This is to certify that the

dissertation entitled

Strongly Dependent Economic Time Series: Theory and Applications

presented by

Margie A. Tieslau

has been accepted towards fulfillment of the requirements for

Ph.D. degree in Economics

R. T. Baillie

Major professor

Date 7.22 1992

STRONGLY DEPENDENT ECONOMIC TIME SERIES: THEORY AND APPLICATIONS

By

Margie A. Tieslau

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

1992

ABSTRACT

STRONGLY DEPENDENT ECONOMIC TIME SERIES: THEORY AND APPLICATIONS

By

Margie A. Tieslau

This dissertation contains three essays which address the issues of persistence and stationarity of macroeconomic time series, estimation of persistent series through the procedure of generalized method of moments, and the existence of stable long-run money demand functions and stable real exchange rates.

The first essay investigates the Autoregressive Fractionally Integrated Moving Average ARFIMA(p,d,q) model as a process for describing strongly persistent time series. The essay proposes the ARFIMA(p,d,q)-GARCH(P,Q) model which offers the flexibility of modelling strongly dependent series that are also characterized by non-homoskedastic error. The ARFIMA-GARCH model is applied to the inflation rate series for both low- and high-inflation economies and the inflation rate is found to be highly persistent but, none the less, mean reverting and thus stationary. In addition, empirical support is found for the relatively high-inflation economies for the Friedman hypothesis which implies a direct association between increased levels of inflation and increased inflation variability.

The second essay presents the theoretical derivation of a generalized method of moments (GMM) estimator for strongly dependent time series. The GMM estimation technique may be preferred to maximum likelihood estimation methods since GMM does not rely on distributional

assumptions. The moment conditions exploited by the GMM estimator make use of theoretical and estimated autocorrelation functions, and the derivation of these functions is presented here. Numerical results from variance calculations using all available moment conditions as well as groups of moment conditions are presented to examine the relative efficiency of the GMM estimator.

The third essay investigates the existence of cointegrating relationships between macroeconomic variables in Canada and the U.S. This essay examines the long-run equilibrium relationship between real money balances, real income, and short-term interest rates for Canada and the U.S., and investigates the existence of cross-country effects between these countries. Evidence of stable long-run money demand functions is found for both countries with estimated long run income elasticities near 1.0 and long-run interest elasticities near -0.5. The stationarity of the nominal exchange rate and the relative price levels for Canada and the U.S. is also investigated with some evidence found in support of this hypothesis for the post-Bretton Woods era.

This dissertation is dedicated to my mother, Herta Tieslau,

and to the memory of my father, Adolf Tieslau.

ACKNOWLEDGEMENTS

The successful completion of this dissertation would not have been possible without the input, advice, assistance and encouragement of many individuals throughout the entire process. I am happy to have this opportunity to express my gratitude to these individuals, although I realize that it may be impossible to express, in these few lines, the depth of my appreciation to the many who are deserving.

My greatest appreciation must go to my dissertation committee chairman, Professor Richard T. Baillie, who introduced me to some of the topics of this dissertation and who inspired my work in this area. I have had the great fortune of working under his expert guidance for the past three years, and during this time I have benefitted greatly from his experience and wisdom. He has endured an endless stream of questions on my part and has read repeated drafts of this dissertation, offering many suggestions for improvement. I am indebted to him for the time and effort he has given to me and to my work, and for the kindness and patience that he has shown me throughout the completion of this dissertation.

I have also had the good fortune of working closely with the two remaining members of my dissertation committee, Professors Robert H. Rasche and Peter Schmidt, whose guidance and direction have been invaluable in completing this dissertation. I have learned a great deal from them and have benefitted greatly from their knowledge and advice. In addition, I would like to express my appreciation to Professor Rowena

Pecchenino who read several earlier drafts of parts of this dissertation and offered many valuable comments. I am also grateful to her for all of the support and advisement that she offered to me throughout my four years at Michigan State University.

I owe a debt of gratitude, too, to all of my colleagues in the Ph.D. program at Michigan State University who shared with me the toils of class work and prelims, and who continuously offered me words of encouragement when necessary (which was often). I am particularly indebted to my office mate and friend, Kel Utendorf, who bravely endured my many moods with a smile and a positive word, even through the darkest hours, and never asked for anything in return.

My success throughout graduate school could not have been possible without the love and support of my wonderful family. My mother has continuously offered me her unwavering support, even when she personally may have preferred a different course for my life. My father stood by me from the beginning and left me with an unconditional belief in myself that has carried me through many rough times. I thank the rest of my family, too, for their support and encouragement. My only regret is that my work on this dissertation has meant many missed hours with them.

Finally, no account of my gratitude could be complete without a special word of thanks to Michael Redfearn whose undying support and encouragement, both academic and emotional, have been a constant source of strength and inspiration to me these past four years.

TABLE OF CONTENTS

CHAPTER IV: A GENERALIZED METHOD OF MOMENTS ESTIMATOR FOR LONG-MEMORY PROCESSES

CHAPTER V: EQUILIBRIUM MONEY DEMAND FUNCTIONS, REAL EXCHANGE RATES, AND PPP: AN ANALYSIS OF COINTEGRATING RELATIONSHIPS IN CANADA AND THE US

LIST OF TABLES

CHAPTER IV

CHAPTER V

I. INTRODUCTION

CHAPTER I

INTRODUCTION

Many economic and financial time series are characterized by strong persistence in their mean and variance such that when expressed in levels the series appear to be nonstationary, or contain a unit root, yet when expressed in first differences the series appear to be overdifferenced. The traditional nonstationary Autoregressive-Integrated Moving Average ARIMA(p,d,q) model of Box and Jenkins (1976), where the parameter of integration can take on only integer values, is not able to account for the long-term persistence which is characteristic of these series. This can be especially relevant in applied analysis since empirical investigation of such series must be done within the framework of a model which is able to account for the substantially high order of correlation present in these series.

The first essay of this dissertation, Chapters II and III, considers the long-run characteristics of the CPI inflation rate series which is one macroeconomic variable that exhibits characteristics common to strongly persistent time series. This variable is considered within the context of the Autoregressive Fractionally Integrated Moving Average ARFIMA(p,d,q) process introduced by Granger and Joyeux (1980) and Hosking (1981), which allows for the order of integration of a series to be fractional, as well as the Generalized Autoregressive Conditionally Heteroskedastic GARCH(P,Q) process of Engle (1982) and Bollerslev (1986). In this way the model is able to allow for simultaneous modelling of both long-term persistence and time varying conditional variance which have been found to be present in the inflation rate series.

The application to the inflation rate series is particularly appealing since this series does not appear to be well described by either the traditional stationary I(0) process or the non-stationary I(1) process. In addition, this essay investigates the validity of the hypothesis posited by Friedman in his 1977 Nobel lecture that increased levels of inflation were likely to be associated with increased levels of the variance of inflation. Empirical support for such a relationship would imply a significant role in the economy for policy directed at maximizing net benefit.

Several estimation techniques have been developed in recent years to estimate the degree of persistence of strongly dependent time series utilizing both ordinary least squares and maximum likelihood estimation procedures (Janacek (1982), Geweke and Porter-Hudak (1983), Hosking (1984), Fox and Taqqu (1986), and Sowell (1992)). The second essay, Chapter IV, proposes a generalized method of moments (GMM) estimator to determine the degree of fractional integration of a strongly dependent series. The estimation technique is set within the context of the general ARFIMA(0,d,0) process, but may be extended to include autoregressive and moving average parameters in the model as well. The GMM estimation technique provides an attractive alternative estimation procedure for the fractionally integrated process since it does not require the distributional assumptions necessary under maximum likelihood estimation techniques.

This essay also examines the asymptotic performance of the generalized method of moments estimator and compares its efficiency relative to that of the maximum likelihood estimator. Attention is restricted, in this analysis, to that range of the parameter of fractional integration for which the moment conditions exploited by the model exhibit the usual $\sqrt{T}$ consistency and normality.

The final essay of this dissertation, Chapter V, addresses the issues of stable empirical money demand functions and the stationarity of the nominal exchange rate and relative price levels for the countries of Canada and the U.S. The investigation is set within the framework of the monetary balance of payments theory and utilizes the methodology proposed by Johansen (1988) and Johansen and Juselius (1989) to identify cointegrating relationships among sets of variables. The investigation of long-run equilibrium relationships within and between the economies of Canada and the U.S. is especially appropriate due to the parallel nature of macroeconomic variables in the two countries.

The framework used in this essay allows for an investigation into the existence of cross-country relationships between monetary and international variables in the two countries and allows for an investigation into the degree of similarity in the dynamics between the two countries. Empirical evidence of cross-country effects or similarity in dynamics between the two countries can have significant implications for the conduct of monetary policy and international transactions within and among the two countries.

REFERENCES

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", Journal of Econometrics, 31, 307-327.

Box, G.E.P. and G.M. Jenkins (1976), Time Series Analysis Forecasting and Control, second edition, San Francisco: Holden Day.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation", Econometrica, 50, 987-1008.

Fox, R. and M.S. Taqqu (1986), "Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time-Series", Annals of Statistics, 14, 517-532.

Friedman, M. (1977), "Nobel Lecture: Inflation and Unemployment", Journal of Political Economy, 85, 451-472.

Geweke, J. and S. Porter-Hudak (1983), "The Estimation and Application of Long Memory Time Series Models", Journal of Time Series Analysis, 4, 221-238.

Granger, C.W.J. and R. Joyeux (1980), "An Introduction to Long Memory Time Series Models and Fractional Differencing", Journal of Time Series Analysis, 1, 15-39.

Hosking, J.R.M. (1981), "Fractional Differencing", Biometrika, 68, 165-176.

Hosking, J.R.M. (1984), "Modelling Persistence in Hydrologic Time Series Using Fractional Differencing", Water Resources Research, 20, 1898-1908.

Janacek, G.J. (1982), "Determining the Degree of Differencing for Time Series via the Log Spectrum", Journal of Time Series Analysis, 3, 177-83.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", Journal of Economic Dynamics and Control, 12, 231-254.

Johansen, S. and K. Juselius (1989), "The Full Information Maximum Likelihood Procedure for Inference on Cointegration with Applications", Preprint No. 4, Institute of Mathematical Statistics, University of Copenhagen.

Sowell, F.B. (1992), "Maximum Likelihood Estimation of Stationary Univariate Fractionally-Integrated Time-Series Models", Journal of Econometrics, forthcoming.

II. THE THEORETICAL CHARACTERISTICS AND ESTIMATION PROCEDURES ASSOCIATED WITH LONG-MEMORY TIME SERIES MODELS

CHAPTER II

THE THEORETICAL CHARACTERISTICS AND ESTIMATION PROCEDURES ASSOCIATED WITH LONG-MEMORY TIME SERIES MODELS

1. INTRODUCTION

The use of time series analysis in the field of economics is instrumental in investigating the long-run properties of various economic series. Often times it is the case when fitting parametric models to a time series that the series under investigation initially requires first differencing in order to achieve stationarity. The standard Box and Jenkins (1976) methodology is to difference a series an integer number of times until it appears stationary. The Autoregressive Moving Average (ARMA) model of Box and Jenkins has been used to model the long-run behavior of certain series by expressing today's realization of a variable as a weighted sum of past observations of that variable and past and present innovations. ARMA models assume that the dependence between observations decays exponentially and therefore is relatively weak. However, many economic and financial time series, while being stationary, are characterized by the property of relatively significant dependence between observations which occur at distant intervals. Such series, although stationary, have been termed "long memory".

Two general classes of models have arisen in modelling series which exhibit long memory. The first is the fractional Brownian motion process, which is a generalization of Brownian motion. The use of fractional Brownian motion was motivated by the existence of long-term persistence which was observed by hydrological engineers in streamflow data.

Consequently, the fractional Brownian motion process has been widely applied in the field of hydrology. The second model is the fractionally integrated process, which is a generalization of the Autoregressive-Integrated Moving Average (ARIMA) model of Box and Jenkins (1976). Use of the fractionally integrated process was motivated by the observation that the conventional stationary models of Box and Jenkins were not able to capture the persistent nature of long-memory time series. This class of model, which was proposed independently by Granger and Joyeux (1980) and Hosking (1981), offered an alternative to the standard ARIMA(p,d,q) process by not restricting the degree of differencing, the parameter d, to be limited to an integer value but rather allowing it to take on fractional values.

Fractionally integrated processes are capable of generating extreme persistence represented by the hyperbolic decline of the impulse response weights and the autocorrelation functions. Hence fractionally integrated processes allow for a much slower decay of past observations of the series, relative to the faster exponential decline of the weights of the traditional ARMA model.

This chapter will provide a comprehensive survey of long-memory processes, outlining the theoretical properties of series which exhibit strong dependence and discussing the techniques used to model these processes. The next section will discuss both sample and population characteristics of long-memory time series and also will discuss the characteristics of the Autoregressive-Fractionally Integrated Moving Average process. Section 3 will survey the Brownian motion process and fractional Brownian motion as a method used to model long-term behavior. This section will examine the estimation procedures involving fractional

Brownian motion which have been used to estimate the degree of persistence of a series exhibiting long memory. Section 4 will survey the fractionally integrated process as an alternative method of modelling long-term behavior, examining various estimation techniques involving the fractionally integrated process which have been proposed to estimate the degree of persistence (the degree of fractional differencing) of a series as well as any other model parameters. Section 5 provides a description of the theoretical properties of the model which is developed in this dissertation, the ARFIMA-GARCH process. This process offers the flexibility of modelling series which exhibit long-term persistence compounded with the presence of non-normal or non-homoskedastic error. In Chapter III, this process is applied in a macroeconomic analysis of the aggregate price level. The final section of the current chapter presents a brief summary and conclusion.

2. THEORETICAL CHARACTERISTICS OF LONG-MEMORY TIME SERIES AND THE ARFIMA(p,d,q) PROCESS

Consider a stationary time series, $\{y_t\}$, which can be expressed in ARIMA(p,d,q) form as

$$(1.) \qquad \phi(L)(1 - L)^d (y_t - \mu) = \theta(L)\epsilon_t,$$

or alternatively as

$$(1.') \qquad (1 - L)^d (y_t - \mu) = \theta(L)/\phi(L)\ \epsilon_t = u_t.$$

In this expression, L represents the lag operator, the polynomial $\phi(L)$

incorporates the autoregressive parameters of the model and is expressed as $\phi(L) = 1 - \phi_1 L^1 - \phi_2 L^2 - \ldots - \phi_p L^p$, the polynomial $\theta(L)$ incorporates the moving average parameters of the model and is expressed as $\theta(L) = 1 - \theta_1 L^1 - \theta_2 L^2 - \ldots - \theta_q L^q$, all roots of $\phi(L)$ and $\theta(L)$ lie outside the unit circle, and u_t is iid. In the case where d is zero the model in (1.') simply reduces to the standard ARMA process. However, in the case where d is assumed to be fractional,[1] the model expressed in (1.) takes on characteristics unique to long-memory time series.

From the expression in (1.), the difference operator $(1 - L)^d$ can be expressed in terms of the difference parameter, d, as:

$$(1 - L)^d = [1 - dL^1 + d(d-1)/2!L^2 - d(d-1)(d-2)/3!L^3 + \ldots] \tag{2.}$$

Similarly, this expression can be defined by the Binomial Theorem as:

$$(1 - L)^d = \sum_{j=0}^{\infty} \begin{bmatrix} d \\ j \end{bmatrix} (-1)^j L^j$$

or as defined by the Maclaurin series:

---

[1] The range of values of d is restricted to $-\frac{1}{2} < d < \frac{1}{2}$. Series containing a value of d within this range are stationary and mean reverting. The value of d may be expressed outside of this range as well, for example a value of d equal to .80, in which case it is assumed that the series requires first differencing in order to achieve a value of d corresponding to the range stated above. That is, a value of d = .80 for a non first differenced series corresponds to a value d = -.20 when the series is appropriately first differenced, etc.

$$(1 - L)^d = \sum_{j=0}^{\infty} \left[\frac{\Gamma(-d+j)}{\Gamma(-d)\Gamma(j+1)} \right] L^j$$

which makes use of the standard gamma function, $\Gamma(z)$, defined as

$$\Gamma(z) = \begin{cases} \int_0^{\infty} s^{z-1} e^{-s} ds, \text{ for } z > 0 \\ \infty \text{ , for } z = 0 \text{ .} \end{cases}$$

Using the formulation expressed in (2.) it can be~~en~~ seen that the model expressed in (1.) reduces to the standard stationary time series models of Box and Jenkins (1976) for integer values of d. In this case, the expression in (2.) has a finite number of non-zero terms indicating the relatively short memory of the process y_t. In the case where d takes on a value of one, the operator generates first differences: $(1 - L)^d y_t = [1 - (1)L^1 + 0]y_t = y_t - y_{t-1}$. In the case where d = 2 the operator generates second differences: $(1 - L)^d y_t = [1 - (2)L^1 + L^2 - 0]y_t = \Delta(y_t - y_{t-1})$.

In the case where d takes on fractional values, however, the expression for $(1 - L)^d$ has an infinite number of non-zero terms such that the current realization of y_t is a function of a long history of observations on y. This model exhibits a relatively slow decline of the weights on observations farther back in time. The model with fractional d is characterized by relatively long-term persistence and series which exhibit this type of behavior are called fractionally integrated time series. Formally, a series is fractionally integrated if after applying the fractional difference operator, $(1 - L)^d$, it follows a finite order ARMA(p,q) process. In such case, the series is said to be integrated of orde

Inte

d*l

stu

Hos

pro

(3.

or

(3.

Exp

aut

res

(4.

and

(5

and

whe

noj

aut

The

order d, I(d), and may be expressed as an Autoregressive-Fractionally Integrated Moving Average, ARFIMA(p,d,q) process where for $-\frac{1}{2} < d < \frac{1}{2}$ and $d \neq 0$ the series is sationary and mean reverting.

The theoretical characteristics of the ARFIMA process have been studied extensively by Granger (1980), Granger and Joyeux (1980), and Hosking (1981). In the simplest case where y_t is an ARFIMA(0,d,0) process, the model may be expressed as

$$(3.) \qquad (1 - L)^d y_t = u_t,$$

or alternatively,

$$(3.') \qquad y_t = (1 - L)^{-d} u_t.$$

Expressions for the one-sided representations that correspond to infinite autoregressive and infinite moving average processes are given, respectively, by

$$(4.) \qquad y_t = \sum_{j=1}^{\infty} \pi_j y_{t-j} + u_t$$

and

$$(5.) \qquad y_t = \sum_{j=0}^{\infty} \psi_j u_{t-j}$$

and are derived from the binomial expansions, (2.). In the special case where u_t is iid $(0, \sigma^2)$, y_t is said to be fractionally integrated white noise and (4.) and (5.) may be interpreted as being infinite autoregressive and infinite moving average representations, respectively. The coefficients on these representations for the fractionally integrated process may be expressed in terms of the gamma function as

$$\pi_j = \frac{\Gamma(\overset{d-j}{\cancel{j-d}})}{\Gamma(-d)\Gamma(j+1)} \quad \text{and} \quad \psi_j = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \quad \text{for } j \geq 1.$$

Granger and Joyeux (1980) have shown that as $j \to \infty$,

$$|\pi_j| \simeq A_1 j^{-d-1} \text{ and } \psi_j \simeq A_2 j^{d-1},$$

where the A_i $i=1,2$ are constants to be determined from initial boundary conditions and are expressed as $A_1 = [1/(d-1)!]$ and $A_2 = [1/(-d-1)!]$. Clearly, $|\pi_j|$ and ψ_j will decay to zero more slowly than the decay exhibited by the moving average coefficients of the standard stationary and invertible ARMA model. From this property it should be apparent that no finite-order ARMA model could adequately approximate a long-memory time series for large lags.

Hosking (1981) derives an expression for the autocovariance function of the fractionally integrated process, $\gamma_j = \text{cov}[X_t, X_{t-j}]$ for $t=0, \pm 1, \pm 2, \ldots$, and shows that as the lag j increases the autocovariance function behaves as $\gamma_j = O(j^{2d-1})$. Granger and Joyeux (1980) show that the fractionally integrated process is covariance stationary for $0 < d < \frac{1}{2}$ and has infinite variance for $-\frac{1}{2} < d < 0$. Additionally, an expression for the autocorrelation function of the fractionally integrated process is given in Granger and Joyeux (1980) and Hosking (1981) as

$$\rho_j = \frac{\gamma_j}{\gamma_0}, \qquad \rho_j = \frac{\Gamma(1-d)\Gamma(j+d)}{\Gamma(d)\Gamma(j+d-1)},$$

where $-\frac{1}{2} < d < \frac{1}{2}$. Using Sheppard's formula,[2] for large j the autocorrelation function of the fractionally integrated process may be expressed as

$$\rho_j \simeq A_3 j^{2d-1},$$

where A_3 is a constant expressed as $[-d!/(d-1)!]$. In contrast, the autocorrelation function of the standard stationary invertible ARMA process may be expressed for large j as

$$\rho_j \simeq A_4 \theta^j,$$

where $|\theta| < 1$, which tends to zero exponentially as $j \to \infty$. As with the impulse response weights expressed in (4.) and (5.), the rate of decline of the autocorrelations of the stationary ARMA process is much faster than that of the fractionally integrated process, further indicating the inability of the finite ARMA process to model the type of persistence present in long-memory data.

3. A SURVEY OF LONG-TERM PERSISTENCE AND BROWNIAN MOTION

The existence of highly persistent, long-memory data can be found not only in the field of hydrology in geophysical data (Hurst (1951, 1956), Mandelbrot and Wallis (1969a), and McLeod and Hipel (1978)) but also in the areas of image processing (Kashyap and Lapsa (1984)) and meteorology (Kashyap and Eom (1988)) as well. In searching for a physical

---

[2] Sheppard's formula indicates that for large lag, j, the expression $\Gamma(j+a)/\Gamma(j+b)$ is well approximated by j^{a-b}.

explanation of the cause of strong persistence in data, Hosking (1985) noted that the existence of long-term dependence in a series quite likely could arise in models which require solving stochastic differential equations. These models generally give rise to the properties of self-similarity and power-law covariance, which are characteristic of long-memory data. Models which utilize stochastic differential equations often arise in hydrological and geophysical applications, as noted by Hosking (1985), since these types of processes offer the potential to provide physically realistic specifications of many variables which are encountered in this field. It should not be surprising, then, that pioneering research in the area of strongly dependent time series originated in the field of hydrology with the work of Hurst (1951, 1956) who noted the long-term dependence in geophysical and hydrological time series in analyzing data on river discharge and reservoir storage capacity.[3]

In analyzing series which displayed long-term dependence Hurst developed the rescaled range statistic, R/S, which expresses the range of partial sums of deviations of a time series rescaled by the standard deviation of the series. Consider the time series $\{x_t\colon t=1,2,\ldots,n\}$, which Hurst assumed to be normal and independently distributed. The R/S statistic is calculated as

$$R/S = s^{-1}\{\max(U_1,\ldots,U_n) - \min(U_1,\ldots,U_n)\}$$

[3] An excellent survey on modelling long-term persistence in hydrological time series can be found in Lawrance and Kottegoda (1977).

$$\text{where } U_j = \sum_{t=1}^{j} (x_t - \bar{x}) \ , \qquad \bar{x} = n^{-1} \sum_{t=1}^{n} x_t$$

$$\text{and } s^2 = n^{-1} \sum_{t=1}^{n} (x_t - \bar{x})^2 .$$

The term $\max(U_1,\ldots,U_n)$ represents the maximum, over j, of the partial sums of the first j deviations of x_t from the sample mean, $\bar{x}$. This quantity will always be nonnegative since the deviation of x from its mean, summed over all observations, will be zero. Similarly, the term $\min(U_1,\ldots,U_n)$ is the minimum, over j, of the partial sums of the first j deviations of x_t from the sample mean, and this quantity will always be nonpositive. The range is defined as the difference between these two quantities. The variable s is the usual maximum likelihood standard deviation estimator and is used to normalize, or "scale", the range.

A notable development of Hurst's work was the finding of an apparent discrepancy between what was predicted by theory for the value of the R/S statistic and what was observed in empirical analysis for many hydrological and geophysical time series. In applied work Hurst observed that the value of R/S could be approximated by $(n/2)^k$, where n represents sample size and $0.6 < k < 0.8$. However, when the data were generated by a random process which assumed the dependence between distant observations to be negligible, the value of the rescaled range was asymptotically proportional to $n^{1/2}$. It was then hypothesized that Hurst's observed phenomenon should not hold for relatively large samples; but this hypothesis could not be confirmed empirically. Hurst investigated data on water flows of several rivers, including the Nile, and various other physical series including rainfall, temperatures, thickness of tree rings, thickness of stratified mud beds, and sunspot numbers. In each case the discrepancy between theoretical expectations and empirical evidence remained. This discrepancy came to be known as the "Hurst phenomenon" and its implication was that the geophysical records which displayed this phenomenon must be considered to posses an infinite span of statistical interdependence.

Subsequent research demonstrated that the presence of the Hurst phenomenon in certain data could in fact be explained by the degree of persistence in a series. That is to say, the Hurst phenomenon would be present in those stationary time series which displayed long-term dependence.[4] When applied to long-memory data, the rescaled range statistic would not behave as a function of $n^{1/2}$ as it would when applied to short-memory series. These findings began to indicate the importance of recognizing the existence of long-memory characteristics in data and indicated the need for the specialized type of analysis that these series require.

Since the work of Hurst in 1951, the rescaled range has been refined by Mandelbrot (1972, 1975) and others [Mandelbrot and Taqqu (1979), and Mandelbrot and Wallis (1968, 1969a, 1969b), for example], and Mandelbrot has suggested the use of the rescaled range statistic as a method for detecting long-range dependence in a series. Several researchers have shown the benefits of using this method over the use of, for example, the autocorrelation structure of a series or some measure of variance ratios. Mandelbrot (1972, 1975), for example, has shown the advantage of the rescaled range statistic by demonstrating the almost-sure convergence of

---

[4] See, for example, Mandelbrot and Wallis (1968), Klemes (1974), McLeod and Hipel (1978) and Hosking (1984).

this statistic for series with non-finite variance. In addition, Monte Carlo simulations by Mandelbrot and Wallis (1969a) have indicated that the rescaled range statistic can be used to identify long-term dependence in time series which are non-Gaussian and are compounded by some degree of skewness and kurtosis. The usefulness of the rescaled range statistic in detecting long-range dependence is discussed more fully in Lo (1991).

The analysis undertaken by Hurst was extended by several researchers in the field including Mandelbrot and Van Ness (1968) who suggested that strong persistence and stationarity were not necessarily mutually exclusive. Mandelbrot and Van Ness (1968) proposed the use of the fractional Brownian motion process to model strongly dependent time series since this process was able to characterize the long correlation structures associated with long-memory series. The attractive feature of fractional Brownian motion which made it useful in modelling long-term dependence was the infinite span of interdependence between the increments of the process. The derivation and statistical properties of the Brownian motion are well documented in the statistical and physical science literature (see Mandelbrot and Van Ness (1968), Hosking (1981), Mandelbrot (1982), and Jonas (1983), for more information and further references) and only a brief summary is given here.

The formal notion of Brownian motion was first brought to the attention of the scientific community in the early 1800s to describe the diffusion processes investigated by botanist Robert Brown. The Brownian motion, represented as $B(t)$, is a continuous-time stochastic process. The increments of the Brownian motion, represented as $B(t+u) - B(t)$, are Gaussian distributed with zero mean and variance u. A generalization of the Brownian motion process is fractional Brownian motion, represented as

$B_H(t)$, which is a continuous time, self-similar process. The characteristic of self-similarity is defined for a continuous-time stochastic process, $\{x_t\}$, which satisfies the condition that for some positive constants α and h, x_t and $\alpha^{-h}x_{\alpha t}$ have the same finite joint distributions. That is to say, the distribution of a self-similar process will be invariant to the scale under which the process is observed; the statistical behavior of an observed phenomenon characterized by self-similarity should have the same probabilistic structure whether the phenomenon is observed daily, weekly, monthly, or otherwise.

Fractional Brownian motion, which is the (½ - H)th fractional derivative of Brownian motion, has stationary Gaussian increments, and these increments are known as "fractional Gaussian noise". The increments of the fractional Brownian motion have hyperbolically decaying correlation functions which enable the process to account for the long-term behavior of strongly dependent time series. This is due to the relatively slowly varying span of interdependence between increments of fractional Brownian motion which vary according to power law. The autocovariance function of fractional Gaussian noise, $\gamma_k = E[x_t, x_{t-k}]$, is given as

$$\gamma_k = C\{(|k| + 1)^{2H} - 2|k|^{2H} + (|k| - 1)^{2H}\},$$

where C is a positive constant and the parameter H satisfies $0 < H < 1$.

The fractional Brownian motion process may be expressed, as

$$B_H(t) = \int_{-\infty}^{t} (t-s)^{H-\frac{1}{2}}\, dB(s), \quad \text{where } -\infty < t < \infty. \tag{6.}$$

The parameter H incorporates the long-memory element of the model, and is known as the Hurst coefficient. To ensure stationarity of the process the parameter H must satisfy the condition $0 < H < 1$, and fractional Brownian motion can be divided into three different families, essentially, depending on the range of values of H. For H in the range of $\frac{1}{2} < H < 1$, the fractional Brownian motion process will be meaningful for empirical applications in hydrology; the case where $0 \leq H \leq \frac{1}{2}$ will not be useful in this setting. When $H = \frac{1}{2}$, the fractional Brownian motion simply reduces to the Brownian motion process, implying that observations separated by a relatively large time span are statistically independent. Since the expression in (6.) above is divergent, it is only feasible to look at stationary increments of $B_H(t)$, the fractional Gaussian noise process, which can be expressed as

$$B_H(t) - B_H(0) = \int_{-\infty}^{t} (t-s)^{H-\frac{1}{2}}\, dB(s) - \int_{-\infty}^{0} (-s)^{H-\frac{1}{2}}\, dB(s). \tag{7.}$$

Estimation of the degree of persistence of the series, then, involves estimation of the parameter H, the Hurst coefficient, from the specification given in (7.). Estimation of the Hurst coefficient has been explored in Hurst (1951, 1956) and in Mandelbrot and Wallis (1969a, 1969b) who suggested constructing "pox diagrams" as a method of estimating the parameter H. The pox diagram plots various values of the R/S statistic of a series along with their average. The scatter of points usually produces a short, convex section and a longer, linear section which will have an average slope equal to H. Thus, the value of H is approximated from the shape of the curve in the pox diagram. Further research utilizing the fractional Brownian motion to model strongly dependent series and to estimate the degree of persistence of a series includes Mandelbrot (1972, 1975), Mandelbrot and Taqqu (1979), and Mandelbrot and Wallis (1968), among others.

The use of fractional Brownian motion to model strongly dependent time series involves some disadvantages, however, and hence may not be ideal in modelling all series exhibiting long-term persistence. For example, Mandelbrot and Wallis (1969b) have found great variability in the estimated values of H when the R/S pox diagram technique is used, and have even noted that in certain cases the researcher must be content to wager an intelligent, though imprecise, guess as to the value of H. Mandelbrot and Wallis (1969b) note that estimation of H through the use of the pox diagram becomes even more difficult in the presence of strong periodic elements in a series. In addition, Wallis and Matalas (1970) have proven that estimates of the Hurst coefficient have relatively large biases and large sampling variances, and have indicated that not much is known about the large sample efficiency of the estimator. An additional shortcoming of the model is its limited flexibility in modelling series which embody other characteristics in addition to the long-memory element. Such series would require multiple parameters in the model yet fractional Brownian motion offers only one parameter, H, with which to describe the behavior of a series. Consequently, this process limits the range of correlation structures which may be represented by fractional Brownian motion. In light of the shortcomings of fractional Brownian motion, an alternative class of models has been proposed which generalizes the ARIMA(p,d,q) models of Box and Jenkins (1976) to allow for non-integer values of d. This class of model, known as the fractionally differenced process, may allow for more flexibility in modelling series which contain long-memory. The next section provides a detailed survey of the fractionally differenced process.

4. A SURVEY OF LONG-TERM PERSISTENCE AND FRACTIONAL DIFFERENCING

It is often the case in macroeconomics and financial economics that time series analysis of certain variables involves the problem of extreme persistence in the mean and variance of these variables. The existence of such long-term dependence in economic data was first noted by Granger (1980) who suggested that the aggregation of certain dynamic economic relationships produced a class of model that exhibited long-memory characteristics.

For example, consider the series x_{1t} and x_{2t} which are generated by the process

$$(8.) \qquad x_{jt} = \alpha_j x_{j,t-1} + \epsilon_{jt}\ , \ j = 1,\ 2$$

and ϵ_{1t}, ϵ_{2t} are independent white noise processes with zero mean. As noted in Granger (1980) the sum of the processes, $\bar{x}_t = x_{1t} + x_{2t}$, will obey an ARMA(2,1) process where the autoregressive part of the model is given by $(1-\alpha_1 L)(1-\alpha_2 L)$. Consider such an aggregation for N independent series each following an AR(1) process as expressed in (8.) and each with differing values for the autoregressive parameter, α. The aggregation of these processes will follow an ARMA(N,N-1) process, provided that no cancellation of roots occurs between autoregressive and moving average parameters in the model. In assuming a specific distribution from which the autoregressive parameters are drawn,[5] Granger (1980) derives an approximation of the k th order autocovariance function of the aggregated process. This approximation is derived from the standard Fourier series expansion of the spectrum of the process and is given as $\bar{\mu}_k = A_5 k^{(1-q)}$ where A_5 is some constant. Recall from section 2 that the k th order autocovariance function of the fractionally integrated process is of order $k^{(2d-1)}$. The parameter of integration in the fractionally differenced model, d, corresponds to (1-q/2) in this case such that the aggregation of the N dynamic processes, $\bar{x}_t$, is said to be integrated of order (1-q/2). The model exhibiting this form of autocovariance function is characteristic of the strongly persistent, long-memory, fractionally integrated process, as discussed in section 2.

Thus for N very large the aggregation of N individual dynamic processes produces series which display those characteristics typical of long-term dependence as represented by the fractionally integrated model. In economics, such models are very likely to arise since many economic variables are aggregates of a large number of smaller series; relevant examples include the aggregate price level or the inflation rate, aggregate consumption data, national income data, etc.

As noted by Künsch (1986), the properties of such long-memory data cannot be explained asymptotically by stationary models characterized by finite variances and weak dependence. The alternative proposed by Granger and Joyeux (1980) and Hosking (1981) was derived by fractionally

[5] For mathematical simplicity Granger assumes a beta distribution for α, considered on the range (0,1). The particular choice of the distribution assumed for α is not crucial for deriving the result of highly persistent series through aggregation of dynamic processes. Granger investigates further generalizations which consider alternative distributions that produce the same result.

differencing the random walk process. The fractionally differenced model offers an alternative to the fractional Gaussian noise process of Mandelbrot and Van Ness (1968) and Mandelbrot and Wallis (1969a), although the two models are not completely without similarities. The fractionally differenced process has a covariance structure similar to that of fractional Gaussian noise and is also asymptotically self-similar. In addition, the parameter of fractional differencing in the model of Granger and Joyeux (1980) and Hosking (1981) is related to the Hurst coefficient as $d = (H - ½)$, where H is the parameter described in (6.) and (7.) (see Granger and Joyeux (1980), Hosking (1981), and Geweke and Porter-Hudak (1983), for a more detailed analysis).

Recently, much attention has been given to estimation of the degree of fractional differencing of series that display forms of long-term persistence. The earliest contributions in this area were made by Janacek (1982), who proposed a technique to estimate the degree of fractional differencing of a series based on the log of the power spectrum of the series, and Geweke and Porter-Hudak (1983) who proposed a semi-parametric, two-step estimation procedure which also utilizes the spectrum of a series. These procedures are based in the frequency domain since long-memory time series can be equally well represented in either the time domain, where series can be expressed in terms of stochastic difference equations, or in the frequency domain, where series can be expressed in terms of the relative importance of various cycles that occur within the series. In the context of the frequency domain a long-memory process as expressed in (1.), which exhibits spectral density function, $f_y(\omega) = [4\sin^2(\omega/2)]^{-d} f_u(\omega)$, will be characterized by an infinitely increasing spectral density function for frequency values, ω, near zero.[6]

The estimation technique proposed by Janacek (1982) was designed to determine the order of fractional integration of a series and is grounded in the idea that the low-frequency end of the log-spectrum of a series gives some information about the degree of persistence of the series. Janacek (1982) considered a univariate series $\{y_t: t = 1, 2, \ldots, N\}$ which when differenced d time gives rise to a stationary process, u_t, with rational spectrum $f_u(\omega)$ where ω represents frequency. Applying the standard procedures of linear filters, an expression for the log of the spectrum of x is given as

$$\log f_x(\omega) = -d \log[2(1-\cos \omega)] = \log f_u(\omega).$$

Introducing a "weighting" function as $W(\omega) = -\frac{1}{2} \log[2(1-\cos \omega)]$ and using the standard results for Fourier series, an integral expression for the weighted log of the spectrum can be given as

$$1/\pi \int_0^\pi W(\theta) \log f_x(\theta)\, d\theta = d \sum_{K=1}^{\infty} 1/K^2 + \sum_{K=1}^{\infty} a_K/(2K)$$

where $f_u(\omega) = a_0 + a_1\cos \omega + a_2\cos 2\omega + \ldots$, such that the a_K, K=1, 2, 3, . . . , are the Fourier coefficients of the spectrum of u_t. This allows for estimation of d without the need for model specification. The minimum-mean square error of prediction, σ^2, is given as

---

[6] This particular characteristic corresponds to hyperbolically decaying correlation functions when the series is expressed in the time domain.

and

b_2^2

c_1^2 /

shor

all

seri

Janad

from

where

persi

large

chose

large

estim

to in

that

withi

$$\log \sigma^2 = \pi^{-1} \int_0^{\pi} \log 2\pi\, f(\omega)\, d\omega$$

and the step mean-square error of prediction, σ_k^2 is given as $\sigma^2(1 + b_1^2 + b_2^2 + \ldots + b_K^2)$. The b_j's can be expressed as $(b_1 = c_1)$, $(b_2 = c_2 + c_1^2/2!)$, . . . , where $2c_K = a_K$. The c_K will decay exponentially for a short-memory, stationary process where the c_K are assumed to be zero for all values of K greater than some finite number, M. For a long-memory series, however, the c_K will decay at the much slower hyperbolic rate. Janacek (1982) proposed estimation of d based on an estimate of the c_K from the following formulation

$$d_M = (S - \sum_M \hat{c}_K / K) / \sum_M K^{-2}$$

where $$S = 1/\pi \int_0^{\pi} W(\omega) \log \hat{f}(\omega)\, d\omega .$$

Clearly, in this specification, the estimate of the degree of persistence will depend on the value of M. Janacek proposed that for M large enough the estimate of d will be unbiased, and suggested that M be chosen as being equal to two standard deviations of d. Choosing M too large, however, does necessitate a trade off in increased variance of the estimated parameter, and Janacek (1982) presented small scale simulations to investigate this trade off. The results of the simulations indicate that for M chosen as suggested above, the variance of the estimate is within a 95% confidence interval of theoretical predictions.

Similarly, the motivation for the estimation technique proposed by

Geweke and Porter-Hudak (1983) is like that proposed by Janacek (1982) in that it utilizes the low-frequency end of the spectrum of a series. The procedure proposed by Geweke and Porter-Hudak is based on the observation that the spectral density function of a long-memory time series when expressed in levels is unbounded at $\omega = 0$ but disappears at $\omega = 0$ when the series is first differenced. Consequently, Geweke and Porter-Hudak (1983) use the low-frequency component of the spectral density function to estimate the degree of persistence of a series.

Consider the simple case of the ARFIMA(0,d,0) process as expressed in (3.) where u_t is a white noise process. Again, applying the property of linear filters, the spectral density function of y_t, $f_y(\omega)$, can be expressed as

$$(9.) \qquad f_y(\omega) = |1 - e^{-i\omega}|^{-2d} f_u(\omega).$$

An equivalent expression may be given as

$$(9.') \qquad f_y(\omega) = [4\sin^2(\omega/2)]^{-d} f_u(\omega)$$

where $f_u(\omega)$ represents the spectral density function of u. Taking logarithms of both sides of equation (9.') and evaluating at the harmonic ordinates $\omega_j = 2\pi j/T$ gives

$$(10.) \qquad \log[f_y(\omega_j)] = \log[f_u(0)] - d\log[4\sin^2(\omega_j/2)] + \log[f_y(\omega_j)/f_u(0)].$$

The motivation for this estimation procedure is based on the similarity of (10.) to an ordinary least squares regression where the last term on the right hand side of (10.) represents the regression error and the slope parameter, -d, can be estimated with the usual least squares techniques. The estimation procedure of Geweke and Porter-Hudak (1983), hereafter referred to as GPH, requires two steps. The first step involves obtaining an estimate of d from equation (10.) using OLS. This regression is performed over the range of low-frequency ordinates of the spectrum, ω_1, ω_2, .. ω_m, and involves replacing $f_y(\omega_j)$ with the periodogram of the series. In this formulation it is critical to estimate the OLS regression using the appropriate number of ordinates, m, since the estimation technique is based on performing the regression over only the low-frequency end of the spectrum. Use of too small a regression sample, that is, inclusion of too few ordinates, can mean that relevant long-memory information contained in the low-frequency end of the spectrum is not being included in estimation. Conversely, use of too many ordinates can lead to erroneous inclusion of medium- and high-frequency ordinates in estimation, which will cause the estimate of d to be "contaminated". Geweke and Porter-Hudak (1983) have suggested that the optimal number of low-frequency ordinates to use in the OLS regression should be a function of the sample size of the series under investigation. They suggest choosing the number of spectral ordinates as being T^{α}, where α would typically take on values such as .50, .60, and .70. Alternatively, Sowell (1990) has suggested that the optimal number of spectral ordinates to use in the OLS regression should not depend on sample size but, rather, should be a function of the annual periodicity of the data and the long-run frequency of the series. Consequently, Sowell (1990) has suggested choosing the optimal number of ordinates, m, as

$$m = \frac{\text{number of years of the data}}{\text{lowest long-run frequency of the series}} .$$

Once a consistent estimate of the parameter d, the degree of fractional integration, is obtained by the GPH procedure outlined above, this value is then used in the second step of the estimation procedure to transform, or filter, the observed series for the appropriate degree of fractional differencing. Once the series has been properly transformed the remaining model parameters may be estimated by standard procedures of time series analysis performed on this transformed series. This second step of estimation will provide estimates for the autoregressive and moving average components of the series. The consistency of the GPH estimator of d in the presence of weak dependency in u_t is proven in Geweke and Porter-Hudak (1983) for $-\frac{1}{2} < d < 0$, and is conjectured for $0 < d < \frac{1}{2}$. The estimator is shown to be asymptotically normal, but is not $\sqrt{T}$ consistent.

To date, there appears to be limited applied work utilizing the estimation procedure outlined by Janacek (1982); however, the GPH estimator has been widely applied in many areas since its introduction (Diebold and Rudebusch (1989), Choi and Wohar (1990), Baillie and Pecchenino (1991), Cheung (1991), Tieslau (1991), and Agiakloglou, Newbold, and Wohar (1992), for example). Some difficulty has been encountered, however, in applying the GPH estimator in practice. Agiakloglou, Newbold and Wohar (1992), for example, have demonstrated with simulation results that the GPH estimator can have substantial bias when y_t is generated by an autoregressive process with a substantially large positive AR parameter, or when y_t is generated by a moving average process with a substantially large positive MA parameter. Agiakloglou, Newbold and Wohar (1992) discuss the reasons for these results noting that when y_t exhibits a large positive value for its AR parameter, the log spectrum is downward sloping in the low-frequency range, rather than constant. As such, there will be significant bias in estimating the slope parameter. Agiakloglou, Newbold and Wohar (1992) present simulation evidence indicating that the bias of the GPH estimate decreases slowly as sample size increases, but this comes at the expense of a trade-off of increasing the standard error of the estimate.

Similarly, the presence of large bias in the GPH estimate when y_t exhibits a large positive moving average coefficient is due to fact that for this type of process the spectral density of y_t is close to zero in the low-frequency range of the spectrum. Agiakloglou, Newbold and Wohar (1992) show that for series exhibiting large positive MA parameters the bias of the GPH estimate for the model also decreases slowly as sample size increases.

In addition, Hosking (1985) noted the limited applicability of the fractionally differenced ARMA process in applied work in hydrology due to the linearity of this model. Hosking noted that many typical hydrological time series, such as daily streamflow data, often embody non-linear characteristics which are unlikely to be well explained by the type of model proposed by Geweke and Porter-Hudak (1983). As a result, the GPH estimation procedure appears to be most appropriate in modelling purely fractionally integrated white noise, leaving little flexibility for practical applications.

Since the seminal work of Geweke and Porter-Hudak (1983) on estimation of fractionally integrated series, several alternative estimation procedures have been proposed for the ARFIMA process described in equation (3.). One such procedure involves maximum likelihood estimation (MLE) of the parameters of the model given in (3.), and three alternative approaches which utilize maximum likelihood estimation techniques have been proposed. The first technique proposed involves approximate MLE based in the time domain, the second involves approximate MLE based in the frequency domain, and the third estimation technique involves exact MLE based in the time domain. Each of these are further discussed below.

The first alternative estimation procedure utilizing maximum likelihood estimation techniques was proposed by Hosking (1984) who, in dealing with many non-linear hydrological and geophysical data series, suggested that methods based on maximizing the likelihood function of a series were the most likely methods for producing efficient parameter estimates for fractionally integrated time series. The method proposed by Hosking (1984) utilizes approximate MLE techniques in the time domain where the log likelihood function of a series, y_t, is expressed as

$$\log f(\lambda,\mu;y) = -\tfrac{1}{2} \log |V(\lambda)| - \tfrac{1}{2}(y-\mu 1)' \{V(\lambda)\}^{-1} (y-\mu 1)$$

where λ represents the vector of model parameters, μ represents the mean of the process, $V(\lambda)$ is the covariance matrix of y, which is assumed to be independent of the mean, and 1 is a vector of ones. Hosking (1984) assumed, for convenience, that the likelihood function followed a normal marginal distribution. Direct maximization of the likelihood above has been proposed in applied work by McLeod and Hipel (1978) in estimating the parameters of the fractional Gaussian noise process of (7.). However, Hosking (1984) noted the computational expense of this method since it requires the inversion of a (T × T) matrix of covariances at each iteration of the procedure which necessitates a number of arithmetic operations of order T^3. The cost of repeated iterations of this process in maximizing the likelihood function, $\mathcal{L}$, can be excessive.

As an alternative, Hosking (1984) suggested replacing the likelihood function with an approximation, $\hat{\mathcal{L}}$, which would be more easily evaluated and then maximizing this approximation. This approximation considers the fractionally integrated parameters of the model separately from the ARMA parameters. Estimation of the model parameters, then, simply involved application of the standard algorithm of Ansley (1979) to the approximation of the likelihood function.

The second alternative estimation procedure, proposed by Fox and Taqqu (1986), also utilizes approximate maximum likelihood estimation techniques but is based in the frequency domain rather than in the time domain. Within this framework Fox and Taqqu (1986) assume a stationary Gaussian series and construct an asymptotic approximation to the likelihood function of an ARFIMA process. Their work follows the approach suggested by Whittle (1951) which involves maximizing

$$\frac{1}{\sigma} \exp \left[- \frac{Z' A_T(\theta) Z}{2T\sigma^2} \right]$$

where $Z = (X_1 - \bar{X}_T, \ldots, X - \bar{X}_T)'$, and $\bar{X}_T$ represents the sample average of the process X. The matrix $A_T(\theta)$ is a (T x T) matrix of covariances which has j,k^{th} element $a_{j-k}(\theta)$ as given by

$$a_j(\theta) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} e^{ijx} [f(x,\theta)]^{-1} dx.$$

The maximum likelihood estimates, $\bar{\theta}_T$ and $\bar{\sigma}_T$, are defined as those values of θ and σ that satisfy the above maximization; maximization of the above expression is equivalent to choosing $\bar{\theta}_T$ to minimize $\sigma_T^2(\theta) = [Z'A_T(\theta)Z]/T$ and then setting $\bar{\sigma}_T^2$ equal to $\sigma_T^2(\bar{\theta}_T)$. Fox and Taqqu (1986) derive consistent and asymptotically normal estimates of the model parameters for the case when d is in the range $0 < d < ½$. Maximization of the likelihood function is based in the frequency domain and the procedure is invariant to the true, but unknown, mean of the process.

The third estimation procedure, proposed by Sowell (1992), utilizes an exact maximum likelihood estimation technique which is based in the time domain. This estimation procedure considers the ARFIMA(p,d,q) process as given in (1.) with $u_t \sim NID\ (0,\ \sigma^2)$ and $d < ½$. The probability density function is given by Sowell to be

$$f(y_t,\Sigma) = (2\pi)^{-T/2}\ |\Sigma|^{-1/2}\ \exp\{-½ y_t \Sigma^{-1} y_t\}$$

where Σ is the covariance matrix of the ARFIMA process y_t, and Σ is of the standard block Toeplitz form. The likelihood function of the process is derived in the usual manner and maximization is achieved with standard computer algorithms and use of the Choleski decomposition in evaluating the inverse of the covariance matrix.

The inversion of the $(T \times T)$ matrix of autocovariances which is required by this time domain exact MLE technique is the main disadvantage of this procedure. Inversion of this covariance matrix is required at each iteration of the estimation process and this becomes particularly cumbersome for fractionally integrated processes since even for fractionally integrated white noise each autocovariance is a ratio of gamma functions. This makes arithmetic operations quite difficult. Furthermore, for the more general ARFIMA(p,d,q) process, a typical autocovariance is a nonlinear function of four hypergeometric functions necessitating even greater computational expense. Clearly, maximum likelihood estimation is difficult even for an ARFIMA process with NID (0, σ^2) disturbances.

The relative efficiency of the two estimation procedures of Sowell (1992) and Fox and Taqqu (1986) has been analyzed in a recent study by Cheung and Diebold (1991). The exact MLE time-domain procedure proposed by Sowell (1992) is asymptotically equivalent to the approximate MLE frequency-domain procedure proposed by Fox and Taqqu (1986), although the properties from estimation in finite samples will not be equivalent. Cheung and Diebold (1991) perform a detailed simulation study which compares both large and small sample properties of the two estimation techniques. They analyzed the fractionally integrated white noise process with non-zero mean, μ, which may be represented as:

$$(1-L)^d(y_t-\mu) = \epsilon_t$$

where $\epsilon_t \sim \text{NID}(0, \sigma^2)$.

The results of this analysis indicate that for finite sample sizes when the mean of the process is known, an unambiguous ranking of the two estimation procedures is evident. That is, the mean-squared error (MSE) of the parameter estimate of the frequency-domain approximate MLE of Fox and Taqqu (1986) is greater than that of the time-domain exact MLE of Sowell (1992) for the case of known mean. By this criterion the exact MLE approach of Sowell (1992) is superior to the approximate MLE approach of Fox and Taqqu (1986) when the true mean of the process is known.

However, in the case of unknown mean in finite sample sizes, Cheung and Diebold (1991) show that the relative efficiency of the approximate MLE increases significantly. This is particularly relevant since most practical applications of these estimation procedures will arise in cases where the true mean of the process is unknown and consequently must be estimated. Such estimation is based on de-meaned data and Cheung and Diebold (1991) have shown that precise estimation of the mean of a long-memory process is often difficult. The frequency-domain estimation procedure of Fox and Taqqu (1986) will be invariant to the true mean of the process; the time-domain estimation procedure of Sowell (1992) will not be. By this criterion, the estimation procedure of Fox and Taqqu (1986) may be preferred to that of Sowell (1992).

In addition Cheung and Diebold (1991) indicate that for the case where the mean of the process must be estimated, the efficiency of the exact MLE time-domain approach of Sowell (1992) decreases as the value of d increases. That is, the rate of convergence of the sample mean of the process is a function of the parameter d in that as the value of d increases the rate of convergence decreases, which further decreases the performance of the exact MLE time-domain approach. This is due to the fact that while the maximum likelihood estimates of the ARMA parameters expressed in (1.) are $\sqrt{T}$ consistent, the rate of convergence of the maximum likelihood estimate of the population mean is $T^{+\frac{1}{2}-d}$ (for further discussion, see Li and McLeod (1986)).

The maximum likelihood estimation procedures discussed above, whether approximate or exact MLE, are each predicated on the assumption of NID innovations. That is to say, the approximate MLE technique of Fox and Taqqu (1986), for example, cannot be applied to data which exhibit forms of time dependent heteroskedasticity, which may well be encountered in practical applications. To date, no attempt has been made to address the problem of obtaining maximum likelihood estimates for a process which involves more complicated data generating processes where the presence of fractional integration may be compounded with non-normality and non-homoskedastic error. This can prove to be a significant factor in assessing the usefulness of an estimation procedure since such data series are likely to arise in practice. The next section will present one such estimation procedure which utilizes the estimation technique of maximum likelihood to estimate the model parameters of a series but yet does not restrict the distribution of the series to satisfy the assumption of normal and independently distributed error.

5. THE THEORETICAL PROPERTIES OF THE ARFIMA-GARCH PROCESS

The model developed in this section merges the ARFIMA(p,d,q) process of Granger and Joyeux (1980) and Hosking (1981) with the GARCH(P,Q) process of Engle (1982) and Bollerslev (1986) to allow for simultaneous modelling of fractionally integrated behavior and time dependent conditional heteroskedasticity. The property of time varying conditional variance has been investigated by Engle (1982, 1983) who proposed the Autoregressive Conditionally Heteroskedastic, ARCH, process as a suitable model for series exhibiting this characteristic. In this model, the conditional variance is assumed to be a linear function of past squared errors; the conditional variance is allowed to change over time as a function of these past squared errors, leaving the unconditional variance constant. The ARCH model, then, is able to characterize series which tend to exhibit periods of extreme values followed by other periods of extreme values, whether of the same sign or not. Engle (1982) proposed this model for inflation realizing that the uncertainty of inflation tended to change over time.

The Autoregressive Conditionally Heteroskedastic model of Engle (1982) was later extended by Bollerslev (1986) who generalized the model to allow for a much more flexible lag structure. The generalization proposed by Bollerslev (1986) allows for an additional parameter in the conditional variance equation; under this specification the conditional variance is assumed to be influenced by the squared residuals of the series as well as lagged values of the conditional variance. This model, known as the Generalized Autoregressive Conditionally Heteroskedastic GARCH(P,Q) model, is used to model series which exhibit non-constant conditional variance over time.

The general model proposed in this analysis will be referred to as the ARFIMA(p,d,q)-GARCH(P,Q) model and can be represented with equations (11.) through (14.) below:

$$(11.) \qquad (1-L)^d y_t = b' x_{1t} + \delta\sigma_t + u_t$$

$$(12.) \qquad \phi(L) u_t = \theta(L)\epsilon_t$$

$$(13.) \qquad \epsilon_t | \Omega_{t-1} \sim D(0,\ \sigma_t^2)$$

$$(14.) \qquad \beta(L)\sigma_t^2 = \omega + \alpha(L)\epsilon_t^2 + \gamma' x_{2t}.$$

In this specification, $d \in (-\frac{1}{2}, \frac{1}{2})$ such that y_t is fractionally integrated of order d, and x_{1t} and x_{2t} are vectors of predetermined variables. The polynomials $\phi(L)$, $\theta(L)$, $\beta(L)$, and $\alpha(L)$ are represented as $\phi(L) = 1-\phi_1 L - \ldots-\phi_p L^p$, $\theta(L) = 1-\theta_1 L - \ldots-\theta_q L^q$, $\beta(L) = 1-\beta_1 L - \ldots-\beta_P L^P$, $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \ldots + \alpha_Q L^Q$, and all the roots of $\phi(L)$, $\theta(L)$, $\beta(L)$ and $\alpha(L)$ are assumed to lie outside the unit circle. The innovations of the model, ϵ_t, are assumed to follow a conditional density D, which may be assumed to be either Normal or Student t. Additionally the time dependent heteroskedasticity of σ_t^2 follows the Generalized Autoregressive Conditionally Heteroskedastic, or GARCH(P,Q), model of Engle (1982) and Bollerslev (1986). In the special case where $\delta = 0$ and $b = 0$, equations (11.) and (12.) describe the ARFIMA process of Granger and Joyeux (1980).

One useful application afforded by the ARFIMA-GARCH model is that the specification will allow for the possibility of testing for the presence of feedback between the standard deviation of y and lagged values of y. In this framework, the parameter δ is used to test for this effect such that when $\delta \neq 0$ volatility is allowed to influence the mean of y_t. The model specification will also allow for a test of whether lagged predetermined variables influence the conditional variance of y_t. This is achieved by allowing for the predetermined variable to enter into the conditional variance equation (14.) by being included in x_{2t}. A positive and significant value for the parameter γ, then, would indicate the existence of an effect of lagged predetermined variables on y_t. Such tests will prove valuable in empirical work in subsequent analysis.

The likelihood function of the full model represented by equations

(11.) through (14.) is given by $f(\lambda,d;y)$ where λ is the parameter vector

$$\lambda' = (b'\delta\ \phi'\theta'\omega\ \alpha'\beta').$$

Bollerslev (1987) presents an extension to the ARCH model of Engle (1982) in which the conditional density of the process y_t is assumed to be Student t with υ degrees of freedom. This specification can by particularly useful for analyzing series which exhibit excessive tail distribution, perhaps due to outliers in the series, since it allows for specific modelling of unconditional excess kurtosis in an observed series. Following Bollerslev (1987), the log likelihood function for the Student t distribution with T observations can be expressed as

$$(15.)\qquad \log f(\lambda,d;y) = T[\log \Gamma(\tfrac{\upsilon+1}{2}) - \log \Gamma(\tfrac{\upsilon}{2}) - \tfrac{1}{2}\log(\upsilon-2)] -$$

$$\frac{1}{2}\sum_{t=1}^{T}[\log \sigma_t^2 + (\nu+1)\log(1+\epsilon_t^2\,\sigma_t^{-2}\,(\nu-2)^{-1}].$$

When ν^{-1} approaches zero, the t distribution approaches a normal distribution; but for $\nu^{-1} > 0$ the t distribution will be quite different from the normal distribution, exhibiting substantial leptokurtosis or fat tail distribution. Consequently, the likelihood function expressed in (15.) will be appropriate for those series for which the inverse of the degrees of freedom parameter, ν^{-1}, takes on values greater than zero.

Estimation of the model represented by the system of equations (11.) through (14.) may be achieved through standard maximum likelihood estimation (MLE) techniques. The process of performing exact MLE on the entire system, however, can be rather difficult especially given the allowance for general conditional densities, D, and the presence of time dependent conditional variance as specified in (14.). Attention may be restricted to D being either Normal or Student t to circumvent this difficulty, and approximate maximum likelihood estimates of all model parameters may be obtained via two separate methods which are described below.

The first method which may be used to obtain approximate maximum likelihood estimates of λ is referred to as Method I and involves direct maximization of equation (15.). This may be achieved numerically through standard maximization of the likelihood function via use of any of the standard computing algorithms such as the Berndt *et. al.* (1974) algorithm. This method provides approximate ML estimates of all model parameters simultaneously and will provide appropriate asymptotic standard errors for the parameter vector λ. Unfortunately, this method entails the problem of setting starting values for the process since the presence of d, the fractional differencing parameter, has the effect of making initialization conditions persist for a longer period than would be the case with the standard ARMA process.

The second method which may be used to obtain approximate maximum likelihood estimates of λ is referred to as Method II and considers the conditional likelihood with respect to the fractionally integrated parameter, d. The likelihood function for this method is given by $\pounds^{*}(\lambda;y)$ and is related to the full likelihood as:

$$\pounds(\lambda,d;y)=J(d)\pounds^{*}(\lambda;y),$$

wher

$(u_1$

aver

y_t

.10

ther

func

adj

der

to s

whic

wher

uni

Onc

est

fir

Thi

where J(d) is the Jacobian of the transformation from $(y_1, y_2, \ldots y_T)$ to $(u_1, u_2, \ldots u_T)$. The transformation involves the use of the moving average filter given by equation (5.). This procedure filters the series y_t for a range of values of d which can be taken at discrete intervals of .10, for example, resulting in the filtered series u. The filtered series then may be used in the estimation technique of maximizing the likelihood function. In making this transformation from y to u, however, an adjustment of the covariance matrix is required, necessitating the derivation of the Jacobian of the transformation. It is relatively easy to show that J(d) is the determinant of a (T x T) lower triangular matrix which has ones along the main diagonal and an (i,j)th element of $\psi_{|i-j|}$, where $i>j$. Consequently, the Jacobian of the transformation, J(d), is unity so that

$$\pounds^{*}(\lambda;y) = \pounds(\lambda,d;y).$$

Once this trivial adjustment is realized, full maximum likelihood estimates of the parameter vector λ may be obtained from the standard first-order conditions which are represented as

$$\frac{\partial \pounds(\lambda,d;y)}{\partial \lambda} = \frac{\partial \pounds^{*}(\lambda;y)}{\partial \lambda} = 0 \ .$$

This will give the maximum likelihood estimates, $\hat{\lambda}(d,y)$.

The concentrated likelihood with respect to λ is then defined as

$$\pounds^{c}(d;y) = \pounds[\hat{\lambda}(d,y),d;y].$$

The model can be estimated numerically by maximizing the concentrated likelihood $\pounds^{c}$ with the use of any standard computer algorithm such as Berndt *et. al.* (1974). In this way, Method II will generate approximate maximum likelihood estimates and will also give the correct value of the maximized log likelihood. The standard errors of the parameter estimates, however, will be conditional on d.

6. SUMMARY AND CONCLUSION

The characteristic of long-term, but not permanent, persistence in a series distinguishes long-memory time series from both non-stationary time series which contain unit roots and are not mean reverting, and stationary time series which exhibit very little or no persistence and require no differencing to achieve stationarity. Clearly, long-memory time series are not well represented as having unit roots nor as requiring no differencing to achieve stationarity. This chapter outlines the population and sample characteristics of long-memory time series and distinguishes these series from those which exhibit short-memory.

Two methods of obtaining approximate maximum likelihood estimates are proposed for ARFIMA processes which are compounded with GARCH innovations and conditional Student t densities. These types of models should prove to be most useful in empirical application since many time series encountered in practice do exhibit forms of long-term persistence and non-normality. Estimation of the model parameters of those series exhibiting long memory is rather difficult especially for those series which also exhibit other non-linear characteristics; the procedures proposed here, which employ conditional likelihood functions, offer a good approach to estimating such models.

Agiak

Ansley

Bailli

Bailli

Berndt

Boller

Boller

Box, G

Cheung

Cheung

Choi,

Diebol

Diebol
A

Diebol
T
I

REFERENCES

Agiakloglou, C., P. Newbold and M. Wohar (1992), "Bias in an Estimator of the Fractional Difference Parameter", Journal of Time Series Analysis, forthcoming.

Ansley, C.F. (1979), "An Algorithm for the Exact Likelihood of an Autoregressive Moving-Average Process", Biometrika, 66, 59-65.

Baillie, R.T. (1989), "Commodity Prices and Aggregate Inflation: Would a Commodity Price Rule be Worthwhile?", Carnegie Rochester Conference Series on Public Policy, 31, 185-240.

Baillie, R.T, and R.A. Pecchenino (1991), "The Search for Equilibrium Relationships in International Finance: The Case of the Monetary Model", Journal of International Money and Finance, 10, 582-93.

Berndt, E.K., Hall, B.H., Hall, R.E. and J.A. Hausman (1974), "Estimation and Inference in Nonlinear Structural Models", Annals of Economic and Social Measurement, 3/4, 653-665.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", Journal of Econometrics, 31, 307-327.

Bollerslev, T. (1987), "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return", Review of Economics and Statistics, 69, 542-547.

Box, G.E.P. and G.M. Jenkins (1976), Time Series Analysis Forecasting and Control, second edition, San Francisco: Holden Day.

Cheung, Y.-W. (1991), "Long Memory in Foreign Exchange Rates", mimeo.

Cheung, Y.-W. and F.X. Diebold (1991), "On Maximum Likelihood Estimation of the Differencing Parameter of Fractionally Integrated Noise with Unknown Mean", University of Pennsylvania working paper.

Choi, S. and M.E. Wohar (1990), "Real Exchange Rates: Evidence From Fractional Integration Tests", mimeo.

Diebold, F.X., S. Husted and M. Rush (1991), "Real Exchange Rates Under the Gold Standard", Journal of Political Economy, 99, 1252-1271.

Diebold, F.X. and G.D. Rudebusch (1989), "Long Memory and Persistence in Aggregate Output", Journal of Monetary Economics, 24, 189-209.

Diebold, F.X. and G.D. Rudebusch (1991), "On the Power of Dickey Fuller Tests Against Fractional Alternatives", Economics Letters, 35, 155-160.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation", Econometrica, 50, 987-1008.

Engle, R.F. (1983), "Estimates of the Variance of US Inflation Based Upon the ARCH Model", Journal of Money, Credit and Banking, 15, 286-301.

Fox, R. and M.S. Taqqu (1986), "Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time-Series", Annals of Statistics, 14, 517-532.

Friedman, M. (1977), "Nobel Lecture: Inflation and Unemployment", Journal of Political Economy, 85, 451-472.

Geweke, J. and S. Porter-Hudak (1983), "The Estimation and Application of Long Memory Time Series Models", Journal of Time Series Analysis, 4, 221-238.

Granger, C.W.J. (1978), "New Classes of Time Series Models", Statistician, 27, 237-53.

Granger, C.W.J. (1980), "Long Memory Relationships and the Aggregation of Dynamic Models", Journal of Econometrics, 14, 227-238.

Granger, C.W.J. and R. Joyeux (1980), "An Introduction to Long Memory Time Series Models and Fractional Differencing", Journal of Time Series Analysis, 1, 15-39.

Hosking, J.R.M. (1981), "Fractional Differencing", Biometrika, 68, 165-176.

Hosking, J.R.M. (1984), "Modelling Persistence in Hydrologic Time Series Using Fractional Differencing", Water Resources Research, 20, 1898-1908.

Hosking, J.R.M. (1985), "Fractional Differencing Modeling in Hydrology", Water Resources Bulletin, 21, 677-82.

Hurst, H.E. (1951), "Long Term Storage Capacity of Reservoirs", Transactions of the American Society of Civil Engineers, 116, 770-99.

Hurst, H.E. (1956), "Methods of Using Long-Term Storage Reservoirs", Proc. Inst. Civil Engrs., 1, 519-43.

Janacek, G.J. (1982), "Determining the Degree of Differencing for Time Series via the Log Spectrum", Journal of Time Series Analysis, 3, 177-83.

Jonas, A.B. (1983), "Persistent Memory Random Processes", Doctoral Dissertation, Department of Statistics, Harvard University.

Klemes, V. (1974), "The Hurst Phenomenon - A Puzzle?", Water Resources Research, 10, 675-88.

Kashyap, R.L. and K.-B. Eom (1988), "Estimation in Long-Memory Time Series Model", Journal of Time Series Analysis, 9, 35-41.

Kashyap, R.L. and P.M. Lapsa (1984), "Synthesis and Estimation of Random Fields Using Long Correlation Models", IEEE Trans. Pattern Anal. Machine Intell., PAMI-6,800-9.

Künsch, H. (1986), "Discrimination Between Monotonic Trends and Long-Range Dependence", Journal of Applied Probability, 23, 1025-30.

Lawrance, A.J. and N.T. Kottegoda (1977), "Stochastic Modelling of Riverflow Time Series", Journal of the Royal Statistical Society, A 140, Part 1, 1-47.

Li, W.K. and A.E. McLeod (1986), "Fractional Time Series Modelling", Biometrika, 73, 217-221.

Lo, A.W. (1991), "Long-Term Memory in Stock Market Prices", Econometrica, 59, 1279-1313.

Mandelbrot, B.B. (1972), "Statistical Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis", Annals of Economic and Social Measurement, 259-90.

Mandelbrot, B.B. (1975), "Limit Theorems on the Self-Normalized Range for Weakly and Strongly Dependent Processes", z. Wahrscheinlichkeits Theorie verw. Gebiete, 31, 271-85.

Mandelbrot, B.B. (1982), Fractals. Freeman Press: San Francisco.

Mandelbrot, B.B. and M. Taqqu (1979), "Robust R/S Analysis of Long-Run Serial Correlation", Bulletin of the International Statistical Institute, 48, Book 2, 59-104.

Mandelbrot, B.B. and J.W. Van Ness (1968), "Fractional Brownian Motions, Fractional Noises and Applications", SIAM Review, 10, 422-37.

Mandelbrot, B.B. and J. Wallis (1968), "Noah, Joseph and Operational Hydrology", Water Resources Research, 4, 909-18.

Mandelbrot, B.B. and J. Wallis (1969a), "Computer Experiments with Fractional Gaussian Noises, Parts 1, 2, 3", Water Resources Research, 5, 228-67.

Mandelbrot, B.B. and J. Wallis (1969b), "Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long-Run Statistical Dependence", Water Resources Research, 5, 967-88.

McLeod, A.I. and K.W. Hipel (1978), "Preservation of the Rescaled Adjusted Range", Water Resources Research, 14, 491-518.

Sowell, F.B. (1990), "Modeling Long Run Behavior with the Fractional ARIMA Model," GSIA Carnegie Mellon University working paper.

Sowell, F.B. (1992), "Maximum Likelihood Estimation of Stationary Univariate Fractionally-Integrated Time-Series Models", Journal of Econometrics, forthcoming.

Tieslau, M.A. (1991), "Long Memory Models and Macroeconomic Time Series", Econometrics and Economic Theory Workshop Paper, 9005, Michigan State University.

Wallis, J.R. and N.C. Matalas (1970), "Small Sample Properties of H and K, Estimators of the Hurst Coefficient h", Water Resources Research, 6, 1583-94.

Whittle, P. (1951), Hypothesis Testing in Time Series Analysis. Hafner: New York.

III. LONG-TERM PERSISTENCE, MEAN REVERSION, AND STATIONARITY: A MODEL OF INFLATION AS AN ARFIMA-GARCH PROCESS

CHAPTER III

LONG-TERM PERSISTENCE, MEAN REVERSION, AND STATIONARITY: A MODEL OF INFLATION AS AN ARFIMA-GARCH PROCESS

1. INTRODUCTION

As discussed previously in Chapter II, time series which exhibit long memory are often encountered in economics. This may be attributable to the large number of aggregations of many dynamic components in economic data. The fractionally integrated process proposed by Granger and Joyeux (1980) and Hosking (1981), which has been described in some detail in Chapter II, is very useful in applied economic analysis of persistent time series. The use of fractional integration in modelling long-memory time series offers many advantages not afforded by other modelling processes; when examining the long-run characteristics of variables, which is often the focus of macroeconomic analysis, the fractionally integrated model is able to account for persistent behavior in the data unlike those models which are based on relatively short correlation structures. This allows for a more appropriate investigation of the true long-run characteristics of variables which contain long memory. In addition, the fractionally integrated model offers the advantage of allowing for simultaneous modelling of both long- and short-term persistence in a series. For these reasons the fractionally integrated process provides an excellent framework in which to investigate the long-run properties of many macroeconomic data.

This chapter uses the fractionally integrated long-memory process to examine the long-run time series properties of the Consumer Price Index

(CPI) inflation rate for both high- and low-inflation economies. The stationary component of the model, represented as an ARMA process, combined with the persistent, fractionally integrated component produces the Autoregressive-Fractionally Integrated Moving Average (ARFIMA) model. An analysis of the inflation series in this context is of particular interest due to the relative importance of the aggregate price level in understanding the workings of the macroeconomy. The conventional methodology thus far applied by macroeconomists in analyzing the question of whether the inflation rate is stationary and mean reverting, or alternatively whether it is non-stationary and contains a unit root, has by no means produced a conclusive answer to this question. An analysis of this issue in the context of persistent, long-memory time series can provide valuable insight into many of the issues of the long-run behavior of the inflation rate and subsequent implications this has for the economy with regard to disinflationary macroeconomic policy, and can even provide a better understanding of other economic variables which are in some way related to the aggregate price level.

The plan of the rest of the chapter is as follows. The next section discusses some of the key issues in macroeconomics which are relevant in the context of long-memory time series with a specific focus given to issues pertaining to the inflation rate. Section 3 discusses the issue of the variability of inflation and the subsequent implications this has for the macroeconomy. This section also provides a discussion of the Friedman Hypothesis, which posits a direct relationship between increased levels of inflation and increased inflation variance, and discusses the subsequent implications for the macroeconomy. Section 4 examines the characteristics of the CPI inflation rate for the ten countries considered in this analysis and provides the motivation for considering the inflation rate in the context of the fractionally integrated, long-memory model. The countries examined include the Group of Seven countries, which are considered to be relatively low-inflation economies, and three additional countries, Argentina, Brazil, and Israel, which are considered to be relatively high-inflation economies. Section 5 discusses estimation of the ARFIMA-GARCH model for these ten countries. The final section presents the conclusions and gives a brief summary of the results.

2. LONG MEMORY, MEAN REVERSION, AND THE PERSISTENCE OF INFLATION

The use of the fractionally integrated ARMA process proposed by Granger and Joyeux (1980) and Hosking (1981) in modelling economic time series has proven to be an attractive framework for dealing with variables which exhibit nonstationarity in the form of long-term persistence. This model is particularly appealing in this context since it does not impose full integer orders of integration, or unit roots, on series. In the field of economics this is especially appropriate since it is not always the case that series which appear to be nonstationary are best described as containing unit roots. In particular, there is some evidence that the imposition of a unit root structure may not be reasonable for many macroeconomic time series. This has important implications since the existence of a unit root in a time series implies that innovations or individual shocks to the system will not die out over time but, rather, will persist indefinitely into the future.

Although long-memory time series exhibit some degree of persistence, indicating that shocks to the system will tend to move the series away from the value of its mean for some period of time, the persistence does decline over time (though slowly) so that eventually there is reversion to the mean of the series. In the field of macroeconomics in particular, the inflation rate appears to be one such variable which exhibits behavior typical of long-memory, fractionally integrated time series. The inferences made about the characteristics of inflation may hinge critically upon the type of model assumed to underlie this variable and as such is the motivating factor behind the importance of considering the inflation rate in the context of the fractionally integrated model.

The nature of the true long-run characteristics of the inflation rate has long been a concern among macroeconomists, especially given the crucial role that this variable plays in understanding the workings of the macroeconomy. A central issue of concern in assessing the long-run characteristics of the inflation rate involves an understanding of how aggregate prices respond to shocks in the economy. If the inflation rate is viewed as being a stationary, or I(0), variable then macroeconomic shocks will be viewed as transitory in that their effect on the rate of inflation will die out relatively quickly. On the other hand, if the inflation rate is thought to be a nonstationary, or I(1), variable possessing a unit root, then shocks to the system are forever incorporated into the rate of inflation. Klein (1976) and Nelson and Schwert (1977) have investigated the inflation rate in this context and have found evidence of a unit root in the series, implying nonstationarity of the inflation rate. Alternatively, Barsky (1987) investigated the characteristics of inflation over two unique time periods and found evidence of a stationary inflation rate prior to 1960, with evidence of nonstationarity thereafter. Subsequent research assessing the characteristics of inflation includes the work of Ball and Cecchetti

(1990) who decomposed inflation into two components, one being a transitory component and the other being a permanent component represented as a random walk. Their findings indicate that, in aggregate, the inflation rate appears to be nonstationary.

The issue of whether or not the inflation rate is stationary has significant implications which extend into many areas of the macroeconomy. The finding of a nonstationary inflation rate, which indicates the existence of a unit root in the series, may not seem reasonable from an economic standpoint since this implies that transitory shocks to the economy have a permanent effect on the inflation rate. This has a wide range of subsequent implications for the economy especially with regard to constructing optimal policy rules (see, for example, McCallum (1988) and Baillie (1989)).

An examination of the true long-run characteristics of the inflation rate can prove to be valuable in many areas of the macroeconomy since there are several variables whose properties are closely tied to those of the inflation rate. Such variables typically include those series which incorporate some form of inflationary expectations or are formed as some linear combination of the inflation rate and other variables. For example, in studying the properties of interest rates Fama and Gibbons (1982) and Mankiw and Miron (1986) found evidence that the nominal rate of interest is nonstationary, or contains a unit root; Fama (1975) found evidence that the real rate of interest is stationary, or does not contain a unit root. These findings imply that the inflation rate is necessarily nonstationary, and so contains a unit root, and should therefore be cointegrated with the nominal interest rate.

The issue of whether the inflation rate is stationary or is nonstationary and contains a unit root has strong implications for research in the consumption based capital asset pricing model (CAPM) as well. Rose (1988) investigates the properties of real interest rates and the implications that this has for the CAPM. Rose (1988) finds evidence that the real rate of interest is non-stationary and gives attention to the possible implications that would arise from the possibility of ex-post real rates being nonstationary. Sweeney (1987) also investigates the CAPM and examines the effect of the characteristics of inflation and inflation variability on the demand for real cash balances and the allocation of asset portfolios. In this analysis evidence is found of an inverse relationship between the variability of inflation and the attractiveness of holding cash balances, and a direct relationship between the variability of inflation and the proportion of assets held in equities. These relationships led Sweeney to conclude that there would be a reduction of equity's beta coefficient in the CAPM model and in the discount rate of the firm.

3. THE VARIABILITY OF INFLATION

One further issue which has long been of interest to economists, especially in assessing the welfare costs associated with increases in expected or unexpected changes in inflation, is the variability of inflation. In analyzing the apparent tradeoff between inflation and unemployment which was implied by the Phillips curve, Okun (1971) first posited a direct link between a high rate of inflation and a high variability of inflation. Okun provided the origin to the notion that stabilization policies in the economy were the driving force behind the positive correlation between the level and the variability of inflation.

His argument, which was based on the non-linearity of the Phillips curve, maintained that countries which concentrated on output or employment stability would be operating at the steep end of the Phillips curve and hence would be more likely to experience not only higher average levels of inflation but also increased variability of inflation.

Since the early work of Okun (1971), numerous studies have been advanced which were directed not only at uncovering empirical evidence of the relationship between the level and variability of inflation, but also at providing some theoretical framework within which this relationship could be grounded [Gordon (1971), Logue and Willett (1976), Foster (1978), Parks (1978), Cukierman and Wachtell (1979), Fischer (1981), Taylor (1981), Pagan, Hall, and Trivedi (1983), Bairam (1988), Demetriades (1988), and Devereux (1989)]. Foster (1978) also noted the positive correlation between the level of inflation and its variance but criticized Okun's treatment for not concentrating on the unanticipated element of inflation. Other studies, such as Pagan, Hall, and Trivedi (1983) and Demetriades (1988), have built upon the intuitive explanation offered by Okun by incorporating a Lucas (1973) - type model to provide a theoretical explanation of the correlation between inflation and its variability.

The relationship between a high level of inflation and increased inflation variability has also been studied at great length by Friedman who, in his 1977 Nobel lecture, hypothesized that higher mean levels of inflation were likely to be associated with increased levels of the variance of inflation. Friedman emphasized the welfare costs of the variability of inflation and noted its effect on the macroeconomy in general. According to Friedman, higher rates of inflation would likely reduce forecastability of future inflation, leading to an increased element of uncertainty in individual decision making and reducing the efficiency of the price system. This would lead to a misallocation of resources and would likely result in decreased output stability and increased unemployment. Increased inflation variability clearly would be associated with increased welfare losses.

An examination into the validity of the Friedman hypothesis using the framework of the fractionally integrated model should provide further insight into the true characteristics of inflation and its effect on the macroeconomy. If a link between increased levels of inflation and increased inflation variability can be shown to exist this should offer the possibility of making inferences into the direction of monetary (or other) policy in maximizing net benefit.

4. INFLATION CONSIDERED AS A LONG-MEMORY PROCESS

As discussed in section 2, the notion of a nonstationary inflation rate does not seem intuitively appealing from an economic perspective since this implies that a one time increase in, say, the growth rate of the money supply, would have the ability to permanently raise the mean level of inflation forever into the future (in the absence of any counter policy maneuvers). From a statistical standpoint, the notion of a nonstationary inflation rate, or the imposition of a unit root structure in the series, does not appear to be reasonable either; the inflation rate appears to occupy the "middle ground" between series which are nonstationary when expressed in levels, and series which are stationary when expressed in first differences. The existence of this property in the inflation rate is the motivation for considering this series as a long-memory fractionally integrated process.

To demonstrate the long-term persistence which is present in the inflation rate series considered in this analysis, Table 1 presents the autocorrelation functions for the CPI inflation rate series, expressed in levels, for the Group of Seven (G-7) countries and for Argentina, Brazil, and Israel. The data are monthly, non-seasonally adjusted, span the approximate 40-year period from 1948 through 1990,[1] and have been expressed in logarithms. Each of the inflation rate series exhibits the clear pattern of slow persistent decay in their autocorrelation functions, which is behavior associated with long memory. This pattern of persistent decay is not typical of stationary processes. Table 2 presents the autocorrelation functions of the first differences of the inflation rate series for these countries. In general, these series appear to be over differenced when full first differences are taken, as indicated by the large negative autocorrelations at the initial lag and the relative absence of correlation thereafter. These results certainly are not typical of series which contain a unit root.

The results presented in Tables 1 and 2 support the hypothesis of inflation being a fractionally integrated series which requires some differencing to achieve stationarity yet is over differenced when first differences are taken. Further evidence of the fractionally integrated nature of the series represented in Tables 1 and 2 can be obtained by "filtering", or differencing, each series and examining the autocorrelation functions of the resulting filtered series. Tables 3 through 12 present the autocorrelation functions of the inflation series which have been filtered for discrete values of d, the fractional

[1] The exact time periods spanned by each individual series are given in the keys to Tables 3-13.

diff

func

full

disc

seri

on t

filt

filt

rang

stat

degr

betw

autoc

rate

range

for

provi

remai

some

diffe

the s

appro

the

approp

perfor

differencing parameter. For each inflation series, the autocorrelation functions are presented for the extremes of no filtering, d = 0.00, and full first differencing, d = 1.00, along with filtering at various discrete levels of d. These results generally indicate that each of the series becomes stationary after some amount of differencing is performed on the series and before a full first difference is imposed.

For example, Table 12 presents the autocorrelation functions for the filtered inflation series for the U.S. Between the extremes of no filtering, d = 0.00, and full first differencing, d = 1.00, there is a range of values of d, .30 to .50, in which the series appears to exhibit stationarity. This would imply for the U.S. inflation rate that the degree or order of fractional differencing, d, should be approximately between .30 and .50. As another example, Table 11 presents the autocorrelation functions for the fractionally differenced U.K. inflation rate. This series clearly becomes stationary when differenced in the range of .20 to .40, thereby implying an order of fractional differencing for the U.K. somewhere in the range of .20 to .40. Similar results are provided in Tables 3 through 10 for each of the inflation series of the remaining eight countries of this analysis. In studying these results, some insight may be gained into the approximate degree of fractional differencing, or value of d, required to achieve stationarity of each of the series studied in this analysis.

The results presented in Tables 1 through 12 strongly indicate the appropriateness of considering the inflation rate series in the context of the fractionally integrated model. To further investigate the appropriateness of this specification, standard tests for stationarity are performed on each of the series. Most conventional tests for stationarity offer a null hypothesis of nonstationarity, implying the presence of a unit root in the data. Due to the theoretical structure of classical statistical hypothesis testing, it is generally the case that the null hypothesis is only rejected when there is particularly strong evidence to the contrary. Recently, however, an alternative approach to testing for stationarity has been proposed by Kwiatkowski, Phillips, Schmidt and Shin (1992), hereafter referred to as KPSS, who have proposed a parameterization which offers a null hypothesis of stationarity. Since the conventional parameterizations of Dickey and Fuller (1979, 1981) and Phillips and Perron (1988) generally are most appropriate for testing series which are strongly believed to be nonstationary, the KPSS specification is most useful to the investigator who has strong priors against nonstationarity and so wishes to test stationarity as the null.

The approach formulated by Kwiatkowski, Phillips, Schmidt, and Shin (1992) to test for stationarity utilizes an unobserved components representation which assumes that the series under investigation may be written as the sum of a deterministic trend, a random walk, and a stationary error process as:

(1.) $$y_t = \xi t + r_t + \epsilon_t \,.$$

In this specification, $r_t = r_{t-1} + u_t$, u_t is iid $(0, \sigma_u^2)$, the parameter ξ is used to test for level stationarity, and the null hypothesis of trend stationarity involves testing whether $\sigma_u^2 = 0$. KPSS define the partial sum process of the residuals from a regression of y_t on [1,t], e_i, as

$$S_t = \sum_{i=1}^{t} e_i \,, \quad t = 1, 2, \ldots, T$$

where T represents the sample size. The score test of the null hypothesis of stationarity, which is an upper tail test, is based on the statistic

$$\eta = T^{-2}\Sigma S_t^2/s^2(k),$$

when the series has been regressed on an intercept and also possibly a time trend. The estimate of the disturbance variance, $s^2(k)$, is computed in the same manner in which its equivalent in the Phillips and Perron (1988) test is computed; a Bartlett window adjustment based on the first k sample autocovariances is used, as suggested by Newey and West (1987). The test statistic $\hat{\eta}_\mu$ represents the statistic when the residuals are computed from a regression with only an intercept. The test statistic $\hat{\eta}_\tau$ represents that for which a time trend is included in the initial regression. Both $\hat{\eta}_\mu$ and $\hat{\eta}_\tau$ are shown to be asymptotic functions of a Brownian bridge under the null of stationarity. The critical values, which are produced in KPSS (1992), for $\hat{\eta}_\mu$ and $\hat{\eta}_\tau$ are .739 and .216 at the .01 level of significance, and .463 and .146 at the .05 level of significance, respectively.

For fractionally integrated series, neither the hypothesis of nonstationarity nor that of stationarity describes the processes well. Consequently, the combined application of both a conventional unit root test such as the Phillips and Perron (1988) test which has a null hypothesis of nonstationarity, and the KPSS (1992) test which has a null hypothesis of stationarity, can be used to detect the presence of fractionally integrated series. Application of the Phillips and Perron test, denoted PP, and the Kwiatkowski, Phillips, Schmidt, and Shin test produces four possible outcomes which are summarized as:

	Reject H_0 PP test: H_0: I(1)	Do Not Reject H_0 PP test: H_0: I(1)
Reject H_0 KPSS test: H_0: I(0)	(i) Inflation not well represented as I(1) or I(0). POSSIBLE EVIDENCE FOR FRACTIONAL INTEGRATION	(ii) Strong evidence of a unit root exists in the data. I(1)
Do Not Reject H_0 KPSS test: H_0: I(0)	(iii) Strong evidence of stationarity exists in the data. I(0)	(iv) The data is insufficiently informative on the long run characteristics of the series.

Rejection of both null hypotheses for a given series, scenario (i) above, indicates that the series is not well described by either an I(1) nonstationary or an I(0) stationary process, and consequently may be evidence of fractionally integrated behavior.

Table 13 presents the results of applying the above tests to the inflation rate series for each of the ten countries represented in Table 1. Scenario (i), in which both null hypotheses are rejected, arises for eight out of ten countries: Argentina, Brazil, Canada, France, Italy, Israel, the U.K. and the U.S. The implication, then, is that the inflation rate series for each of these countries are not well described as being either stationary or nonstationary, which suggests that the series may be fractionally integrated.

The results for Germany and Japan, however, seem to indicate that the inflation rate series for these two countries are stationary. This result is not surprising given the extent to which officials in these countries have intervened, historically, in the operation of their economies to maintain a steady and stable rate of inflation. These results of a stationary inflation rate for Germany and Japan are further confirmed in subsequent estimation. That is, for Japan the estimated value of the parameter of integration, $\hat{d}$, is found to be of the same magnitude as that of the moving average parameter such that cancellation of the two result in a value of d near zero. Similarly for Germany, the estimated value of d is found to be close to zero. These results support the hypothesis of a stationarity inflation rate series for these two countries.

Table 14 presents the results of applying the Geweke and Porter-Hudak (1983) estimation technique, which is described in detail in chapter II, to the inflation series of the ten countries: Argentina, Brazil, Canada, France, Germany, Israel, Italy, Japan, the U.K. and the U.S. The parameter of integration, d, has been estimated over the range of low-frequency ordinates used in the spectral regression, as suggested by Geweke and Porter-Hudak, for the values T^{α}, where α = .50, .525, .55, .575, and .60. The results indicate the extreme sensitivity of the parameter estimate to the number of ordinates used in the spectral regression. The estimated value of d for each country's inflation series varies within a substantial range, depending upon the number of ordinates used during estimation.

Alternatively, Table 14 also provides the results of estimating the parameter of integration, d, for each country's inflation rate using the number of ordinates for the spectral regression as suggested by Sowell (1990). Since each of the low-inflation economies span approximately 41 years of data, and each of the high-inflation economies span approximately 31 years of data, and assuming a low-frequency period of five years for the inflation rate series, the values of m implied by this technique for the low- and high-inflation economies are 6 and 8, respectively. This approach further confirms the extreme sensitivity of the GPH estimation technique to the range over which the spectral regression is estimated.

The results presented in Table 14 of estimating d for the inflation rate series using the Geweke and Porter-Hudak (1983) estimation technique clearly indicate the sensitivity of this procedure to the data used in the analysis. These results indicate the importance of considering alternative estimation procedures when dealing with fractionally integrated series. Consequently, the ARFIMA-GARCH model developed in Chapter II has been applied to the inflation rate series for the G-7 countries, Argentina, Brazil, and Israel to estimate both the parameter of fractional differencing for these series as well as all other model parameters of the series. By applying an appropriate modelling procedure which can account for not only the long-memory characteristics of the series but the time varying homoskedasticity as well, it will be possible to examine the true long-run characteristics of the inflation rate which should aid in our understanding of the macroeconomy.

5. ESTIMATION OF THE ARFIMA-GARCH MODEL

Section 5 of Chapter II detailed the ARFIMA(0,d,1) - GARCH(1,1) model as represented by equations (11.) through (14.) of that chapter and discussed the two estimation procedures termed Method I and Method II. These estimation procedures were applied to the CPI inflation rate series for the G-7 countries as well as Argentina, Brazil, and Israel. The estimation procedure for each of these countries and the results of estimation are outlined below.

The ARFIMA(0,d,1) - GARCH(1,1) model as applied to the U.S. CPI inflation rate can be expressed by the following equations:

(2.) $$100(1-L)^{d}\ \Delta \log CPI_t = b + \epsilon_t + \theta\epsilon_{t-1} + \delta\sigma_t$$

(3.) $$\epsilon_t|\Omega_{t-1} \sim t(0,\ \sigma_t^2,\ \nu^{-1})$$

(4.) $$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\Delta \log CPI_t.$$

In this specification, $y_t = 100\Delta\log CPI_t$ is the consumer price index measure of inflation. The random error, ϵ_t, is assumed to follow a conditional density D, assumed to be either normal or Student t, with mean zero and variance σ_t^2. The error ϵ_t is conditional on the information set at time (t-1), Ω_{t-1}. The model parameters to be estimated include: b, the mean of inflation; δ, the effect of the volatility of inflation on the mean level of inflation; θ, the moving average parameter; ω, the intercept in the conditional variance; α, the effect of lagged squared residual on the conditional variance (the ARCH effect); β, the effect of lagged conditional variance on the current variance (the GARCH effect); γ, the effect of lagged inflation on volatility; and, due to the presence of excess kurtosis in the series, the degrees of freedom from the Student t distribution, ν, must also be estimated.

The specification of the ARFIMA-GARCH model will allow for empirical tests of the Friedman hypothesis as well as tests for the presence of feedback between lagged inflation and the degree of volatility of inflation in the current period. The validity of the Friedman hypothesis is investigated through the parameter δ; when $\delta \neq 0$, volatility is allowed to influence the mean of inflation in that a positive and significant value for this parameter indicates that higher levels of inflation are associated with higher volatility of the series. Such a finding would give positive empirical support to the Friedman hypothesis. The model specification will also allow for a test of whether lagged inflation, which is predetermined, enters the conditional variance equation (4.). A positive and significant value for the parameter γ will lend support to the hypothesis that last period's inflation rate is directly correlated with a higher value of the current period's inflation variance. The results of estimation of δ and γ are presented in Table 25 and are discussed in more detail later in this section.

In estimation of each country's inflation rate the statistics m_4 and m_3, which are measures of the sample kurtosis and skewness, respectively, are estimated numerically by the Berndt *et. al.* (1974) algorithm. In the case where the true distribution of a series is normal, m_4 and m_3 will have asymptotic distributions given by $N(0,24/T)$ and $N(0,6/T)$, respectively. The estimated value of m_4 is used as a diagnostic to determine whether the distribution of the model should be assumed normal or Student t, since the presence of significant kurtosis in the residuals indicates the inappropriateness of the assumption of normality of the unconditional distribution of $100\Delta\log CPI_t$. When evidence of significant kurtosis is present in a series as evidenced by the estimated value of m_4, a Student t distribution, rather than the normal, will be assumed for the model. Use of the Student t distribution will necessitate estimation of the degrees of freedom parameter, ν. The appropriateness of the Student t distribution can be examined by comparing the estimated value of the sample kurtosis, m_4, to the level kurtosis implied by the estimated degrees of freedom parameter. That is, in using the Student t distribution, the estimated value of ν implies a conditional kurtosis of $3(\hat{\nu} - 2)/(\hat{\nu} - 4)$, and this value may be directly compared to the estimated value of m_4.

The estimation procedure employed in this analysis also produces the standard Ljung and Box (1978) test statistic, Q(k), which tests for k th order serial correlation in the estimated residuals. In addition, the statistic $Q^2(k)$ is also calculated and this statistic is used to test for k th order serial correlation in the squared residuals. The $Q^2(k)$ statistic will be used in this analysis to provide an LM test of the ARCH(k) specification where a null hypothesis of no serial correlation in the squared residuals of the inflation series is tested against the alternative hypothesis of k th order correlation. Under the null hypothesis of conditional homoskedasticity, these statistics will be asymptotically distributed as chi-squared with k degrees of freedom. In this particular analysis k will be set equal to 10 so that the Lung-Box statistics Q(10) and $Q^2(10)$ are calculated in estimation of the model for each country.

It is worth noting the problems of interpreting the Ljung-Box statistics in the case of the model for the inflation rate, however, since the power of these statistics can be influenced by the presence of significant ARCH effects in the data generating process. As pointed out by Cumby and Huizinga (1988), when testing residuals for autocorrelation, the presence of heteroskedasticity in a series (as is the case with the inflation rate) will tend to bias the Ljung-Box statistics towards rejecting the null hypothesis.

The ARFIMA-GARCH model described above is used to estimate all model parameters given in equations (2.) through (4.) of this section, for the inflation rate series of each of the ten countries of this analysis. Table 15 presents the results of estimating the ARFIMA(0,d,1)-GARCH(1,1) process for the U.S. The first column of the table presents the full maximum likelihood estimates of the parameters of the model, obtained by Method I which was described in some detail in Chapter II. The MLE of d by this method is .36, which implies that the series is highly persistent but none the less mean reverting. In this way, the inflation series for the U.S. can be considered to be a stationary process since mean reversion implies that shocks to the series will not persist indefinitely into the future but, rather, will die out over time.

The results of estimation of the model by Method II, as described in Chapter II, are also presented in Table 15. This involves estimation over the transformed, or filtered, series for ten values of d, 0.00 through 0.90. That is to say, the maximum likelihood estimation procedure is performed ten times for the inflation rate series and the resulting parameter vector which produces the maximized log likelihood is taken to be the true MLE. In the case of the U.S. model, in observing the values of the maximized log likelihood function produced at each value of d, it appears that the log likelihood is relatively flat in the range of d from .30 to .50. The maximized value occurs when d = .40, which is consistent with the ML estimate of d obtained by Method I, and also with the value predicted by observing the autocorrelation functions of the filtered U.S. inflation series presented in Table 12. The value of the Ljung-Box statistics at the MLE indicates that after fitting the GARCH(1,1) model to the inflation rate series the null hypotheses of uncorrelated residuals and uncorrelated squared residuals cannot be rejected. In addition, the estimated value of ν at the MLE implies a conditional kurtosis of 4.28

which is relatively close to the estimated value of 4.42, indicating the appropriateness of the use of the Student t distribution for the U.S. model.

It is interesting to note that a clear trade off exists for the maximized conditional log likelihood between the estimated moving average parameter, $\hat{\theta}$, and the fractional differencing parameter, $\hat{d}$, as can be observed in Table 15. As the value of the fractional differencing parameter increases the estimated value of the moving average parameter decreases. A similar trade off can be observed for the maximized conditional log likelihood between the estimated mean parameter, $\hat{b}$, and the fractional differencing parameter; the MLE of $\hat{b}$ does appear to change conditional on $\hat{d}$.

The ARFIMA-GARCH model was applied to the inflation rate series of the remaining nine countries: Argentina, Brazil, Canada, France, Germany, Israel, Italy, Japan, and the United Kingdom, and the results of estimation are similar to that of the U.S. For each of these countries, however, some degree of seasonality in the conditional mean of the inflation rate series was apparent. This seasonality can be observed by noting the pattern of decay of the autocorrelation functions of each of the countries presented in Table 1. For each country except the U.S., the correlation decreases steadily to zero as the lag increases, but picks up somewhat around lag twelve and then continues to decline again after this point.

Consequently, an ARFIMA(0,d,13)-GARCH(1,1) process was estimated for each of these countries to account for the seasonality in the data. This model includes moving average terms of lags 1, 12, and 13 to account for this seasonality. The multiplicative seasonal restriction of $\theta_{13} = \theta_1\theta_{12}$, however, was not imposed in estimation. The model for these countries may be represented by the following equations:

$$(5.)\qquad 100(1-L)^{d}\,\Delta\log CPI_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_{12}\epsilon_{t-12} + \theta_{13}\epsilon_{t-13} + \delta\sigma_t$$

$$(6.)\qquad \epsilon_t|\Omega_{t-1} \sim N(0,\ \sigma_t^2)$$

$$(7.)\qquad \sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\,\Delta\log CPI_t,$$

This model specification is very similar to that for the U.S. except that the above specification includes the two additional parameters, θ_{12} and θ_{13}, which account for the seasonality of the model. In addition, the above model assumes a zero mean for the inflation rate.

In the cases of France and Israel, the inflation rate series were found to exhibit a substantial degree of excess kurtosis and so the models for these countries were estimated using the Student t density rather than the normal. Recall that the estimated value of the degrees of freedom parameter estimated under the Student t distribution implies a conditional kurtosis of $3(\hat{\nu} - 2)/(\hat{\nu} - 4)$. In the case of the French model, for example, the estimated value of $\hat{\nu}$ implies a conditional kurtosis of 10.22 which is relatively close to the estimated sample kurtosis for France, $m_4 = 9.22$. This will indicate the appropriateness of the use of the Student t distribution over that of the normal in estimating the models for these countries.

Tables 16 through 24 present the results of estimating the ARFIMA(0,d,13)-GARCH(1,1) process via Method II for the remaining nine countries; results for Method I are not reported due to the excessive computational difficulty of applying direct maximization to the seasonal moving average model. As with the U.S. model, the value of the Ljung-Box statistics for the models of each of the remaining countries indicates that at the MLE there is no case in which the null hypothesis of no serial correlation in the squared residuals can be rejected against the alternative hypothesis of tenth-order serial correlation.[2] For Canada, France, Germany, Italy, and the U.K., which are all considered to be relatively low-inflation economies, the log likelihood function is maximized at a value of d which is greater than zero but less than or equal to ½. This implies that the inflation series for these countries, like that for the U.S., are highly persistent but none the less mean reverting. In this way these series may be viewed as stationary processes, contrary to what many researchers have found in examining the inflation rate within the framework of models other than that of the fractionally integrated model.

The results for the relatively high-inflation economies of Argentina, Brazil, and Israel also indicate the highly persistent and mean reverting behavior of the inflation series for these countries. The log likelihood functions for these countries were maximized when the value of d was less than one, but also greater than ½. This indicates that the unconditional variance of the inflation rate for these relatively high-inflation economies is infinite. The fact that all three high-inflation economies exhibit this characteristic should not be surprising due to the great variability in the inflation rate experienced by these countries

---

[2] In each case, the null hypothesis cannot be rejected at the 10% level of significance and in some cases the null hypothesis also cannot be rejected at the 5% level of significance.

during the time period under investigation. Perhaps somewhat more surprising is the result that for France and Italy the MLE of d is found to be exactly equal to ½. This implies an infinite unconditional variance for these countries as well.

The results of applying Method II to the remaining countries' inflation series appear to be quite similar to the results for the U.S. The same trade off may be observed for the countries represented in Tables 16 through 23 as was observed for the U.S. That is, as the value of the fractional differencing parameter increases, the value of the estimated moving average parameter, $\hat{\theta}_1$, decreases. In addition, all of the estimated models with the exception of the U.K. exhibit strong persistence in their variances as evidenced by the sum of the estimated ARCH and GARCH parameters, α and β, being close to one. Again, this confirms the highly persistent nature of the inflation rate series for these countries.

For each of the ten countries examined in this analysis, likelihood ratio tests were performed to investigate whether the value of d produced by Method II was significantly different from zero or from one. That is, a test of the null hypothesis that $d = 0.00$ versus the alternative hypothesis that d was equal to the MLE produced by Method II was performed for those inflation series for which $0 < \hat{d} < ½$. Similarly, a test of the null hypothesis that $d = 1.00$ versus the alternative hypothesis that $d = \hat{d}_{MLE}$ was performed for those series for which $½ < \hat{d} < 1$. In each case the results indicate that the null hypothesis could be strongly rejected against the MLE value of d. The results of these tests for the U.S. are presented in Table 15.

Table 25 presents the results of likelihood ratio tests which were designed to test the validity of the Friedman hypothesis and also to test for the presence of a feedback relationship between lagged inflation and the conditional variance of inflation in the current period. The first row of Table 25 provides the results of testing for whether lagged volatility Granger causes the mean of inflation. As discussed in section 4, this test allows the volatility of inflation (the standard deviation) to influence the mean of inflation through the parameter δ in equation (2.) for the U.S. model and equation (5.) for the models of the remaining countries. The results indicate that for the low-inflation economies of Canada, France, Italy, Japan, and the U.S., there is no evidence of volatility causing the mean of inflation. The results for the U.S. are consistent with the findings of previous studies by Engle (1983) and Cosimano and Jansen (1988). For the high-inflation economies, and also surprisingly for the U.K., there is strong evidence of joint feedback between the conditional mean and variance of inflation.

The second row of Table 25 provides the results of testing whether lagged inflation Granger causes inflation volatility. This test allows lagged inflation to influence the volatility of inflation through the parameter γ in the conditional variance equation, equation (4.) for the U.S and equation (7.) for the remaining countries. This can be interpreted as a direct test of the Friedman hypothesis which states that the volatility or uncertainty of inflation increases in high inflation regimes. The results indicate that for the low-inflation economies of Canada, France, Germany, Italy, and Japan, there is no support for a valid Friedman hypothesis. These findings are consistent with those of Gordon (1971), Logue and Willett (1976), and Fischer (1981) who failed to find evidence of a positive correlation between the level and variability of inflation for relatively highly industrialized economies. The results for the high-inflation economies and the U.K., on the other hand, indicate that there is strong evidence in support of the Friedman hypothesis. For these countries, the parameter γ is found to be positive and significant. These results indicate that for the high-inflation economies, and again also for the U.K., periods of increased inflation should be expected to be associated with periods of increased inflation variability. This is consistent with the findings of Logue and Willett (1976) and Fischer (1981) who noted that for economies experiencing relative instability, for example in the form of hyperinflation or political unrest, there existed a significant positive correlation between increased levels of inflation and increased inflation variability.

The finding of empirical support for the Friedman hypothesis for only the relatively high-inflation economies should not be surprising if one considers that the hypothesized link between inflation and its variance was most likely directed at high inflation economies (see Friedman 1977). Logue and Willett (1976), who found no link between inflation and its variance for economies experiencing relatively low levels of inflation, suggested that there existed some "threshold" level of inflation below which the Friedman hypothesis was not valid. That is to say, Logue and Willett hypothesized that there was some minimum level of inflation[3] below which an increase in the level of inflation would not lead to a subsequent increase in inflation variability; the higher the average rate of inflation, the more likely there was to be a positive association between the level and variability of inflation. As a result,

[3] Based on calculations performed in their analysis, Logue and Willett (1976) propose that this threshold level of inflation should lie somewhere between two to four percent.

evidence of the Friedman hypothesis should be weakest for relatively low-inflation economies, such as those of the G-7, and strongest for relatively high inflation countries, such as Argentina, Brazil, and Israel.

The apparent inconsistency of the results for the U.K. are somewhat surprising, however. In one respect the inflation rate for the U.K. behaves like the inflation rate series for the relatively low-inflation economies of the G-7 countries, having a value of d which falls within the range of $0 < d < \frac{1}{2}$. This indicates the highly persistent but mean reverting nature of the U.K. inflation series and implies a finite variance. Yet in another respect the inflation rate for the U.K. behaves like the inflation rate series for the relatively high-inflation economies of Argentina, Brazil, and Israel in that empirical support is found for a valid Friedman hypothesis and the existence of a feedback mechanism between the mean of inflation and its variance. These results seem to separate the U.K. inflation rate from the norm of either a relatively low-inflation economy or a relatively high-inflation economy, indicating some sort of atypical behavior on the part of the U.K.

6. SUMMARY AND CONCLUSION

This chapter examines the long-run characteristics of the inflation rate series, which is clearly one of the key variables in understanding the macroeconomy, in the context of the long-memory process. The model applied here allows for a more precise investigation into the true long-run characteristics of the inflation rate series since the series is considered, for the first time, within the framework of the fractionally integrated ARMA model, the ARFIMA process. The additional characteristic of the inflation rate in the form of homoskedastic error is also accounted for in the model by use of the ARFIMA-GARCH process.

Two methods of obtaining approximate maximum likelihood estimates are applied to the inflation rate series for the Group of Seven countries, which are considered to be relatively low-inflation economies, and to Argentina, Brazil, and Israel, which are considered to be relatively high-inflation economies. A distinct difference is observed in the estimated parameters of the two types of economies, implying fundamentally different long-run characteristics for high- and low-inflation economies.

The results of this analysis indicate that the inflation rate series, regardless of whether they represent high-inflation regimes or low-inflation regimes, are all highly persistent but none the less mean reverting and stationary. The inflation rate series for Argentina, Brazil, Israel, and the U.K., however, were found to exhibit infinite variances, which may not be unexpected in each case other than the U.K. due to the volatile nature of the series for these countries during the time period under which this investigation takes place.

For each of the relatively high-inflation economies of Argentina, Brazil, Israel, and also for the U.K., there is strong empirical support for the Friedman hypothesis that high inflation should be expected to be associated with increased inflation volatility. This relationship does not appear to hold for any of the relatively low-inflation economies, with the exception of the U.K. which seemed to exhibit atypical behavior. This finding should be consistent with theoretical expectations if one considers that the Friedman hypothesis was directed at high-inflation regimes only.

One issue of interest in pursuing further the findings of this analysis is to consider the effect of exogenous influences on the long-run characteristics of the inflation rate. For example, Alesina (1989) has considered the impact of political stability on an economy's performance and discusses the degree of autonomy in performing monetary policy of the central banks of several countries. The extent to which a country's central bank is divorced from the fiscal activity of the economy should have some effect on the stability of the inflation rate for that country; that is, central banks which are tied directly to government policy can often increase the volatility of inflation by their monetary policy actions. In addition, Bernake (1992) provides a useful discussion of central bank behavior and the degree of autonomy of the central banks of six industrialized countries. Fischer (1981) also discusses the effect of the use of monetary policy on the part of central banks, noting that the validity of a Friedman-type hypothesis which links the level and variability of inflation may depend heavily on the degree of accommodation in a country's monetary policy.

Two additional areas which might be of interest in examining the effect of exogenous influences on the inflation rate include the degree to which a country's wages are indexed to the inflation rate, and the degree to which central banks engage in interest rate smoothing in maintaining their policy objectives (see for example Gray (1976) and Goodfriend (1987)). These analyses may provide further insight into the questions which economists pose about the characteristics of inflation and its variability and the subsequent implications these issues have for the macroeconomy.

TABLE 1

Autocorrelations of CPI Inflation Series

	Country									
Lag	Argentina	Brazil	Canada	France	Germany	Israel	Italy	Japan	U.K.	U.S.
1	.758	.886	.434	.428	.362	.736	.253	.121	.267	.467
2	.561	.789	.369	.169	.275	.653	.280	.092	.232	.423
3	.368	.698	.409	.120	.210	.671	.316	.180	.197	.399
4	.272	.645	.380	.088	.011	.635	.240	.036	.203	.360
5	.373	.611	.362	-.102	.063	.658	.368	.098	.223	.316
6	.410	.588	.288	-.077	-.026	.658	.204	.082	.313	.305
7	.464	.557	.311	-.078	-.024	.602	.310	.025	.160	.312
8	.510	.531	.349	-.022	.047	.614	.326	.166	.146	.359
9	.420	.510	.317	-.025	.000	.615	.194	.112	.210	.386
10	.355	.506	.272	.057	.044	.600	.220	-.021	.173	.349
11	.293	.531	.311	.155	.069	.575	.194	-.031	.201	.320
12	.281	.549	.419	.222	.055	.624	.370	.151	.403	.278
13	.228	.552	.286	.193	.070	.507	.229	-.076	.156	.234
14	.219	.550	.216	.113	-.016	.470	.193	-.135	.144	.180
15	.192	.512	.235	.140	.095	.456	.246	-.083	.168	.211
16	.173	.498	.225	.102	.021	.445	.188	-.019	.106	.229
17	.164	.773	.195	-.035	.184	.461	.296	-.088	.138	.160
18	.163	.449	.144	-.043	-.135	.487	.199	.013	.192	.129
T	408	409	512	511	507	410	510	511	512	525

Key: The inflation series are defined as 100 $\Delta \log CPI_t$.

TABLE 2

Autocorrelations of First Differenced Inflation Series

	Country									
Lag	Argentina	Brazil	Canada	France	Germany	Israel	Italy	Japan	U.K.	U.S.
1	-.092	-.074	-.437	-.191	-.419	-.343	-.505	-.467	-.471	-.499
2	-.009	-.024	-.094	-.214	-.017	-.191	-.018	-.075	-.003	.112
3	-.201	-.169	.057	-.030	.063	.102	.072	.130	-.024	-.086
4	-.304	-.087	-.004	.150	-.082	-.114	-.138	-.108	-.018	.032
5	-.077	-.015	.038	-.189	.069	.047	.185	.028	-.033	-.050
6	.072	-.019	-.075	-.100	-.025	.104	-.153	.038	.160	-.013
7	.013	.002	-.018	-.018	-.054	-.130	.057	-.119	-.095	-.024
8	.281	-.024	.068	.108	.093	.021	.087	.095	-.053	.028
9	-.050	-.066	.017	-.033	-.074	.030	-.098	.058	.065	.042
10	-.008	-.151	-.070	-.050	-.012	.020	.030	-.069	-.041	-.029
11	-.104	-.094	-.072	.041	.035	-.140	-.132	-.105	-.120	.024
12	.086	.092	.222	.163	.003	.314	.209	.226	.309	.037
13	-.091	.013	-.052	.051	.089	-.151	-.070	-.094	-.161	-.034
14	.039	.204	-.084	-.151	-.035	-.044	-.052	-.065	-.025	-.089
15	-.016	-.102	.025	.062	-.134	-.005	.067	-.005	.050	.073
16	-.021	.069	.018	.082	.178	-.051	-.116	.069	-.056	.008
17	-.018	-.004	.008	-.110	-.127	-.019	.130	-.100	-.016	.084
18	-.033	-.001	-.034	-.051	.050	.159	-.064	.117	.078	-.036

Key: The series are defined as 100 $\Delta^2 \log CPI_t$.

TABLE 3

Autocorrelations of Filtered CPI Inflation Series

Argentina

	Value of d:										
Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.758	.644	.534	.435	.346	.263	.116	.112	.041	-.026	-.091
2	.561	.395	.261	.165	.099	.054	.024	.006	-.005	-.008	-.007
3	.368	.147	-.018	-.127	-.191	-.225	-.240	-.241	-.234	-.221	-.204
4	.272	.031	-.144	-.254	-.315	-.345	-.335	-.352	-.341	-.326	-.307
5	.323	.121	-.019	-.099	-.136	-.148	-.144	-.132	-.116	-.097	-.078
6	.410	.253	.146	.087	.060	.051	.051	.056	.062	.069	.075
7	.464	.326	.229	.169	.133	.109	.089	.071	.052	.033	.014
8	.510	.404	.336	.302	.288	.284	.283	.284	.284	.283	.282
9	.420	.281	.181	.118	.078	.050	.027	.006	-.014	-.033	-.052
10	.355	.207	.104	.045	.015	.000	-.007	-.010	-.010	-.008	-.006
11	.293	.140	.035	-.026	-.059	-.077	-.087	-.094	-.099	-.103	-.107
12	.281	.150	.068	.031	.020	.024	.034	.047	.060	.074	.087
13	.228	.090	.002	-.044	-.066	-.075	-.080	-.082	-.085	-.088	-.092
14	.219	.097	.026	-.003	-.009	-.004	.004	.014	.023	.032	.039
15	.192	.068	-.004	-.034	-.041	-.038	-.032	-.026	-.021	-.017	-.015

Key: The inflation series is $100(1-L)^d \Delta\log CPI_t = (1-L)^d y_t$. The series begins in January of 1957 and runs through December of 1990. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 4

Autocorrelations of Filtered CPI Inflation Series

Brazil

	Value of d:										
Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.886	.811	.706	.591	.479	.375	.197	.190	.107	.030	-.042
2	.789	.662	.512	.368	.248	.158	.091	.042	.007	-.016	-.030
3	.698	.528	.340	.170	.040	-.049	-.107	-.144	-.166	-.178	-.183
4	.645	.464	.274	.114	.005	-.059	-.091	-.104	-.104	-.098	-.088
5	.611	.429	.245	.099	.008	-.037	-.053	-.052	-.044	-.033	-.020
6	.582	.397	.216	.077	-.004	-.040	-.048	-.043	-.032	-.021	-.011
7	.557	.368	.186	.051	-.025	-.055	-.058	-.048	-.034	-.021	-.010
8	.531	.337	.152	.016	-.059	-.085	-.083	-.068	-.050	-.031	-.015
9	.510	.309	.115	-.030	-.113	-.144	-.146	-.132	-.113	-.093	-.073
10	.506	.305	.106	-.050	-.146	-.191	-.206	-.204	-.196	-.184	-.172
11	.531	.359	.185	.046	-.040	-.082	-.096	-.097	-.091	-.084	-.077
12	.549	.412	.275	.168	.103	.072	.062	.063	.067	.072	.076
13	.552	.438	.320	.223	.157	.117	.093	.076	.061	.046	.031
14	.550	.456	.362	.288	.241	.217	.206	.203	.203	.203	.205
15	.512	.405	.289	.187	.108	.051	.009	-.025	-.054	-.080	-.103

Key: The inflation series is $100(1-L)^d \Delta \log CPI_t = (1-L)^d y_t$. The series begins in January of 1957 and runs through January of 1991. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 5

Autocorrelations of Filtered CPI Inflation Series
Canada

	Value of d:										
Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.446	.299	.172	.037	-.092	-.200	-.280	-.338	-.381	-.416	-.445
2	.384	.259	.162	.066	-.018	-.076	-.106	-.116	-.113	-.103	-.090
3	.424	.323	.249	.176	.114	.073	.053	.047	.049	.054	.060
4	.395	.302	.230	.154	.086	.036	.006	-.009	-.016	-.020	-.022
5	.392	.308	.243	.177	.120	.081	.061	.054	.055	.059	.063
6	.320	.219	.140	.060	-.010	-.058	-.084	-.096	-.099	-.098	-.097
7	.355	.272	.205	.135	.072	.029	.004	-.007	-.011	-.011	-.010
8	.401	.337	.280	.219	.163	.122	.099	.087	.082	.080	.078
9	.360	.293	.229	.159	.094	.047	.019	.005	-.001	-.003	-.004
10	.323	.252	.183	.105	.035	-.015	-.044	-.056	-.060	-.059	-.056
11	.349	.283	.212	.132	.059	.003	-.033	-.053	-.065	-.073	-.079
12	.461	.435	.393	.341	.291	.254	.231	.217	.209	.204	.200
13	.351	.307	.248	.176	.109	.057	.024	.004	-.007	-.015	-.021
14	.265	.216	.153	.077	.008	-.042	-.071	-.084	-.089	-.089	-.087
15	.275	.243	.193	.129	.070	.029	.007	-.002	-.004	-.003	-.001

Key: Each series is $100(1-L)^d \Delta\log CPI_t = (1-L)^d y_t$. The series begins in January of 1948 and runs through August of 1990. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 6

Autocorrelations of Filtered CPI Inflation Series

France

Value of d:

Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.428	.380	.251	.139	.042	-.042	-.116	-.181	-.239	-.290	-.335
2	.169	.227	.115	.033	-.025	-.063	-.087	-.101	-.106	-.104	-.098
3	.120	.184	.095	.037	.005	-.010	-.014	-.011	-.003	.007	.018
4	.088	.105	.020	-.031	-.058	-.069	-.070	-.066	-.060	-.053	-.045
5	-.102	.073	-.008	-.055	-.018	-.093	-.097	-.097	-.095	-.094	-.092
6	-.077	.155	.097	.068	.057	.057	.062	.070	.077	.085	.091
7	-.078	.116	.051	.014	-.005	-.013	-.016	-.016	-.016	-.014	-.013
8	-.022	.092	.023	-.020	-.046	-.062	-.072	-.080	-.086	-.092	-.098
9	-.025	.197	.155	.138	.135	.140	.149	.161	.172	.185	.195
10	.057	.047	-.033	-.088	-.125	-.151	-.170	-.185	-.196	-.205	-.212
11	.155	.166	.115	.086	.071	.063	.059	.058	.057	.057	.057
12	.222	.215	.177	.157	.148	.143	.141	.139	.138	.136	.134
13	.193	.102	.049	.018	-.002	-.013	-.021	-.026	-.030	-.032	-.034
14	.113	.032	-.029	-.068	-.094	-.112	-.126	-.137	-.145	-.152	-.157
15	.140	.163	.136	.125	.121	.122	.124	.126	.129	.131	.133

Key: The inflation series is $100(1-L)^d \Delta\log CPI_t = (1-L)^d y_t$. The series begins in January of 1948 and runs through July of 1990. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 7

Autocorrelations of Filtered CPI Inflation Series

Germany

	Value of d:										
Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.362	.183	.076	-.015	-.092	-.158	-.216	-.266	-.311	-.350	-.386
2	.275	.044	-.027	-.073	-.101	-.117	-.123	-.122	-.117	-.109	-.097
3	.210	.027	-.016	-.035	-.039	-.035	-.025	-.013	-.001	.011	.021
4	.011	-.022	-.056	-.067	-.066	-.058	-.048	-.038	-.029	-.021	-.015
5	.063	-.039	-.067	-.074	-.070	-.061	-.051	-.041	-.032	-.025	-.019
6	-.026	-.050	-.079	-.087	-.084	-.076	-.066	-.057	-.047	-.039	-.032
7	-.024	-.004	-.036	-.050	-.056	-.058	-.059	-.060	-.061	-.064	-.067
8	.041	.154	.139	.139	.144	.152	.160	.167	.173	.178	.182
9	.000	.010	-.034	-.062	-.081	-.094	-.105	-.114	-.122	-.129	-.135
10	.044	.063	.026	.007	-.003	-.006	-.006	-.004	.000	.004	.010
11	.069	.108	.068	.042	.024	.010	-.001	-.010	-.019	-.026	-.033
12	.055	.231	.206	.191	.181	.172	.164	.156	.148	.140	.133
13	.070	.148	.122	.107	.097	.090	.083	.077	.070	.065	.059
14	-.016	-.022	-.054	-.069	-.077	-.082	-.084	-.084	-.084	-.083	-.082
15	-.095	-.047	-.069	-.076	-.077	-.075	-.072	-.068	-.065	-.062	-.060

Key: The inflation series is $100(1-L)^d \Delta\log CPI_t = (1-L)^d y_t$. The series begins in January of 1948 and runs through March of 1990. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 8

Autocorrelations of Filtered CPI Inflation Series

Israel

	Value of d:										
Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.736	.570	.374	.199	.059	-.046	-.126	-.129	-.240	-.282	-.319
2	.653	.444	.226	.050	-.070	-.144	-.186	-.108	-.216	-.216	-.211
3	.671	.491	.318	.190	.114	.077	.065	.068	.078	.091	.106
4	.635	.433	.238	.090	-.007	-.063	-.094	-.111	-.119	-.124	-.127
5	.658	.486	.323	.204	.130	.090	.069	.060	.056	.055	.055
6	.658	.492	.341	.232	.167	.132	.116	.109	.108	.108	.109
7	.602	.404	.219	.080	-.009	-.062	-.093	-.110	-.120	-.127	-.131
8	.614	.431	.267	.149	.077	.037	.017	.009	.006	.007	.009
9	.615	.446	.294	.186	.119	.080	.060	.049	.043	.040	.038
10	.600	.431	.279	.169	.100	.058	.035	.021	.014	.010	.008
11	.575	.407	.252	.136	.058	.008	-.027	-.051	-.071	-.087	-.100
12	.624	.483	.381	.318	.287	.274	.270	.270	-.273	.275	.278
13	.507	.316	.155	.043	-.025	-.065	-.089	-.103	-.113	-.121	-.128
14	.470	.266	.106	.001	-.055	-.080	-.088	-.087	-.082	-.076	-.069
15	.456	.269	.123	.035	-.008	-.022	-.022	-.015	-.006	.003	.012

<u>Key</u>: The inflation series is $100(1-L)^d \Delta\log CPI_t = (1-L)^d y$. The series begins in January of 1957 and runs through February of 1991. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 9

Autocorrelations of Filtered CPI Inflation Series

Italy

	Value of d:										
Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.253	.126	-.063	-.191	-.280	-.346	-.398	-.440	-.475	-.505	-.532
2	.280	.227	.104	.041	.012	.003	.005	.013	.023	.036	.050
3	.316	.205	.090	.033	.009	.001	.000	.002	.006	.010	.014
4	.240	.160	.042	-.014	-.039	-.048	-.050	-.049	-.047	-.044	-.042
5	.368	.183	.075	.026	.006	.001	.002	.005	.009	.014	.018
6	.204	.171	.059	.003	-.023	-.035	-.041	-.044	-.046	-.047	-.048
7	.310	.249	.153	.107	.087	.078	.073	.071	.069	.066	.064
8	.326	.212	.112	.065	.044	.035	.031	.029	.027	.026	.025
9	.194	.140	.030	-.021	-.043	-.051	-.054	-.054	-.053	-.052	-.051
10	.220	.147	.043	-.004	-.023	-.030	-.031	-.030	-.029	-.027	-.025
11	.194	.194	.101	.060	.045	.039	.038	.038	.039	.039	.040
12	.370	.172	.075	.031	.012	.004	-.001	-.003	-.005	-.006	-.008
13	.229	.165	.071	.029	.012	.005	.002	.001	.000	.000	-.001
14	.193	.162	.072	.035	.021	.018	.018	.020	.021	.023	.024
15	.246	.114	.017	-.025	-.041	-.046	-.047	-.046	-.045	-.043	-.042

Key: The inflation series is $100(1-L)^d \Delta \log CPI_t = (1-L)^d y_t$. The series begins in January of 1948 and runs through June of 1990. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 10

Autocorrelations of Filtered CPI Inflation Series

Japan

	Value of d:										
Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.121	-.008	-.097	-.169	-.228	-.279	-.322	-.361	-.395	-.425	-.452
2	.092	-.048	-.090	-.111	-.118	-.117	-.110	-.100	-.086	-.071	-.055
3	.180	.010	-.011	-.017	-.014	-.008	.001	.009	.017	.024	.030
4	.036	-.035	-.058	-.066	-.065	-.061	-.055	-.048	-.041	-.034	-.027
5	.098	-.045	-.077	-.095	-.105	-.112	-.116	-.119	-.122	-.124	-.127
6	.082	.190	.177	.173	.175	.178	.183	.187	.191	.195	.198
7	.025	.035	.004	-.016	-.029	-.039	-.046	-.053	-.059	-.065	-.070
8	.166	.041	.017	.005	-.001	-.003	-.004	-.004	-.004	-.003	-.003
9	.112	.065	.050	.047	.049	.053	.058	.063	.068	.072	.075
10	-.021	-.062	-.086	-.098	-.102	-.103	-.101	-.099	-.096	-.093	-.089
11	-.031	-.019	-.049	-.069	-.084	-.095	-.106	-.115	-.124	-.133	-.140
12	.151	.328	.334	.343	.352	.359	.364	.368	.370	.370	.370
13	-.076	-.108	-.138	-.157	-.170	-.181	-.190	-.197	-.204	-.210	-.215
14	-.135	-.086	-.097	-.097	-.092	-.085	-.077	-.068	-.059	-.051	-.042
15	-.083	.013	.010	.016	.024	.033	-.041	.047	.053	.057	.061

Key: The inflation series is $100(1-L)^d \Delta\log CPI_t = (1-L)^d y_t$. The series begins in January of 1948 and runs through July of 1990. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 11

Autocorrelations of Filtered CPI Inflation Series
United Kingdom

	Value of d:										
Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.267	.118	-.016	-.115	-.190	-.251	-.302	-.346	-.384	-.418	-.448
2	.232	.075	-.014	-.061	-.083	-.089	-.087	-.079	-.068	-.054	-.038
3	.197	.074	.001	-.032	-.044	-.044	-.038	-.029	-.020	-.011	-.002
4	.203	.063	-.009	-.043	-.057	-.060	-.059	-.055	-.051	-.047	-.043
5	.223	.118	.051	.017	.000	-.009	-.015	-.020	-.023	-.027	-.032
6	.313	.236	.192	.175	.171	.173	.176	.179	.181	.183	.185
7	.160	.028	-.043	-.078	-.096	-.105	-.110	-.114	-.116	-.118	-.120
8	.146	.036	-.027	-.054	-.064	-.065	-.063	-.059	-.054	-.050	-.045
9	.210	.115	.064	.044	.039	.040	.044	.049	.054	.058	.061
10	.173	.085	.022	-.008	-.023	-.027	-.028	-.027	-.025	-.022	-.019
11	.201	.077	-.001	-.047	-.078	-.101	-.119	-.135	-.149	-.161	-.172
12	.403	.391	.367	.361	.362	.365	-.369	.372	.375	.378	.380
13	.156	.030	-.047	-.092	-.122	-.144	-.162	-.177	-.190	-.202	-.213
14	.144	.063	.010	-.011	-.018	-.018	-.014	-.009	-.003	.003	.009
15	.168	.068	.023	.008	.066	.009	.014	.019	.024	.028	.032

Key: The inflation series is $100(1-L)^d \Delta\log CPI_t = (1-L)^d y_t$. The series begins in January of 1948 and runs through August of 1990. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 12

Autocorrelations of Filtered CPI Inflation Series
United States

	Value of d:										
Lag	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
1	.467	.313	.121	-.028	-.142	-.231	-.302	-.360	-.409	-.451	-.499
2	.423	.353	.220	.137	.090	.067	.058	.058	.064	.073	.112
3	.399	.225	.084	.000	-.045	-.067	-.076	-.078	-.076	-.073	-.086
4	.360	.221	.095	.028	-.003	-.015	-.015	-.011	-.006	.001	.032
5	.316	.172	.046	-.020	-.050	-.060	-.061	-.057	-.051	-.046	-.050
6	.305	.185	.069	.009	-.016	-.024	-.024	-.020	-.014	-.009	-.013
7	.312	.207	.094	.034	.005	-.009	-.014	-.016	-.016	-.016	-.024
8	.359	.257	.154	.097	.006	.049	.039	.032	.027	.023	.028
9	.386	.275	.179	.125	.096	.079	.069	.061	.056	.052	.042
10	.349	.235	.132	.073	.038	.017	.003	-.008	-.016	-.022	-.029
11	.320	.242	.156	.101	.076	.062	.054	.049	.045	.042	.024
12	.278	.213	.122	.076	.054	.043	.037	.034	.031	.030	.037
13	.234	.135	.035	-.016	-.038	-.048	-.050	-.050	-.048	-.045	-.034
14	.180	.103	.001	-.050	-.074	-.083	-.086	-.086	-.084	-.082	-.089
15	.211	.187	.110	.076	.063	.060	.060	.061	.062	.063	.073

Key: The inflation series is $100(1\text{-}L)^d\ \Delta\log CPI_t = (1\text{-}L)^d\ y_t$. The series begins in January of 1948 and runs through August of 1990. The first 30 observations were omitted before the autocorrelations were computed for each filtered series.

TABLE 13

Tests for Order of Integration of Different Countries' Inflation Series

Country	H_0: I(1) $Z(t_{\alpha^*})$	$Z(t_{\tilde{\alpha}})$	H_0: I(0) $\hat{\eta}_\mu$	$\hat{\eta}_\tau$
Argentina	-4.78**	-3.67*	2.56**	0.21*
Brazil	-3.48**	-2.50	1.44**	0.44**
Canada	-10.61**	-6.23**	1.13**	0.26**
France	-12.78**	-10.48**	0.20	0.22**
Germany	-15.11**	-13.12**	0.24	0.14
Italy	-15.19**	-9.06**	2.70**	0.46**
Israel	-4.51**	-3.48*	2.16**	0.36**
Japan	-18.38**	-15.76**	0.33	0.17
UK	-14.35**	-8.76**	0.88**	0.26**
US	-9.64**	-5.66**	1.80**	0.36**

<u>Key</u>: $Z(t_{\alpha^*})$ and $Z(t_{\tilde{\alpha}})$ are the Phillips Perron adjusted t statistics of the lagged dependent variable in a regression with intercept only, and intercept and time trend included respectively. The critical values for $Z(t_{\alpha^*})$ and $Z(t_{\tilde{\alpha}})$ are -2.86 and -3.41 at the .05 level of significance and -3.43 and -3.96 at the .01 level of significance.

$\hat{\eta}_\mu$ and $\hat{\eta}_\tau$ are the KPS test statistics and are based on residuals from regressions with an intercept, and intercept and time trend respectively. The .05 critical values for $\hat{\eta}_\mu$ and $\hat{\eta}_\tau$ are 0.463 and 0.146 respectively; the .01 critical values are 0.739 and 0.216 respectively.

All test statistics reported in this table are based on Newey and West (1987) adjustments using 8 lags. Two asterisks denote calculated test statistics which are significant at the .01 level; one asterisk corresponds to significance at the .05 level.

TABLE 14

TABLE 14

Application of the Geweke Porter-Hudak Method to Estimate d

# of spectral ordinates in regression	Argentina	Brazil	Canada	France	Germany	Israel	Italy	Japan	UK	US
$T^{.50}$	0.82 (0.26)	1.00 (0.22)	0.70 (0.14)	0.40 (0.20)	0.16 (0.20)	1.55 (0.18)	0.70 (0.16)	0.04 (0.13)	0.65 (0.17)	0.90 (0.20)
$T^{.525}$	0.80 (0.20)	0.90 (0.18)	0.70 (0.12)	0.40 (0.18)	0.05 (0.21)	1.33 (0.17)	0.65 (0.15)	0.22 (0.15)	0.60 (0.16)	1.05 (0.18)
$T^{.55}$	0.75 (0.18)	0.70 (0.18)	0.72 (0.11)	0.24 (0.16)	0.05 (0.17)	1.30 (0.15)	0.70 (0.13)	0.33 (0.13)	0.50 (0.14)	0.95 (0.16)
$T^{.575}$	0.70 (0.16)	0.57 (0.16)	0.73 (0.09)	0.15 (0.14)	0.06 (0.14)	1.25 (0.14)	0.65 (0.12)	0.36 (0.11)	0.60 (0.12)	0.85 (0.15)
$T^{.60}$	0.70 (0.13)	0.68 (0.15)	0.64 (0.09)	0.03 (0.13)	0.01 (0.13)	1.08 (0.14)	0.60 (0.11)	0.44 (0.37)	0.55 (0.13)	0.78 (0.13)
Lowest long-run cycle										
m=6	1.17 (0.73)	1.53 (0.29)	----	----	----	1.15 (0.24)	----	----	----	----
m=8		----	---- (0.34)	0.91 (0.25)	0.47 (0.56)	1.16	---- (0.21)	1.30 (0.30)	0.03 (0.42)	1.100.70 (0.42)

TABLE 15

Estimated ARFIMA (0, d, 1) ~ Student t-GARCH (1, 1) Models for US CPI Inflation

$$100(1-L)^{d}\ \Delta\ \log\ CPI_t = b + (1+\theta L)\epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim t(0,\ \sigma_t^2,\ \nu^{-1})$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

	Method I	Method II									
$\hat{d}$	.359	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
	(.063)										
$\hat{b}$	-.328	.171	.094	.053	.029	.015	.010	.006	.003	.002	.002
	(.030)	(.015)	(.013)	(.011)	(.010)	(.008)	(.007)	(.006)	(.004)	(.003)	(.002)
$\hat{\theta}$	.099	.174	.060	-.045	-.148	-.248	-.356	-.475	-.610	-.726	-.799
	(.047)	(.049)	(.047)	(.046)	(.045)	(.043)	(.042)	(.039)	(.034)	(.028)	(.026)
$\hat{\omega}$	.0060	.0046	.0051	.0052	.0053	.0052	.0050	.0050	.0049	.0057	.0040
	(.0034)	(.0024)	(.0027)	(.0029)	(.0031)	(.0032)	(.0029)	(.0039)	(.0029)	(.0034)	(.0032)
$\hat{\alpha}$	.130	.121	.112	.104	.100	.101	.090	.093	.099	.112	.092
	(.059)	(.042)	(.092)	(.041)	(.040)	(.042)	(.035)	(.037)	(.039)	(.046)	(.044)
$\hat{\beta}$	.808	.831	.828	.832	.835	.837	.846	.846	.843	.823	.853
	(.073)	(.056)	(.061)	(.064)	(.066)	(.067)	(.061)	(.062)	(.062)	(.069)	(.056)
$\hat{\nu}^{-1}$	.115	.087	.089	.093	.115	.147	.123	.122	.142	.167	.112
	(.000)	(.004)	(.003)	(.000)	(.024)	(.016)	(.000)	(.000)	(.013)	(.006)	(.001)
Q(10)		118.63	52.24	26.010	18.855	19.845	24.300	29.94	34.56	33.07	26.903
$Q^2(10)$		8.39	9.80	10.570	10.408	9.454	8.382	7.69	7.74	7.48	6.539
m_3		.175	.211	.214	.198	.182	.172	.176	.188	.194	.223
m_4		4.660	4.512	4.441	4.415	4.414	4.411	4.401	4.436	4.573	4.821
£		-91.63	-72.77	-63.588	-60.218	-60.574	-62.180	-64.57	-66.43	-66.75	-66.53

<u>Key</u>: The statistics Q(10) and $Q^2(10)$ are the Ljung-Box tests based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and £ is the maximized value of the log likelihood. The model was estimated with their first 30 observations omitted so that estimation uses these values for initialization.

TABLE 16

Estimated ARFIMA-GARCH Model for CPI Inflation: Argentina

$$100(1-L)^d \, \Delta \log CPI_t = (1 + \theta_1 L + \theta_{12} L^{12} + \theta_{13} L^{13})\epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

d	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
$\hat{\theta}_1$	.401	.303	.232	.161	.081	-.017	-.156	-.297	-.433	-.562
	(.058)	(.660)	(.068)	(.040)	(.060)	(.057)	(.052)	(.046)	(.040)	(.034)
$\hat{\theta}_{12}$	.307	.230	.205	.191	.177	.167	.153	.128	.110	.103
	(.040)	(.038)	(.036)	(.036)	(.035)	(.032)	(.030)	(.030)	(.030)	(.029)
$\hat{\theta}_{13}$	.080	-.008	-.040	-.066	-.089	-.094	-.080	-.071	-.064	-.062
	(.050)	(.050)	(.047)	(.045)	(.045)	(.043)	(.039)	(.038)	(.038)	(.037)
$\hat{\omega}$	2.507	2.799	3.403	3.730	3.673	3.665	3.824	4.127	4.399	4.591
	(.426)	(.457)	(.562)	(.596)	(.572)	(.558)	(.606)	(.646)	(.642)	(.631)
$\hat{\alpha}$	.629	.656	.714	.789	.804	.813	.822	.833	.835	.831
	(.098)	(.102)	(.113)	(.120)	(.111)	(.104)	(.099)	(.097)	(.096)	(.096)
$\hat{\beta}$	.399	.351	.256	.172	.158	.154	.141	.119	.101	.091
	(.042)	(.048)	(.064)	(.069)	(.063)	(.060)	(.063)	(.064)	(.060)	(.057)
Q(10)	48.88	35.76	25.54	18.34	15.44	15.20	17.16	20.09	23.12	25.44
$Q^2(10)$	14.83	16.73	17.94	15.58	14.65	14.49	13.84	12.34	11.07	10.31
m_3	0.501	0.695	0.738	0.693	0.612	0.511	0.431	0.384	0.357	0.356
m_4	4.321	4.421	4.278	4.107	4.099	4.190	4.337	4.539	4.743	4.874
£	-1083.2	-1051.8	-1031.9	-1019.3	-1014.1	-1013.9	-1016.4	-1019.5	-1022.3	-1024.1

Key: All models are estimated by approximate MLE having concentrated out d. Q(10) and $Q^2(10)$ are the Ljung-Box tests based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and £ is the maximized value of the log likelihood.

TABLE 17

Estimated ARFIMA-GARCH Model for CPI Inflation: Brazil

$$100(1-L)^d \Delta \log CPI_t = (1 + \theta_1 L + \theta_{12} L^{12} + \theta_{13} L^{13})\epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

d	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
$\hat{\theta}_1$	.467	.360	.254	.145	-.029	-.098	-.240	-.389	-.544	-.712
	(.049)	(.057)	(.059)	(.058)	(.055)	(.052)	(.047)	(.044)	(.039)	(.032)
$\hat{\theta}_{12}$	.312	.232	.178	.145	.129	.126	.127	.130	.131	.125
	(.041)	(.046)	(.047)	(.046)	(.044)	(.043)	(.042)	(.041)	(.039)	(.039)
$\hat{\theta}_{13}$	.381	.284	.219	.175	.146	.121	.091	.055	.008	-.048
	(.049)	(.050)	(.046)	(.040)	(.036)	(.033)	(.033)	(.036)	(.038)	(.032)
$\hat{\omega}$	.071	.078	.086	.098	.119	.135	.144	.149	.157	.164
	(.018)	(.018)	(.020)	(.022)	(.025)	(.028)	(.028)	(.028)	(.030)	(.034)
$\hat{\alpha}$	.282	.267	.275	.302	.354	.401	.424	.432	.435	.446
	(.055)	(.056)	(.059)	(.061)	(.068)	(.072)	(.070)	(.068)	(.069)	(.081)
$\hat{\beta}$	.741	.745	.732	.704	.656	.618	.599	.593	.585	.570
	(.036)	(.041)	(.045)	(.047)	(.050)	(.051)	(.048)	(.045)	(.045)	(.057)
Q(10)	20.58	14.43	7.83	4.36	5.58	9.87	15.62	21.46	26.88	32.61
$Q^2(10)$	8.14	6.76	7.62	9.26	10.63	10.74	10.30	10.13	10.67	12.35
m_3	0.504	0.568	0.569	0.551	0.529	0.511	0.494	0.485	0.493	0.525
m_4	4.398	4.208	4.026	3.960	3.929	3.928	3.919	3.883	3.788	3.644
£	-829.81	-781.36	-752.07	-735.90	-728.42	-726.46	-727.74	-730.32	-732.61	-732.91

Key: All models are estimated by approximate MLE having concentrated out d. Q(10) and $Q^2(10)$ are the Ljung-Box tests based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and £ is the maximized value of the log likelihood.

TABLE 18

Estimated ARFIMA-GARCH Model for CPI Inflation: Canada

$$100(1-L)^d \, \Delta \log CPI_t = (1 + \theta_1 L + \theta_{12} L^{12} + \theta_{13} L^{13})\epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

d	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
$\hat{\theta}_1$	.160	.014	-.120	-.249	-.368	-.475	-.627	-.737	-.836	-.916
	(.049)	(.048)	(.046)	(.046)	(.044)	(.040)	(.035)	(.029)	(.027)	(.023)
$\hat{\theta}_{12}$	.289	.246	.224	.208	.220	.225	.221	.220	.227	.273
	(.046)	(.047)	(.047)	(.045)	(.045)	(.047)	(.046)	(.038)	(.048)	(.042)
$\hat{\theta}_{13}$	.230	.146	.076	.023	-.033	-.078	-.124	-.144	-.189	-.218
	(.048)	(.049)	(.049)	(.048)	(.047)	(.047)	(.046)	(.039)	(.049)	(.044)
$\hat{\omega}$	.0038	.0036	.0049	.0091	.0211	.0204	.0148	.0059	-.0009	.0200
	(.0025)	(.0026)	(.0038)	(.0068)	(.0079)	(.0056)	(.0053)	(.0024)	(.0004)	(.0017)
$\hat{\alpha}$	.054	.048	.048	.058	.084	.094	.074	.092	.047	.078
	(.019)	(.018)	(.021)	(.028)	(.029)	(.022)	(.023)	(.021)	(.013)	(.021)
$\hat{\beta}$	.916	.920	.908	.859	.723	.720	.792	.841	.958	.712
	(.031)	(.034)	(.048)	(.081)	(.092)	(.065)	(.062)	(.038)	(.008)	(.013)
Q(10)	24.04	13.28	10.06	10.36	11.52	13.40	14.15	13.80	10.96	9.06
$Q^2(10)$	7.97	8.82	9.70	9.43	8.37	8.14	9.77	10.06	14.28	8.34
m_3	0.368	0.383	0.372	0.361	0.356	0.359	0.329	0.399	0.254	0.365
m_4	3.899	3.915	3.932	3.942	3.935	3.937	3.870	3.999	3.666	3.929
$\pounds$	-197.05	-171.43	-160.17	-156.03	-153.58	-154.65	-154.12	-155.57	-158.21	-171.55

Key: All models are estimated by approximate MLE having concentrated out d. Q(10) and $Q^2(10)$ are the Ljung-Box tests based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and $\pounds$ is the maximized value of the log likelihood.

TABLE 19

Estimated ARFIMA-GARCH Model for CPI Inflation: France

$$100(1-L)^d \ \Delta \log CPI_t = (1 + \theta_1 L + \theta_{12} L^{12} + \theta_{13} L^{13})\epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim t(0, \sigma_t^2, \nu^{-1})$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

d	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
$\hat{\theta}_1$	.338 (.043)	.194 (.043)	.061 (.042)	-.062 (.042)	-.177 (.041)	-.288 (.040)	-.401 (.038)	-.518 (.035)	-.645 (.032)	-.766 (.030)
$\hat{\theta}_{12}$	.247 (.045)	.210 (.037)	.194 (.037)	.187 (.037)	.182 (.038)	.179 (.038)	.176 (.039)	.173 (.039)	.167 (.039)	.151 (.040)
$\hat{\theta}_{13}$	.293 (.041)	.194 (.042)	.103 (.041)	.036 (.041)	-.012 (.040)	-.049 (.040)	-.080 (.040)	-.110 (.040)	-.141 (.040)	-.157 (.041)
$\hat{\omega}$	.0012 (.0010)	.0006 (.0008)	.0004 (.0005)	.0004 (.0005)	.0004 (.0005)	.0004 (.0005)	.0004 (.0005)	.0005 (.0005)	.0005 (.0006)	.0010 (.0008)
$\hat{\alpha}$	.136 (.038)	.071 (.023)	.046 (.016)	.042 (.015)	.041 (.016)	.042 (.015)	.043 (.015)	.045 (.016)	.048 (.016)	.078 (.024)
$\hat{\beta}$	.863 (.033)	.962 (.022)	.948 (.015)	.951 (.014)	.951 (.014)	.951 (.014)	.949 (.015)	.947 (.015)	.945 (.016)	.912 (.024)
$\hat{\nu}^{-1}$	.172 (.016)	.196 (.000)	.215 (.085)	.211 (.017)	.207 (.017)	.204 (.015)	.201 (.003)	.198 (.001)	.198 (.019)	.030 (.000)
Q(10)	42.10	34.64	27.58	22.91	22.62	24.66	27.33	29.51	30.72	27.39
$Q^2(10)$	2.15	3.35	4.82	5.43	5.96	6.53	7.13	7.76	8.39	6.16
m_3	0.640	0.876	1.022	1.082	1.093	1.077	1.045	0.998	0.950	0.916
m_4	7.524	8.384	8.905	9.149	9.218	9.133	8.913	8.551	8.117	7.451
£	-261.35	-220.82	-202.92	-196.77	-196.20	-198.08	-200.76	-203.18	-204.46	-205.53

Key: All models are estimated by approximate MLE having concentrated out d. Q(10) and $Q^2(10)$ are the Ljung-Box tests based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and £ is the maximized value of the log likelihood.

TABLE 20

Estimated ARFIMA-GARCH Model for CPI Inflation: Germany

$$100(1-L)^{d}\ \Delta \log CPI_t = (1 + \theta_1 L + \theta_{12}L^{12} + \theta_{13}L^{13})\epsilon_t$$

$$\epsilon_t|\Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

d	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
$\hat{\theta}_1$	.214	.100	-.004	-.104	-.207	-.315	-.444	-.625	-.846	-.895
	(.047)	(.047)	(.045)	(.045)	(.045)	(.044)	(.041)	(.034)	(.024)	(.021)
$\hat{\theta}_{12}$	.228	.201	.191	.190	.189	.193	.202	.215	.232	.234
	(.042)	(.041)	(.041)	(.041)	(.041)	(.041)	(.041)	(.040)	(.038)	(.040)
$\hat{\theta}_{13}$	.208	.164	.134	.108	.082	.052	.014	-.058	-.176	-.186
	(.046)	(.046)	(.047)	(.048)	(.048)	(.048)	(.047)	(.045)	(.041)	(.041)
$\hat{\omega}$	.0004	.0003	.0002	.0002	.0002	.0002	.0003	.0002	.0003	.0040
	(.0004)	(.0003)	(.0003)	(.0004)	(.0004)	(.0004)	(.0004)	(.0004)	(.0004)	(.0004)
$\hat{\alpha}$	.030	.028	.028	.028	.028	.029	.029	.028	.028	.031
	(.007)	(.006)	(.006)	(.006)	(.007)	(.007)	(.007)	(.007)	(.007)	(.008)
$\hat{\beta}$	.963	.966	.967	.967	.966	.966	.966	.967	.967	.962
	(.009)	(.008)	(.008)	(.008)	(.008)	(.008)	(.009)	(.008)	(.008)	(.010)
Q(10)	17.22	16.95	21.95	28.52	34.44	42.34	50.93	65.47	64.09	43.99
$Q^2(10)$	8.55	10.51	12.28	13.87	15.45	16.72	17.76	17.29	14.32	13.93
m_3	0.553	0.587	0.584	0.567	0.545	0.521	0.494	0.450	0.344	0.289
m_4	4.875	4.764	4.653	4.547	4.446	4.351	4.266	4.199	4.196	4.185
£	-170.90	-158.07	-155.10	-157.32	-161.01	-165.22	-169.02	-171.07	-167.15	-163.34

Key: All models are estimated by approximate MLE having concentrated out d. Q(10) and $Q^2(10)$ are the Ljung Box tests based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and £ is the maximized value of the log likelihood.

TABLE 21

Estimated ARFIMA-GARCH Model for CPI Inflation: Israel

$$100(1-L)^{d} \Delta \log CPI_t = (1 + \theta_1 L + \theta_{12} L^{12} + \theta_{13} L^{13})\epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim t(0, \sigma_t^2, \nu^{-1})$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

d	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
$\hat{\theta}_1$	.317 (.056)	.177 (.054)	.083 (.056)	-.067 (.052)	-.201 (.050)	-.353 (.052)	-.475 (.049)	-.613 (.040)	-.732 (.036)	-.762 (.038)
$\hat{\theta}_{12}$	.221 (.039)	.226 (.040)	.212 (.035)	.223 (.039)	.208 (.039)	.235 (.039)	.220 (.040)	.203 (.038)	.220 (.038)	.187 (.039)
$\hat{\theta}_{13}$	.108 (.039)	.075 (.042)	.040 (.034)	.001 (.043)	-.033 (.042)	-.076 (.043)	-.115 (.042)	-.136 (.040)	-.176 (.038)	-.143 (.041)
$\hat{\omega}$	.161 (.069)	.105 (.049)	.403 (.117)	.104 (.050)	.133 (.058)	.141 (.067)	.160 (.084)	.160 (.067)	.153 (.060)	.341 (.133)
$\hat{\alpha}$	.306 (.074)	.174 (.049)	.371 (.091)	.141 (.044)	.138 (.044)	.184 (.062)	.209 (.078)	.126 (.042)	.176 (.055)	.612 (.166)
$\hat{\beta}$	.671 (.054)	.783 (.044)	.490 (.072)	.807 (.045)	.793 (.050)	.774 (.055)	.772 (.060)	.787 (.055)	.750 (.060)	.509 (.076)
$\hat{\upsilon}^{-1}$	.030 (.000)	.132 (.000)	.100 (.000)	.136 (.000)	.115 (.000)	.170 (.000)	.240 (.010)	.128 (.000)	.090 (.000)	.084 (.000)
Q(10)	38.44	37.37	36.61	22.14	21.32	22.09	23.20	23.43	20.80	23.10
$Q^2(10)$	8.22	10.43	13.66	11.49	10.26	8.95	8.04	7.27	7.10	5.97
m_3	1.263	1.375	1.695	1.506	1.524	1.461	1.432	1.427	1.414	1.583
m_4	7.706	7.593	8.890	8.135	8.367	9.253	8.222	8.542	8.352	9.677
£	-777.49	-759.30	-759.25	-746.15	-744.53	-740.80	-737.11	-740.51	-742.01	-745.27

Key: All models are estimated by approximate MLE having concentrated out d. Q(10) and $Q^2(10)$ are the Ljung-Box tests based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and £ is the maximized value of the log likelihood.

TABLE 22

Estimated ARFIMA-GARCH Model for CPI Inflation: Italy

$$100(1-L)^d \, \Delta \log CPI_t = (1 + \theta_1 L + \theta_{12} L^{12} + \theta_{13} L^{13})\epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

d	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
$\hat{\theta}_1$	.156	.019	-.108	-.226	-.343	-.458	-.577	-.694	-.789	-.868
	(.046)	(.046)	(.045)	(.045)	(.043)	(.041)	(.038)	(.034)	(.030)	(.024)
$\hat{\theta}_{12}$	.187	.129	.099	.092	.075	.070	.065	.057	.073	.080
	(.047)	(.047)	(.047)	(.047)	(.047)	(.047)	(.047)	(.048)	(.045)	(.039)
$\hat{\theta}_{13}$	.223	.144	.090	.057	.027	.002	-.024	-.050	-.072	-.080
	(.042)	(.043)	(.044)	(.045)	(.046)	(.046)	(.046)	(.047)	(.044)	(.038)
$\hat{\omega}$	.003	.002	.002	.002	.002	.002	.002	.002	.004	.007
	(.002)	(.001)	(.001)	(.002)	(.002)	(.002)	(.002)	(.002)	(.003)	(.003)
$\hat{\alpha}$	.112	.108	.110	.117	.114	.114	.115	.115	.130	.168
	(.018)	(.016)	(.016)	(.018)	(.018)	(.018)	(.018)	(.018)	(.023)	(.023)
$\hat{\beta}$	.886	.893	.891	.882	.886	.885	.885	.884	.864	.819
	(.015)	(.013)	(.013)	(.015)	(.015)	(.015)	(.015)	(.015)	(.019)	(.022)
Q(10)	22.94	13.84	11.37	11.51	12.53	13.79	14.12	12.86	10.66	8.96
$Q^2(10)$	7.67	8.35	8.15	7.89	6.99	6.35	5.88	5.61	6.36	9.41
m_3	0.536	0.665	0.726	0.756	0.759	0.759	0.762	0.772	0.792	0.829
m_4	3.779	3.864	3.944	4.025	4.059	4.091	4.127	4.161	4.251	4.409
$\mathcal{L}$	-426.0	-392.3	-378.9	-374.9	-374.4	-375.3	-376.2	-376.2	-376.0	-377.4

Key: All models are estimated by approximate MLE having concentrated out d. Q(10) and $Q^2(10)$ are the Ljung-Box statistics based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and $\mathcal{L}$ is the maximized value of the log likelihood.

TABLE 23

Estimated ARFIMA-GARCH Model for CPI Inflation: Japan

$$100(1-L)^d \Delta \log CPI_t = (1 + \theta_1 L + \theta_{12} L^{12} + \theta_{13} L^{13})\epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

d	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
$\hat{\theta}_1$	-.019	-.181	-.348	-.499	-.617	-.709	-.788	-.856	-.901	-.927
	(.403)	(.045)	(.042)	(.038)	(.034)	(.031)	(.028)	(.025)	(.022)	(.021)
$\hat{\theta}_{12}$	.321	.298	.284	.267	.256	.248	.215	.259	.274	.294
	(.043)	(.044)	(.044)	(.044)	(.044)	(.045)	(.045)	(.045)	(.044)	(.043)
$\hat{\theta}_{13}$	-.045	-.142	-.225	-.278	-.301	-.305	-.301	-.291	-.287	-.289
	(.455)	(.045)	(.043)	(.041)	(.041)	(.042)	(.044)	(.044)	(.044)	(.043)
$\hat{\omega}$	.0100	.0106	.0132	.0151	.0145	.0120	.0086	.0065	.0057	.0063
	(.0042)	(.0050)	(.0062)	(.0070)	(.0069)	(.0420)	(.0046)	(.0036)	(.0032)	(.0032)
$\hat{\alpha}$	.080	.082	.090	.099	.098	.087	.073	.062	.057	.060
	(.016)	(.016)	(.017)	(.017)	(.016)	(.014)	(.013)	(.012)	(.013)	(.013)
$\hat{\beta}$	.908	.905	.893	.883	.885	.898	.914	.927	.933	.930
	(.176)	(.018)	(.019)	(.021)	(.020)	(.017)	(.015)	(.014)	(.015)	(.015)
Q(10)	17.70	22.18	27.07	29.52	29.68	28.76	26.95	24.48	22.47	21.03
$Q^2(10)$	9.41	10.88	11.77	12.94	14.24	15.33	16.01	14.76	11.25	8.61
m_3	0.413	0.443	0.461	0.463	0.452	0.431	0.399	0.346	0.269	0.196
m_4	3.973	4.009	4.015	4.035	4.096	4.211	4.360	4.525	4.657	4.662
£	-613.4	-610.1	-610.2	-610.4	-610.2	-609.6	-608.7	-608.0	-609.1	-614.0

Key: All models are estimated by approximate MLE having concentrated out d. Q(10) and $Q^2(10)$ are the Ljung-Box statistics based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and £ is the maximized value of the log likelihood.

TABLE 24

Estimated ARFIMA-GARCH Model for CPI Inflation: United Kingdom

$$100(1-L)^d \Delta \log CPI_t = (1 + \theta_1 L + \theta_{12} L^{12} + \theta_{13} L^{13})\epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

d	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00
$\hat{\theta}_1$	.177 (.059)	.053 (.061)	-.055 (.062)	-.164 (.060)	-.282 (.058)	-.412 (.055)	-.556 (.050)	-.673 (.044)	-.762 (.039)	-.866 (.030)
$\hat{\theta}_{12}$	.336 (.037)	.304 (.037)	.287 (.037)	.278 (.037)	.271 (.037)	.265 (.037)	.254 (.037)	.239 (.037)	.229 (.038)	.227 (.037)
$\hat{\theta}_{13}$	.017 (.045)	-.068 (.047)	-.126 (.048)	-.172 (.049)	-.211 (.049)	-.251 (.049)	-.292 (.048)	-.310 (.045)	-.306 (.044)	-.275 (.042)
$\hat{\omega}$	.093 (.026)	.104 (.028)	.118 (.034)	.135 (.041)	.137 (.044)	.144 (.047)	.142 (.047)	.141 (.047)	.154 (.050)	.107 (.032)
$\hat{\alpha}$	.232 (.044)	.226 (.044)	.207 (.046)	.204 (.050)	.196 (.052)	.200 (.054)	.189 (.052)	.182 (.050)	.192 (.053)	.104 (.041)
$\hat{\beta}$	.590 (.075)	.556 (.081)	.529 (.097)	.486 (.119)	.491 (.128)	.474 (.135)	.488 (.135)	.494 (.135)	.456 (.143)	.610 (.103)
Q(10)	24.29	20.03	20.70	24.16	28.20	31.08	31.46	30.00	27.68	22.14
$Q^2(10)$	18.79	19.93	18.76	17.15	14.51	13.33	12.37	11.87	12.16	11.14
m_3	0.682	0.815	0.870	0.876	0.865	0.855	0.864	0.886	0.899	0.954
m_4	4.186	4.506	4.727	4.807	4.850	4.886	5.001	5.082	5.093	5.398
£	-481.76	-466.34	-461.74	-462.21	-453.64	-465.60	-466.66	-466.62	-466.08	-465.03

Key: All models are estimated by approximate MLE having concentrated out d. Q(10) and $Q^2(10)$ are the Ljung-Box tests based on the residuals and squared residuals, m_3 and m_4 are the sample skewness and kurtosis statistics based on the standardized residuals, and £ is the maximized value of the log likelihood.

TABLE 25

Likelihood Ratio Tests of Relationship Between
Mean and Variability of Inflation, y_t

$$(1-L)^d \; y_t = (1+\theta_1 L+\theta_{12}L^{12}+\theta_{13}L^{13})\epsilon_t+\delta\sigma_t$$

$$\epsilon_t|\Omega_{t-1} \sim D(0,\ \sigma_t^2,\ \nu^{-1})$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma y_{t-1}$$

LR Tests	Argentina	Brazil	France	Germany	Israel	Italy	Japan	UK	US
$\delta=0$	8.90*	4.26*	0.00	2.40	5.46*	0.72	1.44	3.88*	0.40
$\gamma=0$	12.92**	7.14**	1.54	0.80	10.66**	1.12	0.44	9.12**	1.28

Key: y_t is 100 Δlog CPI_t, the conditional density D is student t for France and Israel and is Normal otherwise. Under the null hypothesis all test statistics are distributed as asymptotic χ_1^2, random variables. Two asterisks denotes significance at the .01 level and one asterisk denotes significance at the .05 level.

REFERENCES

Alesina, A. (1989), "Politics and Business Cycles in Industrial Democracies", Economic Policy, 8, 57-98.

Baillie, R.T. (1989), "Commodity Prices and Aggregate Inflation: Would a Commodity Price Rule be Worthwhile?", Carnegie Rochester Conference Series on Public Policy, 31, 185-240.

Bairam, E. (1988), "The Variability of Inflation", Economics Letters, 28, 327-329.

Ball, L. and S.G. Cecchetti (1990), "Inflation and Uncertainty at Short and Long Horizons", Brookings Papers on Economic Activity, 215-254.

Barsky, R.B. (1987), "The Fisher Hypothesis and the Forecastability and Persistence of Inflation", Journal of Monetary Economics, 19, 3-24.

Bernake, B. (1992), "Central Bank Behavior and the Strategy of Monetary Policy: Observations from Six Industrialized Countries", mimeo.

Berndt, E.K., Hall, B.H., Hall, R.E. and J.A. Hausman (1974), "Estimation and Inference in Nonlinear Structural Models", Annals of Economic and Social Measurement, 3/4, 653-665.

Cosimano, T.F. and D.W. Jansen (1988), "Estimates of the Variance of US Inflation Based Upon the ARCH Model", Journal of Money, Credit and Banking, 20, 409-421.

Cukierman, A. and P. Wachtell (1979), "Differential Inflationary Expectations and the Variability of the Rate of Inflation: Theories and Evidence", American Economic Review, 69, 595-609.

Cumby, R.E. and J. Huizinga (1988), "Testing the Auto-Correlation Structure in Disturbances in Ordinary Least Squares and Instrumental Variables Regression", University of Chicago, mimeo.

Demetriades, P. (1988), "Macroeconomic Aspects of the Correlation Between the Level and Variability of Inflation", Economic Letters,26, 121-124.

Devereux, M. (1989), "A Positive Theory of Inflation and Inflation Variance", Economic Inquiry, 27, 105-116.

Dickey, D.A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", Journal of the American Statistical Association, 74, 427-431.

Dickey, D.A. and W.A. Fuller (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", Econometrica, 49, 1057-1072.

Engle, R.F. (1983), "Estimates of the Variance of US Inflation Based Upon the ARCH Model", Journal of Money, Credit and Banking, 15, 286-301.

Fama, E.F. (1975), "Short Term Interest Rates as Predictors of Inflation", American Economic Review, 65, 269-282.

Fama, E.F. and M.R. Gibbons (1982), "Inflation Real Returns and Capital Investment", Journal of Monetary Economics, 9, 297-323.

Fischer, S. (1981), "Relative Shocks, Relative Price Variability, and Inflation", Brookings Papers on Economic Activity, 2, 381-431.

Foster, E. (1978), "The Variability of Inflation", Review of Economics and Statistics, 60, 346-350.

Friedman, M. (1977), "Nobel Lecture: Inflation and Unemployment", Journal of Political Economy, 85, 451-472.

Geweke, J. and S. Porter-Hudak (1983), "The Estimation and Application of Long Memory Time Series Models", Journal of Time Series Analysis, 4, 221-238.

Goodfriend, M. (1987), "Interest Rate Smoothing and Price Level Trend-Stationarity", Journal of Monetary Economics, 19, 335-348.

Gordon, R.J. (1971), "Steady Anticipated Inflation: Mirage or Oasis?", Brookings Papers on Economic Activity, 2, 499-510.

Granger, C.W.J. and R. Joyeux (1980), "An Introduction to Long Memory Time Series Models and Fractional Differencing", Journal of Time Series Analysis, 1, 15-39.

Gray, J. (1976), "Wage Indexation: A Macroeconomic Approach", Journal of Monetary Economics, 2, 221-235.

Hosking, J.R.M. (1981), "Fractional Differencing", Biometrika, 68, 165-176.

Klein, B. (1976), "The Social Costs of the Recent Inflation: The Mirage of Steady 'Anticipated' Inflation", Carnegie-Rochester Conference Series on Public Policy, 3, 185-212.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Are Non Stationary?", Journal of Econometrics, forthcoming.

Ljung, G.M. and E.P. Box (1978), "On a Measure of Lack of Fit in Time Series Models", Biometrika, 65, 297-303.

Logue, D.E. and T.D. Willett (1976), "A Note on the Relation Between the Rate and Variability of Inflation", Economica, 43, 151-158.

Lucas, R.E. (1973), "Some International Evidence on Output-Inflation Trade-Offs", American Economic Review, 63, 326-334.

Mankiw G. and J. Miron (1986), "The Changing Behavior of the Term Structure of Interest Rates", Quarterly Journal of Economics, 101, 211-228.

McCallum, B.T. (1988), "Robustness Properties of a Rule for Monetary Policy", Carnegie Rochester Conference Series on Public Policy, 29, 173-203.

Nelson, C.R. and G.W. Schwert (1977), "Short-Term Interest Rates as Predictors of Inflation: On Testing the Hypothesis that the Real Rate of Interest is Constant", American Economic Review, 67, 478-486.

Newey, W.K., and K.D. West (1987), "A Simple, Positive, Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", Econometrica, 55, 703-708.

Okun, A.M. (1971), "The Mirage of Steady Inflation", Brookings Papers on Economic Activity, 2, 485-498.

Pagan, A.R., Hall, A.D., and P.K. Trivedi (1983), "Assessing the Variability of Inflation", Review of Economic Studies, 50, 585-596.

Parks, R.W. (1978), "Inflation and Relative Price Variability", Journal of Political Economy, 86, 79-96.

Phillips, P.C.B. and P. Perron (1988), "Testing for a Unit Root in Time Series Regression", Biometrika, 75, 335-346.

Rose, A. (1988), "Is the Real Interest Rate Stable?", Journal of Finance, 43, 1095-1112.

Sowell, F.B. (1990), "Modeling Long Run Behavior with the Fractional ARIMA Model", GSIA Carnegie Mellon University working paper.

Sweeney, R.J. (1987), "Some Macro Implications of Risk", Journal of Money, Credit and Banking, 19, 222-234.

Taylor, J.B. (1981), "On the Relation Between the Variability of Inflation and the Average Inflation Rate", Carnegie-Rochester Conference Series on Public Policy, 15, 57-85.

IV. A GENERALIZED METHOD OF MOMENTS ESTIMATOR FOR LONG-MEMORY PROCESSES

CHAPTER IV

A GENERALIZED METHOD OF MOMENTS ESTIMATOR FOR LONG-MEMORY PROCESSES

1. INTRODUCTION

As discussed in Chapters II and III, there currently exist two general types of estimation procedures which have been used to estimate the degree of persistence of a fractionally integrated, long-memory process. The frequency-domain based, two-step estimation procedure of Geweke and Porter-Hudak (1983) utilizes the spectrum of the series in estimating the parameter of fractional integration. Alternatively, the maximum likelihood estimation procedures of Hosking (1984b), Fox and Taqqu (1986), and Sowell (1992), which have been based in both the time and frequency domains, utilize standard first-order conditions in maximizing the log of the likelihood function of the fractionally integrated process. The generalized method of moments (GMM) estimation technique is an attractive alternative framework in which to estimate the parameter of fractional integration of a long-memory process since it does not require the distributional assumptions necessary under maximum likelihood estimation techniques and consequently offers the advantage of robustness in parameter estimation. In addition, as evidenced in Chapter III, approximate MLE can often involve numerically cumbersome techniques which may be avoided, in some part, with the technique of generalized methods of moments. For the fractionally integrated process, the GMM estimation technique exploits the set of moment conditions that equate the expected value of the sample autocorrelations to the corresponding population autocorrelations, evaluated at the true parameter values. In this way a consistent estimate of the parameters can be obtained.

This chapter provides the derivation of a generalized method of moments estimator for the degree of persistence, d, of a long-memory time series and examines its efficiency relative to those estimation procedures discussed in Chapters II and III. The following section discusses the GMM estimation technique in the context of the fractionally integrated model and presents the motivation for the use of GMM in this context. Section 3 presents the derivation of the asymptotic distribution of the estimated autocorrelations under specified assumptions. This section also presents the derivation of the asymptotic variance of the GMM estimator. Section 4 provides an investigation of the estimation procedure by examining the asymptotic efficiency of the estimator for a range of values of the parameter d using various moment conditions as well as subsets of moment conditions. The chapter ends with a brief summary and concluding section.

2. GMM ESTIMATION IN THE CONTEXT OF THE FRACTIONALLY INTEGRATED MODEL

The estimation technique of generalized method of moments makes use of a set of orthogonality conditions that are implied by the model to be estimated such that the expected value of the orthogonality condition is equated to zero at the true parameter value. For the case of the fractionally integrated model, consider the non-zero mean, stationary time series $\{y_t\}$ expressed in ARFIMA(p,d,q) form as introduced in Chapter II as

$$(1 - L)^d (y_t - \mu) = \theta(L)/\phi(L)\ \epsilon_t = u_t \tag{1.}$$

where for $-\frac{1}{2} < d < \frac{1}{2}$, y_t is said to be fractionally integrated of order d, the polynomials $\theta(L)$ and $\phi(L)$ are as defined in Chapter II, and u_t is a stationary and invertible error process. Following the framework of Hansen's (1982) GMM estimator, estimation of a $(p \times 1)$ parameter vector λ via the GMM estimation technique involves the use of m orthogonality restrictions where m is at least as great as p. Defining the $(m \times 1)$ vector of orthogonality conditions as some function $g(y_t, \lambda)$, the GMM estimator of λ is given as that value of the parameter vector which satisfies

$$(2.)\qquad \min_{\lambda} \ \bar{g}(y,\lambda)' \ W \ \bar{g}(y,\lambda),$$

where $\bar{g}(y,\lambda)$ is the standard expression[1] for the orthogonality condition of the GMM estimator written in the form of an average as

$$\bar{g}(y,\lambda) = 1/T \sum_{t=1}^{T} g(y_t,\lambda).$$

In this formulation, W is an $(m \times m)$ positive definite, symmetric weighting matrix defined as that matrix which has the characteristic of minimizing the sample orthogonality conditions. The minimized value of the criterion function (2.) will be asymptotically distributed as Chi-square with (m - p) degrees of freedom. Within the context of the GMM estimation procedure, the expression $\bar{g}(y,\lambda)$ should converge to zero for

[1] In the context of the fractionally integrated process expressed in (1.), use of an orthogonality condition of the form $\bar{g}(y,\lambda)$ is not directly applicable due to the difficulty in expressing the orthogonality conditions in the form of an average. This problem arises because the typical orthogonality condition of the fractionally integrated process is a function of an infinite number of terms. Therefore, as an alternative, the function $g(\cdot)$ is expressed in the form of a moment condition for the fractionally integrated process, as discussed later in this section.

the true parameter vector and not for any other element of the parameter space. Additionally, the optimal weighting matrix, W, is given as

$$W = [\text{cov } g(y_t,\lambda)]^{-1}.$$

Under weak regularity conditions, Hansen (1982) shows that the GMM estimator of the parameter vector λ satisfies

$$\sqrt{T}\ (\hat{\lambda}_{GMM} - \lambda) \sim N(0,[D'C^{-1}D]^{-1})$$

where C^{-1} is the optimal weighting matrix. In this representation, D is defined as the (m $\times$ p) matrix of partial derivatives of the moment conditions with respect to the parameter vector; that is,

$$D = \frac{\partial\ g(y_t,\lambda)}{\partial\ \lambda'}.$$

The GMM estimation procedure may be applied to many standard econometric models, each of which exploits its own unique set of moment conditions and asymptotically optimal weighting matrix. For the case of the fractionally integrated model, the moment conditions exploited make use of the theoretical and estimated autocorrelation functions of the model. Consider, for simplicity, the zero-mean ARFIMA(0,d,0) process

$$(3.)\qquad (1 - L)^d\ y_t = u_t$$

where u_t is a stationary error process, $d \in (-½,½)$, and $\rho_j = \text{corr}(y_t,y_{t-j})$ is defined as the j^{th} autocorrelation function of the process. The simple model expressed in (3.) is a single parameter model such that λ consists of a single element, d.[2] Recall that for the model given by (3.), ρ_j may be expressed simply as a function of d, as given earlier in Chapter II, as

$$\rho_j = \frac{\Gamma(1-d)\Gamma(j+d)}{\Gamma(d)\Gamma(j+d-1)} = \prod_{i=1}^{j} \frac{(d+i-1)}{(i-d)} .$$

The moment condition exploited by the fractionally integrated model, considering the first k moments, may be expressed as $E[\hat{\rho} - \rho(d)] = 0$ where

$$\hat{\rho} = [\hat{\rho}_1, \ . \ . \ . \ , \hat{\rho}_k]' \text{ and}$$

$$\rho(d) = [\rho_1(d), \ . \ . \ . \ , \rho_k(d)]' .$$

Within the context of the fractionally integrated model, the GMM estimator of the parameter vector λ may be expressed as that value of d which satisfies

$$S(d) = \min_{d} \ [\hat{\rho} - \rho(d)]' \ W \ [\hat{\rho} - \rho(d)] \tag{4.}$$

and the asymptotically optimal weighting matrix, W, is given as

$$W = [\mathrm{cov}\{\hat{\rho} - \rho(d)\}]^{-1}.$$

In considering the efficiency of the GMM estimator, it should be the

---

[2] This estimation procedure may be applied to the more general, multi-parameter ARFIMA representation given by (1.) in which case λ would be a vector and would include the parameters of the autoregressive and moving average polynomials.

case that any estimator based on all available moment conditions should be relatively more efficient than that based on only a subset of these moment conditions. However, in the case of the fractionally integrated process there will be some advantage to considering the GMM estimator based on a subset of moment conditions, especially in the case where a stationary ARMA component exists in the series. In such a model, the autocorrelation functions for the lower-order moments of the process will be a function of the autoregressive and moving average parameters of the model as well the parameter d. As such, the autocorrelation functions for the lower-order moments will be quite different from those autocorrelations that exist at higher-order moments, which are simply a function of the parameter d. In this sense the autocorrelation functions for the lower-order moments may be thought of as being "contaminated" when a stationary ARMA component exists in the series. As a result, it would be of interest in this context to determine whether the efficiency of the GMM estimator is maintained when using some subset of the moment conditions, say moments (r + 1) through {(r + 1) + k}, such that the first r moments may be discarded. Simple asymptotic variance calculations may be employed to determine these relative efficiencies, and these operations are discussed further in section 4.

3. ASYMPTOTIC DISTRIBUTION THEORY

In order to determine the asymptotic distribution and optimal weighting matrix of the generalized method of moments estimator for the fractionally integrated process, it is necessary to derive the asymptotic distribution of the moment condition, $[\hat{\rho} - \rho(d)]$. Recall that $\hat{d}_{GMM}$ solves the operation $\partial S(d)/\partial \hat{d} = 0$ as given in equation (4.). This expression may be written in the form of its Taylor-series expansion as

$$\frac{\partial S(d)}{\partial \hat{d}} = \frac{\partial S(d)}{\partial d} + \frac{\partial^2 S(d)}{\partial d_*^2} (\hat{d} - d),$$

where d_* lies between d and $\hat{d}$. Equating the above expansion to zero and solving for $(\hat{d} - d)$ gives

$$(\hat{d} - d) = - \left[\frac{\partial^2 S(d)}{\partial d_*^2}\right]^{-1} \frac{\partial S(d)}{\partial d}$$

where $\quad \partial S(d)/\partial d = -2\ D' W[\hat{\rho} - \rho(d)],$

$$\partial^2 S(d)/\partial d^2 = 2\ D' W D + o_p(1),$$

and D is as defined previously. It follows that the asymptotic distribution of the GMM estimator of d satisfies

$$\sqrt{T}\ (\hat{d} - d) = - \left[\frac{\partial^2 S(d)}{\partial d_*^2}\right]^{-1} \sqrt{T}\ \frac{\partial S(d)}{\partial d}$$

$$\overset{A}{=} (D' W D)^{-1}\ D' W \sqrt{T}[\hat{\rho} - \rho(d)].$$

The asymptotic distribution of $\sqrt{T}[\hat{\rho} - \rho(d)]$ for the fractionally integrated, long-memory process is given in Hosking (1984a). Hosking considers the fractionally integrated ARIMA(p,d,q) process as expressed in (1.) where ϵ_t is an independent and identically, but not necessarily normally, distributed white noise error process with mean zero and variance σ^2, ϵ_t has a finite fourth moment, and y_t has mean μ. The sample autocorrelation function is defined as

$$\hat{\rho}_j = \frac{\sum_{t=1}^{T-j}(y_t - \bar{y})(y_{t+j} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}$$

where $\bar{y} = 1/T \sum_{t-1}^{T} y_t$ is the sample mean of the process. For the standard, stationary, short-memory time series process where d takes on integer values, there are standard results for the asymptotic distribution of the sample autocovariance function. However, in the case of the fractionally integrated, long-memory time series process where $-\frac{1}{2} < d < \frac{1}{2}$, Hosking (1984a) shows that these standard results hold for $d \in [-\frac{1}{2},\frac{1}{4})$ but not for $d \geq \frac{1}{4}$. This discrepancy may be attributed to the treatment of the estimation of the mean of the fractionally integrated process. That is, for $\frac{1}{4} \leq d < \frac{1}{2}$ the effect of replacing μ with $\bar{y}$ is not negligible, even asymptotically, and large bias (of the same order of magnitude as that of the standard deviation) is introduced into the estimate of the autocorrelation function. Consequently, the remaining analysis of this chapter will restrict attention to the range of values of the parameter vector for which $d \in [-\frac{1}{2},\frac{1}{4})$. Within this range[3] the estimated autocorrelation functions will be distributed asymptotically normal with variance of order 1/T.

Following Hosking (1984a), the estimated autocorrelation function, $\hat{\rho}$, has covariance matrix C which has i,j th element as given by

[3] For $d = \frac{1}{4}$, asymptotic normality is retained but the variance of the estimated autocorrelation function is of order 1/T(log T). For $d \in (\frac{1}{4},\frac{1}{2})$, asymptotic normality is not retained and the variance is of order $T^{-2(1-2d)}$.

$$(5.)\qquad c_{i,j} = 1/T\ \{\sum_{s=1}^{\infty}(\rho_{s+i} + \rho_{s-i} - 2\,\rho_i\rho_s)(\rho_{s+j} + \rho_{s-j} - 2\,\rho_j\rho_s)\}$$

and $C = \{c_{ij}\}$. In applying the GMM estimation procedure to the single parameter fractionally integrated process, then, the asymptotic distribution of $[\hat{\rho} - \rho(d)]$ will be given by $\sqrt{T}\,(\hat{\rho} - \rho(d)) \sim N(0,C)$ where the dimension of C will be defined by the number of moments used in estimation, and the asymptotic distribution of the GMM estimator will be given by

$$(6.)\qquad \sqrt{T}\,(\hat{d} - d) \sim N[0,(D'C^{-1}D)^{-1}].$$ [4]

4. ASYMPTOTIC PERFORMANCE OF THE GMM ESTIMATOR FOR THE FRACTIONALLY INTEGRATED MODEL

The asymptotic performance of the generalized method of moments estimator for the fractionally integrated process is examined by calculating the large sample variance of $\hat{d}$ as given in (6.). To perform this calculation it is necessary to compute $[D'C^{-1}D]^{-1}$ where D and C are functions of the parameter d and the number of moments, k, used in estimation. In the general case, the efficiency of the GMM estimator should be greatest when calculations are performed utilizing all available moment conditions. However, in the case of the fractionally integrated process, which uses the estimated autocorrelations in calculation, the possible number of available moment conditions is infinite. Relative efficiency, then, should continue to increase as a greater number of

[4] In this representation the optimal weighting matrix, defined in (2.) as W, is given by C^{-1}.

moments are used in estimation such that more moments will always be preferred. As such, the use of any subset of moments in estimation should provide lower levels of efficiency relative to that in which a greater number of moments are employed.

The calculation of the vector of partial derivatives, D, and the covariance matrix of the estimated autocorrelation functions, C, is as follows. Recall that the $(k \times 1)$ vector $\rho(d)$ is given by

$$\rho(d) = \begin{bmatrix} \frac{d}{1-d} \\ \frac{d}{1-d}\;\frac{d+1}{2-d} \\ \vdots \\ \frac{d}{1-d}\;\frac{d+1}{2-d}\;\frac{d+2}{3-d}\;\cdots\;\frac{d+(k-1)}{k-d} \end{bmatrix} .$$

It follows then that D is given by

$$D = \begin{bmatrix} \frac{1}{(d-1)^2} \\ \frac{-2(-1-2d+2d^2)}{(d-1)^2\,(d-2)^2} \\ \frac{3(4+12d-9d^2-6d^3+3d^4)}{(d-1)^2\,(d-2)^2\,(d-3)^2} \\ \vdots \\ \frac{(-1)^{k+1}\;k * f(\cdot)}{(d-1)^2(d-2)^2(d-3)^2\;\cdots\;(d-k)^2} \end{bmatrix}$$

where $f(\cdot)$ will be a function of d^i and $i = (0, 1, 2, 3, \ldots, 2[k-1])$. From the above expressions and the formula given by (6.), the values of D and C are calculated for various values of $d \in [-\frac{1}{2}, \frac{1}{2})$ taken at discrete intervals, that is d = -.50, -.45, -.40, . . . , .20, .24, and various numbers of moment conditions. Relative efficiency comparisons are provided in Tables 1 through 4 which will each be discussed in turn below.

Table 1 presents the asymptotic variance calculations of the GMM estimator for given values of d using moments 1 through "n" in calculation, where n = 1, 2, 3, . . . , 20. In each case it appears that as the number of moments used in estimation increases, the asymptotic variance of the GMM estimator converges to that of the maximum likelihood estimate, $(\pi^2/6)^{-1} = .6079$, as given in Li and McLeod (1986). For positive values of d it appears quite reasonable to conclude that the relative efficiencies of the GMM estimator and the MLE are comparable when only 10 moments are used, although the efficiency of the GMM estimator decreases slightly as the absolute value of d increases. For negative values of d the convergence of the variance of the GMM estimator to that of the MLE requires the use of additional moment conditions in calculation, and the efficiency of the estimator also decreases over this range as the absolute value of d increases.

As discussed in section 2, there is some interest in employing the GMM estimation technique to the fractionally integrated model since this procedure allows for calculation of the estimator based upon subsets of moments so that earlier moments may be dropped from estimation. This notion is particularly attractive within the framework of the long-memory process since the presence of autoregressive and moving average components in the process may contaminate the autocorrelation functions for lower-order moments. Tables 2 through 4 allow for an examination of the efficiency of the GMM estimator when dropping earlier moments in calculation, and the results of each table are discussed below.

Table 2 presents the asymptotic variance of $\hat{d}$ when using only moment "n" in calculation, where n = 1, 2, 3, . . . , 10. In this way it will be possible to examine the contribution of each individual moment condition to the efficiency of the GMM estimator. Table 2 clearly indicates the sacrifice in efficiency for a given value of d when using only one moment condition in estimation, particularly when using any individual moment after the first moment. For negative values of d, for example, the asymptotic variance of the estimator increases dramatically when using any moment other than the first in calculation. For example, for d = -.05, the asymptotic variance of the GMM estimator based on the use of moment two only is more than five times that based on moment one only. The loss in efficiency when using only the second moment is even more dramatic as the value of d decreases to d = -.49. In addition, Table 2 indicates that similar losses in efficiency are evident when calculation is based on use of only moment three, or only moment four, and so on. The same sacrifice in efficiency in using only one moment is evident for positive values of d as well, although the magnitude of the increase in the asymptotic variance is somewhat smaller. For example, for d = .05, the asymptotic variance of the GMM estimator based on the use of only moment two is approximately three times that of the estimator based on only moment one; recall, as discussed above, that for = -.05 the variance is more than five times greater.

Recall that Table 1 illustrated the trade off that existed between the efficiency of the GMM estimator and the absolute value of the parameter d. That is, the relative efficiency of the GMM estimator based on moments 1 through n increases as the absolute value of d decreases.

The same trade off is evident in Table 2. When using only a single moment to calculate the GMM estimator, the asymptotic variance of the estimator decreases as the absolute value of d decreases. This trade off may be explained for the fractionally integrated process by considering the relative contribution of successive moments to the efficiency of the estimator, for a given value of d. As expressed in (6.), the elements of the vector D represent the derivatives of the moment conditions with respect to the parameter, d. In the case of the fractionally integrated process, there is relatively little change in each element of the vector D beyond the first element. This may be attributed to the relative flatness of the autocorrelation functions beyond the first moment, for a given value of d. In addition, the diagonal elements of the matrix C, as expressed in (6.), show relatively little change beyond the first element. It appears, then, in the case of the fractionally integrated process that, for a given value of d, a significant amount of information is contained in the first moment and thus there exists a sacrifice in the efficiency of the GMM estimator when using any one moment, other than the first, in calculation.

Table 3 presents the results of using subsets of five moment conditions in calculating the asymptotic variance of the GMM estimator in which the first moment used in estimation is equal to "n", and n = 1, 2, 3, . . . , 10. Again, it can be seen that, for a given value of d, the asymptotic efficiency of the GMM estimator decreases significantly when the first moment is dropped from the calculations. For example, for d = -.05, the asymptotic variance using moments 2 through 6 is more than four times that using moments 1 through 5. In addition, the results of Table 3 indicate that calculations of the estimator based on a subset of five moment conditions, dropping earlier moments in calculation, are relatively less efficient than calculations based on more (or all available) moment conditions. As observed in Tables 1 and 2, the same trade off exists between the efficiency of the GMM estimator and the value of d when using a subset of five moments; for a given subset of five moments, the efficiency of the GMM estimator increases as the value of d approaches zero. In addition, Table 3 clearly indicates the sacrifice in the efficiency of the GMM estimator that results from dropping more and more of the earlier moments from the calculations. That is, for any given value of d, the asymptotic variance of the GMM estimator increases as more of the earlier moments are dropped from the calculations. For any given value of d when using a subset of five moment conditions, the relative efficiency of the GMM estimator is the greatest when using the first five moments.

The results of Table 3 should not be surprising given the findings of Table 2 which indicate the relative importance of the first moment condition in estimation. It appears that any calculations which omit the first moment condition result in some loss of efficiency.

Finally, Table 4 presents the results of using a subset of ten moment conditions in calculating the asymptotic variance of the GMM estimator, where the first moment used in estimation is equal to "n" and n = 1, 2, 3, . . . , 10. The results of Table 4 are very similar to those of Table 3 in that they indicate the relative loss in efficiency in using subsets of moment conditions where earlier moments are dropped from estimation. It is evident that calculations based on a subset of ten moment conditions, especially when dropping the first moment, involve significant losses in efficiency, with the greatest loss occurring when the largest number of earlier moments are dropped from the calculations. Again, given the results of Table 2 this should not be surprising since a great deal of information is contained in the first moment. It does appear, however, that the efficiency of the GMM estimator is greater when a larger subset of moment conditions are used in the calculations. That is, for any given value of d, the asymptotic variance of the GMM estimator based on a subset of ten moments is smaller than that based on a subset of five moments. It is still the case, however, that the use of a greater number of moments in calculation of the GMM estimator, as opposed to the use of any subset of moments, dominates in terms of the asymptotic efficiency of the estimator.

5. SUMMARY AND CONCLUSION

This chapter has examined the use of the estimation technique of generalized method of moments in estimating the parameters of the fractionally integrated process. The use of this technique is particularly appealing in this context since it does not require the distributional assumptions encountered in using maximum likelihood estimation techniques, and also because it avoids the computational difficulty often encountered in employing approximate MLE techniques. In addition, the relative efficiencies of the two methods appear to be comparable, asymptotically, as the variance calculations provided in Table 1 indicate convergence of the variance of the GMM estimator to that of the MLE (a value of .6079). As such, the GMM estimation technique appears to be a reasonable procedure to employ in the context of the simple ARFIMA(0,d,0) processes.

It does appear, however, that GMM applied to the fractionally integrated process requires the use of lower-order autocorrelations in order to avoid large losses of efficiency. The results of Tables 1 through 4 demonstrate that the relative efficiency of the GMM estimation technique, when applied to the fractionally integrated process, is greatest when using a greater number of moment conditions in estimation. Table 2 shows the significant loss in efficiency which is encountered when the first moment is dropped from estimation. This apparently is due to the relatively small contribution of information attributable to successively higher moments of the long-memory process. This observation is further confirmed in Tables 3 and 4 where there exists considerable inefficiency in using subsets of moment conditions, particularly as a greater number of the earlier moments are dropped in estimation.

These results are especially relevant if one allows for short-run dynamics in the model, as in the case of the ARFIMA(p,d,q) process. For $p > 0$ or $q > 0$, the lower-order autocorrelations may be substantially different than those for the (0,d,0) part of the process. Since it appears that these lower-order autocorrelations cannot be dropped from estimation without sacrificing efficiency, it is reasonable to consider GMM estimation of the ARFIMA(p,d,q) model in the context in which d is estimated jointly with the autoregressive and moving average parameters of the process. This is an important topic for further research.

TABLE 1

Asymptotic Variances of $\sqrt{T}\,(\hat{d} - d)$ Using Moments 1, 2, 3, . . ., n

	n=1	n=2	n=3	n=4	n=5	n=6	n=7	n=8	n=9	n=10
d=-.49	3.4013	2.0666	1.6219	1.3974	1.2611	1.1690	1.1037	1.0573	1.0219	.9938
d=-.45	3.1353	1.9355	1.5337	1.3303	1.2065	1.1225	1.0615	1.0149	.9780	.9480
d=-.40	2.8206	1.7792	1.4281	1.2482	1.1370	1.0610	1.0053	.9626	.9303	.9047
d=-.35	2.5256	1.6294	1.3256	1.1704	1.0752	1.0102	.9629	.9268	.8932	.8667
d=-.30	2.2507	1.4874	1.2274	1.0943	1.0125	.9567	.9159	.9105	.9035	.8977
d=-.25	1.9946	1.3527	1.1331	1.0204	.9510	.9036	.8690	.8425	.8214	.8043
d=-.20	1.7580	1.2257	1.0434	.9496	.8919	.8524	.8236	.8015	.7840	.7697
d=-.15	1.5399	1.1061	.9577	.8814	.8346	.8026	.7793	.7614	.7473	.7358
d=-.10	1.3407	.9950	.8773	.8172	.7804	.7553	.7372	.7233	.7123	.7034
d=-.05	1.1604	.8924	.8025	.7570	.7294	.7108	.6974	.6873	.6793	.6728
d=.00	1.0000	.8000	.7347	.7024	.6833	.6705	.6615	.6547	.6495	.6453
d=.05	.8613	.7202	.6772	.6561	.6444	.6369	.6317	.6279	.6251	.6229
d=.10	.7491	.6588	.6341	.6237	.6184	.6153	.6133	.6120	.6111	.6105
d=.15	.6769	.6303	.6215	.6186	.6179	.6177	.6176	.6176	.6176	.6176
d=.20	.6884	.6783	.6783	.6775	.6759	.6740	.6719	.6698	.6677	.6657
d=.24	.8629	.8594	.8432	.8252	.8088	.7945	.7822	.7714	.7620	.7538

TABLE 1 (Cont'd)

	n=11	n=12	n=13	n=14	n=15	n=16	n=17	n=18	n=19	n=20
d=-.49	.9712	.9527	.9373	.9244	.9135	.9043	.8963	.8894	.8834	.8783
d=-.45	.9230	.9018	.8834	.8674	.8533	.8407	.8295	.8193	.8101	.8017
d=-.40	.8837	.8663	.8515	.8388	.8280	.8185	.8101	.8026	.7960	.7901
d=-.35	.8446	.8258	.8095	.7908	.7748	.7587	.7442	.7306	.7179	.7062
d=-.30	.8930	.8855	.8806	.8768	.8737	.8712	.8691	.8664	.8643	.8625
d=-.25	.7900	.7779	.7675	.7584	.7505	.7435	.7372	.7316	.7253	.7200
d=-.20	.7579	.7478	.7392	.7317	.7251	.7193	.7140	.7094	.7051	.7013
d=-.15	.7263	.7183	.7113	.7055	.7001	.6934	.6885	.6843	.6806	.6771
d=-.10	.6970	.6910	.6859	.6814	.6775	.6740	.6768	.6764	.6742	.6728
d=-.05	.6675	.6630	.6592	.6559	.6531	.6506	.6534	.6522	.6509	.6496
d=.00	.6418	.6390	.6366	.6345	.6327	.6312	.6298	.6286	.6275	.6265
d=.05	.6212	.6197	.6185	.6175	.6167	.6160	.6153	.6148	.6143	.6139
d=.10	.6100	.6096	.6094	.6092	.6090	.6089	.6088	.6087	.6086	.6085
d=.15	.6176	.6176	.6175	.6174	.6174	.6173	.6172	.6171	.6170	.6170
d=.20	.6639	.6622	.6606	.6590	.6576	.6562	.6549	.6537	.6526	.6515
d=.24	.7464	.7399	.7340	.7287	.7239	.7195	.7155	.7118	.7084	.7052

TABLE 2

Asymptotic Variances of $\sqrt{T}\,(\hat{d} - d)$ Using Moment n Only

	n=1	n=2	n=3	n=4	n=5	n=6	n=7	n=8	n=9	n=10
d=-.49	3.4013	276.54	383.92	706.60	1230.1	2029.5	3132.8	4622.3	6543.2	9126.8
d=-.45	3.1353	516.27	463.86	742.51	1199.8	1868.7	2760.0	3936.7	5430.4	7358.4
d=-.40	2.8206	2586.1	703.37	888.06	1282.2	1824.7	2546.7	3477.4	4598.8	6006.6
d=-.35	2.5256	9569.9	1619.9	1346.3	1611.0	2062.6	2681.0	3431.0	4378.1	5465.4
d=-.30	2.2507	454.92	23430.0	3452.5	2827.7	3039.3	3496.8	4117.2	4916.5	5886.4
d=-.25	1.9946	118.74	2853.1	644700.0	15100.0	8865.4	7585.4	7591.7	7987.4	8718.1
d=-.20	1.7580	46.549	357.98	2059.4	13130.0	220000.0	543400.0	87040.0	48290.0	36590.0
d=-.15	1.5399	22.004	103.09	326.53	847.18	1948.1	4164.0	8604.1	17570.0	35730.0
d=-.10	1.3407	11.620	39.894	96.409	193.24	343.76	563.21	866.31	1282.3	1816.1
d=-.05	1.1604	6.1534	18.073	36.554	63.441	99.403	145.45	202.48	271.58	352.19
d=.00	1.0000	4.0000	9.0000	16.000	25.000	36.000	49.000	64.000	81.000	100.00
d=.05	.8613	2.5586	4.8958	7.7839	11.166	15.014	19.328	24.997	29.081	34.572
d=.10	.7491	1.7451	2.9065	4.1977	5.5978	7.0958	8.7097	10.349	12.085	13.904
d=.15	.6769	1.3053	1.9431	2.5900	3.2467	3.9125	4.6027	5.2683	5.9566	6.6543
d=.20	.6884	1.1628	1.5944	2.0000	2.3914	2.7697	3.1317	3.5001	3.8552	4.2052
d=.24	.8629	1.3540	1.7723	2.1492	2.4982	2.8275	3.1409	3.4418	3.7322	4.0137

TABLE 3

Asymptotic Variances of $\sqrt{T}\,(\hat{d} - d)$ Using Five Moments

	1-5	2-6	3-7	4-8	5-9	6-10	7-11	8-12	9-13	10-14
d=-.49	1.2611	18.783	30.082	50.677	81.124	123.23	178.13	247.71	336.24	445.14
d=-.45	1.2065	25.396	35.049	54.530	82.519	119.99	168.15	228.73	303.11	392.54
d=-.40	1.1370	43.514	47.144	65.464	93.257	129.49	174.34	226.67	285.86	360.89
d=-.35	1.0752	104.33	79.417	93.092	118.20	151.85	193.27	242.42	300.21	365.64
d=-.30	1.0125	243.61	205.21	233.60	293.37	352.72	368.21	315.24	373.03	439.79
d=-.25	.9510	80.067	1100.9	770.22	568.87	528.54	540.65	580.11	635.27	703.36
d=-.20	.8919	23.526	178.61	1179.6	7417.6	7995.3	4483.7	3238.2	2735.5	2517.4
d=-.15	.8346	9.7678	38.425	108.16	259.33	569.49	1197.7	2485.3	5223.9	11492.0
d=-.10	.7804	5.0525	14.130	29.573	53.255	87.237	136.46	201.61	286.87	393.32
d=-.05	.7294	3.0305	6.7501	12.000	18.781	27.314	37.697	50.014	64.412	80.942
d=.00	.6833	1.5866	2.4957	6.0128	8.6091	11.604	15.003	18.808	23.010	27.614
d=.05	.6444	1.5008	2.4652	3.5344	4.7070	5.9816	7.3544	8.8226	10.383	12.034
d=.10	.6184	1.2080	1.7925	2.3886	3.0021	3.6355	4.2890	4.9633	5.6568	6.3707
d=.15	.6179	1.0710	1.4693	1.8734	2.2523	2.6272	2.9992	3.3702	3.7398	4.1180
d=.20	.6759	1.1052	1.4667	1.7913	2.0937	2.3809	2.6566	2.9257	3.1872	3.4432
d=.24	.8088	1.3202	1.7423	2.1097	2.4411	2.7470	3.0341	3.3067	3.5676	3.8189

TABLE 4

Asymptotic Variances of $\sqrt{T}(\hat{d} - d)$ Using Ten Moments

	1-10	2-11	3-12	4-13	5-14	6-15	7-16	8-17	9-18	10-19
d=-.49	.9938	7.6420	11.667	18.301	27.366	39.155	53.889	72.010	94.265	121.09
d=-.45	.9480	9.5782	13.336	19.683	28.216	39.070	52.494	68.820	88.312	111.19
d=-.40	.9047	14.678	17.880	24.570	33.483	46.352	57.196	71.978	88.853	109.44
d=-.35	.8667	26.757	26.655	32.747	40.604	53.169	62.260	76.579	94.173	116.78
d=-.30	.8977	101.99	88.877	89.724	99.346	110.97	114.12	107.84	120.12	135.37
d=-.25	.8043	72.637	211.98	173.74	162.58	166.95	178.92	195.72	215.98	236.49
d=-.20	.7697	22.335	175.82	870.72	1398.0	1128.2	924.65	825.65	783.05	770.91
d=-.15	.7358	8.4917	32.553	90.117	214.99	475.75	1013.1	2167.8	4756.9	10826.
d=-.10	.7034	4.2714	11.275	22.508	38.964	61.773	92.229	146.48	216.36	311.04
d=-.05	.6728	2.5812	5.3867	9.0495	13.592	19.045	25.452	37.398	49.635	54.364
d=.00	.6453	1.7927	3.1769	4.7666	6.5621	8.5587	10.758	13.159	15.758	18.558
d=.05	.6229	1.3797	2.1680	2.9939	3.8625	4.7770	5.7375	6.7442	7.7965	8.8950
d=.10	.6105	1.1597	1.6728	2.1727	2.6698	3.1693	3.6742	4.1864	4.7057	5.2343
d=.15	.6176	1.0629	1.4420	1.8127	2.1530	2.4822	2.8034	3.1193	3.4308	3.7462
d=.20	.6657	1.0989	1.4644	1.7894	2.0880	2.3677	2.6330	2.8890	3.1356	3.3751
d=.24	.7538	1.2490	1.6753	2.0540	2.3980	2.7154	3.0119	3.2918	3.5580	3.8126

REFERENCES

Fox, R. and M.S. Taqqu (1986), "Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time-Series", Annals of Statistics, 14, 517-532.

Geweke, J. and S. Porter-Hudak (1983), "The Estimation and Application of Long Memory Time Series Models", Journal of Time Series Analysis, 4, 221-238.

Hosking, J.R.M. (1984a), "Asymptotic Distributions of the Sample Mean, Autocovariances and Autocorrelations of Long-Memory Time Series", Mathematics Research Center Technical Summary Report #2752, University of Wisconsin-Madison.

Hosking, J.R.M. (1984b), "Modelling Persistence in Hydrologic Time Series Using Fractional Differencing", Water Resources Research, 20, 1898-1908.

Li, W.K. and A.I. McLeod (1986), "Fractional Time Series Modelling", Biometrika, 73, 217-21.

Sowell, F.B. (1992), "Maximum Likelihood Estimation of Stationary Univariate Fractionally-Integrated Time-Series Models", Journal of Econometrics, forthcoming.

V. EQUILIBRIUM MONEY DEMAND FUNCTIONS, REAL EXCHANGE RATES, AND PPP: AN ANALYSIS OF COINTEGRATING RELATIONSHIPS IN CANADA AND THE U.S.

CHAPTER V

EQUILIBRIUM MONEY DEMAND FUNCTIONS, REAL EXCHANGE RATES, AND PPP: AN ANALYSIS OF COINTEGRATING RELATIONSHIPS IN CANADA AND THE U.S.

1. INTRODUCTION

Investigation of empirical money demand functions has long been of interest to macroeconomists seeking to identify the existence of a stable long-run relationship between real money balances and measures of economic activity in an economy. Empirical evidence of a stable money demand function can have far-reaching implications for, among other things, the conduct of monetary disinflationary policy within an economy. In addition, estimation of long-run money demand functions can provide researchers with measures of income and interest elasticities of the demand for real balances which can provide insight into the behavior of individuals in the economy.

Traditional modelling of money demand functions has been grounded in the theory that money demand should be linked, in a predictable manner, to the behavior of some scale variable and to some measure of the opportunity cost of holding money. This tradition is now augmented by recent advances in econometrics, in the form of cointegration tests, which allow for a more rigorous analysis of long-run economic relationships. These advances allow researchers to identify stable long-run relationships among a set of non-stationary variables, providing for nearly an ideal framework with which to examine the stability of money demand functions.

In addition, the recent advances in the area of cointegration provide a useful methodology in which to analyze the stability of the real exchange rate between two countries. These techniques can be used to test for the existence of such relationships as the interest rate parity condition or the validity of a long-run Purchasing Power Parity (PPP) between two countries, which are intuitively appealing relationships from an economic standpoint. The issue of stable exchange rates and valid PPP is by no means clear cut among many nations' currencies, and insight into this area can provide valuable information in the areas of international policy and trade.

This chapter investigates the existence of stable money demand functions and exchange rates for Canada and the U.S., utilizing the technique of cointegration as developed by Engle and Granger (1987). The analysis is made within the context of the monetary balance of payments theory which provides a useful method for analyzing issues of the balance of payments and exchange rates that stress the interaction of the supply and demand for money. This theory is grounded on several key assumptions and the validity of these assumptions may be investigated empirically through the cointegration methodology. The contribution of this chapter is two-fold. First, the long-run equilibrium relationship between real money balances, real output, and short-term interest rates in Canada and the U.S. is examined using the cointegration techniques of Johansen (1988) and Johansen and Juselius (1989). This approach will allow for a unique investigation into the existence of cross-country relationships in the monetary model for these two countries. Second, the stability of the Canadian/U.S. exchange rate is examined, also using cointegration techniques, and an investigation is made into the validity of long-run PPP and interest rate parity for these two countries.

The use of the monetary balance of payments framework is of

particular interest in the case of Canada and the U.S. due to the relationship that exists between these two countries with respect to their position as trading partners and the nature of their trade operations. The usual criteria which often cause the theory of the monetary balance of payments to break down do not exist in this case given the unique relationship between Canada and the U.S. or, at least, exist in rather weak form. In this sense, the theory of the monetary balance of payments provides a useful framework in which to examine the issues of interest in this chapter.

The plan of the rest of the chapter is as follows. Section 2 provides a survey of the theoretical underpinnings of the monetary balance of payments framework and the equilibrium monetary model. Section 3 reviews the Johansen and Juselius (1988) methodology for detecting the presence of cointegration and presents the estimation of the equilibrium money demand function for Canada. Particular attention is given, in this section, to the apparent collapse of the money demand relationship in the Canadian data after 1980. Section 4 examines the existence of a stable equilibrium money demand function for the U.S. and confirms its existence for the sample period of this analysis. Section 5 presents the joint estimation of the Canadian and U.S. money demand functions to examine the similarity of the dynamics in the two countries and to investigate the existence of cross-country effects in the model. An alternative to the double-logarithmic specification of the money demand equation for the joint model is also considered in this section to investigate the robustness of the equilibrium money demand function to alternative functional forms. Section 6 examines the stationarity of nominal exchange rates and relative prices and addresses the issue of long-run purchasing power parity. A brief concluding section follows.

2. THE MONETARY BALANCE OF PAYMENTS THEORY

The monetary approach to the balance of payments concentrates on the direct relationship between the money market and the balance of payments. This analytical framework deals exclusively with long-run equilibrium relationships which focus on the connection between prices, output, interest rates, and the balance of payments, and is based on a few central assumptions. Under these assumptions, the balance of payments will reflect any disequilibrium that emerges in the money market.[1]

The primary assumption of the monetary approach to the balance of payments is that the demand for money is a stable function of a given set of variables. This implies that there exists a stable, long-run, equilibrium money demand function for each country in the model. Another assumption is that there is perfect mobility of goods and financial assets between countries so that there is perfect substitutability of these goods and assets. The implication of this assumption is that the market should ensure a single price for each commodity and a single rate of interest. That is, changes in relative goods prices should be proportional to changes in the nominal exchange rate so that the law of one price holds in the long run. In addition, the expected return on interest bearing securities and assets which are denominated in different currencies should

---

[1] The origin of this theory began in the 18th century with the work of David Hume. Contemporary revival of the theory came about with the work of James Meade in the early 1950s and further development is attributed to Polak and his associates at the International Monetary Fund in the late 1950s. Interest in this area was greatly expanded with the work of Mundell (1971) and Johnson (1972). An excellent survey of many of the contributions of the theory can be found in Kreinin and Officer (1978).

be the same. In addition, it is assumed that output and employment tend to full-employment levels, at least in the long run. The specific nature of each of these assumptions will be discussed more fully below.

2.1 LONG-RUN MONEY DEMAND FUNCTIONS

The basic premise of the monetary approach to the balance of payments is that in the long run there exists a stable demand for the stock of money balances as a function of a given set of variables. In particular, the demand for real money balances is posited to be a function of real income and nominal interest rates. To derive this relationship, consider that the demand for nominal money balances, M^d, is a function of nominal income, Q, short-term interest rates, R, and the price level, P. This relationship can be expressed as: $M^d = f(Q,R,P)$. Assuming that individuals in the economy are not subject to money illusion, the money demand function can be written as homogeneous of degree one in prices, and can be re-written as the demand for *real* money balances: $M^d/P = f(Q/P,R)$. In this representation, the demand for real money balances, M^d/P, is a function of real income, Q/P, and interest rates, R. An increase in real income, ceteris paribus, would be expected to increase the demand for money balances since this increases the amount of transactions one is able to finance. An increase in the interest rate, on the other hand, is expected to decrease the demand for money balances since this increases the opportunity cost of holding money.

If one assumes that the income velocity of money, k, varies with the interest rate, and assumes further that the interest elasticity of money demand is unity, then the above money demand relationship can be expressed in the form of the familiar Cambridge cash balances equation as $M^d/P =$

$k(r)Q/P$. Expressed in this manner, the demand for real money balances is homogeneous of degree one in real income.

2.2 REAL EXCHANGE RATES AND THE LAW OF ONE PRICE

As formulated in the monetary approach to the balance of payments, real exchange rates are simply relative prices of the foreign and domestic countries' currencies. That is, the real exchange rate is the foreign currency price of a unit of domestic currency. The determination of this price occurs via the equilibrium supply and demand for the stock of foreign and domestic monies. Under this approach, exchange rates are considered to be predominately a monetary phenomenon, but "real" factors, operating through monetary channels, are also assumed to influence the equilibrium exchange rate.

Given the assumption under the monetary approach to the balance of payments of perfect substitutability across countries in both the goods and capital markets, a natural outcome of the model is that the law of one price should hold for all countries in the model. This implies the existence of a single integrated market for all traded goods and capital where the actions of the market ensure a single price for each commodity and a unique interest rate. That is, the monetary approach to the balance of payments is based on the assumption that interest rate parity holds. Due to efficient arbitrage in assets with similar characteristics, the interest rate differential between two countries will be reflected in the forward premium of the exchange rate so that securities that share the same characteristics will yield the same return in equilibrium.

One further implication of the framework of the monetary balance of payments is that in the long run a unit of domestic currency is expected to command the same purchasing power in the foreign country, when converted into the foreign currency, as it would command in the domestic country. This is the so called Purchasing Power Parity (PPP) theory. The implication of PPP is that the nominal exchange rate, s, which is expressed in terms of units of foreign currency per unit of domestic currency, will be proportional to the ratio of the foreign price level, P^*, and the domestic price level, P: $s = P^*/P$. This relationship suggests that, adjusted for changes in exchange rates, the price levels of the domestic and foreign countries move in step with one another. The resulting implication, assuming that s, P, and P^* are all non stationary, is that a linear combination of the logarithms of the nominal exchange rate and relative prices defines a stationary time series.

2.3 THE LINK BETWEEN THE CANADIAN AND U.S. ECONOMIES

The economies of Canada and the U.S. are so closely linked to one another that fluctuations in one country's macroeconomic variables often closely parallel fluctuations in the macroeconomic variables of the other country. Poole (1967) investigated the strength of the relationship between Canadian and U.S. macroeconomic variables from 1950 through 1962 and identified a strong linear relationship between many of these variables. In addition, Poole noted that the magnitude of movements in the two economies were very similar. These similarities make the Canadian and U.S. economies quite interesting from the perspective of the monetary approach to the balance of payments. The U.S. is by far Canada's largest trading partner with approximately eighty percent of all Canadian exports going to the U.S. and approximately twenty-five percent of all U.S. exports going to Canada. In fact, the trade relationship that exists between Canada and the U.S. is the largest of any two trading partners in the world. In addition, the flow of foreign direct investment between Canada and the U.S. is the largest two-way flow of foreign direct investment anywhere in the world. (See, for example, Rugman (1990) and Hill and Whalley (1985).)

The international trade relationship that exists between Canada and the U.S. can no doubt be attributed partly to the close geographic location of the two countries and partly to the large size of the U.S. relative to Canada in the international economy. The close proximity of the two countries allows for relatively easy mobility of goods and financial assets across borders. In addition, trade relationships between Canada and the U.S. have been steadily improving over the past decade with the advent of actions such as the General Agreement on Tariffs and Trade (GATT) and the Canada/U.S. Free Trade Agreement, such that tariff and non-tariff barriers to trade between the two countries have been substantially lowered. With the end of the Tokyo Round of the GATT negotiations in 1987, Canadian and U.S. tariffs had been reduced to such a level that approximately 80% of Canadian exports entered the U.S. duty free and approximately another 15% entered at tariff rates of less than 5%. For Canada, approximately 65% of U.S. exports entered Canada duty free by 1987 with another 25% entering at rates of less than 5%.

2.4 THE EQUILIBRIUM MONETARY MODEL

Given the conditions set out in the monetary approach to the balance of payments theory, the equilibrium monetary model of money demand and exchange rates may be formalized as shown in equations (1) through (3).

(1) $$\log(M_t/P_t) = \delta_0 + \delta_1 \log(Q_t/P_t) - \delta_2 \log(R_t) + \epsilon_{1t}$$

(2) $$\log(M_t^*/P_t^*) = \zeta_0 + \zeta_1 \log(Q_t^*/P_t^*) - \zeta_2 \log(R_t^*) + \epsilon_{2t}$$

(3) $$\log(s_t) = \gamma_1 \log(P_t^*) - \gamma_2 \log(P_t) + \epsilon_{3t}$$

In this formulation, M is the U.S. nominal money stock, P is the U.S. price variable, Q is U.S. nominal income, R is the short-term U.S. interest rate, s is the Canadian dollar per U.S. dollar nominal exchange rate. Asterisks on variables indicate the Canadian equivalents of the variables. In addition, if $\gamma_1 = \gamma_2 = 1$, then ϵ_{3t} represents the Canadian/U.S. real exchange rate.

The first and second equations specify the demand for the real stock of money in the U.S. and Canada, respectively, as outlined in the monetary balance of payments framework. The disturbance terms about the money demand equations, ϵ_{1t} and ϵ_{2t}, are assumed to be stationary, I(0), processes such that these two equations represent long-run equilibrium money demand functions. When $\gamma_1 = \gamma_2 = 1$, the third equation expresses the formulation of the real exchange rate as the nominal exchange rate times the ratio of the U.S. and Canadian price levels. The term ϵ_{3t}, the real exchange rate if $\gamma_1 = \gamma_2 = 1$, is also assumed to be a stationary, I(0), process such that equation (3) represents a stable linear relationship between the nominal exchange rate and the relative prices of the two countries.

3. CANADIAN EQUILIBRIUM MONEY DEMAND FUNCTIONS

The notion that the demand for money should be a stable function of a given set of relevant macroeconomic variables is intuitively appealing, but empirical evidence to support this notion has been relatively weak over the past several decades [see for example Brenton (1968), Goodhart (1969), Shearer (1970), Clinton (1973), Foot (1977), Cameron (1979) and Poloz (1980)]. Indeed one of the most significant issues in the field of monetary economics centers around the notion that a stable, functional relationship exists between some measure of money and key macroeconomic variables in an economy. The existence of a stable money demand function coupled with an ability on the part of monetary authorities to exert influence over monetary aggregates presents far reaching implications for successful implementation of monetary policy.

In the case of the Canadian economy, several researchers have analyzed the existence of a stable money demand function but with only limited success. Breton (1968) reports to have isolated a stable velocity function for Canada but this work has been heavily disputed by Goodhart (1969) and Shearer (1970). Clinton (1973) has identified a stable money demand function for Canada for the period of 1955 to 1970 using the stock adjustment principle, but finds that the relationship is rather tenuous for all forms of monetary aggregates but the most narrowly defined. Foot (1977) extends the analysis of Clinton (1973) to include the time period form 1970 to 1975 which is characterized by Canada's return to a floating exchange rate regime. These results are also rather tenuous for non-narrowly defined money.

Recent theoretical developments in econometric methods concerning estimation of long-run equilibrium relationships between economic variables have provided a more satisfactory framework in which to analyze the existence of a stable money demand function. The framework, as presented in this paper, uses the notion of cointegration between a set of variables to identify the existence of a stable, long-run equilibrium relationship between these variables.

A cointegrating relationship, as developed in Engle and Granger (1987), implies that there exists stable linear combinations of a set of variables which themselves, independently, are non-stationary. That is, a group of variables is said to be cointegrated if linear combination of the levels of these variables are stationary even though the individual series themselves are stationary only after differencing. The existence of a cointegrating relationship between certain series seems intuitively appealing since there is often theoretical economic support for the notion that certain economic variables should move together over time and obey certain equilibrium conditions, despite the fact that these series may wander from each other in the short run. If a cointegrating relationship does not exist between a set of variables then a long-run link between the variables would not exist.

Analyzing the existence of a stable money demand function in the framework of cointegration is particularly appropriate since the primary assumption of the monetary approach is the existence of an underlying equilibrium relationship between real money balances, real income, and interest rates.

In investigating the existence of a stable money demand function for Canada all data used, with the exception of the Canadian monetary aggregate, are taken from the Main Economic Indicators Historical Statistics Yearbooks of the OECD. The monetary aggregate for Canada is Bank of Canada data on the narrowly defined measure, M1. The income measure used for Canada is real GNP. The interest rate measure is the 3 month Treasury bill rate. The Canadian price variable used to adjust to real quantities is the Canadian consumer price index since the Canadian equivalent of the personal consumption deflator is not currently available. Again, the frequency of the GNP variable is quarterly and all other data series were aggregated to quarterly from monthly through an arithmetic average. All variables were converted to logarithms for this analysis. The full sample period of analysis for the Canadian money supply equation runs from 1956.1 through 1989.3.

A key assumption that must be satisfied before testing for the presence of a cointegrating relationship between variables is that the variables themselves be non-stationary when expressed in levels. The results of unit root tests performed on the fundamentals of the equations (1) through (3) for the U.S. and Canadian variables are presented in Table 1. The conventional unit root tests of Phillips and Perron (1988) are used in this analysis. In all cases, the null hypothesis of a unit root in the data cannot be rejected. The finding of a unit root in the individual series of equations (1) through (3) indicates that the series themselves are non-stationary I(1) processes, which satisfies the initial stage of testing for cointegration among a set of variables.

The methodology used in this chapter to test for the number of cointegrating vectors that exist in a model is the Johansen (1988) and Johansen and Juselius (1989) full information maximum likelihood procedure. This procedure tests for cointegration in a multivariate setting. Statistical evidence of a single cointegrating vector implies that a unique stationary linear combination of the variables of the model exists. In particular, evidence of one cointegrating vector among real money balances, real income, and interest rates implies a unique and stable long-run money demand function.

The estimation technique of Johansen and Juselius is given in detail in Johansen and Juselius (1989), but a brief summary will be given here. Consider a vector process, X_t, containing p random variables, all of which are non-stationary I(1) variables. This vector process may be written in the form of a vector autoregression of order k, VAR(k), as:

$$X_t = \sum_{i=1}^{k} \Pi_i X_{t-i} + \epsilon_t$$

where the Π_i matrices are $(p \times p)$ dimensional coefficient matrices and ϵ_t is a vector white noise process. This expression can represented as:

$$\Delta X_t = \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} - \Pi X_{t-k} + \epsilon_t .$$

In this representation: $\Gamma_i = -I + \Pi_1 + . . . + \Pi_i$, where $i = 1, 2, . . .,$ $k-1$, and $\Pi = I - \Pi_1 - \Pi_2 - . . . . - \Pi_k$. The matrix Π embodies the long-run information of the data. If the matrix Π is of full rank such that $\text{rank}(\Pi)=p$, then it follows that each element of X_t is stationary. If the matrix Π is the null matrix, then it follows that each element of X_t is non-stationary. That is, the equation reduces to a traditional differenced vector time series model. The intermediate case is the one in which Π is not the null matrix and is of less than full rank. In this case, $0 < \text{rank}(\Pi) = r$, where $r < p$. Under this circumstance the rank of Π indicates the number of linearly independent cointegrating relationships among the variables contained in the vector process, X_t. When this condition holds the matrix Π may be expressed as $\Pi = \alpha\beta'$. In this representation, the β matrix is interpreted as the matrix of cointegrating vectors and is of dimension $(p \times r)$ while the α matrix is interpreted as the matrix of vector error correction parameters and is also of dimension $(p \times r)$. Derivation of the maximum likelihood estimates of α, β, and Γ_i is developed in Johansen (1988). In addition, the exact number of cointegrating relationships, r, that exist in the data can be tested using the test developed in Johansen (1988).

With respect to the Canadian model of money demand, the vector process, X_t, contains three random variables and may be expressed as:

$$X = \begin{bmatrix} \log\ (M^*/P^*) \\ \log\ (Q^*/P^*) \\ \log\ (R^*) \end{bmatrix}$$

where (M^*/P^*) represents Canadian real M1 balances, (Q^*/P^*) represents Canadian real GNP, and (R^*) is the Canadian 3 month Treasury bill rate. Recall from equation (2) that the expression for the Canadian money demand function is given by:

$$\log(M_t^*/P_t^*) = \zeta_0 + \zeta_1 \log(Q_t^*/P_t^*) - \zeta_2 \log(R_t^*) + \epsilon_{2t} \tag{2}$$

Given that the variables of the money demand equation have been found to be non-stationary I(1) processes as indicated in Table 1, it is logical to investigate whether a cointegrating relationship exists between the variables of the money demand equation. This would be consistent with finding a long-run equilibrium demand for real balances equation.

The Johansen test for cointegration was applied to Canadian real M1, real GNP, and the 3-month Treasury bill rate to investigate the existence of a long-run money demand function for Canada. These tests initially were performed over the full sample period without imposing any velocity restrictions or including any dummy variables to account for the breakdown in the relationship between M1, income, and interest rates. These results are given in Table 3.

Clearly, the hypothesis that there exists one unique cointegrating vector for the Canadian variables cannot be rejected. There is, however, very little meaning to the estimated income and interest elasticities in this model. The estimated income elasticity is on the order of 15.01 and the estimated interest elasticity is on the order of -16.98. This is consistent with finding a long run equilibrium money demand function for Canada, but indicates some type of structural problem in estimation.

Under the hypothesis that Canadian M1 velocity has been subject to the same conditions which lead to the apparent instability of U.S. velocity in the early 1980s, an investigation is made into the nature of the Canadian velocity relationship. There is evidence of some type of structural change in the Canadian velocity relationship as demonstrated by the results of Table 3. The tests were performed over various sample periods beginning with the period from 1956.1 to 1979.4 on up to the period from 1956.1 to 1983.4. These results clearly indicate the existence in Canada of the same type of "shift in the drift" of the velocity of M1 which has been observed to be present in the U.S. and also in the Japanese data.

The model is re-estimated with the inclusion of a dummy variable which accounts for the break in the velocity relationship for the Canadian M1 data. The dummy variable takes on a value of 1.0 beginning in the third quarter of 1981, and is equal to zero in all previous quarters. Estimation of the model with the inclusion of this dummy variable allows for a shift in the drift parameter of the velocity of M1 in the Canadian model. When specified in this way, the estimated income and interest elasticities for the Canadian money demand function are much closer to the values of 1.0 and -0.50, respectively, which have been found to exist in previous studies.

The magnitude of the estimated income and interest elasticities are of central importance in this analysis. If the estimated income elasticity is truly equal to unity, as is implied in Table 3, then it should be possible to test this restriction with the use of a likelihood ratio test. The next phase of estimation was to impose this velocity restriction that the coefficient on real income is the negative of that on real money balances in the Canadian money demand equation. These results are presented in Table 4. The model is again estimated for various sample periods to account for the break in the velocity of M1. In each case, the estimated values of the likelihood ratio tests for the velocity restriction fail to reject the maintained hypothesis of a unitary income elasticity of demand for real balances.

The results presented in Table 4 indicate the significance of the inclusion of the dummy variable identifying the break in Canadian velocity. In each case, with and without the dummy variable, the hypothesis of a unitary income elasticity of the demand for real balances cannot be rejected. The magnitude of the interest elasticity of demand does seem to vary with the inclusion of the dummy variable, however. With the inclusion of the dummy variable, the estimated interest elasticity is on the order of -0.57.

The results of this analysis indicate that it is appropriate to interpret the estimated cointegrating vector of Table 4, for the full sample period with velocity restrictions and the inclusion of a dummy variable to indicate the break in M1 velocity, as a stationary linear combination of money, income, and interest rates in Canada. These results are strikingly similar to those found for the money demand equations for not only the U.S. but also for Japan[2], which experienced similar phenomena in their M1 velocity relationship. This may suggest some similarity in the dynamics between the U.S. and Canada with regard to the monetary model. The next section will examine joint estimation of the U.S. and Canadian money demand equations to address the issue of the degree of similarity between the two countries.

4. U.S. EQUILIBRIUM MONEY DEMAND FUNCTIONS

Over the period of approximately the last decade and a half, studies investigating the existence of a stable money demand function for the U.S. have not always found satisfactory empirical evidence of an equilibrium relationship between real money balances, some measure of wealth, and some measure of the opportunity cost of holding money. In particular, estimation of this relationship over a post 1980 sample period has indicated an apparent structural change in the underlying relationship among the variables, leading many to question the stability of the long-run money demand function. (See, for example, Judd and Scadding (1982) and Rasche (1987).)

Recent studies by Hoffman and Rasche (1991a and 1991b) and Hafer and Jansen (1991) examine the existence of a stable long-run money demand

---

[2] Rasche (1990) investigates the equilibrium relationship between real Japanese M1 balances, real GNP, and short-term interest rates since 1955 and finds evidence of a stable long-run equilibrium relationship. The estimated long-run income elasticity is found to be insignificantly different from unity and the estimated long-run interest elasticity is found to be near -0.50.

function for the U.S. in the context of a cointegrating relationship between the variables involved. These analyses find strong evidence in support of the existence of a stable equilibrium demand function for real balances in the U.S. economy. In particular, Hoffman and Rasche (1991a) find evidence of a stable, long-run equilibrium relationship between real balances, real income, and interest rates in the U.S during the post World War II period. They examine the narrowly defined monetary aggregate, M1, as well as the monetary base measure of money, and use both long- and short-term interest rates in their analysis. Hoffman and Rasche report a long-run income elasticity of the demand for real balances that is not significantly different from one, and a long-run interest elasticity on the order of about -0.50 for this period. No significant economic difference is found between long- and short-term interest rates in this analysis.

One important aspect of the Hoffman and Rasche (1991a) study is that the model is estimated over various sample periods which allows for an examination of the apparent breakdown of the stable relationship between M1 and measures of economic activity in the U.S. The two sample periods of the analysis run from 1953:1 to 1988:12, which encompasses the full sample period, and from 1953:1 to 1981:12, which includes only that period associated with the existence of a stable M1 velocity relationship in the U.S. In this way, Hoffman and Rasche are able to examine whether the stability or instability of the M1 money demand function changes with the stability or instability of the velocity relationship.

The stability of M1 velocity during the post-Accord period in the U.S. is well documented. The behavior of velocity over this period is described as a random walk with deterministic drift. However, a significant change in the drift parameter is observed in post 1981 M1 velocity data. This apparent "shift in the drift" of the velocity of M1 has lead many researchers to conclude that a stable relationship between the nominal money supply and nominal measures of economic activity no longer exists in the U.S. Hoffman and Rasche identify this break in M1 velocity in the third quarter of 1981 and include a dummy variable, D82, for the post 1981 period of the sample to account for this occurrence. Hoffman and Rasche interpret the D82 variable in the error correction model as representing the shift in the deterministic trend of the real balance series. (For a more complete exposition on the nature of this break in the drift of velocity after 1981, see Hoffman and Rasche (1991b).)

Hafer and Jansen (1991) investigate the existence of a stable money demand function using both the narrowly defined monetary aggregate, M1, and the broader measure, M2, using both long- and short-term interest rates over the periods of 1915 to 1988 and 1953 to 1988. They present evidence of a unique cointegrating relationship for the broader monetary aggregate but fail to conclusively identify such a relationship for the narrowly defined aggregate, M1. Results of estimation on the M1 aggregate over the 1915 to 1988 sample period produce a long-run income elasticity on the order of 0.89 and a long-run interest elasticity of -0.36. Results using the M2 aggregate produce an income elasticity that is much closer to unity and an interest elasticity that is much smaller, in absolute value, than that for the M1 aggregate.

There has been a long-standing debate over the issue of which monetary aggregate measure is the most appropriate for use in measuring money balances, whether it be the narrowly defined measure, M1, or the broader measure, M2. The difficulty, in the Hafer and Jansen (1991) analysis, of finding a long-run equilibrium money demand relationship using the M1 measure is consistent with the findings of previous studies examining this relationship.[3] However, this analysis does not allow for a change in the drift parameter of the M1 velocity relationship, and so no account is taken of the apparent break in M1 velocity which occurred at the end of 1981. This may account for the apparent inability to identify a stable long-run money demand function using the M1 monetary aggregate.

In order to confirm the existence of a stable U.S. money demand function, estimates of the long-run equilibrium interest and income elasticities of the demand for real balances are reproduced in this analysis, consistent with the findings of Hoffman and Rasche (1991a). The full sample period under investigation runs from the first quarter of 1959 through the third quarter of 1990. The data used in this analysis, with the exception of the monetary aggregate, were taken from the Citibase macroeconomic data base. The monetary aggregate used in this analysis is the narrow measure, M1, and was taken from OECD data. The income variable is real GNP. The interest rate measure is the short-term 3 month Treasury bill rate. The price measure used to adjust to real quantities is the personal consumption deflator. The frequency of the GNP variable is quarterly and all other series were aggregated to quarterly frequency from monthly by using a standard arithmetic average. All variables were converted to logarithms for the analysis.

Hoffman and Rasche (1991a) and Hafer and Jansen (1991) present

---

[3] For example, Hallman, Porter, and Small (1989) prefer the use of the M2 aggregate (per unit of potential GNP) over the use of the M1 aggregate, as do Moore, Porter, and Small (1990), and Gavin and Dewald (1989).

evidence on the statistical properties of the data used in their analyses, using the unit root test of Dickey and Fuller (1979). Upon confirming the criterion that the hypothesis of a unit root in all the individual data series can be rejected, tests for a cointegrating relationship among the variables of the model may be performed.

The results of estimating the U.S. money demand equation are presented in Table 2. The results of this analysis find an estimated long-run income elasticity of 0.76 for the narrowly defined monetary aggregate and an estimated long run interest elasticity of -0.40 is found for the short-term interest rate. The D82 variable was included in this analysis to account for the presence of a structural change in the underlying income velocity relationship. Estimation of this model over the full sample period without accounting for the break in M1 velocity produces unsatisfactory results in the sense that the size of the estimated elasticities give them little economic meaning.

In addition, the U.S. money demand function was also estimated with the inclusion of the restriction that the coefficient of real balances plus that on real income sum to zero. This is the velocity restriction used in Hoffman and Rasche (1991a). When this restriction is imposed, the estimated long-run interest elasticity is on the order of -0.52, which is not statistically different from a value of -0.50. In addition, the calculated value of the likelihood ratio test statistic of this restriction fails to reject the hypothesis that the coefficient on real income is unity. This corresponds to an estimated long-run income elasticity of real income of 1.0.

5. JOINT ESTIMATION OF THE CANADIAN AND U.S. MONEY DEMAND FUNCTIONS

This section will examine the joint estimation of the money demand equations for Canada and the U.S. The variables used in estimation are as defined previously and the sample period under investigation runs from the first quarter of 1956 to the third quarter of 1989. In estimating the Canadian and U.S. money demand functions in a combined model it will be possible to examine whether or not the income and interest elasticities of the demand for real balances in the two countries are the same, to identify the magnitude of these elasticities, and to test for the presence of any "cross-country" effects in this two-country model. Identifying of the existence and magnitude of cross-country effects can have important implications for monetary policy within the two countries. That is, estimating joint Canadian and U.S. money demand functions will allow for an examination of the influence of lagged changes in the U.S. variables on the Canadian money demand equation as well as an examination of the influence of lagged changes in the Canadian variables on the U.S. money demand equation. In using the estimation techniques proposed by Johansen (1988) and Johansen and Juselius (1989) it will be possible to examine whether these cross-country effects are present in the long run or whether they are merely a short-run phenomenon. In addition, joint estimation of the model will allow for an analysis of the similarity of the dynamics between the two countries.

Recall from equations (1) and (2) the expressions for the U.S. and Canadian long-run demand for real balances equations:

$$(1) \qquad \log(M_t/P_t) = \delta_0 + \delta_1\log(Q_t/P_t) - \delta_2\log(R_t) + \epsilon_{1t}$$

$$(2) \qquad \log(M_t^*/P_t^*) = \zeta_0 + \zeta_1 \log(Q_t^*/P_t^*) - \zeta_2 \log(R_t^*) + \epsilon_{2t}$$

In estimating the U.S. and Canadian models, the VAR(k) equations for each model may be expressed as:

$$\Delta Xus_t = \sum_{i=1}^{k-1} \Gamma_i \Delta Xus_{t-i} - \Pi \; Xus_{t-k} + \epsilon_t$$

and

$$\Delta Xcan_t = \sum_{i=1}^{k-1} \Gamma_i \Delta Xcan_{t-i} - \Pi \; Xcan_{t-k} + \epsilon_t .$$

In estimating the joint model, this produces the standard VAR(k) expression:

$$\Delta \; X_t = \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} - \Pi \; X_{t-k} + \epsilon_t .$$

In this case, the vector process of the joint model, X_t, may be written as

$$X_t = \begin{bmatrix} Xus_t \\ Xcan_t \end{bmatrix}$$

where

$$Xcan = \begin{bmatrix} \log\,(M^*/P^*) \\ \log\,(Q^*/P^*) \\ \log\,(R^*) \end{bmatrix} \quad \text{and} \quad Xus = \begin{bmatrix} \log\,(M/P) \\ \log\,(Q/P) \\ \log\,(R) \end{bmatrix} .$$

The joint model, then, may be represented as:

$$\begin{bmatrix} \Delta Xus_t \\ \Delta Xcan_t \end{bmatrix} = \Gamma(B) \begin{bmatrix} B\Delta Xus_t \\ B\Delta Xcan_t \end{bmatrix} + \Pi \begin{bmatrix} Xus_{t-k} \\ Xcan_{t-k} \end{bmatrix} + \epsilon_t .$$

The Γ matrices for the joint model are (6×6) dimensional matrices which can be represented in partitioned form as

$$\begin{bmatrix} \Gamma_{11i} & \Gamma_{12i} \\ \Gamma_{21i} & \Gamma_{22i} \end{bmatrix}$$

such that $\sum \Gamma_i \, \Delta X_{t-i}$ is expressed as

$$\begin{bmatrix} \Gamma_{11i} & \Gamma_{12i} \\ \Gamma_{21i} & \Gamma_{22i} \end{bmatrix} \begin{bmatrix} \Delta Xus_{t-i} \\ \Delta Xcan_{t-i} \end{bmatrix} .$$

The Π matrix is also of dimension (6×6) and is expressed, as described in section 3, as $\Pi = \alpha\beta'$. In the joint model, β and α are (6×2) dimensional matrices. As described earlier, the β matrix is interpreted as the matrix of cointegrating vectors and is discussed in more detail later in this section.

Estimation of the joint model is of interest since the Γ matrix may be used to examine the nature of short-run cross-country effects present in the model. In this representation, Γ_{12i} represents the influence of lagged changes in the U.S. variables on the Canadian money demand function in the short run. Similarly, Γ_{21i} represents the influence of lagged changes in the Canadian variables on the U.S. money demand function. These cross-country effects are of interest since they represent the influence that one country's economy has on the demand for real balances in the other country, in the short run. In addition, estimation of Γ_{11i} and Γ_{22i} is of interest since this will allow for an examination of whether or not the dynamics in the two countries are similar.

Estimation of the joint model also is of interest since it allows for an examination of cross-country effects which exist between Canada and the U.S. in the long run. Recall that when $0 < \text{rank}(\Pi) = r$ for $r < p$, the matrix Π may be expressed as $\Pi = \alpha\beta'$ where α is interpreted as the matrix of vector error correction parameters and β is interpreted as the matrix of cointegrating vectors. In the joint model, β' is expressed as

$$\begin{bmatrix} 1 & 0 & \beta_{13} & \beta_{14} & \beta_{15} & \beta_{16} \\ 0 & 1 & \beta_{23} & \beta_{24} & \beta_{25} & \beta_{26} \end{bmatrix}^{4} .$$

In this representation, the β matrix contains information on the effect of the U.S. variables on the Canadian money demand function in the long run and also the effect of the Canadian variables on the U.S. money demand function in the long run. For example, the parameter β_{13} represents the effect of (Q/P) on the U.S. money demand equation in the long run, and the parameter β_{14} represents the effect of (Q^*/P^*) on the U.S. money demand function in the long run. The parameters β_{15} and β_{16} represent the effects of R and R^* on the U.S. money demand function, respectively, in the long run. Similarly, the parameter β_{23} represents the effect of (Q/P) on the Canadian money demand function in the long run and the parameter β_{24} represents the effect of (Q^*/P^*) on the Canadian money demand function in the long run, and so on. The coefficients β_{11} and β_{22} represent the coefficients on real U.S. and Canadian money balances, respectively, and are normalized to one in this analysis.

An interesting result of estimating the joint model is that cross-

---

[4] Note: The ordering of the variables of the β' matrix are as follows:

$$M/P,\ M^*/P^*,\ Q/P\ Q^*/P^*,\ R, \text{ and } R^*,$$

with the first row of the matrix representing the U.S. money demand equation, and the second row of the matrix representing the Canadian money demand equation.

country effects between the two countries in both the long and short run appear to be small. Estimation of the Γ and β matrices of the model indicate that there is little influence of lagged variables in the U.S. on the Canadian money demand functions and relatively little influence of lagged variables in Canada on the U.S. money demand function. For example, in estimating the matrix of normalized cointegrating vectors (normalized β), which represents the long-run information in the data, the following representations of the U.S. and Canadian money demand equations results:

$$\log(M/P)_t = \underset{(0.34)}{1.69} \log(Q/P)_t - \underset{(0.20)}{0.67} \log(Q^*/P^*)_t - \underset{(0.50)}{0.19} \log(R)_t - \underset{(0.11)}{0.14} \log(R*)_t$$

and

$$\log(M^*/P^*)_t = \underset{(0.62)}{-0.19} \log(Q/P)_t + \underset{(0.36)}{1.13} \log(Q^*/P^*)_t + \underset{(0.04)}{0.04} \log(R)_t - \underset{(0.20)}{0.45} \log(R*)_t$$

where the standard errors appear in parentheses beneath the parameter estimates. For Canada, the long-run cross-country effects of U.S. income and interest rates on the demand for Canadian real balances appear to be insignificantly different from zero. For the U.S., the long-run cross-country effect of the Canadian interest rate on the demand for U.S. real balances also appears to be insignificant. There does appear to be a significant relationship between Canadian real income and the demand for real U.S. balances, however. With respect to short-run cross-country effects, a likelihood ratio test of the null hypothesis that all short-run cross-country effects are equal to zero fails to reject this hypothesis.[5] As such, there should be little reason to expect that events related to the monetary model that occur in one economy would be useful in predicting the demand for the stock of real balances in the other economy.

Given the above findings of relatively small cross-country effects in the combined Canadian/U.S. money demand model, the next step of estimation was to examine the model in "stacked" form where all cross-country effects are assumed to be zero. This is represented in the Γ and β matrices as:

$$\Gamma = \begin{bmatrix} \Gamma_{11} & 0 \\ 0 & \Gamma_{22} \end{bmatrix} \qquad \beta' = \begin{bmatrix} 1.0 & 0 & \beta_{13} & 0 & \beta_{15} & 0 \\ 0 & 1.0 & 0 & \beta_{24} & 0 & \beta_{26} \end{bmatrix}$$

The results of estimation of the stacked model are presented in Table 5. The results from the Johansen trace tests clearly indicate the existence of 2 distinct cointegrating relationships which are interpreted as a stable long-run Canadian money demand function and a stable long-run U.S. money demand function. When no velocity restrictions are imposed in estimation, the estimated long-run income elasticities for Canada and the U.S. are 0.73 and 1.16, respectively. In addition, the estimated interest elasticities are -0.36 and -0.59 for Canada and the U.S., respectively.

When velocity restrictions are imposed on the coefficients of real income in both money demand equations, the estimated interest elasticities for Canada and the U.S. are -0.50 and -0.45, respectively. In addition,

---

[5] That is, in estimating the Γ matrix a likelihood ratio test is performed where under the null hypothesis $\Gamma_{12} = \Gamma_{21} = 0$, and under the alternative hypothesis none of the parameters are constrained to be zero.

the value of the likelihood ratio test statistic for the hypothesis that the income elasticity is unity fails to reject the null hypothesis that the velocity restrictions are true. Consequently, estimation of the joint Canadian/U.S. model produces income and interest elasticities that are not significantly different from 1.0 and -0.50, respectively, for both countries.

An interesting alternative specification to consider in estimating the stacked model for Canada and the U.S. is the semi-logarithmic specification of the money demand equation. There has been, for some time now, a dialogue among macroeconomists concerning the appropriate form of the money demand equation and which variables are most appropriate for inclusion in this equation. Among competing alternatives of variables to include in the equation are narrowly and broadly defined monetary aggregates, and long- and short-term interest rates. A competing alternative with regard to the functional form of the money demand function is the semi-logarithmic specification in which the interest rate variable appears in levels rather that in logarithmic form. That is, the money demand functions for the U.S. and Canada would be expressed, respectively, as

$$\log(M_t/P_t) = \theta_0 + \theta_1 \log(Q_t/P_t) - \theta_2(R_t) + \epsilon_{4t} \tag{4}$$

$$\log(M_t^*/P_t^*) = \psi_0 + \psi_1 \log(Q_t^*/P_t^*) - \psi_2(R_t^*) + \epsilon_{5t} \tag{5}$$

Recall that estimation of the double-logarithmic specification of the stacked model clearly produces two cointegrating vectors which may be interpreted as stable long-run velocity functions for Canada and the U.S.

Table 6 presents the trace test statistics and cointegrating vector from estimation of the semi-logarithmic specification of the money demand functions for the joint Canadian/U.S. model with velocity restrictions imposed. These results indicate the existence of three cointegrating vectors, implying the presence of an additional long-run equilibrium relationship in the model.

The presence of this third cointegrating vector in the joint, semi-logarithmic specification of the model may be accounted for by the interest rate parity condition that results from profit-seeking arbitrage activities. More specifically, consider the joint model written in terms of the velocity restrictions as

$$(6) \quad V_t + \theta_2 R_t = \epsilon_{4t}$$

$$(7) \quad V^*_t + \psi_2 R^*_t = \epsilon_{5t}$$

$$(8) \quad R_t = \alpha + \phi R^*_t + \epsilon_{6t}.$$

In this representation, V embodies the velocity restriction imposed on the model and equations (6) through (8) represent three unique equilibrium relationships which correspond to those identified by the Johansen and Juselius procedure for the semi-logarithmic specification of the joint model. The parameters θ and ψ represent the interest elasticities of the U.S. and Canada, respectively; the parameter ϕ embodies the uncovered interest parity condition.

According to the theory of uncovered interest parity, the difference between the domestic and foreign interest rates of two countries should be stationary. This implies a value for ϕ in equation (8) of unity, which corresponds to the finding of a third cointegrating vector for the model estimated via the semi-logarithmic specification. That is, under uncovered interest parity,

$$(R_t - R_t^*) = \alpha + \epsilon_{6t}$$

should be stationary. If uncovered interest parity holds, and therefor accounts for the presence of the third cointegrating vector identified in estimation, then the coefficient on the foreign interest rate should be the negative of that on the domestic interest rate.

In order to test the hypothesis that the third cointegrating vector identified by the Johansen and Juselius procedure is indeed attributable to the condition of uncovered interest parity, a chi-square test of the restriction that $\hat{\phi}$ is equal to -1 was performed. The calculated test statistic indicates that the null hypothesis cannot be rejected; that is, uncovered interest rate parity holds. As such, the third cointegrating vector identified in the semi-logarithmic specification of the joint model has the interpretation of being a stationary difference between the Canadian and U.S. interest rates. This combined with the existence of two stable long-run velocity functions for Canada and the U.S. accounts for three unique, long-run equilibrium relationships among these key macroeconomic variables in the joint Canadian/U.S. model.

6. STATIONARITY OF REAL EXCHANGE RATES

Although much of the recent empirical literature on exchange rates tends to support the notion that exchange rates behave as a random walk, the possibility still remains that exchange rates may belong to a larger equilibrium system in which short-run deviations from equilibrium are possible. Such deviations are discussed by Dornbusch (1976) who noted that exchange rates often times were experiencing temporary overshooting.

Examination of the Canadian/U.S. exchange rate in the context of a long-run equilibrium relationship is a natural setting in which to apply the techniques of cointegration. Use of this methodology enables one to investigate the possible existence a stable long-run equilibrium relationship for the real exchange rate between any two countries. The possible existence of such a relationship between the economies of Canada and the U.S. is particularly interesting given the parallel nature of the two economies.

Before proceeding with the investigation of a cointegrating relationship for the Canadian/U.S. exchange rate it is necessary to determine whether the nominal Canadian/U.S. exchange rate and the price levels of the two countries are non-stationary I(1) variables. Table 1 indicates that each of the variables is indeed I(1) and as such it is possible to test whether a particular linear combination of these variables, the real exchange rate, is stationary. Evidence of a single cointegrating vector among these variables would indicate the existence of a stable long-run relationship between the nominal exchange rate and the relative price levels.

Recall from equation (3) the relationship between the nominal Canadian/U.S. exchange rate and the Canadian and U.S. price levels:

$$(3) \qquad \log(s_t*) = \gamma_1 \log(P_t^*) - \gamma_2 \log(P_t^*) + \epsilon_{3t}$$

If it is found that $\gamma_1 = \gamma_2 = 1$ then equation (3) may be interpreted as the real Canadian/U.S. exchange rate. Evidence of a stable long-run relationship in this formulation may be interpreted as the existence of a stationary real Canadian/U.S. exchange rate.

The vector process, X_t, for the analysis of the real exchange rate may be represented as

$$X = \begin{bmatrix} \log s \\ \log P/P^* \end{bmatrix} .$$

This representation uses the log of the ratio of the price levels of the U.S. and Canada rather than the individual price levels so that the hypothesis of long-run Purchasing Power Parity, PPP, may be examined later.

In estimating the real Canadian/U.S. exchange rate the period considered in this investigation corresponds to the post-Bretton Woods period in which most of the Group of Seven countries abandoned their systems of fixed exchange rates. The data run from January of 1974 through September of 1990. The nominal exchange rate was taken from International Monetary Fund data which contains end-of-month observations. The nominal exchange rate is expressed in terms of Canadian dollars per U.S. dollar. Results of estimation of the exchange rate model are provided in Table 7.

The results of the Johansen trace tests for cointegration clearly indicate the presence of one unique cointegrating relationship between the nominal exchange rate and the ratio of the U.S. and Canadian price levels over the period of 1974 to 1990. This is consistent with finding a stationary real Canadian/U.S. exchange rate for this period. The results of this analysis indicate that the estimated parameter on the ratio of the price levels is 1.40.

The model was also estimated over a sample containing the period in which Canada initially abandoned their system of fixed exchange rates, the period beginning in June of 1970, but the results over this period were inconclusive on the long-run properties of the data. The trace test statistics of the analysis over this sample period do not indicate a stationary long-run real Canadian/U.S. real exchange rate for the period beginning in 1970. This apparent tenuousness of the model indicates the sensitivity of the stability of the Canadian/U.S. real exchange rate to the time period under consideration. This may not be a surprising result, however, if one considers that the period immediately following the collapse of the Bretton Woods system (June of 1970) was one in which there was still substantial adjustment occurring in many of the exchange rates which began to float at this time.

An interesting implication of the 1974-1990 estimation result can be made with regard to the validity of the long-run PPP relationship. If an estimated parameter of unity is found for the ratio of real prices in the model, this indicates that changes in relative goods prices between the two countries will be proportional to changes in the nominal exchange rate. This is consistent with the assumptions laid out in the monetary approach to the balance of payments.

The existence of a valid PPP relationship may not be a reasonable expectation for all pairs of countries in the world, but is particularly appealing for the U.S. and Canada. These two countries seem to satisfy, most nearly, the conditions described by the monetary approach to the balance of payments for which PPP should be expected to hold. Canada is relatively small, compared to the U.S., and the two countries are the largest trading partners of any two countries in the world. The close proximity of the two countries allows for relatively easy mobility of goods and capital across borders with very limited restrictions to trade between the two countries. In addition, evidence of stable, long-run money demand functions for both countries has been found, as indicated in sections 3 through 5.

Given these conditions, one would expect that the nominal Canadian/U.S. exchange rate and the relative price levels of the two countries, although non-stationary by themselves, should not move too far from one another in the long run and as such should have a stable linear representation. That is, it should be reasonable to expect that the variables comprising the Canadian/U.S. real exchange rate should move together over time and that the relative goods prices between the two countries, adjusted for changes in the exchange rate, should move in step with one another.

The validity of the PPP hypothesis can be tested within the framework of the Johansen model by imposing the restriction of a unitary coefficient on the ratio of the price levels in the real exchange rate equation. The results of this investigation are also presented in Table 7. The value of the likelihood ratio test statistic for the hypothesis that the coefficient on the ratio of the price levels is the opposite of that on the nominal exchange rate fails to reject the maintained hypothesis at the 1% level, although the null hypothesis is rejected at the 5% level. This may be interpreted as evidence in favor of the long-run PPP relationship between Canada and the U.S., although this evidence is somewhat weak. It may be appropriate, therefore, to conclude that a valid long-run PPP relationship exists between the U.S. and Canada, consistent with theoretical expectations.

7. SUMMARY AND CONCLUSION

This chapter has investigated key elements of the theory of the monetary balance of payments for the economies of Canada and the U.S., using the methodology proposed by Johansen (1988) and Johansen and Juselius (1989). The primary element of the monetary balance of payments theory, the existence of a stable long-run demand for real money balances, is found to exist for both Canada and the U.S. for the period of the late 1950s through 1989. With respect to the monetary model, the dynamics between the two countries are found to be very similar in that the estimated interest elasticities of velocity are found to be insignificantly different from .50 for both countries. In addition, there appear to be no long- or short-run cross-country effects between the two countries with regard to the formulation of the demand for real money balances. A test of the null hypothesis that all cross-country effects in the model are zero could not be rejected, despite the finding of some long-run influence in the U.S. money demand equation with respect to Canadian real income. In general, then, it may not be reasonable to expect that influences from U.S. monetary conditions would be useful in predicting the demand for the stock of real money balances in Canada, and vice versa.

A further component of the theory of the monetary balance of payments, the existence of a stable Canadian/U.S. real exchange rate, is found in this analysis as well. Empirical evidence of a stable real exchange rate is found for the period after the breakdown of the Bretton

Woods system of fixed exchange rates, which spans the period from January of 1974 to the present. This indicates that the Canadian/U.S. nominal exchange rate and the relative price levels of the two countries tend to move together in the long run, even though individually they may tend to wander without bound.

In addition, strong support is also found for the model of Canada and the U.S. for the validity of the law of one price, especially with regard to the capital market. That is, uncovered interest rate parity is found to hold for these two countries for the period from the late 1950s through 1989. With respect to the goods market, some evidence is found for the validity of long-run purchasing power parity between Canada and the U.S., although the sensitivity of this relationship to the time period under investigation must be noted. The null hypothesis of a valid PPP relationship between Canada and the U.S. is maintained at the 1% level of significance but not at the 5% level, and the only time period for which the relationship holds is between January of 1974 and September of 1990. Despite the tenuousness of the long-run purchasing power parity relationship for the Canadian and U.S. economies, it does appear that each of the key components of the theory of the monetary balance of payments are well-grounded for the case of Canada and the U.S.

TABLE 1

Phillips and Perron Unit Root Tests

test statistic:		$Z(t_{\alpha^*})$		$Z(t_{\tilde{\alpha}})$	
lag length:		4	12	4	12
series:	P	1.46	4.78	8.92	7.42
	R	-2.07	-2.65	-2.18	-1.95
	M/P	1.77	1.25	-3.05	-2.22
	Q/P	-2.32	-2.18	1.43	1.43
	P^*	-2.80	-2.14	6.89	4.32
	R^*	-2.22	-2.08	-1.83	-1.71
	M^*/P^*	-2.36	-2.28	-1.67	-1.46
	Q^*/P^*	-1.90	-1.73	1.56	1.65
	s	-0.38	-0.69	-1.02	-1.10

Key: $Z(t_{\alpha^*})$ and $Z(t_{\tilde{\alpha}})$ are the Phillips and Perron adjusted t statistics used to test the parameter on the lagged dependent variable in a regression with an intercept, and with an intercept and a time trend, respectively. The test statistic $Z(t_{\alpha^*})$ tests the null hypothesis of a unit root, H_0: $\alpha^* = 1$, in the regression $y_t = u^* + \alpha^* y_{t-1} + \epsilon_t^*$. The test statistic $Z(t_{\tilde{\alpha}})$ tests the null hypothesis of a unit root, H_0: $\tilde{\alpha} = 1$, in the regression $y_t = \tilde{u} + \beta[(t-n)/2] + \tilde{\alpha} y_{t-1} + \tilde{\epsilon}_t$. The critical values for these test statistics at the .05 level are -2.86 and -3.41, respectively, and at the .10 level are -3.43 and -3.91, respectively.

TABLE 2

Cointegration Tests, U.S. Data

Real M1, Real GNP, and Treasury Bill Rate
(double-logarithmic specification)

	Johansen Trace Test Statistics			Normalized Cointegrating Vector		
sample	r≤0	r≤1	r≤2	M/P	Q/P	R
1959.1 - 90.3 (with dummy)	37.41	15.77	5.60	1.0	-0.76 (0.12)	0.40 (0.07)

with velocity restriction:

Johansen Trace Test Statistics r≤0	Johansen Trace Test Statistics r≤1	LR test of Velocity Restriction	Estimated Interest Elasticity
27.84	7.51	1.34	0.52 (0.06)

Key: The standard errors are given beneath the parameter estimates in parentheses. Critical values for trace tests at 5% level are 31.53, 17.95, and 8.18 for r≤0, r≤1, and r≤2 cointegrating vectors, respectively. Results are based on k=4 lag specification. The critical value for the likelihood ratio test, $\chi^2_{(1)}$, at 5% level is equal to 3.84.

TABLE 3

Cointegration Tests, Canadian Data

Real M1, Real GNP, and Treasury Bill Rate
(double-logarithmic specification)

	Johansen Trace Test Statistics			Normalized Cointegrating Vector		
sample	r≤0	r≤1	r≤2	M/P	Q/P	R
1956.1 - 89.3	40.41	8.04	0.71	1.0	-15.01 (81.16)	16.98 (94.09)
1956.1 - 89.3 (with dummy)	42.97	8.15	0.41	1.0	-1.06 (0.26)	0.59 (0.27)
1956.1 - 79.4	30.89	11.74	3.83	1.0	-1.05 (0.26)	0.49 (0.33)
1956.1 - 80.4	28.31	8.70	1.51	1.0	-1.01 (0.21)	0.46 (0.31)
1956.1 - 81.1	31.11	7.70	0.26	1.0	-1.15 (0.31)	0.63 (0.39)
1956.1 - 81.2	32.21	7.57	0.00	1.0	-1.17 (0.49)	0.66 (0.41)
1956.1 - 81.3	33.49	7.56	0.05	1.0	-1.43 (0.84)	1.03 (1.11)
1956.1 - 81.4	37.07	8.67	0.05	1.0	2.14 (12.62)	-3.65 (16.27)
1956.1 - 82.4	37.59	7.59	0.99	1.0	3.20 (4.33)	-2.89 (4.83)
1956.1 - 83.4	34.34	7.07	0.38	1.0	87.57 (137.1)	99.63 (157.2)

Key: The standard errors are given beneath the parameter estimates in parentheses. Critical values for trace tests at 5% level are 31.53, 17.95, and 8.18 for r≤0, r≤1, and r≤2, respectively. Results are based on k=4 lag specification.

TABLE 4

Cointegration Tests with Velocity Restrictions, Canadian Data

Real M1, Real GNP, and Treasury Bill Rate
(double-logarithmic specification)

sample	Johansen Trace Test Statistics $r \leq 0$	$r \leq 1$	LR test of Velocity Restriction	Estimated Interest Elasticity
1956.1 - 89.3	26.57	1.03	0.06	1.02 (0.22)
1956.1 - 89.3 (with dummy)	29.89	1.89	0.03	0.57 (0.12)
1956.1 - 79.4	24.02	4.89	0.39	0.43 (0.08)
1956.1 - 80.4	24.96	5.36	0.16	0.45 (0.08)
1956.1 - 81.4	26.53	1.87	1.12	0.63 (0.20)
1956.1 - 82.4	24.86	3.57	0.15	1.09 (0.87)
1956.1 - 83.4	22.54	0.86	1.12	0.72 (0.22)

Key: The standard errors are given beneath the parameter estimates in parentheses. Critical values for trace tests at 5% level are 17.95 and 8.18 for $r \leq 0$ and $r \leq 1$, respectively. Results are based on k=4 lag specification. The critical value for the Likelihood Ratio test, $\chi^2(1)$, at 5% level of significance is 3.84.

TABLE 5

Cointegration Tests for the Joint Model
double-logarithmic specification

Real M1, Real GNP, and Treasury Bill Rate

Johansen Trace Test Statistics

sample	r≤0	r≤1	r≤2	r≤3	r≤4	r≤5
1956.1 - 89.3 (with dummies)	143.12	67.25	32.47	15.01	3.60	0.29

Cointegrating Vector

M/P	M^*/P^*	Q/P	Q^*/P^*	R	R^*
1.0	1.0	-1.16	-0.73	0.59	0.36

With Velocity Restrictions:

Johansen Trace Test Statistics				LR test of Velocity Restrictions	Estimated Interest Elasticities for	
r≤0	r≤1	r≤2	r≤3		Canada	U.S.
118.61	55.30	10.12	4.12	4.05	-0.48	-0.49
					(0.16)	(0.19)

Key: The standard errors are given beneath the parameter estimates in parentheses. Critical values for trace tests at 5% level are

r=0	r=1	r=2	r=3	r=4	r=5
95.18	70.60	48.28	31.53	17.95	8.18

when testing without imposing velocity restrictions, and

r=0	r=1	r=2	r≤3
48.28	31.53	17.95	8.18

when imposing velocity restrictions. Results are based on k=4 lag specification. The critical value for the Likelihood Ratio test, $\chi^2_{(2)}$, at 5% level of significance is 5.99.

TABLE 6

Cointegration Tests for the Joint Model

Semi-Logarithmic Specification

	Johansen Trace Test Statistics			
sample	r≤0	r≤1	r≤2	r≤3
1956.1 - 89.3 (with dummies)	72.56	32.26	17.43	3.35

Cointegrating Vector

θ	ψ	ϕ
.0813	.0457	-.1727
(.0017)	(.0013)	(.4253)

Likelihood Ratio Statistic: 1.61

Key: The standard errors are given beneath the parameter estimates in parentheses. Critical values for trace tests at 5% and 10% levels are:

	r≤0	r≤1	r≤2	r≤3
5%	48.28	31.53	17.95	8.18
10%	45.23	28.71	15.66	6.50

Results are based on k=4 lag specification. The critical value for the Likelihood Ratio test, $\chi^2_{(1)}$ at 5% level of significance is 3.85.

TABLE 7

Cointegration Tests for Canadian/U.S. Real Exchange Rate

sample period	Johansen Trace Test Statistics $r \leq 0$	$r \leq 1$	LR test of H_0: $\gamma_1/\gamma_2=-1$	Estimated Coefficient
1970.6 - 1984.12	4.40	0.15		1.91 (0.45)
1970.6 - 1985.12	4.79	0.01		2.12 (0.51)
1970.6 - 1986.12	5.21	0.10		1.89 (0.31)
1970.6 - 1987.12	4.16	0.01		1.63 (0.26)
1970.6 - 1988.12	1.82	0.11		1.14 (0.19)
1970.6 - 1989.12	10.44	1.50		1.40 (0.17)
1974.1 - 1987.12	5.89	2.19		0.88 (0.95)
1974.1 - 1988.12	7.04	1.05		11.16 (51.00)
1974.1 - 1989.12	13.09	3.83		1.45 (0.25)
1974.1 - 1990.09	18.51	5.52	5.18	1.40 (0.24)

Key: The standard errors are given in parentheses beneath the parameter estimates. Critical vaules for the trace tests at the 5% level are 17.95 and 8.17 for $r \leq 0$ and $r \leq 1$, respectively. The critical value for the Likelihood Ratio test, $\chi^2_{(1)}$ at the 1% level of significance is 6.64; the critical value at the 5% level of significance is 3.85.

REFERENCES

Blejer, M.I. and J.A. Frenkel (1989), "Monetary Approach to the Balance of Payments", in The New Palgrave, J. Eatwell, M. Milgate, and P. Newman, editors, New York: W.W. Norton, 497-500.

Boothe, P.M. and S.S. Poloz (1988), "Unstable Money Demand and the Monetary Model of the Exchange Rate", Canadian Journal of Economics, 21, 785-798.

Breton, A. (1968), "A Stable Velocity Function for Canada?", Economica, 35, 451-453.

Burdekin, R.C.K. and P. Burkett (1990), "A Re-Examination of the Monetary Model of Exchange Market Pressure: Canada, 1963-1988", Review of Economics and Statistics, 4, 677-681.

Cameron, N. (1979), "The Stability of Canadian Demand for Money Functions 1954-75", Canadian Journal of Economics, 10, 259-281.

Clinton, K. (1973), "The Demand for Money in Canada, 1955-70: Some Single-Equation Estimates and Stability Tests", Canadian Journal of Economics, 6, 53-60.

Dickey, D.A. and W.A. Fuller (1979), "Distribution of the Estimates for Autoregressive Time Series with a Unit Root", Journal of the American Statistical Association, 74, 427-431.

Dickey, D.A., D.W. Jansen, and D.L. Thornton (1991), "A Primer on Cointegration with an Application to Money and Income", Review, Federal Reserve Bank of St. Louis, 73, 58-70.

Dornbusch, R. (1976), "Expectations and Exchange Rate Dynamics", Journal of Political Economy, 84, 1161-1176.

Engle, R.F. and C.W. Granger (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing", Econometrica, 55, 251-276.

Foot, D.K. (1977), "The Demand for Money in Canada: Some Additional Evidence", Canadian Journal of Economics, 10, 475-485.

Gavin, W.T. and W.G. Dewald (1989), "The Effect of Disinflationary Policies on Monetary Velocity", Cato Journal, 9, 149-164.

Goodhart, C.A.E. (1969), "A Stable Velocity Function for Canada? A Note", Economica, 36, 314-315.

Hallman, J.J., Porter, R.D, and D.H. Small (1989), "M2 per Unit of Potential GNP as an Anchor for the Price Level", Staff Study, 157, Board of Governors of the Federal Reserve System.

Hafer, R.W. and D.W. Jansen (1991), "The Demand for Money in the United States: Evidence from Cointegration Tests", Journal of Money, Credit, and Banking, 23, 155-68.

Hill, R. and J. Whalley (1985), "A Possible Canada-U.S. Free Trade Arrangement", Summary of the Proceedings of a Research Symposium, in Canada-United States Free Trade, J. Whalley, research coordinator, Toronto: University of Toronto Press, 43-66.

Hoffman, D. and R.H. Rasche (1991a), "Long Run Income and Interest Elasticities of Money Demand in the United States", Review of Economics and Statistics, 73, 665-674.

Hoffman, D. and R.H. Rasche (1991b) "The Demand for Money in the U.S. During the Great Depression and Post War Periods: Identifying the Source of Shifts in Velocity Drifts", mimeo.

Hong, K. and R.H. Rasche (1991), "Multivariate Cointegration Tests and Long-Run Purchasing Power Parity Theory", mimeo.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", Journal of Economic Dynamics and Control, 12, 231-254.

Johansen, S. and K. Juselius (1989), "The Full Information Maximum Likelihood Procedure for Inference on Cointegration with Applications", Preprint No. 4, Institute of Mathematical Statistics, University of Copenhagen.

Johnson, H.G. (1972), "The Monetary Approach to the Balance of Payments", in Further Essays in Monetary Economics, London: Allen and Unwin.

Judd, J.P. and J.L. Scadding (1982), "The Search for a Stable Money Demand Function: A Survey of the Post-1973 Literature", Journal of Economic Literature, 20, 993-1023.

Kreinin, M.E. and L.H. Officer (1978), "The Monetary Approach to the Balance of Payments: A Survey", Princeton Studies in International Finance, 43.

Laidler, D. (1985), The Demand for Money: Theories, Evidence, and Problems, 3rd edition, London: University of Western Ontario.

Laidler, D. (1986), "International Monetary Economics in Theory and Practice", in Postwar Macroeconomic Developments, J. Sargent, editor, University of Toronto Press, Toronto, 225-270.

MacDonald, R. and T.S. Torrance (1988), "Covered Interest Parity and UK Monetary 'News'", Economics Letters, 26, 53-56.

McNown, R. and M. Wallace (1989), "Co-Integration Tests for Long Run Equilibrium in the Monetary Exchange Rate Model", Economics Letters, 31, 263-267.

Miller, S.M. (1991), "Monetary Dynamics: An Application of Cointegration and Error-Correction Modeling", Journal of Money, Credit, and Banking, 23, 140-154.

Moore, G.R., Porter, R.D., and D.H. Small (1990), "Modelling Disaggregated Demands for M2 and M1 in the 1980s: The U.S. Experience", in Board of Governors of the Federal Reserve System, Financial Sectors in Open Economies: Empirical Analysis and Policy Issues, Washington, D.C.

Morici, P. (1985), "Trends in U.S. Trade Policy and Non-Tariff Barriers", in Canada-United States Free Trade, J. Whalley, research coordinator, Toronto: University of Toronto Press, 223-238.

Moroz, A.R. (1985), "Some Observations on Non-Tariff Barriers and Their Use in Canada", in Canada-United States Free Trade, J. Whalley, research coordinator, Toronto: University of Toronto Press, 239-266.

Mundell, R. (1963), "Capital Mobility and Stabilisation Policy Under Fixed and Flexible Exchange Rates", Canadian Journal of Economics and Political Science, 29, 475-485.

Phillips, P.C.B. and P.Perron (1988), "Testing for a Unite Root in Time Series Regression", Biometrika, 75, 335-346.

Poloz, S.S. (1980), "Simultaneity and the Demand for Money in Canada", Canadian Journal of Economics, 13, 407-420.

Poole, W. (1967), "The Stability of the Canadian Flexible Exchange Rate, 1950-1962", Canadian Journal of Economics and Political Science, 33, 205-217.

Rasche, R.H. (1987), "M1-Velocity and Money Demand Functions: Do Stable Relationships Exist?", in Empirical Studies of Velocity, Real Exchange Rates, Unemployment, and Productivity, K. Brunner and A. Meltzer editors, Carnegie-Rochester Conference Series on Public Policy, 27, 9-88.

Rasche, R.H. (1990), "Equilibrium Income and Interest Elasticities of the Demand for M1 in Japan", Bank of Japan Monetary and Economic Studies, 8, 31-58.

Rugman, A. (1990), Multinationals and Canada-United States Free Trade, Columbia: University of South Carolina Press.