The Pennsylvania State University

The Graduate School

Department of Economics

ECONOMETRICS ANALYSIS OF A CASH-IN-ADVANCE MODEL

A Thesis in

Economics

by

Hung-Pin Lai

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2003

The thesis of Hung-Pin Lai has been reviewed and approved* by the following:

Herman J. Bierens
Professor of Economics
Thesis Advisor
Chair of Committee

Neil Wallace
Professor of Economics
Graduate Program Officer

Joris Pinkse
Associate Professor of Economics

Gustavo Ventura
Assistant Professor of Economics

Darren Frechette
Assistant Professor of Agricultural Economics

*Signatures are on file in the Graduate School

ABSTRACT

The goal of this dissertation is to illustrate how the standard cash-in-advance model can be adapted to study the role of monetary and technology shocks in business cycles, and the analysis presented can also be easily extended to variety of dynamic stochastic equilibrium models. I derived a linearized CIA model that provides a link between the deep parameters and model coefficients. Through the linearized model I then assess the empirical performance of the theoretical model and estimate the deep parameters. The 90% confidence intervals of the deep parameters are also obtained by a bootstrap approach. The effects of the monetary and technology shocks are further analyzed by innovation response analysis. Finally, the distortional effects of the Hodrick-Prescott and rational square-wave filters are also examined.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am most grateful and indebted to my thesis advisor, Herman Bierens, for all of the patience, guidance and encouragement he has shown me during my time here at Penn State. I also thank my committee members for their comments and administering the Comprehensive and Final Oral Examinations. There are so many colleagues, friends and relatives to thank here in USA and in my country. I would like to give my sincerely thanks to all of them. Finally, I would like to thank my husband for all of his support.

Chapter 1

Introduction

The goal of this thesis is to investigate the empirical performance of a cash-in-advance (CIA) model and to illustrate how it can be adapted to study the role of monetary and technology shocks in business cycles using an econometric approach. The most direct motivation for this analysis comes from Lucas and Stokey (1987). They state that a theoretical model should permit the calculation of a predicted theoretical joint distribution of shocks and endongenous variables that can be compared to observed distributions. Although macroeconomic theorists have proposd a variety of theoretical models, there is little empirical investigation of the joint distribution of shocks and endogenous variables of these models. Another motivation for this analysis is the empirical failure of nonmonetary, dynamic, stochastic models in model specifications tests. Therefore, I consider a monetary model, where money is introduced by a cash-in-advance constraint, that may provide a better explanation for the observed data than a barter economy.

The most common approach used in assessing the performance of CIA models is calibration. Calibration compares the moments of artificial data generated by the theoretical model with those of the observed data. However, Dejong, Ingram, and Whiteman (1996 and 2000) criticized this approach for lacking a formal statistical foundation because the deep parameters are chosen instead of estimated.

There are few empirical studies which have evaluated the suitability of the CIA model or suggest how it can be used as an analytical tool for business cycles. Part of the reason is that it is difficult to obtain an appropriately estimable model that not only represents the theoretical model but also provides a link between the deep parameters and the model coefficients. Moreover, since the theoretical models usually have only one or a few stochastic shocks, the joint distribution of the model endogenous variables is singular. Singularity means that the variance of the model distribution in some particular directions is zero, so the joint density does not exist in this case. Hence, even if a suitable model is used, it is still difficult to use econometric methods, such as maximum likelihood, to investigate the empirical implications of these models.

Now I discuss how one can investigate the empirical performance of a CIA model and illustrate how a CIA model can be adapted to study the role of monetary and technology shocks in business cycles using econometrics. First, I use the methods suggested by Kydland and Prescott (1982), Hansen and Prescott (1995), and McGrattan (1990) to derive a linearized CIA model that links the deep parameters of the theoretical model to an estimable theoretical model with a singular distribution. Different econometric techniques are then used to estimate the linearized version of the model. The role of stochastic shocks in business cycles can then be examined.

I first investigate the empirical relevance of this model using the approach of Bierens and Swanson (2000). I evaluate the CIA model in the nonsingular direction by the Bierens-Swanson average conditional reality bound. In other words, how much information is contained in the theoretical model is evaluated without penalizing the possibility of model misspecification or missing variables.

Furthermore, I apply the Bierens multiplicative conditional reality bound to estimate the deep parameters of the theoretical model. The linearized version of the CIA model can be used to provide predictions about the joint stochastic properties of the endogenous variables, e.g. investment, consumption, inflation rate, and the output growth rate. Their innovation responses to monetary and technology shocks are also investigated. Using the bootstrap, confidence intervals for the deep parameters can also be obtained.

Finally, I discuss the effects of applying the Hodrick-Prescott (H-P) filter and a rational square-wave filter on the CIA model. In the last stage of calibration, it is common to use the H-P filter to remove the stochastic trend and then compare the moments of the artificial and observed data. However, some studies, such as Harvey and Jaeger (1993) and Cogley and Nason (1995), show that the H-P filter can generate spurious business cycles even if none are present in the original data. In order to understand the distortional effect of the filters, the Bierens-Swanson average conditional reality bound is applied to evaluate the possible damage of the filters.

The rest of this thesis is organized as follows. Chapter 2 reviews the relevant empirical and theoretical work and discuss their findings. Chapter 3 introduces the cash-in-advance model and shows that the linearized model implies a cointegrated error correction model. The linearized model will be used to represent the CIA model in our application and analysis. Chapter 4 introduces the two econometric approaches, the Bierens-Swanson average conditional reality bound and the Bierens multiplicative conditional reality bound, which will be used to evaluate and estimate the CIA model. Chapter 5 presents the empirical results and discusses the effects of technology and monetary shocks on inflation and output growth. Chapter 6 considers the H-P filter and Pollock's rational square-wave filter and examines the effects of the filters. Chapter 7 summarizes and concludes.

Chapter 2

Review of the literature

This chapter reviews the relevant theoretical and empirical literature. Section 2.1 reviews some empirical work of cash-in-advance models. Section 2.2 and 2.3 discuss the recent work on measurement of the reality of theoretical models and estimation of the deep parameters. Section 2.4 discusses the motivation of using the average and multiplicative conditional reality bound. Finally, Section 2.5 highlights and summarizes the findings about applying the Hordick-Prescott filter on data.

2.1 Money and growth

Contemporary thought on the topic of money and growth originally begins with the work of Tobin (1965). His model predicts that higher inflation is associated with a larger capital stock and also higher per capita output because saving can only be divided into money and capital, and inflation lowers the real return on money. In Sidrauski's (1967) model, money is superneutral. Monetary growth rate and inflation have no effects on steady-state output and capital since the long-run capital stock depends only on its depreciation rate, population growth, and discount rate. Moreover, in the CIA model studied by Stockman (1981), inflation acts as a tax on productivity and therefore impedes capital accumulation. Subsequent work on money and growth has focused on the effects of inflation on capital and output. For instance, Ireland (1994) discusses the issue of how sustained capital accumulation affects money's role and growth. Dotsey and Sarte (2000) use a CIA model to analyze the effects of inflation variability on output growth.

Despite abundant discussion of the cash-in-advance models, there is only little empirical investigation of them. Hordick, Kocherlakota, and Lucas (1991) use a calibration technique to study CIA models. They found that the first and second moments of the artificial data generated by such models cannot match that in the empirical data, and thus did not proceed with a further formal estimation. Finn, Hoffman, and Schlagenhauff (1990) use serval versions of CIA models to explore whether liquid services and nonsuperneutral effects of money are important for asset pricing. Sill (1998) uses a generalized method of moments to estimate parameters from first-order conditions of a CIA model. The responses of money demand and velocity to exogenous shocks to monetary growth and income growth are also examined and compared with that of an unrestricted VAR.

2.2 Measurement of the reality of a theoretical model

Lately the issue in assessing the empirical relevance of theoretical models, has been widely debated. Calibration of models is markedly different from the standard econometrics approach, and has been widely used in evaluating business cycle models by most macroeconomists.

However, this technique has been criticized for a few of its aspects. First, a model is evaluated by examining the moments of the artificial data generated by calibration experiments, but matching the major moments of the data is not necessary the most desirable feature of a model, because even a complicated model is still a simplified description of reality. Secondly, use of a calibration technique to evaluate models is not a test against an alternative; therefore, it cannot be used to compare whether another different model can match the data better. Third, it lacks a formal statistical foundation. The parameter values of the theoretical model are selected, instead of estimated, on the basis of data evidence. For example, the rate of population growth in a business cycle model is set according to data statistics, rather than estimated by an econometric method. In other words, the parameter values selected in the beginning of calibration are not necessarily the ones that provide the best fit of the model in some statistical sense.

Dejong, Ingram, and Whiteman (1996 and 2000) (hereafter DIW) and Geweke (1999) develop a Bayesian approach, different from calibration, that provides a statistical framework for combining the theoretical and empirical models. Since calibration does not involve realistic specifications of uncertainty regarding the parameterization of the model in question, they suggest incorporating the uncertainty regarding the parameters of the theoretical model in the calibration technique. They use prior probability distributions to represent the uncertainty of the parameters of the theoretical model and then compute the implied distributions over statistical properties of artificial data simulated by the model.

Alternatively, Watson (1993) also suggests a procedure for evaluating the fit of a calibrated theoretical model. By augmenting the variables in the model with enough stochastic error, the second moment of the simulated data can be made to exactly match that of the observed data. The greater the required error, the worse is the fit of the theoretical model.

2.3 Estimation of the deep parameters based on data

In addition to evaluating the theoretical models, estimating the deep parameters of the theoretical model has also lately become an important issue in the analysis of theoretical macro-models. In standard calibration experiments, it has been criticized that the deep parameters are chosen based on the data evidence and thus lack statistical foundation.

DIW and Geweke (1999) suggest a Bayesian approach to incorporate the calibration technique into a statistical framework. However, two main limitations of the Bayesian approach are pointed out by Bierens (2003). First, the Bayesian approach assumes the theoretical model represents the data generating process (DGP), which is not a realistic assumption. Second, this approach requires the existence of the conditional densities of the observables, but it is not the case if the theoretical model is driven by only a few stochastic shocks, which leads to singularity of the model distribution.

Traditional econometric approaches, such as the maximal likelihood method, thus no longer apply because the conditional density of the theoretical model does not exist under this circumstance. DIW and Geweke (1999) did not deal with singularity when applying the Bayesian approach. DIW focus on a subset of the variables such that their distribution is not singular and Geweke (1999) uses a risk premium model and concentrate on one variable only.

Christiano and Eichenbaum (1992) suggest to use a generalized method of moments (GMM) to estimate the deep parameters by the first order conditions in the first stage of calibration. They ignore the singularity problem by implicitly adding error terms in the first order conditions directly when applying GMM. Moreover, their approach still relies on the unrealistic assumption that the theoretical model represents the DGP.

2.4 Motivation to use the average and multiplicative conditional reality bound

Most model specification tests rely on a null hypothesis that the theoretical model is equal to the DGP. Since the true data comes from a much more complex real economy and observables are often restricted to a degenerate space of lower dimension in the theoretical model, it is not surprising that business cycle models are easily to be rejected by classical model specification tests. Therefore, a model that fits data well along every dimension except one unimportant one might be overwhelmingly rejected statistically. Interpretation of the criteria for rejection or lack of rejection is thus not so obvious under this situation.

The average and multiplicative conditional reality bounds introduced in Bierens and Swanson (2000) and Bierens (2003) have some good properties which other measures do not have. First, it not a test against an alternative. The theoretical model is treated as only an approximation of the DGP, instead of the DGP itself. It can be used as a criterion in choosing models.

Second, this measure represents the probability that the theoretical model is correctly specified or a criterion of how much information about the data-generating process is contained in the theoretical model. The conditional density of a theoretical model is embedded in the true conditional density by a probability, which is estimated by comparing the theoretical and the true densities. Moreover, the scale is between 0 and 1.

Third, the multiplicative reality bound in Bierens (2003) can be estimated in a way similar to the maximal likelihood method. The deep parameters of the theoretical model link to the coefficients in the econometric model derived from the theoretical model, thus then can be estimated by maximizing the objective function. It not only provides an alternative to calibration, but also provides the statistical inference of the parameters.

The main advantage of Bierens'(2003) approach over Bierens and Swanson's (2000) is that the latter focuses on the nonsingular part of the model, so that some links between the deep parameters and model coefficients will be lost. In other words, the deep parameters are in fact estimated by using a subset of the equations which results in lack of identification of some deep parameters. Although the average conditional reality bound can still be obtained by maximizing the objective function, lack of identification cannot be detected during the whole process in estimation.

In order to identify the deep parameters and further solve for them, Bierens (2003) then suggests using convolutions of the distributions of the theoretical model and the data-generating process with a non-singular distribution. The same non-singular noise is added to both distributions, so that they are non-singular and can be compared in a way similar to the previous approach without marginalization in the singular direction. This approach is related to Watson's (1993) approach in the sense that Watson (1993) suggests augmenting the theoretical variables with enough stochastic errors such that the model can match the second moments of the actual data.

2.5 The Hodrick-Prescott Filter

Since many statistical procedures assumes stationarity, it is often considered necessary to transform data into stationary series by removing the trends. For example, the second moments of the series should be finite if we are interested in their cross-correlations. Among a number of transformations H-P filter is the most popular method and has been widely applied in most real business cycle analysis. For example, Kydland and Prescott (1982) and Cooley and Hansen (1989) use this technique to remove trend in the data before calibration.

In the work of Harvey and Jaeger (1993) and Cogley and Nason (1995), they show that detrending based on the H-P filter can lead investigators to report spurious cyclical behavior. Cogley and Nason (1995) further argue that standard real business cycle models do not generate business cycle dynamics in pre-filtered data and that the business cycles observed in H-P filtered data are due to the filter. The 'stylized facts' are in fact 'stylized artifacts', resulted from the filter, not the time series itself.

Similarly, Pollock (2000) and Pedersen (2001) criticized that the H-P filter has some leakage. Pedersen (2001) argues that the H-P filter passes some frequencies which it was supposed to impede and dampen the component which it was supposed to pass since it is only a close approximation of a high-pass filter, which cuts off components with frequencies below the specified cut-off frequency and leaves components with higher frequencies unchanged. This deficiency is the main reason that the H-P filter induces spurious cycles in detrended data.

Chapter 3

The cash-in-advance model

This chapter develops an estimable econometric model from Cooley and Hansen's (1989) cash-in-advance model. The linearized model is obtained by applying the Taylor's expansion to the nonlinear objective function and constraints. The linearized model will be used to analyze the effect of technology and monetary shocks on inflation and output growth.

Section 2.1 introduces the Cooley and Hansen model where money is introduced by a cash-in-advance constraint. Section 2.2 defines the equilibrium. Section 2.3 explains the approach to solve this dynamic model and introduces part of Cooley and Hansen's result. Section 2.4 linearizes the cash-in-advance model and section 2.5 shows that the linearized model is equivalent to a cointegrated error correction model.

3.1 The Cooley and Hansen model

One of the major differences between the Cooley and Hansen model and the standard business cycle model is that the former includes money, introduced via a cash-in-advance constraint[1]. Another distinction is that, in addition to a technology shock, a monetary shock is also included in this model. Further, labor is assumed to be indivisible instead of a continuous variable between 0 and 1, which follows the assumptions in Hansen (1985).

To begin with, let E_t be the conditional expectation operator given the information at time t and let $\bar{\beta} \in (0, 1)$ be the discount rate. All agents in this economy maximize their lifetime utilities based on all information available at time t. Each agent's utility depends on nonstorable consumption c_t and working hours h_t. The utility function takes the form

$$u(c_t, h_t) = \ln c_t - Bh_{t,}$$

[1] The cash-in-advance constraint is based on the assumption that cash balances are required for the purchase of cash goods. In this model economy, two categories of goods are considered: cash goods and credit goods. Cash goods include only consumption goods here. We can imagine that consumption goods are exchanged in circumstances where the buyer is unknown to the seller, so that the seller is unwilling to accept invoices or trade credit as payments issued in securities trading. Such goods, if purchased at all, must be paid for with currency acquired in advance at the market. The other goods, such as labor, are treated as credit goods.

where $B > 0$ is a constant and u is strictly concave, continuous, differentiable, increasing in c_t and, of course, decreasing in h_t.

The capital stock follows the law of motion

$$k_{t+1} = (1-\delta)k_t + x_t$$

with a depreciation rate $\delta \in (0,1)$, where k_t represents capital stock in period t and x_t represents investment, the portion of output that is not consumed. The capital stock k_t is owned by the household, and the investment is used to replace the depreciation and the augmentation of capital stock. Therefore, $c_t + x_t \leq Y_t$, where Y_t is the per capita output.

Given the price level p_t, the real wage rate w_t, and the real capital rental rate r_t, each agent can spend his income on consumption and investment, or keep some cash and carry it to the next period. His income mainly comes from his wage payroll, capital return, government's lump-sum transfer, so his budget constraint can be written as

$$c_t + x_t + m_t/p_t \leq w_t h_t + r_t k_t + (m_{t-1} + (g_t - 1)M_{t-1})/p_t.$$

The individual specific variables includes c_t, x_t, and k_t, and $h_t \in (0,1)$ is the normalized working hours. Moreover, in period t the nominal money balance of each household, m_{t-1}, is augmented with a lump-sum transfer $(g_t - 1)M_{t-1}$ from government, where $M_t = g_t M_{t-1}$ is the per capita money supply in period t, g_t is the gross growth rate of money, and M_{t-1} be the per capita money supply in period $t-1$.

After g_t is realized in the beginning of each period, each agent uses the money balances $m_{t-1}+(g_t-1)M_{t-1}$ to purchase nonstorable consumption goods. The purchase of cash goods should satisfy the cash-in-advance constraint

$$p_t c_t \leq m_{t-1} + (g_t - 1)M_{t-1}.$$

It has been shown by Cooley and Hansen (1989) that the cash-in-advance constraint will be binding if and only if $E_t(1/g_{t+1}) < 1/\overline{\beta}$. Thus, a sufficient condition for this constraint to be binding is that the gross growth rate of money g_t always exceeds the discount factor $\overline{\beta}$. Here, we only concentrate on the case in which this constraint always holds with equality.

Therefore, the representative agent's problem can be summarized as below:

$$\max E_0 \sum_{t=0}^{\infty} \overline{\beta}^t u(c_t, h_t), \tag{3.1}$$

subject to

$$p_t c_t \leq m_{t-1} + (g_t - 1)M_{t-1}, \tag{3.2}$$

$$c_t + x_t + m_t/p_t \leq w_t h_t + r_t k_t + (m_{t-1} + (g_t - 1)M_{t-1})/p_t, \tag{3.3}$$

$$k_{t+1} = (1-\delta)k_t + x_t. \tag{3.4}$$

Two stochastic shocks, z_t and g_t, are assumed in this cash-in-advance model, and both of them are revealed to all agents at the beginning of period t. One of the exogenous random shocks considered in this model is the technology shock z_t. The technology shock z_t obeys a law of motion

$$z_{t+1} = \overline{\gamma} z_t + \epsilon_{t+1}, \ \epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2), \text{ where } 0 < \overline{\gamma} < 1. \tag{3.5}$$

In other words, the technology shock is a stationary AR process. Since $E(z_t) = 0$, technology progress is not considered; otherwise, $E(z_t)$ would be positive.

Another random disturbance is the monetary shock g_t:

$$\ln g_{t+1} = \bar{\alpha} \ln g_t + \xi_{t+1}, \ \xi_t \sim i.i.d.(\ln \bar{g}(1-\bar{\alpha}), \sigma_\xi^2), \text{ where } 0 < \overline{\alpha} < 1. \tag{3.6}$$

The monetary shock is assumed to follow the law of motion given by (3.6), which is also a stationary AR process, where $\ln \bar{g}$ is the unconditional mean of $\ln g_t$.

The firms' production technology is assumed to be a constant return-to-scale function

$$Y_t = \exp(z_t) K_t^\rho H_t^{1-\rho},$$

where $0 \leq \rho \leq 1$. Here, the fixed cost of each individual firm is ignored; therefore, the assumptions of the competitive market and the constant return-to-scale technology imply that firms make zero profit in equilibrium. We can assume, without loss of generality, that there is only one competitive firm in the market and it makes zero profit in equilibrium. By the first-order condition of maximizing the firm's profit,

$Y_t - w_t H_t - r_t K_t$, it is easy to show that the wage rate and capital rental rate in the budget constraint (3.3) are

$$w(z_t, K_t, H_t) = (1-\rho)\exp(z_t)(\frac{K_t}{H_t})^{\rho}, \tag{3.7}$$

and

$$r(z_t, K_t, H_t) = \rho \exp(z_t)(\frac{K_t}{H_t})^{\rho-1}, \tag{3.8}$$

and they are equal to the marginal productivity of labor and capital, respectively.

3.2 The equilibrium

Given the business cycle model introduced in the previous section, in a manner similar to Cooley and Hansen's (1989) calibration steps we first define the equilibrium and then solve for the decision rules by iterating the quadratic version of Bellman's equation. The procedure is outlined in Section 3.3. I will also show that the monetary business cycle model implies a cointegrated error correction model by linearizing the constraints.

In order to obtain the stationary solution, a change of variable is introduced since m_t, M_t, and p_t are nonstationary. Let $\hat{m}_t = m_t/M_t$ and $\hat{p}_t = p_t/M_t$. Also let the maximized expected present value of the agent's lifetime utility at equilibrium be $V(z_t, g_t, \hat{m}_{t-1}, K_t, k_t)$, where z, g, and K are the aggregate state variables and $\hat{m}$ and k

are individual state variables. Variables with primes represent the next period values.

Then the representative agent's dynamic problem can be rewritten as

$$\begin{aligned} V(z, g, \hat{m}, K, k) =& \max_{c,h} \{U(c, h) \\ & +\bar{\beta}\mathrm{E}[V(z', g', \hat{m}', K', k')|z, g, \hat{m}, K, k]\} \end{aligned} \tag{3.9}$$

$$\text{s.t. } z' = \bar{\gamma} z + \epsilon \tag{3.10}$$

$$\ln g' = \bar{\alpha} \ln g + \xi \tag{3.11}$$

$$c = (\hat{m} + g - 1)/(\hat{p}g) \tag{3.12}$$

$$\begin{aligned} c + x + \hat{m}'/\hat{p} = & w(z, K, H)h + r(z, K, H)k \\ & + (\hat{m} + g - 1)/(\hat{p}g) \end{aligned} \tag{3.13}$$

$$k' = (1 - \delta)k + x \tag{3.14}$$

$$K' = (1 - \delta)K + X. \tag{3.15}$$

Denote the state vector by $s = (z, g, \widehat{m}, K, k)^T$ and $S = (z, g, K)^T$. The stationary competitive equilibrium of this model contains a set of decision rules for the household $c(s)$, $x(s)$, $\widehat{m}'(s)$, and $h(s)$; a set of aggregate decision rules, $X(S)$ and $H(S)$; pricing functions $\widehat{p}(S), w(S)$,and $r(S)$, and a value function, $V(s)$, such that:

(i) The household's problems are solved. In other words, given the pricing functions and the aggregate decision rules, $v(s)$ solves the functional equation in (3.9), and $c(s), x(s)$, $\widehat{m}'(s)$, and $h(s)$ are the associated decision rules.

(ii) The firm's profit is maximized. In a competitive market, the constant return-to-scale technology implies zero profit; therefore, the wage and rental rate of capital are decided by their marginal productivity, (3.7) and (3.8), respectively.

(iii) By symmetry, each individual's decision is consistent with aggregate outcomes: $x = X$, $h = H$, and $\widehat{m}' = 1$, when $k = K$ and $\widehat{m} = 1$.

(iv) The equilibrium of commodity goods should satisfy $c(s) + x(s) = Y(S)$ for all s.

3.3 Solution method

The system equation (3.9)-(3.15) can be solved by a dynamic programming technique[2]. Since the utility function is concave and the constraint set is convex, the value function is also concave. Therefore, there exists a unique continuous function $V : s \to \mathbf{R}$ that satisfies the Bellman equation. However, the maximization problem cannot be solved analytically but only numerically, because there is no explicit functional form for the

[2]See Stocky and Lucas (1989) for further detail.

value function V.

In order to compute the equilibria of business cycle models, the class of models that fluctuate around a steady state and display local dynamics is approximated by a set of linear functions in most macroeconomic research. In other words, a quadratic approximation of the objective function is formed by a Taylor's expansion of the function at the deterministic steady-state values. It gives an approximated version of the maximization problem, which then can be solved by a numerical method. The approximation method used here and section 6.2 is based on the work of Kydland and Prescott (1982), Hansen and Prescott (1995), and McGrattan (1990).

The maximization in (3.9) is over c and h, so we first solve for the per capita working hours and then substitute it into the utility function. By so doing, we can eliminate the nonlinearity in the constraint set. The next step is to use the second-order Taylor's expansion to approximate the utility function and solve the quadratic version of Bellman's equation.

We first combine (3.12) and (3.13), so the budget constraint can be rewritten as

$$x + \hat{m}'/\hat{p} = w(z, K, H)h + r(z, K, H)k. \tag{3.16}$$

Since both labor and investment are credit goods, the cash-in-advance constraint only applies to nonstorable consumption goods. It implies that agents can only reduce their money holdings by reducing their consumption. Next, by aggregating (3.16) and using (3.7) and (3.8), we can solve for H :

$$H = \left[\frac{X + (1/\hat{p})}{\exp(z)K^{\rho}}\right]^{\frac{1}{1-\rho}}. \tag{3.17}$$

Using (3.16), (3.17), and the utility function $u(c, h)$, we obtain

$$u(c,h) = \ln\left(\frac{\hat{m} + g - 1}{\hat{p}g}\right) - B\left[\frac{\left(x + \hat{m}/\hat{p} - \rho(X + 1/\hat{p})\frac{k}{K}\right)(X + 1/\hat{p})^{\frac{\rho}{1-\rho}}}{(1-\rho)(\exp(z)K^{\rho})^{\frac{1}{1-\rho}}}\right]. \tag{3.18}$$

In order to linearize the nonlinear functions, we work with the logarithm of each variable for the sake of convenience. We let $\tilde{g} = \ln g$, $\tilde{p} = \ln \hat{p}$, $\tilde{x} = \ln x$, $\tilde{X} = \ln X$, $\tilde{k} = \ln k$, $\tilde{K} = \ln K$, then we replace p, x, X, k, and K with $\exp(\widetilde{p})$, $\exp(\widetilde{x})$, $\exp(\widetilde{X})$, $\exp(\widetilde{k})$, and $\exp(\widetilde{K})$, respectively. Furthermore, let the **state vector** be $\tilde{s} = (1, z, \tilde{g}, \hat{m}, \tilde{K}, \tilde{k})^T$ and $\tilde{S} = (1, z, \tilde{g}, \tilde{K})^T$, the **individual's decision vector** be $\tilde{u} = (\hat{m}', \tilde{x})^T$, the **economy-wide variables** be $\tilde{U} = (\tilde{p}, \tilde{X})^T$, $W = (z, \tilde{g}, \hat{m}, \tilde{K}, \tilde{k}, \hat{m}', \tilde{x}, \tilde{p}, \tilde{X})^T$, and let the variables with superscript "*" represent the values at steady state. Then the utility function (3.18) can be approximated by a second-order Taylor's expansion around the steady state as follows:

$$\begin{aligned} u(c,h) = {} & \ln\left(\frac{\hat{m} + \exp(\widetilde{g}) - 1}{\exp(\widetilde{p})\exp(\widetilde{g})}\right) \\ & -B\left[\frac{\left[\exp(\widetilde{x}) + \frac{\hat{m}'}{\exp(\widetilde{p})} - \rho\left(\exp(\widetilde{X}) + \frac{1}{\exp(\widetilde{p})}\right)\frac{\exp(\widetilde{k})}{\exp(\widetilde{K})}\right]\left(\exp(\widetilde{X}) + \frac{1}{\exp(\widetilde{p})}\right)^{\frac{\rho}{1-\rho}}}{(1-\rho)\left(\exp(z)\exp(\widetilde{K}^{\rho})\right)^{\frac{1}{1-\rho}}}\right], \end{aligned}$$

$$\simeq \quad u(W^*) + D^T u(W^*)(W - W^*) + \frac{1}{2}(W - W^*)^T D^2 u(W^*)(W - W^*),$$

$$= \quad (1, W - W^*)^T \begin{pmatrix} u(W^*) & \frac{1}{2}D^T u(W^*) \\ \frac{1}{2}Du(W^*) & \frac{1}{2}D^2 u(W^*) \end{pmatrix} \begin{pmatrix} 1 \\ W - W^* \end{pmatrix},$$

$$= \quad (1, W^{*T})Q \begin{pmatrix} 1 \\ W^* \end{pmatrix},$$

where Q is a 10×10 symmetric matrix with elements[3]:

$$Q_{1,i+1} = Q_{i+1,1} = \tfrac{1}{2}[D_i u(W^*) - \textstyle\sum_{j=1}^{9} D^2_{ij} u(W^*) W^*_j],$$

$$Q_{i+1,j+1} = Q_{j+1,i+1} = \tfrac{1}{2} D^2_{ij} u(W^*),$$

$$Q_{11} = u(W^*) - \textstyle\sum_{j=1}^{9} D_j u(W^*) + \tfrac{1}{2} \sum_{i=1}^{9} \sum_{j=1}^{9} D^2_{ij} u(W^*) W^*_i W^*_j,$$

for $i, j = 1, \cdots, 9$. Therefore, the approximation of the representative agent's problem can be summarized as

$$\tilde{s}^T V \tilde{s} \quad = \quad \max[\tilde{s}^t \quad \tilde{u}^T \quad \tilde{U}^T] Q \begin{bmatrix} \tilde{s} \\ \tilde{u} \\ \tilde{U} \end{bmatrix} + \beta \tilde{s}'^T V \tilde{s}', \tag{3.19}$$

$$s.t. \ \tilde{g}' = \bar{\alpha}\tilde{g} + \epsilon, \tag{3.20}$$

$$z' = \bar{\gamma} z + \epsilon, \tag{3.21}$$

[3]See the details in Hansen and Prescott (1995).

$$e^{\tilde{K}'} = (1-\delta)e^{\tilde{K}} + e^{\tilde{X}}, \tag{3.22}$$

$$\tilde{U} = \mathbf{U}(S), \tag{3.23}$$

where $\mathbf{U}$ is a linear function which describes the relationship between $\tilde{U}$ and $\tilde{S}$ realized by the households in the beginning of each period.

The equilibrium process for the approximate economy maximizes the representative's lifetime utility (3.19) subject to the constraints (3.20) through (3.23). To obtain the decision rules, we start with a guess of the matrix V, called V_0, and also choose a candidate for the function $\mathbf{U}$. Then we iterate on the quadratic version of Bellman's equation until V_n and V_{n+1} are sufficiently close.

Substituting (3.20) through (3.23) into (3.19), we can obtain the linear decision rule by using the first-order condition for $\tilde{u}$,

$$\tilde{u} = D_1\tilde{s} + D_2\tilde{U},$$

or

$$\begin{pmatrix} \hat{m}' \\ \ln x \end{pmatrix} = D_1 \begin{pmatrix} 1 \\ z \\ \ln g \\ \hat{m} \\ \ln K \\ \ln k \end{pmatrix} + D_2 \begin{pmatrix} \ln \widehat{p} \\ \ln X \end{pmatrix},$$

where D_1 and D_2 are 2×6 and 2×2 matrices, respectively.

Imposing the equilibrium conditions, $\tilde{x} = \tilde{X}$, $\hat{m}' = \hat{m} = 1$, and $\tilde{k} = \tilde{K}$, we further have

$$\tilde{U} = D_3 \tilde{S},$$

where D_3 is a 2×4 matrix. Using the results obtained from the previous steps, we can further compute the value function for the next iteration. The procedure is repeated until V_{n+1} is sufficiently close to V_n. When the iteration converges, we obtain the decision rules for equilibrium price and investment, which are the same as those in Cooley and Hansen (1989):

$$\ln \hat{p} = d_{11} + d_{12} z + d_{13} \ln g + d_{14} \ln K, \tag{3.24}$$

$$\ln X = d_{21} + d_{22} z + d_{23} \ln g + d_{24} \ln K. \tag{3.25}$$

Equations (3.24) and (3.25) show that, at equilibrium, the price and investment are completely decided by the two stochastic shocks and capital stock.

3.4 Linearization

This section uses the above decision rules and Taylor's expansion to linearize the non-linear constraint around the steady state to obtain the linearized version of the Cooley and Hansen model.

Since in this model cash goods contains only consumption goods, the binding cash-in-advance constraint implies that all money is used on consumption goods. The marginal utility of consumption $\lim_{c \to 0} \partial u(c,h)/\partial c = \infty$, guaranteeing that money holding must be positive. Combined with the equilibrium condition $\widehat{m} = 1$, the binding cash-in-advance constraint (3.12) implies that price level and per capita consumption have an inverse relationship,

$$C = \frac{1}{\hat{p}},$$

or

$$\ln C = -\ln \hat{p}.$$

Replacing $\hat{p}$ by equation (3.24), we then have

$$\ln C = -d_{11} - d_{12} z - d_{13} \ln g - d_{14} \ln K.$$

Here d_{11}, d_{12}, $d_{13,}$ and d_{14} are the coefficients.

Next, taking the logarithms with respect to the per capita working hours, (3.17) yields

$$\begin{aligned}\ln H &= \tfrac{1}{1-\rho}\ln\left(X+\tfrac{1}{\hat{p}}\right)+\tfrac{z}{1-\rho}+\tfrac{\rho}{1-\rho}\ln K,\\ &= \tfrac{1}{1-\rho}\ln\left(\exp(\widetilde{X})+\tfrac{1}{\exp(\widetilde{p})}\right)+\tfrac{z}{1-\rho}+\tfrac{\rho}{1-\rho}\ln K,\end{aligned}$$

which can be approximated by the first-order Taylor's expansion in the neighborhood of the steady state. Therefore,

$$\begin{aligned}\ln H \simeq \frac{1}{1-\rho}&\left[\ln\left(\exp(\widetilde{X}^*)+\frac{1}{\exp(\widetilde{p}^*)}\right)+\frac{(\tilde{X}-\tilde{X}^*)\exp(\widetilde{X}^*)}{\exp(\widetilde{X}^*)+\dfrac{1}{\exp(\widetilde{p}^*)}}\right.\\ &\left.+\frac{(\tilde{p}-\tilde{p}^*)\exp(\widetilde{p}^*)}{\exp(\widetilde{X}^*)+\dfrac{1}{\exp(\widetilde{p}^*)}}\right]+\frac{z}{1-\rho}+\frac{\rho}{1-\rho}\ln K,\\ &= h_1+h_2\ln X+h_3\ln\hat{p}+\frac{z}{1-\rho}+\frac{\rho}{1-\rho}\ln K,\end{aligned}$$

where

$$h_1=\frac{1}{1-\rho}\left[\ln\left(X^*+\frac{1}{p^*}\right)-\frac{X^*\ln X^*+p^*\ln p^*}{X^*+\frac{1}{p^*}}\right],$$

$$h_2=\frac{1}{1-\rho}\frac{X^*}{X^*+\frac{1}{p^*}},$$

and

$$h_3=-\frac{1}{1-\rho}\frac{p^*}{X^*+\frac{1}{p^*}}.$$

Here h_1, h_2, and h_3 are all constants dependent on the deterministic steady-state values $\widetilde{X}^*$ and $\tilde{p}^*$, as well as the coefficient ρ.

Replacing $\ln \hat{p}$ and $\ln X$ by (3.24) and (3.25) yields

$$\ln H = \alpha_1 + \alpha_2 z + \alpha_3 \ln g + \alpha_4 \ln K,$$

where

$$\alpha_1 = h_1 + h_2 d_{21} + h_3 d_{11},$$

$$\alpha_2 = h_2 d_{22} + h_3 d_{12} + 1/(1-\rho),$$

$$\alpha_3 = h_2 d_{23} + h_3 d_{13}, \text{ and}$$

$$\alpha_4 = h_2 d_{24} + h_3 d_{14} + \rho/(1-\rho).$$

Similarly, taking logarithms with respect to the constant return-to-scale production function $Y_t = \exp(z_t) K_t^{\rho} H_t^{1-\rho}$,yields

$$\begin{aligned} \ln Y &= z + \rho \ln K + (1-\rho) \ln H, \\ &\simeq \beta_1 + \beta_2 z + \beta_3 \ln g + \beta_4 \ln K, \end{aligned}$$

where

$$\beta_1 = (1-\rho)\alpha_1,$$

$$\beta_2 = 1 + (1-\rho)\alpha_2,$$

$$\beta_3 = (1-\rho)\alpha_3, \text{ and}$$

$$\beta_4 = \rho + (1-\rho)\alpha_4.$$

Since capital stock is a state variable which is given in the beginning of each period, the variation of total output is affected by the two stochastic shocks only through the per capita working hours H.

Furthermore, using the law of motion of capital accumulation (3.22), yields

$$
\begin{aligned}
0 &= \exp(\widetilde{K}') - (1-\delta)\exp(\widetilde{K}) - \exp(\widetilde{X}) \\
&\simeq \exp(\widetilde{K}'^{*}) + \exp(\widetilde{K}'^{*})(\tilde{K}' - \tilde{K}'^{*}) - (1-\delta)[\exp(\widetilde{K}^{*}) + \exp(\widetilde{K}^{*})(\tilde{K} - \widetilde{K}^{*})] \\
&\quad -[\exp(\widetilde{X}^{*}) + \exp(\widetilde{X}^{*})(\tilde{X} - \tilde{X}^{*})], \\
&= \exp(\widetilde{K}'^{*})\Bigg\{\tilde{K}' - (1-\delta)\frac{\exp(\widetilde{K}^{*})}{\exp(\widetilde{K}'^{*})}\tilde{K} - \frac{\exp(\widetilde{X}^{*})}{\exp(\widetilde{K}'^{*})}\tilde{X} \\
&\quad + \left[(1-\tilde{K}'^{*}) - (1-\delta)\frac{\exp(\widetilde{K}^{*})}{\exp(\widetilde{K}'^{*})}(1-\tilde{K}^{*}) + \frac{\exp(\widetilde{X}^{*})}{\exp(\widetilde{K}'^{*})}(1-\tilde{X}^{*})\right]\Bigg\}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\tilde{K}' - (1-\delta)\frac{\exp(\tilde{K}^{*})}{\exp(\tilde{K}'^{*})}\tilde{K} - \frac{\exp(\tilde{X}^{*})}{\exp(\tilde{K}'^{*})}\tilde{X} \\
&+ \left[(1-\tilde{K}'^{*}) - (1-\delta)\frac{\exp(\tilde{K}^{*})}{\exp(\tilde{K}'^{*})}(1-\tilde{K}^{*}) + \frac{\exp(\tilde{X}^{*})}{\exp(\tilde{K}'^{*})}(1-\tilde{X}^{*})\right] \simeq 0.
\end{aligned}
$$

Under the constant return-to-scale technology, the marginal product of capital is positive, but it declines as capital rises. When the capital stock reaches the steady state, we have $\tilde{K}'^{*} = \tilde{K}^{*}$ such that

$$
\begin{aligned}
\tilde{K}' &\simeq (1-\delta)\tilde{K} + \frac{\exp(\widetilde{X}^{*})}{\exp(\widetilde{K}'^{*})}\tilde{X} \\
&\quad - \left[(1-\tilde{K}'^{*}) - (1-\delta)(1-\tilde{K}^{*}) + \frac{\exp(\widetilde{X}^{*})}{\exp(\widetilde{K}'^{*})}(1-\tilde{X}^{*})\right], \\
&\simeq k_1 + k_2 \ln K + k_3 \ln X, \\
&\simeq \gamma_1 + \gamma_2 z + \gamma_3 \ln g + \gamma_4 \ln K,
\end{aligned}
$$

where

$$\gamma_1 = k_1 + k_3 d_{21},$$

$$\gamma_2 = k_3 d_{22},$$

$$\gamma_3 = k_3 d_{23},$$

$$\gamma_4 = k_2 + k_3 d_{24},$$

$$k_1 = \delta\left[2 - (1+\delta)\ln K^*\right],$$

$$k_2 = (1-\delta), \text{ and}$$

$$k_3 = \delta.$$

k_1, k_2, and k_3 are constants and depend only on the steady state values $\tilde{K}^*$, $\tilde{X}^*$, and the depreciation rate δ.

Moreover, $M_t = g_t M_{t-1}$ can also be written as

$$\begin{aligned} \ln M_t &= \ln g_t + \ln M_{t-1}, \\ &= \sum_{j=0}^{t-1} \ln g_{t-j} + M_0, \end{aligned} \tag{3.26}$$

where M_0 is the initial money stock. We may, without loss of generality, set $M_0 = 1$. Since $\ln \hat{p}_t = \ln p_t - \ln M_t$, replacing $\ln M_t$ in (3.24) by (3.26) yields

$$\ln p_t = d_{11} + 1 + d_{12} z_t + d_{13} \ln g_t + \sum_{j=0}^{t-1} \ln g_{t-j} + d_{14} \ln K_t.$$

The above results can be summarized as the following linearized version of the business cycle model:

$$
\begin{aligned}
\ln p_t &= d_{11} + 1 + d_{12} z_t + (d_{13}+1)\ln g_t + \sum_{j=1}^{t-1} \ln g_{t-j} + d_{14}\ln K_t, \\
\ln X_t &= d_{21} + d_{22} z_t + d_{23}\ln g_t + d_{24}\ln K_t, \\
\ln C_t &= -d_{11} - d_{12} z_t - d_{13}\ln g_t - d_{14}\ln K_t, \\
\ln H_t &= \alpha_1 + \alpha_2 z_t + \alpha_3 \ln g_t + \alpha_4 \ln K_t, \\
\ln Y_t &= \beta_1 + \beta_2 z_t + \beta_3 \ln g_t + \beta_4 \ln K_t, \\
\ln K_{t+1} &= \gamma_1 + \gamma_2 z_t + \gamma_3 \ln g_t + \gamma_4 \ln K_t,
\end{aligned}
\tag{3.27}
$$

with two stochastic shocks

$$
\begin{aligned}
z_{t+1} &= \bar{\gamma} z_t + \epsilon_{t+1}, \\
\ln g_{t+1} &= \bar{\alpha} \ln g_t + \xi_{t+1}.
\end{aligned}
$$

Since (3.27) is a linear approximation of the deterministic steady state when $\epsilon_t = 0$ and $\xi_t = 0$, it follows that $z_t = \overline{\gamma}^t z_0$ and $\ln g_t = \overline{\alpha}^t \ln g_0$. Therefore, $K_{t+1} = K_t = K^*$ implies that the coefficients γ_1 and γ_4 equal 0 and 1, respectively, which means that $\ln K$ is an $I(1)$ process. Moreover, $X^* = \delta K^*$ implies that $d_{21} = \ln \delta$ and $d_{24} = 1$. Since $\gamma_1 = k_1 + k_3 d_{21} = 0$, together with $d_{21} = \ln \delta$, it follows that

$$
\ln K^* = \frac{2 + \ln \delta}{1 + \delta}.
$$

Furthermore, both $\bar{\gamma}$ and $\bar{\alpha}$ are less than one in absolute value, so both z and $\ln g$ are stationary. The system of equations (3.27) allow all the variables to be written as functions of z_t, $\ln g_t$, and $\ln K_t$. Since $\ln K_t$ is an $I(1)$ process, it is clear that $\ln P_t$, $\ln X_t$, $\ln C_t$, $\ln H_t$, and $\ln Y_t$ are also $I(1)$ processes according to (3.27).

3.5 The Error Correction Model

In this section I will show that the linearized Cooley and Hansen model implies a cointegrated error correction model. For the sake of convenience, we rewrite the system of equations in (3.27) as

$$W_t = \delta_0 + D_1 u_t + \delta_3 \sum_{j=1}^{t-1} \ln g_{t-j} + \delta_4 \ln K_t, \tag{3.28}$$

$$\Delta \ln K_t = \eta^T u_{t-1}, \tag{3.29}$$

$$u_t = \Gamma u_{t-1} + e_t, \tag{3.30}$$

where

$$W_t = \begin{pmatrix} \ln p_t \\ \ln X_t \\ \ln C_t \\ \ln H_t \\ \ln Y_t \end{pmatrix},$$

$$\delta_0 = \begin{pmatrix} d_{11} + 1 \\ \ln \delta \\ -d_{11} \\ h_1 + h_2 \ln \delta + h_3 d_{11} \\ (1-\rho)(h_1 + h_2 \ln \delta + h_3 d_{11}) \end{pmatrix},$$

$$\delta_1 = \begin{pmatrix} d_{12} \\ d_{22} \\ -d_{12} \\ h_2 d_{22} + h_3 d_{12} + 1/(1-\rho) \\ 2 + (1-\rho)(h_2 d_{22} + h_3 d_{12}) \end{pmatrix},$$

$$\delta_2 = \begin{pmatrix} d_{13} + 1 \\ d_{23} \\ -d_{13} \\ h_2 d_{23} + h_3 d_{13} \\ (1-\rho)(h_2 d_{23} + h_3 d_{13}) \end{pmatrix},$$

$$\delta_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\delta_4 = \begin{pmatrix} d_{14} \\ 1 \\ -d_{14} \\ h_2 + h_3 d_{14} + \rho/(1-\rho) \\ 2\rho + (1-\rho)(h_2 + h_3 d_{14}) \end{pmatrix}$$

$$D_1 = (\delta_1, \delta_2),$$

$$\eta = \begin{pmatrix} \gamma_2 \\ \gamma_3 \end{pmatrix},$$

$$u_t = \begin{pmatrix} z_t \\ \ln g_t \end{pmatrix},$$

$$\Gamma = \begin{pmatrix} \bar{\gamma} & 0 \\ 0 & \bar{\alpha} \end{pmatrix},$$

and

$$e_t = \begin{pmatrix} \epsilon_t \\ \\ \xi_t \end{pmatrix} \sim i.i.d.\ N\left(\begin{pmatrix} 0 \\ \\ (1-\bar{\alpha})\ln\bar{g} \end{pmatrix}, \begin{pmatrix} \sigma_\epsilon^2 & 0 \\ \\ 0 & \sigma_\xi^2 \end{pmatrix}\right).$$

Since the two stochastic shocks are independent, their joint distribution has zero covariance.

Let $D_2 = (\delta_{1,}\delta_2, \delta_3, \delta_4)$, then (3.28) can be rewritten as

$$D_2^T W_{t-1} = D_2^T \delta_0 + D_2^T D_2 \begin{pmatrix} u_{t-1} \\ \\ \sum_{j=1}^{t-1} \ln g_{t-j} \\ \\ \ln K_{t-1} \end{pmatrix}.$$

Furthermore, let

$$I_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

so that

$$u_{t-1} = I_0 (D_2^T D_2)^{-1} D_2^T (W_{t-1} - \delta_0). \tag{3.31}$$

Taking the first difference with respect to equation (3.28) and then substituting (3.29) and (3.30) into it yields:

$$\Delta W_t = D_1 \Delta u_t + \delta_3 \ln g_{t-1} + \delta_4 \Delta \ln K_t.$$

Let $\delta_5 = \delta_3(0,1)$; then

$$\Delta W_t \quad = \quad D_1(u_t - u_{t-1}) + \delta_5 u_{t-1} + \delta_4 \eta^T u_{t-1},$$

$$= D_1(\Gamma u_{t-1} + e_t) - D_1 u_{t-1} + \delta_5 u_{t-1} + \delta_4 \eta^T u_{t-1},$$

$$= [D_1(\Gamma - I_2) + \delta_4 \eta^T + \delta_5] u_{t-1} + D_1 e_t.$$

Replacing u_{t-1} by (3.31), then (3.28) through (3.30) can be rewritten in the form of an error correction model:

$$\begin{aligned}
\Delta W_t &= [D_1(\Gamma - I_2) + \delta_4 \eta^T + \delta_5] I_0 (D_2^T D_2)^{-1} D_2^T (W_{t-1} - \delta_0) + D_1 e_t, \\
&= \nu_0 + \alpha \beta^T (W_{t-1} - \delta_0) + \varepsilon_t, \qquad (3.32)
\end{aligned}$$

where

$$\begin{aligned}
\nu_0 &= D_1 \bar{e}, \\
\alpha &= D_1(\Gamma - I_2) + \delta_4 \eta^T + \delta_5, \\
\beta &= D_2 (D_2^T D_2)^{-1} I_0^T
\end{aligned}$$

and

$$\bar{e} = \begin{pmatrix} 0 \\ (1 - \overline{\alpha}) \ln \bar{g} \end{pmatrix}.$$

Here, ν_0 is a 5×1 matrix, and both α and β are 5×2 matrices.

Moreover, the error term has distribution

$$\varepsilon_t = D_1(e_t - \bar{e}) \sim i.i.d.\ N(0, \Sigma(\theta)),$$

and

$$\Sigma(\theta) = D_1 \Xi D_1^T = \sigma_\epsilon^2 \delta_1 \delta_1^T + \sigma_\xi^2 \delta_2 \delta_2^T,$$

where

$$\Xi = \begin{pmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_\xi^2 \end{pmatrix}.$$

and θ is the set of parameters. The system of equations (3.27) show that the model is driven by only two stochastic shocks, so the variance of the error term ε_t in (3.32) has only rank two. In other words, the randomness of W_t is solely due to the two random variables because there are only two stochastic shocks assumed in this CIA model. Therefore, the conditional distribution of the theoretical model is a singular multivariate normal distribution. The main consequence of the singularity is that the density function of the theoretical model does not exist. Bierens and Swanson (2000) suggest working on the transformed theoretical and empirical distributions and then concentrate on only the marginal distributions. Bierens (2003) proposes to add some noise to both the theoretical and empirical models using convolution. Further details are stated in sections 4.2 and 4.4.

On the other hand, both α and β are 5×2 matrices, and thus $\alpha\beta^T$ is a 5×5 matrix with rank two. Accordingly, two cointegrating vectors are implied by the theoretical model. In the error correction model, $\beta^T W_{t-1}$ equal to a constant represents an economic equilibrium relation with α the adjustment coefficient, and β the cointegration vector. In other words, the representative agent reacts to the disequilibrium error or the deviation of $\beta^t W_{t-1}$ from the constant through the adjustment coefficient α in order to bring the variables back on the right track such that the economic relationship is satisfied.

In summary, in this section we have shown that the business cycle model with a cash-in-advance constraint implies a cointegrated error correction model. The theoretical model turns out to be an error correction model with two cointegration vectors and singular variance with rank two, and the model conditional distribution of ΔW_t is

$$\Delta W_t | W_{t-1} \sim N_5(\mu_t(\theta), \Sigma(\theta)),$$

where

$$\mu_t(\theta) = \nu_0 + \alpha\beta^T(W_{t-1} - \delta_0).$$

Chapter 4

The average and multiplicative conditional reality bound

This chapter introduces the two econometric approaches that will be used to analyze the cash-in-advance model. The first approach is proposed by Bierens and Swanson (2000). They suggest an information measure, called average conditional reality bound, to evaluate a theoretical model. The second approach is proposed by Bierens (2003). He suggests using convolutions of the theoretical and DGP distributions with a non-singular distribution. The deep parameters in the theoretical model can then be estimated, instead of calibrated.

Section 3.1 discusses the ceteris paribus assumption and its consequences on model specification. Section 3.2 and 3.3 introduces the Bierens-Swanson average conditional reality bound and its implementation. Section 3.4 and 3.5 introduces the Bierens multiplicative conditional reality bound and its implementation, respectively.

4.1 The ceteris paribus assumption

Economic theorists and econometricians have different ways to link the economic theories and econometric models. One major difference is the way they deal with a ceteris paribus condition. Theorists simplifies the environment of the real world to construct a theoretical model by imposing ceteris paribus assumptions. In other words, they set the variables outside the scope of theoretical models to be constant. However, a ceteris paribus condition might have two major undesirable consequences on the theoretical model: model misspecification and missing variables.

The most common interpretation of a ceteris paribus assumption is that the state of world is fixed as a constant. For example, if Y is a vector of endogenous variables, X is a vector of exogenous variables in the theoretical model, and W is a vector of variables outside the scope of the theoretical model, then the conditional distribution of Y given $X = x$ and $W = w$ is $f(y|x, w)$. Imposing a ceteris paribus condition is equivalent to setting W as a constant, say $W = 0$. Therefore, the conditional mean of the theoretical model is

$$E[Y|x, W=0] = \int y f(y|x, W=0) dy.$$

On the other hand, the econometricians take the uncertainty of the variables outside the scope of the theoretical model into account by integrating it out. Thus the conditional distribution of y, given x is

$$f(y|x) = \int f(y|x, w) dG(w),$$

and the true conditional mean of Y given $X = x$ is

$$E[Y|x] = \int \int y f(y|x, w) dG(w) dy,$$

where $G(\cdot)$ is the distribution function of W. Generally, $E[Y|x, W = 0] \neq E[Y|x]$, which not only means that the theoretical model does not represent the data generating process, but also means that a ceteris paribus assumption causes model misspecification.

Furthermore, a ceteris paribus assumption can also be interpreted as omitted variables. For example, according to the DGP y_t is generated by

$$y_t = \alpha_0 + \alpha_1' x_t + \alpha_2' w_t + u_t,$$

with the conditional distribution

$$y_t | x_t, w_t \sim i.i.d. N(\alpha_0 + \alpha_1' x_t + \alpha_2' w_t, \sigma^2),$$

and the conditional density $f_{DGP}(y|x,w)$, where $u_t \sim i.i.d.N(0,\sigma^2)$. Under the ceteris paribus assumption, the variables outside the scope of the theoretical model represented by w_t are assumed to be a constant vector, say $w_t = w = 0$. The theoretical model then becomes

$$y_t = \alpha_0 + \alpha_1' x_t + u_t^*,$$

with the conditional distribution

$$y_t|x_t \sim i.i.d.N(\alpha_0 + \alpha_1' x_t, \sigma_*^2),$$

and the conditional density $f_{TM}(y|x)$, where $u_t^* \sim i.i.d.N(0,\sigma_*^2)$. Generally, the conditional mean of the DGP is not equal to that of the theoretical model

$$E_{DGP}(y|x) = \int y f_{DGP}(y|x,w)dG(w)dy \neq E_{TM}(y|x) = \int y f_{TM}(y|x)dy.$$

With possible misspecification and missing variables in the setting of models, it is not surprising that theoretical models do not fit data well and are often rejected by classical specification tests. To measure the reality content of the theoretical model under a ceteris paribus assumption, we introduce the approaches suggested by Bierens and Swanson (2000) and Bierens (2003) in Section 3.2 and 3.4.

4.2 The Bierens-Swanson average conditional reality bound

Bierens and Swanson (2000) suggests to measure the goodness of fit of the theoretical model by comparing the joint densities of the theoretical model and DGP. Consider two densities, say $f_0(y)$ and $f(y)$, where $f_0(y)$ represents the 'model version' of $f(y)$. Then, if both $f_{TM}(y)$ and $f(y)$ have the same support, there exists a $p_0 \in (0,1)$ such that $p_0 f_0(y) \leq f(y)$, $\forall y$. In other words, the theoretical density is squeezed under the true density by multiplying it by a measure p_0, which can be treated as a measure of the reality content of the theoretical model. An illustrative graphical example is given in Figure 4.1.

Therefore, if the true density $f(y)$ and the theoretical density $f_0(y)$ have the same supports, then one can write the true density as

$$f(y) = p_0 f_0(y) + (1 - p_0) f_1(y),$$

where $f_1(y) = (f(y) - p_0 f_0(y)/(1 - p_0)$, and $f_1(y) = f(y)$ if $p_0 = 0$. Therefore, the maximal p_0 can be further written as

$$p_0 = \inf_y \frac{f(y)}{f_0(y)}.$$

Here p_0 can be interpreted either as an upper bound of the probability that

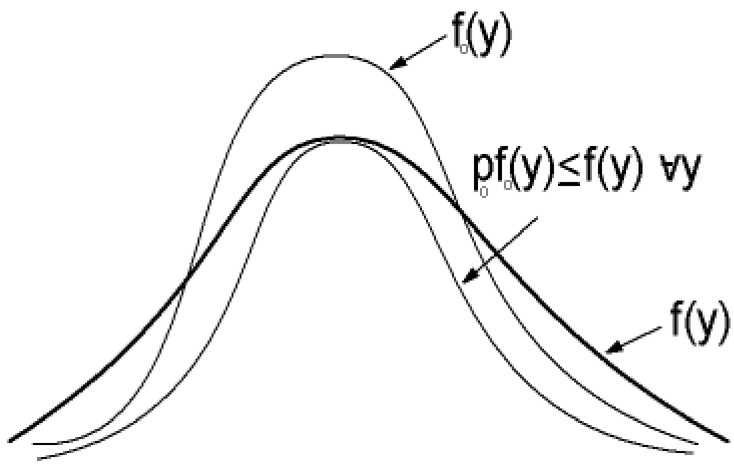

Figure 4.1: The theoretical and true densities.

$f_0(y)$ is correct, or as the maximal Bayesian prior that the model $f_0(y)$ is correctly specified.

In the case that $f(y) = \int f(y|x,w)dG(w) = f(y|x)$ and $f_0(y) = f(y|x, W = 0;\theta)$, both p_0 and θ depend on x, where θ is the parameters in the theoretical model. Then, p_0 can be rewritten as

$$p_0(x) = \sup_{\theta} \inf_{y} \frac{f(y|x)}{f_0(y|x,\theta)}, \tag{4.1}$$

and θ as

$$\theta(x) = \arg\max_{\theta} \inf_{y} \frac{f(y|x)}{f_0(y|x,\theta)}.$$

Therefore, in order to make p_0 and θ independent of x, Bierens and Swanson (2000) further define the 'average conditional reality bound'

$$p_0 = \sup_{\theta} \mathbf{E}[p_0(x)],$$

which can be estimated by

$$\widehat{p_0} = \sup \theta \frac{1}{n} \sum_{t=1}^{n} p_0(x_t).$$

4.3 Implementation of the average conditional reality bound

The average conditional reality bound not only compares the true density and the density of the theoretical model (hereafter called the theoretical density), but also evaluates the difference between them. However, we can never know the real DGP, so we replace the true density in (4.1) by the density implied by the econometric model (hereafter called the empirical density).

Supposing that the empirical model is a Gaussian vector autocorrelation regression (VAR) model with lag length q, $Y_t \in \mathbf{R}^k$ is the dependent variable, and the regressor $Z_t = (Y_{t-1}^T, Y_{t-2}^T, ..., Y_{t-q}^T)^T$ is a $kq \times 1$ vector, then the empirical model can be written as

$$Y_t = \Gamma Z_t + V_t,\ V_t \sim i.i.d.N_k(0, \Omega),\ \det(\Omega) \neq 0, \tag{4.2}$$

where Γ is a $k \times (kq)$ matrix of coefficients in the linear econometric model and $V_t \in \mathbf{R}^k$ is a stochastic shock process with nonsingular variance. It has a conditional distribution

$$Y_t | Z_t \sim N_k(\Gamma Z_t, \Omega).$$

The linearized version of the theoretical model takes the form

$$Y_t = A(\theta)X_t + B(\theta)U_t,\ U_t \sim i.i.d.N_m(0, I_m), \tag{4.3}$$

where $X_t \in \mathbf{R}^p$ is a vector of stacked lags of Y_t, $A(\theta)$ is a $k \times p$ matrix, $B(\theta)$ is a $k \times m$ matrix, and $U_t \in \mathbf{R}^m$ is a stochastic shock process with singular variance. It is usually the case that $m < k$, instead of $m = k$, because in dynamic structural economic models the assumption that the model is driven by one or only a few shocks leads to singularity of the model distribution. In other words, the number of stochastic shocks is less than the number of dependent variables in the model. The theoretical model has a conditional distribution

$$Y_t|A(\theta)X_t \sim N_m(A(\theta)X_t, B(\theta)B(\theta)^T), \text{ where } rank(B(\theta)B(\theta)^T) = m < k.$$

Since $\det(B(\theta)B(\theta)^T) = 0$, the variance of the model distribution in some particular dimensions is zero. In other words, the joint density of Y_t does not exist because of the ceteris paribus condition, which fixes some state of the world as constants. Therefore, one cannot compare the theoretical and empirical joint distribution directly.

One way to solve this problem is to compare the theoretical marginal distribution with the empirical marginal distribution. In other words, one needs to integrate out the empirical distribution in the singular dimensions where a ceteris paribus condition is imposed. Bierens and Swanson (2000) suggest working with the transformed distribution instead. By using the orthogonal matrix of eigenvectors of the singular conditional variance of the theoretical model, the theoretical and empirical distributions can be rotated so that the singular directions are on some of the principal axis. Then the variables in the distribution of the empirical model in the singular direction can easily be integrated out.

To begin with, one needs to find the orthogonal matrix of eigenvectors of $\Sigma(\theta)$, $\Pi(\theta) = (\Pi_1(\theta), \Pi_2(\theta))$, where $\Sigma(\theta) = B(\theta)B(\theta)^T$, $\Pi_1(\theta)$ is the $k \times (k-m)$ matrix of eigenvectors corresponding to the zero eigenvalues of $\Sigma(\theta)$, and $\Pi_2(\theta)$ is the $k \times m$ matrix of eigenvectors corresponding to the two positive eigenvalues $(\lambda_1(\theta), \lambda_2(\theta))$. In other words,

$$\Sigma(\theta) = (\Pi_1(\theta), \Pi_2(\theta)) \times \; diag(0, 0, 0, \lambda_1(\theta), \lambda_2(\theta)) \; \times \begin{pmatrix} \Pi_1^T(\theta) \\ \\ \Pi_2^T(\theta) \end{pmatrix}.$$

Multiplying both distributions by the orthogonal matrix $\Pi(\theta)$ gives the orthogonal transformations, which implies that the shapes of the distribution functions are invariant under rotation and location shifts[1].

Let $y^* = \Pi(\theta)y + \mu(\theta)$ represent the transformed vector of variables in the empirical model, and $y^{**} = \Pi(\theta)y + \mu(\theta)$ represent the transformed vector of variables in the theoretical model. Also, let $y^* = (y_1^*, y_2^*)^T \in \mathbf{R}^{k-m} \times \mathbf{R}^m$, where y_1^* represents the variables on the singular direction. Then the empirical marginal density is

$$f_y(y_2^*) \quad = \quad \int f(y_1^*, y_2^*) dy_1^*,$$

[1]See Bierens and Swanson (2000) for further details.

$$= \frac{\exp[-\frac{1}{2}(y_2^* - \Pi_2(\theta)^T\mu(\theta))^T(\Pi_2(\theta)^T\Omega\Pi_2(\theta))^{-1}(y_2^* - \Pi_2(\theta)^T\mu(\theta))]}{\left(\sqrt{2\pi}\right)^m\sqrt{\det(\Pi_2(\theta)^T\Omega\Pi_2(\theta))}}.$$

and the theoretical distribution on the nonsingular dimensions can be obtained by setting $y_1^{**} = 0$.

$$f_{y,0}(y_2^{**}|\theta) = \frac{\exp\left[-\frac{1}{2}y_2^{**T}\Lambda^{-1}y_2^{**}\right]}{\left(\sqrt{2\pi}\right)^m\sqrt{\det(\Lambda)}},$$

where $\Lambda = diag(\lambda_1(\theta), \lambda_2(\theta))$. Therefore, the reality bound is

$$p_0 = \sup_\theta \inf_y \frac{f_y(y_2^*)}{f_{y,0}(y_2^{**}|\theta)}. \tag{4.4}$$

It is shown by Bierens and Swanson (2000) that, under some mild condition[2], the reality bound in (4.4) can be further written as

$$\begin{aligned} p_0 &= \sup_{\theta,\lambda_{\max}[\Psi]<1}\left\{\sqrt{\det\Psi}\exp\left[-\frac{1}{2}\vartheta_t(\theta)'\Psi\vartheta_t(\theta)\right]\right. \qquad (4.5)\\ &\left.\times\exp\left[-\frac{1}{2}\vartheta_t(\theta)'\Psi(I-\Psi)^{-1}\Psi\vartheta_t(\theta)\right]\right\}, \end{aligned}$$

where

$$\Psi = \Lambda^{1/2}(\Pi_2'\Omega\Pi_2)^{-1}\Lambda^{1/2},$$

and

$$\vartheta_t(\theta) = \Lambda^{-1/2}\Pi_2'(\omega_t - \mu_t(\theta)).$$

[2]Bierens and Swanson (2000) show that the infimum in p_0 can be taken out if the maximum eigenvalue of the matrix $\Lambda(\theta)^{1/2}(\Pi_2(\theta)'\Omega\Pi_2(\theta))^{-1}\Lambda(\theta)^{1/2}]$ is less than one.

Λ is a 2×2 diagonal matrix of positive eigenvalues of $\Sigma(\theta)$, and $\lambda_{\max}[\Psi]$ represents the maximal eigenvalue of Ψ. The maximal p_0 then can be estimated according to (4.5).

In the case of conditional densities, p_0 depends on observations x_t. Therefore, taking the average of $p_0(x_t)$ over t gives the Bierens-Swanson average conditional reality bound

$$\begin{aligned} \widehat{p} &= \sup_{\theta, \lambda_{\max}[\Psi]<1} \frac{1}{n} \sum_{t=1}^{n} \left\{ \sqrt{\det \Psi} \exp\left[-\frac{1}{2} \vartheta_t(\theta)' \Psi \vartheta_t(\theta) \right] \right. \\ & \left. \times \exp\left[-\frac{1}{2} \vartheta_t(\theta)' \Psi (I - \Psi)^{-1} \Psi \vartheta_t(\theta) \right] \right\} \end{aligned} \tag{4.6}$$

In this approach, the variables of the empirical model on the singular direction are integrated out and thus not compared when one estimates p_0. In other words, the consequence of imposing the ceteris paribus condition, such as model misspecification, in the theoretical model is not penalized. In Section 4.4, I introduce another approach where the singular part of the theoretical model is considered.

4.4 The Bierens multiplicative conditional reality bound

As mentioned earlier, using the Bierens-Swanson reality bound to estimate the deep parameters may have the problem of lack of identification because it focuses only on the distributions on the nonsingular direction by marginalizing the distributions in the singular direction. Moreover, the singular part of the model is not penalized because the distribution of DGP in the nonsingular direction is integrated out when compared with the theoretical distribution. Bierens (2003) then suggests augmenting both the theoretical and empirical models by the same independent nonsingular stochastic disturbances, when estimating the reality bound.

Let R be a k-variate independent nonsingular normal noise, $R \sim i.i.d.N(0, \Omega)$, and $\det(\Omega) \neq 0$. Let $y^* = R + y$ represent the transformed vector of variables in the empirical model and $y^{**} = R+y$ be the transformed vector of variables in the theoretical model, then both distributions of the theoretical and empirical models are nonsingular. Similar to the Bierens-Swanson reality bound, the reality bound based on convolution can be defined as

$$p_0(x|\theta) = \inf_y \frac{f(y^*|x)}{f_0(y^{**}|x, \theta)},$$

where $p_0(x)$ can be interpreted as the probability that the distribution of the theoretical model is equal to that of the empirical model.

Moreover, given the data $x_1, ..., x_T$, the probability that the joint distribution of the theoretical model is the same as the joint distribution of DGP can be written as the product of $p_0(x_t|\theta)$

$$\prod_{t=1}^{T} p_0(x_t|\theta). \tag{4.7}$$

By maximizing (4.7), one can further estimate the parameters θ

$$\theta(x) = \arg\max_{\theta} \sum_{t=1}^{T} \ln p_0(x_t|\theta), \tag{4.8}$$

which provides an alternative to calibration.

4.5 Implementation of the multiplicative reality bound

In the case of the $VAR(q)$ model in Section 3.3, for instance, let $R_t = \tau\Omega$, where $\tau \in \mathbf{R}^+$ is a constant that decides the degree of shock to be added into both the theoretical and empirical models, and Ω is the variance matrix of the empirical model. Then (4.3) becomes

$$Y_t^{TM} = A(\theta)X_t + B(\theta)U_t + R_t \sim N_k\left(\mu_t(\theta), \Sigma(\theta) + \tau\Omega\right), \tag{4.9}$$

where μ_t represents the conditional mean of Y_t^{TM}, and it depends on both the observation X_t and the parameter θ. The density of theoretical model is

$$f_{TM}(y|X_t,\theta) = \frac{\exp\left[-\frac{1}{2}(y_t-\mu_t(\theta))^T(\Sigma(\theta)+\tau\Omega)^{-1}(y_t-\mu_t(\theta))\right]}{\left(\sqrt{2\pi}\right)^k\sqrt{\det\left(\Sigma(\theta)+\tau\Omega\right)}}.$$

Furthermore, the distribution of the empirical model after convolution in (4.2) becomes

$$Y_t^{EM} = \Gamma Z_t + V_t + R_t,\ \ V_t \sim N_k\left(\omega_t,(1+\tau)\Omega\right), \tag{4.10}$$

where ω_t is the conditional mean of Y_t^{EM}, and the empirical conditional density is

$$f_{EM}(y|Z_t,\Gamma,\Omega) = \frac{\exp\left[-\frac{1}{2(1+\tau)}(y_t-\omega_t)^T\Omega^{-1}(y_t-\omega_t)\right]}{\left(\sqrt{2\pi}\right)^k\sqrt{(1+\tau)^k\det\left(\Omega\right)}}.$$

Therefore, similar to the Bierens-Swanson reality bound, the reality bound based on convolution can be written as

$$\begin{aligned}
p_t(\Gamma,\Omega,\theta|\tau) &= \inf_y \frac{f_{EM}(y|Z_t,\Gamma,\Omega)}{f_{TM}(y|X_t,\theta)}, \\
&= \sqrt{\frac{\det\left(\Sigma(\theta)+\tau\Omega\right)}{(1+\tau)^k\det\left(\Omega\right)}}\times\exp\left\{-\frac{1}{2}\left[y_t^T\left((\Sigma(\theta)+\tau\Omega)^{-1}-\frac{1}{1+\tau}\Omega^{-1}\right)y_t\right]\right. \\
&\quad -2y^T\left((\Sigma(\theta)+\tau\Omega)^{-1}-\frac{1}{1+\tau}\Omega^{-1}\right)\mu_t(\theta) \qquad (4.11) \\
&\quad +\frac{2}{1+\tau}y^T\Omega^{-1}\left(\omega_t-\mu_t(\theta)\right) \\
&\quad \left.+\mu_t(\theta)^T(\Sigma(\theta)+\tau\Omega)^{-1}\mu_t(\theta)-\frac{1}{1+\tau}\omega_t\Omega^{-1}\omega_t\right\},
\end{aligned}$$

which can be interpreted as the probability that Y_t^{TM} and Y_t^{EM} have the same conditional distributions.

It has been proven by Bierens (2003) that if all the eigenvalues of $\Omega^{-1/2}\Sigma(\theta)\Omega^{-1/2}$ are less than 1, $\left[(\Sigma(\theta)+\tau\Omega)^{-1}-\frac{1}{1+\tau}\Omega^{-1}\right]$ is positive definite and thus (4.11) is minimized when

$$y=\mu_t(\theta)-\frac{1}{1+\tau}\left[(\Sigma(\theta)+\tau\Omega)^{-1}-\frac{1}{1+\tau}\Omega^{-1}\right]\Omega^{-1}\left(\omega_t-\mu_t(\theta)\right). \tag{4.12}$$

Substituting (4.12) into (4.11), the multiplicative reality bound takes the form

$$\begin{aligned}
p_t(\Gamma,\Omega,\theta|\tau) =& \inf_y \frac{f_{EM}(y|Z_t,\Gamma,\Omega)}{f_{TM}(y|X_t,\theta)}, \\
= & \sqrt{\frac{\tau^{k-m}\det(\Lambda_1(\theta)+\tau I_m)}{(1+\tau)^k}} \qquad (4.13)\\
& \times\exp\left\{-\frac{1}{2}(\omega_t-\mu_t(\theta))^T\Omega^{-1/2}Q(\theta)(I_k-\Lambda(\theta))^{-1}Q(\theta)\Omega^{-1/2}(\omega_t-\mu_t(\theta))\right\},
\end{aligned}$$

where

$$\begin{aligned}
\Omega^{-1/2}\Sigma(\theta)\Omega^{-1/2} &= Q(\theta)\Lambda(\theta)Q(\theta)^T \\
&= (Q_1(\theta),Q_2(\theta))\begin{pmatrix}\Lambda_1(\theta) & 0\\ 0 & 0\end{pmatrix}(Q_1(\theta),Q_2(\theta))^T,
\end{aligned}$$

$$\Lambda_1(\theta)=diag(\lambda_1(\theta),...,\lambda m(\theta)),\ 0<\lambda_m(\theta)\le...\le\lambda_1(\theta)<1,$$

and $Q_1(\theta)$ and $Q_2(\theta)$ are $k \times m$ and $k \times (k-m)$ matrices of the corresponding orthogonal eigenvectors.

Moreover,

$$\prod_{t=1}^{n} p_t(\Gamma, \Omega, \theta | \tau) \tag{4.14}$$

is the probability that, conditional on the data, the joint distribution of $Y_1^{TM}, ..., Y_n^{TM}$ is the same as the joint distribution of $Y_1^{EM}, ...Y_n^{EM}$. Maximizing (4.14) provides an alternative way to estimate the parameters in the linearized theoretical model by

$$\theta(\Gamma, \Omega, \tau) = \arg \ \max \frac{1}{n} \sum_{t=1}^{n} \ln p_t(\Gamma, \Omega, \theta | \tau). \tag{4.15}$$

Consequently, after some tedious calculation, the log of the multiplicative reality bound can be written as

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^{n} \ln p_t(\Gamma, \Omega, \theta, Q, \Lambda_1 | \tau) \quad &= \quad \frac{k-m}{2} \ln \tau - \frac{k}{2} \ln(1+\tau) + \frac{1}{2} \ln \left(\det(\Lambda_1(\theta) + \tau I_m) \right) \\ &\quad -\frac{1}{2} \text{trace} \left((I_k - \Lambda(\theta))^{-1} Q^T \Gamma_n(\theta) Q \right), \end{aligned} \tag{4.16}$$

where

$$\Gamma_n(\theta) = \Omega^{-1/2} \left(\frac{1}{n} \sum_{t=1}^{n} (\omega_t - \mu_t(\theta))(\omega_t - \mu_t(\theta))^T \right) \Omega^{-1/2},$$

which can be decomposed as a product of matrices of eigenvalues and eigenvectors

$$\Gamma_n(\theta) = S(\theta) \Xi(\theta) S(\theta)^T,$$

where $\Xi(\theta) = diag(\xi_{n,1}(\theta), ..., \xi_{n,k}(\theta))$, $S(\theta)S(\theta)^T = I$, and $\xi_{n,1}(\theta) \leq ... \leq \xi_{n,k}(\theta)$ are the eigenvalues of $\Gamma_n(\theta)$.

Furthermore, Bierens (2003) shows that (4.16) is maximized if $Q(\theta) = S(\theta)$ and $\lambda_j(\theta) = 1 + \xi_{n,j}(\theta)/2 - 1/2\sqrt{\xi_{n,j}^2(\theta) + 4(1+\tau)\xi_{n,j}(\theta)}$, which satisfies $\lambda_j(\theta) < 1$, if $\xi_{n,j}(\theta) > 0$, and $\lambda_j(\theta) > 0$, if $\xi_{n,j}^{-1}(\theta) > \tau$, for $j = 1, ..., m$. Then

$$\begin{aligned} & \max_{Q,\Lambda_1,\lambda_{\max}[\Lambda_1]<1} \frac{1}{n}\sum_{t=1}^{n} \ln p_t(\Gamma, \Omega, \theta, Q, \Lambda_1|\tau) \\ = \; & \frac{k-m}{2}\ln\tau - \frac{k}{2}\ln(1+\tau) + \frac{1}{2}\sum_{j=1}^{m}(\ln\lambda_j(\theta,\tau) + \tau) \\ & -\frac{1}{2}\sum_{j=1}^{m}\frac{\xi_{n,j}(\theta)}{1-\lambda_j(\theta,\tau)} - \frac{1}{2}\sum_{j=m+1}^{k}\xi_{n,j}(\theta), \end{aligned} \tag{4.17}$$

and the deep parameters of the theoretical model can be obtained by

$$\widehat{\theta} = \arg \max_{Q,\Lambda_1,\lambda_{\max}[\Lambda_1]<1} \frac{1}{n}\sum_{t=1}^{n} \ln p_t(\Gamma, \Omega, \theta, Q, \Lambda_1|\tau). \tag{4.18}$$

Chapter 5

The empirical results

This chapter estimates the linearized version of the cash-in-advance model that is developed in Chapter 2. When applying the Bierens-Swanson and Bierens reality bounds, the empirical model based on the data evidence is used to replace the DGP and compared with the theoretical model since the DGP is unknown. We also use an approach different from that used in Bierens and Swanson (2000) to estimate the average conditional reality bound to simplify the whole process in estimation. The effects of technology and monetary shocks on inflation and output growth are also discussed.

Section 4.1 discusses the data properties. Section 4.2 estimates the empirical model that will be used to represent the DGP. Section 4.3 applies the Bierens-Swanson reality bound to evaluate the CIA model. Section 4.4 applies the Bierens multiplicative reality bound to estimate the deep parameters. Section 4.5 uses innovation response analysis to examine the effects of the stochastic shocks.

5.1 The data

The data set that is used for estimating the theoretical and empirical models are taken from Citibase database. The sample period is from 1948:1 to 2001:2, so there are 214 quarterly observations. The five U.S. time-series used include CPI, real GNP, consumption of nondurables and services, gross private domestic investment (all in 1996 dollars), and also working hours. The working hours used here only include the employee-hours in nonagricultural industries. All the variables used are already seasonally adjusted and natural logged. These five series are shown in Figure 5.1. All of the series have linear trends, except the working hours $\ln H$. The main difference between the data used here and the data used in Cooley and Hansen (1989) is that in the latter the Hodrick-Prescott filter is used to remove the time trend.

The augmented Dickey-Fuller and Phillips-Perron unit root tests for each series are conducted. The test results show that $\ln P$, $\ln C$, $\ln H$, and $\ln Y$ are all $I(1)$ processes, except the result of $\ln X$ is mixed. The unit root test of $\ln X$ is not rejected by the augmented Dickey-Fuller test, but is rejected by the Phillips-Perron test

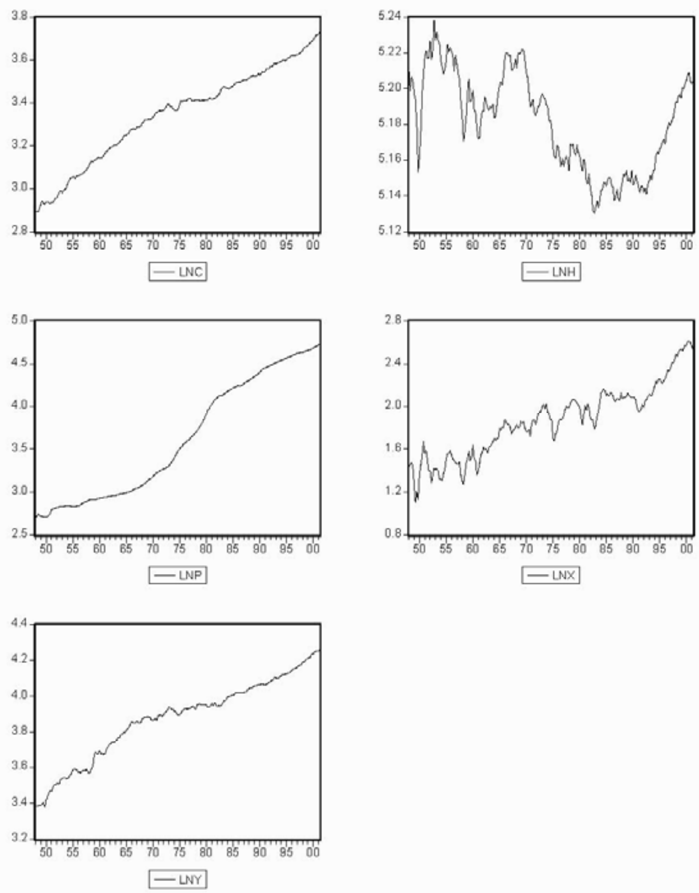

Figure 5.1: The five U.S. series

at a 5% significance level. Since the test results are not consistent, the simulation of the actual p-value of the Phillips-Perron test is used. The simulated p-value based on 1000 simulations is 0.368, which means that the unit root hypothesis is not rejected. Therefore, based on these test results, we conclude that all variables are $I(1)$ processes.

5.2 The empirical model

Since the DGP is unknown, the empirical model is used to replace the DGP and compared with the theoretical model. Given the data properties found out in Section 5.1, I first parameterize the empirical model as an error correction model with p lags as follows:

$$\begin{aligned}\Delta W_t &= \Gamma Z_t + V_t \\ &= \pi_0 + \sum_{j=1}^{p-1} \pi_j \Delta W_{t-j} + \alpha\beta^T W_{t-1} + V_t,\end{aligned}$$

where $V_t \sim N_5(0, \Omega)$ and $\det(\Omega) \neq 0$. Therefore, the conditional distribution of ΔW_t is

$$\Delta W_t | Z_t \sim i.i.d.\ N_5(\omega_t, \Omega),$$

where

$$Z_t = (1, \Delta W'_{t-1}, \cdots, \Delta W'_{t-p+1}, W'_{t-1})',$$

and

$$\omega_t = \pi_0 + \sum_{j=1}^{p-1} \pi_j \Delta W_{t-j} + \alpha\beta^T W_{t-1}.$$

Two consistent criteria, the Hannan-Quinn criterion (HQ) and the Schwartz criteria (SC), have been used widely in applied work lately, so I use them to decide the number of lags for the empirical model. I choose the lag order for the error correction model by examining values of HQ and SC for lag order from one to ten. The HQ suggests two lags, while the SC selects one lag. Since the SC is more conservative than the HQ, and it has been shown by Lutkepohl (1993) that $\widehat{p}(SC) \leq \widehat{p}(HQ)$ for all T, where $\widehat{p}$ represents the number of lags suggested by the criterion and T is the sample size. Accordingly, I choose the error correction model with two lags.

Next, as Figure 5.1 shows, all of the five series have linear trends, except $\ln H$, so an error correction model with the intercept term is considered. Moreover, I consider two error correction models. One has cointegrating restriction on the intercept parameters in the error correction equation, i.e., $\Delta W_t = \pi_1 \Delta W_{t-1} + \alpha\beta^T (W_{t-1} - \delta_0)$. The other one has no restriction on the intercept term, i.e., $\Delta W_t = \pi_0 + \pi_1 \Delta W_{t-1} + \alpha\beta^T W_{t-1}$. Given these two models, I also conduct Johansen's lambda max and trace tests to decide the cointegration rank. The test of imposing the restriction in the model is rejected and both of Johansen's tests suggest one cointegrating vector under the second model.

Therefore, an error correction model with two lags is chosen as the empirical model:

$$\Delta W_t = \pi_0 + \pi_1 \Delta W_{t-1} + \alpha\beta^T W_{t-1} + \epsilon_t,$$

where α is the adjustment coefficient and β is the cointegration coefficient. Since there is only one cointegrating relationship, both α and β are 5×1 matrices and $\alpha\beta^T$ is a 5×5 matrix with rank one.

The estimated results are shown in Table 5.1. The empirical model has the error correction term

$$\ln C_t - 0.014515 \ln P_t - 0.150909 \ln X_t + 0.526272 \ln H_t - 0.709652 \ln Y_t,$$

which represents the equilibrium relationship among the observables. This empirical model will be used to represent the DGP in following analysis.

5.3 Estimation of the Bierens-Swanson average conditional reality bound

The Bierens-Swanson average reality bound is a measure of the information about the DGP contained in the theoretical model. It assumes that the theoretical model is only an approximation of the DGP; therefore, this reality bound does not rely on a null hypothesis that the theoretical model is equal to the DGP. Moreover, it is an

Table 5.1: Estimated results of the empirical error correction model.

	$\Delta \ln p_t$	$\Delta \ln X_t$	$\Delta \ln C_t$	$\Delta \ln H_t$	$\Delta \ln Y_t$
π_0^T	0.083804	-2.594961	0.119320	0.133220	-0.286075
	0.773510	-0.183231	-0.140898	-0.012110	-0.144154
	0.019726	0.001275	-0.000686	0.037889	-0.009099
π_1^T	-0.051514	0.780754	0.126317	0.126547	0.160191
	-0.126529	1.583506	-0.299232	-0.063446	0.293419
	0.022300	-0.821750	0.006262	0.017970	-0.109158
α^T	-0.027337	0.870706	-0.038333	-0.044787	0.097519
β^T	-0.014515	-0.150909	1	0.526272	-0.709652
$s.e.$	5.04734×10^{-3}	5.19469×10^{-2}	5.31770×10^{-3}	5.69996×10^{-3}	8.91614×10^{-3}
R^2	0.6358	0.1272	0.1509	0.0952	0.0959

information measure, which summarizes the ability of a theoretical model in explaining the data. In this section, the approach suggested by Bierens and Swanson (2000) is applied, but the estimation method is different from theirs due to the complexity of the CIA model.

We now follow the lines in Bierens and Swanson (2000) to compute the average reality bound. According to the results in section 6.3, the error correction model derived from the theoretical model has one lag and can be written as

$$\Delta W_t = \nu_0 + \alpha\beta^T W_{t-1} + \varepsilon_t,$$

and the model conditional distribution of ΔW_t is

$$\Delta W_t | W_{t-1} \sim N_5(\mu_t(\theta), \Sigma(\theta)).$$

Since there are only two, instead of five, stochastic shocks in the Cooley and Hansen model, the theoretical model has a 5-variate singular normal distribution $N_5(\mu_t(\theta), \Sigma(\theta))$, where $rank(\Sigma(\theta)) = 2$. Moreover, both α and β are 5×2 matrices, and thus the theoretical model implies two cointegrating relationships.

Furthermore, as shown in section 5.2, the empirical model has two lags and one cointegrating relation:

$$\Delta W_t = \pi_0 + \pi_1 \Delta W_{t-1} + \alpha \beta^T W_{t-1} + V_t,$$

with the conditional distribution of ΔW_t is

$$\Delta W_t | \Delta W_{t-1}, W_{t-1} \sim N_5(\omega_t, \Omega).$$

Recall that the Bierens-Swanson average conditional reality bound can be estimated by

$$\begin{aligned} \widehat{p} &= \sup_{\theta, \lambda_{\max}[\Psi]<1} \frac{1}{n} \sum_{t=1}^{n} \left\{ \sqrt{\det \Psi} \exp\left[-\frac{1}{2} \vartheta_t(\theta)' \Psi \vartheta_t(\theta) \right] \right. \\ &\quad \left. \times \exp\left[-\frac{1}{2} \vartheta_t(\theta)' \Psi (I - \Psi)^{-1} \Psi \vartheta_t(\theta) \right] \right\}, \end{aligned} \tag{3.6}$$

where

$$\Psi = \Lambda^{1/2} (\Pi_2' \Omega \Pi_2)^{-1} \Lambda^{1/2},$$

and

$$\vartheta_t(\theta) = \Lambda^{-1/2}\Pi_2'(\omega_t - \mu_t(\theta)).$$

One way to estimate p_0 is to obtain the parameter values in θ and then substitute them into (4.6). However, due to the complexity of Ψ in (4.6), we suggest a different approach to simplify the computation by using the information contained in the empirical model.

First, we decompose Ω as

$$\Omega = Q\Lambda^* Q^T = (Q_1, Q_2)\begin{pmatrix} \Lambda_1^* & 0 \\ 0 & \Lambda_2^* \end{pmatrix}\begin{pmatrix} Q_1^T \\ Q_2^T \end{pmatrix},$$

where

$$\Lambda^* = \begin{pmatrix} \Lambda_1^* & 0 \\ 0 & \Lambda_2^* \end{pmatrix}$$

contains the eigenvalues of Ω, and the eigenvalues are arranged in ascending order. Q_1 and Q_2 are 5×3 and 5×2 matrices, respectively. Λ_1^* and Λ_2^* are 3×3 and 2×2 matrices with the eigenvalues of Ω as their diagonal elements. Then, replacing Ω in Ψ by $Q\Lambda^* Q^T$ yields

$$\Psi = \Lambda^{1/2}(\Pi_2' Q\Lambda^* Q^T \Pi_2)^{-1}\Lambda^{1/2}.$$

Let $R = Q^T\Pi_2$, which is a 5×2 matrix, and $R^T R = I$. Then

$$\Psi = \Lambda^{1/2}(R^T\Lambda^* R)^{-1}\Lambda^{1/2}.$$

The key of this approach is to simplify Ψ by imposing the assumption $\Pi_2 = Q_2$. In other words, we let that two of the eigenvectors in the empirical and theoretical models are equal. Imposing this condition not only simplifies the computation, but more importantly makes the theoretical variance-covariance matrix closer to the variance in reality since Q_2 is estimated from the empirical model.

However, one drawback of imposing this condition is that some link between the parameters in $\Sigma(\theta)$ and the deep parameters in the theoretical model is broken; therefore, some parameters cannot be estimated or recovered by this method because of losing some information contained in $\Sigma(\theta)$. In addition, this condition adds a restriction on the variance matrix $\Sigma(\theta)$, and thus the estimated reality bound will actually be a lower bound of p_0. Imposing the restriction $\Pi_2 = Q_2$ yields

$$R = Q^T \Pi_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and thus

$$R^T \Lambda^* R = \Lambda_2^*.$$

Consequently, Ψ can be rewritten as

$$\Psi = \Lambda^{1/2}\Lambda_2^{*-1}\Lambda^{1/2}.$$

In addition, the constraint in (4.6), that $\lambda_{\max}[\Psi]$ is less than 1, is equivalent to restricting the maximal root of the eigenvalue problem

$$\det(\Lambda_2^{*-1} - \lambda\Lambda^{-1}) = 0 \tag{5.1}$$

to a value less than 1. Moreover, (5.1) can also be written as

$$\det(\Lambda - \lambda\Lambda_2^*) = 0,$$

which is equivalent to

$$\prod_{j=1}^{2} (\lambda_j/\lambda_j^* - \lambda) = 0.$$

Therefore,

$$\lambda_{\max}[\Psi] = \max_{1\leq j\leq 2} \frac{\lambda_j}{\lambda_j^*}, \tag{5.2}$$

which should be less than one.

Using the above results, the maximization of the average reality bound in (4.5) can be further simplified. Let $\lambda_1 = c_1\lambda_1^*$ and $\lambda_2 = c_2\lambda_2^*$ be the optimal solution in maximizing (4.5), where both c_1 and $c_2 \in (0, 1)$ are constants. Then Ψ becomes

$$\Psi = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix},$$

and (4.5) can be rewritten as

$$
\begin{aligned}
\widehat{p} &= \sup_{\theta,\lambda_{\max}[\Psi]<1} \frac{1}{n}\sum_{t=1}^{n}\left\{\sqrt{\det\Psi}\exp\left[-\frac{1}{2}a_t(\theta)'\Lambda^{-1/2}\Psi\Lambda^{-1/2}a_t(\theta)\right]\right. \\
&\quad \left.\times\exp\left[-\frac{1}{2}a_t(\theta)'\Lambda^{-1/2}\Psi(I-\Psi)^{-1}\Psi\Lambda^{-1/2}a_t(\theta)\right]\right\}, \\
&= \sup_{\theta,0<c_1,c_2<1} \frac{1}{n}\sum_{t=1}^{n}\left\{\sqrt{\det\Psi}\exp\left[-\frac{1}{2}\left(a_{1t}^2(\theta)/\lambda_1^* + a_{2t}^2(\theta)/\lambda_2^*\right)\right]\right. \\
&\quad \left.\times\exp\left[-\frac{1}{2}\left(\frac{c_1}{(1-c_1)\lambda_1^*}a_{1t}^2(\theta) + \frac{c_2}{(1-c_2)\lambda_2^*}a_{2t}^2(\theta)\right)\right]\right\}, \qquad (5.3)
\end{aligned}
$$

where

$$
a_t(\theta) = \begin{pmatrix} a_{1t}(\theta) \\ a_{2t}(\theta) \end{pmatrix} = \Pi_2'(\omega_t - \mu_t(\theta)).
$$

The restriction $\lambda_{\max}[\Psi] < 1$ in (4.5) while maximizing $\widehat{p}$ over θ is equivalent to the restrictions $0 < c_1 < 1$ and $0 < c_2 < 1$. The estimated result of the theoretical model is shown in the second and third columns of Table 5.2.

Since Π_2 is a 5×2 matrix and α is a 5×2 vector, there are five unknowns in α but only two equations. α cannot be solved by α^* because of underidentifying. For the same reason, μ cannot be solved by μ^*. Substituting the result into (4.5), we obtain the average reality bound

$$
p_0 = 0.473989,
$$

Table 5.2: The estimated results of the Bierens-Swanson average conditional reality bound. $\alpha^* = \Pi_2'\alpha$, $\nu_i^* = \Pi_2'\nu_i$, $i = 1, 2$.

	ECM(1)		ECM(2)	
α^*	0.086405	−0.000203	−0.035978	−0.033471
	−0.157441	0.3437208	0.068734	0.768659
	0.067988	−0.113876	0.049727	−0.027650
	−0.107506	−0.028441	−0.105175	−0.129882
β	0.004592	1	0.017184	1
	1	−2.147265	1	0.239859
	−0.12338	−0.821485	0.052610	−0.721728
ν_0^*	−0.447942	–	0.249045	–
	4.739475	–	−1.479041	–
	–	–	−0.308585	−0.316561
	–	–	0.004048	0.023821
ν_1^{*T}	–	–	0.183906	0.908812
	–	–	−0.070266	1.394924
	–	–	−0.009533	−0.837362
c_1	0.739475	–	0.949792	–
c_2	0.649947	–	0.944327	–
p_0	0.473989		0.897437	

which suggests that the nonsingular part of the theoretical model can explain about 47% of the corresponding part of the data-generating process. Comparing with the result in Bierens and Swanson (2000), where $p_0 = 0.77$, the estimated p_0 is relatively low and suggests that the reality content of the Cooley and Hansen model is relatively lower than the KPR model. Comparing the results in Table 5.1 and 5.2, the standardized cointegrating vectors of the theoretical and empirical models are quite a bit different from the cointegrating vector of the empirical model, which is mainly caused by the lag and rank differences between these two models since only the nonsingular part of the distributions of the empirical and theoretical models are compared. In Section 5.4, we use an approach similar to that used in Section 3.5 to show that the theoretical model implies an error correction model with two lags under a different assumption of stochastic structure. Whether the lag difference can explain the difference between the empirical and theoretical models is further examined.

5.4 Interpretation of the discrepancy between the empirical and the theoretical model

Since the assumption that the two stochastic shocks follow AR(1) processes is not essential in the CIA model, in this section we consider the case where the shocks follow

AR(2) processes, in order to examine whether the main penalty of the Bierens-Swanson reality bound is caused by the lag difference between the empirical and theoretical models. We change the assumption of the stochastic shocks in (3.30) to

$$u_t = \Gamma_1 u_{t-1} + \Gamma_2 u_{t-2} + e_t, \tag{5.4}$$

and (3.28) and (3.29) remain unchanged:

$$W_t = \delta_0 + D_1 u_t + \delta_4 \ln K_t + \delta_5 u_{t-1},$$

and

$$\Delta \ln K_t = \eta^T u_{t-1}.$$

Using a procedure similar to that used in section 6.3, u_{t-1} can be written as

$$u_{t-1} = I_0 (D_2^T D_2)^{-1} D_2^T (W_{t-1} - \delta_0). \tag{5.5}$$

Then

$$\begin{aligned} \Delta W_t &= D_1 \Delta u_t + \delta_4 \Delta \ln K_t + \delta_5 \Delta u_{t-1}, \\ &= D_1(\Gamma_1 u_{t-1} + \Gamma_2 u_{t-2} + e_t) + (\delta_4 \eta^T + \delta_5 - D_1) u_{t-1} - \delta_5 u_{t-2}, \\ &= (D_1(\Gamma_1 - I_2) + \delta_4 \eta^T + \delta_5) u_{t-1} + (D_1 \Gamma_2 - \delta_5)\, u_{t-2} + D_1 e_t, \\ &= (D_1(\Gamma_1 - I_2) + \delta_4 \eta^T + \delta_5) I_0 (D_2^T D_2)^{-1} D_2^T (W_{t-1} - \delta_0) \end{aligned}$$

$$+ (D_1\Gamma_2 - \delta_5)\, I_0(D_2^T D_2)^{-1} D_2^T (W_{t-2} - \delta_0) + D_1 e_t,$$

$$= \; - (D_1\Gamma_2 - \delta_5)\, I_0(D_2^T D_2)^{-1} D_2^T \Delta W_{t-1}$$

$$+ (D_1(\Gamma_1 + \Gamma_2 - I_2) + \delta_4 \eta^T) I_0(D_2^T D_2)^{-1} D_2^T \,(W_{t-1} - \delta_0) + D_1 e_t;$$

therefore,

$$\Delta W_t = \nu_0 + \nu_1 \Delta W_{t-1} + \alpha\beta^T W_{t-1} + v_t, \tag{5.6}$$

where

$$\nu_0 = [-(D_1(\Gamma_1 + \Gamma_2 - I_2) + \delta_4 \eta^T) I_0(D_2^T D_2)^{-1} D_2^T \delta_0],$$

$$\nu_1 = -\,(D_1\Gamma_2 - \delta_5)\, I_0(D_2^T D_2)^{-1} D_2^T,$$

$$\alpha = D_1(\Gamma_1 + \Gamma_2 - I_2) + \delta_4 \eta^T,$$

$$\beta = D_2(D_2^T D_2)^{-1} I_0^T.$$

Consequently, the CIA model implies a cointegrated error correction model with two lags, as shown in (5.6), given the assumption that both of the shocks follow AR(2) processes.

Applying the same estimation procedure used in (5.3), the estimated results are shown in the fourth and fifth columns of Table 5.2. In this case, the average reality bound is 0.897437, which means that the information content in the theoretical model is about 90% of that found by the empirical model. Comparing this with the previous result, it is clear that the lag difference leads to the major penalty of the Bierens-Swanson reality bound. The remaining 10% penalty may be due to the rank difference. Moreover, the second cointegration vector estimated by the theoretical model is also very close to the cointegration vector of the empirical model.

Our finding shows that along the nonsingular direction the Cooley and Hansen model cannot successfully account for the data, mainly because of the lag difference. The CIA model accounts for 47% of the reality in the nonsingular dimension. 43% of the penalty comes from the lag difference, which is implied by the stochastic process. The other 10% of the penalty can be accounted by the rank difference between the theoretical and empirical model.

5.5 Estimation of the Bierens multiplicative conditional reality bound

In order to estimate, instead of calibrate, the deep parameters in the theoretical model, Bierens (2003) suggests augmenting both the theoretical and empirical models by the same independent stochastic disturbances, R_t^* and R_t^{**}. For instance, let R_t^* and $R_t^{**} \sim i.i.d.N(0, \tau\Omega)$, where $\det(\Omega) \neq 0$. Given the error correction model in Section 3.5, the theoretical model becomes

$$\Delta W_t^* = \nu_0 + \alpha\beta^T W_{t-1} + \varepsilon_t + R_t^*,$$

with the conditional distribution

$$\Delta W_t^* | W_{t-1} \sim N_5(\mu_t(\theta), \Sigma(\theta) + \tau\Omega),$$

and the density

$$f_{TM}(y|\theta,\tau) = \frac{1}{(\sqrt{2\pi})^5\sqrt{\det(\Sigma(\theta)+\tau\Omega)}} \exp\left\{-\frac{1}{2}(y-\mu_t(\theta))^T(\Sigma(\theta)+\tau\Omega)^{-1}(y-\mu_t(\theta))\right\}.$$

Similarly, the empirical model after augmenting with stochastic errors can be written as

$$\begin{aligned}\Delta W_t^{**} &= \Gamma Z_t + V_t,\\ &= \pi_0 + \pi_1\Delta W_{t-1} + \alpha\beta^T W_{t-1} + V_t + R_t^{**},\end{aligned}$$

with the conditional distribution

$$\Delta W_t^{**} | \Delta W_{t-1}, W_{t-1} \sim N_5(\omega_t, (1+\tau)\Omega)$$

and the density

$$f_{EM}(y|\theta,\tau) = \frac{1}{(\sqrt{2\pi})^5\sqrt{(1+\tau)^5\det(\Omega)}} \exp\left\{-\frac{1}{2(1+\tau)}(y-\omega_t)^T\Omega^{-1}(y-\omega_t)\right\}.$$

Using (4.13), if all the eigenvalues of $\Omega^{-1/2}\Sigma(\theta)\Omega^{-1/2}$ are less than 1, the multiplicative reality bound takes the form

$$
\begin{aligned}
p_t(\Gamma,\Omega,\theta,Q,\Lambda_1|\tau) &= \inf_y \frac{f_{EM}(y|\tau)}{f_{TM}(y|\theta,\tau)}, \\
&= \sqrt{\frac{\tau^3 \det(\Lambda_1(\theta)+\tau I_2)}{(1+\tau)^5}} \\
&\quad \times \exp\left\{-\frac{1}{2}(\omega_t-\mu_t(\theta))^T\Omega^{-1/2}Q(\theta)(I_5-\Lambda(\theta))^{-1}Q(\theta)\Omega^{-1/2}(\omega_t-\mu_t(\theta))\right\},
\end{aligned}
$$

where

$$
\begin{aligned}
\Omega^{-1/2}\Sigma(\theta)\Omega^{-1/2} &= Q(\theta)\Lambda(\theta)Q(\theta)^T \\
&= (Q_1(\theta),Q_2(\theta))\begin{pmatrix} \Lambda_1(\theta) & 0 \\ 0 & 0 \end{pmatrix}(Q_1(\theta),Q_2(\theta))^T,
\end{aligned}
$$

$$
\Lambda_1(\theta) = diag(\lambda_1(\theta),\lambda_2(\theta)),\ 0 < \lambda_2(\theta) \le \lambda_1(\theta) < 1,
$$

and $Q_1(\theta)$ and $Q_2(\theta)$ are 5×2 and 5×3 matrices of the corresponding orthogonal eigenvectors. Consequently, (4.17) becomes

$$
\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\ln p_t(\Gamma,\Omega,\theta,Q,\Lambda_1|\tau) &= \frac{3}{2}\ln\tau - \frac{5}{2}\ln(1+\tau)+\frac{1}{2}\ln\left(\det(\Lambda_1(\theta)+\tau I_2)\right) \\
&\quad -\frac{1}{2}\mathrm{trace}\left((I_5-\Lambda(\theta))^{-1}Q^T\Gamma_n(\theta)Q\right), \qquad (5.7)
\end{aligned}
$$

and the parameters can be estimated by (5.7).

Given different values of τ, we maximize (5.7) over $\theta = (d_{11},\ d_{12},\ d_{13},\ d_{14},\ d_{22},\ d_{23},\ p^*,\ \overline{\alpha},\ \overline{\gamma},\ \rho,\ \delta,\ \ln \overline{g})^T$, where $p^*, \ln \overline{g} > 0$, and $\overline{\alpha}, \overline{\gamma}, \rho, \delta \in (0, 1)$. Since the set of parameters, θ, is in fact a subset of all deep parameters, we cannot obtain all values of the deep parameters here. The main reason is that in the process of linearization, some links between the model coefficients and the deep parameters, such as the discount rate in the utility function, are lost.

The estimated parameters shown in Table 5.3 and 5.4 are stable with only little fluctuation under different values of τ, but values of both ρ and δ are much higher than the usual calibrated values. For example, Cooley and Hansen (1989) use $\rho = 0.36$ and $\delta = 0.025$ in calibration. This suggests that some important features of the economy are missing from the model. In other words, the Cooley and Hansen model does not fit the data well although the failure is not so immediately evident. In spite of that the parameters may be restricted to a reasonable area or an area close to the calibrated values, I choose to leave the parameters free and let the data speak. Moreover, the estimated $\ln \overline{g}$ implies an average annual monetary growth rate between 3.8% to 5.7% for different τ values.

The 90% confidence interval of θ is also simulated by a bootstrap approach. Under the assumption that the empirical model is correctly specified, by Cholesky

Table 5.3: The estimation results of the mulitplicative reality bound. α is a 5×2 matrix, so the ten elements of α represent the first and second columns of α, respectively. So is β.

τ	0.1	0.2	0.3	0.4	0.5
d_{11}	-1.619234	-1.512823	-1.616875	-1.458609	-1.668025
d_{12}	-0.10916	-0.071137	-0.223706	-0.079165	-0.066965
d_{13}	0.296333	0.310649	0.333028	0.293085	0.273301
d_{14}	0.450385	0.503996	0.486075	0.479013	0.450597
d_{22}	-1.787301	-1.608197	-2.219154	-1.919055	-1.416365
d_{23}	1.255176	1.210487	1.318714	1.308205	1.08544
p^*	2.909936	2.756174	2.871515	2.510822	3.33487
$\overline{\alpha}$	0.024367	0.016469	0.02759	0.022449	0.002292
$\overline{\gamma}$	0.000276	0.000045	0.00069	0.000045	0.000078
ρ	0.685083	0.724187	0.693556	0.726532	0.686299
δ	0.728562	0.746744	0.759689	0.622586	0.87312
$\ln\overline{g}$	0.009448	0.010802	0.007361	0.010945	0.013176
v_0	0.01195	0.013924	0.009542	0.013835	0.016738
	0.011571	0.01286	0.009439	0.013997	0.014269
	-0.002732	-0.0033	-0.002384	-0.003136	-0.003593
	0.000785	0.000975	0.000553	0.0012	0.001142
	0.000247	0.000269	0.000169	0.000328	0.000358

Table 5.4: (Continue the above table) The estimation results of the mulitplicative reality bound. α is a 5×2 matrix, so the ten elements of α represent the first and second columns of α, respectively. So is β.

τ	0.1	0.2	0.3	0.4	0.5
	-0.477343	-0.534121	-0.595906	-0.493152	-0.490275
	0.484649	0.407213	0.531756	0.72419	0.179598
	0.477343	0.534121	0.595906	0.493152	0.490275
	-0.782309	-0.909008	-0.785469	-0.652465	-1.008403
	-2.138174	-2.120355	-2.409255	-2.046427	-2.164976
α					
	0.147119	0.166511	0.190706	0.126086	0.156657
	-0.310119	-0.286628	-0.280518	-0.464366	-0.135232
	-0.122752	-0.150042	-0.163116	-0.103637	-0.154365
	0.248924	0.30735	0.24522	0.186018	0.348953
	0.70488	0.739382	0.769959	0.642609	0.759886
	0	0	0	0	0
	0.015946	0.020352	0.02621	-0.009813	0.017823
	0.328613	0.347141	0.318413	0.313028	0.350173
	1	1	1	1	1
	-0.288901	-0.337351	-0.24148	-0.246081	-0.384911
β					
	0	0	0	0	0
	0.048525	0.058627	0.082314	-0.029656	0.051106
	1	1	1	1	1
	3.043089	2.880674	3.140578	3.193549	2.855649
	-0.879151	-0.9718	-0.758388	-0.787716	-1.099408
	-0.619234	-0.512823	-0.616875	-0.458609	-0.668025
	-0.316683	-0.292033	-0.274847	-0.473873	-0.135682
δ_0	1.619234	1.512823	1.616875	1.458609	1.668025
	5.056568	4.938796	5.152627	5.094518	4.724877
	1.592398	1.362182	1.57899	1.393189	1.482196

Table 5.5: The 90% confidence intervals of the estimated parameters when $\tau = 0.1$.

	L	U	$\tau = 0.1$		L	U	$\tau = 0.1$
d_{11}	-2.193731	-1.036577	-1.619234		-1.116254	-0.111663	-0.477343
d_{12}	-0.554408	0.297962	-0.10916		-0.006407	0.883481	0.484649
d_{13}	0.147071	0.646388	0.296333		0.111663	1.116254	0.477343
d_{14}	0.263902	0.721664	0.450385		-2.362188	-0.097241	-0.782309
d_{22}	-3.377627	0.277362	-1.787301		-3.38308	-1.137422	-2.138174
d_{23}	0.635488	2.192187	1.255176	α			
p^*	2.178594	3.94882	2.909936		0.030684	0.381615	0.147119
$\overline{\alpha}$	0.000695	0.031071	0.024367		-0.578817	-0.012813	-0.310119
$\overline{\gamma}$	3.90E-05	0.011439	0.000276		-0.380407	-0.015095	-0.122752
ρ	0.411646	0.852235	0.685083		0.004677	1.349834	0.248924
δ	0.59951	0.975084	0.728562		0.394092	1.654245	0.70488
$\ln \overline{g}$	0.002248	0.015261	0.009448				
					0	0	0
	0.003037	0.021012	0.01195		-0.118079	0.172992	0.015946
	0.001774	0.025537	0.011571		0.214314	0.673088	0.328613
v_0	-0.007328	-0.000502	-0.002732		1	1	1
	-5.00E-05	0.004891	0.000785		-0.862711	-0.014307	-0.288901
	-9.00E-06	0.00171	0.000247	β			
					0	0	0
	-1.193731	-0.036577	-0.619234		-0.421385	0.39417	0.048525
	-0.511642	-0.025231	-0.316683		1	1	1
δ_0	1.036577	2.193731	1.619234		1.48558	4.66604	3.043089
	3.893588	5.460241	5.056568		-1.640804	-0.051375	-0.879151
	0.580104	2.934035	1.592398				

decomposition one can write the variance matrix of the empirical model as

$$\widehat{\Omega} = P^T P.$$

Next, I randomly draw a 5×1 vector z_t from $N_5(0, I_5)$, and generate the disturbance $z_t^* = P^T z_t$, such that $E(z_t^* z_t^{*'}) = \widehat{\Omega}$, where $t = 1, ..., T$. The artificial data W_t^* can then be generated by

$$W_t^* = \widehat{\pi}_0 + (I_5 + \widehat{\pi}_1 + \widehat{\alpha}\widehat{\beta}^T)W_{t-1}^* - \widehat{\pi}_1 W_{t-2}^* + z_t^*, \qquad t = 3, ..., T,$$

where $W_t^* = W_t$ for $t = 1, 2$, and $\widehat{\pi}_0$, $\widehat{\pi}_1$, $\widehat{\alpha}$,and $\widehat{\beta}$ are the estimated coefficients in the empirical model. The deep parameters and the model coefficients can be estimated according to (4.18). Repeating the same steps 500 times, the bootstrap distribution of θ can then be obtained. The choice of $\tau = 0.1$ is arbitrary and, as shown in Table 7.3, the estimated parameters appear to be about the same under different values of τ. The bootstrap results at $\tau = 0.1$ are given in Table 5.5.

Furthermore, according to Table 5.5, the signs of the coefficients in the decision rules also show that monetary growth rate has positive effects on price level as well as investment. Our finding is consistent with the prediction of the cash-in-advance constraint. In other words, increasing money supply causes inflation. Under the cash-in-advance constraint, $\ln \widehat{p} = -\ln C$ implies that consumption decreases because of the higher price level. Since output can be only used either in consumption or investment, increasing of money supply then indirectly causes the decrease in investment. On the other hand, the effects of technology shocks on price level and investment are not significant. The bootstrap confidence interval also shows that $\ln X$ does not play a significant role in the cointegrating relationships.

5.6 Inflation and economic growth

The coefficients in the system of equations (5.7) for the case $\tau = 0.1$ can also be obtained by the Bierens multiplicative reality bound, as follows:

$$\begin{aligned}
\ln P_t &= -0.619234 - 0.109160 z_t + 1.296333 \ln g_t + \sum_{j=1}^{t-1} \ln g_{t-j} \qquad (5.8)\\
&\quad +0.450385 \ln K_t,\\
\ln X_t &= -0.316683 - 1.787301 z_t + 1.255176 \ln g_t + \ln K_t, \qquad (5.9)\\
\ln C_t &= 1.619234 + 0.109160 z_t - 0.296333 \ln g_t - 0.450385 \ln K_t, \qquad (5.10)\\
\ln H_t &= 5.056568 - 0.309707 z_t + 0.085107 \ln g_t + 0.363003 \ln K_t, \qquad (5.11)\\
\ln Y_t &= 1.592398 + 1.097532 z_t + 0.026802 \ln g_t + 0.799399 \ln K_t, \qquad (5.12)\\
\ln K_{t+1} &= -1.302159 z_t + 0.914473 \ln g_t + \ln K_t, \qquad (5.13)
\end{aligned}$$

and two stochastic processes

$$z_t = 0.000276 z_{t-1} + \epsilon_t, \tag{5.14}$$

$$\ln g_t = 0.024367 \ln g_{t-1} + \xi_t. \tag{5.15}$$

Equation (5.13) is equivalent to

$$\Delta \ln K_t = -1.302159 z_{t-1} + 0.914473 \ln g_{t-1}. \tag{5.16}$$

Taking the first difference of (5.8) gives the inflation rate. Together with (5.16), we have

$$\Delta \ln P_t = -0.109160 z_t - 4.773129 z_{t-1} + 1.296333 \ln g_t + 0.914473 \ln g_{t-1}. \tag{5.17}$$

Similarly, differencing (5.12) gives the output growth rate, which is

$$\Delta \ln Y_t = 1.097532 z_t - 2.138477 z_{t-1} + 0.026802 \ln g_t + 0.704223 \ln g_{t-1}. \tag{5.18}$$

The estimated parameters can be used to generate predictions about the responses of technology and monetary shocks. The innovation responses for both shocks for 15 periods are shown in Figure 5.2. The first panel in Figure 5.2 shows the innovation response of $\Delta \ln p$ corresponding to one unit of technology shock and the second panel shows the innovation response of $\Delta \ln Y$ corresponding to the same technology shock.

Since a technology shock has no permanent effect on productivity and no technology progress is assumed, a positive technology shock only boosts the productivity in current and the consecutive periods. It only reduces the inflation rate in the short run as panel 1 shows. On the other hand, a positive technology also has a beneficial effect on output growth in the current period, but a negative effect in the second period. Both (5.9) and (5.13) suggest that the technology shock has a negative effect on investment and thus reduces the level of capital stock in the next period. Therefore, in the second period the positive effect of the technology shock on productivity is dominated by its negative effect on capital stock. The combination of the two effects almost wears off after three periods as shown in panel 2.

The third and fourth panels shows the innovation response of $\Delta \ln p$ and $\Delta \ln Y$ corresponding to one unit of monetary shock. The third panel suggests that increasing money only causes inflation in the current and consecutively two periods. The effect dies out after three periods and the inflation rate goes back to the original level afterward. Panel 4 shows that an expansionary monetary also has positive effect on output growth in short run, but has no effect in the long run.

Although the two stochastic shocks have no permanent effects on inflation

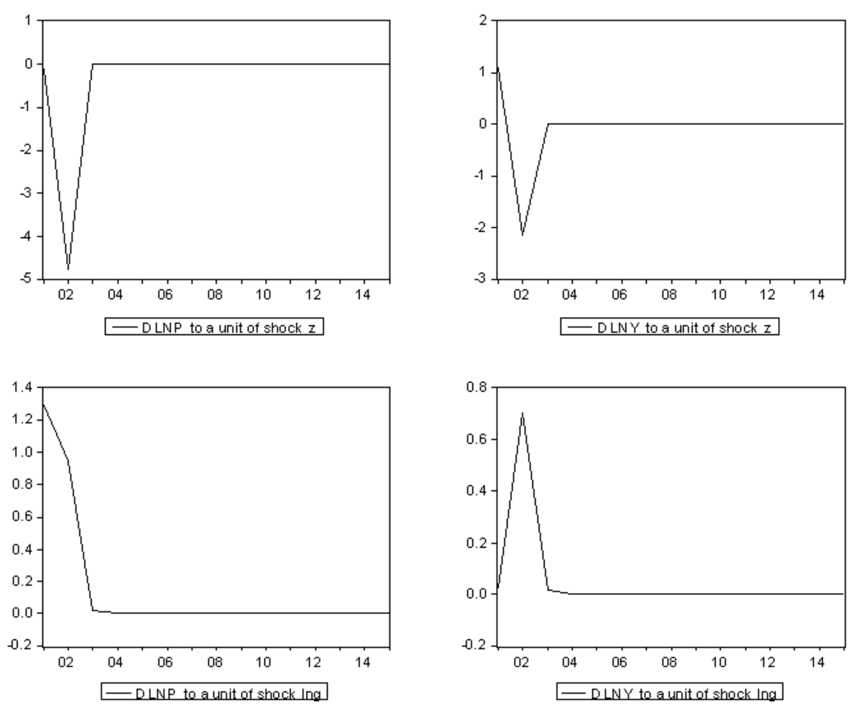

Figure 5.2: Innovation responses for the technology and monetary shocks. Panel 1 and 2 show the effects of one unit of technology shock on inflation rate and output growth rate. Panel 3 and 4 show the effects of one unit of monetary shock on inflation rate and output growth rate.

and output growth rate, they do have permanent effects on the levels of the variables through the change of capital stock. The innovation responses of the levels of the variables to the two stochastic shocks are shown in Figure 5.3 and Figure 5.4. Figure 5.3 shows the effects of a unit of technology shock on the levels of the variables. Panel 1 in Figure 5.3 shows that the price level $\ln p$ becomes lower after a positive technology shock and panel 6 shows that the capital stock also drops to a lower level. Since $\ln X$, $\ln H$, and $\ln Y$ are positive correlated with $\ln K$, as (5.9), (5.11), and (5.12) predicted, they all drop to lower levels after the shock. On the contrary, (5.10) shows that $\ln C$ positively depends on $\ln K$, so $\ln C$ jumps to a higher level after the shock.

The stochastic shock has no effect on the level of capital stock during the shock period, because $\ln K$ is already decided in the previous period before a shock happens. Panel 5 shows that $\ln Y$ has a peak during the shock period and drops to a lower level later because the positive effect of a technology shock is dominated by the effect of lower capital stock. On the other hand, a technology shock has significantly negative effect on investment, so $\ln X$ has a trough during the shock period and still stays at a lower level than that before the shock because of the lower level of $\ln K$.

The innovation responses of the levels of the variables to a unit of monetary shock are shown in Figure 5.4. Panel 1 in Figure 5.4 shows that the price level jumps to a higher level after a positive monetary shock and panel 6 shows that the capital

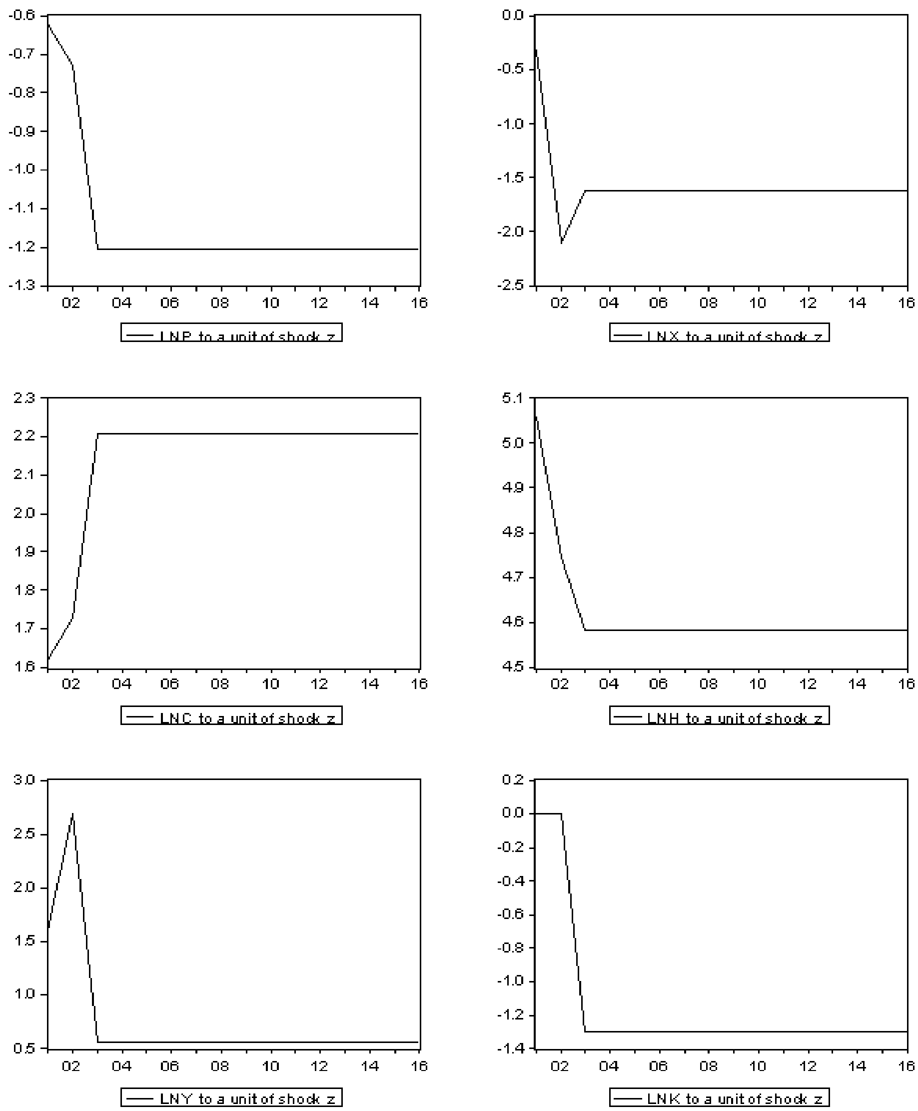

Figure 5.3: The innovation responses of the levels of the variables to one unit of technology shock for 15 periods. The unit shock happens in period 2.

stock also jumps to a higher level. Similarly, the new levels of $\ln X$, $\ln H$, and $\ln Y$ are all higher than that before the shock. $\ln C$ drops to a lower level after the shock. $\ln X$ has a peak during the shock period and still stays at a higher level after the effect of the shock wears off because of the higher level of $\ln K$.

In summary, our results suggest that an expansionary monetary policy has permanent positive effects on the level of investment, working hours, and a temporary effect on output growth. Its deficiency implied by (5.17) is that it induces inflation. Moreover, a beneficial technology shock decreases the levels of investment and working hours by increasing the productivity during the shock period, but the level of output stays at a lower level after the shock period because of the lower capital stock.

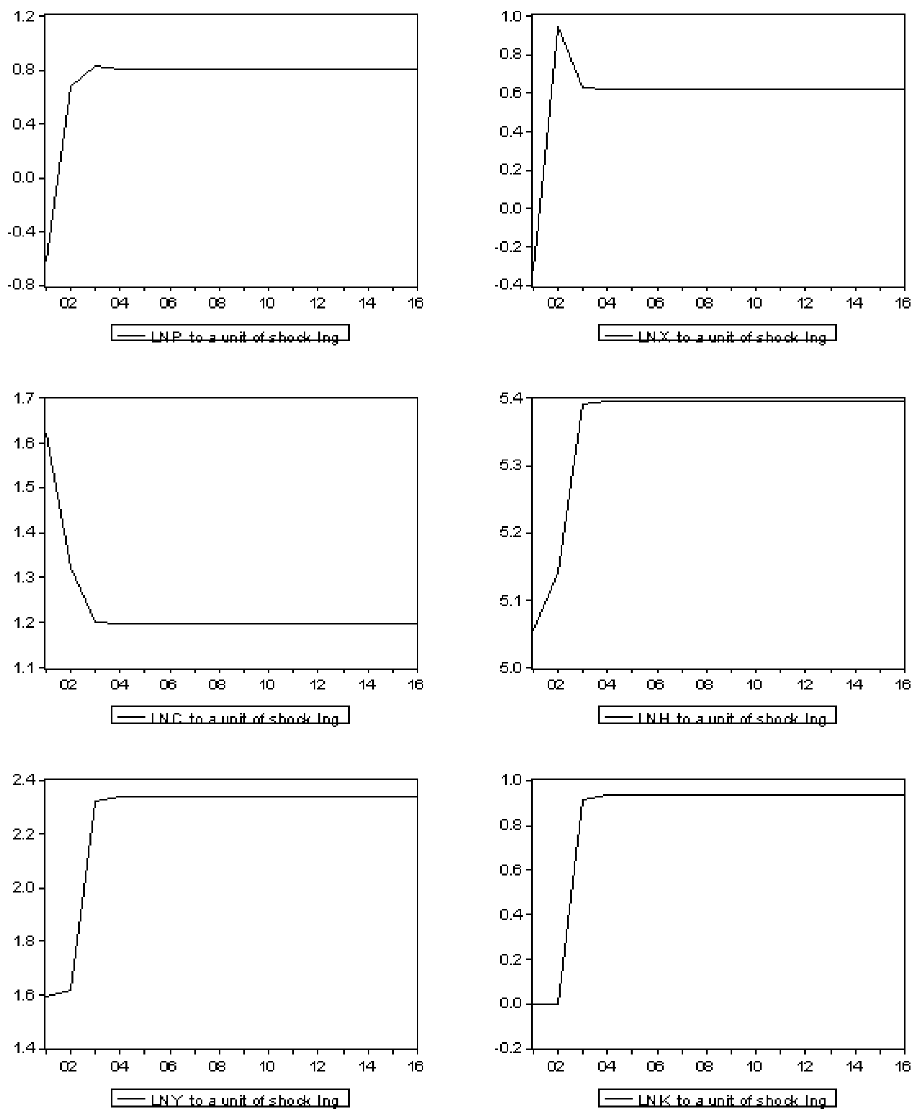

Figure 5.4: The innovation responses of the levels of the variables to one unit of monetary shock for 15 periods. The unit shock happens in period 2.

Chapter 6

Filters

This chapter discusses the effects of the Hodrick-Prescott and rational square-wave filters on macroeconmic analysis. Section 6.1 introduces some notation and concepts from spectral analysis and filtering. Section 6.2 and 6.3 introduce the H-P filter and the rational square-wave filters. Section 6.4 examines the distortional effects of these two filters. Section 6.5 compares the effects of the filters.

6.1 Linear filters

In most macroeconomic research, filters are use to achieve two main objectives. One is to extract a component, such as cyclical or seasonal components. The other is to transform a nonstationary series into a stationary one. Here, we focus on the first motivation. In order to examine the effect of filters, we start with the introduction of some notation and concepts in spectral analysis and linear filters.

Let y_t be a stationary series and γ_j be the jth autocovariance, the autocovariance generating function of y is then given by

$$a_y(z) = \sum_{j=-\infty}^{\infty} \tau_j z^j, \tag{6.1}$$

where z is complex scalar. Dividing (6.1) by 2π and replacing z by $e^{-i\omega}$, we obtain the power spectrum (or the power spectral density function) of y_t

$$f_y(\omega) = \frac{1}{2\pi} a_y(e^{-i\omega}),$$

where $-\pi < \omega < \pi$ is the angular frequency measured in radians and $i = \sqrt{-1}$. The area under the spectrum over the interval $[-\pi, \pi]$ represents the variance of y_t. On the other hand, given the spectrum of y the autocovariance can also be obtained by

$$\tau_j = \int_0 \pi 2 f_y(\omega) e^{i\omega j} d\omega. \tag{6.2}$$

A linear filter $F(L)$ decomposes a time series into cyclical and growth components,

$$\begin{aligned} y_t &= g_t + c_t, \\ &= [1 - F(L)]\, y_t + F(L) y_t, \end{aligned}$$

where $g_t = [1 - F(L)]\, y_t$ is the growth component and $c_t = F(L)y_t$ is the cyclical component. It implies that the power transfer function is

$$f_c(\omega) = |F(e^{-i\omega})|^2 f_y(\omega),$$

where $F(e^{-i\omega})$ is the frequency response function (or transfer function) and $|F(e^{-i\omega})|^2$ is the power transfer function, which multiplies the spectrum of the input series y at frequency ω and generates the output series at the same frequency.

The effect of a linear filter can be decomposed into the gain effect and phase effect. The gain effect of a filter is the change of the amplitudes of a signal and the phase effect is the advance or delay of the signal in time. Both the effect can jointly be refereed as the frequency response of the filter. By the polar decomposition, the frequency response function becomes

$$F(e^{-i\omega}) = G(\omega)e^{-iP(\omega)},$$

where $|G(\omega)|$ is called the gain of the filter, $|G(\omega)|^2$ is call the square gain of the filter, and $P(\omega)$ is the phase. The phase shift is zero if a linear filter is symmetric, or $F(L) = F(-L)$.

A low pass filter is a filter which allows frequencies below the cutoff point to pass and impede the frequencies above the cutoff point. On the other way, a high pass filter allows frequencies higher than the cutoff frequency to pass and impede the frequencies below the cutoff point. For example, the ideal high-pass filter has a power transfer function which satisfies

$$|F(e^{-i\omega})|^2 = \begin{cases} 0 \text{ if } |\omega| < \omega_0, \\ 1 \text{ if } |\omega| \geqslant \omega_0, \end{cases} \tag{6.3}$$

where ω_0 is the cutoff frequency.

Furthermore, applying a linear filter to a stationary series usually changes the variance and autocovariances of the original series. The original series y_t has the autocovariance γ_j^* given in (6.2) which in general is not equal to the autocovariance of the filtered series $F(L)y_t$

$$\gamma_j^* = \int_0 \pi 2|F(e^{-i\omega})|^2 f_y(\omega) e^{i\omega j} d\omega, \tag{6.4}$$

Similarly, so is the cross-covariance of the filtered and unfiltered series.

6.2 The Hodrick-Prescott filter

In this section I will discuss the effect of the Hodrick-Prescott filter on business cycle research. The H-P filter is an approach of fitting a smooth curve through a set of points. It decomposes a time series y_t into a cyclical component c_t and a growth component g_t, such that $y_t = g_t + c_t$. More specifically, the H-P filter computes a trend $\{g_t\}_{t=1}^T$ by minimizing the sum of squared deviations of a time series from its trend $(y_t - g_t)^2$

subject to the constraint that the sum of the squared second differences is not too large, which can be written as

$$\min_{\{g_t\}_{t=1}^{T}} \sum_{t=1}^{T} (y_t - g_t)^2 + \lambda \sum_{t=2}^{T-1} \left[(g_{t+1} - g_t) - (g_t - g_{t-1})\right]^2, \ \lambda > 0,$$

where T is the sample size, and λ is the penalty parameter. The first term is a measure of 'goodness-of-fit' and the second term is a measure of 'degree-of-smoothness' which penalizes decelerations in the growth rate of the trend component. The penalty parameter λ controls the smoothness of the series g_t and also alters the trade-off between the goodness-of-fit and degree-of-smoothness. The larger the λ, the smoother the g_t. As λ approaches infinity, g_t approaches a linear trend.

According to Harvey and Jaeger (1993), the detrended observations c_t can be written as

$$\begin{aligned} c_t &= \left[\frac{(1-L)^2(1-L^{-2})^2}{1/\lambda + (1-L)^2(1-L^{-2})^2}\right] y_t, \qquad (6.5) \\ &= HP(L) y_t, \end{aligned}$$

where

$$HP(L) = \frac{(1-L)^2(1-L^{-1})^2}{1/\lambda + (1-L)^2(1-L^{-1})^2},$$

is a lag polynomial and L is the lag operator. Similarly, the growth component can be written as

$$g_t = [1 - HP(L)]\, y_t.$$

Both $HP(L)$ and $1 - HP(L)$ are the so called linear filters.

6.3 The rational square-wave filter

Pollock (2000) suggests rational square-wave filter, which is applied by passing forward and backward through the series. In other words, the same filter is used in both directions. Let y_t be the time series, then the filtering operation can be described by

$$\begin{aligned} \gamma(L)q_t &= \delta(L)y_t, \\ \gamma(L^{-1})c_t &= \delta(L^{-1})q_t, \end{aligned}$$

where L is the backwards-shift operator, L^{-1} is the forwards-shift operator, q_t is the intermediate output, and c_t is the final output generated by the filter. The stability condition of the filter should satisfies that the roots of $\gamma(z) = 0$ lie outside the unit circle. The above two filters can be combined as a symmetric two-side rational filter

$$R(L) = \frac{\delta(L^{-1})\delta(L)}{\gamma(L^{-1})\gamma(L)},$$

and it is shown by Pollock (2000) that the optimal high-pass filter takes the form

$$R(L) = \frac{1}{1 + \frac{1}{\lambda}\left[i(1+L)/(1-L)\right]^{2n}},$$

where λ decides the cut-off frequency and n is a positive integer. The rate of transition from the passband to the stopband of a filter can be increased by increasing the value of n. Decreasing the cut-off frequency also has similar effect in increasing the rate of transition. Therefore, the detrended observation of the cyclical component is

$$c_t = R(L)y_t.$$

6.4 The distortional effect of the H-P Filter

In this section we analyze the distortionary effect of the H-P filter by its frequency response function. According to (6.5), the H-P filtered y_t or the cyclical component c_t can be written as

$$\begin{aligned} c_t &= HP(L)y_t, \qquad (6.6) \\ &= HP_1(L)(1-L)y_t, \end{aligned}$$

where

$$HP_1(L) = \frac{(1-L)(1-L^{-2})^2}{1/\lambda + (1-L)^2(1-L^{-2})^2}$$

is a lag polynomial. If y_t is stationary, then, according to Cramer representation, when applying the H-P filter to y_t the spectrum for the input and output series the following equality should hold

$$f_c(\omega) = |HP(\omega)|^2 f_y(\omega),$$

The frequency response functions of the H-P filter and the rational square-wave filter are shown in Figure 6.1. Both filters are close approximations of the ideal filter, but the ideal filter cuts off low frequencies sharper as (6.3) suggested. The frequency response functions smoothly move from 0 to 1 in the neighborhood of the cutoff frequency. Therefore, both filters pass some frequencies which it was supposed to suppress and also remove part of the frequencies which it was supposed to pass around the cutoff frequency. Figure 6.1 shows that the rational square-wave cuts off low frequencies relatively sharper than the H-P filter does.

Cogley and Nason (1995) and Park (1996) pointed out that when the H-P filter is applied to an $I(0)$ series, $HP(L)$ is a symmetric moving average filter that operates like a high pass filter as intended. But, when applied to an $I(1)$ process, $HP_1(L)$ becomes an asymmetric moving average filter that amplifies certain frequency components and dampens all the other high and low frequency components, which creates the spurious cycles. Similar argument applies to Pollock's rational square-wave filter.

One obvious example is provided by Figure 6.2, which shows the unfiltered and filtered series of $\ln H$ and $\ln Y$. The first panel in Figure 6.2 shows the two standardized series. There is no obvious cycles exist between these two series. The second panel shows that there exists spurious cycles between the same two series in the H-P filtered data.

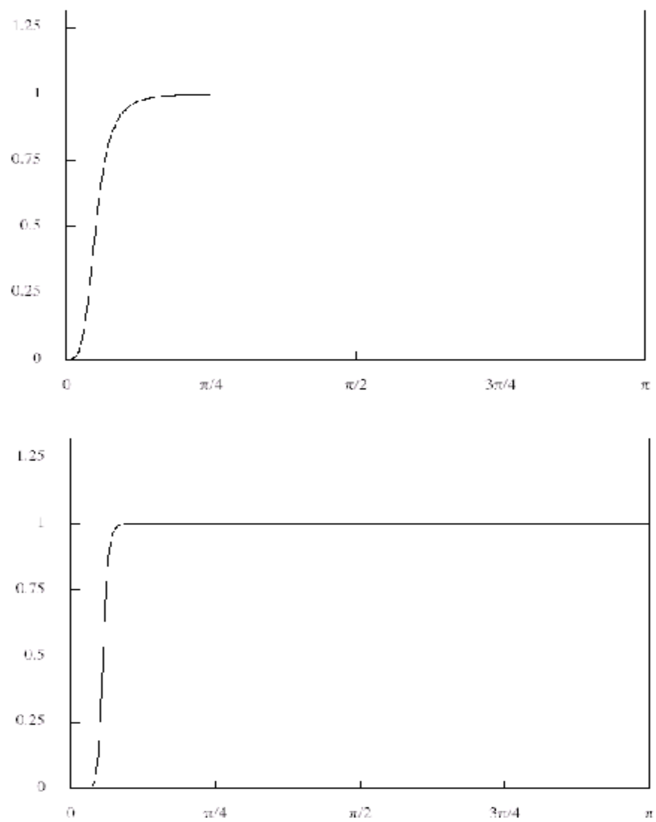

Figure 6.1: The frequency responses of the H-P filter and the rational square-wave filter. The upper graph shows the frequency response of the H-P filter and the lower graph shows that of the rational square-wave filter.

The third shows the same result in the rational square-wave filtered data.

6.5 Comparison of the rational square-wave filter and the H-P filter

The graphs of the H-P and the rational square-wave filtered data are shown in figure 6.3 and 6.4, respectively. The tests for unit roots for all filtered series are rejected. Both filters have similar effects on all series, except $\ln p$. Table 5.2 also suggests the same results. The CCF are similar in both filtered series. Notably, the cross-correlations between $\ln p$ and $\ln Y$ have dramatically differences. In the H-P filtered data, the correlation of $\ln p$ with $\ln Y$ is -0.2177 while it is 0.0342 in the rational square-wave filtered data.

Periodicity and comovement between the variables are the main focuses in business cycle research. In the case of calibration, periodicity is measured by the autocorrelation function (ACF) and comovements are summarized by cross-correlation functions (CCF) between the variables. The ACF and CCF of the H-P filtered and rational square-wave filtered data are shown in Table 5.1 and Table 5.2.

The autocorrelation of the original and unfiltered series are shown in Table 5.1. It is clear that both filters alter the moments of the series. All ACF values in the

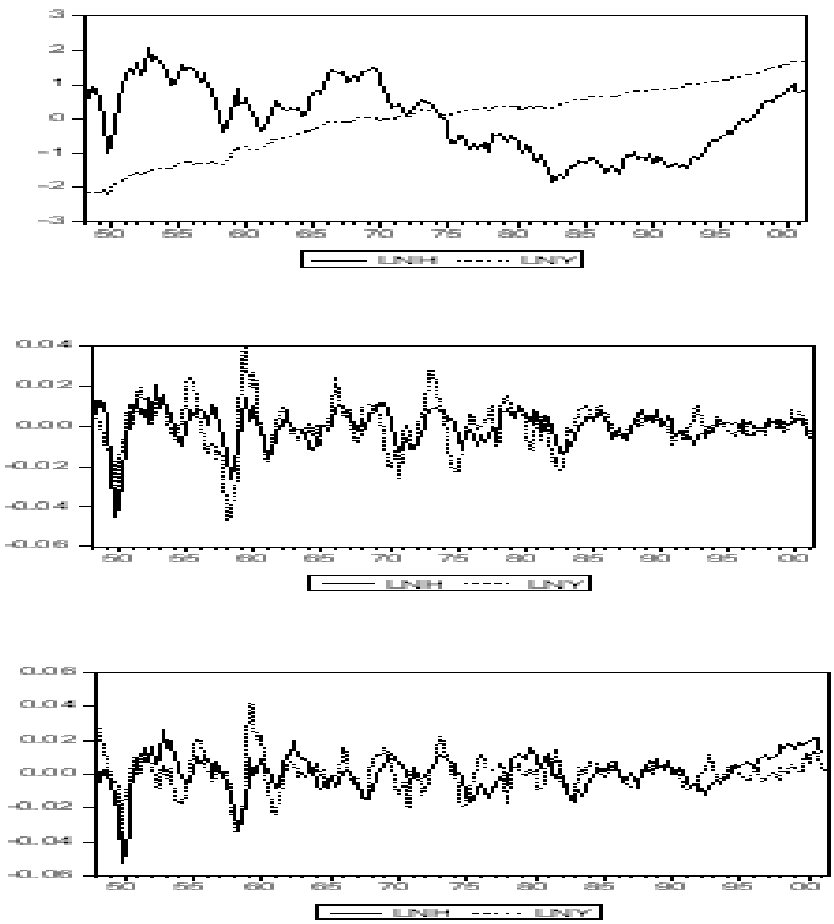

Figure 6.2: The spurious cycles between lnY and lnH. The first graph shows the standardized original series, the second graph shows the H-P filteres series, and the third graph shows the rational square-wave filtered series.

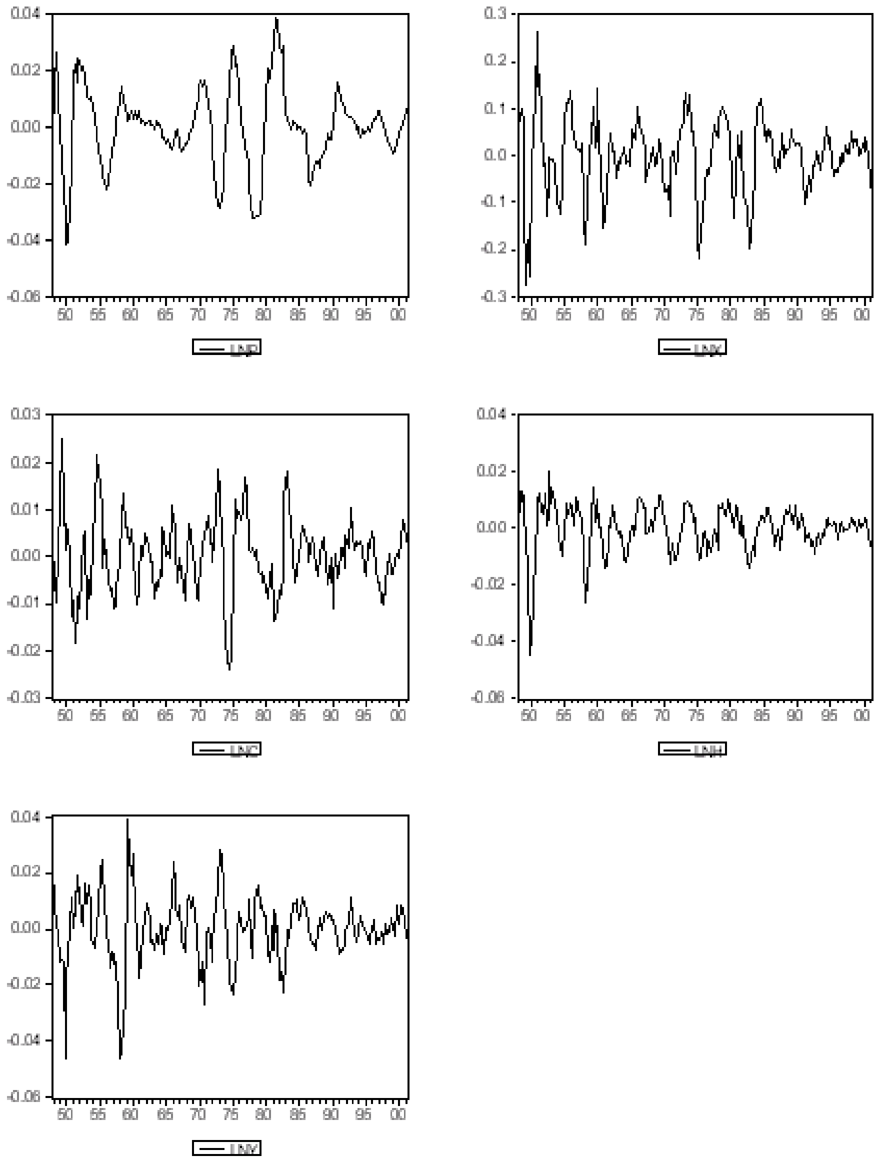

Figure 6.3: The five series filteres by the H-P filter

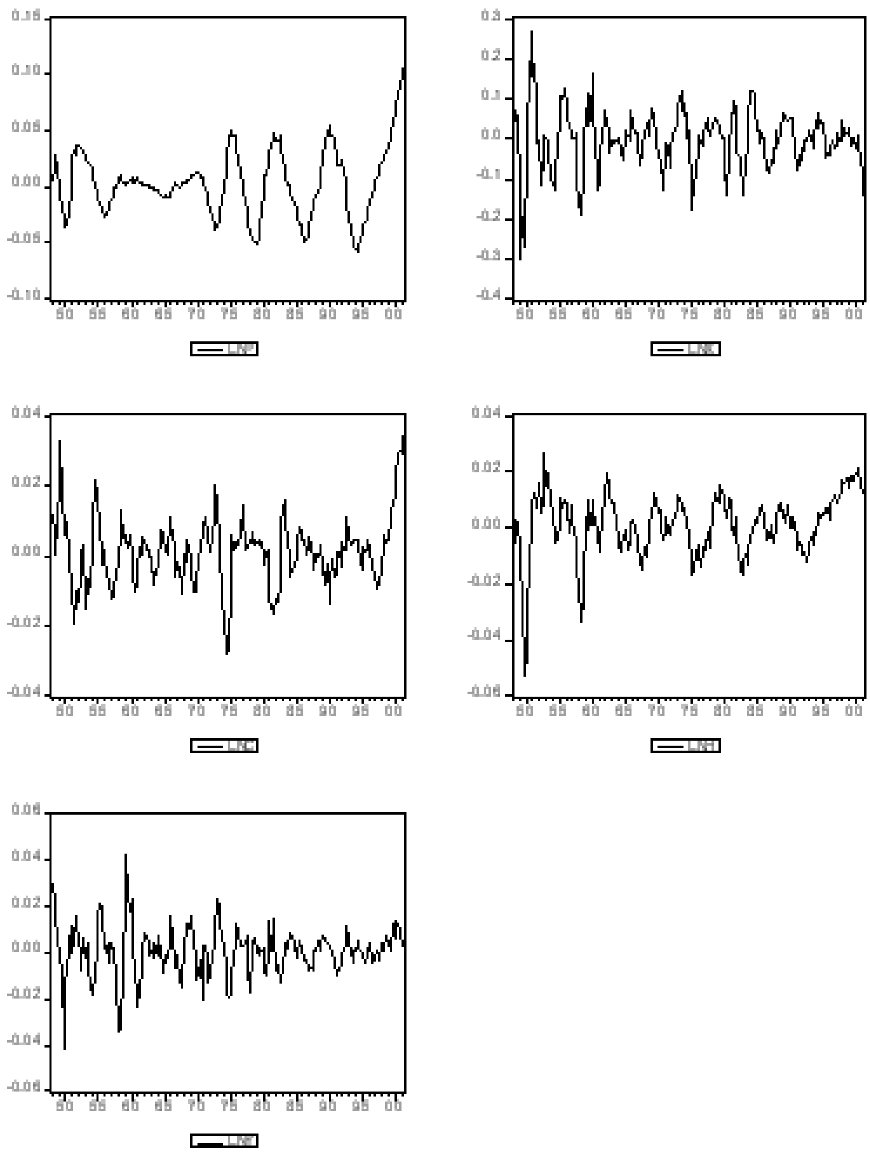

Figure 6.4: The five series filtered by Pollock's rational square-wave filter

filtered series become lower than that in the original data. In long periods, the ACF values have negative values while they are all positive in the original series.

The cross-correlation of the original and unfiltered series are shown in Table 5.2. As shown by (6.2) and (6.4) in Section 6.1, since the stochastic trend has been removed by the filters, the cross-correlation functions that measures the cyclical behavior are altered by both filters.

The effects of the filters can be further examined by applying the Bierens-Swanson average conditional reality bound. The estimation results by using the filtered series are shown in Table 5.3. It shows that the Bierens-Swanson reality bounds estimated by using both filtered data are lower than that estimated by using the original series. In other words, it implies that some useful information included in the stochastic trend has been removed by both the H-P filter and the rational square-wave filter.

However, in this application we have to point out that in the Bierens and Swanson's approach the theoretical density is embedded in the empirical density, which requires that both the theoretical and empirical densities are estimated by the same data. Since we focus on the distortionary effect of the filters, we use the filtered data in the theoretical model and unfiltered data in empirical models as a comparison.

Moreover, the filters remove the interesting part of the theoretical model. As shown in Section 3.5, the theoretical model is equivalent to a cointegrated error

Table 6.1: The autocorrelation of the original and filtered series.

Variable	1	2	4	8	12
			The original series		
y	0.981	0.962	0.924	0.847	0.778
p	0.990	0.980	0.959	0.916	0.870
x	0.973	0.944	0.876	0.771	0.706
c	0.983	0.965	0.931	0.867	0.804
h	0.972	0.939	0.866	0.777	0.709
			The H-P filtered data		
y	0.728	0.457	-0.080	-0.233	-0.179
p	0.926	0.778	0.374	-0.374	-0.504
x	0.751	0.474	-0.146	-0.273	-0.177
c	0.746	0.445	-0.056	-0.321	-0.174
h	0.746	0.468	-0.065	-0.325	-0.120
			The rational square-wave filtered data		
y	0.648	0.313	-0.304	-0.226	0.032
p	0.941	0.845	0.585	-0.002	-0.409
x	0.717	0.409	-0.256	-0.305	-0.073
c	0.799	0.568	0.148	-0.221	-0.225
h	0.854	0.677	0.314	-0.058	-0.146

Table 6.2: The cross-correlation of the original and filtered series.

Cross-correlations with y_{t-j}, $j = 0, \pm1, \pm2, \pm4, \pm8, \pm12$											
					Moments of the original series						
Variable	-12	-8	-4	-2	-1	0	1	2	4	8	12
y	0.778	0.847	0.924	0.962	0.981	1	0.981	0.962	0.924	0.847	0.778
p	0.7947	0.8335	0.8735	0.8929	0.9025	0.9124	0.8963	0.8802	0.8480	0.7832	0.7220
x	0.7807	0.8216	0.9001	0.9249	0.9363	0.9483	0.9265	0.9037	0.8537	0.7637	0.6809
c	0.8006	0.8644	0.9272	0.9611	0.9780	0.9950	0.9970	0.9590	0.9221	0.8475	0.7802
h	-0.5368	-0.5175	-0.4908	-0.4959	-0.4972	-0.5053	-0.5138	-0.5223	-0.5394	-0.5574	-0.5681
					Moments of the H-P filtered data						
y	-0.179	-0.233	-0.080	0.457	0.728	1	0.728	0.457	-0.080	-0.233	-0.179
p	0.1162	0.0461	-0.0329	-0.0924	-0.1547	-0.2177	-0.2407	-0.2400	-0.1407	0.1519	0.1821
x	-0.1017	-0.2355	-0.0588	0.3585	0.5384	0.6878	0.5037	0.3173	-0.1552	-0.1794	-0.1287
c	0.0956	-0.0145	-0.3045	-0.2577	-0.1351	0.0933	0.2438	0.3317	0.3108	-0.0651	-0.1189
h	-0.1159	-0.1218	0.2230	0.5316	0.6231	0.6417	0.3533	0.0516	-0.3725	-0.2596	-0.0023
					Moments of the rational square-wave filtered data						
y	0.032	-0.226	-0.304	0.313	0.648	1	0.648	0.313	-0.304	-0.226	0.032
p	0.0377	0.0087	0.0539	0.0814	0.0622	0.0342	0.0186	0.0071	0.0142	0.0637	0.0366
x	0.0269	-0.2143	-0.2191	0.2590	0.4767	0.6621	0.4597	0.2586	-0.2536	-0.1615	-0.0314
c	0.0847	0.0853	-0.1237	-0.1167	-0.0364	0.1434	0.2372	0.2814	0.2177	-0.1086	-0.0785
h	-0.0560	-0.2125	0.0342	0.3423	0.4414	0.4874	0.2785	0.0556	-0.2470	-0.0890	0.1542

Table 6.3: The estimated results of the Bierens-Swanson average conditional reality bound by using the H-P and rational square-wave filtered data. $\alpha^* = \Pi_2^{'}\alpha$, $\nu* = \Pi_2^{'}\nu$.

	The H-P filter		The rational filter	
α^*	-0.042027	-0.098711	-0.040869	-0.067609
	0.240875	-0.049976	-0.006381	-0.130503
	-0.005697	-0.0365778	-0.082767	-0.041228
	-0.052834	0.048628	-0.197311	0.151988
β	0.025507	1	-0.454957	1
	1	0.185909	1	-0.340934
	-0.523602	0.090135	0.076582	0.100724
ν_0^*	0.004897	–	0.004954	–
	0.00496	–	0.004991	–
c_1	0.641691	–	0.632687	–
c_2	0.544205	–	0.500316	–
p_0	0.355856		0.325611	

correction model. With the unit roots removed by the filters, the cointegrating relations are removed as well.

Chapter 7

Summary and conclusion

The goal of this dissertation was to illustrate how the standard cash-in-advance model can be adapted to study the role of monetary and technology shocks in business cycles, and the analysis presented can also be extended to a variety of dynamic stochastic equilibrium models. I derived a linearized cash-in-advance model that provides a link between the deep parameters and model coefficients. Through the linearized model I then assessed the empirical performance of the theoretical model and estimated the deep parameters. I also obtained confidence intervals for the estimated parameters using the bootstrap and the effects of the stochastic shocks are further analyzed using the innovation response analysis. Finally, I also examined the distortional effects of the Hodrick-Prescott and rational square-wave filters.

My findings can be summarized as follows. First, the business cycle model under review provides an explanation for some, but not all, observable macroeconomic fluctuation. If the possible model misspecification or missing variables in the CIA model are not penalized, the model can account for about 47% of the information contained in the empirical model. Moreover, most of the penalty is due to the lag difference between these two models and the rest can be accounted by their rank difference. The theoretical model implies a cointegrated error correction with one lag and two cointegrating relations, and the empirical model representing the data-generating process is an error correction model with two lags and only one cointegrating relation.

My second finding is that the cash-in-advance model does not do well in explaining the observed data if model misspecification is allowed for. For instance, the estimated capital depreciation rate and capital share are much higher than the usual calibrated values. These high values imply that some important features of the economy are missing from the model. The question as to how to build a better model should be further explored.

Third, the innovation response analysis suggests that an expansionary monetary policy has a positive effect on economic growth but the effect almost wears off after three periods. The main deficiency is that it also causes inflation. Moreover, it has permanent positive effects on the level of investment, working hours, and output and negative effects on the level of consumption because of the permanent change in the level of capital stock. In contrast, a positive technology shock decreases the levels of investment and working hours by increasing productivity during the shock period, but the level of output is lower after the shock period because of the lower capital stock. Positive technology shocks have a positive effect on consumption.

Finally, in addition to generating spurious cycles, both the Hodrick-Prescott and rational square-wave filters remove the unit roots of the series. I showed that the theoretical model is equivalent to a cointegrated error correction model. With the unit roots removed by the filters, important features of the model economy implied by the cointegrating relations are removed as well. The interesting part of the theoretical model, contained in the stochastic trend, no longer exists in the filtered series. Researchers should be careful when applying filters because they can influence the model predictions substantially.

This application is based on the assumption that the estimated parameters have kept the same values throughout the sample period from 1948 to 2001. This analysis is not suitable for studying and predicting how macroeconomic aggregates reacts towards the structural changes. However, the stability of the parameters can be further examined by decomposing the sample data into subsamples.

The main limitation of the econometric approaches used in this study is that since some links between the model coefficients and the deep parameters, such as the discount rate in the utility function, are lost in the process of linearization, only a subset of the deep parameters can be estimated. However, since the estimated coefficients of the decision rules are in fact complicated functions of the deep parameters in the utility function, in principle they can be solved using those functions. Therefore, further analysis that requires the full specification of the utility function, such as the analysis of welfare cost of inflation, can be done in future research.

Bibliography

[1] Barrow, R., (1996), "Determinants of Economic growth: a cross-country empirical study," National Bureau of Economics Research, working paper No. 5698.

[2] Bierens, H.J. (1994), *Topics in Advanced Econometrics*, Cambridge University Press.

[3] Bierens, H.J. (2002), *EasyReg International*, Department of Economics, Pennsylvania State University (free econometric software downloadable from URL http://econ.la.psu.edu/~hbierens/EASYREH.HTM)

[4] Bierens, H.J. (2003), *Econometric Analysis of Singular Macro-Economic Models, with an Application to the King-Plosser-Rebelo Stochastic Growth Model,* Department of Economics, Penn State University and Tilburg University, working paper.

[5] Bierens, H.J. and N.R. Swanson (2000), "The Econometric Consequences of the

Ceteris Paribus Condition in Economic Theory," *Journal of Econometrics*, 95:223-253.

[6] Chow, G.C. and Y.K. Kwan (1998), "How the Basic RBC Model Fails to Explain US Time Series," *Journal of Monetary Economics*, 41:301-318.

[7] Christiano, L.J. (1990), "Solving the Stochastic Growth Model by Linear-Quadratic Approximation and by Value-Function Iteration," *Journal of Business and Statistics*, 8:23-26.

[8] Christiano, L.J. and M. Eichenbaum (1992), "Current Business Cycle Theories and Aggregate Labor Market Fluctuations," *American Economic Review*, 82:430-50.

[9] Cogley, T., and J.M. Nason (1995), "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series," *Journal of Economic and Dynamic and Control*, 19: 253-278.

[10] Cooley, T.F. and G.D. Hansen (1989), "The Inflation Tax in a Real Business Cycle," *American Economic Review*, 79:733-748.

[11] Cooley, T.F. and G.D. Hansen (1992), "Tax Distortions in a Neoclassical Monetary Economy," *Journal of Economic Theory*, 58:290-316.

[12] Cooley, T.F. and G.D. Hansen (1995), "Money and the Business Cycle," in T.F. Cooley (ed.), *Frontiers of Business Research*, Princeton University Press, Princeton, NJ.

[13] Cooley, T.F. and E.C. Prescott (1995), "Economic Growth and Business Cycles," in T.F. Cooley (ed.), *Frontiers of Business Research*, Princeton University Press, Princeton, NJ.

[14] DeJong, D.N., B.F. Ingram, and C.H. Whiteman (1996), "A Bayesian Approach to Calibration, " *Journal of Business and Statistics*, 14:1-9.

[15] DeJong, D.N., B.F. Ingram, and C.H. Whiteman (2000), "A Bayesian Approach to Dynamic Macroeconomics, " *Journal of Econometrics*, 98:203-223.

[16] Dotsey, M. and Sarte P.D. (2000), "Inflation uncertainty and growth in a cash-in-advance economy," Journal of Monetary Economics, 45:631-655.

[17] Fuller, W.A. (1996), *Introduction to Statistical Time Series*, New York: John Wiley.

[18] Gregory A.W. and G.W. Smith (1991), "Calibration as Testing: Inference in Simulated Macroeconomics Models," *Journal of Business and Economic Statistics*, 9(3): 297-303.

[19] Geweke, J. (1999), “Computational Experiments and Reality,” Department of Economics, University of Minnesota and Federal Reserve Bank of Minneapolis working paper.

[20] Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, NJ.

[21] Hansen, G.D. (1985), “Indivisible Labor and the Business Cycle,” *Journal of Monetary Economics*, 16:309-327.

[22] Hansen, L.P. and J.J. Heckman (1996), “The Empirical Foundations of Calibration,” *Journal of Economic Perspectives*, 10:87-104.

[23] Harvey, A.C. and A. Jaeger (1993), “Detrending, Stylized Facts and the Business Cycle,” *Journal of Applied Econometrics*, 8: 231-247.

[24] Haavelmo (1944), “ The Probability approach in econometrics,” *Econometrica,* 12: 1-115.

[25] Hordick, R.J. and E.C. Prescott (1997), “Postwar U.S. Business Cycles: An Empirical Investigation,” Journal of Money, Credit, and Banking, 29:1-16. Carnegie-mellon working paper, 1980.

[26] Hodrick, Kocherlakota, and Lucas (1991), "The Variability of Velocity in Cash-in-Advance Models," Journal of Political Economy, 99(2):358-384.

[27] Ireland, P. (1994), "Money and Growth," *The American Economic Review,* 84(1): 47-65.

[28] Johansen, S. (1995), *Likelihood-Based Inference in Cointegration Vector Autoregressive Models*, Oxford University Press.

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

[30] Johansen, S. and K. Juselius (1988), "Maximum Likelihood Estimation and Inference on Cointegration with Application to the Demand of Money," *Oxford Bulletin of Economics and Statistics*, 52:109-210.

[31] King, R.G., C.I. Plosser, and S.T. Rebelo (2001), "Production, Growth and Business Cycles: Technical Appendix," *Computational Economics.*

[32] King, R.G., C.I. Plosser, and S.T. Rebelo (1988), "Production, Growth, and Business Cycles: I. The Basic Neoclassical Model," *Journal of Monetary Economics*, 21:195-232.

[33] King, R.G., C.I. Plosser, and S.T. Rebelo (1988), "Production, Growth, and Business Cycles: II. New Directions," *Journal of Monetary Economics*, 21:309-341.

[34] Kozicki, S. (1999), "Multivariate Detrending Under Common Trend Restrictions: Implications for Business Cycle Research," *Journal of Economic Dynamics and Control*, 23: 997-1028.

[35] Kydland, F.E., and E.C. Prescott (1982), "Time to Build and Aggregate Fluctuations," *Econometrica*, 50:1345-1370.

[36] Kydland, F.E., and E.C. Prescott (1996), "The Computational Experiment: An Econometric Tool," *Journal of Economic Perspectives*, 10:69-85.

[37] Labadie, P. (1989), "Stochastic Inflation and the Equity Premium," *Journal of Monetary Economics,* 24:277-298.

[38] Lucas, R.E. (1984), "Money in a Theory of Finance," *Carnegie-Rochester Conference Series Public Policy* 21:9-45.

[39] Lucas, R.E. (1987), *Model of Business Cycles*, New York: Basil Blackwell.

[40] Lucas, R.E. and N.L. Stocky (1987), "Money and Interest in a Cash-In-Advance Economy," *Econometrica*, 55:491-513.

[41] Lutkepohl, H. (1993), *Introduction to Multiple Time Series Analysis*, Berlin: Springer-Verlag.

[42] Malinvaud, E. (2000), *Macroeconomic Theory, Volume C: Inflation, Employment and Business Fluctuations,* North-Holland.

[43] Mehra, R., and E.C. Prescott (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics,* 15:145-161.

[44] McGrattan, E.R. (1990), "Solving the Stochastic Growth Model by Linear-Quadratic Approximation," *Journal of Business and Statistics*, 8:41-44.

[45] Park, G. (1996), "The Role of Detrending Methods in a Model of Real Business Cycle," *Journal of Macroeconomics,* 18:479-501.

[46] Pedersen, T.M. (2001), "The Hodrick-Prescott Filter, the Slutzky Effect, and the Distortionary Effects of Filters," *Journal of Economic Dynamics and Control*, 25: 1081-1101.

[47] Pollock, D.S.G. (1999), *A Handbook of Time-Series Analysis, Signal Processing and Dynamics*, California: Academic Press.

[48] Pollock, D.S.G. (2000), "Trend Estimation and De-trending via Rational Square-Wave Filters," *Journal of Econometrics*, 99: 317-334.

[49] Rietz, T.A. (1988), "The Equity Premium: A Solution," *Journal of Monetary Economics*, 22:17-33.

[50] Romer, David (1996), *Advanced Macroeconomics*, New York: McGraw-Hill.

[51] Schott, J.R. (1997), *Matrix Analysis for Statistics*, New York: John Wiley.

[52] Sidrauski, M. (1967), "Rational Choice and Patterns of Growth in a Monetary Economy," *The American Economic Review*, 57(2): 534-544.

[53] Sill, K. (1998), "An Empirical Investigation of Money Demand: Evidence from a Cash-In-Advance Model," *The Canadian Journal of Economics*, 31(1)125-147.

[54] Stocky, N.L. and R.E. Lucas (1989), *Recursive Methods in Economic Dynamics*, Harvard University Press.

[55] Taylor, J.B. and H. Uhlig (1990), "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods," *Journal of Business and Statistics*, 8:1-17.

[56] Tobin, J. (1965), "Money and Economic Growth," 33(4): 671-684.

[57] Watson, M.W. (1993), "Measure of Fit for Calibrated Models," *Journal of Political Economy*, 101:1011-1041.

[58] Wang, P. and C.K. Yip (1992), "Alternative Approaches to Money and Growth," *Journal of Money, Credit, and Banking*, 24(4): 553-562.

Hung-Pin Lai

Address: 425 Waupelani Dr. Apt. 304
State College, PA 16801
USA

Home: (814) 235-8251
Office: (814) 865-1108
Email: hplai@psu.edu

Curriculum Vitae

CITIZENSHIP:
- Republic of China

EDUCATION:
- Ph.D., Economics, The Pennsylvania State University, University Park, PA, August 2003
- M.A., Economics, National Taiwan University, May 1996
- B.A., Agricultural Economics, National Taiwan University, May 1994

Ph.D. THESIS:
- Econometric Analysis of a Cash-In-Advance Model
  Thesis Advisor: Herman J. Bierens

FIELDS:
- Econometrics
- Macroeconomic Theory

TEACHING EXPERIENCE:
- General Education Teaching Assistant, Introductory Microeconomics, Fall 2001- Spring 2003
- Teaching Assistant: Introductory Microeconomics Summer 2002
- Teaching Assistant: Introductory Macroeconomics Fall 2000 – Spring 2001
- Teaching Assistant: Public Finance, Spring 2000
- Teaching Assistant: Intermediate Macroeconomics, Fall 1999
- Teaching Assistant: Intermediate Econometrics, Fall 1998, Spring 1999, Summer 1999
- Teaching Assistant: Statistics, Fall 1997, Spring 1998

RESEARCH EXPERIENCE:
- Research Assistant, Hermen Bierens, Summer 2000-2002
- Research Assistant, Institute of Economics, Academia Sinica, July 1996- June 1997