the two components of the variance of inflation forecast for a particular forecast horizon. Evans and Wachtel (1993) de-

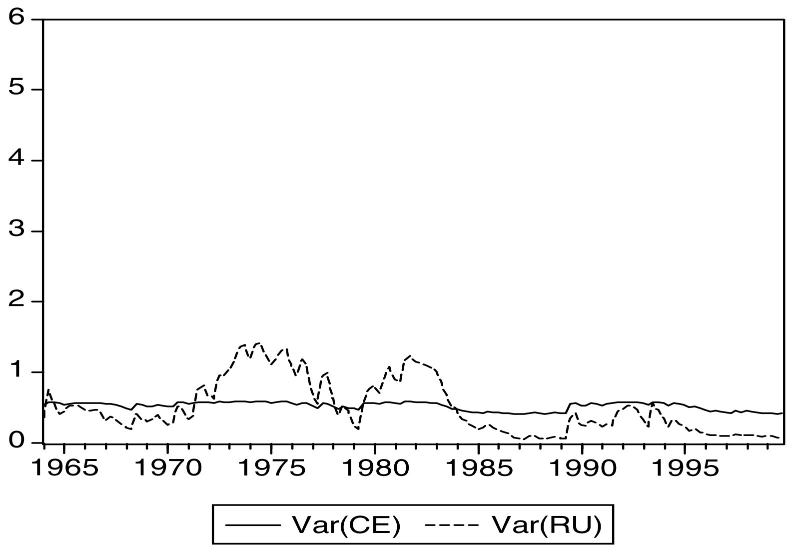

Figure 5.17. Components of forecast variance at $k = 2$: Germany

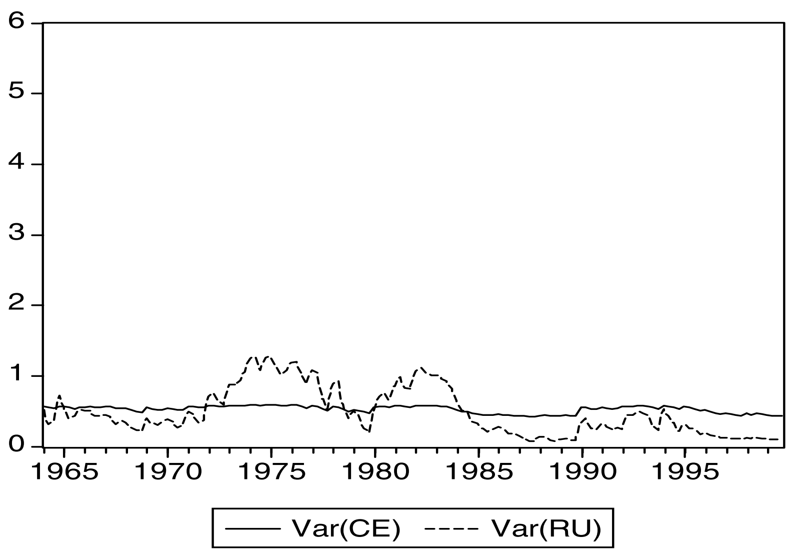

Figure 5.18. Components of forecast variance at $k = 4$: Germany

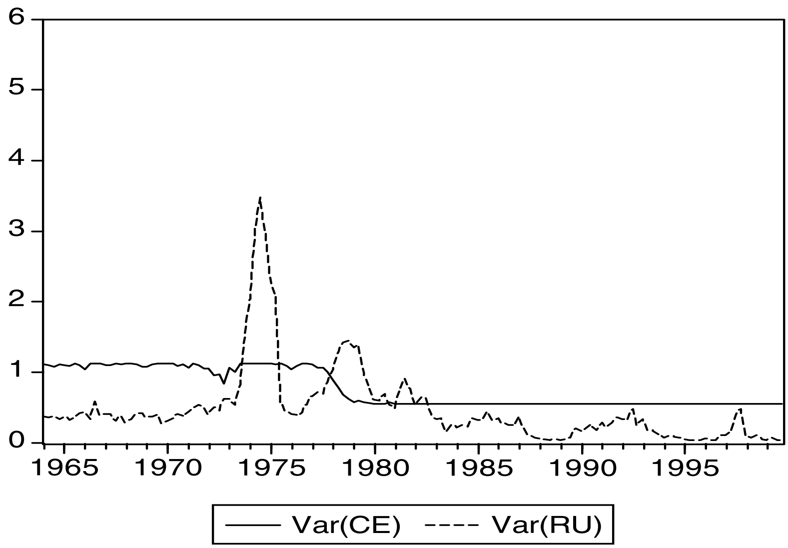

Figure 5.19. Components of forecast variance at $k = 2$: Japan

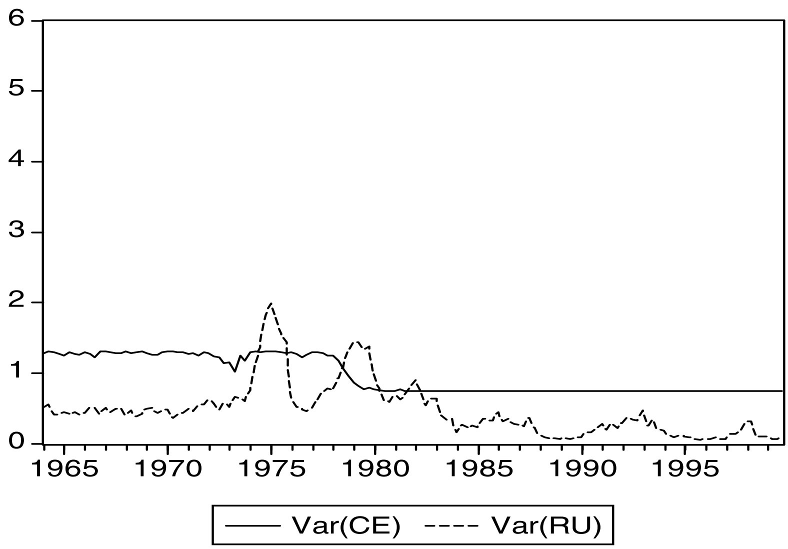

Figure 5.20. Components of forecast variance at $k = 4$: Japan

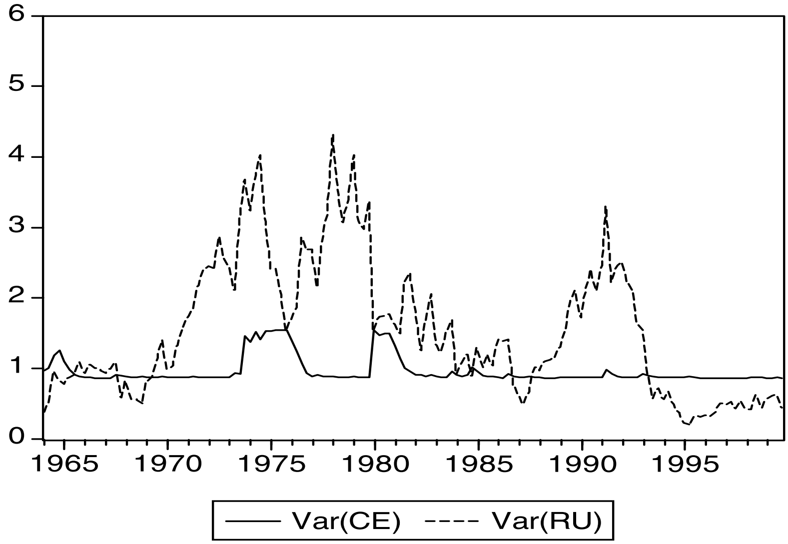

Figure 5.21. Components of forecast variance at $k = 2$: UK

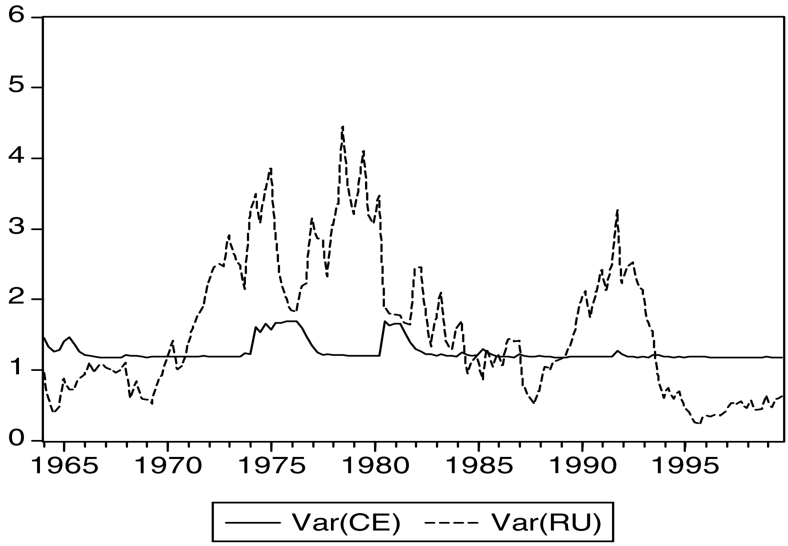

Figure 5.22. Components of forecast variance at $k = 4$: UK

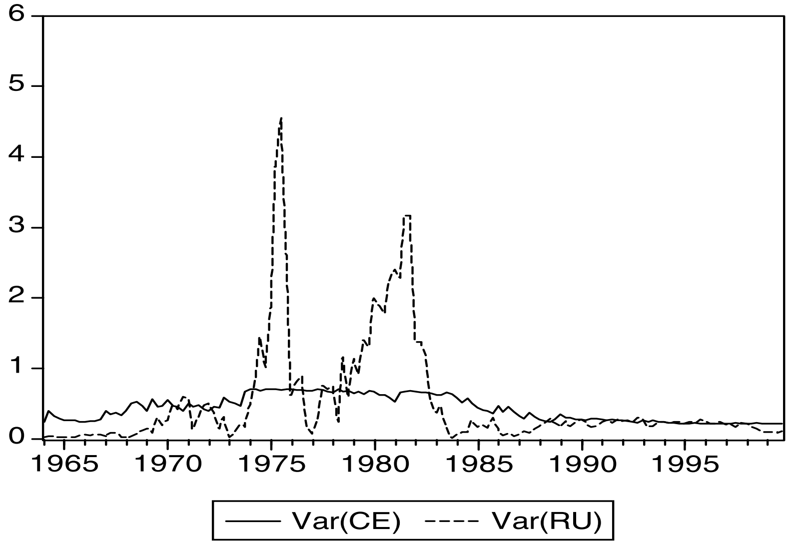

Figure 5.23. Components of forecast variance at $k = 2$: USA

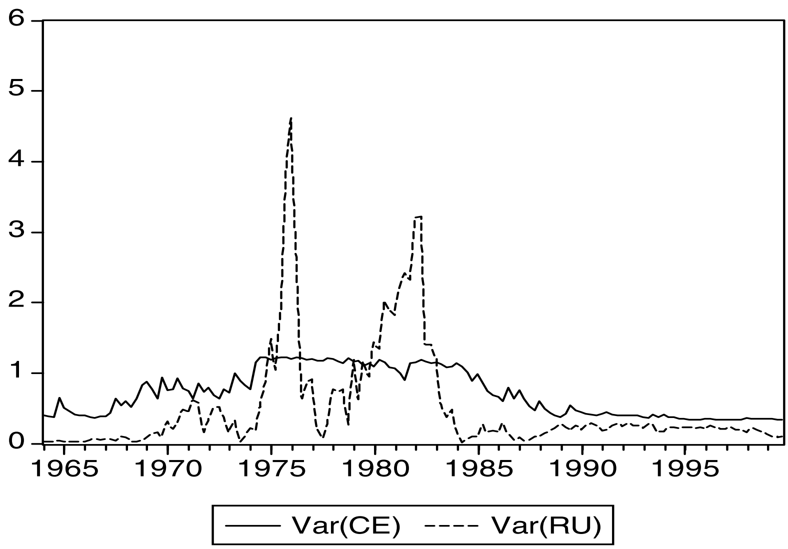

Figure 5.24. Components of forecast variance at $k = 4$: USA

scribe these two as the certainty equivalent (Var(CE)) and the regime uncertainty components (Var(RU)). They suggest that it is intuitive to analyze this way since there is uncertainty in the inflation process that is likely to be present at a future date. These two components help us analyze the interaction between these two parts of inflation uncertainty and the realized inflation for different horizon. Appendix C explains the details.

Figures 5.17 through 5.24 show the two components of variance forecast for forecast horizons of two ($k = 2$) and four quarters ($k = 4$) for each country. Generally, the figures show that inflation uncertainty increases at all horizons in the middle of 1970s and return to the low level in the middle of 1980s. This result is consistent with Evans and Wachtel (1993) and these periods correspond to the first and the second oil crises. It is interesting to see that the adjustment is found to be relatively fast for Japan. Inflation uncertainty return to the original low level in a few years for the case of Japan, whereas it takes more years for other countries.

5. Conclusion

In this chapter we adopt the modeling approach of Kim (1993), in which the mean inflation rate and volatility are subject to regime changes and driven by two independent Markov processes of order one. We apply this model to the inflation series for Germany, Japan, the UK, and the USA, and examine the interaction between inflation rate and its uncertainty over both the short- and long-run. The data sets examined in this study are also more up-to-date than those that were available in December 1999.

Grier and Perry (1998) fit the GARCH model to generate a measure of inflation uncertainty and use the Granger causality approach to determine the relationship between average inflation and inflation uncertainty. They found that inflation significantly raises inflation uncertainty, fulfilling the prediction by Friedman (1977), in four countries. However, the Markov-switching model estimation is superior to the GARCH model in at least two respects. Firstly, as emphasized by Lastrapes (1989) and Lamoureux and Lastrapes (1990), the Markov-switching model explicitly accounts for the possibility of regime shifts, whereas the GARCH model does not. Secondly, the Markov-switching model can decompose the shock into two components: permanent shock and transitory shock. Our specification for inflation rates are empirically supported by various diagnostics tests for each country. As pointed out by Cosimano and Jansen (1988) and Raymond and Rich (1992), regime changes seem to be an important source of persistence in the conditional variance of inflation. By introducing regime shifts in both mean and variance structures, we carried out a direct test of the link between inflation and its uncertainty over different time horizons for four countries.

Empirical results show that the relationship between inflation and inflation uncertainty depends on whether the shock is permanent or transitory. The relationship also differs from country to country. High uncertainty about long-run inflation is associated with a significant positive shift in inflation for Germany and Japan. This positive association between long-run uncertainty and the inflation level is consistent with Ball and Cecchetti (1990) and Evans (1991). High uncertainty about short-run inflation is associated with a significant positive shift in inflation for Germany and the USA. This is consistent with Kim (1993).

We adopt a more recent test for model selection in order to establish the superiority of our proposed model for inflation. Specifically, we test the non-nested model selection procedure developed by Vuong (1989). The Vuong test statistic is used to compare the Markov switching model with two alternative GARCH specifications that are clearly non-nested. The results of this testing and the RCM measure empirically support our modeling approach.

We also derive the two components of the variance of inflation forecast for a particular forecast horizon. These two components help us analyze the interaction between these two parts of inflation uncertainty and the realized inflation for different horizons. The figures generally show that inflation uncertainty increases at all horizons in the middle of 1970s and returns to a low level in the mid 1980s. This result is consistent with Evans and Wachtel (1993), and these periods correspond to the first and second oil crises. It is interesting to note that the adjustment is relatively fast for Japan. The inflation uncertainty drops back to its formerly low level in a few years in Japan, whereas those in the other countries subside much more gradually.

Notes

1 Grier and Perry (1998), p. 675, footnote 10.
2 Grier and Perry (1996, p.393).
3 For application to finance and economics, see Harvey (1991), Shumway and Stoffer (2000).
4 See Fuller (1976) and Dickey and Fuller (1979).
5 AIC is used to choose the order of augmented terms.
6 The programming for this model has benefited from the support provided through the website associated with the book by Kim and Nelson (1999).
7 See Kim and Nelson (1999, p. 164)
8 The programming for the BDS test is non-trivial, particularly when the speed of operation is important. We gratefully acknowledge the code available from W. Dechert, University of Houston, one of the authors of the BDS statistic.

APPENDIX 5.A
Data

GDP deflator is obtained from OECD Main Economic Indicator, whereas consumer price index is obtained from the International Financial Statistics of the International Monetary Fund. The series code of price level in each country is shown as follows:

Germany: 13464...ZF (International Financial Statistics, IMF),
Japan: JPNDEFLS (OECD Main Economic Indicator),
UK: GBRDEFLS (OECD Main Economic Indicator),
USA: USADEFLS (OECD Main Economic Indicator).

APPENDIX 5.B
Unobserved Component Model: Markov Switching Variance in State and Observation Equations

We discuss the problem of estimation of a dynamic linear system that is represented in term of a state space formulation. The general nature of such a system gives equation (5.B.1) as the measurement equation and the equation (5.B.2) as the state equation. In our problem for modeling inflation (π_t) series, the measurement vector is one-dimensional and the state vector may be constructed as consisting of two elements, e.g. the permanent component and the transitory component. In order to express the essential elements of the algorithm succinctly we express the model of the return with convenient notations as follows. However, first we consider the state space system (without the switching variables),

$$y_t = H\beta_t + M + e_t, \tag{5.B.1}$$

$$\beta_t = F\beta_{t-1} + v_t, \tag{5.B.2}$$

$$\begin{bmatrix} e_t \\ v_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R_t & 0 \\ 0 & Q_t \end{bmatrix} \right), \tag{5.B.3}$$

where equation (5.B.1) describes the observations made along with an error, and equation (5.B.2) describes the dynamic of the unobserved state(s). For example, if the permanent component consists of a random-walk and the transitory component is simply a white noise process then the elements of the system are: $y_t = \pi_t$, $H = 1$. The state vector consists of the permanent component (T_t), and $F = 1$ while ν_t represents the state noise. The measurement equation noise e_t will correspond to transitory components, which simply the white noise under the assumption. In the absence of any other specification, $M = 0$.

As given by equation (5.B.3) the variances of the two noise sources are given by R_t and Q_t and these are independent of each other. It is instructive to go through the recursive relations without the switching variables first. This implies that the variances of the measurement and the state equations are simply h_0 and Q_0 respectively. In the usual applications of state space system in finance and economics the unknown parameters are embedded in H and F in addition to the variance parameters. Using Kalman filter the prediction error form of likelihood function can be built which when maximized gives the estimates of these unknown parameters. Further details of this procedure can be found in Harvey (1991). Here we just summaries the main prediction and updating equations for the filter. Prediction equations move the system from period $(t-1)$ to t based upon all the information at $(t-1)$ and the updating equations modify the system parameters once an observation has been made.

Prediction:

$$\beta_{t|t-1} = F\beta_{t-1|t-1}, \tag{5.B.4}$$

$$P_{t|t-1} = FP_{t-1|t-1}F' + Q_t, \tag{5.B.5}$$

$$\eta_{t|t-1} = y_t - y_{t|t-1} = y_t - H\beta_{t|t-1} - M, \tag{5.B.6}$$

$$f_{t|t-1} = HP_{t|t-1}H' + R_t, \tag{5.B.7}$$

where, $P_{t|t-1}$ is the covariance of the state conditional on information at $(t-1)$, $\eta_{t|t-1}$ is the prediction error, $f_{t|t-1}$ is the conditional variance of the forecast error.

Updating:

$$\beta_{t|t} = \beta_{t|t-1} + K_t\eta_{t|t-1}, \tag{5.B.8}$$

$$P_{t|t} = P_{t|t-1} + K_tHP_{t|t-1}, \tag{5.B.9}$$

$$K_t = P_{t|t-1}H'f_{t|t-1}^{-1}, \tag{5.B.10}$$

where, K_t is known as the Kalman gain.

As required by our model in this paper we describe the extension of the basic state space model by assuming that an unobserved discrete-time, discrete state, first order Markov-switching process drives the variances. These two Markov processes are assumed to be independent. This helps account for the heteroscedasticity in the residual series. Application of ARCH type errors in state space system requires augmenting the state space as described in Harvey, Ruiz, and Sentana (1992). The basic idea for Markov switching processes to account for the heteroscedasticity originate from the works of Lamoureaux and Lastrapes (1990), Raymond and Rich (1992). In addition to this, since we are interested in exploring the impact of the Markov switching states on the mean inflation rate, we augment the model mean, M in equation (5.B.1) as $M = \mu_2 S_{1,t} + \mu_3 S_{2,t} + \mu_4 S_{1,t} S_{2,t}$.

With this background we define the variances as,

$$Q_{s_{1,t}} = Q_1 \Theta_{1,0t} + Q_2 \Theta_{1,1t}, \tag{5.B.11}$$
$$\Theta_{1,0t} = \begin{cases} 1 & \text{if} \quad S_{1,t} = 0, \\ 0 & \text{if} \quad S_{1,t} = 1, \end{cases}$$
$$\Theta_{1,1t} = \begin{cases} 1 & \text{if} \quad S_{1,t} = 1, \\ 0 & \text{if} \quad S_{1,t} = 0, \end{cases}$$

$$R_{s_{2,t}} = R_1 \Theta_{2,0t} + R_2 \Theta_{2,1t}, \tag{5.B.12}$$
$$\Theta_{2,0t} = \begin{cases} 1 & \text{if} \quad S_{2,t} = 0, \\ 0 & \text{if} \quad S_{2,t} = 1, \end{cases}$$
$$\Theta_{2,1t} = \begin{cases} 1 & \text{if} \quad S_{2,t} = 1, \\ 0 & \text{if} \quad S_{2,t} = 0, \end{cases}$$

where $S_{1,t}$ and $S_{2,t}$ are two independent (two state) Markov processes, which evolve according to the transition probability matrices defined below:

$$P_{S1} = \begin{bmatrix} p_{00} & 1 - p_{11} \\ 1 - p_{00} & p_{11} \end{bmatrix}, \tag{5.B.13}$$

$$P_{S2} = \begin{bmatrix} q_{00} & 1 - q_{11} \\ 1 - q_{00} & q_{11} \end{bmatrix}. \tag{5.B.14}$$

Comparing the above specification with that for standard ARCH type models we note that for the Markov switching variance model, where the unconditional variances are different in different states, past values of disturbances are not needed. The presence of state dependent variances in the process complicates the direct application of the prediction and updating equations described earlier. A simplification based on an approximation algorithm suggested by Kim (1993) greatly reduces the computational complexity. We describe this next.

Consider the realizations of the state variables at time $(t-1)$ and t as $S_{1,t-1} = m$, $S_{2,t-1} = n$, and $S_{1,t} = m'$ and $S_{2,t} = n'$, where $m, m' = 0, 1$, and $n, n' = 0, 1$ representing the possible combinations of states in the two consecutive time slots. Given these realizations, we can rewrite the prediction and updating equations as below:

Prediction:

$$\beta^{m,n}_{t|t-1} = F\beta^{m,n}_{t-1|t-1}, \tag{5.B.15}$$

$$P^{m,n,m'}_{t|t-1} = FP^{m,n}_{t-1|t-1}F' + Q^{m'}_t, \tag{5.B.16}$$

$$\eta^{m,n}_{t|t-1} = y_t - y^{m,n}_{t|t-1} = y_t - H\beta^{m,n}_{t|t-1} - M, \tag{5.B.17}$$

$$f^{m,n,m',n'}_{t|t-1} = HP^{m,n,m'}_{t|t-1}H' + R^{n'}_t, \tag{5.B.18}$$

Updating:

$$\beta^{m,n,m',n'}_{t|t} = \beta^{m,n}_{t|t-1} + K^{m,n,m',n'}_t \eta^{m,n}_{t|t-1}, \tag{5.B.19}$$

$$P^{m,n,m',n'}_{t|t} = P^{m,n,m'}_{t|t-1} + K^{m,n,m',n'}_t HP^{m,n,m'}_{t|t-1}, \tag{5.B.20}$$

$$K^{m,n,m',n'}_t = P^{m,n,m'}_{t|t-1}{}' H'(f^{m,n,m',n'}_{t|t-1})^{-1}. \tag{5.B.21}$$

This recursion in the filter produces $(2 \times 2 \times 2 \times 2)$ posteriors for $\beta^{m,n,m',n'}_{t|t}$ and $P^{m,n,m',n'}_{t|t}$ when moving from $(t-1)$ to t. Depending on the observations to consider this becomes an almost impossible task. Kim (1993) develops the following approximation where by taking appropriate weighted average over the states at $(t-1)$ this can be reduced to (2×2). We define the probability weighting as,

$$\Gamma^{m,n,m',n'} = \frac{\Pr(S_{1,t-1} = m, S_{2,t-1} = n, S_{1,t} = m', S_{2,t} = n'|\Psi_t)}{\Pr(S_{1,t} = m', S_{2,t} = n'|\Psi_t)}, \tag{5.B.22}$$

where Ψ_t is the information available at time t. The approximation for $\beta^{m,n,m',n'}_{t|t}$ is,

$$\beta^{m',n'}_{t|t} = \sum_{m=0}^{l}\sum_{n=0}^{l} \beta^{m,n,m',n'}_{t|t} \times \Gamma^{m,n,m',n'}, \tag{5.B.23}$$

and the approximation for $P^{m,n,m',n'}_{t|t}$ is,

$$\begin{aligned} P^{m',n'}_{t|t} = \sum_{m=0}^{l}\sum_{n=0}^{l} [P^{m,n,m',n'}_{t|t} \\ + (\beta^{m',n'}_{t|t} - \beta^{m,n,m',n'}_{t|t})(\beta^{m',n'}_{t|t} - \beta^{m,n,m',n'}_{t|t})'] \times \Gamma^{m,n,m',n'}, \end{aligned} \tag{5.B.24}$$

The equations (5.B.23) and (5.B.24) describe the nature of approximation applied to collapse the $(2 \times 2 \times 2 \times 2)$ posteriors to (2×2) posteriors with the help of the probability weighting factor. The probability terms necessary to achieve this can be obtained as follows:

$$\begin{aligned}\Gamma^{m,n,m',n'} &= \frac{\Pr(S_{1,t-1}=m,S_{2,t-1}=n,S_{1,t}=m',S_{2,t}=n'|\Psi_t)}{\Pr(S_{1,t}=m',S_{2,t}=n'|\Psi_t)}\\ &= \frac{\Pr(y_t,S_{1,t-1}=m,S_{2,t-1}=n,S_{1,t}=m',S_{2,t}=n'|\Psi_{t-1})}{\Pr(y_t|\Psi_{t-1})}\\ &= (\frac{\Pr(y_t|S_{1,t-1}=m,S_{2,t-1}=n,S_{1,t}=m',S_{2,t}=n',\Psi_{t-1})}{\Pr(y_t|\Psi_{t-1})})\\ &\quad\times \Pr(S_{1,t-1}=m,S_{2,t-1}=n,S_{1,t}=m',S_{2,t}=n'|\Psi_{t-1}),\end{aligned} \tag{5.B.25}$$

where as before $m, m' = 0, 1$ and $n, n' = 0, 1$. With the help of the forecast error in the prediction relations we can now construct the numerator of the last term of equation (5.B.25) as,

$$\begin{aligned}\Pr(y_t \mid & S_{1,t-1}=m,S_{2,t-1}=n,S_{1,t}=m',S_{2,t}=n',\Psi_{t-1})\\ &= \frac{1}{\sqrt{2\pi(f_{t|t-1}^{m,n,m',n'})}} exp[-\frac{1}{2}(\eta_{t|t-1}^{m,n'})'(f_{t|t-1}^{m,n,m',n'})^{-1}(\eta_{t|t-1}^{m,n'})],\end{aligned} \tag{5.B.26}$$

and $\Pr(y_t|\Psi_{t-1})$ may be expressed as,

$$\begin{aligned}\Pr(y_t|\Psi_{t-1}) = \sum_0^l\sum_0^l\sum_0^l\sum_0^l \Pr(y_t, S_{1,t-1}=m,\\ S_{2,t-1}=n,S_{1,t}=m',S_{2,t}=m'|\Psi_{t-1}).\end{aligned} \tag{5.B.27}$$

The equation (5.B.27) is the basis for generating the log likelihood functions and shows that in this model there are sixteen combinations of state possible. To move the inference on probabilities forward we need to specify the last multiplier term in equation (5.B.25). Based on the assumption of independence of the two Markov processes, we have

$$\begin{aligned}\Pr(S_{1,t-1} &= m,S_{2,t-1}=n,S_{1,t}=m',S_{2,t}=n'|\Psi_{t-1})\\ &= \Pr(S_{1,t}=m'|S_{1,t-1}=m)\times\Pr(S_{2,t}=n'|S_{2,t-1}=n)\\ &\quad\times\Pr(S_{1,t-1}=m,S_{2,t-1}=n|\Psi_{t-1}),\end{aligned} \tag{5.B.28}$$

where the first two terms on the right hand side are recognizable as the corresponding transition probabilities. The third term on the right hand side of equation (5.B.28) can be separated in term of the realizations of the states in the previous time step. It, therefore, follows that,

$$\Pr(S_{1,t-1}=m\ ,S_{2,t-1}=n|\Psi_{t-1})$$

$$= \sum_{i=0}^{l}\sum_{i=0}^{l} \Pr(S_{1,t-2} = i, S_{1,t-1} = m, S_{2,t-2} = j, S_{2,t-1} = n|\Psi_{t-1}). \tag{5.B.29}$$

The log likelihood function constructed from equation (5.B.27) may be maximized with respect to the unknown parameters using suitable numerical optimization algorithm. We implemented this in GAUSS$^{\text{TM}}$and we started with the program code available as part of book, Kim and Nelson (1999). However, to make the filter operational we need starting values, . These are obtained following the suggestions in Kim and Nelson (1999, pp. 28-29, 70-71). The Markov switching heteroscedasticity model involving the unobserved component is suitable for constructing the conditional variance of the stock return. This is simply achieved by multiplying the filtered probability of the states with the conditional forecast error variance given by equation (5.B.18). The conditional variance is thus given by,

$$\sum_{0}^{l}\sum_{0}^{l}\sum_{0}^{l}\sum_{0}^{l} \Pr(S_{1,t-1} = m, S_{2,t-1} = n, S_{1,t} = m', S_{2,t} = n'|\Psi_{t-1}) \times f_{t|t-1}^{m,n,m',n'}. \tag{5.B.30}$$

In a similar manner the estimate of probability weighted forecast error could be generated using (5.B.17). This generated error series may then be analyzed for model diagnostics tests.

APPENDIX 5.C
Components of Variance of Inflation Forecast With Process Switching

In this appendix we derive the two components of the variance of inflation forecast for a particular forecast horizon. Evans and Wachtel (1993) describe these two as the certainty equivalent and the regime uncertainty components. They suggest that it is intuitive to analyze this way since there is uncertainty in the inflation process that is likely to be present at a future date. These two components help us analyze the interaction between these two parts of inflation uncertainty and the realized inflation for different horizon. Our basic inflation model may be restated as:

$$\pi_t = T_t + M_t, \tag{5.C.1}$$

$$T_t = T_{t-1} + (Q_0 + Q_1 S_{1,t})v_t, \tag{5.C.2}$$

$$M_t = \mu_2 S_{1,t} + \mu_3 S_{2,t} + \mu_4 S_{1,t} S_{2,t} + (h_0 + h_1 S_{2,t})e_t. \tag{5.C.3}$$

Denoting the model expectation by E^M and the information set at time t by Ω_t, we can write $k-$quarter ahead expectation as,

$$E^M(\pi_{t+k}|\Omega_t) = T_t + \mu_2 \Pr(S_{1,t+k} = 1||\Omega_t) + \mu_3 \Pr(S_{2,t+k} = 1|\Omega_t)$$

$$+ \mu_4 \Pr(S_{1,t+k} = 1, S_{2,t+k} = 1|\Omega_t), \tag{5.C.4}$$

where the probability terms may be obtained as,

$$\begin{bmatrix} \Pr(S_{1,t+k} = 0|\Omega_t) \\ \Pr(S_{1,t+k} = 1|\Omega_t) \end{bmatrix} = \begin{bmatrix} p_{00} & 1 - p_{11} \\ 1 - p_{00} & p_{11} \end{bmatrix}^k \begin{bmatrix} \Pr(S_{1,t} = 0||\Omega_t) \\ \Pr(S_{1,t} = 1|\Omega_t) \end{bmatrix}, \tag{5.C.5}$$

similarly,

$$\begin{bmatrix} \Pr(S_{2,t+k} = 0|\Omega_t) \\ \Pr(S_{2,t+k} = 1|\Omega_t) \end{bmatrix} = \begin{bmatrix} q_{00} & 1 - q_{11} \\ 1 - q_{00} & q_{11} \end{bmatrix}^k \begin{bmatrix} \Pr(S_{2,t} = 0||\Omega_t) \\ \Pr(S_{2,t} = 1|\Omega_t) \end{bmatrix}. \tag{5.C.6}$$

Following the definition in Evans and Wachtel (1993), the certainty equivalent of the variance of inflation at k quarters in future is given by,

$$Var(k)_{CE} = E[Var(\pi_{t+k})|\Omega_t, S_{1,t+k}|\Omega_t, S_{2,t+k}|\Omega_t], \tag{5.C.7}$$

and

$$\begin{aligned} Var(k)_{CE} = {} & k(Q_0 + Q_1) \Pr(S_{1,t+k} = 1|\Omega_t) \\ & +k \Pr(S_{1,t+k} = 0|\Omega_t) \\ & +(h_0 + h_1) \Pr(S_{2,t+k} = 1|\Omega_t) \\ & +h_0 \Pr(S_{2,t+k} = 0|\Omega_t). \end{aligned} \tag{5.C.8}$$

In a manner similar to that in Evans and Wachtel (1993), the regime uncertainty component of the forecast inflation variance is given by,

$$\begin{aligned} Var(k)_{RU} = {} & (T_t - \mu_2)^2 \Pr(S_{1,t+k} = 1|\Omega_t) \Pr(S_{2,t+k} = 0|\Omega_t) \\ & +(T_t - \mu_2 - \mu_3 - \mu_4)^2 \Pr(S_{1,t+k} = 1|\Omega_t) \Pr(S_{2,t+k} = 0|\Omega_t) \\ & +(T_t - 0)^2 \Pr(S_{1,t+k} = 0|\Omega_t) \Pr(S_{2,t+k} = 0|\Omega_t) \\ & +(T_t - \mu_3)^2 \Pr(S_{1,t+k} = 0|\Omega_t) \Pr(S_{2,t+k} = 1|\Omega_t). \end{aligned} \tag{5.C.9}$$

Chapter 6

EXPLORING PERMANENT AND TRANSITORY COMPONENTS OF STOCK RETURN

1. Introduction

An economic variable can naturally be considered in terms of a permanent component and a transitory component. Several approaches have been proposed in the related literature to decompose the univariate economic time series into permanent and transitory components. Nelson and Plosser (1982) use autocorrelation functions of a model comprising permanent and temporary components to ascertain the relative size of each component. Other researchers, for example, Watson (1986) and Clark (1987), have attempted this type of decomposition using unobserved component models of gross national product. Campbell and Mankiw (1987) use the parameters of a low order ARMA model to determine the impact of a shock on the long-run forecasts. Their technique enables them to decide the relative importance of the two components.

Another group of researchers focus on the mean reversion in stock prices. Based on an analysis of mean reversion using an autoregressive test, Fama and French (1988) report that stock prices in the US have a transitory mean-reverting component in addition to a permanent component. Porterba and Summers (1988), Lo and MacKinlay (1988) and Kim, Nelson and Startz (1991) employ the variance ratio methodology of Cochrane (1988) to detect mean reversion in stock prices [1]. These authors report mixed evidence on the presence of mean reversion in stock prices. More recently, Chou and Ng (1995) have applied the canonical decomposition method of Tsay and Chou (1994) to decompose equity indices into permanent and temporary components. They assert, based on their findings, that this multivariate model is relatively easy to estimate.

Beveridge and Nelson (1981) adopt a decomposition method that can conveniently estimate the permanent and transitory components of a time series using a forecasting model of the first difference of the series. As pointed out by Morley (2002), however, the need to evaluate infinite sums makes their method somewhat unwieldy. As a clever alternative, Morley (2002) suggests setting up the problem in a state space framework, a very convenient framework for dealing with unobserved component modeling.

In this chapter we adapt the empirical model of stock return proposed by Kim and Kim (1996) to decompose stock returns into permanent and transitory components, and then analyze the movements of stock returns for Germany, Japan, the UK and the USA. We also make use of state space structures with enhancements that allow different shocks to drive the two components. Two questions are specifically addressed. (1) How different are the speeds of the mean reversion of the temporary components? (2) How different are the patterns of correlation between permanent and temporal components across international markets?

Our analysis is carried out in two stages. First, we apply the Markov switching heteroskedasticity model. This approach has been adopted as an alternative method for dealing with ARCH effects in economic data.[2] The behavior of the unconditional variance constitutes the main difference between the ARCH type conditional heteroskedasticity and the Markov switching variance model. Specifically, the unconditional variance remains constant in the case of the former, whereas it changes with the state of the economy in the latter. Kim, Nelson and Startz (1998) demonstrate good fitting of the Markov switching variance model to monthly stock return data, particularly in terms of normality of the standardized return. In the framework of Kim and Kim (1996) the two components of the return are driven by different unobserved Markov processes, and this separation of stock return allows us to study how each of the two components impacts the variance of the return. Also, the estimated Markov switching parameters indicate the probability of each of these components to be in a high variance state and thus link them to different stages in the stock return process.

Second, we apply the model to stock return data for the four countries previously mentioned and examine how interactions between the stock return components and return variance differ from country to country. While many of the studies cited above use only US data for empirical analysis, our extended analysis attempts to detect the behavior of other influential economies of the world.

2. Markov Switching Heteroscedasticity Model of Stock Return

Under the empirical model of stock return proposed by Kim and Kim (1996) that we adapted for this study, the stock return series for a particular country is assumed to consist of a permanent component (pm_t) and a stationary component. We also assume that the return r_t consists of a constant plus noise and an autoregressive stationary component, x_t. Although the most commonly used method in modeling the well-known variance clustering in stock return is by a GARCH process, we adopt an alternative specification more suitable at the monthly frequency. An important feature of this model is the incorporation of the shocks to the transitory component. This allows us to examine an impact on the overall return variance.

In a GARCH framework, the unconditional variance does not change, whereas in a Markov switching specification it changes depending on the state of volatility. This is the main difference between these two approaches to capture the empirical characteristics of the stock return volatility. Hamilton and Susmel (1994) propose a switching ARCH model in which they allow the parameters of the ARCH process to come from one of several different regimes. Although the ARCH process controls the short-run dynamics, the long-run dynamics are governed by regime shifts in unconditional variance, while an unobserved Markov switching process drives the regime changes. These authors apply the model to weekly return data and show that the ARCH effects almost completely diminish after a month. This tends to indicate that in modeling monthly return an ARCH term may not be necessary.

The regime switching approach has also been popular to examine the mean reversion in stock return, as in, for example, Kim and Nelson (1998) and Graflund (2000). These authors argue that the variance ratio test that is often used for analyzing mean reversion may need to be modified to take into account the changes in variance due to changes in regimes. In this present study we assume that the return series is drawn from a mixture of normal distributions, as in Kim, Nelson and Startz (1998). These authors have shown that the Markov switching heteroscedasticity model of stock return is a good approximation of the underlying data generating process. This leads us to formulate the return series as

$$r_t = pm_t + x_t, \tag{6.1}$$

$$pm_t = \mu + (Q_0 + Q_1 S_{1,t})\, v_t, \tag{6.2}$$

$$x_t = \phi x_{t-1} + (h_0 + h_1 S_{2,t})\, e_t, \tag{6.3}$$

where $e_t \sim N(0,1)$, and $v_t \sim N(0,1)$. In this model we use x_t to represent the temporary part of the return and not the prices directly. We include ϕ simply reflecting the fact that the temporary component of the return could be auto-correlated.[3] $S_{1,t}$ and $S_{2,t}$ are unobserved state variables that evolve independently as two state Markov processes. These state variables determine the underlying regime at any given time. Their associated transition probability matrices govern the evolution of these state variables. The parameters h_1 and Q_1 help us identify any shift in variance during periods of high uncertainty. The estimation of this model allows us to comment on the time series behavior of the return volatility for a particular country and how this is influenced by the switching probability of the transitory component.

The two Markov switching variables are independent of each other and the respective transition probabilities are defined as:

$$\Pr(S_{1,t} = 0 | S_{1,t-1} = 0) = p_{00}, \;\; \Pr(S_{1,t} = 1 | S_{1,t-1} = 1) = p_{11}, \tag{6.4}$$

$$\Pr(S_{2,t} = 0 | S_{2,t-1} = 0) = q_{00}, \;\; \Pr(S_{2,t} = 1 | S_{2,t-1} = 1) = q_{11}. \tag{6.5}$$

In order to estimate such a model that involves unobserved components and is subject to Markov switching shocks, we utilize the procedure discussed by Kim and Nelson (1999). This involves generating a probability weighted likelihood function and a recursive algorithm to update the probabilities as new observations becomes available. This has been written with computer programming in mind. The parameters to be estimated are, therefore, $[p_{11}, p_{00}, Q_0, Q_1, \mu, q_{11}, q_{00}, h_0, h_1, \phi]$.

3. Data

This analysis uses the data from the monthly stock price indexes of Germany, Japan, the UK and the USA. The sample period spans the 32 years from December 1969 to March 2001. Each stock price index is obtained from the Morgan Stanley Capital International Index and represents the end-of-month index. The rate of return on stocks for each country is calculated as $y_t = (P_t - P_{t-1}) \times 100/P_{t-1}$, where P_t is the stock price index at time t. Thus, the rates of return on stocks are obtained for the period from January 1970 to March 2001.

Table 6.1 summarizes statistics on the rate of return in each country, including descriptive statistics on the mean, standard deviation (Std. Dev.), skewness, kurtosis and the $P-$value of the Jarque-Bera test statistic (JB test) for testing the normality of the series. Under the null hypothesis, the Jarque-Bera statistic has a chi-square distribution with two

Table 6.1. Summary statistics

	Germany	Japan	UK	USA
Mean (%)	0.9101	0.8351	1.2973	1.0416
Std. Dev.	5.3472	5.3987	6.1799	4.4699
Skewness	-0.3007	-0.0290	1.4162	-0.3095
Kurtosis	4.6979	4.2131	18.3381	4.9873
JB test	0.0000	0.0000	0.0000	0.0000

Note: The hypothesis of normal distribution is rejected at the 5% (1%) level of significance if the $P-$value for the JB test is less than 0.05 (0.01).

degrees of freedom. When the reported probability for the Jarque-Bera statistic is small, the null hypothesis of a normal distribution is rejected. As clearly demonstrated in Table 1, the mean and standard deviation of stock return are relatively high for the UK, but relatively low for Germany and Japan. The hypothesis of normal distribution is rejected at the 1% significance level for all of the countries.

4. Empirical Results

Table 6.2 shows the parameter estimates of the Markov switching heteroskedasticity model for the sample period from March 1971 to March 2001. The results are computed using the algorithm discussed in the previous chapter. The initial values for the filter are obtained from the observations recorded during the first year, i.e. between January 1970 and February 1971. The estimates of transition probability p_{11} (high-variance state of the permanent component) and the probability p_{00} (low-variance state of the permanent component) are both highly significant for all of the countries.

The variance estimates of the permanent component of the return also significantly differ between the countries. The low-variance state estimate (Q_0) is highly significant for both the UK and USA, but not for Germany or Japan. In contrast, the additional variance (Q_1) of the permanent component due to the high volatility regime is significant for all four countries. This means that the variance of the permanent component of the return increases for all of these markets as they enter the high volatility state. We were also interested to note that the magnitude of the overall variance of the permanent component during the

Table 6.2. Permanent and transitory components of equity return: Markov switching heteroscedasticity framework

	Germany	Japan	UK	USA
p_{11}	0.9922*	0.9963*	0.9945*	0.9948*
	(0.0077)	(0.0051)	(0.0055)	(0.0055)
p_{00}	0.9746*	0.9785*	0.9448*	0.9695*
	(0.0256)	(0.0141)	(0.0499)	(0.0266)
Q_0	0.1192	0.0045	0.8048*	1.5918*
	(1.4704)	(0.3146)	(0.1082)	(0.1870)
Q_1	2.4398**	3.3458*	2.3961*	1.3405*
	(1.4703)	(0.5074)	(0.2237)	(0.2163)
μ	0.8443*	1.0070*	1.5258*	1.2070*
	(0.2330)	(0.2289)	(0.2239)	(0.1932)
q_{11}	0.9515*	0.8070*	0.8683*	0.7078*
	(0.0394)	(0.1708)	(0.0780)	(0.1702)
q_{00}	0.9792*	0.9325*	0.9896*	0.9855*
	(0.0172)	(0.0511)	(0.0078)	(0.0113)
h_0	2.7861*	2.3276*	1.6005*	0.0270
	(0.3919)	(0.4436)	(0.3311)	(0.3744)
h_1	3.8856*	3.6222*	13.6552*	9.1604*
	(0.7958)	(1.2882)	(2.5826)	(2.8335)
ϕ	0.0902	0.2887	0.0047	-0.1434
	(0.2031)	(0.2052)	(0.1266)	(0.1286)

Note: The parameters are described in the text. Standard errors are given in parentheses below the parameter estimates. Significance at the 10% level is indicated by ** and at the 5% level is indicated by *.

high volatility state, i.e. $Q_0 + Q_1$, diverges very little among the four markets.

The parameters of particular interest in this study are those relating to the transitory component of the return. The transition probabilities q_{11} (high-variance state of the transitory component) and q_{00} (low-variance state of the transitory component) are both is highly significant for all of the countries. This is an indication that the low volatility state dominates in each country. The expected duration of the low volatility state ranges from a low of 14.81 months for Japan to a high of 96.15 months for the UK, while that for the USA falls closer to the latter, at 68.97 months. The expected duration of the high volatility state also ranges quite widely, from a low of 3.42 months for the USA to a high of 20.62 months for Germany. The average duration of the low volatility state

Table 6.3. Residual diagnostics and model adequacy tests

	Portmanteau	ARCH	KS Test	MNR	Recursive T
Germany	0.362	0.276	0.060	0.632	0.506
Japan	0.179	0.679	0.050	0.258	0.669
UK	0.064	0.067	0.083	0.957	0.572
USA	0.769	0.811	0.043	0.747	0.633

Note: Entries are $P-$values for the respective statistics except for the KS statistic. These diagnostics are computed from the recursive residual of the measurement equation. The null hypothesis in portmanteau test is that the residuals are serially uncorrelated. The ARCH test checks for no serial correlations in the squared residual up to lag 26. Both these tests are applicable to recursive residuals as explained in Wells (1996, p. 27). MNR is the modified Von Neumann ratio test using recursive residual for model adequacy (see Harvey, 1990, chapter 5). Similarly, if the model is correctly specified then Recursive T has a Student's $t-$distribution (see Harvey, 1990, p. 157). KS statistic represents the Kolmogorov-Smirnov test statistic for normality. 95% significance level in this test is 0.072. When KS statistic is less than 0.072 the null hypothesis of normality cannot be rejected at the indicated level of significance.

Table 6.4. Correlations: return, permanent and temporary components

Return				
	Germany	Japan	UK	USA
Germany	1.0000			
Japan	0.3166	1.0000		
UK	0.3930	0.3013	1.0000	
U.S.A.	0.4212	0.3356	0.5676	1.0000
Permanent component				
	Germany	Japan	UK	USA
Germany	1.0000			
Japan	0.3154	1.0000		
UK	0.3925	0.2987	1.0000	
USA	0.4274	0.3434	0.5680	1.0000
Temporary component				
	Germany	Japan	UK	USA
Germany	1.0000			
Japan	0.2545	1.0000		
UK	0.3651	0.2783	1.0000	
USA	-0.3085	-0.2900	-0.5214	1.0000

among the four countries is 57.00 months, while the average duration of the high volatility state is 9.20 months. The latter value means that the high volatility transitory state fades in about 9 months on average for this group of countries.

The variance measure for the low volatility state of the transitory component (h_0) is significant for Germany, Japan and the UK, but not the USA. On the other hand, h_1, the additional variance measure for the high volatility state of the transitory component, is highly significant for all of the countries. Moreover, h_1 is relatively high for all the countries. It is interesting to observe that the overall variance in the high volatility regime ($h_0 + h_1$) is quite comparable among all of the countries, and the highest for the UK. This, coupled with the fairly long duration of the high volatility state (average of 9 months), indicates that the high variance of the transitory component tends to be relatively short lived.

In order to check the performance of the model, we analyze the residual from the model using a variety of diagnostics tests. Our test results are presented in Table 6.3. The portmanteau tests support the whiteness of the residuals, the ARCH tests indicate no remaining heteroskedasticity in the residuals, and the Kolmogorov-Smirnov tests support the normality of the residuals. The test results taken collectively overwhelmingly support the modeling approach adopted, and on this basis they validate the significance of the statistical conclusions drawn from this approach.

In addition to the three tests outlined above, we also apply a pair of tests specifically designed for the recursive residuals produced by the state space system adopted in this study. The former, the modified von-Neumann ratio, tests against serial correlations in the residuals, while the latter, the recursive t-test, checks for correct model specifications. As the entries in Table 6.3 indicate, the model applied to different country datasets performs extremely well in both these tests. Again, there is overwhelming support for the adequacy of the model in describing the price process involving the permanent and transitory components.

Table 6.4 shows the correlation patterns for the return itself, for the permanent component, and for the temporal component across international markets.

The returns of all four markets are positively correlated with each other. The highest correlation is 0.5676, between the USA and UK, and the lowest correlation is 0.3013, between the UK and Japan. The USA, UK and German markets correlate strongly with each other, but relatively weakly with the market in Japan. The inter-country correlations for the temporary component tend to be slightly smaller than those for the permanent component. The correlations between Germany and the

UK, for example, are 0.3925 for the permanent component versus 0.3651 for the temporary component. It is of interest to note that all of the correlation coefficients between the USA and other three countries are negative for temporary components and positive for the permanent ones. This implies that the USA market moves in the same direction as the other three markets in the long run, but in an opposite direction in the short run.

5. Conclusion

In this chapter we apply the Markov switching heteroskedasticity model to stock returns for Germany, Japan, the UK, and USA, and decompose these stock returns into permanent and transitory components. This modeling approach is superior to the GARCH model, as emphasized by Lamoureux and Lastrapes (1990). In particular, the Markov-switching model explicitly considers the possibility of regime shift, whereas the GARCH model does not.

According to our evaluation of the differences between the speeds of mean reversion of the temporary components, the high-variance state of the transitory component lasts for an average of only 3.42 months for the USA, versus 20.62 months for Germany. The time periods for the other two countries fall between these values.

Regarding the second objective of this study, to determine differences in the patterns of correlation between permanent and temporal components across international markets, we find that the US market correlates positively with the other markets for the permanent component, but negatively for the temporal component. This implies that the USA moves in the same direction as the other three countries in the long run, and in an opposite direction in the short run.[4]

Notes

1 Porterba and Summers (1988) test for the presence of a transitory component, though they do not formally decompose stock prices into permanent and transitory components.

2 See for example Kim, Nelson and Startz (1998).

3 In Kim and Kim (1996) the state equation consists of the model for the temporary component of the prices and the authors obtain relatively high values of the AR parameters. Moreover, the permanent component contains only constant and regime dependent volatility terms. In that context, we believe our model is quite parsimonious.

4 This chapter is an edited version of Bhar and Hamori (2004) with permission from Elsevier.

Chapter 7

EXPLORING THE RELATIONSHIP BETWEEN COINCIDENT FINANCIAL MARKET INDICATORS

1. Introduction

The time series properties of expectations and volatility of stock returns have attracted attention in the financial literature. Driven by the strong intuition that risk and return should be positively correlated, researchers have empirically examined the covariance between the mean and volatility of returns in search of a positive relationship between expected returns and conditional volatility. However, these efforts have yet to yield conclusive results. In fact, prior empirical investigations into the contemporaneous correlation between the first two moments of stock market returns have yielded decidedly mixed results.[1] Further, Backus and Gregory (1993) point out that a negative and even non-monotonic relationship is consistent with equilibrium. Researchers nonetheless continue to investigate stock return data within the framework of ARCH-M and its generalizations, descriptive empirical models that impose a linear relationship between the conditional expectation and conditional volatility of returns.

However, the mixed empirical results in the literature indicate that inferences are sensitive to the way in which the moments and the relationships between them are modeled. Thus, some empirical studies have focused on the non-linearity between stock returns and proxies for market risk. Whitelaw (1994), for example, empirically investigates the comovements of the conditional mean and volatility of stock returns using monthly US data from April 1953 to March 1989. His results show that the conditional mean and volatility exhibit an asymmetric relationship that contrasts with the contemporaneous relationship found earlier. As a consequence, Whitelaw (1994) questions the value of modeling expected returns as a constant function of conditional volatility and concludes that imposing a constant linear relationship between the mean and volatility may lead to erroneous inferences.

Many studies also show that stock returns can be predicted by means of publicly available information, such as time series data on financial and macroeconomic variables. Pesaran and Timmermann (1995), for example, assess the economic significance of the predictability of US stock returns at a monthly frequency over the period between January 1954 and December 1992. According to their results, the power of economic factors to predict stock returns changes through time and tends to vary with the volatility of the returns. Noting the close links between important episodes of stock return predictability and the incidence of sudden shocks to the economy, they recommend the use of forecasting procedures that allow for possible regime changes when analyzing stock return predictability.

These two lines of research suggest the importance of considering non-linearity when grouping the behavior of excess returns according to the state of the business cycle.

Following the example of Chauvet and Potter (2000), we attempt to represent stock market fluctuations in this chapter by constructing a non-linear coincident financial indicator. The indicator we use is based on a broad information set of market conditions generated from expectations about changes in future economic activity, including the overall state of financial markets.

Chauvet and Potter (2000) formulate swings in the stock market as a function of investor reactions towards changes in unobserved market risk factors. They construct a non-linear proxy for market risk premia as a latent factor, whose first moment and conditional second moment are driven by a two-state Markov variable. Excess returns and their volatility may reflect changes in average opinion about financial information as investors learn more about the data. Chauvet and Potter (2000) stress the fact that when econometricians try to evaluate how stock returns affect changes in systematic risk, they cannot directly see the information that market participants use. Moreover, investors may respond asymmetrically when a broad information set is adopted, depending on how the investors perceive business conditions.[2] In addition to capturing the asymmetric behavior of investors toward risk and the inclinations of investors to hedge against noise, the Markov-switching introduced in the unobservable factor can generate higher levels of volatility when an economy moves into certain states.

2. Markov Switching Coincidence Index Model

Many economic and/or financial time series are found to be highly correlated, but nor necessarily cointegrated. Economists have attempted to make use such observation to develop models that might be able to predict the likely direction of the particular market being investigated. It has also been found useful to be able to summarize a number of series into a smaller number and use that for prediction purposes. In this context the factor structure has been employed. Stock and Watson (1991) model of coincident economic indicator in this context is a classic example. Since such a factor is essentially unobservable, it is well suited for modeling by an appropriate state space system.

More recently, Chauvet (1998) and Chauvet and Potter (2000) have enhanced such a modeling strategy by incorporating heteroscedastic innovations. This not only allows them to incorporate different regimes the economy might pass through over a long period of time, say twenty years or more. In the following paragraphs we outline this modeling approach and apply that to infer the indicator variable for the financial markets for three countries, e.g. Japan, the UK and the USA. We then study the dependence pattern of these indicator variables, which, in turn, shed light on comovement between these markets. Chauvet and Potter (2000) exploit this inferred variable to alter portfolio composition in order to earn additional return. In our case, all three markets could be involved in such a portfolio strategy, although we do not pursue this line of investigation in this chapter.

We focus on four financial market variables that have been reported in the literature to be commoving and these are, equity market excess return, proxy for the market volatility, short-term interest rate and the price-earning ratio. As we find the short-term interest rate and the price-earning ratios are non-stationary, we use the first difference of the log of these variables in the model. The preliminary analysis of the principal component of these four variables indicate one dominant factor, hence we set up the model to account for this unobserved factor with a given dynamic characteristic. Additional details of the data and their sources are given the data section.

The model is, thus, based around the unobserved dynamic factor wit the following structure:

$$\lambda_t = \alpha_0 + \alpha_1 S_t + \phi\lambda_{t-1} + \eta_{S_t}, \quad \eta_{S_t} \sim N(0, \sigma^2_{S_t}), \tag{7.1}$$

where, λ_t is the coincident financial indicator, and $S_t = \{0, 1\}$ is the Markov switching variable indicating the state of the financial market at any given month. This evolves with its own transition probability property. This is discussed further in the appendix. The innovation variance also depends on the state of the financial market as does the mean of the unobserved factor.

Our model assumes that the observations of the four financial market variables are related to this dynamic factor as well as its own idiosyncratic innovations. Under this proposal the measurement process is given by,

$$\begin{bmatrix} y_{ret} \\ y_{vol} \\ y_{ir} \\ y_{pe} \end{bmatrix} = \begin{bmatrix} \beta_{ret} \\ \beta_{vol} \\ \beta_{ir} \\ \beta_{pe} \end{bmatrix} \lambda_t + \begin{bmatrix} v_{ret} \\ v_{vol} \\ v_{ir} \\ v_{pe} \end{bmatrix}. \tag{7.2}$$

The idiosyncratic innovations are assumed to have an AR(1) structure of their own as described by,

$$v_{i,t} = \theta_i v_{i,t-1} + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_i^2). \tag{7.3}$$

Here i represents either return, volatility, interest rate or the price-earning variables. We also assume that these innovations are uncorrelated between themselves as well as with the innovation of the factor. With a little thought we could put these set of equations in the state space form which then directly comparable to the model estimation procedure discussed in the appendix. The measurement equation of the state space form is, therefore,

$$\begin{bmatrix} y_{ret} \\ y_{vol} \\ y_{ir} \\ y_{pe} \end{bmatrix} = \begin{bmatrix} \beta_{ret} & 1 & 0 & 0 & 0 \\ \beta_{vol} & 0 & 1 & 0 & 0 \\ \beta_{ir} & 0 & 0 & 1 & 0 \\ \beta_{pe} & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_t \\ v_{ret} \\ v_{vol} \\ v_{ir} \\ v_{pe} \end{bmatrix}, \tag{7.4}$$

and the state transition equation becomes,

$$\begin{aligned} \begin{bmatrix} \lambda_t \\ v_{ret,t} \\ v_{vol,t} \\ v_{ir,t} \\ v_{pe,t} \end{bmatrix} &= \begin{bmatrix} \alpha_0 + \alpha_1 S_t \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ &+ \begin{bmatrix} \phi & 0 & 0 & 0 & 0 \\ 0 & \theta_{ret} & 0 & 0 & 0 \\ 0 & 0 & \theta_{vol} & 0 & 0 \\ 0 & 0 & 0 & \theta_{ir} & 0 \\ 0 & 0 & 0 & 0 & \theta_{pe} \end{bmatrix} \begin{bmatrix} \lambda_{t-1} \\ v_{ret,t-1} \\ v_{vol,t-1} \\ v_{ir,t-1} \\ v_{pe,t-1} \end{bmatrix} + \begin{bmatrix} \eta_{S_t} \\ \epsilon_{ret} \\ \epsilon_{vol} \\ \epsilon_{ir} \\ \epsilon_{pe} \end{bmatrix}. \end{aligned} \tag{7.5}$$

The estimation of the unknown parameters of this model is achieved via the numerical maximization of the prediction error form of the likelihood function as described in the appendix.

3. Data

To carry out empirical analysis, we need a set of financial variables that vary contemporaneously with stock market cycles. This includes equity market excess returns, proxies for market volatility, short-term interest rates, and price-earnings ratios. These variables contain information on the underlying systematic risks within the economy, and they have been successfully used in previous empirical studies. We use these data for three countries: Japan, the UK, and the USA. The data were taken from Data Stream and International Financial Statistics (International Monetary Fund). The sample period spans the approximately 30 year period from January 1973 through June 2001. As the short-term interest rates and price-earnings ratios turn out to be non-stationary, we use the first difference of the log of these variables in the model. Thus, the data are obtained for the period between February 1973 and June 2001.

4. Empirical Results

We estimate the monthly coincident indicators for Japan, the UK, and the USA. Table 7.1 reports the results for the dynamic factor model when both its mean and variances are assumed to switch regimes.

State 0 is characterized by a positive mean rate ($\alpha_0 = 0.07770$ for Japan, 0.05629 for the UK, and 0.05844 for the USA) and low variance ($\sigma^2_{\eta 0} = 0.09764$ for Japan, 0.21467 for the UK, and 0.15178 for the USA), conditions that describe bull markets. State 1 is characterized by a negative mean rate ($\alpha_1 = -0.10613$ for Japan, -0.25420 for the UK, and -0.41292 for the USA) and high variance ($\sigma^2_{\eta 1} = 0.57381$ for Japan, 1.97569 for the UK, and 1.10593 for the USA), conditions that prevail during bear markets. Notice that the volatility in bear markets is 5.88 times (Japan), 9.20 times (UK), and 7.29 times (USA) higher than the volatility during bull markets, respectively. In other words, the model captures empirical observation of asymmetry during the stages of stock market cycles: bear markets are shown to be volatile and associated with steep and short contractions, while bull markets are shown to change more gradually.

The transition probabilities $\Pr(S_t = i | S_{t-1} = i) = p_{ii},\ i = 0, 1$ are that a financial market will remain in state i, given that it is in state i. The probability of staying in a bull market (p_{00}) is estimated at 0.93744

Table 7.1. Parameter estimates for the Markov switching coincidence index model

	Japan	UK	USA
α_0	0.07770 (0.0315)	0.05629 (0.0295)	0.05844 (0.0273)
α_1	-0.10613 (0.0737)	-0.25420 (0.2540)	-0.41292 (0.3512)
ϕ	0.02937 (0.0515)	0.05190 (0.0594)	0.10522 (0.0596)
$\sigma^2_{\eta 0}$	0.09764 (0.0254)	0.21467 (0.0230)	0.15178 (0.0229)
$\sigma^2_{\eta 1}$	0.57381 (0.0739)	1.97569 (0.5407)	1.10593 (0.5176)
p_{00}	0.93744 (0.0461)	0.98392 (0.0108)	0.97618 (0.0194)
p_{11}	0.96187 (0.0302)	0.84466 (0.0785)	0.60952 (0.1851)
β_{vol}	0.02799 (0.0374)	-0.05287 (0.0399)	-0.08797 (0.0550)
β_{ir}	-0.15221 (0.1590)	-0.09190 (0.0493)	0.01678 (0.0746)
β_{pe}	0.84269 (0.0323)	1.03485 (0.0283)	0.97917 (0.0891)
θ_{ret}	-0.25824 (0.4442)	-0.06042 (0.0732)	-0.29387 (0.1060)
θ_{vol}	0.61852 (0.0435)	0.58351 (0.0414)	0.60788 (0.0420)
θ_{ir}	0.23427 (0.0535)	0.36696 (0.0436)	0.23691 (0.0498)
θ_{pe}	0.19116 (0.0694)	0.84117 (0.0751)	-0.11752 (0.0866)
σ^2_{ret}	0.00806 (0.0100)	0.07525 (0.0076)	0.05791 (0.0203)
σ^2_{vol}	0.24708 (0.0194)	0.32301 (0.0231)	0.16828 (0.0128)
σ^2_{ir}	3.62191 (0.2583)	0.31148 (0.0210)	0.42581 (0.0290)
σ^2_{pe}	0.07112 (0.0092)	0.00485 (0.0028)	0.07204 (0.0192)

Note: Maximum likelihood estimates of the parameters are reported here. The numbers in parentheses are the standard errors computed from the diagonal elements of the final covariance matrix.

Table 7.2. Diagnostic statistics for the residuals of measurement equations

	Portmanteau	ARCH	KS Test	Recursive T
Japan				
Return	0.137	0.531	0.054	0.972
Volatility	0.188	0.486	0.053	0.988
T Bill	0.186	0.434	0.052	0.995
P/E	0.167	0.443	0.051	0.914
UK				
Return	0.431	0.366	0.097	0.479
Volatility	0.424	0.329	0.097	0.471
T Bill	0.41	0.336	0.097	0.473
P/E	0.349	0.279	0.097	0.477
USA				
Return	0.451	0.572	0.047	0.501
Volatility	0.699	0.527	0.047	0.475
T Bill	0.681	0.537	0.046	0.468
P/E	0.784	0.440	0.046	0.447

Note: Entries are $P-$values for the respective statistics except for the KS statistic. These diagnostics are computed from the recursive residual of the corresponding measurement equation. The null hypothesis in portmanteau test is that the residuals are serially uncorrelated. The ARCH test checks for no serial correlations in the squared residual up to lag 26. Both these test are applicable to recursive residuals as explained in Wells (1996, p. 27). If the model is correctly specified then Recursive T has a Student's t-distribution (Harvey, 1990, p. 157). KS statistic represents the Kolmogorov-Smirnov test statistic for normality. 95% significance level in this test is 0.074. When KS statistic is less than 0.074 the null hypothesis of normality cannot be rejected at the indicated level of significance.

Table 7.3. Correlation between the coincident financial indicator and its components

	Excess Return	Volatility	T-bill	P/E
Japan	0.9989	-0.1339	-0.0747	0.8988
UK	0.9295	0.0606	-0.1518	0.9945
USA	0.9447	-0.3111	0.0020	0.9258

Table 7.4. Granger causality tests

Country	Null Hypothesis	$P-$value
Japan	Excess return does not cause coincident indicator.	0.4597
	Coincident indicator does not cause excess return.	0.5021
	Volatility does not cause coincident indicator.	0.1314
	Coincident indicator does not cause volatility.	0.0000
	T-bill does not cause coincident indicator.	0.0265
	Coincident indicator does not cause T-bill.	0.3841
	P/E does not cause coincident indicator.	0.5063
	Coincident indicator does not cause P/E.	0.0062
UK	Excess return does not cause coincident indicator.	0.0663
	Coincident indicator does not cause excess return.	0.0814
	Volatility does not cause coincident indicator.	0.0533
	Coincident indicator does not cause volatility.	0.0007
	T-bill does not cause coincident indicator.	0.0193
	Coincident indicator does not cause T-bill.	0.4096
	P/E does not cause coincident indicator.	0.7083
	Coincident indicator does not cause P/E.	0.3506
USA	Excess return does not cause coincident indicator.	0.0599
	Coincident indicator does not cause excess return.	0.0224
	Volatility does not cause coincident indicator.	0.0534
	Coincident indicator does not cause volatility.	0.1075
	T-bill does not cause coincident indicator.	0.0000
	Coincident indicator does not cause T-bill.	0.0048
	P/E does not cause coincident indicator.	0.0836
	Coincident indicator does not cause P/E.	0.0012

for Japan, 0.98392 for the UK, and 0.97618 for the USA, while the probability of staying in a bear market (p_{11}) is estimated at 0.96187 for Japan, 0.84466 for the UK, and 0.60952 for the USA. These estimates are highly significant for all countries. Chauvet and Potter (2000) showed that p_{00} is higher than p_{11} in the USA, which means that the average duration of bull markets is longer than that of bear markets. According to our own results, the findings of Chauvet and Potter hold true for the US data, but not for data from Japan or the UK.

The factor loadings characterize the direct structural relations between the unobservable variable λ_t and the observable variables Y_{it}. The β_i coefficients measure the sensitivity of Y_{it} to a one-unit change in λ_t.

The financial market indicator has the same scale as the excess return, since we set its factor loading to one. Thus, the sign of the other factor loadings indicates the direction of correlation of the financial variables with the financial market indicator. Using the US data, Chauvet and Potter (2000) find that the financial market indicator is positively related to growth in the price-earning ratio but negatively related to both market volatility and changes in interest rates. However, our own empirical results are only consistent with their findings in the case of the UK. According to the data from our study, the financial market indicator in Japan is negatively related to changes in interest rates and positively related to market volatility and growth in the price-earnings ratio, while that in the USA is negatively related to the financial market indicator and positively related to changes in interest rates and growth in the price-earnings ratio.

The adequacy of the model specification is verified through the diagnostic analysis outlined in Table 7.2. The analysis consists of the Portmanteau test, ARCH test, KS test, and Recursive test. These diagnostics are computed from the recursive residual of the corresponding measurement equation. The null hypothesis in the Portmanteau test is that the residuals are serially uncorrelated. The ARCH test checks for an absence of serial correlations in the squared residual up to lag 26.[3] If the model is correctly specified, then Recursive T has a Student's t-distribution.[4] The KS statistic represents the Kolmogorov-Smirnov test statistic for normality. The 95% significance level in this test is 0.074. When the KS statistic is less than 0.074, the null hypothesis of normality cannot be rejected at the indicated level of significance. This table basically shows that our specification improvement in Table 7.1 is empirically supported.

Figures 7.1, 7.2, and 7.3 show the estimated coincident indicator and the excess return for the three countries. The time series path of the financial indicator is remarkably similar to the excess return series for all countries. Figures 7.4, 7.5, and 7.6 show the inferred probabilities of the low variance state in Japan, the UK, and the USA, respectively.

Table 7.3 shows the correlation coefficient between the coincident financial indicator and its components. The financial market indicators have particularly high correlations with excess returns, i.e. 0.9989 for Japan, 0.9295 for the UK, and 0.9447 for the USA. The growth rates of the price-earnings ratios are also highly correlated with the financial indicators, i.e. 0.8988 for Japan, 0.9945 for the UK, and 0.9258 for the USA. These results are consistent with those of Chauvet and Potter (2000), indicating that the structure of the financial market indicator

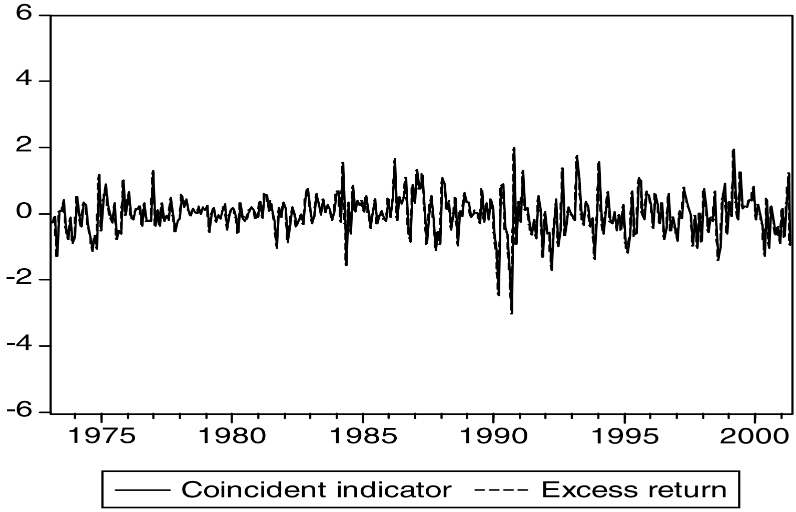

Figure 7.1. Estimated coincident indicator and excess return: Japan

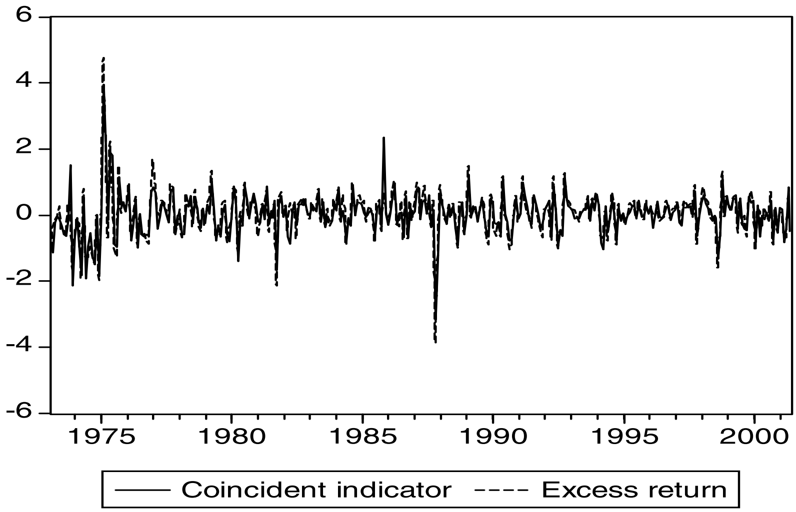

Figure 7.2. Estimated coincident indicator and excess return: UK

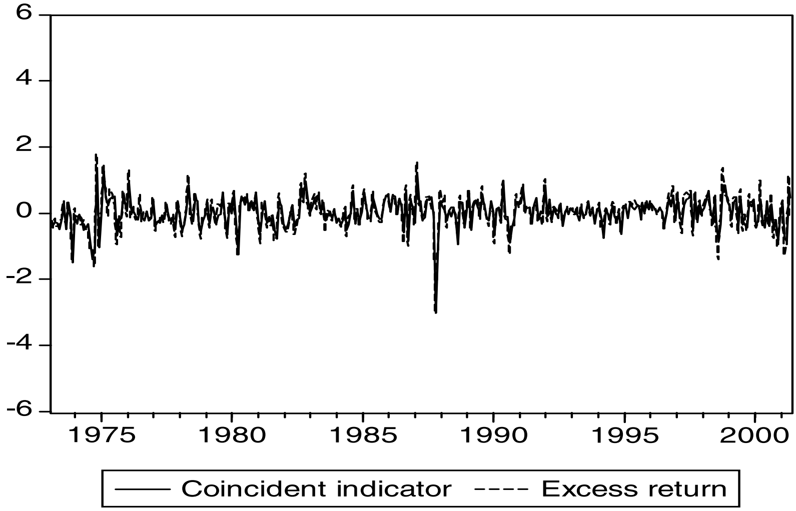

Figure 7.3. Estimated coincident indicator and excess return: USA

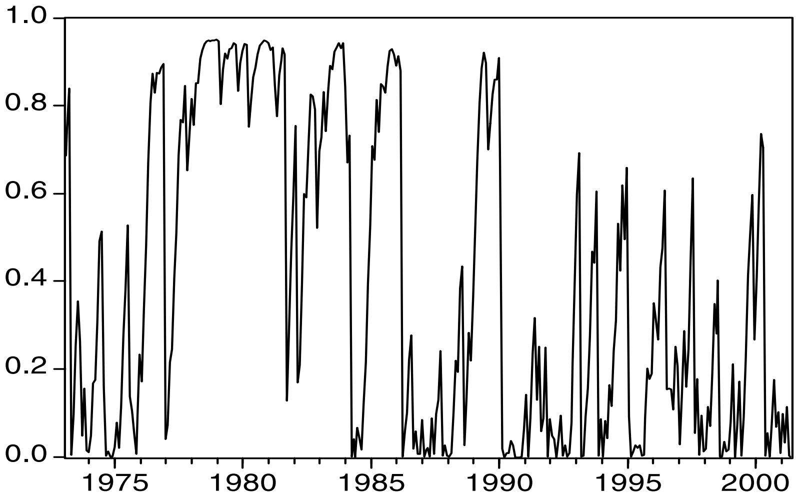

Figure 7.4. Inferred probability of low variance state: Japan

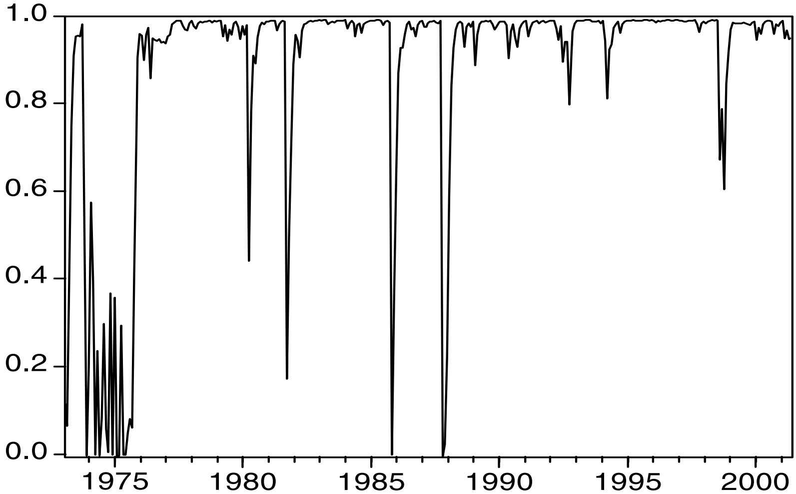

Figure 7.5. Inferred probability of low variance state: UK

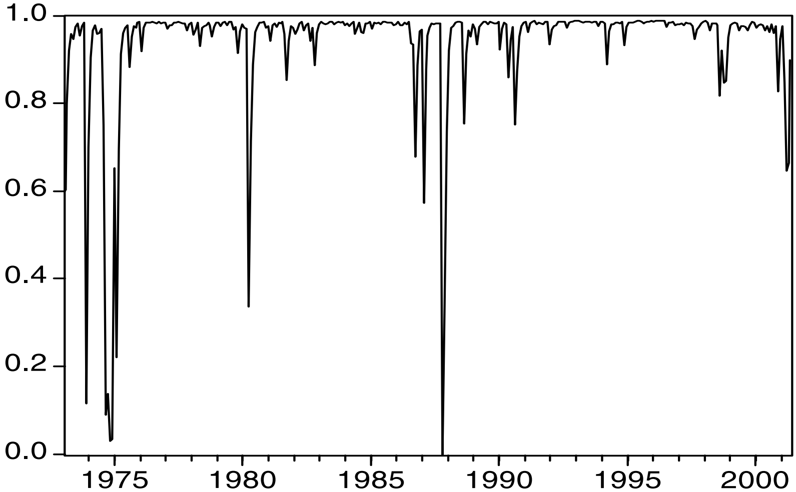

Figure 7.6. Inferred probability of low variance state: USA

cannot be simply imposed on the financial variables by assuming large idiosyncratic noise terms.

Table 7.4 shows the results of the Granger causality test of the bivariate system between the coincident financial indicator and other components. Here, we use the AIC to choose the optimal lag length of the system. In Japan, the coincident index causes the volatility and price-earning ratio, while the interest rate causes the financial indicator. In the UK, the coincident index causes the volatility and the interest rate causes the financial indicator. In the USA, the financial indicator causes the excess return, price-earning ratio, and interest rate, while the interest rate causes the coincident indicator. Thus, the interest rate is found to lead the coincident indicator in each of the three countries.

5. Conclusion

Following Chauvet and Potter (2000), we represented the monthly stock market fluctuations by constructing a non-linear coincident financial indicator in this chapter. The indicator is constructed as an unobservable factor whose first moment and conditional volatility are driven by a two-state Markov variable. The model is applied to economic data from Japan, the UK, and the USA, and successfully estimated for each of these industrialized countries. The coincident indicator has a particularly high correlation with excess returns, and its time series path is remarkably similar to the excess return series in Japan, the UK, and the USA. The estimated indicator can be interpreted as the investor's real-time belief about the state of financial conditions. In other words, it can be viewed as a coincident indicator of movements in the stock market reflecting common assessments of the implications of given sets of financial information.[5]

Notes

1. French, Schwert, and Stambaugh (1993) find a statistically significant positive relationship between expected returns and anticipated volatility. Glosten, Jagannathan, and Runkle (1987) conclude that there is a negative relationship, or no relationship, between expected returns and anticipated volatility. Chan, Karolyi, and Stulz (1992) conclude that the expected return on the US market is not related to its own conditional variance, but positively related to the conditional covariance with a foreign index.
2. For example, changes in interest rates may be interpreted as bad or good news, depending on whether the economy is in a recession or a boom.
3. Both these tests are applicable to recursive residuals as explained in Wells (1996, p. 27).
4. See Harvey (1990, p. 157).
5. See Chauvet and Potter (2000, p.89).

APPENDIX 7.A
Data

The main sources of the data are tabulated below.

Data Set Codes from DataStream and IFS

	Price Index (PI)	Dividend Yield (DY)	P/E Ratio (PE)	Interest Rate (IR)
Japan	TOTMKJP(PI)	TOTMKJP(DY)	TOTMKJP(PE)	15860B..ZF...
UK	TOTMKUK(PI)	TOTMKUK(DY)	TOTMKUK(PE)	11260C..ZF...
USA	S&PCOMP(PI)	S&PCOMP(DY)	S&PCOMP(PE)	USCOD3M

Excess return and volatility are calculated as follows:

$$\text{Excess return}(xs_t) = [\ln(PI_t + PI_{t-1} \times DY_t) - \ln(PI_{t-1}) - \ln(1 + IR_t/12)]^{12},$$
$$\text{Volatility} = \sqrt{(xs_t - xs_{avg})^2}.$$

APPENDIX 7.B
State Space Model Estimation Algorithm
State Dynamic Subject to First Order Markov Chain Evolution

We discuss the problem of estimation of a state space model when only the state equation is subjected to multiple regimes and the switch between regimes take place according to a first order Markov chain. Both the mean and the variance part of the state equation may be subjected to this influence. Following the steps of this algorithm is important in understanding the accompanying program logic. Without the presence of this Markov chain, the state space system can be estimated using the standard Kalman filter recursion and updating algorithms that would be used to derive the prediction error form of the likelihood function.

The main issues that have to addressed due to the Markov chain driving the state transition are: 1) The probabilities of being in a particular state assuming that the system were in a given state in the previous time step, 2) How to cope with the exploding number of states to be accounted for as each observation is processed. For example, given that there are only two states possible, then at each time step there is a two-fold increase of number of state to account for. This implies that for a 100-time step observation of a given system there will be at the end 2^{100} states to be dealt with. This is clearly impractical. Hence, there is a need for approximating the system with a sensible approach. The algorithm discussed below deals with the first issue following the approach suggested in Hamilton (1989) and the second issue is addressed by the algorithm suggested in Kim (1994). Kim's procedure collapses the number of states to the previous number by a probability weighting scheme. Thus for a two-state Markov chain, we will always deal with two states after each observation input is processed.

The general structures to reference with respect these algorithms we focus on the system given in equation (7.B.1) as the measurement equation and the equation (7.B.2) as the state equation. Obviously, all the matrices and vectors are of compatible dimensions. The 2-state Markov chain $S_t = \{0, 1\}$ is used as a suffix to explicitly recognize those variables that may depend on the state we are in at any time. The equation (7.B.3) also states that the innovations of the measurement and the state equations are uncorrelated.

$$y_t = H\beta_t + Az_t + e_t, \tag{7.B.1}$$

$$\beta_t = M_{S_t} + F_{S_t}\beta_{t-1} + G_{S_t}v_t, \tag{7.B.2}$$

and

$$\begin{bmatrix} e_t \\ v_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R & 0 \\ 0 & Q_{S_t} \end{bmatrix}\right). \tag{7.B.3}$$

The covariance matrix of the innovations in measurement equation is given by R and that of the state equation is given by Q_{S_t}. This representation is somewhat general and not all the elements of the system could be present for a given problem. For example, the component Az_t is the measurement equation suggests possible presence of endogenous variables z_t entering the measurement process through the coefficient matrix A.

We next focus on the prediction and the updating equation of the basic Kalman filter assuming that in the previous time slot $S_{t-1} = i$ and the next time step it is

$S_t = j$. At this stage we define the probability transition matrix for the Markov chain variable. This is given by,

$$P = \begin{bmatrix} p_{00} & 1 - p_{11} \\ 1 - p_{00} & p_{11} \end{bmatrix}. \tag{7.B.4}$$

Following Harvey (1991) and Kim and Nelson (1999) and assuming that we are moving from state realization of i to state realization j, the relevant equations are given below.

Prediction:

$$\beta_{t|t-1}^{i,j} = M_j \beta_{t-1|t-1}^{i}, \tag{7.B.5}$$

$$P_{t|t-1}^{i,j} = F_j P_{t-1|t-1}^{i} F_j' + G_j Q_j G_j', \tag{7.B.6}$$

$$\eta_{t|t-1}^{i,j} = y_t - H\beta_{t|t-1}^{i,j} - Az_t, \tag{7.B.7}$$

$$f_{t|t-1}^{i,j} = HP_{t|t-1}^{i,j} H' + R, \tag{7.B.8}$$

Updating:

$$\beta_{t|t}^{i,j} = \beta_{t|t-1}^{i,j} + P_{t|t-1}^{i,j} H (f_{t|t-1}^{i,j})^{-1} \eta_{t|t-1}^{i,j}, \tag{7.B.9}$$

$$P_{t|t}^{i,j} = [I - P_{t|t-1}^{i,j} H' (f_{t|t-1}^{i,j})^{-1} H] P_{t|t-1}^{i,j}, \tag{7.B.10}$$

In the above equations, $\beta_{t-1|t-1}^{i}$ is the state vector estimated based upon information at time $(t-1)$, and the equation (7.B.5) states how it would evolve if at time t the state realization happens to be j. Similar interpretation applies to the estimate of the state covariance matrix, $P_{t-1|t-1}^{i}$ with respect to the equation (7.B.6). The equation (7.B.7) describes the forecast error at time t when the state realization is j assuming the previous state was j at time $(t-1)$. The equation (7.B.8) gives the covariance of the forecast error just discussed above. Thus, the equations (7.B.7) and (7.B.8) would provide the input required to build the state dependent conditional density of observations. The updating equations propel the equations (7.B.5) and (7.B.6) based upon the observations just made and makes it ready for use at the next time step. Thus, the basic nature of the Kalman filter is preserved; only these are now state contingent. Obviously, for this recursive procedure to work, we need to supply the prior starting values for $\beta_{0|0}$, $P_{0|0}$. We use the method discussed in Kim and Nelson (1999, p. 27).

For a two state Markov process, this recursion in the filter produces (2×2) posteriors for $\beta_{t|t}^{i,j}$ and $P_{t|t}^{i,j}$ when moving from $(t-1)$ to t. Kim (1994) develops the following approximation where by taking appropriate weighted average over the states at $(t-1)$ from which the particular state at t could be reached, this can be reduced to (2). We define the probability weighting as,

$$\Gamma^{i,j} = \frac{\Pr(S_{t-1} = i, S_t = j|\Psi_t)}{\Pr(S_t = j|\Psi_t)}, \tag{7.B.11}$$

where Ψ_t is the information available at time t. Therefore, the approximation for the state vector is,

$$\beta^j_{t|t} = \sum_{i=0}^{l} \beta^{i,j}_{t|t} \times \Gamma^{i,j}, \tag{7.B.12}$$

and the approximation for $P^j_{t|t}$ is,

$$P^j_{t|t} = \sum_{i=0}^{l} [P^{i,j}_{t|t} + (\beta^j_{t|t} - \beta^{i,j}_{t|t})(\beta^j_{t|t} - \beta^{i,j}_{t|t})'] \times \Gamma^{i,j}. \tag{7.B.13}$$

The equations (7.B.12) and (7.B.13) describe the nature of approximation applied to collapse the (2×2) posteriors to (2) posteriors with the help of the probability weighting factor. The detailed derivation of this could be found in Kim and Nelson (1999, p. 101). The probability terms necessary to achieve this can be obtained as follows:

$$\begin{aligned}
\Gamma^{i,j} &= \frac{\Pr(S_{t-1} = i, S_t = j|\Psi_t)}{\Pr(S_t = j|\Psi_t)} \\
&= \frac{\Pr(y_t, S_{t-1} = i, S_t = j|\Psi_{t-1})}{\Pr(y_t|\Psi_{t-1})} \\
&= (\frac{\Pr(y_t|S_{t-1} = i, S_t = j, |\Psi_{t-1})}{\Pr(y_t|\Psi_{t-1})}) \times \Pr(S_{t-1} = i, S_t = j|\Psi_{t-1}),
\end{aligned} \tag{7.B.14}$$

where as before $i = 0, 1$ and $j = 0, 1$. With the help of the forecast error in the prediction relations we can now construct the numerator (in the parentheses) of the last term of equation (7.B.14) as,

$$\Pr(y_t|S_{t-1} = i, S_t = j|\Psi_{t-1}) = \frac{1}{\sqrt{2\pi(f^{i,j}_{t|t-1})}} exp[-\frac{1}{2}(\eta^{i,j}_{t|t-1})'(f^{i,j}_{t|t-1})^{-1}(\eta^{i,j}_{t|t-1})], \tag{7.B.15}$$

and $\Pr(y_t|\Psi_{t-1})$ may be expressed as,

$$\Pr(y_t|\Psi_{t-1}) = \sum_{i=0}^{l}\sum_{j=0}^{l} \Pr(y_t, S_{t-1} = i, S_t = j|\Psi_{t-1}) \times \Pr(S_{t-1} = i, S_t = j|\Psi_{t-1}). \tag{7.B.16}$$

It may be recognized that the last product term in equation (7.B.14) is the transition probability. Furthermore, the separation of the joint probability in equation (7.B.14) is possible due to the Markov assumption. The equation (7.B.16) shows how to propagate the probability information as new observation is processed. This also gives the log likelihood function that has to be maximized with respect to all the unknown parameters in the model by using some suitable numerical optimization routine.

The logic of propagation of the probability information requires starting values at time 0. This is based on steady state probabilities of the assumed ergodic Markov chain. In this context we adopt the steps outlined in Kim and Nelson (1999, p. 71).

This Markov switching state space model generate, during the estimation process, the conditional variance of the forecast error given by equation (7.B.8) based upon a given state realization. Using the probability of the state occurring as discussed above, we could easily construct the conditional variance of the state process. The conditional variance is thus given by,

$$\sum_{0}^{l}\sum_{0}^{l} \Pr(S_{t-1} = i, S_t = j|\Psi_{t-1}) \times f_{t|t-1}^{i,j}.$$

In a similar manner the estimate of probability weighted forecast error could be generated using (7.B.7). This generated error series may then be analyzed for model diagnostics tests. Furthermore, we make inference of the expected state vector based on the relation given by equation (7.B.16) but the last term is replace by $\beta_{t|t}^{i,j}$ from the equation (7.B.9).

References

Ang, A. and Bekaert, G., (2002), Regime switches in interest rates, *Journal of Business and Economic Statistics*, 20, 163-182.

Ahlgren, N. and Antell, J., (2002), Testing for cointegration between international stock prices, *Applied Financial Economics*, 12, 851-861.

Baca, S., Garbe, B. and Weiss, R., (2000), The rise of sector effects in major equity mrkets, *Financial Analysts Journal*, September/October, 35-40.

Backus, D. K., and Gregory, A. W., (1993), Theoretical relations between risk premiums and conditional variances, *Journal of Business & Economic Statistics*, 11, 177-180.

Ball, C. A. and Torous, W. N., (1999), The stochastic volatility of short-term interest rate: Some international evidence, *Journal of Finance*, 54, 2339-2359.

Ball, L. and Cecchetti, S. G., (1990), Inflation and uncertainty at short and long horizons, *Brookings Papers on Economic Activity*, I, 215-254.

Bar-Shalom, Y. and Li, X-R., (1993), *Estimation and Tracking: Principles, Techniques and Software*, Artech House, Boston.

Baum, L. E., Petrie, T., Soules,, G. and Weiss, N., (1970), A maximisation technique occurring in statistical estimation of probabilistic functions of Markov chains, *Annals of Mathematical Statistics*, 41, 164-171.

Beveridge, S. and Nelson, C. R., (1981), A new approach to the decomposition of economic time series into permanent and transitory component with particular attention to measurement of the business cycle, *Journal of Monetary Economics*, 7, 151-174.

Bhar, R. and Chiarella, C., (1997), Transformation of the Heath-Jarrow-Morton models to Markovian system, *The European Journal of Finance*, 3, 1-26.

Bhar, R. and Hamori, S., (2003a), Alternative characterization of the volatility in the growth rate of real GDP, *Japan and the World Economy*, 15, 223-231.

Bhar, R. and Hamori, S., (2003b), New evidence of linkages among G7 stock markets, *Finance Letters*, 1, 35-40.

Bhar, R. and Hamori, S., (2003c), The link between inflation and inflation uncertainty, forthcoming in *Empirical Economics*.

Bhar, R. and Hamori, S., (2004), Empirical characteristics of the permanent and transitory component of stock return: analysis in a Markov switching heteroskedasticity framework, *Economics Letters*, 82, 157-165.

Bollerslev, T., (1986) Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307-327.

Bollerslev, T. Chou, R. Y and Kroner, K. F., (1992), ARCH modelling in finance: a review of the theory and empirical evidence, *Journal of Econometrics*, 52, 5-39.

Brock, W., Dechert, W., Scheinkman, J. and LeBaron, J., (1996), A test for independence based on the correlation dimension, *Econometric Review*, 15, 197-235.

Brooks, R. and Del Negro, M., (2002), The rise in comovement across national stock markets: market integration or IT bubble? Federal Reserve Bank of Atlanta, Working Paper 2002-17a.

Campbell, J. Y. and Mankiew, N. G., (1987) Are output fluctuation transitory?, *Quarterly Journal of Economics*, 102, 857-880.

Cashin, P., McDermott, C. J. and Scott, A., (1999), The myth of co-moving commodity prices, Reserve Bank of New Zealand Discussion Paper G99/9, Wellington.

Cecchetti, S. G., Lam, P. S. and Mark, N. C., (1990), Mean reversion in equilibrium asset prices, *American Economic Review*, 80, 398-418.

Cha, B. and Oh, S., (2000), The relationship between developed equity markets and the Pacific Basin's emerging equity markets, *International Review of Economics and Finance*, 9, 299-322.

Chan, K. C., Karolyi, G. A. and Stulz, R. M., (1992), Global financial markets and the risk premium on U.S. equity, *Journal of Financial Economics*, 32, 137-167.

Chauvet, M., (1998), An econometric characterization of business cycle dynamics with factor structure and regime switching, *International Economic Review*, 39, 969-996.

Chauvet, M., (1998/1999), Stock market fluctuations and the business cycle, *Journal of Economic and Social Measurement*, 25, 235-257.

Chauvet, M. and Potter, S., (2000), Coincident and leading indicators of the stock market, *Empirical Finance*, 7, 87-111.

Chou, R. Y. and Ng, V. K., (1995), Correlation structure of the permanent and temporary components o international stock market pries, Discussion paper No. 9512, Institute of Economics, Academia Sinica, Narking, Taipei, Taiwan.

Chu, C. J., Santoni, G. J. and Liu, T., (1996), Stock market volatility and regime shifts in returns, *Information Sciences*, 94, 179-190.

Clark, P. K., (1987), The cyclical component of U.S. economic activity, *Quarterly Journal of Economics*, 102, 797-814.

Cochrane, J. H., (1988), How big is the random walk in GNP?, *Journal of Political Economy*, 96, 893-920.

Cooley, T. and Hansen, G., (1991), The welfare costs of moderate inflation, *Journal of Money, Credit and Banking*, 23, 483-503.

Corhay, A., Rad, A. T., and Urbin, J-P., (1993), Common stochastic trends in European stock markets, *Economics Letters*, 42, 385-390.

Cosimano, T. F. and Jansen, D. W., (1988), Estimates of the variance of U.S. inflation based upon the ARCH model: Comment, *Journal of Money, Credit and Banking*, 20, 409-421.

Cox, J. C., Ingersoll, J. E. and Ross, S. A., (1985), A theory of the term structure of interest rates, *Econometrica*, 53, 385-407.

Danielsson, J., (1998), Multivariate stochastic volatility models: Estimation and a comparison with VGARCH models, *Journal of Empirical Finance*, 5, 155-173.

Das, S. R., (1996), Revisiting Markov chain term structure models: Extensions and applications, *Financial Practice and Education*, Spring/Summer.

Dempster, A. P., Laird, N. M. and Rubin, D. B., (1977), Maximum likelihood from incomplete data via the EM algorithm, *Journal of Royal Statistical Society*, B 39, 1-38.

Dickey, D. and Fuller, W. A., (1979), Distribution of the estimates for autoregressive time series with a unit root, *Journal of he American Statistical Society*, 74, 427-431.

Diebold F., (1986), Comment on modelling the persistence of conditional variance, *Econometric Reviews*, 5, 51-56.

Diebold, F. X., Gunther, T. A. and Tay, A. S., (1998), Evaluating density forecasts, with application to risk management, *International Economic Review*, 39, 863-883.

Drost, F. C. and Nijman, T. E., (1993), Temporal aggregation of GARCH processes, *Econometrica*, 61, 909-927.

Engle, R. F., (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* 50, 987-1008.

Engsted, T. and Lund, J., (1997), Common stochastic trends in international stock prices and dividends: an example of testing overidentifying restrictions on multiple cointegrating vectors, *Applied Financial Economics*, 7, 659-665.

Eun, C. and Shim, S., (1989), International Transmission of Stock Market Movements, *Journal of Financial and Quantitative Analysis*, 24, 241-256.

Evans, M., (1991), Discovering the link between inflation rates and inflation uncertainty, *Journal of Money, Credit and Banking*, 23, 169-184.

Evans, M. and Wachtel, P., (1993), Inflation Regimes and Sources of Inflation Uncertainty, *Journal of Money, Credit and Banking*, 25, (3), 475-511.

Fama, E. F. and French, K. R., (1988), Permanent and temporary components of stock prices, *Journal of Political Economy*, 96, 246-273.

French, K. R., Schwert, G. W. and Stambaugh, R. F., (1987), Expected stock returns and volatility, *Journal of Financial Economics*, 19, 3-29.

Friedman, M., (1977), Nobel lecture: Inflation and unemployment, *Journal of Political Economy*, 85, 451-472.

Fuller, W. A., (1976), *Introduction to statistical time series*, John Wiley and Sons.

Glosten, L. R., Jagannathan R. and Runkle D. E., (1993), On the relation between the expected value of and the volatility of nominal excess return on stocks, *Journal of Finance*, 53, 1779-1801.

Graflund, A., (2000), A Bayesian inference approach to testing mean reversion in the Swedish stock market, Department of Economics Working Paper 8, Lund University, Sweden.

Gray, S., (1996), Modelling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics*, 42, 27-62.

Grier, K. and Perry, M. J., (1996), Inflation, inflation uncertainty, and relative price dispersion: Evidence from bivariate GARCH-M models, *Journal of Monetary Economics*, 38, 391-405.

Grier, K. and Perry, M. J., (1998), On inflation and inflation uncertainty in the G7 countries, *Journal of International Money and Finance*, 17, 671-689.

Guha, D. and Banerji, A., (1998/1999), Testing for regional cycles: a Markov-switching approach, *Journal of Economic and Social Measurement*, 25, 163-182.

Hamilton, J. D., (1989), A new approach to the economic analysis of non-stationary time series and the business cycle, *Econometrica*, 57, 357-384.

Hamilton, J. D., (1990), Analysis of time series subject to changes in regime, *Journal of Econometrics*, 45, 39-70.

Hamilton J. D., (1994), *Time Series Analysis*, Princeton University Press.

Hamilton, J. D. and Susmel, R., (1994), Autoregressive conditional heteroscedasticity and changes in regime, *Journal of Econometrics*, 64, 307-333.

Hamilton J. D. and Lin, G., (1996), Stock market volatility and the business cycle, *Journal of Applied Econometrics*, 11, 573-593.

Hamori, S., (2000), Volatility of real GDP: some evidence from the United States, the United Kingdom and Japan, *Japan and the World Economy*, 12, 143-152.

Harding, D. and Pagan, A., (1999), Dissecting the cycle, Melbourne Institute Working Paper No. 13/99, Melbourne: University of Melbourne, Australia.

Harvey, A. C., (1990), *The Econometric Analysis of Time Series*, second edition, Cambridge, Massachusetts: The MIT Press.

Harvey, A. C., (1991), *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge, U.K.

Harvey, A. C., Ruiz, E. and Sentana, E., (1992), Unobserved Component Time Series Models with ARCH Disturbances, *Journal of Econometrics*, 52, 129-157.

Heath, D., Jarrow, R. A. and Morton, A. J., (1992), Bond pricing and the term structure of interest rates: a new methodology for contingent claim valuation, *Econometrica*, 60, 77-105.

Hess, G. D. and Morris, C. S., (1996), The Long-Tun Costs of Moderate Inflation, *Economic Review*, Federal Reserve Bank of Kansas City, Second Quarter.

Holland, S., (1995), Inflation and Uncertainty: Tests for Temporal Ordering, *Journal of Money, Credit and Banking*, 27 (3), 827-837.

Jarque C. M. and Bera A. K., (1987) Tests for Normality of Observations and Regression Residuals, *International Statistical Review* 55, 163-172.

Jochum, C., (1999), Volatility spillovers and the price of risk: evidence from the Swiss stock market, *Empirical Economics*, 24, 303-322.

Jones, C. P., and Wilson, J. M., (1989), Is stock price volatility increasing?, *Financial Analysts Journal*, 45, 20-26.

Judson, R. and Orphanides, A., (1996), Inflation, volatility and growth, Working paper, Board of Governors of the Federal Reserve System, May.

Kasa, K., (1992), Common stochastic trends in international stock markets, *Journal of Monetary Economics*, 29, 95-124.

Kim, C-J., (1993), Unobserved-Component Models With Markov Switching Heteroscedasticity: Changes in Regime and the Link Between Inflation rates and Inflation Uncertainty, *Journal of Business and Economic Statistics*, 11, 341-349.

Kim, C-J., (1994), Dynamic linear models with Markov switching, *Journal of Econometrics*, 60, 1-22.

Kim, C-J. and Kim, M-J., (1996), Transient fads and the crash of '87, *Journal of Applied Econometrics*, 11, 41-58.

Kim, D. and Kon, S. J., (1996), Sequential parameter nonstationarity in stock market returns, *Review of Quantitative Finance and Accounting*, 6, 103-131.

Kim, D. and Kon, S. J., (1999), Structural change and time dependence in models of stock returns, *Journal of Empirical Finance*, 6, 283-308.

Kim, C-J, and Nelson, C. R., (1998), Testing for mean-reversion in heteroscedastic data II: Autocorrelation tests based on Gibbs-sampling augmented randomization, *Journal of Empirical Finance*, 5, 385-396.

Kim, C-J, and Nelson, C. R., (1999), *State Space Models with Regime Switching, Classical and Gibbs Sampling Approaches with Applications*, The MIT Press, Cambridge, Massachusetts.

Kim, C-J. Nelson, C. R. and Piger, J., (2001), The less volatile U.S. economy: a Bayesian investigation of timing, breadth and potential explanations", International Finance Discussion Papers, No. 707, August, Board of Governors of the Federal Reserve System.

Kim, M-J., Nelson, C. R. and Startz, R., (1991), Mean reversion in stock prices? a reappraisal of the empirical evidence, *Review of Economic Studies*, 58, 515-528.

Kim, C-J., Nelson, C. R. and Startz, R., (1998), Testing for mean reversion in heteroscedastic data based on Gibbs sampling augmented randomisation, *Journal of Empirical Finance*, 5, 131-154.

Kloeden, P. E. and Platen, E., (1992), *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag.

Kunsch, H. R., (1989), The jackknife and the bootstrap for general stationary observations, *Annals of Statistics*, 17, 1217-1241.

Lamoureux, C. G. and Lastrapes, W., (1990), Persistence in variance, structural change, and the GARCH model, *Journal of Business and Economic Statistics*, 8, 225-234.

Lastrapes, W., (1989), Exchange rate volatility and U.S. monetary policy: An ARCH application, *Journal of Money, Credit and Banking*, 21, 66-77.

Lo, A. W. and MacKinlay, A. C., (1988), Stock market prices do not follow random walks: evidence from a simple specification test, *Review of Financial Studies*, 1, 41-66.

Ljung G. and Box G., (1979), On a measure of lack of fit in time series models, *Biometrika*, 66, 265-270.

Lyn C. T., Allen, D. E., Morkel-Kingsbury N., (2002), A hidden Markov chain model for the term structure of bond credit risk spreads, *International Review of Financial Analysis*, Vol. 11, pp. 311-329.

MacDonald, I. L. and Zucchini, W., (1997), *Hidden Markov and Other Discrete-valued Time Series*, Chapman and Hall, London.

Malliaris, A. G. and Urrutia, J. L., (1992), The International crash of October 1987: Causality Tests, *Journal of Financial and Quantitative Analysis*, 27, 353-364.

Malliaris, A. G. and Urrutia, J. L., (1997), Equity and Oil Markets Under External Shocks in Ghosh, D. and Ortiz, E., eds. *Global Structure of Financial Markets*, London, U.K., Routlegde Publishers, 103-116.

Manton, J., Muscatelli, A., Krishnamurthy, V. and Hurn, S., (1998), Modelling stock market excess returns by Markov modulated Gaussian noise, Working paper, Department of Economics, University of Glasgow.

McCarthy, J. and Najand, M., (1995), State Space Modeling of Linkages Among International markets, *Journal of Multinational Financial Management*, 5, 1-9.

McConnell, M. M. and Quiros, G. P., (2000), Output fluctuations in the United States: what has changed since the early 1980's?, *American Economic Review*, 90, 1464-1476.

McDermott, C. J. and Scott, A., (1999), Concordance in business cycles, Reserve Bank of New Zealand Discussion Paper G99/7, Wellington.

Morley, J. C., (2002), A state-space approach to calculating the Beveridge-Nelson decomposition, *Economics Letters*, 75, 123-127.

Nelson, D. B., (1991), Conditional heteroskedasticity in asset returns: a new approach, *Econometrica*, 59, 347-370.

Nelson, C. R. and Plosser, C. I., (1982), Trends and random walks in macroeconomic time series: some evidence and implications, *Journal of Monetary Economics*, 10, 139-162.

Pesaran, M. H. and Timmermann, A., (1995), Predictqbility of stock returns: robustness and economic significance, *Journal of Finance*, 50, 1201-1228.

Porterba, J. M. and Summers, L. H., (1988), Mean reversion in stock prices: evidence and implications, *Journal of Financial Economics*, 22, 27-59.

Perron P., (1989), The great crash, the oil crisis, and the unit root hypothesis, *Econometrica*, 57, 1361-1401.

Pliska, S. R., (1997), *Introduction to Mathematical Finance, Discrete Time Models*, Blackwell Publishers, Oxford, U.K.

Rabiner, L. A., (1989), A tutorial on Hidden Markov Models and selected applications in speech recognition, *Proceedings of the IEEE*, 77, 257-286.

Raymond, J. and Rich, R. W., (1992), Changes in Regime and the Behaviour of Inflation, Working paper, Vanderbilt University, Department of Economics.

Richards, A. J., (1995), Comovements in National Stock Market Returns: Evidence of Predictability, but Not Cointegration, *Journal of Monetary Economics*, 36, 631-654.

Ross, S. M., (2000), *Introduction to Probability Models*, Seventh Edition, Academic Press, New York.

Sahani, M., (1999), *Latent variable models for neural data analysis*, PhD Thesis, California Institute of Technology.

Schaller, H. and Van Norden, S., (1997), Regime switching in stock market returns, *Applied Financial Economics*, 7, 177-191.

Schwert, G. W., (1989), Why does stock market volatility change over time? *Journal of Finance*, 44, 1115-1153.

Schwert, G. W., (1990), Stock Market Volatility, *Financial Analysts Journal*, 46, 23-34.

Shiller, R. J., (1993), *Market Volatility*, The MIT Press, Cambridge, Massachusetts.

Shumway, R.H. and Stoffer, D. S., (2000), *Time Series Analysis and Its Applications*, Springer Text in Statistics, New York.

Smith, D. R., (2002), Markov switching and stochastic volatility diffusion models of short-term interest rates, *Journal of Business and Economic Statistics*, 20, 183-197.

Stock, J. H. and Watson, M. W., (1991), A probability model of the coincident economic indicator, in *Leading Economic Indicators: New Approaches and Forecasting Records*, ed. K. Lahiri and G. H. Moore, Cambridge University Press, 63-89.

Tapeiro, C. S., (1998), *Applied Stochastic Models and Control for Finance and Insurance*, Kluwer Academic Press, Dordrecht.

Tsay, R. and Chou, R., (1994), Correlation transformation and components of stock prices, Working Paper, Graduate School of Business, University of Chicago.

Thomas, L. C., Allen, D. E. and Morkel-Kingsbury, N., (2002), A hidden Markov chain model for the term structure of bond credit risk spreads, *International Review of Financial Analysis*, 11, 311-329.

Turner, C. M., Startz, R. and Nelson, C. R., (1989), A Markov model of heteroscedasticity, risk, and learning in the stock market, *Journal of Financial Economics*, 25, 3-32.

Vuong, Q. H., (1989), Likelihood ratio test for model selection and non-nested hypotheses, *Econometrica*, 57 (2), 307-333.

Watson, M. W., (1986), Univariate detrending methods with stochastic trends, *Journal of Monetary Economics*, 18, 49-75.

Weigend, A. S., Mangeas, M. and Srivastava, A. N., (1995), Nonlinear gated experts for time series: Discovering regimes and avoiding overfitting, *International Journal of Neural Systems*, 6, 373-399.

Weigend, A. S. and Shi, S., (1998), Predicting daily probability distributions of S&P500 returns, Working paper IS-98-23, Leonard N. Stern School of Business, New York University.

Wells, C., (1996), *The Kalman Filter in Finance*, Amsterdam, Kluwer Academic Publishers.

Whitelaw, R. F., (1994), Time variations and covariations in the expectations and volatility of stock market returns, *Journal of Finance*, 49, 515-541.

This page intentionally left blank

Index

ARCH model, 39

Baum-Welch algorithm, 18

coincident financial market indicators, 127
components of forecast variance: Germany, 103
components of forecast variance: Japan, 104
components of forecast variance: UK, 105
components of forecast variance: USA, 106
components of variance of inflation forecast with process switching, 114
concordance measure, 45
concordance statistics, 75
concordance statistics between probability of high variance state, 50
correlation between probabilities of expansion states, 74
correlation between the coincident financial indicator and its components, 133
correlation of stock returns, 47
correlation statistics between probability of high variance state, 50
correlations: return, permanent and temporary components, 123

diagnostic statistics for the residuals of measurement equations, 133
diagnostics using standardized residuals from excess return equation, 66
diagnostics using standardized residuals from GARCH(1,1) Model, 34
diagnostics using standardized residuals from Markov switching heteroskedasticity model, 35
diagnostics using standardized returns from Markov switching heteroscedasticity model, 49

EM (expectation maximization) algorithm, 18
emission probability matrix, 16
estimated coincident indicator and excess return: Japan, 136
estimated coincident indicator and excess return: UK, 136
estimated coincident indicator and excess return: USA, 137
estimated filtered probability: Canada, 67
estimated filtered probability: France, 68
estimated filtered probability: Germany, 69
estimated filtered probability: Italy, 70
estimated filtered probability: Japan, 71
estimated filtered probability: UK, 72
estimated filtered probability: USA, 73
estimated Markov probabilities of staying in the same state for the G7 countries, 49
estimated variance from the Markov switching heteroskedasticity model: Japan, 36
estimated variance from the Markov switching heteroskedasticity model: UK, 37
estimated variance from the Markov switching heteroskedasticity model: USA, 37
excess return: Canada, 67
excess return: France, 68
excess return: Germany, 69
excess return: Italy, 70
excess return: Japan, 71
excess return: UK, 72
excess return: USA, 73

filtering algorithm, 13

GARCH model, 31, 39
Granger causality tests, 134
growth rate in real GDP: GARCH(1,1) estimation, 34
growth rate in real GDP: Markov switching heteroskedasticity estimation, 35

hidden Markov experts, 13
hidden Markov model (HMM), 13
hidden states, 16
HMM Estimation Algorithm, 16
HMM illustrative examples, 24
HMM parameter estimation, 18

industrial production growth: Canada, 67
industrial production growth: France, 68
industrial production growth: Germany, 69
industrial production growth: Italy, 70
industrial production growth: Japan, 71
industrial production growth: UK, 72
industrial production growth: USA, 73
inferred probability of low variance state: Japan, 137
inferred probability of low variance state: UK, 138
inferred probability of low variance state: USA, 138
inflation, 81
inflation rate: Germany, 88
inflation rate: Japan, 88
inflation rate: UK, 89
inflation rate: USA, 89
inflation uncertainty, 81
interplay between industrial production and stock market, 55

Kalman Filter, 11
Kalman smoother, 12

linkages among G7 stock markets, 41

Markov chains, 1
Markov Switching Coincidence Index Model, 129
Markov switching heteroscedasticity model of output and equity, 58
Markov switching heteroscedasticity model of stock return, 119
Markov switching heteroscedasticity model of the inflation rate, 85
Markov switching stock return model, 44
Markov switching variance model, 32

observable states, 16

parameter estimates for the Markov switching coincidence index model, 132
parameter estimates: bivariate Markov switching heteroscedasticity model of output and equity, 64
parameter estimates: Markov switching heteroscedasticity model of inflation rate, 94
parameter estimates: Markov switching heteroscedasticity model of stock return, 48
passage time, 5
permanent and transitory components of equity return: Markov switching heteroscedasticity framework, 122
permanent components of stock return, 117
pi-vector, 16
probability of high variance state for permanent shocks: Germany, 96
probability of high variance state for permanent shocks: Japan, 96
probability of high variance state for permanent shocks: UK, 97
probability of high variance state for permanent shocks: USA, 97
probability of high variance state for transitory shocks: Germany, 98
probability of high variance state for transitory shocks: Japan, 98
probability of high variance state for transitory shocks: UK, 99
probability of high variance state for transitory shocks: USA, 99

regime classification measure (RCM), 101
residual diagnostics and model adequacy tests, 100, 123

state space model (SSM), 11
state space model estimation algorithm, 141
state transition matrix, 16
summary statistics, 87, 121
summary statistics on stock return, 47
summary statistics: monthly excess return, 62
summary statistics: monthly industrial production growth, 63

term structure of interest rates, 6
transitory components of stock return, 117

unobserved component model, 109

Viterbi algorithm, 22

volatility in growth rate of real GDP, 29

Vuong statistic, 86, 101