THREE ESSAYS ON EXCHANGE RATE MODELS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Min-Seok Yang, M.A.

* * * * *

The Ohio State University

1997

Dissertation Committee: Approved by

Professor Masao Ogaki, Advisor

Professor G.S. Maddala

Professor Pok-Sang Lam

Advisor

Department of Economics
Copyright by
Min-Seok Yang
1997
Vector Autoregression (VAR) models for stationary random variables are reduced form models. Structural interpretations of VAR models require restrictions on structural form models. When random variables are unit root nonstationary, Error Correction Models (ECM) are widely used. We argue that the standard ECMs are reduced form models just as VAR models are.

We consider a structural ECM with restrictions which are based on economic theory. In particular, we consider structural ECMs in which at least one variable adjusts slowly to its long-run equilibrium level, with a constant speed of adjustment. We discuss sufficient conditions under which the structural speed of adjustment coefficient is consistently estimated by standard estimation methods for ECMs. We show that these conditions are not satisfied by an exchange rate model with slow price adjustment. We then propose an instrumental variable (IV) method to estimate the structural speed of adjustment coefficient.

First, we perform a single equation method, where an IV method is applied to a slow adjustment equation. We obtain positive estimates for the structural speed of adjustment coefficient in most cases. Second we employ a system method, where Hansen and Sargent's method that applies GMM to linear rational expectation models is
combined with our single equation method. The speed of adjustment coefficient is estimated from a) the slow adjustment equation for the domestic price and b) the rational expectations equation for the exchange rate. We form a specification test by comparing the estimates for the speed of adjustment coefficient from these two equations.

So far we have assumed that the unbiased interest parity holds. As a first step towards combining risk premium models with the sticky price exchange rate model, we compare risk premium models nested in a generalized factor pricing model to discover which model is more consistent with the forward exchange rate data. With various instruments, the traditional CAPM usually fits the data better than the simple consumption CAPM. However, whether the CAPM or the habit formation model fits the data better depends on the choice of instruments.
Dedicated to my family
ACKNOWLEDGMENTS

I thank Professors Stephen Cecchetti, Stephen Cosslett, Mario Crucini, Paul Evans, Peter Howitt, Nelson Mark, J. Huston McCulloch, James Peck, René Stulz and Alan Viard for their teaching, discussion, or comments.

I especially thank Professors G.S. Maddala and Pok-Sang Lam for their kindness, encouragement, teaching, and valuable comments. In particular, I express my gratitude to Professor G.S. Maddala for his fundamental insight.

Words are insufficient to express my thanks to my advisor, Professor Masao Ogaki, for his enthusiasm, intellectual support, and especially inspiring in me more advanced way of thinking, all of which made this dissertation possible, and for his patience in correcting my scientific errors.
VITA

November 11, 1960 ...................................... Born - Seoul, Korea
1982 ......................................................... B.A. Economics, Korea University
1990 .......................................................... M.A. Economics, Korea University
1988 - 1992 .................................................. Researcher,
                                             Korea Energy Economics Institute
1992 - present ............................................ Graduate Teaching and Research
                                             Associate, The Ohio State University

FIELDS OF STUDY

Major Field : Economics
Studies in International Finance
             Econometrics
             Monetary Economics
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>vi</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Structural Error Correction Models: A Single Equation Method</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>4</td>
</tr>
<tr>
<td>2.2 An Exchange Rate Model with Sticky Prices</td>
<td>6</td>
</tr>
<tr>
<td>2.3 Structural Models and Error Correction Models</td>
<td>12</td>
</tr>
<tr>
<td>2.4 A Single Equation Method</td>
<td>15</td>
</tr>
<tr>
<td>2.5 A Measurement Error Model</td>
<td>17</td>
</tr>
<tr>
<td>2.6 Econometric Methods</td>
<td>19</td>
</tr>
<tr>
<td>2.7 Empirical Results</td>
<td>25</td>
</tr>
<tr>
<td>2.8 Conclusion</td>
<td>37</td>
</tr>
<tr>
<td>3. Structural Error Correction Models: A System Method</td>
<td>42</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>42</td>
</tr>
<tr>
<td>3.2 Literature Survey</td>
<td>44</td>
</tr>
<tr>
<td>3.3 A System Method for Rational Expectation Models</td>
<td>49</td>
</tr>
<tr>
<td>3.4 An Exchange Rate Model with Sticky Prices</td>
<td>54</td>
</tr>
<tr>
<td>3.4.1 Exchange Rate Model</td>
<td>54</td>
</tr>
<tr>
<td>3.4.2 Econometric Model</td>
<td>56</td>
</tr>
<tr>
<td>3.4.3 Stable Money Demand</td>
<td>59</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Purchasing Power Parity</td>
<td>39</td>
</tr>
<tr>
<td>2.2 Instrumental Variable Estimation: Single Equation</td>
<td>40</td>
</tr>
<tr>
<td>3.1 Money Demand Equation</td>
<td>69</td>
</tr>
<tr>
<td>3.2 Instrumental Variable Estimation: System</td>
<td>71</td>
</tr>
<tr>
<td>4.1 CAPM with US stock Returns</td>
<td>100</td>
</tr>
<tr>
<td>4.2 Consumption CAPM: time separable ((\beta=1))</td>
<td>101</td>
</tr>
<tr>
<td>4.3 Consumption CAPM: time separable ((\beta=0.99))</td>
<td>102</td>
</tr>
<tr>
<td>4.4 Consumption CAPM: time separable ((\beta=0.995))</td>
<td>103</td>
</tr>
<tr>
<td>4.5 Consumption CAPM: time separable ((\beta=1.005))</td>
<td>104</td>
</tr>
<tr>
<td>4.6 Consumption CAPM: time non-separable</td>
<td>105</td>
</tr>
<tr>
<td>4.7 Consumption CAPM: time non-separable ((\beta=1))</td>
<td>106</td>
</tr>
<tr>
<td>4.8 Consumption CAPM: time non-separable ((\beta=1, \gamma&lt;10))</td>
<td>107</td>
</tr>
<tr>
<td>4.9 LR-type test: CAPM vs. Consumption CAPM (time separable, (\beta=1))</td>
<td>108</td>
</tr>
<tr>
<td>4.10 LR-type test: CAPM vs. Consumption CAPM (time separable, (\beta=0.995))</td>
<td>109</td>
</tr>
<tr>
<td>4.11 LR-type test: CAPM vs. Consumption CAPM (time non-separable, (\beta=1))</td>
<td>110</td>
</tr>
<tr>
<td>4.12 LR-type test: CAPM vs. Consumption CAPM (time non-separable, (\beta=0.99))</td>
<td>111</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

As Cooley and Leroy (1985) point out, Vector Autoregression (VAR) models for stationary random variables are reduced form models. Structural interpretations of VAR models require restrictions on structural form models. Since the appearance of the seminal paper by Davidson et al. (1978), Error Correction Models (ECM) are widely used, when random variables are unit root nonstationary. We argue that the standard ECMs are reduced form models just as VAR models are.

In chapter 2 and 3, we consider a structural ECM which reflects restrictions based on the economic theory. In particular, we consider structural ECMs in which at least one variable adjusts slowly to its long-run equilibrium level, with a constant speed of adjustment. We discuss sufficient conditions under which the structural speed of adjustment coefficient is consistently estimated by standard estimation methods for ECMs. We show that these conditions are not satisfied by an exchange rate model with slow price adjustment. We then propose an instrumental variables (IV) method to estimate the structural speed of adjustment coefficient.

In chapter 2, we deal with a single equation method where an IV method is applied to a slow adjustment equation that describes how a variable adjusts...
slowly to the long-run equilibrium level in the structural ECM. We apply this method to a one good version of Mussa’s (1982) model, which may be viewed as a stochastic discrete time version of Dornbusch’s (1976) model. Using the single equation method, we obtain positive estimates for the structural speed of adjustment coefficient in most cases.

In chapter 3, we present the system method, where Hansen and Sargent’s (1982) method that applies Hansen’s (1982) Generalized Method of Moments (GMM) to linear rational expectations models is combined with our single equation method. The system method is more efficient than the single-equation method when the restrictions implied by linear rational expectations models are true. On the other hand, the single equation method is robust to misspecification in the other equations of the structural ECM.

In the system method, the speed of adjustment coefficient can be estimated from the slow adjustment equation for the domestic price and the rational expectations equation for the exchange rate. We form a specification test by comparing the estimates for the speed of adjustment coefficient from these two equations.

In this chapter, we assume that the unbiased interest rate parity holds. Fama (1984) and many other studies, however, show that the forward foreign exchange rate is not an unbiased predictor of the future spot exchange rate. One explanation for this bias is based on the existence of risk premiums. The risk premium models assume market efficiency and the possibility that forward rates contain time-varying risk premiums as well as market forecasts of future spot rates.

In chapter 4, we consider models dealing with the bias in the forward
exchange market. As a first step towards combining the risk premium models with the sticky price exchange rate model, we compare the risk premium models with each other to find out which model is more consistent with the forward rate data. In particular, we compare all these models within a nested framework. Since the popular risk premium models, such as the traditional CAPM and consumption-based CAPMs, can be regarded as particular cases of arbitrage pricing theory, and these differ only in the specification of the stochastic discount factor, we can compare all these models in a generalized multifactor model.
CHAPTER 2

STRUCTURAL ERROR CORRECTION MODEL:
A SINGLE EQUATION METHOD

2.1 Introduction

As Cooley and Leroy (1985) point out, Vector Autoregression (VAR) models for stationary random variables are reduced form models. Structural interpretations of VAR models require restrictions on structural form models. Since the appearance of the seminal paper by Davidson et al. (1978), Error Correction Models (ECM) are widely used, when random variables are unit root nonstationary. As Engle and Granger (1988) show, an ECM representation exists, when the variables are cointegrated and vice versa. An ECM is a stable VAR in the first differences of the variables, augmented by one lag of the cointegrating vector itself, which represents the equilibrium error. Thus, the ECM is a particular representation of a vector autoregression appropriate for cointegrated vectors. We argue that the standard ECMs are reduced form models, just as VAR models are. Hence, the standard ECM may not recover the structural parameters, without imposing particular restrictions.
In this chapter, we consider structural ECMs, in which at least one variable adjusts slowly to its long-run equilibrium level, with a constant speed of adjustment. We discuss sufficient conditions, under which the structural speed of adjustment coefficient is consistently estimated by standard estimation methods for ECMs. We show that these conditions are not satisfied by an exchange rate model with slow price adjustment. We then propose an instrumental variables (IV) method to estimate the structural speed of adjustment coefficient. In a single equation method, which we deal with in this chapter, an IV method is applied to a slow adjustment equation, that describes how a variable adjusts slowly to its long-run equilibrium level in the structural ECM.

These methods are applied to an exchange rate model with sticky prices. The model we use is a one-good version of Mussa’s (1982) model, which may be viewed as a stochastic discrete time version of Dornbusch’s (1976) model. Data are for the exchange rates of currencies of Canada, France, Germany, Italy, Japan, Netherlands, Switzerland, United Kingdom, against the U.S. dollar. Using the single equation method, we obtain positive estimates for the structural speed of adjustment coefficient for most cases.

The rest of this chapter is organized as follows. Section 2.2 presents an exchange rate model in which the domestic price adjusts slowly toward the Purchasing Power Parity (PPP) level. In Section 2.3, a structural ECM is presented, and its relationship to a reduced form ECM is discussed. Section 2.4 proposes a single equation method for the structural ECM. In Section 2.5, the model of Section 2.2 is augmented to include measurement errors. In section 2.6, we describe the econometric methods that we employ in this paper. Section 2.7 presents our empirical results for the single equation method. Section 2.8
concludes the first chapter.

2.2 An Exchange Rate Model with Sticky Prices

In general, the dynamics of exchange rates can be characterized by two factors, exogenous shocks and the transmission mechanism of the exogenous shocks. Exogenous shocks are assumed to have their own dynamics which is called exogenous dynamics. Various time-series assumptions on the exogenous dynamics are given to capture the dynamics of the endogenous variables in the economic model. On the other hand, various transmission mechanisms themselves of the exogenous shocks bring about particular types of the dynamics of endogenous variables. The dynamics of target variables such as exchange rates which reflects a particular mechanism of transmission will be called endogenous dynamics. In this chapter we focus on the endogenous dynamics of exchange rates, which occurs when we assume that prices in the goods market are sticky.

There are various types of sticky price models in the literature. They differ in the source of the stickiness of the prices. Kiley (1996) compares three different sticky price models: the partial adjustment model, the staggered price setting model and the P-bar model. In these models output is usually assumed to be at its natural level to see the endogenous dynamics of the sticky price models. The partial adjustment model for prices is represented by

\[ p_t = (1-\pi)p_{t-1} + \pi p^*_t, \]  

(2.1)
where \( p_t \) is the log aggregate price level, \( 0 < \pi < 1 \) is the speed of price adjustment, and \( p_t^* \) is the target price. The target price is obtained by

\[
p_t^* = (1-(1-\pi)p)\sum_{j=0}^{\infty}(1-\pi)^j E_t p_{t+j}.
\]  (2.2)

Kiley distinguishes two interpretations of the partial adjustment model. The first interpretation is based on the quadratic cost function. Since quadratic costs of adjustment bring about partial adjustment, the speed of price adjustment is determined by the parameters of the adjustment cost function. In this case, the price setters change their prices in each period, but only adjust these partially towards the equilibrium levels. The second interpretation is based on the uncertainty that firms face. The fraction \( \pi \) of the price setters change their prices in period \( t \) to the target price, while the fraction \( (1-\pi) \) of the price setters remain unchanged at the price of period \( t-1 \) in period \( t \). In this case, although every individual firm does not adjust in every period even partially, the aggregate price level, nevertheless, follows the partial adjustment process.

In the staggered price setting model, the economy consists of two types of price setters. Both types choose a fixed nominal price for two periods, but differ in the period in which they change price. This model is very similar to the partial adjustment model. However, Kiley argues that the staggered price setting model differs from the partial adjustment model, because it has an important imperfect information component that does not enter the partial adjustment model.

Mussa’s (1982) model which we consider in this paper belongs to the P-bar model of price adjustment. The P-bar model assumes that price adjustment depends
on movements in the equilibrium price level, \( p_t^e \),

\[
p_t - p_{t-1} = \pi(p_{t-1}^e - p_{t-1}^e) + E_{t-1}(p_t^e - p_{t-1}^e),
\]  

(2.3)

In this model, the price level at \( t \) adjusts based on the degree of disequilibrium in \( t-1 \), and adjusts for any expected movements in the equilibrium price. Kiley (1996) points out that the degree of disequilibrium indicates inflationary pressure in a sticky price world, and that prices are set before trading in a period, so that price setters can only adjust at \( t \) to the movements in equilibrium prices expected at \( t-1 \). Therefore, he asserts that this model is an imperfect information model. This imperfect information brings about the effect of monetary shocks on the economy.

Mussa’s (1982) model incorporates both exogenous and endogenous dynamics. Obstfeld and Stockman (1985) clarify this, pointing out that they are the equilibrium adjustment of prices to current and anticipated future movements in exogenous variables, and the adjustment of prices and quantities as goods market disequilibrium is eliminated over time. In Mussa’s model, the nominal price of domestic output adjusts over time in response to deviations between aggregate demand and the full-employment level of output. The economic agents in these dynamic models are endowed with rational expectations of exchange rate and price movement. Obstfeld and Stockman point out that one implication of the sticky price models is that slower adjustment in the goods market implies a more volatile exchange rate.

Researchers have tested sticky price models in various ways. Frankel (1979) derives an equation for exchange rate determination, in which the spot rate is
expressed as a function of the relative money supply, the relative income level, the nominal interest rate differential, and the expected long-run inflation rate differential. His model also allows a role for differences in the secular rates of inflation. Since his work, many economists have studied the relationship between the real exchange rate and the real interest rate differentials.¹

Another direction of research is to specify the flexible-price monetary models of exchange rate determination as ECMs. MacDonald and Taylor (1993) use the multivariate cointegration technique proposed by Johansen (1988,1989) and Johansen and Juselius (1990) to test the long run relationship between the exchange rate and the monetary variables. Although their exchange rate equation of a dynamic error correction form gives an estimate for the coefficient on the equilibrium error, it does not necessarily mean that it gives the estimate for the speed of adjustment in goods market. Papell (1995) derives a reduced form from the Dornbusch’s (1976) model. His reduced form with cointegration does not recover all of the structural parameters either.

The failure of recovering the structural parameters results from their reduced forms. Cooley and LeRoy (1985) show that Vector Autoregression (VAR) models for stationary random variables are reduced form models. Structural interpretations of VAR models require restrictions on structural models, such as triangularity or orthogonality assumption. To treat VAR models as structural, however, we need explicit justification for whatever triangularity or orthogonality assumption made. These assumptions are not arbitrary normalization, but substantive restrictions on the parameter space that must be justified from theory. On the other hand, if we interpret VAR models as nonstructural, then

triangulization and orthogonalization are in fact arbitrary normalizations not requiring theoretical justification. But the model cannot be interpreted as doing more than merely summarizing the correlations in the data. Similar interpretations can be given to the traditional ECM.

Our purpose in this chapter is to estimate the speed of price adjustment in a structural model, developing a new approach of estimation. Especially we focus on the ECM in the literature, and suggest a way of solving the problem in applying the traditional ECM to the sticky price models.

In order to motivate a particular form of a structural ECM, we consider a simple exchange rate model in which the domestic price adjusts slowly toward the long-run equilibrium level implied by Purchasing Power Parity (PPP). Let $p(t)$ be the log domestic price level, $p^*(t)$ be the log foreign price level, and $e(t)$ be the log nominal exchange rate (the price of one unit of the foreign currency in terms of the domestic currency). We assume that these variables are first difference stationary. We also assume that PPP holds in the long run, so that the real exchange rate, $p(t) - p^*(t) - e(t)$, is stationary, or $y(t) = (p(t), e(t), p^*(t))'$ is cointegrated with a cointegrating vector $(1, -1, -1)$. Let $\mu = E[(p(t)-p^*(t) - e(t))$. Then, $\mu$ can be nonzero when different units are used to measure prices in the two countries.

We use a one-good version of Mussa’s (1982) model, where the domestic price level is assumed to be sticky in the short-run and to adjust to the foreign price level which is given in the long-run and converted into a domestic price level

\[ p^e(t) = \mu + e(t) + p^*(t) \]  \hspace{1cm} (2.4A)

\[ p(t+1) - p(t) = b[p^e(t) - p(t)] + E[p^e(t+1) - p^e(t)], \]  \hspace{1cm} (2.4B)
where \( p^e(t) \) is the long-run equilibrium level of domestic price at period \( t \) and \( b \) is the speed of price adjustment. This specification implies that the domestic price level is assumed to adjust slowly to the PPP level through

\[
\Delta p(t+1) = b \left[ \mu + p^*(t) + e(t) - p(t) \right] + E[p^*(t+1) + e(t+1)]
- [p^*(t) + e(t)]
\]  

(2.5)

where \( \Delta x(t+1) = x(t+1) - x(t) \) for any variable \( x(t) \), \( E(\cdot | I_t) \) is the expectation operator conditional on \( I_t \), the information available to the economic agents at time \( t \), and a positive constant \( b < 1 \) is an adjustment coefficient. The idea behind Equation (2.5) is that the domestic price level adjusts slowly toward its long-run PPP level of \( p^*(t) + e(t) \). The adjustment speed is slow when \( b \) is close to zero, and the adjustment speed is fast when \( b \) is close to one. From Equation (2.5), we obtain

\[
\Delta p(t+1) = d + b[ p^*(t) + e(t) - p(t)] + \Delta p^*(t+1) + \Delta e(t+1)
+ \epsilon(t+1)
\]  

(2.6)

where \( d = b\mu, \epsilon(t+1) = E[p^*(t+1) + e(t+1)] - [p^*(t+1) + e(t+1)] \). Hence \( \epsilon(t+1) \) is a one-period ahead forecasting error, and \( E[\epsilon(t+1) | I_t] = 0 \). Equation (2.6) motivates the form of structural ECM used in the next section.
2.3 Structural Models and Error Correction Models

In this section, we discuss the relationship between structural models and ECMs. Let $y(t)$ be an $n$-dimensional vector of first difference stationary random variables. We assume that there exist $\rho$ linearly independent cointegrating vectors, so that $A'y(t)$ is stationary, where $A'$ is a $(\rho \times n)$ matrix of real numbers whose rows are linearly independent cointegrating vectors. Consider a standard ECM

$$\Delta y(t+1) = k + GA'y(t) + F_1 \Delta y(t) + F_2 \Delta y(t-1) +$$

$$\ldots + F_p y(t-p+1) + v(t+1), \quad (2.7)$$

where $k$ is an $(n \times 1)$ vector, $G$ is an $(n \times \rho)$ matrix of real numbers, and $v(t)$ is a stationary $n$-dimensional vector of random variables with $E(v(t+1)|I_{t-1})=0$. In many applications $\tau=0$, but we will give examples of applications in which $\tau>0$. There exist many ways to estimate (2.7). For example, Engle and Granger’s (1987) two step method or Johansen’s (1988) Maximum Likelihood methods can be used.

In many applications of standard ECMs, elements in $G$ are given structural interpretations as parameters of the speed of adjustment toward the long-run equilibrium represented by $A'y(t)$. It is of interest to study conditions under

---

2 We will treat more general cases in which the expectation of $v(t+1)$ conditional on the economic agents’ information is not zero, but the linear projection of $v(t+1)$ onto an econometrician’s information set (which is smaller than the economic agents’ information set) is zero.
which the elements in $G$ can be given such a structural interpretation. In the model of the previous section, the domestic price level adjusts gradually to its PPP level, with a speed of adjustment parameter $b$. We will investigate conditions under which $b$ can be estimated as an element in $G$ from (2.7).

In our view, (2.7) is a reduced form model. A class of structural models can be written in the following form of a structural ECM:

\[
C_0 \Delta y(t+1) = d + BA' y(t) + C_1 \Delta y(t) + C_2 \Delta y(t-1) + \\
... + C_p \Delta y(t-p+1) + u(t+1) \tag{2.8}
\]

where $C_i$ is a $(n \times n)$ matrix, $d$ is a $(n \times 1)$ vector, and $B$ is a $(n \times p)$ matrix of real numbers. Here, $C_0$ is a nonsingular matrix of real numbers with ones along its principal diagonal, and $u(t)$ is a stationary $n$-dimensional vector of random variables with $E[u(t+1)|I_{t-\tau}] = 0$. Even though the cointegrating vectors are not unique, we assume that there is a normalization that uniquely determines $A$, so that parameters in $B$ have structural meanings.

When $y(t) = ((p(t), e(t), p^*(t)))$, Equation (2.6) based on the model in the previous section gives the first row of Equation (2.8) with $A' = (I, -I, -I)$, the first row of $B$ is $b$, the first rows of $C_1$, $C_2$, ..., $C_p$ are zero vectors, and the first row of $u(t)$ is $\varepsilon(t)$. Note that for any nonzero constant $\psi$, $\psi(I, -I, -I)'$ is also a cointegrating vector. However, the first row of $B$ is $b$ only when $\psi$ is normalized to one.

We assume that $C_0$ is nonsingular, and premultiply both sides of (2.8) by $C_0^{-1}$ to obtain the standard ECM (2.7), where $k = C_0^{-1}d$, $G = C_0^{-1}B$, $F_i = C_0^{-1}C_i$, and $v(t) = C_0^{-1}u(t)$. Thus, the standard ECM, estimated by Engle and Granger's (1987) two
step method or Johansen's (1988) Maximum Likelihood methods, is a reduced form model. Hence, it cannot be used to recover structural parameters in $B$, nor can the impulse response functions based on $v(t)$ be interpreted in a structural way, unless some restrictions are imposed on $C_0$.

As in a VAR, various restrictions are possible for $C_0$. One example is to assume that $C_0$ is lower triangular. If $C_0$ is lower triangular, then the first row of $G$ is equal to the first row of $B$, and structural parameters in the first row of $B$ are estimated by the standard methods to estimate an ECM.

In the exchange rate model in the previous section, $b$ is a structural parameter of interest. For the purpose of estimating $b$ in the model, the restriction that $C_0$ is lower triangular is not attractive. When $p(t)$ is taken as the first element of $y(t)$, the model implies that $C_0$ is not lower triangular because the first row of $C_0$ is $(1, -1, -1)$. In Chapter 3, we will close the model of the previous section by adding a money market clearing condition and the Unconditional Interest Parity condition. As we will show, in the representation (2.8) for $y(t) = ((p(t), e(t), p^*(t)))'$,

$$C_0 = \begin{bmatrix} 1 & -1 & -1 \\ -(1/bh) & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

(2.9)

$A' = (1, -1, -1)$, and $B = (-b, 0, 0)'$, where $h$ is the interest elasticity of demand for money. Hence, the structural ECM from the one-good version of the exchange rate model does not satisfy the restriction that $C_0$ is lower triangular for any ordering of the variables. Even though some structural models may be written in lower triangular form, this example suggests that many structural
models cannot be written in that particular form.

It is instructive to observe the relationship between the structural ECM and the reduced form ECM in the exchange rate model. Because

\[
C_0^{-1} = \begin{bmatrix}
\frac{bh}{(bh-1)} & \frac{bh}{(bh-1)} & 0 \\
\frac{1}{(bh-1)} & \frac{bh}{(bh-1)} & -1 \\
0 & 0 & 1
\end{bmatrix},
\] (2.10)

\[G = C_0^{-1}B = [-b^2h/(bh-1), -b/(bh-1), 0]^\prime.\] Comparing \(G\) and \(B\), observe the way in which the contemporaneous interactions between the domestic price and the exchange rate affect the speed of adjustment coefficients. The speed of adjustment coefficient for the domestic price is \(b\) in the structural model, while it is \(b^2h/(bh-1)\) in the reduced form model. The error correction term does not appear in the second equation for the exchange rate in the structural ECM, while it appears with the speed of adjustment coefficient of \(b/(bh-1)\) in the reduced form model.

2.4 A Single Equation Method

Because standard methods of estimating Equation (2.7) may not recover the structural parameters of interest in \(B\), we propose a single equation procedure to estimate a structural ECM (2.8) which does not require restrictions on \(C_0\). This method applies an IV method to Equation (2.8).

Suppose we are interested in estimating the first row of Equation (2.8). In some applications, the cointegrating vectors are known, and thus the values of \(A\)
are known. In other applications, the values of $\mathbf{A}$ are unknown. In the case of unknown cointegrating vectors, a two step method similar to Engle and Granger’s (1987) or Cooley and Ogaki’s (1996) methods can be used. In this two-step method, the cointegrating vectors are estimated consistently in the first step.

In the first step, we estimate $\mathbf{A}$, using a method to consistently estimate cointegrating vectors. There exist many methods to estimate cointegrating vectors. Johansen’s Maximum Likelihood (ML) Estimators for Equation (2.7) can be used for this purpose. If $\rho$ is equal to one, estimators based on regressions that are as efficient as Johansen’s ML estimators such as Phillips and Hansen’s (1990) Fully Modified Estimation Method, Park’s (1992) Canonical Cointegrating Regression, or Stock and Watson’s (1993) estimator can be used. Ordinary Least Squares estimators are also consistent when $\rho$ is equal to one, but not as efficient as these estimators. We assume that $\mathbf{A}_T$ is the first step estimator, where $T$ is the sample size, and that $\mathbf{A}_T$ converges to $\mathbf{A}$ at a faster rate than $T^{1/2}$.

In the single equation method, an IV method is applied to

$$
\Delta y_1(t+I) = d_1 - c^1_{02} \Delta y_2(t+I) - \ldots - c^1_{0n} \Delta y_n(t+I) + b^1_1 \mathbf{A}' y(t) \\
+ c^1_1 \Delta y(t) + c^1_2 \Delta y(t-I) + \ldots + c^1_p \Delta y(t-p+I) + u^1_1(t+I),
$$

(2.11)

where $y_i(t)$ is the $i$-th element of $y(t)$, $d_1$ is the first element in $\mathbf{d}$, $c^1_{0i}$ is the $i$-th element of the first row of $\mathbf{C}_0$, $b^1_1$ is the first row of $\mathbf{B}$, $c^1_1$ is the first row of $\mathbf{C}_1$, and $u^1_1(t)$ is the first element of $\mathbf{u}(t)$. When $\mathbb{E}[u^1_1(t+I)|I_{t-T}]=0$, any

---

3 Usually, $\mathbf{A}_T$ converges at the rate of $T$, but there are cases where $\mathbf{A}_T$ converges at the rate of $T^{3/2}$ (see West (1988)).
stationary variable in the information set available at time \( t-\tau \) that is correlated with variables in the right hand side of Equation (2.11) can be used as an instrumental variable. In the case of known cointegrating vectors, the known values of \( A \) are used in (2.11). In the case of unknown cointegrating vectors, \( A \) in Equation (2.11) is replaced by \( A_T \) obtained in the first step. Because \( A_T \) converges to \( A \) at a faster rate than \( T^{1/2} \), the first step estimation does not affect the asymptotic distributions of the second step estimator as in Engle and Granger’s (1987) two step procedure.

2.5 A Measurement Error Model

We apply the single equation procedure to the exchange rate model of Section 2.2, using monthly exchange rate and aggregate price data for Canada, France, Germany, Italy, Japan, Netherlands, Switzerland, the United Kingdom, and the United States from January 1974 to June 1995. In the model, we assume that \( y(t) = (p(t), e(t), p^*(t)) \) is cointegrated with a known cointegrating vector \((1,-1,-1)\). This assumption may cause a problem in applications of the model to data in the post-Bretton Woods period because many researchers have failed to reject the null hypothesis of no cointegration using similar data sets. Because more favorable evidence for the assumption is often found when a longer sample period is used, the failure to reject no cointegration may be due to low power of the no cointegration tests in small samples. Because the evidence is mixed, a sensitivity analysis with respect to this assumption is in order.

For the purpose of a sensitivity analysis, we employ Cheung and Lai (1993)
and Fisher and Park's (1991) model with measurement errors to allow the cointegrating vector to be different from \((1,-1,-1)\). Let \(p^m(t)\) and \(p^*_m(t)\) be the log measured domestic and foreign prices, which are related to the true prices by

\[
p^m(t) = \theta + \phi p(t) + \nu(t) \quad (2.12A)
\]

\[
p^*_m(t) = \theta^* + \phi^* p^*(t) + \nu^*(t) \quad (2.12B)
\]

where \(E_{t-1}\nu(t) = 0\) and \(E_{t-1}\nu^*(t) = 0\). We assume that true prices follow the model of Section 2.2 and satisfy PPP in the long-run. Then

\[
p^n(t) - \phi e(t) - (\phi/\phi^*) p^m(t) = (\theta-\theta^* \phi/\phi^*) + \phi [p(t)-e(t)-p^*(t)]
+ [\nu(t)-(\phi/\phi^*) \nu^*(t)] \quad (2.13)
\]

is stationary. Hence, \(y(t) = (p^n(t), e(t), p^m(t))^t\) is cointegrated with a cointegrating vector \((1,-\phi,-\phi/\phi^*)\). In the first step, we run a cointegrating regression of the form

\[
p^n(t) = \psi_0 + \psi_1 e(t) + \psi_2 p^m(t) + \zeta(t), \quad (2.14)
\]

where \(\psi_1 = \phi, \psi_2 = \phi/\phi^*\), and \(\zeta(t)\) is stationary with mean zero.

In order to obtain the second step estimator, we use Equation (2.6) and 
\[
\Delta p^n(t+1) = \phi \Delta p(t+1) + \Delta \nu(t+1)
\]

to obtain
\[ \Delta p^m(t+1) = d - b \left[ p^m(t) - \phi e(t) - (\phi/\phi^*)p^m(t) \right] \\
+ (\phi/\phi^*)\Delta p^m(t+1) + \phi \Delta e(t+1) + w(t+1) \]  
(2.15)

where \( d = b(\mu + \theta - \phi^*) \phi/\phi^* \), and

\[ w(t+1) = \phi e(t+1) + \nu(t+1) - (I-b)\nu(t) - (b\phi/\phi^*)v^*(t+1) \\
+ (I-b)(\phi/\phi^*)v^*(t). \]  
(2.16)

Because \( E_{t+1}w(t+1) = 0 \), we can apply the two step procedure from the last section as long as the instrumental variables are chosen from the information set available at \( t-I \). In this case, the second step is to apply an IV estimation method to Equation (2.15), where \( \phi \) and \( \phi^* \) are obtained in the first step estimation. Because \( E_{t+1}w(t+1) = 0 \) and \( w(t+1) \) is in the information set available at \( t+1 \), \( w(t+1) \) has a moving average (MA) representation of order one, and this serial correlation structure needs to be taken into account (see, e.g., Ogaki (1993a) for an explanation of methods which treat this type of serial correlation in GMM).

2.6 Econometric Methods

In this section we summarize the econometric methods we employ (See Ogaki, 1993a and 1993b). For the first stage estimation, we use Park(1992)'s Canonical Cointegration Regressions (CCR) method, and for the second stage one we use

Park suggests a procedure for statistical inference in cointegrating regressions. His canonical cointegrating regressions are basically the regressions formulated with the transformed data. When we apply the least square procedure to the transformed model, we obtain asymptotically efficient estimators and chi-square tests. Consider a cointegrating system

\[ y_t = X_t' \gamma + \varepsilon_t \quad (2.17A) \]
\[ \Delta X_t = \nu_t \quad (2.17B) \]

where \( y_t \) and \( X_t \) are difference stationary, and \( \varepsilon_t \) and \( \nu_t \) are stationary with zero mean. Here \( y_t \) is a scalar and \( X_t \) is a \((n-1) \times 1\) random vector. Let

\[ w_t = (\varepsilon_t, y_t') \]

Define \( \Phi(i) = E(w_t w_{t-1}') \), \( \Sigma = \Phi(0) \), \( \Gamma = \sum_{i=0}^{\infty} \Phi(i) \), and \( \Omega = \sum_{i=-\infty}^{\infty} \Phi(i) \). Here \( \Omega \) is the long run covariance matrix of \( w_t \). Partition of \( \Omega \) as

\[ \Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \]

and partition \( \Gamma \) conformably. Define \( \Omega_{11,2} = \Omega_{11} - \Omega_{12} \Omega_{22}^{-1} \Omega_{21} \) and \( \Gamma_2 = (\Gamma'_{12}, \Gamma'_{22})' \). The CCR procedure assumes that \( \Omega_{22} \) is positive definite, implying that \( X_t \) is not itself cointegrated. This assumption assures that \((1, -\gamma)\) is the unique cointegrating vector (up to a scale factor). The OLS estimator in equation
(2.17A) is super-consistent because the estimator converges to $\beta$ at the rate of $T$ (sample size) even when $\Delta x(t)$ and $u(t)$ are correlated. The OLS estimator, however, is not asymptotically efficient. To obtain asymptotically efficient OLS estimator Park suggest a transformed model:

$$y_t^* = y_t + \Pi_y' w_t,$$

$$X_t^* = X_t + \Pi_x' w_t.$$

Because $w_t$ is stationary, $y_t^*$ and $X_t^*$ are cointegrated with the same cointegrating vector $(1, -\beta)$ as $y_t$ and $X_t$ for any $\Pi_y$ and $\Pi_x$. The idea of CCR is to choose $\Pi_y$ and $\Pi_x$, so that the OLS estimator is asymptotically efficient when $y_t^*$ is regressed on $X_t^*$. This requires

$$\Pi_y = \Sigma^{-1} \Gamma_2 \gamma + (0, \Omega_{12} \Omega_{22}^{-1})',$$

$$\Pi_x = \Sigma^{-1} \Gamma_2.$$

In practice we need to estimate long-run covariance parameters. And then the estimated $\Pi_y$ and $\Pi_x$ are use to transform $y_t$ and $X_t$. As long as the long run parameters are estimated consistently, the resultant CCR estimator is asymptotically efficient. To estimate long-run covariance parameters $\Omega$ and $\Gamma$, we use Andrews(1991)'s quadratic spectral(QS) kernel. Let $\Phi(\tau) = E(w_t w_{t-\tau})$,

$$\Phi_F(\tau) = \frac{1}{T} \sum_{t=\tau+1}^{T} \hat{w}_t \hat{w}_{t-\tau} \text{ for } \tau \geq 0,$$
and $\Phi_T(\tau) = \Phi_T(-\tau)'$ for $\tau < 0$, where $\hat{w}_t$ is constructed from a consistent estimate of the cointegrating vector. The estimator for $\Omega$ has the form

$$\Omega_t = \frac{T}{T-p} \sum_{\tau = -T+1}^{T-1} k\left(\frac{T}{S_T}\right) \Phi_T(\tau),$$

where $k(\cdot)$ is a real-valued kernel, and $S_T$ is a bandwidth parameter. The factor $T/(T-p)$ is a small sample degrees of freedom adjustment. Similarly, $\Gamma$ is estimated by

$$\Gamma_t = \frac{T}{T-p} \sum_{\tau = 0}^{T-1} k\left(\frac{T}{S_T}\right) \Phi_T(\tau).$$

To estimate the unknown bandwidth parameter for a variety of kernels, Andrews proposes automatic bandwidth estimators in which these unknown parameters are estimated from the data. The first step is to use a parametric approximation to estimate the law of motion of the disturbance $w_t$. The second step is to calculate the parameters for the optimal bandwidth parameter from the estimated law of motion. Andrews uses a AR(1) parameterization for each term of the disturbance. Andrews’s QS kernel has certain asymptotic optimal properties.

The second step of our estimation procedure is to use Hansen(1982)’s GMM. Let $\{X_t: t=1,2,\ldots\}$ be a collection of random vectors $X_t$, $\beta_0$ be a p-dimensional vector of the parameters to be estimated, and $f(X_t, \beta)$ a q-dimensional vector of functions. Assume that $X_t$ is stationary. We refer to $u_t = f(X_t, \beta)$ as the disturbance of GMM.
Consider the unconditional moment restrictions

\[ E(f(X_t, \beta)) = 0 \tag{2.18} \]

Suppose that a law of large numbers can be applied to \( f(X_t, \beta) \) for all admissible \( \beta \), so that the sample mean of \( f(X_t, \beta) \) converges to its population mean:

\[ \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} f(X_t, \beta) = E(f(X_t, \beta)) \]

with probability one. The basic idea of GMM estimation is to mimic the moment restrictions (2.18) by minimizing a quadratic form of the sample means

\[ J_T(\beta) = \left\{ \frac{1}{T} \sum_{t=1}^{T} f(X_t, \beta) \right\}^\prime \ W_T \left\{ \frac{1}{T} \sum_{t=1}^{T} f(X_t, \beta) \right\} \tag{2.19} \]

with respect to \( \beta \). \( W_T \) is a positive semidefinite matrix, which satisfies

\[ \lim_{T \to \infty} W_T = W_0, \]

with probability one for a positive definite matrix \( W_0 \).

\[ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t \] has an asymptotic normal distribution with mean zero and the covariance matrix \( \Omega \) in large samples. If \( u_t \) is serially uncorrelated, \( \Omega = E(u_t u_t^\prime) \). If \( u_t \) is serially correlated,
\[ \Omega = \lim_{j \to \infty} \sum_{j} E(u_j u_j'). \]

When \( W_0 = \Omega^{-1} \), \( \sqrt{T}(\beta_T - \beta_0) \) approximately has a normal distribution with mean zero and the covariance matrix

\[ \text{Cov}(\Omega^{-1}) = \left\{ \Gamma' \Omega^{-1} \Gamma \right\}^{-1}, \]

in the large samples, where \( \Gamma = E(\partial f(X_t, \beta)/\partial \beta') \) which is \( q \times p \) matrix and has a full column rank.

The first stage GMM estimator is obtained by setting \( W_0 = I \), and then \( \Omega_T \) is estimated from the first stage GMM estimate \( \beta_T \). The second stage GMM estimator is formed by setting \( W_T = \Omega_T^{-1} \). This procedure can be iterated by using the second stage GMM estimate to form the weighting matrix for the third stage GMM estimator, and so on. It is known that the GMM estimator and test statistics have better small sample properties when this procedure is iterated. In some nonlinear model, it is recommended that the third stage GMM estimate be used because the gains from further iterations may be small.

In the case where there are overidentifying restrictions \( q > p \), a chi-square statistics can be used to the overidentifying restrictions. Hansen shows that \( T \) times the minimized value of the objective function, \( T J_T(\hat{\beta}_T) \), has an asymptotic chi-square distribution with \( q-p \) degrees of freedom if \( W_0 = \Omega^{-1} \).
2.7 Empirical Results

In this section, we present empirical results for the single equation procedure. Monthly end-of-period foreign exchange rates from the International Financial Statistics (IFS) are used. The foreign exchange rates are the domestic price of one unit of foreign currency. In each regression, the United States is regarded as the foreign country, while other countries are the domestic countries. Monthly CPI is used to measure prices in the model. The sample period is from January 1974 to June 1995.

For each country, we report results for two cases. The first case is when prices are measured without error, which leads to the case of the known cointegrating vector. The second case is that of the measurement error model of Section 2.5, in which the cointegrating vector for domestic prices, exchange rates and foreign prices is not restricted to be (1,-1,-1). For the latter case, the two-step method described in Section 2.4 is used. In the first step, we use CCR to obtain long-run coefficients in PPP relations, and then in the second step, we apply GMM to estimate the short-run coefficient.

Since the first case of long-run relationship assumes that PPP holds with the cointegrating vector of \((I,-I,-I)\), we present the literature about the results of empirical tests for this relationship. Rogoff (1995) gives an excellent survey.

The law of one price states that once prices are converted to a common currency, the same good should sell for the same price in different countries. The no arbitrage principle leads to the law of one price. If goods market arbitrage enforces broad parity in prices across a sufficient range of individual
goods according to the law of one price, then there should be a high correlation in the aggregate price level. This is regarded as a long-run equilibrium, which implies PPP.

To test PPP, that is, the long-run relationship between domestic prices, foreign prices and exchange rates, many empirical studies use the real exchange rate, defined as the nominal exchange rate adjusted for differences in national price levels. If real exchange rates are stationary with the cointegrating vector \((1,-1,-1)\), then it is regarded as an evidence that supports PPP. Although many studies show that PPP does not hold, recent research finds favorable results for PPP. However, there is a problem with these recent results. Many studies, which involve a short-run adjustment term, report that the speed of convergence to PPP is extremely slow, and the deviations from PPP appear to damp out at a rate of roughly 15% per year. Furthermore, the short-run deviations from PPP are large and volatile.

One explanation for the evidence against PPP is based on the empirical tests of the law of one price. With examples of traded goods and nontraded goods, Rogoff (1995) shows that the law of one price does not hold if we apply it to nontraded goods. Rogoff shows an example with McDonald’s hamburgers. They have some components which are traded goods, but the restaurant space and labor inputs are not traded internationally. There can be many tariff and nontariff barriers which cause PPP not to hold. He cites data which shows that the prices of Big Mac hamburger range from $5.20 in Switzerland to $1.05 in China. However, for gold, which is a highly traded good, the law of one price holds very well. The price of gold ranges from $379.35 in Hong Kong to $378.81 in Paris, per one troy ounce. Many studies show that disaggregated data of standard industrial classification
categories of exports for US, Canada, Japan, Germany finds evidence against the law of one price. Rogoff argues that the deviations from the law of one price are large, persistent, and are highly correlated with exchange rate movements. Citing other studies, he shows that international price volatility is larger than intranational price volatility. Relative price differentials for similar goods are more volatile and persistent internationally. Another study finds evidence that the volatility of deviations from the law of one price, even among highly traded goods, has been large over several hundred years. From Rogoff’s survey, we can conclude that, in general, the deviations from the law of one price have been large and persistent between countries for a very long time. This result implies that absolute PPP may not hold because absolute PPP is built on the law of one price.

Another explanation for the evidence against PPP is based on examination of price measurements. For a broader measure of international price differential, absolute PPP measures are used. Consumer price index, wholesale price index and production price index represent the aggregate measure of general price levels in a country. These indices are not constructed for an internationally standardized basket of goods, and they do not give any information on the deviation from the PPP in the base year. Summers and Heston’s (1991) International Comparison Program (ICP) dataset is a result of developing more reasonable absolute PPP measure. Rogoff (1995), however, contends that there remain problems with ICP dataset because these are gathered infrequently at five-year intervals for benchmark surveys and countries, and the data is filled in largely by extrapolation for nonbenchmark years and countries. Another problem is the long lag between the time the data is gathered and the time it is made widely
available.

We are more interested in recent studies which show that PPP holds in the long-run, although there are short-run deviations from the PPP. In our study in this chapter, the long-run relationships are assumed to exist. Otherwise, we would have nothing to talk about the short-run adjustment. In the literature on time-series analysis, long-run equilibrium is defined as a systematic co-movement among economic variables. A shock to the system causes a deviation from the long-run relationship. But if there is no more shock, then the system converges to the long-run relationship over time. Therefore, short-run can be defined as a period in which a deviation from the long-run relationship exists without there being any more shock. PPP is usually regarded as a long-run relationship rather than a short-run relationship. Thus, if there is no long-run equilibrium we cannot define the short-run adjustment. In our specification of Mussa’s (9182) model we allow short-run deviations from the long-run relationship, implicitly assuming that there exist long-run relationships. There are three recent approaches that obtain favorable results for PPP.

One approach argues that if PPP deviations damp out sufficiently slowly, then one may require many decades of data to be able to reliably reject the existence of a unit root in real exchange rates, because one may fail to reject a random walk for real exchange rate due to a lack of power. Abuaf and Jorion (1990) used 1900-1972 annual data for eight currencies, and found strong rejections of the random walk model for Canada, France, Italy, Japan, Netherlands and UK. For Germany and Switzerland, they fail to find the cointegrating vector to be (1, -1, -1). Their estimates suggest a half life for PPP deviations of 3.3 years, using the average slope coefficient of an AR(1) specification for real
exchange rates. Glen (1992) finds similar results for nine bilateral rates over the years 1900-1987. Diebold, Husted and Rush (1991) look at data from the gold standard period, with data samples ranging from 73 to 123 years. They reject the random walk model and suggests an average half life of 2.8 years for exchange rates across the six countries in their sample. Lothian and Taylor (1995) test the random walk hypothesis on two centuries of data for dollar/pound (1791-1990) and the franc-pound (1803-1990) exchange rates. They reject the random walk model, and find strong evidence of mean reversion in both rates with an estimated half life of 4.7 years for the dollar/pound and 2.5 years for the franc/pound rate. Oh (1996) also finds evidence that PPP holds in the lon-run, using panel data.

Rogoff (1995) argues that, since these results are based on the mixed data of fixed and floating regimes, they may not be enough to confirm PPP over only the floating exchange rate period. He cites Mussa (1986), who demonstrates that real exchange rates tend to be much more volatile under floating than under fixed exchange rates, and that the econometric implications of mixing data from the two regimes is unclear.

Another approach to test for PPP is to use panel data to enhance the power of unit root tests. Frankel and Rose (1996) examine a panel data set, including annual data for the years 1948-1992 for 150 countries. They reject the random walk model, even using only post-1973 floating data. Their results suggest an estimated half life for PPP deviations of about 4 years.

The third approach exploits the implication of Samuelson-Balassa model which states that the existence of non-traded goods can allow for cointegrating vector different from (1, -1, -1). Samuelson-Balassa model can be regarded as a theoretical background for the measurement error models.
Taylor (1996) investigates PPP since the late nineteenth century for a sample of twenty countries. He obtains test results for each country using ECMs with annual panel data. We cite four of his specifications of the ECMs.

In his first specification, the left-hand side variable is the difference of the dollarized price index, defined as the difference between the logarithm of domestic price and the logarithm of exchange rate. The right-hand side variables include a constant, one period lagged real exchange rate and one period lagged difference of the dollarized price. The test depends on the t-statistic of the error correction coefficient. The cointegrating vector is prespecified as (1, -1, -1). For Canada, France and UK, the estimate for the error correction coefficient is significant at the 5% level. However, Germany, Italy, Japan, Netherlands and Switzerland fail the test.

In his second specification, the left-hand side variable is a vector which involves domestic price, exchange rate and foreign price. The right-hand side variables include p lags of the left-hand side vector. In this test, Canada fails the test, but all other countries including France, Germany, Italy, Japan, Netherlands, Switzerland and UK pass the test.

Taylor also tests a measurement error model in an ECM nesting the no measurement error model with the cointegrating vector (1, -1, -1). He finds cointegration for Canada, France, Italy, Japan, and UK. Canada shows that the measurement error model is more appropriate, but France, Italy, Japan, and UK have measurement errors which are not significant at the 5% level. Germany, Netherlands and Switzerland do not show cointegrating relationships regardless of the measurement errors. However, with twenty countries, he concludes that the global restriction of the cointegrating vector (1, -1, -1) appears to be
unjustified.

He adds another test, arguing that if cross-country heterogeneity must be
admitted to the analysis of panel dynamics, then attempts to raise the power of
cointegration tests by using a panel instead of individual time series must take
into account the modifications necessary to construct valid test statistics and
finite sample distributions for such special cases. Thus, he uses cointegration
tests for panels with heterogeneous slopes following Pedroni (1995). In this
study, he shows that the slope terms are indeed heterogeneous and differ from
unity, which implies the cointegrating vector \((1, -1, -1)\). In particular, Japan
and Switzerland have estimates most different from unity.

Cheung and Lai (1993) examine PPP, allowing for measurement errors, during
the recent floating exchange rate period. In their study, US is the home country,
and Canada, France, Germany, Switzerland and UK are the foreign countries. The
data covers the period from January 1974 to December 1989. They estimate the
equation

\[
e(t) = \alpha_0 + \alpha_1 p^m(t) + \alpha_2 p^*_m(t) + \varepsilon(t). \tag{2.20}
\]

With residual-based tests for cointegration, they fail to reject the null of
no cointegration. However, using Johansen's likelihood test for the hypothesis of
at most \(r\) linearly independent cointegrating vectors, the hypothesis of no
cointegrating vector \((r=0)\) can be rejected at the 5% level. The estimates for \(\alpha_1\)
and \(\alpha_2\) are, respectively, 25.422 and 23.714 for Canada, 3.826 and 3.521 for
France, 7.636 and 14.972 for Germany, 1.445 and 5.079 for Switzerland, and 1.426
and 1.035 for UK. This result shows that each country has different measurement
error coefficients.

MacDonald (1993) uses data from January 1974 to June 1990 to reexamine PPP, allowing for measurement errors. His specification of PPP is the same as that of Cheung and Lai’s, but no constant is included implying the absolute version of PPP. Using Johansen tests, he shows that there would be two cointegrating vectors for UK, and one cointegrating vector each for France, Germany and Japan, when consumer prices are used. For Canada, there is no cointegration when consumer prices are used, but one cointegrating vector when wholesale prices are used.

Table 2.1 presents the results of PPP cointegrating regressions with measurement error. We report the third stage estimates of CCR for the coefficients and the fourth stage test results. We use Andrews’s (1991) QS kernel to estimate long-run covariance parameters.

For the Canadian dollar, H(0,1) test shows that the deterministic cointegrating restriction is rejected at the 5 percent level. But H(1,2) test is not rejected at the 5 percent level, although H(1,3) test rejects the long-run relationship with measurement error. Taylor (1996) shows that the cointegrating relationship of the measurement error model exists for the Canadian dollar. Cheung and Lai (1993) also reject the null of no cointegration and support the measurement error model. MacDonald (1993) finds no cointegration when he uses CPI in a measurement error model, but detects one cointegrating vector when the wholesale price index is used.

For France, PPP between US and France is strongly rejected when we use the measurement error model. Taylor (1996) presents that for France there exists the cointegrating relationship with cointegrating vector (1, -1, -1) rather than with the measurement error vector. Cheung and Lai (1993) reject the null of no
cointegration. MacDonalld (1993) finds one cointegrating vector when he uses CPI in a measurement error model.

For the German mark, both the deterministic and stochastic cointegrating restrictions are rejected at the 5 percent level. However, they are marginally accepted at the 1 percent level. Taylor (1996) fails to find the cointegrating relationship of the measurement error model, which nests the cointegrating vector \( (1, -1, -1) \). However, he finds the cointegrating relationship when he uses an ECM with prespecified cointegrating vector \( (1, -1, -1) \). Cheung and Lai (1993) also reject the null of no cointegration, and their test results support the measurement error model. MacDonalld (1993) detects one cointegrating vector when he uses CPI in a measurement error model.

For Italy, the stochastic cointegrating restrictions are satisfied at the 5 percent level while the deterministic cointegrating restriction is satisfied at 1 percent level of p-value. Taylor (1996) shows that Italian lira has a cointegrating relationship with US dollar when the measurement error model is used. However the measure ment error is not significant implying that the cointegrating vector \( (1, -1, -1) \) is appropriate. In a ECM specification with this prespecified vector, he finds the evidence which is favorite to the long-run relationship.

For the Japanese yen, the deterministic cointegrating restriction is satisfied, but the stochastic cointegrating restriction is rejected. Taylor’s (1996) test results support the cointegrating relationship with cointegrating vector \( (1, -1, -1) \), when he uses an ECM with the prespecified cointegrating vector. The measurement error model has a cointegrating relationship, but the measurement error coefficient is not significant, which confirms the test results.
of an ECM with prespecified cointegrating vector. However, when he applies Pedroni's (1995) approach using panel data, he finds a cointegrating vector which is different from the cointegrating vector (1, -1, -1). MacDonald (1993) finds one cointegrating vector, using CPI in a measurement error model.

For Netherlands, PPP between US and Netherlands prices is not rejected at the 5 percent level when we apply H(0,1) and H(1,2) tests, but H(1,3) test rejects the stochastic cointegrating restriction. Taylor (1996) fails to find the cointegrating relationship when the measurement error model is used. Using an ECM with the prespecified cointegrating vector (1, -1, -1), however, he finds the long-run relationship.

For Switzerland, PPP between US and Switzerland holds when measurement errors are assumed to exist. Taylor (1996) obtains a similar results for Switzerland to those for Netherlands. He fails to find the cointegrating relationship when the measurement error model is used. Using an ECM with the prespecified cointegrating vector (1, -1, -1), however, he finds the long-run relationship. His test results are mixed because in a panel data approach he finds the cointegrating relationship which is very different from the cointegrating vector (1, -1, -1). Cheung and Lai (1993) support the cointegration relationship using a measurement error model.

For UK, There exists a long-run relationship. The deterministic and the stochastic cointegrating restrictions are satisfied at the 5 percent level. The coefficients of measurement errors for the UK and US prices are precisely measured. Taylor (1996) finds the cointegrating relationship between British pound and US dollar in any specification he uses. However, his test results are more favorable to the cointegrating vector (1, -1, -1) than to the measurement
error model. Cheung and Lai(1993) also reject the null of no cointegration and their test results support the measurement error model. MacDonald(1993) finds that there is one cointegrating vector when he uses CPI in a measurement error model.

In our test results, only US-Switzerland and US-UK exchange rates do not reject the PPP relationship with measurement errors, implying that, in general, PPP is rejected over the floating exchange rate period.

However, since our primary interest is to present a procedure of estimating the structural speed of price adjustment in a simple economic model, we use the estimates for the cointegrating vectors as given in order to do a sensitivity analysis. To show our instrumental variable approach, we present an illustration of an exchange rate model with the known cointegrating vector as a base case. The measurement error model that we estimate can be used for a sensitivity analysis. On the other hand, the estimates for the short-run coefficient, given long-run coefficient, can have some impication for the long-run relationship. If we can obtain reasonable estimates for the speed of price adjustment in the second step with the given estimates for the long-run relationship, it can be interpreted as an evidence for the PPP. If PPP does not hold with the given estimates for the cointegrating vector, then we expect to obtain a negative estimate for the short-run coefficient. This reasoning also applies to the case of the known cointegrating vector of (1, -1, -1). If PPP holds with this known cointegrating vector, then we expect a positive estimates for the short-run coefficient.

Table 2.2 reports the results of the GMM estimation for the short-run adjustment coefficient in the structural ECM. The insrumental variables are $p^m(t-3)$ and $p^m(t-4)$, which are US prices. For each country, the first row in
Table 2.2 reports the GMM results for the case of the known cointegrating vector (1, -1, -1), and the second row shows the GMM results for the case of the unknown cointegrating vector. To obtain the half life of each estimate for \( b \), we rearrange the ECM equation as an AR(1) process for the real exchange rates.

In most cases, we obtain positive point estimates for the structural speed of adjustment coefficient \( b \). In cases where the point estimates are negative, they are not significantly different from zero at the 5 percent level.

For Canada, the half life is 4.6 years when the cointegrating vector is known, and 3.5 years when it is estimated. The estimate for \( b \) is significant at the 5 percent level when we use the estimated cointegrating vector.

For France, the half life is 3.8 years for the known cointegrating vector (1, -1, -1), and 1.7 years for the estimated cointegrating vector. Neither model is rejected, giving similar p-values when we apply Hansen's (1982) J-test. But the coefficient of the short-run dynamics is not measured precisely when we use the known cointegrating vector, although the estimate is positive.

For Germany, the half life is 5.2 years when the known cointegrating vector is used. With the measurement error model, we obtain an estimate with wrong sign for the short-run coefficient, but the large standard error implies statistical insignificance of the estimate.

For Italy, the estimates for the speed of adjustment are negative, but they are insignificant for both cases.

For Japan, we obtain a half life of 15.6 years for the case of the known cointegrating vector, and 1.4 years for the case of the unknown cointegrating vector. However, the estimate in the first case is insignificant, whereas the estimate in the second case is significant at the 5 percent level.
For Netherlands, the half life is 3.1 years when we use the cointegrating vector (1, -1, -1), and 1.3 years when we assume measurement error. The measurement error model gives a statistically significant estimate.

For Switzerland, the half life is 2.3 years when we do not assume measurement error. When we use the measurement error model, the estimate for the short-run coefficient is negative. Both estimates are insignificant.

For UK, the half life is 2.8 years for the cointegrating vector (1, -1, -1), and 5.4 years for the estimated cointegrating vector, although neither estimate is measured precisely.

2.8 Conclusion

To estimate the speed of price adjustment in dynamic models of exchange rate determination many studies employ Error Correction Models. We point out that the use of the traditional ECM may fail to identify the structural parameters governing the short-run dynamics, unless strong assumptions on the coefficients of the ECM are imposed. To overcome this shortcoming, we suggest a structural ECM. This exploits the different speeds of convergence of long-run and short-run estimators, as in Engle and Granger (1987). The structural ECM does not impose any arbitrary assumptions on the coefficient matrix of ECM, but instead reflects theoretical restrictions in a structural economic model. Thus, it avoids the Cooley and Leroy's (1985) critique on estimating structural parameters from reduced form VARs.

We apply an instrumental variable method to the structural ECM, derived
from a version of Mussa's (1982) dynamic model of exchange rate determination. When the cointegrating vector is known, we substitute it into the structural ECM, and apply an instrumental variable method to obtain an estimate for the structural speed of price adjustment. When the cointegrating vector is unknown, we first estimate the long-run coefficients in the structural model, and then apply an instrumental variable method to the ECM with the estimated long-run coefficients.

With a known cointegrating vector, we obtain positive estimates for the short-run coefficient for most countries, except Italy. The estimates imply half lives of 2.3 to 5.2 years, except Japan. With the unknown cointegrating vectors, which reflect measurement errors, we also obtain positive estimates for the short-run coefficients for most countries. When the estimates are positive, these are significant and show shorter half lives ranging 1.3 to 3.5 years except UK, for which the estimate is insignificant. For both known and unknown cointegrating vectors, negative estimates for the short-run coefficient are always insignificant.
<table>
<thead>
<tr>
<th>Country</th>
<th>$\psi_0$</th>
<th>$\phi$</th>
<th>$\phi/\phi^*$</th>
<th>$H(0,1)$</th>
<th>$H(1,2)$</th>
<th>$H(1,3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>Canada</td>
<td>0.005</td>
<td>0.029</td>
<td>1.014</td>
<td>5.377</td>
<td>2.911</td>
<td>8.443</td>
</tr>
<tr>
<td></td>
<td>(0.169)</td>
<td>(0.122)</td>
<td>(0.024)</td>
<td>(0.020)</td>
<td>(0.088)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>France</td>
<td>-0.899</td>
<td>0.221</td>
<td>1.106</td>
<td>104.0</td>
<td>30.17</td>
<td>41.67</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.017)</td>
<td>(0.014)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Germany</td>
<td>2.796</td>
<td>0.073</td>
<td>0.376</td>
<td>5.402</td>
<td>5.571</td>
<td>6.068</td>
</tr>
<tr>
<td></td>
<td>(0.220)</td>
<td>(0.022)</td>
<td>(0.054)</td>
<td>(0.020)</td>
<td>(0.018)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>Italy</td>
<td>-4.283</td>
<td>0.192</td>
<td>1.621</td>
<td>4.134</td>
<td>1.355</td>
<td>3.264</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.020)</td>
<td>(0.028)</td>
<td>(0.042)</td>
<td>(0.244)</td>
<td>(0.196)</td>
</tr>
<tr>
<td>Japan</td>
<td>0.509</td>
<td>0.156</td>
<td>0.728</td>
<td>0.204</td>
<td>6.963</td>
<td>22.45</td>
</tr>
<tr>
<td></td>
<td>(0.370)</td>
<td>(0.039)</td>
<td>(0.042)</td>
<td>(0.652)</td>
<td>(0.008)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1.548</td>
<td>0.141</td>
<td>0.651</td>
<td>2.871</td>
<td>0.208</td>
<td>19.81</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.036)</td>
<td>(0.021)</td>
<td>(0.090)</td>
<td>(0.648)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Switzerland</td>
<td>2.929</td>
<td>0.027</td>
<td>0.346</td>
<td>1.763</td>
<td>3.460</td>
<td>3.473</td>
</tr>
<tr>
<td></td>
<td>(0.221)</td>
<td>(0.031)</td>
<td>(0.052)</td>
<td>(0.184)</td>
<td>(0.063)</td>
<td>(0.176)</td>
</tr>
<tr>
<td>U.K.</td>
<td>-1.902</td>
<td>0.369</td>
<td>1.458</td>
<td>0.003</td>
<td>2.686</td>
<td>5.078</td>
</tr>
<tr>
<td></td>
<td>(0.144)</td>
<td>(0.071)</td>
<td>(0.030)</td>
<td>(0.958)</td>
<td>(0.101)</td>
<td>(0.079)</td>
</tr>
</tbody>
</table>

\[ p^m(t) = \psi_0 + \phi \epsilon(t) + (\phi/\phi^*)p^m(t) + \zeta(t) \]

Column (1) : domestic countries
Column (2)-(4) : Standard errors are in parentheses.
Column (5)-(7) : P-values are in parentheses.

Table 2.1: Purchasing Power Parity
<table>
<thead>
<tr>
<th>Country (1)</th>
<th>$\phi$ (2)</th>
<th>$\phi/\phi^*$ (3)</th>
<th>$d$ (4)</th>
<th>$b$ (5)</th>
<th>$\chi^2$ (6)</th>
<th>DW (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>1</td>
<td>1</td>
<td>-0.0036</td>
<td>0.0126</td>
<td>1.712</td>
<td>2.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0024)</td>
<td>(0.0104)</td>
<td>(0.634)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.029</td>
<td>1.014</td>
<td>0.0005</td>
<td>0.0166</td>
<td>4.688</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td></td>
<td>(0.0050)</td>
<td></td>
<td>(0.196)</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>1</td>
<td>1</td>
<td>-0.0268</td>
<td>0.0151</td>
<td>2.113</td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0224)</td>
<td>(0.0125)</td>
<td>(0.549)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.221</td>
<td>1.106</td>
<td>-0.0300</td>
<td>0.0335</td>
<td>2.324</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>(0.0131)</td>
<td></td>
<td>(0.0146)</td>
<td></td>
<td>(0.508)</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>1</td>
<td>1</td>
<td>-0.0042</td>
<td>0.0110</td>
<td>5.793</td>
<td>1.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0063)</td>
<td>(0.0110)</td>
<td>(0.122)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.073</td>
<td>0.376</td>
<td>-0.0039</td>
<td>-0.0018</td>
<td>6.531</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>(0.0093)</td>
<td></td>
<td>(0.0033)</td>
<td></td>
<td>(0.088)</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>1</td>
<td>1</td>
<td>0.0701</td>
<td>-0.0094</td>
<td>1.612</td>
<td>1.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.1880)</td>
<td>(0.0256)</td>
<td>(0.657)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.192</td>
<td>1.621</td>
<td>0.0058</td>
<td>-0.0014</td>
<td>9.025</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>(0.0423)</td>
<td></td>
<td>(0.0099)</td>
<td></td>
<td>(0.029)</td>
<td></td>
</tr>
</tbody>
</table>

$$\Delta p^m(t+1) = d - b /p^m(t) - \phi e(t) - (\phi/\phi^*) p^m(t) + (\phi/\phi^*) \Delta p^m(t+1) + \phi \Delta e(t+1) + w(t+1)$$

Column (1): domestic countries
Column (2)-(3): Standard errors are in parentheses.
Column (4): P-values are in parentheses.
Column (5): Durbin-Watson test for GMM residuals

Table 2.2: Instrumental Variable Estimation: Single Equation
(continued on the next page)
Table 2.2 (continued)

<table>
<thead>
<tr>
<th>Country</th>
<th>( \phi )</th>
<th>( \phi/\phi^* )</th>
<th>( d ) (2)</th>
<th>( b ) (3)</th>
<th>( \chi^2 ) (4)</th>
<th>DW (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan</td>
<td>1</td>
<td>1</td>
<td>-0.0143</td>
<td>0.0037</td>
<td>5.552</td>
<td>1.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0488)</td>
<td>(0.0096)</td>
<td>(0.136)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.156</td>
<td>0.728</td>
<td>0.0209</td>
<td>0.0398</td>
<td>6.374</td>
<td>1.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0083)</td>
<td>(0.0164)</td>
<td>(0.095)</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>1</td>
<td>1</td>
<td>-0.0110</td>
<td>0.0182</td>
<td>0.074</td>
<td>1.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0071)</td>
<td>(0.0106)</td>
<td>(0.995)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.141</td>
<td>0.651</td>
<td>0.0694</td>
<td>0.0445</td>
<td>2.788</td>
<td>1.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0197)</td>
<td>(0.0127)</td>
<td>(0.426)</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>1</td>
<td>1</td>
<td>-0.0098</td>
<td>0.0248</td>
<td>1.471</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0081)</td>
<td>(0.0162)</td>
<td>(0.689)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.027</td>
<td>0.346</td>
<td>-0.0108</td>
<td>-0.0041</td>
<td>11.70</td>
<td>1.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0071)</td>
<td>(0.0024)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>U.K.</td>
<td>1</td>
<td>1</td>
<td>0.0105</td>
<td>0.0206</td>
<td>2.980</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0085)</td>
<td>(0.0197)</td>
<td>(0.395)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.369</td>
<td>1.458</td>
<td>-0.0208</td>
<td>0.0106</td>
<td>4.806</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0217)</td>
<td>(0.0115)</td>
<td>(0.187)</td>
<td></td>
</tr>
</tbody>
</table>

\[
\Delta p^m(t+1) = d - b \left[ p^m(t) - \phi e(t) - (\phi/\phi^*)p^* m(t) \right] \\
+ (\phi/\phi^*)\Delta p^* m(t+1) + \phi \Delta e(t+1) + w(t+1)
\]

Column (1) : domestic countries
Column (2)-(3) : Standard errors are in parentheses.
Column (4) : P-values are in parentheses.
Column (5) : Durbin-Watson test for GMM residuals
CHAPTER 3

STRUCTURAL ERROR CORRECTION MODELS:
A SYSTEM METHOD

3.1 Introduction

In the single equation method, we consider only the goods market imposing no theoretical restrictions on other markets, such as money and bond markets. In the system method, we consider all these markets together. Hansen and Sargent’s (1982) method, which applies Hansen’s (1982) Generalized Method of Moments (GMM) to linear rational expectations models, is combined with our single equation method. The system method is more efficient than the single-equation method, when the restrictions implied by linear rational expectations models are true. On the other hand, the single equation method is robust to misspecification in the other equations of the structural ECM.

Structural ECMs have been considered by several authors. Urbain (1992) investigates sufficient conditions for weak exogeneity for structural ECMs that are similar to ours. Hsiao (1995) discusses the relationship between the ECM and structural simultaneous equations models. However, unlike Urbain and Hsiao, we do not assume that exogenous variables are observed by the econometrician. In
our empirical application, it is not attractive to assume the exogeneity of any variable in the cointegrated system. Papell (1995) derives a reduced form ECM from an exchange rate model that is similar to ours. He applies Phillips (1991) ML estimator to the reduced form ECM.

Dolado, Galbraith and Banerjee (1991) and Gregory, Pagan and Smith (1993) derive structural ECMs from linear quadratic models. They discuss the difficulties associated with the application of standard estimation methods such as Engle and Granger’s (1987) two-step method and Johansen’s (1988, 1991) maximum likelihood (ML) method to the ECM. Their IV methods are similar to our single equation method in that a cointegrating vector is estimated in the first step, but are different in that they apply IV methods to the Euler equation of linear quadratic models, whereas our single equation method applies IV methods to the slow adjustment equation. As a result of this difference, our method can be applied to a broader range of structural ECMs. Another difference is that they do not combine their method with Hansen and Sargent’s (1982) IV method for linear rational expectations models, whereas our system method combines the single equation method with Hansen and Sargent’s procedure.

We apply the system method to the same data set as in chapter 2. In the system method, the speed of adjustment coefficient can be estimated from the slow adjustment equation for the domestic price, and from the rational expectations equation for the exchange rate. We form a specification test by comparing the estimates for the speed of adjustment coefficient from these two equations.

The rest of this chapter is organized as follows. Section 3.2 surveys the relevant literature. In section 3.3, the system method is proposed. In section 3.4, the model in section 2.2 is closed by adding the money market equilibrium
condition and the uncovered interest rate parity condition. In section 3.5, the system method is applied to the model in section 3.4 and empirical results are presented. Section 3.6 contains concluding remarks.

3.2 Literature Survey

Dolado, Galbraith and Banerjee (1991) examines the stochastic properties of the variables and parameter estimates which appear in linear Euler equations derived from the simple intertemporal quadratic adjustment model. Especially, they concentrate on the consequences of assuming that the forcing variables contain unit roots rather than being stationary about deterministic trends, and on the implications of this hypothesis for the use of consistent, but not asymptotically efficient, methods of estimation. They show that, where series of interest are integrated of order one or greater, we can use the information gained from pre-testing for the order of integration to improve the specification and estimation procedure. In particular, they show that we can avoid assumptions such as those of knowledge of the discount factor, or the forms of the process generating the forcing variables. They argue that Euler equation estimation need not yield accurate estimates of critical parameters, especially if the integration properties of the data are disregarded. By assuming that the forcing variables are integrated, they characterize the order of the control variable and of the deviations from the target. Their approach is based on the two step approach in Engle and Granger (1987). They exploit different convergence rates for estimators of parameters of variables exhibiting different degrees of
integration. They first consistently estimate the long-run coefficients and then apply instrumental variable methods to estimate other variables. They assume that the forcing variables are strictly exogenous.

Gregory, Pagan and Smith (1992) criticize that Dolado, Galbraith and Banerjee's approach cannot identify other parameters even though one parameter is assumed to be known. Gregory, Pagan and Smith also study estimation in the linear-quadratic optimization model when the forcing variables are integrated. They demonstrate that there is frequently a lack of identification for some of the parameters in linear quadratic optimization models with integrated forcing variables. It is shown that the ability to identify parameters is dependent on the nature of the forcing variables, and that it is quite likely that one of the parameters, the discount factor, cannot be estimated with much accuracy. Hence, they argue that the common practice of pre-setting this parameter has theoretical support. They fix one of the parameters and proceed to look at how estimators of the remaining parameters can be derived.

They show that by the nature of Euler equation solutions there are two sets of parameters to be estimated; one associated with I(1) variables and the other with I(0) variables. Their study is based on the literature that Euler equations derived from quadratic objective functions can be represented as error correction models. While in Engle and Granger (1987) the ECM is derived as a consequence of cointegration between variables, they derive an ECM as a consequence of a theoretical specification. So, while in Engle and Granger the cointegrating vector is of prime importance and all other parameters are regarded as nuisance parameters, in their study the nuisance parameters become of interest. They argue that previous literature largely concentrate upon the
estimation of the parameters attached to I(1) variables, the cointegrating vector, and that standard procedures for this have undesirable side effect on the estimation of the I(0) variable parameters.

They devise two general classes of estimators for the unknown parameters, corresponding to the two procedures normally used for estimating models with rational expectations. For the first class, they replace the unknown expectation in the Euler equation with the observed value of the variables expectations are being formed about. For the second class, they solve for the expectation and obtain an error correction model. In a simple simulation study they find that the ECM estimator of the long run response generally works well, while Euler equation estimators are much less reliable. Although an Euler equation estimator is preferred in many rational expectation estimation contexts since it does not require the specification of an expectation generating process and hence it exhibits a degree of robustness, they argue that the robustness property is much less advantageous for determining the long-run response when there are integrated regressors. Since the best estimator of the adjustment cost parameters frequently seems to come from the Euler equation, while the best estimator for the long-run response is that from ECM, they conclude that using each estimator for the task at which it has a comparative advantage is recommended.

Hsiao (1995) provides a basic framework linking the multiple time series model and the dynamic simultaneous equation model. He also presents conditions for identifying both the short-run dynamics and long-run equilibrium conditions. He demonstrates that conditions for identification derived for stationary variables also hold for integrated variables under appropriate assumptions. He argues that nonstationarity is not a substitute for exogeneity and the classical
concern of simultaneous bias remains a legitimate one despite the regressors being integrated. The long-run reduced form coefficients obtained by solving the structural parameters converge at the speed of T and the limiting distribution is a mixed normal. He shows that cointegration makes the limiting distribution of the conventional estimators for structural parameters singular, and provides results which can be used as a basis to transform structural parameters such that the transformed parameterization with appropriate scaling will have a nonsingular limiting distribution. He proves that the dynamic simultaneous equation model is identified, if and only if its structural error correction representation and its implied long-run equilibrium and short-run dynamics are identified because both the long-run coefficients and the short-run response coefficients are linear transformations of the dynamic simultaneous equations model. However, his arguments are based on the assumption that in the structural model there exist exogenous variables which do not respond to the deviation from equilibrium.

Papell (1995) derives a reduced form of Dornbusch (1976) model under the assumption that the exchange rate is cointegrated with its fundamentals. The reduced form is a highly nonlinear combination of the structural coefficients, with the constraints caused by the interplay between the structural equations, the assumptions necessary to derive a unique rational expectations solution, and cointegration. He argues that if the variables which determine the real exchange rate, which are exchange rates, domestic prices and foreign prices, have unit roots, long-run equilibrium requires cointegration between the exchange rate and its fundamentals. He contends that the standard open economy price adjustment specifications, for example Dornbusch (1976) and Mussa (1981), are not consistent with a unit root in the price level. In Mussa (1981), the rate of inflation is
specified as equal to the expected rate of change of the equilibrium price level plus some portion of the difference between the equilibrium and actual price levels. In order to impose a unit root in the price level, he modifies Musa's specification to eliminate the second term. That is, in the long-run, the equilibrium price level equals the exchange rate plus the foreign price level. The model is completed by specifying the exogenous variables, which are foreign prices, foreign output, foreign interest rate and domestic money supply, as unit root processes. By solving the model by the method of undetermined coefficients, he obtains a reduced form which consists of the cointegrating vector, an autoregressive component, and moving average terms. He estimates the cointegrating vector and coefficients for the stationary variables jointly using full system maximum likelihood methods. He investigates the implications of relaxing the assumption of rational expectations, assuming theories consistent expectations, where expected depreciation is qualitatively, but not quantitatively, consistent with the theoretical model. The rational expectation model is always rejected in favor of the theories consistent expectations model at standard significance levels.

We use the two-step estimation procedure following Dolado, Galbraith and Banerjee (1991) and Gregory, Pagan and Smith (1992). However, we derive a structural ECM from a dynamic model of exchange rate determination with sticky price in goods market instead of using a quadratic optimization model. In our specification the forcing variables are not necessarily exogenous and they are allowed to be endogenous policy variables. Our approach is easier to use than Hsiao's (1995) in the sense that our approach exploits the theoretical assumptions on the long-run and short-run dynamics directly allowing for more
intuitive interpretations on the dynamics. Furthermore, our approach does not have the problem of simultaneous bias or because we allow the forcing variables to be endogenous. Our specification is simpler than Papell’s (1995) and has a more general form of ECM for dynamic models.

### 3.3 A System Method for Rational Expectations Models

In this section, we propose an econometric method that combines our single equation method with Hansen and Sargent’s (1982) procedure, imposing nonlinear restrictions implied by rational expectations models.

Let $y(t) = (y_1(t), y_2(t), y_3(t), y_4(t))'$ be a $4 \times 1$ vector of random variables, with an structural ECM representation (2.8) and only one linearly independent cointegrating vector $A$ such that $A'y(t)$ is stationary. In the following, $y(t)$ is partitioned into four subvectors, and each subvector is given a different role. For expositional simplicity, we assume that each subvector is one dimensional so that $y(t)$ is a $4 \times 1$ vector, and that $y(t)$ has only one cointegrating vector.

The first element of $y(t)$ represents a slow adjustment, as in Equation (2.6), so that

$$
\Delta y_1(t+1) = d_1 - c_{02}^1 \Delta y_2(t+1) - \ldots - c_{04}^1 \Delta y_4(t+1) + b_1 A'y(t)
$$

$$
+ c_1^1 \Delta y(t) + c_2^1 \Delta y(t-1) + \ldots + c_p^1 \Delta y(t-p+1) + u_1(t+1),
$$

(3.1)

with nonzero $b_1$, and $E[u_1(t+1)|I_{t-1}] = 0$.

As pointed out by Hansen and Sargent (1982), many linear rational
expectations models imply that one variable is a geometrically declining weighted sum of expected future values of other variables. We assume that the second element of \( y(t) \) is related to a discounted sum of expected future values of the fourth element in the following form:

\[
\Delta y_2(t+1) = d_2 - c_{01}^2 \Delta y_1(t+1) - c_{03}^2 \Delta y_3(t+1) - c_{04}^2 \Delta y_4(t+1)
\]

\[
+ \alpha E \left[ \sum_{j=0}^{\infty} \delta^j \Delta y_4(t+j+1) | I_t \right] + \epsilon_e(t+1)
\]  
(3.2)

where \( \delta \) is a positive constant that is smaller than one, and \( \alpha \) is a constant.

Hansen and Sargent's (1982) methodology is to project the conditional expectation of the discounted sum, \( E[\sum \delta^j \Delta y_4(t+j+1) | I_t] \), onto an information set \( H_t \), which is a subset of \( I_t \), the economic agents' information set. Let \( \hat{E}(\cdot | H_t) \) be the linear projection operator, conditional on an information set \( H_t \), which is a subset of \( I_t \). Replacing the conditional expectation by the linear projection gives

\[
\Delta y_2(t+1) = d_2 - c_{01}^2 \Delta y_1(t+1) - c_{03}^2 \Delta y_3(t+1) - c_{04}^2 \Delta y_4(t+1)
\]

\[
+ \alpha \hat{E} \left[ \sum_{j=0}^{\infty} \delta^j \Delta y_4(t+j+1) | H_t \right] + u_2(t+1),
\]  
(3.3)

where

\[
u_2(t+1) = \epsilon_e(t+1) + E[\sum \delta^j \Delta y_4(t+j+1) | I_t] - \hat{E}\left[ \sum_{j=0}^{\infty} \delta^j \Delta y_4(t+j+1) | H_t \right]. \]  
(3.4)
Because $H_t$ is a subset of $I_t$, we obtain $E[u_2(t+1)\mid H_t] = 0$.

The current and past values of the first difference of the third element of $y(t)$ are used to form the econometrician's information set $H_t$. Since $E[\cdot \mid H_t]$ is the linear projection operator onto $H_t$, there exist possibly infinite order lag polynomials $\beta(L)$, $\gamma(L)$, and $\xi(L)$, such that

\begin{align*}
\hat{E}[\Delta y_3(t+1)\mid H(t)] &= \beta(L)\Delta y_3(t) \\
\hat{E}[\Delta y_4(t+1)\mid H(t)] &= \gamma(L)\Delta y_3(t) \\
\hat{E}\left[\sum_{j=0}^{\infty} \delta^j \Delta y_4(t+j+1)\mid H(t)\right] &= \xi(L)\Delta y_3(t).
\end{align*}

(3.5)  
(3.6)  
(3.7)

Then, following Hansen and Sargent (1980, Appendix A), we obtain the restrictions imposed by (3.7) on $\xi(L)$:

\[
\xi(L) = \frac{\gamma(L) - \delta L^{-1} \gamma(\delta) (I - \delta \beta(\delta))^{-1} (I - L \beta(L))}{I - \delta L^{-1}}
\]

(3.8)

Substituting (3.7) into (3.3) gives the equation

\[
\Delta y_2(t+1) = d_2 - c_{01}^2 \Delta y_1(t+1) - c_{03}^2 \Delta y_3(t+1) - c_{04}^2 \Delta y_4(t+1) \\
+ \alpha \xi(L) \Delta y_3(t) + u_2(t+1),
\]

(3.9)

where $\xi(L)$ is given by (3.8). We now make the additional assumption that the lag
polynomials $\beta(L)$ and $\gamma(L)$ are finite order polynomials, so that

$$\Delta y_3(t+1) = \beta_1 \Delta y_3(t) + \beta_2 \Delta y_3(t-1) + \ldots + \beta_p \Delta y_3(t-p+1) + u_3(t+1) \quad (3.10)$$

$$\Delta y_4(t+1) = \gamma_1 \Delta y_3(t) + \gamma_2 \Delta y_3(t-1) + \ldots + \gamma_p \Delta y_3(t-p+2) + u_4(t+1) \quad (3.11)$$

where $\hat{E}[u_i(t+1)|H_t]=0$ for $i=3,4$. Here we assume $\beta(L)$ is of order $p$ and $\gamma(L)$ is of order $p-1$ in order to simplify the exposition, but we do not lose generality because any of $\beta_i$ and $\gamma_i$ can be zero. Then as in Hansen and Sargent (1982), (3.8) implies

$$\xi_0 = \gamma(\delta)(I-\delta\beta(\delta))^{-1}$$

$$\xi_j = \gamma(\delta)(I-\delta\beta(\delta))^{-1}(\beta_{j+1} + \delta\beta_{j+2} + \ldots + \delta^{p-1}\beta_p)$$

$$+ (\gamma_j + \delta\gamma_j + \ldots + \delta^{p-1}\gamma_p) \quad \text{for } j=1,\ldots,p. \quad (3.12)$$

In the structural ECM form (2.8), we have $B=[-b,0,0,0]', A=[1, -1, -1, 0]'$, $C_0 = \begin{bmatrix} 1 & c_{02}^1 & c_{03}^1 & c_{04}^1 \\ c_{01}^2 & 1 & c_{03}^2 & c_{04}^2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.13)$

and

$$C_j = \begin{bmatrix} c_{j1}^1 & c_{j2}^1 & c_{j3}^1 & c_{j4}^1 \\ 0 & 0 & \alpha^2 & 0 \\ 0 & 0 & \beta_j & 0 \\ 0 & 0 & \gamma_j & 0 \end{bmatrix} \quad (3.14)$$
for \(j = 1, \ldots, p\), where \(\gamma_p = 0\).

We have now obtained a system of four equations that consist of (3.1), (3.9), (3.10), and (3.11). Because \(E(u_1(t) | I_{\tau - T}) = 0\) and \(\hat{E}(u_i(t) | H_i) = 0\), we can obtain a vector of instrumental variables \(z_1(t)\) in \(I_{\tau - T}\) for \(u_1(t)\) and \(z_i(t)\) in \(H_i\) for \(u_i(t), i = 2, 3, 4\).

Because the speed of adjustment \(b\) for \(y_1(t)\) affects the dynamics of the other variables,\(^1\) there are cross-equation restrictions involving \(b\) in many applications, in addition to the restrictions in (3.12). Using the moment conditions \(E[z_i(t)u_i(t)] = 0\) for \(i = 1, \ldots, 4\), we form a GMM estimator, imposing the restrictions (3.12) and the other cross-equation restrictions implied by the model.

Given estimates of cointegrating vectors from the first step, this system method provides more efficient estimators than the single equation two step method proposed in Section 2.3, as long as the restrictions implied by the model are true. On the other hand, the single equation two step method estimators are more robust because misspecification in the other equations does not affect their consistency. The cross-equation restrictions can be tested by Wald, Likelihood Ratio type, and Lagrange Multiplier tests in a GMM framework (see, e.g., Ogaki (1993a)). When the restrictions are nonlinear, Likelihood Ratio type and Lagrange Multiplier tests are known to be more reliable than Wald tests.

\(^1\) Note that only \(y_1(t)\) adjusts slowly, but \(b\) affects the dynamics of the other variables because of the interactions of \(y_1(t)\) with those variables.
3.4 An Exchange Rate Model with sticky prices

In this section, we use a one-good version of Mussa’s (1982) model to illustrate the econometric method proposed in the previous section. In the first subsection, we present the exchange rate model. In the second subsection, an econometric model is derived from the exchange rate model.

3.4.1 The Exchange Rate Model

We close the model presented in Section 2.2 by adding the money demand equation and the Uncovered Interest Parity condition. Let

\[ m(t) = \theta_m + p(t) - hi(t) \]  \hspace{1cm} (3.15)

\[ i(t) = i^*(t) + E[e(t+1)|I_I] - e(t) \]  \hspace{1cm} (3.16)

where \( m(t) \) is the log nominal money supply minus the log real national income, \( i(t) \) is the nominal interest rate in the domestic country, and \( i^*(t) \) is the nominal interest rate in the foreign country. In (3.15), we are assuming that the income elasticity of money is one. From (3.15) and (3.16), we obtain

\[ E[e(t+1)|I_I-e(t) = (1/h) \theta_m + p(t) - \omega(t) \]

\[ - h[E[p^*(t+1)-p^*(t)]|I_I] \]  \hspace{1cm} (3.17)
where
\[ \omega(t) = m(t) + hr^*(t) \] (3.18)

, and \( r^*(t) \) is the foreign real interest rate:
\[ r^*(t) = i^*(t) - E[p^*(t+1)|I_t] + p^*(t). \] (3.19)

Following Mussa, solving (2.5) and (3.17) as a system of stochastic difference equation for \( E[p(t+j)|I_t] \) and \( E[e(t+j)|I_t] \) for fixed \( t \) results in
\[ p(t) = E[F(t)|I_{t-1}] - \sum_{j=0}^{\infty} (1-h)^j \{E[F(t-j)|I_{t-j}] - E[F(t-j)|I_{t-j-1}]\} \] (3.20)
\[ e(t) = \frac{bh-1}{bh} E[F(t)|I_t] - p^*(t) + \frac{1}{bh} p(t) \] (3.21)

where
\[ F(t) = (1-\delta) \sum_{j=0}^{\infty} \delta^j \omega(t+j) \] (3.22)

and \( \delta = h/(1+h) \).

We assume that \( \omega(t) \) is first difference stationary. Since \( \delta \) is a positive constant smaller than one, this implies that \( F(t) \) is also first difference stationary. From (3.20) and (3.21),
\[ e(t) + p^*(t) - p(t) = \sum_{j=0}^{\infty} (1-b)^j [E(F(t-j)|I_{t-j}) - E[F(t-j)|I_{t-j-1}]]. \quad (3.23) \]

Since the right hand side of (3.23) is stationary,\(^2\) \( e(t) + p^*(t) - p(t) \) is stationary. Hence Equation (3.23) implies that \( (p(t), e(t), p^*(t)) \) is cointegrated, with the cointegrating vector \( (1,-1,-1) \).

3.4.2 The Econometric Model

From (3.21), we obtain

\[ \Delta e(t+1) = \frac{bh-1}{bh} (I-\delta) \sum_{j=0}^{\infty} \bar{\Delta} \omega(t+j+1)|I_t| + \frac{1}{bh} \Delta p(t+1) - \Delta p^*(t+1) + \epsilon_e(t+1) \quad (3.24) \]

where \( \epsilon_e(t+1) = \frac{bh-1}{bh} [E[F_{t+1}|I_{t+1}] - E[F_{t+1}|I_t]] \), so that the law of iterated expectations implies \( E[\epsilon_e(t+1)|I_t] = 0 \). Because this equation involves a discounted sum of expected future values of \( \Delta \omega(t) \), the system method in Section 3.3 is applicable.

We take the econometrician's information set at \( t, H_t \), to be the one generated by linear functions of the current and past values of \( \Delta p^*(t) \). Then, replacing the best forecast of the economic agents, \( E[\sum_{j=0}^{\infty} \bar{\Delta} \omega(t+j+1)|I_t] \) by the econometrician's linear forecast based on \( H(t) \) in Equation (3.24), we obtain

\(^2\) This assumes that \( E[F(t)|F(t-1)/F(t)] \) is stationary, which is true for a large class of first difference stationary variable \( F(t) \) and information sets.

56
\[
\Delta e(t+1) = \frac{bh-1}{bh} (1-\delta) \sum_{j=0}^{\infty} \mathbb{E}[\Delta \omega(t+j+1)|H(t)] + \frac{1}{bh} \Delta p(t+1) - \Delta p^*(t+1) + u_2(t+1) \\
(3.25)
\]

where
\[
u_2(t+1) = \varepsilon_e(t+1) + \frac{bh-1}{bh} (1-\delta) \mathbb{E}\left[\sum \Delta \omega(t+j+1)|H(t)\right].
\]

Hence, \(\mathbb{E}[u_2(t+1)|H(t)] = 0\). Assume that linear projections of \(\Delta p^*(t+1)\) and \(\Delta \omega(t+1)\) onto \(H(t)\) have only a finite number of \(\Delta p^*(t)\) terms:

\[
\mathbb{E}[\Delta \omega(t+1)|H(t)] = \beta_1 \Delta p^*(t) + \beta_2 \Delta p^*(t-1) + \ldots + \beta_p \Delta p^*(t-p+1) \\
(3.26)
\]

\[
\mathbb{E}[\Delta \omega(t+1)|H(t)] = \gamma_1 \Delta p^*(t) + \gamma_2 \Delta p^*(t-1) + \ldots + \gamma_p \Delta p^*(t-p+2) \\
(3.27)
\]

Then,

\[
\mathbb{E}\left[\sum_{j=0}^{\infty} \Delta \omega(t+j+1)|H(t)\right] = \xi_1 \Delta p^*(t) + \xi_2 \Delta p^*(t-1) + \ldots + \xi_p \Delta p^*(t-p+1) \\
(3.28)
\]

where \(\xi_i\) is given by Equation (3.12).

Combining (2.6), (3.25), (3.26), and (3.27) with (3.28), we obtain a system of four equations:

\[
\Delta p(t+1) = d + \Delta p^*(t+1) + \Delta e(t+1) - b[p(t) - p^*(t) - e(t)] + u_1(t+1) \\
(3.29)
\]
\[ \Delta e(t+1) = \frac{1}{bh} \Delta p(t+1) - \Delta p^*(t+1) + \]

\[ \alpha \xi_1 \Delta p^*(t) + \alpha \xi_2 \Delta p^*(t-1) + \ldots + \alpha \xi_p \Delta p^*(t-p+1) + u_2(t+1) \quad (3.30) \]

\[ \Delta p^*(t+1) = \beta_1 \Delta p^*(t) + \beta_2 \Delta p^*(t-1) + \ldots + \beta_p \Delta p^*(t-p+1) + u_3(t+1) \quad (3.31) \]

\[ \Delta \omega(t+1) = \gamma_1 \Delta p^*(t) + \gamma_2 \Delta p^*(t-1) + \ldots + \gamma_{p+1} \Delta p^*(t-p+2) + u_4(t+1) \quad (3.32) \]

where \( \alpha = \frac{bh-1}{bh} (1-\delta) \) and \( u_1(t+1) = \varepsilon(t+1) \). Given the data for \([\Delta p(t+1), \Delta e(t+1), \Delta p^*(t+1), \Delta \omega(t+1)]'\), the system method in Section 3.3 can be applied to these four equations. There exist additional complications for obtaining data for \( \Delta \omega(t+1) \), as discussed in the next subsection.

In the structural ECM form (2.8), we have \( y(t) = [\Delta p(t+1), \Delta e(t+1), \Delta p^*(t+1), \Delta \omega(t+1)]' \), \( B = [-b,0,0,0]' \), \( A = [1, -1, -1, 0]' \),

\[
C_0 = \begin{bmatrix}
1 & -1 & -1 & 0 \\
-(1/bh) & 1 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}, \quad (3.33)
\]

and

\[
C_j = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & \alpha \xi_j & 0 \\
0 & 0 & \beta_j & 0 \\
0 & 0 & \gamma_j & 0
\end{bmatrix}, \quad (3.34)
\]

for \( j = 1, \ldots, p \).
3.4.3 Stable Money Demand

In order to implement the system method, we need data for $\Delta \omega(t)$, which requires knowledge of $h$. Even though $h$ is unknown, a cointegrating regression can be applied to money demand, if money demand is stable in the long-run, as in Stock and Watson (1993). For this purpose, we augment the model as follows:

$$m(t) = \theta_m + p(t) - ht + \zeta_m(t)$$  \hspace{1cm} (3.35)

where $\zeta_m(t)$ is the money demand shock, assumed to be stationary, so that money demand is stable.

By redefining $m(t)$ in the previous section as $m(t) - \zeta_m(t)$, the same equations, as in the previous section are obtained. For the measurement of $\Delta \omega(t)$, we note that the ex ante foreign real interest rate can be replaced by the ex post real foreign real interest rate in (3.20) and (3.21) because of the Law of Iterated Expectations. Using the money market clearing condition (3.35), we obtain

$$\Delta \omega(t+1) = \Delta p(t+1) - h\Delta i(t+1) + h\Delta i^*(t+1) - h [\Delta p^*(t+2) - \Delta p^*(t+1)].$$  \hspace{1cm} (3.36)

With this expression, $\Delta \omega(t)$ can be measured from price and interest rate data without data for monetary aggregate and national income once $h$ is obtained. This
is useful because the latter are not available at the monthly frequency for many countries.

We use the measurement model of Section 2.4 for the purpose of a sensitivity analysis with respect to PPP, as in the case of the single equation method. Again, it is assumed that the model is true for the true price levels, but that only measured prices that follow (2.12A) and (2.12B) are observed. Since \( p^m(t) \) and \( p^*(t) \) are observed, instead of \( p(t) \) and \( p^*(t) \), (2.12A) and (2.12B) are substituted into Equations (3.29)-(3.32) in order to express these equations in terms of measured prices. It is also assumed that \( H_t \) is the information set generated by current and past values of \( \Delta p^*(t) \), instead of \( \Delta p^*(t) \).

As for the adjustment to the PPP level, (3.29) is replaced by (2.15). For \( \Delta \omega(t) \), we use

\[
\Delta \omega^m(t+1) = \frac{1}{\phi} \Delta p^m(t+1) - h \Delta i(t+1) + h \Delta i^*(t+1) \\
- \frac{h}{\phi} [\Delta p^m(t+2)-\Delta p^m(t+1)],
\]

so that

\[
\Delta \epsilon(t+1) = d_z + \frac{bh-I}{bh} (I-\delta) \hat{E} \sum_{j=0}^{\infty} \delta^j \Delta \omega^m(t+j+1) |H(t)| \\
+ \frac{1}{bh\phi} \Delta p^m(t+1) - \frac{1}{\phi} \Delta p^m(t+1) + u^m_2(t+1)
\]

where

\[
u^m_2(t+1) = u_2(t+1) - \frac{bh-I}{bh} (I-\delta) \hat{E} \left\{ \frac{1}{\phi} \nu^m(t) - \frac{h}{\phi} \nu^*(t) \right\} |H(t)| \\
+ \frac{1}{bh\phi} \Delta \nu(t+1) - \frac{1}{\phi} \Delta \nu^*(t+1),
\]
so that \( E[u_2^m(t+1)|H_{t-1}] = 0. \)

Because the price level is assumed to be measured with errors as in (2.8),

\[
m(t) = \theta_2 + (I/\phi)p^m(t) - hi(t) + \zeta_2(t),
\]

where \( \theta_2 = \theta_m - \theta/\phi \) and \( \zeta_2(t) = \zeta_m(t) - \nu(t)/\phi. \) Because \( \zeta_2(t) \) is stationary, a cointegrating regression is applied to (3.40), assuming \( m(t) \) and \( i(t) \) are first difference stationary.

Thus, we run two cointegrating regressions, (2.14) and (3.40), in the first step. In the second step, GMM is applied to the system of four equations that consist of (2.15),

\[
\Delta e(t+1) = \frac{1}{bh\phi} \Delta p^m(t+1) - \frac{1}{\phi} \Delta p_\phi(t+1)
\]
\[
+ \alpha_1 \Delta p^m(t) + \alpha_2 \Delta p^m(t-1) + \ldots + \alpha_p \Delta p^m(t-p+1) + u^m_2(t+1)
\]

\[
\Delta p^m(t+1) = \beta_1 \Delta p^m(t) + \beta_2 \Delta p^m(t-1) + \ldots + \beta_p \Delta p^m(t-p+1) + u_3(t+1)
\]

\[
\Delta \omega^m(t+1) = \gamma_1 \Delta p^m(t) + \gamma_2 \Delta p^m(t-1) + \ldots + \gamma_p \Delta p^m(t-p+2) + u_4(t+1)
\]

where \( h \) is replaced by its estimate from (3.40), and \( \phi \) and \( \phi^* \) are replaced by their estimates from (2.14). As before, because the first step estimators are super consistent, the first step estimation does not affect asymptotic distributions of the second step GMM estimators.
3.5 Empirical Results

Our estimation procedure has two steps. First, we obtain the measurement error coefficients, exploiting the long-run relationship between domestic prices, foreign prices and exchange rates. We also estimate the monetary equilibrium equation to obtain interest elasticity of money demand. At the second step, the speed of price adjustment is estimated applying GMM to the structural ECM.

The estimates for the measurement error coefficients are from chapter 2, where we use CCR to estimate the long-run coefficients in a single equation method, which are allowed to differ from the cointegrating vector (1, -1, -1).

To estimate the interest elasticity of money demand, we use the sum of M1 and Quasi Money as the measure of money stock, called M2, as IFS suggests. The data for interest rates are the three month T-bill rates, but three month deposit rates are employed for Japan because T-bill rates are not available. We used nominal and real gross domestic product data in IFS dataset for all countries except UK, for which we use DRI data. All data series are seasonally unadjusted. Sample periods vary across countries, according to the availability of data.\footnote{Sample periods used to estimate money demand equations are 1975:1 - 1995:1 for Canada, 1986:3 - 1995:1 for France, 1978:3 - 1995:1 for Germany, 1977:1 - 1993:3 for Italy, 1974:1 - 1993:4 for Japan, 1978:2 - 1990:2 for Netherlands, 1980:1 - 1995:2 for Switzerland, and 1974:1 - 1994:1 for U.K.}

The semi-elasticity of interest rates in money demand equation is estimated using quarterly data. However, since we use monthly data to estimate the system, we transform the quarterly estimated elasticities into monthly elasticities. We multiply the elasticities from the quarterly data by 3, because the quarterly...
interest rates are divided by 3 to obtain monthly interest rates.

Table 3.1 shows the CCR results for the money demand equations. We assume that the income elasticity of money demand is one. For each country, the first row reports the results when the coefficient of the log price is restricted to be one, and the second row reports the results when the coefficient is allowed to be different from one. When we employ the measurement error model, we use the results reported in the second row.

The null of stochastic cointegration is not rejected for all countries except Germany, regardless of the assumption of measurement errors at the 5 percent level. The deterministic cointegrating restriction is rejected for Germany, Italy and Japan at the 5 percent level, when we allow for the measurement errors. With the prespecified cointegrating vector (1, -1, -1), France and Switzerland reject the deterministic cointegrating restriction at the 5 percent level, but do not reject it at the 1 percent level of significance.

In all cases, the signs of the estimates for the interest elasticity of money demand are negative, as expected from the economic model. For Canada, France and Switzerland, the specification of measurement errors does not affect the estimates for the interest elasticities. However, for Germany, Italy, Japan, Netherlands and UK, the estimates from the measurement error models have smaller values than those from the models without measurement error. Interestingly, these range from one fourth to one fifth of the estimates from the no measurement error models for each country.

When we restrict the cointegrating vector to (1, -1, -1), the measurement error coefficients are no longer free parameters. In this case, we have no problem when we run separately two cointegrating regressions which include a
common coefficient. But, if we allow for measurement errors in price indices, then we have two estimates for the measurement error coefficient on the domestic prices. One set of estimates is obtained from the PPP regression and the other set from the money demand equation. There is no guarantee in practice for the two estimates to be the same. If the estimates from the two equations are significantly different, it might imply misspecification of the simple exchange rate model. Although this is the case, we use the estimates from the PPP equation in chapter 2 because we are more interested in PPP than in the money demand equation. Park and Ogaki (1991) suggest seemingly unrelated canonical cointegrating regressions (SUCCR) method to deal with cross equation restrictions, when there are cointegrating vectors in the equations. However, since the small sample properties of their estimator are not better than CCR, we use, in this chapter, the estimates from PPP.

Table 3.2 reports the results of GMM estimation using the system method. The instrumental variables are $p^*_{m(t-3)}$ and $p^*_{m(t-4)}$, which are US prices. For each country, we report results for the known cointegrating vector case and the unknown cointegrating vector case. In the system method, the structural speed of adjustment coefficient $b$, appears in two equations: the slow adjustment equation, (3.29) or (2.15), or the Hansen-Sargent equation, (3.30) or (3.38). The model imposes the restriction that the coefficient $b$ in the slow adjustment equation is the same as the coefficient $b$ in the Hansen-Sargent equation. We report results with and without this restriction imposed for the system method of estimation. In the case of unrestricted estimation, $b_{hs}$ is the estimate of $b$ from the Hansen-Sargent equation, and $b_{ss}$ is the estimate of $b$ from the slow adjustment equation. The restricted estimate is denoted by $b_r$. The likelihood ratio type test
statistic (see, e.g., Ogaki (1993a) for an explanation of this test), denoted by $C$, is used to test the restriction. In most cases, this restriction is not rejected at the five percent level.

For Canada, the likelihood ratio type test does not reject the restriction that the speed of adjustment in the price adjustment equation and the equation derived from the Hansen and Sargent formula are not different. The estimate for the restricted coefficient implies a half life of 5.7 years and 5.9 years for the prespecified cointegration vector and the estimated cointegrating vector, respectively. The half lives are 4.6 and 3.5 years, respectively, when we use the single equation method. As expected, the estimates from the system method are measured more precisely.

For France, the restriction across equations in the system is not rejected when we use the pre specified cointegrating vector (1, -1, -1). The estimate for the restricted short-run dynamic coefficient has a half life of 2.6 years, which is shorter than 3.8 years from the single equation method and measured more precisely. When we use the measurement error model, we obtain a negative estimate for the speed of adjustment, implying that the forcing variables may follow nonstationary processes, if PPP holds and the money demand equation is stationary. The single equation method, with measurement errors, suggests a half life of 1.7 years.

For Germany, we obtain negative values for the speed of adjustment, with both known and unknown cointegrating vectors. But, estimates have large standard errors. The single equation method shows a half life of 5.2 years for the known cointegrating vector, although it is not precisely estimated.

For Italy, the estimates are negative regardless of whether the single
equation method or the system method is used.

For Japan, we obtain a negative estimate, which has a large standard error when we prespecify the cointegrating vector. However, with the measurement error model, Japanese yen shows a half life of 1.7 years, which is similar to the half life of 1.4 years from the single equation method. Both these values are calculated from the significant estimates for the speed of adjustment. The system method gives a relatively more precise estimate.

For Netherlands, with the measurement error model, we obtain a positive estimate, which implies a half life of 9.9 years. This is not significant. On the other hand, the estimate from the single equation is significant and has a half life of 1.3 years. With the known cointegrating vector, we obtain an estimate which is negative but insignificant.

For Switzerland, both the known and unknown cointegrating vectors give positive estimates for the short-run coefficient. However, the likelihood ratio type test rejects the restriction across equations in the system at the 5 percent level.

For UK, the likelihood ratio type test rejects the restriction across equations in the system at the 5 percent level and the estimate for the speed of adjustment is negative when we assume the cointegrating vector is (1, -1, -1). Using the measurement error model, we obtain positive and significant estimates, implying a half life of 5.8 years. In this case, the likelihood ratio type test does not reject the restriction across equations in the system. In the single equation method, the implied half life is 5.4 years.

In general, the system method gives more efficient estimates for the coefficient on the equilibrium error term than those from the single equation.
method. However, in the system method, the number of negative estimates is eight out of sixteen, while it is four out of sixteen in the single equation method. This may imply that the additional restrictions in the structural model, such as the stationary money demand equation, uncovered interest rate parity without risk premium, and the assumption of rational expectations, are invalid. The single equation method gives more robust results.

3.6 Conclusion

This chapter compares reduced form ECMs with structural form ECMs. The speed of adjustment coefficients in reduced form ECMs are very different from those in structural form ECMs in general, and also in our example of an exchange rate model with sticky prices. We propose a single equation IV method and a system IV method to estimate structural speed of adjustment coefficients. These IV methods do not require exogeneity assumptions, and can be applied to a broad range of structural ECMs.

When the system method is applied to the exchange rate model, the speed of adjustment coefficient is estimated from both the slow adjustment equation for the domestic price, and the rational expectations equation for the exchange rate. In most cases, this restriction is not rejected at the five percent level. In some cases, the restricted estimate is significantly negative, indicating that the restricted model is misspecified. For the purpose of estimating the structural speed of adjustment coefficient, the results for Canada and France, in the case of a known cointegrating vector, and those for Canada, Japan, and the
U.K., in the case of an unknown cointegrating vector, are encouraging. In each of these cases, the restriction is not rejected and the restricted estimate is positive and significant at the five percent level.

A potential source of misspecification could be the assumption that uncovered interest parity holds. Since Fama (1984), many studies have shown that this parity condition does not hold in the forward exchange rate market. In the next chapter, we deal with this problem.
<table>
<thead>
<tr>
<th>Country</th>
<th>(\theta_2)</th>
<th>1/(\phi)</th>
<th>(h)</th>
<th>(H(0,1))</th>
<th>(H(1,2))</th>
<th>(H(1,3))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Canada</td>
<td>-0.046</td>
<td>1</td>
<td>30.03</td>
<td>1.019</td>
<td>0.445</td>
<td>1.098</td>
</tr>
<tr>
<td></td>
<td>(0.247)</td>
<td></td>
<td>(9.875)</td>
<td>(0.313)</td>
<td>(0.505)</td>
<td>(0.578)</td>
</tr>
<tr>
<td></td>
<td>0.464</td>
<td>1.899</td>
<td>40.79</td>
<td>0.523</td>
<td>4.681</td>
<td>4.792</td>
</tr>
<tr>
<td></td>
<td>(0.456)</td>
<td>(0.240)</td>
<td>(16.20)</td>
<td>(0.469)</td>
<td>(0.030)</td>
<td>(0.091)</td>
</tr>
<tr>
<td>France</td>
<td>-0.341</td>
<td>1</td>
<td>5.661</td>
<td>4.823</td>
<td>0.963</td>
<td>1.183</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td></td>
<td>(3.182)</td>
<td>(0.028)</td>
<td>(0.326)</td>
<td>(0.554)</td>
</tr>
<tr>
<td></td>
<td>-0.337</td>
<td>0.253</td>
<td>5.560</td>
<td>2.036</td>
<td>0.097</td>
<td>0.321</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.038)</td>
<td>(0.636)</td>
<td>(0.153)</td>
<td>(0.755)</td>
<td>(0.851)</td>
</tr>
<tr>
<td>Germany</td>
<td>-0.272</td>
<td>1</td>
<td>17.88</td>
<td>1.537</td>
<td>1.600</td>
<td>4.575</td>
</tr>
<tr>
<td></td>
<td>(0.165)</td>
<td></td>
<td>(9.856)</td>
<td>(0.215)</td>
<td>(0.206)</td>
<td>(0.102)</td>
</tr>
<tr>
<td></td>
<td>-0.454</td>
<td>1.597</td>
<td>3.003</td>
<td>11.50</td>
<td>18.61</td>
<td>18.65</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.050)</td>
<td>(1.307)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Italy</td>
<td>-0.205</td>
<td>1</td>
<td>7.847</td>
<td>3.755</td>
<td>0.753</td>
<td>2.424</td>
</tr>
<tr>
<td></td>
<td>(0.172)</td>
<td></td>
<td>(4.807)</td>
<td>(0.053)</td>
<td>(0.386)</td>
<td>(0.298)</td>
</tr>
<tr>
<td></td>
<td>-0.508</td>
<td>0.789</td>
<td>1.994</td>
<td>9.247</td>
<td>1.560</td>
<td>5.979</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.013)</td>
<td>(1.073)</td>
<td>(0.002)</td>
<td>(0.212)</td>
<td>(0.050)</td>
</tr>
</tbody>
</table>

\[ m(t) = \theta_2 + (1/\phi)p^m(t) - h_i(t) + \zeta_2(t) \]

Column (1) : domestic countries
Column (2)-(4) : Standard errors are in parentheses.
Column (5)-(7) : P-values are in parentheses.

Table 3.1: Money Demand Equation
(continued on the next page)
Table 3.1 (continued)

<table>
<thead>
<tr>
<th>Country</th>
<th>$\theta_2$</th>
<th>$1/\phi$</th>
<th>$h$</th>
<th>$H(0,1)$</th>
<th>$H(1,2)$</th>
<th>$H(1,3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Japan</td>
<td>-11.31</td>
<td>1</td>
<td>39.66</td>
<td>3.274</td>
<td>3.229</td>
<td>5.353</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td></td>
<td>(5.827)</td>
<td>(0.070)</td>
<td>(0.072)</td>
<td>(0.069)</td>
</tr>
<tr>
<td></td>
<td>-13.95</td>
<td>1.520</td>
<td>8.089</td>
<td>13.02</td>
<td>0.035</td>
<td>0.674</td>
</tr>
<tr>
<td></td>
<td>(0.784)</td>
<td>(0.171)</td>
<td>(5.228)</td>
<td>(0.000)</td>
<td>(0.852)</td>
<td>(0.714)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>-0.178</td>
<td>1</td>
<td>6.658</td>
<td>10.16</td>
<td>0.292</td>
<td>0.257</td>
</tr>
<tr>
<td></td>
<td>(0.075)</td>
<td></td>
<td>(4.124)</td>
<td>(0.001)</td>
<td>(0.864)</td>
<td>(0.879)</td>
</tr>
<tr>
<td></td>
<td>-1.451</td>
<td>1.778</td>
<td>1.731</td>
<td>0.565</td>
<td>3.729</td>
<td>4.826</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.150)</td>
<td>(3.006)</td>
<td>(0.452)</td>
<td>(0.053)</td>
<td>(0.090)</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.282</td>
<td>1</td>
<td>9.858</td>
<td>4.029</td>
<td>0.876</td>
<td>1.751</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td></td>
<td>(2.468)</td>
<td>(0.045)</td>
<td>(0.349)</td>
<td>(0.417)</td>
</tr>
<tr>
<td></td>
<td>0.287</td>
<td>1.155</td>
<td>8.987</td>
<td>0.085</td>
<td>2.296</td>
<td>3.859</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.056)</td>
<td>(1.276)</td>
<td>(0.771)</td>
<td>(0.129)</td>
<td>(0.145)</td>
</tr>
<tr>
<td>U.K.</td>
<td>10.66</td>
<td>1</td>
<td>111.3</td>
<td>1.853</td>
<td>0.088</td>
<td>0.597</td>
</tr>
<tr>
<td></td>
<td>(3.410)</td>
<td></td>
<td>(128.3)</td>
<td>(0.173)</td>
<td>(0.766)</td>
<td>(0.742)</td>
</tr>
<tr>
<td></td>
<td>8.982</td>
<td>2.265</td>
<td>27.56</td>
<td>2.942</td>
<td>0.129</td>
<td>5.562</td>
</tr>
<tr>
<td></td>
<td>(0.273)</td>
<td>(0.116)</td>
<td>(9.306)</td>
<td>(0.086)</td>
<td>(0.720)</td>
<td>(0.062)</td>
</tr>
</tbody>
</table>

$m(t) = \theta_2 + (1/\phi)p^m(t) - h(t) + \zeta_2(t)$

Column (1) : domestic countries
Column (2)-(4) : Standard errors are in parentheses.
Column (5)-(7) : P-values are in parentheses.
<table>
<thead>
<tr>
<th>Country</th>
<th>$\phi$</th>
<th>$\phi/\phi^*$</th>
<th>$b_{u,hs}$</th>
<th>$b_{u,sa}$</th>
<th>$J_u$</th>
<th>$b_r$</th>
<th>$J_r$</th>
<th>$C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>1</td>
<td>1</td>
<td>0.0107</td>
<td>0.0035</td>
<td>1.782</td>
<td>0.0101</td>
<td>1.947</td>
<td>0.165</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0015)</td>
<td>(0.0164)</td>
<td>(0.410)</td>
<td>(0.0004)</td>
<td>(0.584)</td>
<td>(0.685)</td>
</tr>
<tr>
<td></td>
<td>0.029</td>
<td>1.014</td>
<td>0.0097</td>
<td>0.0098</td>
<td>0.689</td>
<td>0.0097</td>
<td>0.689</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0013)</td>
<td>(0.0071)</td>
<td>(0.709)</td>
<td>(0.0009)</td>
<td>(0.876)</td>
<td>(0.987)</td>
</tr>
<tr>
<td>France</td>
<td>1</td>
<td>1</td>
<td>0.0293</td>
<td>-0.0283</td>
<td>0.348</td>
<td>0.0218</td>
<td>0.397</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0476)</td>
<td>(0.0394)</td>
<td>(0.840)</td>
<td>(0.0050)</td>
<td>(0.941)</td>
<td>(0.824)</td>
</tr>
<tr>
<td></td>
<td>0.221</td>
<td>1.106</td>
<td>-0.0450</td>
<td>-0.0123</td>
<td>2.996</td>
<td>-0.0228</td>
<td>4.363</td>
<td>1.367</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0570)</td>
<td>(0.0133)</td>
<td>(0.224)</td>
<td>(0.0099)</td>
<td>(0.225)</td>
<td>(0.242)</td>
</tr>
<tr>
<td>Germany</td>
<td>1</td>
<td>1</td>
<td>0.7699</td>
<td>-0.0525</td>
<td>0.629</td>
<td>-0.0067</td>
<td>3.249</td>
<td>2.620</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.2349)</td>
<td>(0.0578)</td>
<td>(0.730)</td>
<td>(0.0332)</td>
<td>(0.354)</td>
<td>(0.106)</td>
</tr>
<tr>
<td></td>
<td>0.073</td>
<td>0.376</td>
<td>-0.0051</td>
<td>-0.0052</td>
<td>1.151</td>
<td>-0.0052</td>
<td>1.155</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0407)</td>
<td>(0.0058)</td>
<td>(0.562)</td>
<td>(0.0057)</td>
<td>(0.764)</td>
<td>(0.951)</td>
</tr>
<tr>
<td>Italy</td>
<td>1</td>
<td>1</td>
<td>-0.3210</td>
<td>-0.0350</td>
<td>1.714</td>
<td>-0.0412</td>
<td>1.942</td>
<td>0.228</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(3.6246)</td>
<td>(0.0332)</td>
<td>(0.424)</td>
<td>(0.0308)</td>
<td>(0.585)</td>
<td>(0.633)</td>
</tr>
<tr>
<td></td>
<td>0.192</td>
<td>1.621</td>
<td>-0.6262</td>
<td>-0.0219</td>
<td>3.305</td>
<td>-0.0284</td>
<td>6.953</td>
<td>3.648</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2.9938)</td>
<td>(0.0098)</td>
<td>(0.192)</td>
<td>(0.0082)</td>
<td>(0.073)</td>
<td>(0.056)</td>
</tr>
</tbody>
</table>

Column (1) : domestic countries  
Column (4),(5) & (7) : Standard errors are in parentheses.  
Column (6),(8) & (9) : P-values are in parentheses.  
Column (2) and (3) are from Table 2.  
$b_{u,hs}$ is the estimate for the speed of adjustment, obtained from Hansen and Sargent equations.  
$b_{u,sa}$ is the estimate for the speed of adjustment, obtained from the price adjustment equation.

Table 3.2: Instrumental Variable Estimation: System  
(continued on the next page)
<table>
<thead>
<tr>
<th>Country</th>
<th>$\phi$</th>
<th>$\phi/\phi^*$</th>
<th>$b_{u,hs}$</th>
<th>$b_{u,sa}$</th>
<th>$J_u$</th>
<th>$b_r$</th>
<th>$J_r$</th>
<th>$C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan</td>
<td>1</td>
<td>1</td>
<td>0.0498</td>
<td>-0.0568</td>
<td>2.240</td>
<td>-0.0111</td>
<td>14.58</td>
<td>12.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.1478)</td>
<td>(0.0305)</td>
<td>(0.326)</td>
<td>(0.0278)</td>
<td>(0.002)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>0.156</td>
<td>0.728</td>
<td>0.0316</td>
<td>0.0602</td>
<td>5.324</td>
<td>0.0327</td>
<td>5.891</td>
<td>0.567</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0055)</td>
<td>(0.0343)</td>
<td>(0.070)</td>
<td>(0.0047)</td>
<td>(0.117)</td>
<td>(0.452)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1</td>
<td>1</td>
<td>-135.25</td>
<td>-0.0378</td>
<td>1.653</td>
<td>-0.0428</td>
<td>2.992</td>
<td>1.339</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(5223.5)</td>
<td>(0.0461)</td>
<td>(0.438)</td>
<td>(0.0458)</td>
<td>(0.393)</td>
<td>(0.247)</td>
</tr>
<tr>
<td></td>
<td>0.141</td>
<td>0.651</td>
<td>31.734</td>
<td>0.0099</td>
<td>1.060</td>
<td>0.0058</td>
<td>4.437</td>
<td>3.377</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(297.31)</td>
<td>(0.0124)</td>
<td>(0.589)</td>
<td>(0.0123)</td>
<td>(0.218)</td>
<td>(0.066)</td>
</tr>
<tr>
<td>Switzerland</td>
<td>1</td>
<td>1</td>
<td>0.5907</td>
<td>0.0115</td>
<td>0.483</td>
<td>0.0150</td>
<td>11.12</td>
<td>10.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2.2681)</td>
<td>(0.0220)</td>
<td>(0.786)</td>
<td>(0.0183)</td>
<td>(0.011)</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>0.027</td>
<td>0.346</td>
<td>-0.5660</td>
<td>0.0121</td>
<td>0.355</td>
<td>0.0395</td>
<td>8.152</td>
<td>7.798</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.4543)</td>
<td>(0.0175)</td>
<td>(0.837)</td>
<td>(0.0170)</td>
<td>(0.043)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>U.K.</td>
<td>1</td>
<td>1</td>
<td>1.7222</td>
<td>-0.0090</td>
<td>1.228</td>
<td>-0.0051</td>
<td>16.44</td>
<td>15.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(7.9049)</td>
<td>(0.0611)</td>
<td>(0.541)</td>
<td>(0.0002)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>0.369</td>
<td>1.458</td>
<td>0.0131</td>
<td>-0.0026</td>
<td>0.645</td>
<td>0.0099</td>
<td>3.439</td>
<td>2.796</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0026)</td>
<td>(0.0075)</td>
<td>(0.725)</td>
<td>(0.0004)</td>
<td>(0.329)</td>
<td>(0.095)</td>
</tr>
</tbody>
</table>

Column (1) : domestic countries
Column (4),(5) & (7) : Standard errors are in parentheses.
Column (6),(8) & (9) : P-values are in parentheses.
Column (2) and (3) are from Table 2.

$b_{u,hs}$ is the estimate for the speed of adjustment, obtained from Hansen and Sargent equations.

$b_{u,sa}$ is the estimate for the speed of adjustment, obtained from the price adjustment equation.
CHAPTER 4

A COMPARISON OF RISK PREMIUM MODELS FOR FORWARD EXCHANGE RATE CONTRACTS

4.1 Introduction

We assume in chapter 3 that the unbiased interest parity condition holds. Fama (1984) and many other studies, however, show that the forward foreign exchange rate is not an unbiased predictor of the future spot exchange rate. One explanation for this bias is based on risk premiums. The risk premium models assume market efficiency, and the possibility that forward rates contain time-varying risk premiums as well as market forecasts of future spot rates.

As a first step towards incorporating a risk premium into the sticky price model, we compare the popular risk premium models with each other to find out which model is more consistent with the forward rate data.

We compare all these models in a general framework. Since the traditional CAPM and the consumption-based CAPMs can be regarded as particular cases of APT, and they differ only in the specification of the stochastic discount factor, we can compare all these models in a generalized multifactor model.

We apply Hansen's (1982) GMM and J-test to the traditional CAPM, a simple
consumption-based CAPM and a consumption-based CAPM with habit formation. We treat them as special cases of an arbitrage theory, which includes the traditional CAPM beta and the consumption-based CAPM beta as factors. First, we estimate each risk premium model in a factor pricing model. Nested tests show which model fits the forward return data better.

In section 4.2, we review various risk premium models of the forward returns. Section 4.3 explains the factor pricing model that we use in this paper. Section 4.4 describes the data, and presents estimation results for each risk premium model. Section 4.5 deals with nested tests to compare the models. Section 4.6 concludes.

4.2 Risk Premium Models

Fama (1984) suggests that there exist risk premiums in the forward exchange rate markets. An exchange rate is the price of one currency in terms of another currency, and a forward rate is a contractual exchange rate established at a point in time for a transaction that will take place at the maturity date on the contract in the future. Well-organized forward markets exist for all major currencies of the world for various maturities.

Although there exist well-organized forward markets, one of the notable characteristics of flexible exchange rate has been the magnitude of forward rate forecast error. Many studies show evidence against the hypothesis of simple market efficiency in which either the forward rate or its logarithm is equated with the conditional expectation of the level or its logarithm of the future spot
exchange rate. Fama (1984) tests a model for joint measurement of variation in
the premium and expected future spot rate components of forward rates.
Conditional on the hypothesis that the forward market is efficient or rational
and using a regression approach, Fama finds reliable evidence that both
components of forward rates vary through time. Furthermore he concludes that most
of the variation in forward rates is variation in premiums. Fama shows that, with
the Fisher equation, interest rate parity and purchasing power parity, the
premium in the forward foreign exchange market is the difference between the
expected real returns on the nominal bonds of the two countries. Therefore, the
variables that determine the difference between the expected real returns on the
nominal bonds also explain the premium in the forward rate. He argues that this
interpretation applies to any model of international capital market equilibrium
characterized by interest rate parity, purchasing power parity, and the Fisher
equation for nominal interest rates.

Instead of the regression approach, Hansen and Hodrick (1983) employ an
intertemporal asset pricing model to determine the nature of the risk premium in
a way that leads to statistical representations with testable hypotheses. They
develop testable hypotheses that incorporate the relevant intertemporal risk
considerations because using a traditional approach to measuring risk with a
static capital asset pricing model cannot adequately characterize the
intertemporal movements in risk premium. They first derive the first-order
conditions of an economic agent who has the opportunity to trade forward foreign
exchange contracts in competitive equilibrium. And then they develop linear
econometric models of the risk premiums.

They also develop an approach which is related to the traditional CAPM. The
traditional CAPM characterizes the risk return trade-off facing investors with a single beta model. In this model the riskiness of any asset is measured by the covariance of the excess return on the asset with the excess return from some benchmark portfolio. They argue that the empirical content of the traditional static capital asset pricing model is based on the assumption that observations on a vector of returns, including the return on the aggregate wealth portfolio, are normally distributed with probability distributions that are independent and identical over time. Hansen and Hodrick relax the requirement that returns be temporarily independent.

Since the intertemporal asset pricing models do not have the implication that the return on the aggregate wealth portfolio be mean variance efficient, they do not require that observations on a benchmark return for a single beta model be available a priori. Instead they treat the benchmark return as a latent variable. In this sense their method is called a latent variable model. They assume that all of the time variation in risk premiums in the foreign exchange market and in a suitable benchmark portfolio can be captured by movements in conditional means. They also assume that conditional betas are constant. With these assumptions their model incorporates the ingredients of the static CAPM without its restrictive assumptions of zero temporal covariances and observability of the return on a benchmark portfolio. They conclude that the latent variable model captures most of the significant variation in the deviation of spot rates from forward rates.

Mark (1988) also specifies the pricing of forward foreign exchange contracts as a single beta capital asset pricing model. He specifies the CAPM with the beta which is parameterized following ARCH. That is, the model is specified in a
conditional setting that explicitly models time variation in the betas. In the standard specification of the static capital asset pricing model in an unconditional environment, the beta is the ratio of the unconditional covariance and variance, and therefore constant. Mark argues that a constant beta is consistent with time variation in the risk premium in his specification of the CAPM, due possibly to time variation in the ex ante excess return on the reference portfolio. He assumes that the benchmark return is a weighted average of the return on stock market indexes in the U.S., Germany, Switzerland, Japan and U.K.. This specification can be regarded as a special case of the Hansen and Hodrick's (1983) latent variable model (see Engel (1995)). He concludes that the evidence supports the idea that deviations of the forward exchange rate from the expected future spot rate are due to a risk premium and not to irrationality among market participants. Engel (1995) argues that when the appropriate benchmark is the intertemporal marginal rate of substitution in consumption, then the tests of Mark’s version of the intertemporal CAPM, which uses a measure of equity returns as the benchmark return, would be misspecified unless that benchmark return happened to be perfectly correlated with the marginal rate of substitution. Engel also contends that there is no model of the risk premium that has been explicitly derived that implies that the return on equities would be the correct benchmark.

Cumby (1988) uses the CAPM-type latent variable model to determine if ex ante returns to forward speculation can be explained by models of a foreign exchange risk premium. He specifies a latent variable model which requires that the conditional covariances of real returns to forward speculation and the rate of change of real consumption move together over time for all currencies. Cumby
substitutes out the expected benchmark return by using an arbitrarily chosen reference asset and derives a set of proportionality restrictions. Since the ex ante returns are unobservable, he works with observable realized returns, which can be decomposed into expected returns and a forecast error. By assuming that expectations are rational, he can make inferences about the behavior of ex ante returns based on the observed behavior of realized returns because forecast errors are assumed to be unforecastable given information available at the time the forecast is made. He finds that the restrictions implied by the consumption-based model and proportionality of the conditional covariances are rejected by the data.

The latent variable models exploits the CAPM specification although they are derived from the consumption-based intertemporal model. Mark (1988) shows this relationship explicitly by using stock market return data. None of these models use consumption data. On the other hand, there are studies exploiting the information in the consumption data. We call these models consumption-based CAPMs.

Mark (1985) works directly with the first order conditions of the intertemporal capital asset pricing model with a representative agent to examine time varying risk premiums in the forward foreign exchange market. In discrete-time intertemporal asset pricing models such as Lucas (1982), investors in the forward market are assumed to maximize discounted expected utility defined on consumption subject to sequential budget constraints. The first order conditions require that the conditional first cross moment of the marginal rate of substitution of consumption between periods $t$ and $t+k$ and the $k$-period real return from investing in an asset equals zero. In this model, the risk premium on
forward foreign exchange is proportional to the conditional covariance of the k-period return on foreign exchange speculation and the marginal rate of substitution of money between periods t and t+k. Mark argues that time variation of the risk premium can be explained through the time variation of this conditional covariance. Mark derives the fundamental restrictions imposed by the intertemporal model by taking the difference between uncovered and covered investments in the nominally risk-free asset and dividing by the known foreign nominal return and by the constant discount factor that is known to the agent but not to the econometrician. Mark concludes that the evidence against the model is not overwhelming.

Backus et al. (1993) deal with the issue of predictable returns from currency speculation from the perspective of the representative agent theory of asset pricing. The variability of intertemporal prices in this model, measured by the intertemporal marginal rate of return, is too small to account for mean excess returns on equity and long-term bond. Introducing habit formation into this representative agent model, Ferson and Constantinides (1991) argue that this model can explain many of the properties of expected returns on equity and bonds. Backus et al. examine the potential of such a model to account for expected returns from speculation in foreign exchange market. A nonseparability of the representative agent’s preferences is introduced to the simple representative agent model by assuming habit formation or durability which depends on the sign of a parameter. This habit formation model produces nonzero risk premiums even when the representative aggregate consumer is risk neutral. Backus et al. conclude that the habit formation model strongly rejects the overidentifying restrictions and there exists durability rather than habit formation.
There are a few studies that compare CAPM with consumption CAPM in a closed economy.

Kocherlakota (1996) shows a comparison of the traditional CAPM to the consumption-based CAPM. The real returns paid by different financial securities differ considerably over long period of time. These differences in average returns are typically explained by differences in the degree to which a security’s return covaries with the typical investors’s consumption. If this covariance is high, selling off the security would greatly reduce the variance of the typical investor’s consumption stream. In equilibrium, the investor must be deterred from reducing his risk in this fashion by the security’s paying a high average return. He argues that the crucial problem in this qualitative explanation of cross-sectional differences in asset returns is to figure out what the typical investor’s consumption exactly is. The traditional CAPM assumes that the typical investor’s consumption is perfectly correlated with the return to the stock market. In this model the risk of a financial security is measured by its covariance with the return to the stock market. Representative agent models of asset returns assume that per capita consumption is perfectly correlated with the consumption of the typical investors. In this case the risk of a financial security is measured using the covariance of its return with per capita consumption. Although the representative agent model is not widely used as the traditional CAPM in real world applications, he asserts, the representative agent models are more important that the traditional CAPM because they are an integral part of modern macroeconomics and international economics’.

Mankiw and Shapiro (1986) compare the traditional CAPM and a representative agent model, which is a consumption-based CAPM. Their understanding of the
consumption CAPM is that while the covariance with the return on a stock market index may be the standard measure of systematic risk, the consumption CAPM suggests that a better measure is the covariance with aggregate consumption growth. They argue that the consumption beta appears preferable on theoretical grounds because it takes account of the intertemporal nature of the portfolio decision and it implicitly incorporates the many forms of wealth beyond stock market wealth that are in principle relevant for measuring systematic risk. Although the overidentifying restrictions implied by the consumption-based model are rejected, they believe that it is possible that in economic terms the model is approximately true, but the strict tests of overidentification fail. Therefore, they examine whether the consumption CAPM provides an empirically more useful framework for understanding cross-sectional stock returns. They construct a test that nests an alternative hypothesis motivated by economic theory to tell whether the consumption CAPM or the traditional CAPM is more consistent with the data. They derive a linear relation between the stock return on an asset and the market beta of the traditional CAPM and another linear relation between the stock return on an asset and a consumption beta. They regress the return on an asset on its market beta and its consumption beta to see which measure of risk is a better explanation of return. They use quarterly data from 1959 to 1982 to calculate the return and covariances for each of 464 stocks. The market return they use is the return on the Standard and Poor composite. The consumption measure is real consumer expenditure per capita on non-durables and services during the first month of the quarter. After they obtain the sample estimates of a measure of the systematic risk of each asset, they use them as a variable in the regression which involves the market beta and the consumption beta.
Since the traditional CAPM implies that assets with high systematic risk earn high average return, they examine whether the positive association holds true. The slope coefficient on the market beta, which should be the spread between the market return and the risk-free return, is always positive, significant, and of reasonable size. The estimated constant, which should be the risk-free return, is always insignificantly different from one or from zero. However, the coefficient on the consumption beta is insignificant. The constant term implies a risk-free return of 4% which is higher than the theory suggests it would be. When they include the market beta and the consumption beta together, the coefficient on the market beta is always far larger and far more significant than is the coefficient on the consumption beta. They conclude that the market rewards systematic risk with higher return, but the relevant measure of systematic risk appears to be the market beta rather than the consumption beta.

4.3 Factor Pricing Models

In this section we present the factor pricing model we employ in our empirical study of forward exchange rate returns. Our specification is based on Hansen and Jaganathan (1991) and Hansen and Richard (1987). Hansen and Jaganathan characterize the properties of the discount factors that are consistent with the behavior of the asset market payoffs and prices. Hansen and Richard show that stochastic discount factors provide a convenient vehicle for summarizing the implications of dynamic economic models for security market pricing because alternative models can imply differing stochastic discount factors. Their
argument begins with the principle of no arbitrage.

Dybvig and Ross (1989) define an arbitrage opportunity as an investment strategy that guarantees a positive payoff in some contingency with no possibility of a negative payoff and with no net investment. They emphasize that the most important implication of the absence of arbitrage is the existence of a positive linear pricing rule, which in many spaces including finite spaces, is the same as the existence of positive state prices that correctly price all assets. The principle of no arbitrage implies that alternative ways of constructing the same payoff must have the same cost or price, as long as there is a nontrivial, nonnegative portfolio payoff. The law of one price asserts that two perfect substitutes must trade at the same price. In other words, the law of one price means that each portfolio payoff must have a unique price. Although the absence of arbitrage implies the law of one price, the law of one price is not equivalent to the absence of arbitrage. The law of one price is less restrictive than the absence of arbitrage because it deals only with the case in which two assets are identical but have different prices. It does not cover cases in which one asset dominates another but may do so by different amounts in different states. The law of one price can be tested by looking for two portfolios with the same payoffs, but different prices. The absence of arbitrage can be tested by looking for a portfolio with a nonnegative and nontrivial payoff with a nonpositive price. When the law of one price holds, investors can purchase a claim to a linear combination of any two security market payoffs by simply purchasing the corresponding linear combination of the securities. Thus, the unique assignment of prices to portfolio payoffs depends on this linearity.

Cochrane and Hansen (1992) give an example that shows the relationship
between the absence of arbitrage and the stochastic discount factor. Suppose there are \( n \) primitive payoffs and \( x \) is an \( n \)-dimensional random vector with a finite second moment, obtained by stacking the \( n \) payoffs. A space of payoffs, \( P \), which is used in econometric analyses consists of constant weighted portfolios of the primitive payoffs:

\[ P = \{ p: p = cf \text{ for some } c \in \mathbb{R}^n \}, \]

where \( c \) is a vector of portfolio weight. Then, we can construct a candidate price of a portfolio payoff \( cx \), from prices of the original \( n \) payoffs \( q \), via

\[ \pi(cx) = cq. \]

The law of one price says that this price assignment depends on only the payoff \( cx \) itself and not necessarily on the choice of \( c \) used to construct this payoff. If \( E(xx') \) is nonsingular, the law of one price is satisfied because there is only one portfolio weight that achieves any attainable payoff.

Cochrane and Hansen define a stochastic discount factor as any random variable \( y \) that correctly represent the prices of payoffs via the formula:

\[ \pi(p) = E(yp) \text{ for all } p \text{ in } P, \]

where \( \pi \) is a pricing function and \( P \) is a space of payoffs. The Riesz representation theorem guarantees the existence of a stochastic discount factor, as long as the the law of one price is satisfied. They show that the discount factor is not unique. Thus, the intertemporal marginal rate of substitution of consumers in a consumption-based asset pricing model can be a stochastic discount factor generated from a theoretical device. Consumption-based models generate positive stochastic discount factors, while the traditional CAPM implies that stochastic discount factors need not be positive. They assert that whether the family of all discount factors or the family of all discount factors that are

84
strictly positive is the relevant family of discount factors depends on the economic models being studied. Hansen and Jaganathan (1991) show that the discount factor can be represented as a linear function of factors. We follow their approach in this paper. Thus, using the no arbitrage condition, we can express any asset pricing model as an unconditional linear factor pricing model,

\[ E(m\bar{r}) = 1, \quad E(mrf) = 1, \quad m = \sum_{f} b_{f}r_{f} \]  \hspace{1cm} (4.1)

where \( m \) is a stochastic discount factor, \( \bar{r} \) is an asset return, and \( r_{f} \) is a factor.

Hansen and Richard (1987) introduce conditioning information into the setup above. In intertemporal models, information accumulates over time and the accumulated information becomes imbedded in asset prices. Consequently, they say, a pricing function maps payoffs modeled as random variables into prices that are also modeled as random variables, but are constrained to be in the information set of traders at the time prices are quoted. They show that the alternative asset pricing functions that embody conditioning information can be represented using alternative random variables among the collection of payoffs from portfolios. They derive this result from a conditional counterpart to the Riesz representation of a linear functional on a Hilbert space. Establishing conditions under which equilibrium pricing functions can be viewed as conditional linear functions with conditional inner product representation, they use pricing functions to construct linear functionals defined on a standard Hilbert space of portfolio payoffs. The linear functionals are used to deduce testable restrictions in terms of unconditional moments. Since the payoffs can be
constructed that are conditional linear combinations of some initial collection of payoffs, the prices of the resulting payoffs are just the conditional linear combinations of the corresponding prices of the initial collection of payoffs. The conditional weights used in forming payoffs and prices correspond to the instrumental variables in Hansen and Singleton (1982).

Following this approach, a conditional asset pricing model is

$$E[m_{t+1} r_{t+1} | I_t] = 1, \quad m = \mathbf{b}' \mathbf{f}$$

(4.2)

where $I_t$ is the information set at time $t$. This conditional asset pricing model can be tested by including returns scaled by instruments,

$$E[m_{t+1} (r_{t+1} \otimes z_t) | I_t] = 1, \quad z_t \in I_t$$

(4.3)

where $\otimes$ denotes the Kronecker product, and $z_t$ represents a vector of instrumental variables included in $I_t$. The scaled returns $r_{t+1} \otimes z_t$ are returns on a managed portfolio. $m$ can be expressed as a linear function of factors.

Bansal and Viswanathan (1993) use a nonlinear arbitrage pricing model to price assets. The arbitrage pricing model implies that only risks related to a few factors are relevant in determining asset prices. Ross's (1976) unconditional version of the arbitrage pricing theory assumes that payoffs are linear in the factors and the idiosyncratic noise. This linearity of payoffs and the no arbitrage restriction leads to an approximate linear relationship between the unconditional mean returns and the factor loadings. Recent studies provide conditions under which an equilibrium arbitrage-pricing relationship holds. In
contrast to Ross's original linear arbitrage pricing model, these models imply that individual security prices are exactly linear in the factor prices. To obtain this result, a linear payoff structure with fixed factor loadings is assumed. However, Bansal and Viswanathan argue that the assumptions of the linear arbitrage pricing theory are unnecessarily restrictive. In particular, the linear payoff assumption implies that only primitive securities having this linear structure can be priced. Derivative securities that are nonlinear functions of the linear payoffs cannot be priced by the linear APT model.

Their approach is also based on Hansen and Jagannathan (1991) who show the existence of a unique minimum variance pricing kernel that is a payoff and characterize it. This pricing kernel is the projection of any valid pricing kernel on the space of payoffs and is the linear kernel. In general, the minimum variance pricing kernel is a conditional linear combination of all payoffs. In the linear APT, this pricing kernel is a conditional linear combination of only the factor payoffs and embodies the refutable restrictions of linearity and low dimensionality.

Bansal and Viswanathan argue that if the payoffs are nonlinear in the factors, these simple representations do not obtain. Nonlinearity could arise because of derivative securities or because the primitive payoffs themselves are nonlinear functions of the factors. In this case, the minimum variance pricing kernel is not a conditional linear combination of the factor payoffs themselves and must include the securities that are nonlinear functions of the factor payoffs.

To accommodate nonlinear payoff structures and retain the restriction of low dimensionality on the pricing kernel, they exploit the theoretical underpinnings
of equilibrium pricing models. The equilibrium restriction says that the intertemporal marginal rate of substitution of some agent is a function of factor risk only. They insist that this approach to asset pricing is a parsimonious one and embeds versions of a number of asset-pricing models like the traditional CAPM, the linear APT, the Merton-Constantinides discrete time ICAPM and the Rubinstein CAPM.

Since the exact form of this nonlinear pricing kernel is unknown, they use the semi-nonparametric estimation approach developed by Gallant and Tauchen (1989) to estimate the nonlinear pricing kernel. They also insist that this approach of estimation embeds a large class of parametric asset pricing models including the traditional CAPM, the linear APT, the Merton-Constantinides discrete time ICAPM and the Rubinstein CAPM.

Bansal, Hsieh and Viswanathan (1993) apply this nonlinear arbitrage pricing model to price international equities, bonds, and forward currency contracts.

Cochrane (1996) examines a factor pricing model for stock returns to see whether cross-sectional and time-series variation in expected stock returns can be explained by investment returns, inferred from investment data via adjustment cost production function. His primary concern is to identify macroeconomic risks that drive asset prices and expected returns. He argues that reduced form models that explain an asset's expected return by its covariance with other asset returns rather than covariance with macroeconomic risks may successfully describe variation in expected returns but they will never explain it. He studies the hypothesis that a factor pricing model holds, that is, the investment returns are factors for the asset returns. While the law of one price implies that there is always a discount factor that is a linear combination of the investment and asset
returns and that prices both, the factor pricing model restriction means that the asset returns can be excluded from this construction. Thus, an investment return factor pricing model says that there exists a discount factor that is a function of only the investment returns and yet prices both asset and investment returns.

Cochrane (1996) distinguishes scaling factors from scaling returns. He suggests a scaled factor pricing model which is represented by \( m \) such that

\[
m_{t+1} = b'(f_{t+1} \otimes z_t)
\]  

(4.4)

where \( b \) is a vector of coefficients, \( f_{t+1} \) is a vector of factors and \( z_t \) is a vector of instrumental variables. The discount factor in this specification is expressed by a linear combination of a given set of factors with weights that vary as a vector of instruments \( z \) varies across different information set. That is, \( m_{t+1} = b(z_t)'f_{t+1} \). The scaling returns is Hansen and Singleton's (1982) instruments in a moment condition. He argues that the inclusion of scaled factor and returns captures variation in conditional betas and factor risk premiums in a very simple structure. While most tests of factor pricing models include auxiliary assumptions, such as constant conditional betas, constant conditional factor risk premiums, constant covariances, or complex time-series models for these quantities, Cochrane argues that his approach does not require any of these assumptions. In Cochrane's approach the factors do not have to be conditionally mean zero, conditionally orthogonal, or conditionally homoskedastic, as is often assumed. Cochrane also proves that the statement that the stochastic discount factor is a linear function of factors is equivalent to the conventional statement of factor pricing models in terms of betas and factor risk premiums.
We apply these conditional models to the analysis of forward risk premium. We first define real forward returns and forward premium. They are respectively

\[ \rho^j = \frac{(S_{t+1}^j - F^j_t)/S_t}{(p_t/p_{t+1})}, \]

\[ \pi^j = \frac{(F^j_t - S_t)/S_t}{(p_t/p_{t+1})}, \]

where \( S_t \) is a nominal spot exchange rate and \( F^j \) is a nominal forward exchange rate for currency \( j \). Exchange rates are US dollar price of each foreign currency. In equilibrium foreign exchange market satisfies

\[ E[m_{t+1} \rho_{t+1} | I,J] = 0, \quad m_{t+1} = b' \tau_{t+1} \quad \text{for all} \ j. \quad (4.5) \]

We use GMM to estimate \( b \). The instrumental variables are lagged forward returns and lagged forward premiums. To obtain a unique vector of estimates for \( b \), we normalize \( b \) by

\[ E[m_{t+1} r_{t+1}^f | I,J] = 1 \quad (4.6) \]

where \( r^f \) is a risk-free return. All variables are real-valued in estimation.

4.4 Estimation of a factor pricing model

We use the spot and forward exchange rate data published in the International Financial Statistics. Spot rates are quarterly end-of-period data
for Japan, Germany, France and Canada. The 3-month forward rates are used. 3-month US Treasury bill rates are assumed to be risk free. Quarterly T-bill rates are obtained by dividing the annualized rates by 4. T-bill rates and the value weighted US stock return data are from the Center for Research of Stock Prices (CRSP) dataset. These stock return data include capital gains and dividends. Monthly data are transformed into quarterly data by multiplying three monthly returns of each quarter. Since all these variables are nominal we deflated them using an implicit US consumption price index. Quarterly US consumption data include nondurable consumption goods and services which we extracted from the Citybase dataset. Per capita consumption data are obtained using population data from the Citybase dataset. We used the third month population of each quarter, excluding army forces abroad. The time period of data series is from 1975:2 to 1990:4.

CAPM implies that \( m \) can be expressed by a linear function of the market return. We specify

\[
m_{t+1} = b_1 + b_2 r^m_{t+1},
\]

(4.7)

where \( r^m_{t+1} \) is the value weighted US stock market return. We assume that US stock market return is a proxy for the world stock market return, \( r^m_{t+1} \).

Table 4.1 shows the results of CAPM estimation. We use different instrumental variables to see the effect of the choice of instrumental variables on the estimation results. They are one-period lagged forward return, two-period lagged forward return, one-period lagged forward premium and two-period lagged forward premium. For any choice of instrumental variables except one-period
lagged forward premium the CAPM of which benchmark return is US stock market returns is not rejected. Lagged forward premiums give significant estimates for the coefficient on the US stock market return, while lagged forward returns do not.

The simple consumption-based CAPM is derived from the equilibrium condition for the representative consumer:

\[
E \left[ \frac{(U'(c_{i+1}) / U'(c_i)) \, \rho_{i+1} \, J}{U'(c_i) \, \rho_{i+1} \, J} \right] = 0,
\]

where \( U' \) is a marginal utility function and \( c_{i+1} \) is real per capita consumption in period \( t+1 \). When we assume a simple CRRA utility function, we have

\[
E \left[ \beta(c_{t+1}/c_t)^{-\gamma} \, \rho_{i+1} \, J \right] = 0, \tag{4.8}
\]

where \( \gamma \) is the risk aversion coefficient. \( \rho_{i+1} \) is the real return to forward speculation in currency \( j \). Table 4.2 shows the estimation results, restricting \( \beta = 1 \). The consumption includes nondurable goods and services. The model is not rejected for any of the instrument variables. Lagged consumptions are included as instrumental variables to compare the results to those of Mark (1985)'s. Inclusion of lagged consumption, not reported here, does not improve the results significantly. One problem with the simple consumption CAPM is the unrealistically large estimate of risk aversion coefficient.\(^1\)

Table 4.3 and 4.4 show the results for \( \beta \) less than 1. In both cases, we have wrong signs for the risk aversion coefficient. Smaller values of \( \beta \) give larger

\(^1\) In our estimation, \( \gamma \) is around 1. These numerically small estimates come from the normalization equation for the risk free rate.
negative and significant estimates for the risk aversion coefficient. On the other hand, \( \beta \) which is larger than 1 brings about larger positive and significant estimates for the risk aversion coefficient. This result is shown in Table 4.5. From these results, we can say that the value of the time preference parameter strongly affects the estimates for the risk aversion coefficient. One reason for these results could be the normalization equation for risk free rates. To identify the coefficients in the discount factors, we need a normalization. Our normalization equation implies that the discount factors should satisfy the Euler equations for the risk free rates. This normalization dramatically reduces the value of the estimated risk aversion coefficient. Without this normalization equation, however, we obtain unrealistically large estimates for the risk aversion coefficient, which are around 80 (See Mehra and Prescott(1985) for the reasonable estimates). However, the restriction \( \beta = 1 \) does not affect the LR-type test results, although we do not report the table here. With the restriction \( \beta = 1 \) and the normalization equation, we obtain more precise estimates for the risk aversion coefficient. Other values for \( \beta \) also give qualitatively similar results for the LR-type tests.

The consumption CAPM with habit formation is applied to the asset return data by Ferson and Constantinides (1991) and Backus, Gregory and Telmer (1993). When we introduce nonseparability into the representative agent's preferences, the agent's expected utility function takes the form

\[
U_t = E_t \sum_{k=0}^{\infty} \beta^k u(d_{t+k})
\]  

(4.9A)

\[
u(d_0) = \frac{(d^{1-\gamma} - 1)}{(1-\gamma)}
\]

(4.9B)
\[ d_t = c_t + \delta c_{t-1}, \quad (4.9C) \]

where \( c_t \) is the current expenditure on consumption goods and \( d_t \) is the service flow. \( \gamma \) is the curvature parameter, which however is not the risk aversion coefficient any more. \( \delta \) is the habit parameter. The utility function is defined only for \( \delta \) that guarantee positive \( d \). For \( \delta > 0 \), preferences exhibit habit formation, in the sense that higher current consumption requires higher future consumption to maintain the same level of future utility. For \( \delta < 0 \), there is durability, since current expenditures raise future utility. If \( \delta = 0 \), we have the additively separable power utility function.

The moment condition takes the form

\[
\frac{E_d[\beta((c_{t+1} + \delta c_t)^{-\gamma} + \beta \delta (c_{t+2} + \delta c_{t+1})^{-\gamma})]ho_{t+1}}{E_d(c_t + \delta c_{t-1})^{-\gamma} + \beta \delta (c_{t+1} + \delta c_t)^{-\gamma}} = 0. \quad (4.10)
\]

By multiplying the denominator to both sides and normalizing both sides to avoid the trivial solution, \( \gamma = 0 \) and \( \delta = 1/\beta \), we obtain the discount factor which can be used for our empirical study:

\[
m_{t+1} = \frac{\beta ((c_{t+1} + \delta c_t)^{-\gamma} + \beta \delta (c_{t+2} + \delta c_{t+1})^{-\gamma})}{(1 + \beta \delta) (c_t + \delta c_{t-1})^{-\gamma}}. \quad (4.11)
\]

To obtain the unique vector of estimates, we use the moment condition for the risk free returns. In this case the disturbance term, \( e_{t+1} \), satisfies
\[ e_{t+1} = \frac{\beta((c_{t+1} + \delta c_{t+2})^{\gamma} + \beta \delta(c_{t+1} + \delta c_{t+1})^{\gamma}) r^f}{(1+\beta \delta)(c_i + \delta c_{t-1})^{\gamma}} - \frac{(c_i + \delta c_{t-1})^{\gamma} + \beta \delta(c_{t+1} + \delta c_{t})^{\gamma}}{(1+\beta \delta)(c_i + \delta c_{t-1})^{\gamma}}. \]  

(4.12B)

Table 4.6 shows the estimation results for the consumption-based CAPM with a time-nonparallel utility. In this table, we do not restrict \( \beta \). The estimates for the time preference parameter \( \beta \) range from 0.157 to 1.067 according to the choice of the instrumental variables, and the estimates for the habit formation coefficient \( \delta \) are negative implying habit formation rather than durability. In Table 4.7, we restrict \( \beta \) to 1. As in table 4.6, the model is not rejected at the 5 percent level, regardless of the choice of the instrumental variables. The estimates for the habit formation coefficient have negative signs. However, the estimates for the habit formation coefficient are -0.98, when the estimates for the curvature parameter are almost zero. It is well known that the curvature parameter becomes smaller than the risk aversion parameter when habit formation is true. But the estimates we obtain are unrealistically small. Thus, in the next experiment, we restrict the curvature parameter to be less than 10 whenever we iterate the estimation until the estimates converge. Table 4.8 presents the results. The estimates for the curvature parameter are bounded by the restriction. In this case, we obtain more reasonable estimates for the habit formation coefficient, which range from -0.609 to -0.637. With the bounded curvature parameter, the model is still not rejected, with the p-values similar to those in the tables for the nonseparable utility. Backus et al. (1993) obtain
positive estimates for the habit formation coefficient, although they are
estimated imprecisely, implying durability rather than habit formation. In our
results they are negative and significant for all values of \( \beta \). Our results are
consistent with Cecchetti, Lam, and Mark (1994) who find that negative \( \delta \) gives
relatively small values for the curvature parameter, while positive \( \delta \), implying
durability, gives unrealistically large values.

4.5 Nested Tests for Risk Premium Models

We compare the risk premium models in section 4.3 using an LR-type test.
Newey and West (1987) provide an LR-type test statistic for the following
question: given factors \( f_1 \), are factors \( f_2 \) important for pricing assets? When we
use the same weighting matrix to estimate the unrestricted and restricted models,
we have the test statistic

\[
T \times \min J_T(\text{restricted}) - T \times \min J_T(\text{unrestricted}) \sim X^2
\]  

(4.13)

where \( T \) is the number of observations, and the degrees of freedom for the chi-
square distribution equals the number of restrictions. We use the weighting
matrix of the unrestricted model for the restricted one.

For the comparison of CAPM and the simple consumption CAPM we estimate

\[
m_{t+1} = b_1 + \beta (c_{t+1}/c_t)^{-b_2} + b_3 m_{t+1} \quad \text{(4.14A)}
\]

\[
m_{t+1} = \beta (c_{t+1}/c_t)^{-b_2} \quad \text{(4.14B)}
\]

96
\[ m_{t+1} = b_1 + b_2 r_{t+1}^{m}, \quad (4.14C) \]

where \( \beta = 1 \), and the second and third equations are restricted models. When US stock market return data is omitted, the restricted model is significantly different from the unrestricted model. However, when we set a restriction to the consumption variable the restricted model is not different from the unrestricted one. This implies that US stock market return can explain what a simple consumption-based CAPM can do about the variation of forward returns, but not vice versa. This might imply that the choice of macro economic variable or the functional form of the stochastic discount factor is not settled yet.

To compare the habit formation model with the CAPM, we use the following specifications for the nested tests.

\[ m_{t+1} = b_1 + b_2 r_{t+1}^{m} + \frac{\beta((c_{t+1} + \delta c_t)^{-\gamma} + \beta \delta(c_{t+2} + \delta c_{t+1})^{-\gamma})}{(1+\beta \delta)(c_t + \delta c_{t-1})^{-\gamma}} \quad (4.15A) \]

\[ m_{t+1} = \frac{\beta((c_{t+1} + \delta c_t)^{-\gamma} + \beta \delta(c_{t+2} + \delta c_{t+1})^{-\gamma})}{(1+\beta \delta)(c_t + \delta c_{t-1})^{-\gamma}} \quad (4.15B) \]

\[ m_{t+1} = b_1 + b_2 r_{t+1}^{m}, \quad (4.15C) \]

where \( \beta, \delta \), and \( \gamma \) are estimated. For the risk-free rate, the GMM disturbance term is the equation \((4.12B)\).

Tables 4.9 and 4.10 show the nested test results for the CAPM and the simple consumption CAPM, assuming \( \beta = 1 \) and \( \beta = 0.995 \), respectively. In both cases CAPM
beats the simple consumption CAPM, in the sense that, given CAPM, the consumption CAPM is not important in pricing the forward returns, but, given the consumption-based CAPM, the traditional CAPM is important. In other words, CAPM has additional explanatory power for asset returns over the simple consumption CAPM does.

However, as shown in table 4.11 and 4.12, the results are partly reversed when we compare the habit formation model to the traditional CAPM. With lagged forward returns as the instrumental variables, the habit formation model explains the forward return better than CAPM does. With lagged forward premiums as the instrumental variables, both are rejected, implying that CAPM does not explain what the habit formation model cannot, and the habit formation model does not explain what CAPM can not.

The traditional CAPM, which uses the US stock market return data, is better in fitting the forward return data than the simple consumption CAPM. With a habit formation model, the results are mixed. When economic agents use the lagged forward returns to form portfolios in the forward exchange rate markets, the habit formation model describes the forward return data more accurately than does CAPM. When agents exploit only the information on the forward premium, CAPM and the habit formation model are exclusive in explaining forward returns.

4.6 Conclusion

In this chapter, we compare the traditional CAPM with a simple consumption-based CAPM and a habit formation model. All asset pricing models differ only in
their formulation of the stochastic discount factors, which can be represented as linear or non-linear functions of factors. Since the traditional CAPM, the consumption-based CAPMs with time-separable utility and the consumption-based CAPMs with time-nonseparable utility can all be regarded as particular cases of an arbitrage pricing model, we can compare the asset pricing models within a nested framework.

One advantage of the consumption-based CAPM is that it enables us to find the relationship between macroeconomic variables and asset returns. Through this relationship, we can explain asset prices rather than merely describing them. Mankiw and Shapiro (1985) show that CAPM fits the domestic stock return data better than the simple consumption CAPM does. They use a simple linear regression model. Unlike them, we use a more general specification for the asset pricing models. We compare the traditional CAPM to the consumption CAPMs with time-separable and time-nonseparable utility functions and find which model explains better the forward returns in the forward exchange rate market.

Each risk premium model is not rejected under particular choice of instrument variables. Nested tests show that the CAPM, with US stock market return data, is better than the simple consumption CAPM, with time-separable utility, for any choice of the instrumental variables. However, when we compare the traditional CAPM to a habit formation model, the results are mixed, depending on the choice of the instrumental variables. This result implies that the variations of macroeconomic variables, which are not reflected in stock market return data, can explain a portion of the variations of the forward exchange rate returns. On the other hand, the stock return data has information that is not present in the consumption data.
<table>
<thead>
<tr>
<th>IV</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$\chi^2$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td></td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>0.949</td>
<td>3.387</td>
<td>21.09</td>
<td>0.134</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(2.620)</td>
<td>(15)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{t-1}$</td>
<td>1.243</td>
<td>-10.99</td>
<td>15.73</td>
<td>0.400</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(2.596)</td>
<td>(15)</td>
<td></td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>0.928</td>
<td>4.540</td>
<td>27.18</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(2.649)</td>
<td>(15)</td>
<td></td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>0.904</td>
<td>8.504</td>
<td>21.69</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>(0.078)</td>
<td>(2.890)</td>
<td>(15)</td>
<td></td>
</tr>
</tbody>
</table>

Column (1): instrumental variables ($z_i$)  
Column (2),(3): standard errors in parentheses  
Column (4): degree of freedom in parentheses  
$E[m_{t+1} | \rho_{t+1} | I_t] = 0$, $m_{t+1} = b_1 + b_2 r_{t+1}$,  
$E[m_{t+1} r_{t+1} | I_t] = 1$

Table 4.1: CAPM with US stock returns
<table>
<thead>
<tr>
<th>IV (1)</th>
<th>( b_I ) (2)</th>
<th>( \chi^2 ) (3)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_t )</td>
<td>1.436 (0.237)</td>
<td>22.91 (16)</td>
<td>0.116</td>
</tr>
<tr>
<td>( \rho_{t-1} )</td>
<td>1.616 (0.258)</td>
<td>17.33 (16)</td>
<td>0.364</td>
</tr>
<tr>
<td>( \pi_t )</td>
<td>0.997 (0.206)</td>
<td>31.34 (16)</td>
<td>0.012</td>
</tr>
<tr>
<td>( \pi_{t-1} )</td>
<td>1.117 (0.212)</td>
<td>26.26 (16)</td>
<td>0.050</td>
</tr>
</tbody>
</table>

| Column (1) | instrumental variables (\( z_i \)) |
| Column (2) | standard errors in parentheses |
| Column (3) | degree of freedom in parentheses |

\[
\begin{align*}
E[m_{t+1}\rho_{t+1}z_t | I_t] &= 0, \\
m_{t+1} = \beta(c_{t+1}/c_t)^{-\gamma}, \\
E[m_{t+1}r_{t+1}^I | I_t] &= 1
\end{align*}
\]

Table 4.2: Consumption CAPM: time-separable(\( \beta = 1 \))
<table>
<thead>
<tr>
<th>IV</th>
<th>$b_I$</th>
<th>$\chi^2$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_t$</td>
<td>-1.522</td>
<td>21.82</td>
<td>0.149</td>
</tr>
<tr>
<td></td>
<td>(0.259)</td>
<td>(16)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{-1}$</td>
<td>-1.248</td>
<td>17.63</td>
<td>0.346</td>
</tr>
<tr>
<td></td>
<td>(0.263)</td>
<td>(16)</td>
<td></td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>-1.496</td>
<td>32.89</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.274)</td>
<td>(16)</td>
<td></td>
</tr>
<tr>
<td>$\pi_{-1}$</td>
<td>-1.236</td>
<td>23.64</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td>(0.257)</td>
<td>(16)</td>
<td></td>
</tr>
</tbody>
</table>

Column (1): instrumental variables ($z_i$)
Column (2): standard errors in parentheses
Column (3): degree of freedom in parentheses

$E[m_{t+1}\rho_{i+1}I_i] = 0, m_{t+1} = \beta(c_{t+1}/c_t)^{\beta_1}$,
$E[m_{t+1}r_{t+1}I_i] = 1$

Table 4.3: Consumption CAPM: time-separable ($\beta = 0.99$)
<table>
<thead>
<tr>
<th>IV</th>
<th>$b_1$</th>
<th>$\chi^2$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_t$</td>
<td>-0.094</td>
<td>21.77</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>(0.181)</td>
<td>(16)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{t-1}$</td>
<td>0.166</td>
<td>17.65</td>
<td>0.344</td>
</tr>
<tr>
<td></td>
<td>(0.201)</td>
<td>(16)</td>
<td></td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>-0.249</td>
<td>32.89</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.192)</td>
<td>(16)</td>
<td></td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>-0.026</td>
<td>23.62</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.182)</td>
<td>(16)</td>
<td></td>
</tr>
</tbody>
</table>

Column (1) : instrumental variables ($z_t$)  
Column (2) : standard errors in parentheses  
Column (3) : degree of freedom in parentheses  

$E[m_{t+1}|p_{t+1}x_{t+1}\mid I_t] = 0$,  
$E[r_{t+1}|c_{t+1}\mid I_t] = 0$,  
$E[(c_{t+1}/c_t)^\beta_{t+1}] = 1$  

Table 4.4: Consumption CAPM: time-separable ($\beta=0.995$)
<table>
<thead>
<tr>
<th>IV (1)</th>
<th>$b_I$ (2)</th>
<th>$\chi^2$ (3)</th>
<th>P-value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_t$</td>
<td>2.781 (0.353)</td>
<td>21.67 (16)</td>
<td>0.154</td>
<td></td>
</tr>
<tr>
<td>$\rho_{t-1}$</td>
<td>2.999 (0.394)</td>
<td>17.69 (16)</td>
<td>0.342</td>
<td></td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>2.251 (0.310)</td>
<td>32.90 (16)</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>2.401 (0.331)</td>
<td>23.57 (16)</td>
<td>0.099</td>
<td></td>
</tr>
</tbody>
</table>

Column (1): instrumental variables ($z_t$)
Column (2): standard errors in parentheses
Column (3): degree of freedom in parentheses

$E[m_{t+1}|\pi_{t+1}|I_t] = 0$, $m_{t+1} = \beta(c_{t+1}/c_t)^{-\beta_1}$,
$E[m_{t+1}|r_{t+1}|I_t] = 1$  

Table 4.5: Consumption CAPM: time-separable ($\beta = 1.005$)
<table>
<thead>
<tr>
<th>IV (1)</th>
<th>$\beta$ (2)</th>
<th>$\delta$ (3)</th>
<th>$\gamma$ (4)</th>
<th>$\chi^2$ (5)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_t$</td>
<td>1.067 (0.092)</td>
<td>-0.242 (0.176)</td>
<td>80.30 (47.12)</td>
<td>12.74 (14)</td>
<td>0.547</td>
</tr>
<tr>
<td>$\rho_{t-1}$</td>
<td>0.157 (0.061)</td>
<td>-0.713 (0.038)</td>
<td>100.0 (25.41)</td>
<td>12.79 (14)</td>
<td>0.543</td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>0.997 (0.022)</td>
<td>-0.982 (0.002)</td>
<td>0.0001 (0.014)</td>
<td>19.82 (14)</td>
<td>0.136</td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>1.033 (0.020)</td>
<td>-0.982 (0.000)</td>
<td>-0.002 (0.003)</td>
<td>16.83 (14)</td>
<td>0.265</td>
</tr>
</tbody>
</table>

Column (1): instrumental variables ($z_t$)
Column (2), (3) & (4): standard errors in parentheses
Column (4): degree of freedom in parentheses
$E[m_{t+1}|\rho_{t+1}, z_t, L_t] = 0,$

$$m_{t+1} = \frac{\beta((c_{t+1} + \delta c_{t+2})^{\gamma} + \beta \delta (c_{t+1} + \delta c_{t+2})^{\gamma})}{(1 + \beta \delta)(c_t + \delta c_{t-1})^{\gamma}}$$

For the risk free rate, see section 4.4.

Table 4.6: Consumption CAPM: time-nonseparable
<table>
<thead>
<tr>
<th>IV</th>
<th>$\delta$</th>
<th>$\gamma$</th>
<th>$\chi^2$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_t$</td>
<td>-0.574 (0.346)</td>
<td>6.182 (6.661)</td>
<td>13.42 (15)</td>
<td>0.569</td>
</tr>
<tr>
<td>$\rho_{t-1}$</td>
<td>-0.982 (0.000)</td>
<td>0.001 (0.001)</td>
<td>13.37 (15)</td>
<td>0.574</td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>-0.982 (0.000)</td>
<td>0.002 (0.001)</td>
<td>18.06 (15)</td>
<td>0.259</td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>-0.982 (0.000)</td>
<td>0.004 (0.001)</td>
<td>15.94 (15)</td>
<td>0.386</td>
</tr>
</tbody>
</table>

Column (1) : instrumental variables ($z_t$)
Column (2) & (3) : standard errors in parentheses
Column (4) : degree of freedom in parentheses

$E[m_{t+1} | \pi_{t+1} z_t | 1_t] = 0$,

$$m_{t+1} = \frac{\beta((c_{t+1} + \delta c_1)^{-\gamma} + \beta \delta (c_{t+2} + \delta c_{t+1})^{-\gamma})}{(1 + \beta \delta)(c_t + \delta c_{t-1})^{-\gamma}}$$

For the risk free rate, see section 4.4.

Table 4.7: Consumption CAPM: time-nonseparable ($\beta = 1$)
<table>
<thead>
<tr>
<th>IV (1)</th>
<th>$\delta$ (2)</th>
<th>$\gamma$ (3)</th>
<th>$\chi^2$ (4)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_i$</td>
<td>-0.625 (0.081)</td>
<td>9.999 (0.000)</td>
<td>16.38 (15)</td>
<td>0.357</td>
</tr>
<tr>
<td>$\rho_{t-1}$</td>
<td>-0.609 (0.082)</td>
<td>9.999 (0.000)</td>
<td>13.82 (15)</td>
<td>0.539</td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>-0.637 (0.053)</td>
<td>9.999 (0.000)</td>
<td>18.06 (15)</td>
<td>0.259</td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>-0.615 (0.096)</td>
<td>9.999 (0.000)</td>
<td>15.95 (15)</td>
<td>0.385</td>
</tr>
</tbody>
</table>

Column (1) : instrumental variables ($z_i$)
Column (2) & (3) : standard errors in parentheses
Column (4) : degree of freedom in parentheses

$$E[m_{t+1} \mid \rho_{t+1}, z_i] = 0,$$

$$m_{t+1} = \frac{\beta((c_{t+1} + \delta c_t)^{-\gamma} + \beta \delta (c_{t+2} + \delta c_{t+1})^{-\gamma})}{(1+\beta \delta)(c_t + \delta c_{t-1})^{-\gamma}}.$$

For the risk free rate, see section 4.4.

Table 4.8: Consumption CAPM: time-nonseparable ($\beta=1$, $\gamma<10$)
<table>
<thead>
<tr>
<th>IV</th>
<th>unrestricted</th>
<th></th>
<th>restricted(A)</th>
<th></th>
<th>restricted(B)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>$P$-value</td>
<td>$C_A$</td>
<td>$P$-value</td>
<td>$C_B$</td>
<td>$P$-value</td>
</tr>
<tr>
<td>$\rho_t$</td>
<td>21.50</td>
<td>0.089</td>
<td>4.65</td>
<td>0.036</td>
<td>3.72</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>(14)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{t-1}$</td>
<td>14.08</td>
<td>0.443</td>
<td>8.53</td>
<td>0.003</td>
<td>0.91</td>
<td>0.339</td>
</tr>
<tr>
<td></td>
<td>(14)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>26.79</td>
<td>0.020</td>
<td>3.05</td>
<td>0.080</td>
<td>0.11</td>
<td>0.731</td>
</tr>
<tr>
<td></td>
<td>(14)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>17.89</td>
<td>0.211</td>
<td>7.53</td>
<td>0.006</td>
<td>3.21</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>(14)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(1)</td>
<td></td>
</tr>
</tbody>
</table>

Degree of freedom in parentheses
$E[m_{t+1}|p_{t+1}|z_t|I_t] = 0$, $E[m_{t+1}|r_{t+1}|I_t] = 1$
unrestricted model: $m_{t+1} = b_1 + \beta (c_{t+1}/c) - b_2 + b_3 r_{t+1}$
restricted model(A): $m_{t+1} = \beta (c_{t+1}/c) - b_1$
restricted model(B): $m_{t+1} = b_1 + b_2 r_{t+1}$

Table 4.9: LR-type Test : CAPM vs. Consumption CAPM
(time-separable, $\beta = 1$)
<table>
<thead>
<tr>
<th>IV</th>
<th>unrestricted</th>
<th></th>
<th>restricted</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>$P$-value</td>
<td>$C_A$</td>
<td>$P$-value</td>
<td>$C_B$</td>
<td>$P$-value</td>
</tr>
<tr>
<td>$\rho_t$</td>
<td>15.64</td>
<td>0.336</td>
<td>7.61</td>
<td>0.022</td>
<td>5.14</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td>(14)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{t-1}$</td>
<td>12.86</td>
<td>0.537</td>
<td>7.56</td>
<td>0.006</td>
<td>0.71</td>
<td>0.398</td>
</tr>
<tr>
<td></td>
<td>(14)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>28.36</td>
<td>0.013</td>
<td>3.75</td>
<td>0.052</td>
<td>0.01</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td>(14)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>17.37</td>
<td>0.237</td>
<td>6.49</td>
<td>0.011</td>
<td>0.06</td>
<td>0.799</td>
</tr>
<tr>
<td></td>
<td>(14)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(1)</td>
<td></td>
</tr>
</tbody>
</table>

Degree of freedom in parentheses
$E[m_{t+1}|i_{t+1}|I_t] = 0$, $E[\beta r_{t+1}|I_t] = 1$
unrestricted model: $m_{t+1} = b_1 + \beta(c_{t+1}/c_t)^{-b_2} + b_3 r_{t+1}$
restricted model(A): $m_{t+1} = \beta(c_{t+1}/c_t)^{-b_1}$
restricted model(B): $m_{t+1} = b_1 + b_2 r_{t+1}$

Table 4.10: LR-type Test: CAPM vs. Consumption CAPM (time-separable, $\beta=0.995$)
<table>
<thead>
<tr>
<th>IV</th>
<th>unrestricted</th>
<th></th>
<th>restricted(A)</th>
<th></th>
<th>restricted(B)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>$P$-value</td>
<td>$C_A$</td>
<td>$P$-value</td>
<td>$C_B$</td>
<td>$P$-value</td>
</tr>
<tr>
<td>$\rho_t$</td>
<td>9.23</td>
<td>0.755</td>
<td>0.96</td>
<td>0.616</td>
<td>20.5</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{t-1}$</td>
<td>12.58</td>
<td>0.480</td>
<td>5.17</td>
<td>0.075</td>
<td>6.74</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>17.68</td>
<td>0.169</td>
<td>18.4</td>
<td>0.0001</td>
<td>8.74</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>15.29</td>
<td>0.289</td>
<td>18.6</td>
<td>0.000</td>
<td>22.8</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
</tbody>
</table>

Degrees of freedom in parentheses
$E[m_{t+1}|p_{t+1}z_{t}|I_t] = 0$.
unrestricted model: $m_{t+1} = b_1 + b_2r_{t+1}$

\[
+ \frac{\beta((c_{t+1} + \delta c_{t})^{\gamma} + \beta \delta (c_{t+2} + \delta c_{t+1})^{\gamma})}{(1 + \beta \delta)(c_t + \delta c_{t-1})^{\gamma}}
\]

restricted model(A): $m_{t+1} = \frac{\beta((c_{t+1} + \delta c_{t})^{\gamma} + \beta \delta (c_{t+2} + \delta c_{t+1})^{\gamma})}{(1 + \beta \delta)(c_t + \delta c_{t-1})^{\gamma}}$

restricted model(B): $m_{t+1} = b_1 + b_2r_{t+1}$

Table 4.11: LR-type Test: CAPM vs. Consumption CAPM
(time-nonseparable, $\beta=1$)
<table>
<thead>
<tr>
<th>IV</th>
<th>unrestricted</th>
<th></th>
<th>restricted(A)</th>
<th></th>
<th>restricted(B)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>$P$-value</td>
<td>$C_A$</td>
<td>$P$-value</td>
<td>$C_B$</td>
<td>$P$-value</td>
</tr>
<tr>
<td>$p_t$</td>
<td>9.87</td>
<td>0.704</td>
<td>1.98</td>
<td>0.372</td>
<td>19.50</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>$p_{t-1}$</td>
<td>12.68</td>
<td>0.472</td>
<td>5.17</td>
<td>0.075</td>
<td>6.53</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>17.50</td>
<td>0.177</td>
<td>18.65</td>
<td>0.000</td>
<td>8.90</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>$\pi_{t-1}$</td>
<td>15.33</td>
<td>0.287</td>
<td>18.56</td>
<td>0.000</td>
<td>22.90</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td></td>
<td>(2)</td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
</tbody>
</table>

* degree of freedom in parentheses

$E[m_{t+1}|p_{t+1}^j|I_t] = 0$,

unrestricted model: $m_{t+1} = b_1 + b_2 r_{t+1}^m$

$$+ \frac{\beta((c_{t+1} + \delta c^m_{t+1})^{\gamma} + \beta \delta (c_{t+2} + \delta c_{t+1})^{\gamma})}{(1+\beta \delta)(c_t + \delta c_{t-1})^{\gamma}}$$

restricted model(A): $m_{t+1} = \frac{\beta((c_{t+1} + \delta c^m)^{\gamma} + \beta \delta (c_{t+2} + \delta c_{t+1})^{\gamma})}{(1+\beta \delta)(c_t + \delta c_{t-1})^{\gamma}}$

restricted model(B): $m_{t+1} = b_1 + b_2 r_{t+1}^m$

Table 4.12: LR-type Test : CAPM vs. Consumption CAPM
(time-nonseparable, $\beta=0.99$)
CHAPTER 5

SUMMARY AND CONCLUSION

To estimate the speed of price adjustment in dynamic models many studies employ error correction models. We point out that, since ECM is a reduced form, the use of the traditional ECM fails to identify the structural parameter for the short-run dynamics, unless some strong assumptions are imposed on the coefficients of the ECM. To overcome this weakness of the traditional ECM, we consider a structural ECM and present instrumental variable methods to estimate the speed of adjustment in the structural ECM, exploiting the different speeds of convergence of long-run and short-run estimators as in Engle and Granger (1987).

As a simple illustration, we derive a structural ECM from a version of Mussa's (1982) dynamic model of exchange rate determination. We apply an instrumental variable method to estimate the speed of adjustment in the structural ECM. When the long-run coefficient is assumed to be known, we apply the instrumental variable method to the structural ECM with given long-run coefficient. When the long-run coefficient is not known, we first estimate long-run coefficients in the structural model and then apply an instrumental variable method to the error correction model to obtain an estimate for the structural speed of price adjustment.
In the single equation method, with a known cointegrating vector, we obtain positive estimates for the short-run coefficient for most countries, implying half lives of 2.3 to 5.2 years which are consistent with the literature. For the sensitivity analysis, we assume that the cointegrating vectors are unknown. With the estimated cointegrating vector, we also obtain positive estimates for the short-run coefficients for most countries. In such cases the estimates are significant, and imply shorter half lives ranging from 1.3 to 3.5 years. Negative estimates for the short-run coefficient are always insignificant, whether the cointegrating vectors are known or unknown.

When the system method, which imposes money market equilibrium and uncovered interest rate parity, is applied to the exchange rate model, the speed of adjustment coefficient is estimated from both the slow adjustment equation for the domestic price and from the rational expectations equation for the exchange rate. In most cases, the restriction that these two rates of adjustment are equal is not rejected at the five percent level. In some cases, the restricted estimate is significantly negative, indicating that the restricted model is misspecified.

In chapter 4, we relax the assumption of uncovered interest rate parity. However, instead of adding a risk premium term to the sticky price exchange rate model in chapter 3, we consider risk premium in a more general asset pricing setting consistent with intertemporal choice models. We compare various risk premium models to find which model is more consistent with the forward return data. In particular, we compare the traditional CAPM with both a simple consumption-based CAPM and a habit formation model, treating them as special cases of a factor pricing model. Our results show that the CAPM with US stock
market return data is better than the simple consumption CAPM with time-separable utility. When we compare the traditional CAPM to a habit formation model, however, the results are mixed, depending on the choice of the instrumental variables. This result implies that the variations of macroeconomic variables, which are not reflected in stock market return data can explain a portion of the variations of the forward exchange rate returns.

For a future study, we can consider a way of combining risk premium with a sticky price model, which is derived from an intertemporal choice model, overcoming the lack of microfoundations in the sticky price model considered in chapter 2 and 3. Linearization can allow for the use of the instrumental variable methods applied to a structural ECM.


Dornbusch, Rudiger (1976), "Expectations and Exchange Rate Dynamics," *Journal of Political Economy* 84, (December), 1161-76.


Hsiao, Chen (1995), "Cointegration and Dynamic Simultaneous Equations Model," manuscript, University of Southern California.


