Econometric Analysis of a Cash-In-Advance Model

Herman J. Bierens* and Hung-Pin Lai†

January 2007

Abstract

In this paper we analyze and estimate a linearized version of the cash-in-advance (CIA) model of Cooley and Hansen (1989) using the approach of Bierens and Swanson (2000) and Bierens (2006). In this approach the parameters of the CIA model are linked to the parameters of a corresponding empirical model representing the data generating process, rather than by fitting the CIA model directly to the data. We show that the linearized version of the CIA model takes the form of a vector error correction model (VECM) with two cointegrating vectors and singular error variance matrix. The empirical counterpart is a VECM with one cointegrated vector. Using the latter model, the estimated parameters of the CIA-VECM are endowed with 90% bootstrap confidence intervals. The effects of the monetary and technology shocks are examined by innovation response analysis.

Keywords: Dynamic general equilibrium, singularity, cash-in-advance

JEL classifications: C13, C32, C52

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models specify explicitly the dynamic decision problems of households and firms and their interde-

*Department of Economics, Pennsylvania State University, University Park, PA 16802, USA. Email: hbierens@psu.edu. Fax: 814-863-4775 Tel.: 814-865-4921

†Corresponding author. Department of Economics, National Chung Cheng University, Taiwan. Email: ecdhpl@ccu.edu.tw. This paper is based on Hung-Pin Lai's Ph.D. thesis at Penn State University under the supervision of Herman Bierens.

pendence. They form the leading paradigm in theoretical macroeconomic analysis. However, as these models become increasingly complex and stylized it is difficult to estimate DSGE models econometrically. Therefore, most macroeconomists calibrate their models, by simulating the decisions of the economic agents given the model parameters, and trim these parameters to replicate the pattern of co-movement among key-macroeconomic variables (the stylized facts). Although the procedure is useful for understanding the dynamic properties of DSGE models, it has been criticized for lacking a formal statistical foundation (see DeJong et al. 1996 and 2000).

Recently some studies have attempted to estimate DSGE models by using econometric methods. For instance, since calibration does not involve realistic specifications of uncertainty regarding the deep parameters, Dejong et al. (1996 and 2000) and Geweke (1999) suggest a Bayesian approach. They assume prior probability distributions centered at the calibrated values for the uncertainty of the parameters and then compute the posterior distributions implied by the model and data. Christiano and Eichenbaum (1992) and Burnside et al. (1993) propose to use generalized method of moments (GMM) to estimate the deep parameters based on the first order conditions of linearized DSGE models. McGrattan et. al (1997) and Ireland (2004) use maximum likelihood (ML) to estimated a linearized DSGE model. Since the model variables have a singular distribution, they add some measurement error to the model.

DSGE models are designed to gain insight in particular economic issues, and to explain some, but not all, features of key macroeconomic variables. Therefore, these models do not represent the data generating process (DGP). However, the Bayesian, GMM, and ML approaches rely on the assumption that the theoretical model represents the DGP, which is too strong an assumption. Moreover, usually these models are driven by only a few stochastic shock processes, so that their implied conditional distribution is singular. Therefore, empirical applications of the Bayesian or ML analysis require either focusing on a subset of the macroeconomic variables or adding additional errors either implicitly or explicitly to overcome the singularity problem. However, adding additional errors has more or less influences on the analysis of DSGE models or makes changes to the setting of the theoretical DSGE model, which are in general being ignored.

Lately Del Negro and Schorfheide (2004) and Del Negro et al. (2006) attempt to abandon the assumption that the DSGE represents the DGP by considering a Bayesian mixture model of unrestricted VAR and DSGE, where the model is constructed by augmenting the actual data with the calibrated data. The dependence of the VAR coefficient matrices on the deep parameters is controlled by a hyperparameter (i.e. the ratio of the two types of data) which lies between 0 and ∞, where the two boundary values correspond to the VAR and DSGE models, respectively. The marginal likelihood of the hyperparameter is then used to provide an overall assessment of the DSGE model restrictions; however, whether the hyperparameter is a proper index and can be used for comparison of the DSGE models is not so intuitive since its upper bound is unbounded.

Rather than linking the model directly to the data, Bierens and Swanson (2000) and Bierens (2006) propose to link linearized DSGE models to non-structural VAR models via the so-called reality bound, which lies between 0 and 1. Bierens and Swanson (2000) suggest to compare the non-singular part of the conditional distribution of a linearized DSGE model with the corresponding conditional distribution of the VAR model, where the latter is treated as the DGP. Bierens (2006) follows similar lines and further proposes to augment both the linearized DSGE model and the empirical VAR model with the same non-degenerated errors in order to make the theoretical convoluted conditional distribution non-singular while keeping it comparable with the convoluted conditional distribution implied by the VAR model. Then the comparison of these two densities by the reality bound can be used to link the deep parameters of DSGE models to the parameters of the VAR model, and endow the estimated deep parameters with confidence intervals via the delta method. The benchmark model involved is the King, Plosser and Rebelo (1988) stochastic growth model, which is driven by a single shock process.

In this paper we extend this approach by considering a more complicated model driven by two shock processes, namely the cash-in-advance model of Cooley and Hansen (1989). We choose this model because it is a typical and general model used in monetary analysis, but has undergone relatively little empirical investigation. We will also show how to predict the joint stochastic properties of endogenous variables. Our extension is non-trivial, because the linearized version of this model turns out the be a cointegrated vector error correction model (VECM), and therefore the empirical model has to be a cointegrated VECM as well.

The rest of this paper is organized as follows. Section 2 introduces the Cooley and Hansen model and shows that the linearized model is equivalent to a cointegrated VECM. Section 3 introduces the conditional reality bound of Bierens and Swanson (2000) and Bierens (2006). Section 4 demonstrates how to apply the conditional reality bound to analyze the Cooley and Hansen model, including model evaluation, estimation, and innovation response analysis. Section 5 summarizes and concludes.

2 The Cooley and Hansen model

2.1 The Model

The key features of the Cooley and Hansen (1989) are the cash-in-advance constraint and the indivisible labor assumption. All agents in this economy maximize their lifetime utilities based on all information available at time t:

$$\max E_0 \left[\sum_{t=0}^{\infty} \bar{\beta}^t \ln c_t - B.h_{t,} \right] \tag{1}$$

subject to

$$p_t c_t \leq m_{t-1} + (g_t - 1) M_{t-1}, \tag{2}$$

$$c_t + x_t + m_t / p_t \leq w_t h_t + r_t k_t + (m_{t-1} + (g_t - 1) M_{t-1}) / p_t, \tag{3}$$

$$k_{t+1} = (1 - \delta) k_t + x_t, \tag{4}$$

$$z_{t+1} = \overline{\gamma} z_t + \epsilon_{t+1},\ \ \epsilon_t \sim i.i.d.N(0, \sigma_\epsilon^2),\ \text{with } 0 < \overline{\gamma} < 1, \tag{5}$$

$$\ln g_{t+1} = \bar{\alpha} \ln g_t + \xi_{t+1},\ \ \xi_t \sim i.i.d.N(\ln \bar{g}(1 - \bar{\alpha}), \sigma_\xi^2),\ \text{with } 0 < \overline{\alpha} < 1, \tag{6}$$

where E_t represents the conditional expectation operator given the information at time t, $\bar{\beta} \in (0, 1)$ is the discount rate, c_t denotes consumption, h_t is working hours, m_{t-1} is money holding at time t, k_t is capital stock, x_t is investment, p_t is the price level, w_t is the wage rate, r_t is the capital return, and $B > 0$ is a constant. Equation (2) represents the cash-in-advance constraint, (3) is the budget constraint and (4) is the law of motion of capital stock.

Moreover, the firms' production technology is assumed to be a constant return-to-scale function

$$Y_t = \exp(z_t) K_t^\rho H_t^{1-\rho}, \tag{7}$$

where $0 \leq \rho \leq 1$. The assumptions of the competitive market and the constant return-to-scale technology imply that firms make zero profit in equilibrium. The variables z_t and g_t are the stochastic technology and money supply shocks, respectively. Both are assumed to be stationary AR(1) processes. The variables M_t, K_t and H_t are aggregates of m_t, k_t and h_t, respectively, and Y_t is the total output in this economy.

Since 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 $V(z_t, g_t, \hat{m}_{t-1}, K_t, k_t)$. Indicating the next period value of a variables by a prime, the representative agent's dynamic problem can be written as

$$
\begin{aligned}
V(z, g, \hat{m}, K, k) = \ & \max_{c,h} \{u(c, h) \\
& +\bar{\beta} E[V(z', g', \hat{m}', K', k')|z, g, \hat{m}, K, k]\} && (8) \\
\text{s.t. } & z' = \bar{\gamma} z + \epsilon && (9) \\
& \ln g' = \bar{\alpha} \ln g + \xi && (10) \\
& c = (\hat{m} + g - 1)/(\hat{p} g) && (11) \\
& c + x + \hat{m}'/\hat{p} = w(z, K, H)h + r(z, K, H)k && (12) \\
& \qquad\qquad + (\hat{m} + g - 1)/(\hat{p} g) \\
& k' = (1 - \delta)k + x && (13) \\
& K' = (1 - \delta)K + X. && (14)
\end{aligned}
$$

The system equation (8)-(14) can be solved by a dynamic programming technique. 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 $V : s \rightarrow \mathbf{R}$ 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 value function V. Nevertheless, the decision rules can be solved by iterating the quadratic version of Bellman's equation.

2.2 The linearized version of the Cooley and Hansen Model

In order to compute the equilibrium of a DSGE model, the local dynamics around a steady state is approximated by a set of linear functions. Similar to the approaches suggested Kydland and Prescott (1982), Hansen and

Prescott (1995), and McGrattan (1990), a quadratic approximation of the objective function is formed by a Taylor expansion of the function at the deterministic steady-state values, and the corresponding solution is iterated until convergence occurs.

For the sake of convenience, we work with the logarithm of each variable. Denote variables in logs by a tilde, $\widetilde{A} = \ln A$, say, where $A = g$, p, x, X, k and K. Let $\tilde{s} = (1, z, \tilde{g}, \hat{m}, \tilde{K}, \tilde{k})^{\mathrm{T}}$ and $\tilde{S} = (1, z, \tilde{g}, \tilde{K})^{\mathrm{T}}$ be the **state vectors,** $\tilde{u} = (\hat{m}', \tilde{x})^{\mathrm{T}}$ be the **individual's decision vector**, and $\tilde{U} = (\tilde{p}, \tilde{X})^{\mathrm{T}}$, $W = (z, \tilde{g}, \hat{m}, \tilde{K}, \tilde{k}, \hat{m}', \tilde{x}, \tilde{p}, \tilde{X})^{\mathrm{T}}$ be the **economy-wide variable of vectors**. Furthermore, let the variables with superscript asterix "*" represent the values at steady state.

Applying the second order Taylor's expansion to the utility function, the approximated problem of the representative agent can be written as

$$\tilde{s}^{\mathrm{T}} V \tilde{s} = \max[\tilde{s}^{\mathrm{T}}, \tilde{u}^{\mathrm{T}}, \tilde{U}^{\mathrm{T}}] Q \begin{bmatrix} \tilde{s} \\ \tilde{u} \\ \tilde{U} \end{bmatrix} + \beta \tilde{s}'^{\mathrm{T}} V \tilde{s}', \tag{15}$$

$$s.t.\ \tilde{g}' = \bar{\alpha}\tilde{g} + \epsilon, \tag{16}$$

$$z' = \bar{\gamma} z + \epsilon, \tag{17}$$

$$e^{\tilde{K}'} = (1-\delta) e^{\tilde{K}} + e^{\tilde{X}}, \tag{18}$$

$$\tilde{U} = \mathbf{U}(S), \tag{19}$$

where Q is a 10×10 symmetric matrix with elements[1]:

$$Q_{1,i+1} = Q_{i+1,1} = \tfrac{1}{2}[D_i u(W^*) - \textstyle\sum_{j=1}^{9} D_{ij}^2 u(W^*) W_j^*],$$
$$Q_{i+1,j+1} = Q_{j+1,i+1} = \tfrac{1}{2} D_{ij}^2 u(W^*),$$
$$Q_{11} = u(W^*) - \textstyle\sum_{j=1}^{9} D_j u(W^*) + \tfrac{1}{2} \sum_{i=1}^{9} \sum_{j=1}^{9} D_{ij}^2 u(W^*) W_i^* W_j^*,$$

for $i, j = 1, \cdots, 9$; and $\mathbf{U}$ is a linear function which describes the relationship between $\tilde{U}$ and $\tilde{S}$ realized by the households at the beginning of each period. Cooley and Hansen (1989) have shown that the linear decision rules of this model have the following forms

$$\ln \hat{p} = d_{11} + d_{12} z + d_{13} \ln g + d_{14} \ln K, \tag{20}$$

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

where the coefficients d_{ij} are complicated functions of the steady state values of the variables.

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

Next, we use the decision rules in (20) and (21), together with the Taylor's expansion to linearize the nonlinear constraint around the steady state to obtain the linearized version of the Cooley and Hansen model. Since here cash goods contains only consumption goods, the binding cash-in-advance constraint implies that all money is used on consumption goods. The infinite marginal utility of consumption at zero consumption level, $\lim_{c \to 0} \partial u(c, h)/\partial c = \infty$, guarantees that money holding must be positive. Together with the equilibrium condition $\widehat{m} = 1$, the binding cash-in-advance constraint (11) implies that price level and per capita consumption have an inverse relationship, $C = 1/\hat{p}$. Replacing $\hat{p}$ by equation (20), we then have

$$\ln C = -d_{11} - d_{12} z - d_{13} \ln g - d_{14} \ln K. \tag{22}$$

Next, by aggregating the budget constraint and using the equilibrium wage and rental rate implied by (7) we obtain

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

which represents the per capita working hours. Taking the logarithms with respect to H yields

$$\ln H = \alpha_1 + \alpha_2 z_t + \alpha_3 \ln g_t + \alpha_4 \ln K_t, \tag{23}$$

where the coefficients $\alpha_i' s$ are functions of d_{ij} and the steady state values of the variables.[2]

Similarly, taking logarithms with respect to the constant return-to-scale production function $Y_t = \exp(z_t) K_t^{\rho} H_t^{1-\rho}$ and substitute (23) into it, yield

$$\ln Y = \beta_1 + \beta_2 z_t + \beta_3 \ln g_t + \beta_4 \ln K_t, \tag{24}$$

where $\beta_i' s$ are also the functions of d_{ij} and the steady state values of the variables.[3]

Furthermore, linearizing the law of motion of capital accumulation (18) and using the condition $\tilde{K}'^{*} = \tilde{K}^{*}$ at the steady state yield

$$\ln K_{t+1} = \gamma_1 + \gamma_2 z_t + \gamma_3 \ln g_t + \gamma_4 \ln K_t, \tag{25}$$

[2] See appendix for the details.

[3] See appendix.

where $\gamma_i's$ depend d_{ij} and the steady state values of the variables.[4]

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

$$\begin{aligned} \ln M_t &= \ln g_t + \ln M_{t-1}, \\ &= \sum_{j=0}^{t-1} \ln g_{t-j} + M_0, \end{aligned} \quad (26)$$

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 (20) by (26) yields

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

Therefore, the linearized Cooley and Hansen model can be summarized as

$$\begin{aligned} \ln p_t &= d_{11} + 1 + d_{12} z_t + (d_{13} + 1) \ln g_t + \sum_{j=1}^{t-1} \ln g_{t-j} + d_{14} \ln K_t, \\ \ln X_t &= d_{21} + d_{22} z_t + d_{23} \ln g_t + d_{24} \ln K_t, \\ \ln C_t &= -d_{11} - d_{12} z_t - d_{13} \ln g_t - d_{14} \ln K_t, \\ \ln H_t &= \alpha_1 + \alpha_2 z_t + \alpha_3 \ln g_t + \alpha_4 \ln K_t, \\ \ln Y_t &= \beta_1 + \beta_2 z_t + \beta_3 \ln g_t + \beta_4 \ln K_t, \\ \ln K_{t+1} &= \gamma_1 + \gamma_2 z_t + \gamma_3 \ln g_t + \gamma_4 \ln K_t, \end{aligned} \quad (27)$$

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 (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 are equal to 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

$$\ln K^* = (2 + \ln \delta) / (1 + \delta).$$

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 in (27) allows all the

[4] See appendix.

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 $I(1)$ according to (27).

Denoting $W_t = (\ln p_t, \ln X_t, \ln C_t, \ln H_t, \ln Y_t)^{\mathrm{T}}$, the system of equations in (27) can be rewritten in the form of a VAR model

$$W_t = \delta_0 + D_1 u_t + \delta_3 \sum_{j=1}^{t-1} \ln g_{t-j} + \delta_4 \ln K_t, \quad (28)$$

$$\Delta \ln K_t = \eta^{T} u_{t-1}, \quad (29)$$

$$u_t = \Gamma u_{t-1} + e_t, \quad (30)$$

and

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

where δ_0, δ_1, δ_2, δ_3, δ_4 are 5×1 vectors, and the links between the elements in δ_i and the deep parameters and parameters in the decision rules are stated in the appendix. Further, $D_1 = (\delta_1, \delta_2)$, $\eta = (\gamma_2, \gamma_3)^{\mathrm{T}}$, $u_t = (z_t, \ln g_t)$, and $\Gamma = \mathrm{diag}(\bar{\gamma}, \bar{\alpha})$. The diagonal variance matrix of e_t is due to the assumption that two stochastic shocks are independent.

2.3 The Error Correction Model

Now we show that the linearized model implies a cointegrated error correction model. First, let $D_2 = (\delta_{1,}\delta_2, \delta_3, \delta_4)$ and

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

Then multiplying D_2^T on both sides of (28) and rearranging the equation gives

$$u_{t-1} = I_0 (D_2^{\mathrm{T}} D_2)^{-1} D_2^{\mathrm{T}} (W_{t-1} - \delta_0). \quad (31)$$

Taking the first difference with respect to equation (28) and then substituting (29) and (30) into it yield

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

Let $\delta_5 = \delta_3(0,1)$. Then

$$\begin{aligned}\Delta W_t &= D_1(u_t - u_{t-1}) + \delta_5 u_{t-1} + \delta_4 \eta^{\mathrm{T}} u_{t-1}, \\ &= D_1(\Gamma u_{t-1} + e_t) - D_1 u_{t-1} + \delta_5 u_{t-1} + \delta_4 \eta^{\mathrm{T}} u_{t-1}, \\ &= [D_1(\Gamma - I_2) + \delta_4 \eta^{\mathrm{T}} + \delta_5] u_{t-1} + D_1 e_t.\end{aligned}$$

Replacing u_{t-1} by (31), (28) through (30) can be now be written in the form of an error correction model:

$$\begin{aligned}\Delta W_t &= [D_1(\Gamma - I_2) + \delta_4 \eta^{\mathrm{T}} + \delta_5] I_0 (D_2^{\mathrm{T}} D_2)^{-1} D_2^{\mathrm{T}} (W_{t-1} - \delta_0) + D_1 e_t, \\ &= \nu_0 + \alpha \beta^{\mathrm{T}} (W_{t-1} - \delta_0) + \varepsilon_t, \qquad (32)\end{aligned}$$

where

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

and $\bar{e} = (0, (1-\overline{\alpha}) \ln \overline{g})^{\mathrm{T}}$. Here, ν_0 is a 5×1 vector, and both α and β are 5×2 matrices.

Moreover, the error term $\varepsilon_t = D_1(e_t - \bar{e})$ is distributed as i.i.d. $N(0, \Sigma(\theta))$, where

$$\Sigma(\theta) = D_1 \Xi D_1^{\mathrm{T}} = \sigma_\epsilon^2 \delta_1 \delta_1^{\mathrm{T}} + \sigma_\xi^2 \delta_2 \delta_2^{\mathrm{T}}, \ \Xi = \begin{pmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_\xi^2 \end{pmatrix},$$

and θ represents the set of deep parameters.

Since the model is driven by only two stochastic shocks, the 5×5 variance-covariance matrix of the error term ε_t in (32) is of rank two. Therefore, the conditional distribution of the linearized theoretical model is a singular multivariate normal distribution. The main consequence of the singularity is that the density function of ΔW_t in the theoretical model does not exist.

Moreover, both α and β are 5×2 matrices, and thus $\alpha \beta^{\mathrm{T}}$ is a 5×5 matrix with rank two. Thus, two cointegrating vectors are implied by the theoretical model. In the error correction model, $\beta^{\mathrm{T}} W_{t-1}$ represents an economic equilibrium relation with the adjustment coefficient α, and matrix β of cointegration vectors.

In summary, the linearized Cooley and Hansen model is equivalent to a cointegrated error correction model with two cointegration vectors and singular error variance matrix of rank two. The conditional distribution of ΔW_t is

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

where the superscript "TM" denote the theoretical model and $\mu_t(\theta) = \nu_0 + \alpha \beta^{\mathrm{T}} (W_{t-1} - \delta_0)$.

3 The conditional reality bound

Next we introduce the conditional reality bound of Bierens and Swanson (2000) and Bierens (2006). Let $f(y)$ denote the true density of a vector y, and $f_{\text{TM}}(y)$ be the approximated density of $f(y)$ derived from a theoretical model. Suppose that both $f(y)$and $f_{\text{TM}}(y)$ have the same support. Then we can find a $p \in [0,1]$ satisfying

$$f(y) = pf_{\text{TM}}(y) + (1-p)f_1(y), \tag{33}$$

where $pf_{\text{TM}}(y) \leq f(y)$ for all values of y, and $f_1(y) = (f(y) - pf_{\text{TM}}(y)/(1-p)$. Thus, if $p \in (0,1)$, $f(y)$ is a mixture of a density $f_{\text{TM}}(y)$ with associated probability p, and another density $f_1(y)$ with associated probability $1-p$. Therefore, p may be interpreted as the probability that $f(y)$ is generated by $f_{\text{TM}}(y)$, so that it measures the reality content of the theoretical model.

The maximal p satisfying (33) is

$$p = \inf_y \frac{f(y)}{f_{\text{TM}}(y)}. \tag{34}$$

Geometrically, the approximated density is squeezed under the true density by multiplying it by the p in (34).

In the case of conditional densities, given predetermined variables X, (34) becomes

$$p(X,\theta) = \inf_y \frac{f(y|X)}{f_{\text{TM}}(y|X,\theta)}, \tag{35}$$

where θ is a vector the model parameters. Since p_0 depends on X and θ, we can then define the maximal possible p as

$$p = \sup_\theta E\left[p(X,\theta)\right],$$

which is called **the average conditional reality bound** in Bierens and Swanson (2000). Empirically, it can be estimated by

$$\widehat{p}_{\text{avg}} = \sup_\theta \frac{1}{n}\sum_{t=1}^{n} p(X_t,\theta). \tag{36}$$

On the other hand, if we interpret $p(X_t,\theta)$ as the probability that $f_{\text{TM}}(y|X_t,\theta)$ is equal to $f(y|X_t)$, then given the data for time $t = 1, ..., n$,

$$\prod_{t=1}^{n} p(X_t,\theta) \tag{37}$$

may be interpreted as the probability that the true conditional joint density is the same as the conditional joint density of $Y_1, ..., Y_n$ implied by the theoretical model. The maximum of (37) over the parameter space, $\max \theta \prod_{t=1}^{n} p(X_t, \theta)$, is called by Bierens (2006) **the multiplicative conditional reality bound**, with corresponding estimator

$$\widehat{\theta} = \arg\max_{\theta} \sum_{t=1}^{T} \ln p(X_t, \theta). \tag{38}$$

4 Model Estimation and the Empirical Results

4.1 The Data and Empirical Model

The conditional reality bound in (35) compares the theoretical density with the true density. However, the true density or the data-generating process is usually unknown. Therefore, the true density is replaced by a conditional density of an empirical econometric model (hereafter called **the empirical density**) which is the best fit for the data. In this section, we introduce the data used here and illustrate how to find an empirical model that can be used to represent the DGP in our analysis.

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 US. 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 taken in natural logs. These five series are shown in Figure 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. However, this filter will effectively remove one of the interesting features of the model, namely cointegration. Therefore, we will use the unfiltered time series.

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 $\ln Y$

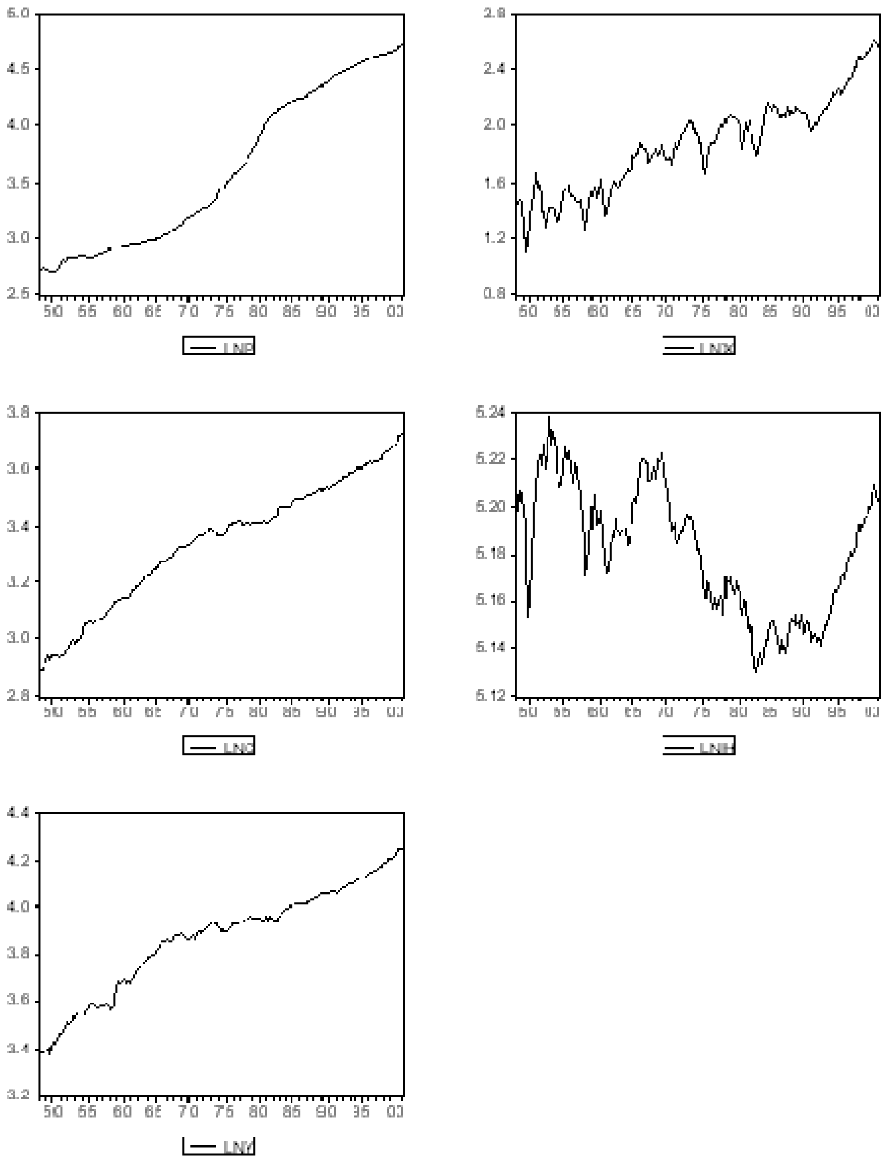

Figure 1: The five U.S. series

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 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.

Given the above results, we first parameterize the empirical model as a VECM of order p:

$$\Delta W_t^{\mathrm{EM}} = \pi_0 + \sum_{j=1}^{p-1} \Pi_j \Delta W_{t-j} + \alpha\beta^{\mathrm{T}} W_{t-1} + \xi_t,$$

where the superscript "EM" denote the empirical model, $\xi_t \sim N_5(0, \Omega)$ and $\det(\Omega) \neq 0$. Therefore, the conditional distribution of ΔW_t^{EM} is

$$\Delta W_t^{\mathrm{EM}} | \mathcal{F}_{t-1} \sim i.i.d.\ N_5(\omega_t, \Omega),$$

where $\mathcal{F}_{t-1}$ represents the information available at time $t-1$,[5] and

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

As Figure 1 shows, all but one of the series have drift. The exception is $\ln H$. Therefore, we need to include a vector π_0 of intercept terms in the VECM.

Two consistent information criteria, the Hannan-Quinn (1979) criterion (HQ) and the Schwarz (1978) criteria (SC), are used to decide the number of lags for the empirical model. We choose the lag order for the VECM by examining values of HQ and SC for lag orders from one to ten. The HQ suggests two lags, whereas 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, we choose the error correction model with two lags.

Johansen's lambda max and trace tests indicate that β has rank 1, so that there is only one cointegrating vector. If π_0 can be written as $\pi_0 = -\alpha\delta_0$,

[5] More formally, $\mathcal{F}_{t-1}$ is the σ-algebra generated by $W_{t-1}, \Delta W_{t-1}, \Delta W_{t-2}, \Delta W_{t-3}, \ldots$.

Table 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
$s.e.$	0.005048	0.051947	0.005318	0.005700	0.008916
R^2	0.6358	0.1272	0.1509	0.0952	0.0959

then the model becomes $\Delta W_t^{\mathrm{EM}} = \pi_1 \Delta W_{t-1} + \alpha\beta^{\mathrm{T}}(W_{t-1} - \delta_0)$. We also have conducted the Johansen tests under this restriction, but this hypothesis is rejected. Therefore, a VECM with two lags without cointegrating restrictions on intercepts is chosen as the empirical model:

$$\Delta W_t^{\mathrm{EM}} = \pi_0 + \Pi_1 \Delta W_{t-1} + \alpha\beta^{\mathrm{T}} W_{t-1} + \xi_t, \tag{39}$$

where α is the vector the adjustment coefficient and β is the cointegration vector.

The estimated results are shown in Table 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.

4.2 The average conditional reality bound

We now follow the lines in Bierens and Swanson (2000) to compute the average reality bound. According to the results in section 2.3, the error correction model derived from the theoretical model has one lag and can be written as

$$\Delta W_t^{\mathrm{TM}} = \nu_0 + \alpha_0 \beta_0^{\mathrm{T}} W_{t-1} + \varepsilon_t,$$

where ΔW_t has a singular conditional normal distribution

$$\Delta W_t^{\mathrm{TM}} | W_{t-1} \sim N_5(\mu_t(\theta), \Sigma(\theta)),\ \operatorname{rank}(\Sigma(\theta)) = 2, \tag{40}$$

with conditional mean $\mu_t(\theta) = \nu_0 + \alpha_0 \beta_0^{\mathrm{T}} W_{t-1}$. Recall that both α_0 and β_0 are 5×2 matrices, and thus the theoretical model implies two cointegrating relationships. The rank condition is due to the fact that there are only two, instead of five, stochastic shocks in the Cooley and Hansen model. Model (40) will be referred to as the **theoretical model**.

On the other hand, the empirical model (39) has two lags and one cointegrating vector, with nonsingular conditional normal distribution of ΔW_t,

$$\Delta W_t^{\mathrm{EM}} | \mathcal{F}_{t-1} \sim N_5(\omega_t, \Omega), \tag{41}$$

where $\omega_t = \pi_0 + \Pi_1 \Delta W_{t-1} + \alpha \beta^{\mathrm{T}} W_{t-1}$.

Due to the singularity of the conditional variance matrix of ΔW_t^{TM}, the conditional density of ΔW_t^{TM} does not exist and thust he conditional reality bound (35) cannot be applied directly. Bierens and Swanson (2000) propose to rotate ΔW_t in both models via an orthogonal transformation $Q(\theta)^{\mathrm{T}} \Delta W_t$ such that the singular directions of $Q(\theta)^{\mathrm{T}} \Delta W_t$ in the case of the theoretical model are along the principle axes, and then marginalize both distributions in the nonsingular directions. Thus, decompose $\Sigma(\theta)$ as

$$\begin{aligned} \Sigma(\theta) &= (Q_1(\theta), Q_2(\theta)) \begin{pmatrix} O & O \\ O & \Lambda(\theta) \end{pmatrix} \begin{pmatrix} Q_1^{\mathrm{T}}(\theta) \\ Q_2^{\mathrm{T}}(\theta) \end{pmatrix}, \\ \Lambda(\theta) &= \operatorname{diag}(\lambda_1(\theta), \lambda_2(\theta)), \end{aligned}$$

where $Q_1(\theta)$ is the 5×3 matrix of eigenvectors corresponding to the three zero eigenvalues of $\Sigma(\theta)$, and $Q_2(\theta)$ is the 5×2 matrix of eigenvectors corresponding to the two positive eigenvalues, $\lambda_1(\theta)$ and $\lambda_2(\theta)$. Pre-multiplying both distributions by the orthogonal matrix $Q(\theta)^{\mathrm{T}}$, the non-singular part of model (40) is

$$Q_2(\theta)^{\mathrm{T}} \Delta W_t^{\mathrm{TM}} | \mathcal{F}_{t-1} \sim N_2(Q_2(\theta)^{\mathrm{T}} \mu_t(\theta), \Lambda(\theta)), \tag{42}$$

with corresponding marginal empirical conditional distribution

$$Q_2(\theta)^{\mathrm{T}} \Delta W_t^{\mathrm{EM}} | \mathcal{F}_{t-1} \sim N_2\left(Q_2(\theta)^{\mathrm{T}} \omega_t, Q_2(\theta)^{\mathrm{T}} \Omega Q_2(\theta)\right). \tag{43}$$

The conditional densities in (42) and (43) can now be used to compute the conditional reality bound in (35), as follows. Denote

$$\begin{aligned} Y_t &= Q_2(\theta)^{\mathrm{T}} \Delta W_t^{\mathrm{TM}} - Q_2(\theta)^{\mathrm{T}} \mu_t(\theta), \\ \eta_t(\theta) &= Q_2(\theta)^{\mathrm{T}} \left(\omega_t - \mu_t(\theta)\right), \end{aligned}$$

and

$$X_t = \left(W_{t-1}^{\mathrm{T}}, W_{t-2}^{\mathrm{T}}\right)^{\mathrm{T}}. \tag{44}$$

The conditional density of Y_t in the case (42) is

$$f_{\mathrm{TM}}(y|X_t,\theta) = \frac{\exp\left[-\frac{1}{2} y^{\mathrm{T}} \Lambda(\theta)^{-1} y\right]}{2\pi\sqrt{\det\left(\Lambda(\theta)\right)}}, \tag{45}$$

and in the case (43),

$$\begin{aligned} f_{\mathrm{EM}}(y|X_t) &= \frac{1}{2\pi\sqrt{\det(Q_2(\theta)^{\mathrm{T}}\Omega Q_2(\theta))}} \qquad (46) \\ &\times \exp[-\frac{1}{2}(y-\eta_t(\theta))^{\mathrm{T}} \left(Q_2(\theta)^{\mathrm{T}}\Omega Q_2(\theta)\right)^{-1}(y-\eta_t(\theta)]. \end{aligned}$$

Under some mild condition[6], the conditional reality bound in (35) can be written as

$$\begin{aligned} p(X_t,\theta) &= \inf_y \frac{f_{\mathrm{EM}}(y|X_t)}{f_{\mathrm{TM}}(y|X_t,\theta)} \\ &= \sqrt{\det \Psi(\theta)} \exp\left[-\frac{1}{2}\vartheta_t(\theta)^{\mathrm{T}}\Psi(\theta)\,\vartheta_t(\theta)\right] \\ &\times \exp\left[-\frac{1}{2}\vartheta_t(\theta)^{\mathrm{T}}\Psi(\theta)\left(I-\Psi(\theta)\right)^{-1}\Psi(\theta)\,\vartheta_t(\theta)\right], \end{aligned} \tag{47}$$

where

$$\begin{aligned} \Psi(\theta) &= \Lambda(\theta)^{1/2}\left(Q_2(\theta)^{\mathrm{T}}\Omega Q_2(\theta)\right)^{-1}\Lambda(\theta)^{1/2}, \\ \vartheta_t(\theta) &= \Lambda(\theta)^{-1/2} Q_2(\theta)^{\mathrm{T}}(\omega_t - \mu_t(\theta)), \end{aligned}$$

[6] 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.

and $\lambda_{\max}[\Psi(\theta)]$ is the maximal eigenvalue of $\Psi(\theta)$. The average conditional reality bound can now be estimated by

$$\widehat{p}_{\text{avg}} = \sup_{\theta, \lambda_{\max}[\Psi(\theta)]<1} \frac{1}{n} \sum_{t=1}^{n} p(X_t, \theta). \tag{48}$$

The matrix $Q_2(\theta)^{\mathrm{T}}$ is chosen such that in the case of the theoretical model the two components of Y_t are independent conditional on X_t. It is quite natural to require the same for the empirical model. Therefore, we will choose $Q_2(\theta)$ to be the 5×2 matrix of orthonormal eigenvectors corresponding to the largest two eigenvalues of Ω. Thus, arranging the eigenvalues λ_i^* of Ω in descending order, we impose the restriction

$$Q_2(\theta) = Q_*, \text{ where } Q_*^{\mathrm{T}}\Omega Q_* = \Lambda_* = \begin{pmatrix} \lambda_1^* & 0 \\ 0 & \lambda_2^* \end{pmatrix}.$$

Under this restriction, the non-singular part of the theoretical model can be written as

$$\begin{aligned} \Delta Q_*^{\mathrm{T}} W_t^{\mathrm{TM}} &= Q_*^{\mathrm{T}}\nu_0 + Q_*^{\mathrm{T}}\alpha_0\beta_0^T W_{t-1} + Q_*^{\mathrm{T}}\varepsilon_t \\ &= \nu_* + \alpha_*\beta_0^{\mathrm{T}} W_{t-1} + Q_*^{\mathrm{T}}\varepsilon_t, \end{aligned}$$

for instance, where

$$Q_*^{\mathrm{T}}\varepsilon_t \sim N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda_1(\theta) & 0 \\ 0 & \lambda_2(\theta) \end{pmatrix}\right).$$

Then $\Psi(\theta)$ simplifies to

$$\Psi(\theta) = \begin{pmatrix} \lambda_1(\theta)/\lambda_1^* & 0 \\ 0 & \lambda_2(\theta)/\lambda_2^* \end{pmatrix},$$

so that the condition $\lambda_{\max}[\Psi(\theta)] < 1$ becomes

$$\lambda_1(\theta) < \lambda_1^*, \ \lambda_2(\theta) < \lambda_2^*.$$

Consequently, (47) can now be written as

$$
\begin{aligned}
p_t(\theta) &= \sqrt{\frac{\lambda_1(\theta).\lambda_2(\theta)}{\lambda_1^*.\lambda_2^*}} \\
&\quad \times \exp\left[-\frac{1}{2}\eta_t(\theta)^{\mathrm{T}}\left(\Lambda_* - \Lambda(\theta)\right)^{-1}\eta_t(\theta))\right] \\
&= \sqrt{\frac{\lambda_1(\theta).\lambda_2(\theta)}{\lambda_1^*.\lambda_2^*}}\exp\left(-\frac{1}{2}\sum_{i=1}^{2}\left(\frac{\left(\eta_{i,t}(\theta)\right)^2}{\lambda_i^* - \lambda_i(\theta)}\right)\right)
\end{aligned}
$$

where

$$
\eta_t(\theta) = Q_*^{\mathrm{T}}\left(\omega_t - \mu_t(\theta)\right) = \begin{pmatrix} \eta_{1,t}(\theta) \\ \eta_{2,t}(\theta) \end{pmatrix}
$$

However, this simplification comes at the cost of losing the link between $(\lambda_1(\theta), \lambda_2(\theta))$ and the deep parameters in θ, so that then $\lambda_1 = \lambda_1(\theta)$ and $\lambda_2 = \lambda_2(\theta)$ are no longer functions of θ, but are now have to be treated as free parameters subject to the restrictions $\lambda_1 < \lambda_1^*, \lambda_2 < \lambda_2^*$. Therefore, reparametrize λ_1 and λ_2 as $\lambda_1 = c_1\lambda_1^*$ and $\lambda_2 = c_2\lambda_2^*$, respectively, where $c_1, c_2 \in (0, 1)$. Then (48) is equivalent to

$$
\widehat{p}_{\mathrm{avg}} = \sup_{\theta, 0<c_1<1, 0<c_2<1}\frac{1}{n}\sum_{t=1}^{n}\exp\left(-\frac{1}{2}\sum_{i=1}^{2}\frac{\sqrt{c_1.c_2}}{1-c_1}\left(\eta_{i,t}(\theta)\right)^2/\lambda_i^*\right) \tag{49}
$$

The estimation result involved is

$$
\widehat{p}_{\mathrm{avg}} = 0.473989,
$$

which suggests that the nonsingular part of the theoretical model can explain about 47% of the corresponding part of the empirical model.

The assumption that the two stochastic shocks follow AR(1) processes is not essential in the Cooley and Hansen model. Therefore, we also consider the case where the shocks follow AR(2) processes in order to examine whether the main penalty of the average conditional reality bound is caused by the lag difference between the empirical and theoretical models. Thus, we change the assumption of the stochastic shocks in (30) to

$$
u_t = \Gamma_1 u_{t-1} + \Gamma_2 u_{t-2} + e_t, \tag{50}
$$

but (28) and (29) remain unchanged.

Using a procedure similar to that used in section 2.3, it can be shown that (50), (28) and (29) implies

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

where

$$\begin{aligned}
\nu_0 &= [-(D_1(\Gamma_1 + \Gamma_2 - I_2) + \delta_4\eta^{\mathrm{T}})I_0(D_2^{\mathrm{T}} D_2)^{-1} D_2^{\mathrm{T}} \delta_0], \\
\nu_1 &= -(D_1\Gamma_2 - \delta_5)\, I_0(D_2^{\mathrm{T}} D_2)^{-1} D_2^{\mathrm{T}}, \\
\alpha &= D_1(\Gamma_1 + \Gamma_2 - I_2) + \delta_4\eta^{\mathrm{T}}, \\
\beta &= D_2(D_2^{\mathrm{T}} D_2)^{-1} I_0^{\mathrm{T}}.
\end{aligned}$$

Given the assumption that both of the shocks follow AR(2) processes, the Cooley and Hansen model implies a two-lag error correction model.

Applying the same estimation procedure as in (49), the estimated results for the average conditional reality bound now becomes

$$\widehat{p}_{\mathrm{avg}} = 0.897437.$$

This result suggests that,with the VAR(2) error specification, the non-singular part of the theoretical model can explain about 90% of the corresponding part of the empirical model. Thus the lag difference is the major penalty in the average conditional reality bound. The remaining 10% penalty may be due to the rank difference.

4.3 Estimation of the deep parameters

In order to eliminate the singularity the theoretical model, Bierens (2006) suggests to augment both the theoretical and empirical models by the same independent stochastic disturbances, say R_t^{TM} and R_t^{EM}, both distributed as $N(0, \tau\Omega)$ for some $\tau > 0$, where Ω is the error variance matrix of the original empirical model. Then adding $R_{T,t}$ to the linearized theoretical model yields

$$Y_t = \Delta W_t^{\mathrm{TM}} + R_t^{\mathrm{TM}} = \nu_0 + \alpha_0\beta_0^{\mathrm{T}} W_{t-1} + \varepsilon_t + R_t^{\mathrm{TM}},$$

which has conditional density

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

where X_t is defined in (44).

To keep the convoluted theoretical density $f_{\mathrm{TM}}(y|X_t, \theta, \tau)$ comparable with the conditional density of the empirical model (39), one should also augment the latter model in the same way:

$$Y_t = \Delta W_t^{\mathrm{EM}} + R_t^{\mathrm{EM}} = \pi_0 + \pi_1 \Delta W_{t-1} + \alpha\beta^{\mathrm{T}} W_{t-1} + \xi_t + R_t^{\mathrm{EM}},$$

which has conditional density

$$\begin{aligned} f_{\mathrm{EM}}(y|X_t, \tau) &= \frac{1}{(\sqrt{2\pi})^5 \sqrt{(1+\tau)^5 \det(\Omega)}} \\ \times \exp &\left\{ -\frac{1}{2(1+\tau)} (y - \omega_t)^{\mathrm{T}} \Omega^{-1} (y - \omega_t) \right\}. \end{aligned}$$

If all the eigenvalues of $\Omega^{-1/2}\Sigma(\theta)\Omega^{-1/2}$ are less than 1[7], the conditional reality bound in (35) takes the form

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

where

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

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

and $Q_1(\theta)$ and $Q_2(\theta)$ are 5×2 and 5×3 matrices of the corresponding orthogonal eigenvectors. Consequently, (38) becomes

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^{n} \ln p_t(\theta, \tau) &= \frac{3}{2} \ln \tau - \frac{5}{2} \ln(1+\tau) + \frac{1}{2} \ln\left(\det(\Lambda_1(\theta) + \tau I_2)\right) \\ -\frac{1}{2} \mathrm{trace} &\left((I_5 - \Lambda(\theta))^{-1} Q(\theta)^{\mathrm{T}} \Gamma_n(\theta) Q(\theta) \right). \end{aligned} \tag{51}$$

[7] See Bierens (2006) for the details.

where

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

Given different values of τ, (51) is maximized over

$$\theta = (d_{11}, d_{12}, d_{13}, d_{14}, d_{22}, d_{23}, p^*, \overline{\alpha}, \overline{\gamma}, \rho, \delta, \ln \overline{g})',$$

where $p^*, \ln \overline{g} > 0$, and $\overline{\alpha}, \overline{\gamma}, \rho, \delta \in (0, 1)$. Although our approach also applies if the parameters are restricted to a reasonable area or an area close to the calibrated values, we have chosen to leave the parameters free and let the data speak.

Since the parameter vector θ only contains a subset of the deep parameters, not all of the deep parameters can be estimated. Some links between the coefficients in (51) and the deep parameters, such as the discount rate in the utility function, are lost in the process of linearization.

Given $\tau = 0.1, 0.2, ..., 0.5$, the estimated parameters are shown in Tables 2 and 3. We see that the parameters are stable with only little fluctuation under different values of τ, but the values of both ρ and δ are much higher than the usual calibrated values. For example, Cooley and Hansen (1989) calibrated $\rho = 0.36$ and $\delta = 0.025$. 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 it is difficult to pinpoint where the failure occurs. Moreover, the estimated $\ln \overline{g}$ implies an average annual monetary growth rate between 3.8% to 5.7% for different τ values.

Because the estimated deep parameters θ are functions of the parameters of the empirical model, Bierens (2006) proposes to compute confidence intervals around the estimated parameters by using the delta-method and the asymptotic normality of the empirical model parameter estimators. However, due to cointegration not all the parameter estimators of the empirical model have a normal limiting distribution. To overcome this problem, we have computed 90% confidence intervals of the deep parameters via a bootstrap approach. Under the assumption that the empirical model represents the data generating process, 500 artificial time series W_t^* have been generated from the estimated empirical model, by replacing the estimated errors by random drawings z_t^* from the normal distribution with variance the estimated error variance matrix $\widehat{\Omega}$:

$$W_t^* = \widehat{\pi}_0 + \left(I_5 + \widehat{\pi}_1 + \widehat{\alpha}\widehat{\beta}'\right) W_{t-1}^* - \widehat{\pi}_1 W_{t-2}^* + z_t^*,\ t = 3, ..., T,$$

Table 2: Estimation results: Part 1.

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}	-0.109160	-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	1.085440
p^*	2.909936	2.756174	2.871515	2.510822	3.334870
$\overline{\alpha}$	0.024367	0.016469	0.027590	0.022449	0.002292
$\overline{\gamma}$	0.000276	0.000045	0.000690	0.000045	0.000078
ρ	0.685083	0.724187	0.693556	0.726532	0.686299
δ	0.728562	0.746744	0.759689	0.622586	0.873120
$\ln\overline{g}$	0.009448	0.010802	0.007361	0.010945	0.013176
	0.011950	0.013924	0.009542	0.013835	0.016738
	0.011571	0.012860	0.009439	0.013997	0.014269
v_0	-0.002732	-0.003300	-0.002384	-0.003136	-0.003593
	0.000785	0.000975	0.000553	0.001200	0.001142
	0.000247	0.000269	0.000169	0.000328	0.000358

Table 3: Estimation results: Part 2. (α is a 5×2 matrix, so the ten elements of α represent the first and second columns of α, respectively. Similar for β.)

τ	0.1	0.2	0.3	0.4	0.5
α	−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
	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
β	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
	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
δ_0	−0.619234	−0.512823	−0.616875	−0.458609	−0.668025
	−0.316683	−0.292033	−0.274847	−0.473873	−0.135682
	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 4: The 90% bootstrap confidence intervals of the estimated parameters for the case $\tau = 0.1$.

	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	−0.109160		−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		−3.383080	−1.137422	−2.138174
d_{23}	0.635488	2.192187	1.255176	α			
p^*	2.178594	3.948820	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	0.599510	0.975084	0.728562		0.394092	1.654245	0.704880
$\ln \overline{g}$	0.002248	0.015261	0.009448				
					0	0	0
	0.003037	0.021012	0.011950		−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	0.001710	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				

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. For each of these 500 artificial time series W_t^* the empirical model and the deep parameters are re-estimated for $\tau = 0.1$, which yields the bootstrap distribution of θ. The choice of $\tau = 0.1$ is arbitrary and, as shown in Tables 2 and 3, the estimated parameters appear to be about the same under different values of τ. The bootstrap results for $\tau = 0.1$ are given in Table 4.

The signs of the coefficients in the decision rules in Table 4 show that the monetary growth rate has a positive effects on the price level as well as on investment. This finding is consistent with the prediction of the cash-in-advance constraint: 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 only be 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.

4.4 Innovation Response Analysis

Our estimation results for the case $\tau = 0.1$ imply the following numerical values for the coefficients in the system of equations (27):

$$\begin{aligned}
\ln P_t &= -0.619234 - 0.109160 z_t + 1.296333 \ln g_t + \sum_{j=1}^{t-1} \ln g_{t-j} \qquad (52)\\
&\quad +0.450385 \ln K_t,\\
\ln X_t &= -0.316683 - 1.787301 z_t + 1.255176 \ln g_t + \ln K_t, \qquad (53)\\
\ln C_t &= 1.619234 + 0.109160 z_t - 0.296333 \ln g_t - 0.450385 \ln K_t, \qquad (54)\\
\ln H_t &= 5.056568 - 0.309707 z_t + 0.085107 \ln g_t + 0.363003 \ln K_t, \qquad (55)\\
\ln Y_t &= 1.592398 + 1.097532 z_t + 0.026802 \ln g_t + 0.799399 \ln K_t, \qquad (56)\\
\ln K_{t+1} &= -1.302159 z_t + 0.914473 \ln g_t + \ln K_t, \qquad (57)
\end{aligned}$$

together with the two stochastic shocks processes

$$z_t = 0.000276 z_{t-1} + \epsilon_t, \qquad (58)$$

$$\ln g_t \;=\; 0.024367 \ln g_{t-1} + \xi_t. \tag{59}$$

Equation (57) is equivalent to

$$\Delta \ln K_t = -1.302159 z_{t-1} + 0.914473 \ln g_{t-1}. \tag{60}$$

Taking the first difference of (52) and substituting (60) yield the equation for the inflation rate:

$$\Delta \ln P_t = -0.109160 z_t - 4.773129 z_{t-1} + 1.296333 \ln g_t + 0.914473 \ln g_{t-1}. \tag{61}$$

Similarly, differencing (56) gives the output growth rate:

$$\Delta \ln Y_t = 1.097532 z_t - 2.138477 z_{t-1} + 0.026802 \ln g_t + 0.704223 \ln g_{t-1}. \tag{62}$$

These estimated equations can be used to generate predictions about the responses to technology and monetary shocks. The innovation responses for both shocks for 15 periods are shown in Figure 2. The first panel in Figure 2 shows the innovation response of $\Delta \ln p$ corresponding to a unit 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 technological progress is assumed, a positive technology shock only boosts the productivity in the 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 shock has a positive effect on output growth in the current period, but a negative effect in the second period. Both (53) and (57) 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 wears off almost completely after three periods, as shown in panel 2.

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

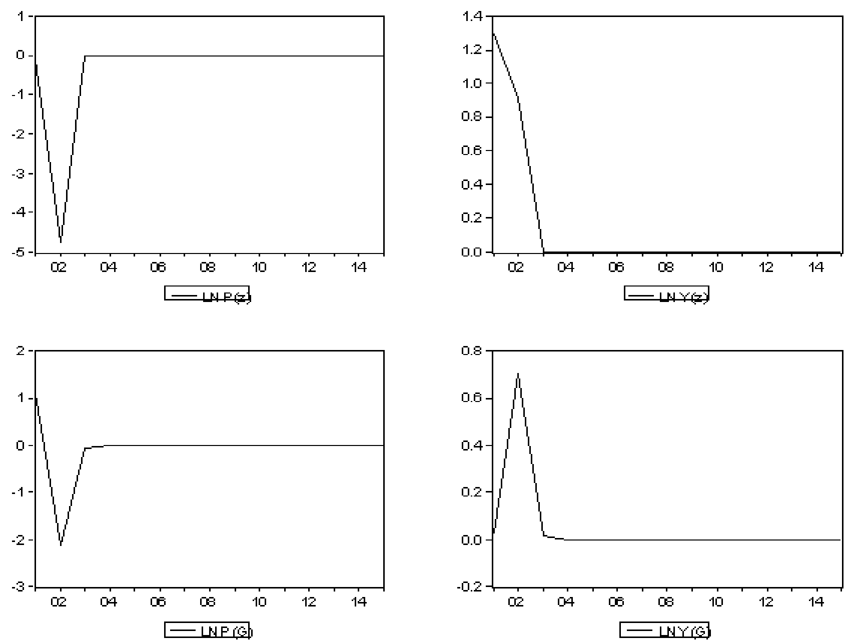

Figure 2: Innovation responses to technology and monetary shocks. Panel 1 and 2 show the effects of one unit of technology shock on inflation rate and output growth rate. Panel 3 and 4 show the effects of one unit of monetary shock on inflation rate and output growth rate.

Although the two stochastic shocks have no permanent effect on inflation and output growth, 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 3 and Figure 4. Figure 3 shows the effects of a unit technology shock on the levels of the variables. Panel 1 in Figure 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 (53), (55), and (56) predict, they all drop to lower levels after the shock. On the other hand, (54) shows that $\ln C$ positively depends on $\ln K$, so $\ln C$ jumps to a higher level after the shock.

The technology shock has no effect on the level of capital stock at 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 at 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 a significantly negative effect on investment, so that $\ln X$ drops at the shock period and stays at a lower level onwards because of the lower level of $\ln K$.

The innovation responses of the levels of the variables to a unit monetary shock are shown in Figure 4. Panel 1 in Figure 4 shows that the price level jumps to a higher level after a positive monetary shock and panel 6 shows that the capital 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 at the shock period and 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 (61) 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.

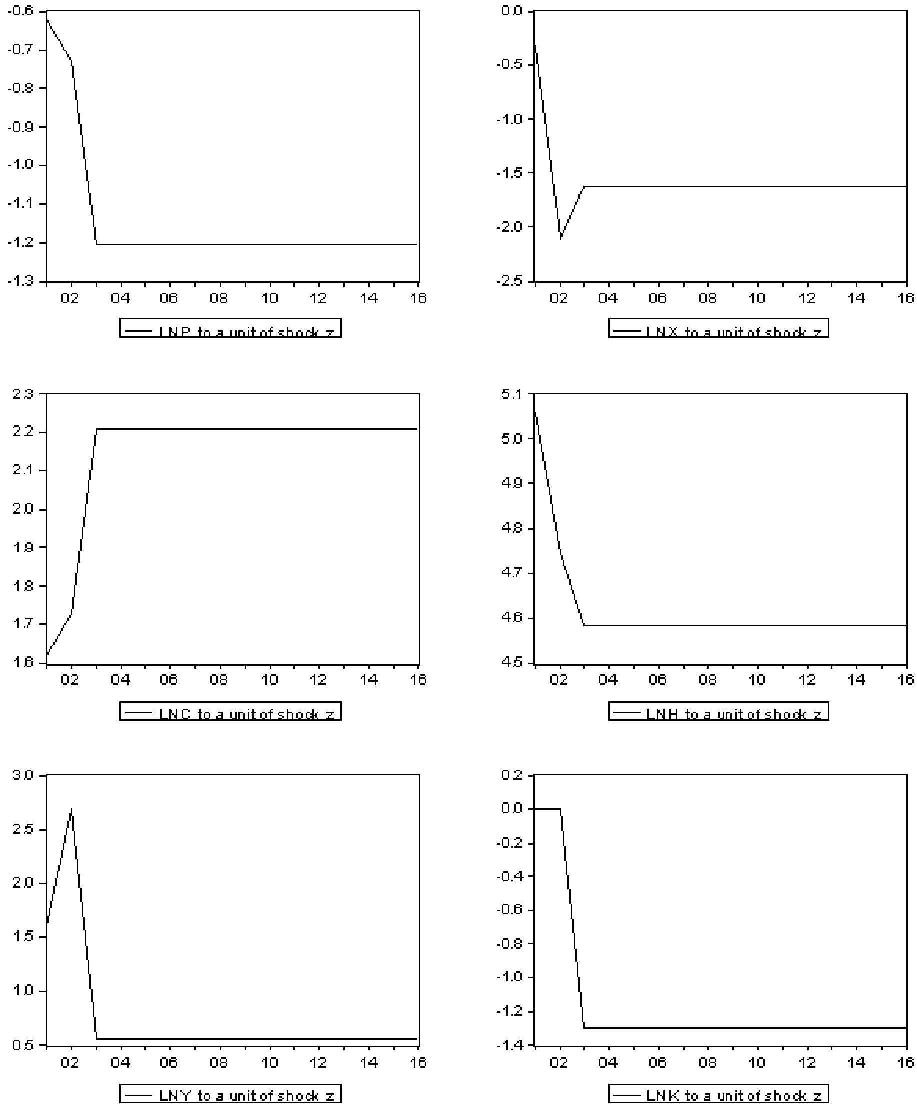

Figure 3: The innovation responses of the levels of the variables to a unit technology shock over a 15 periods horizon. The unit shock happens in period 2.

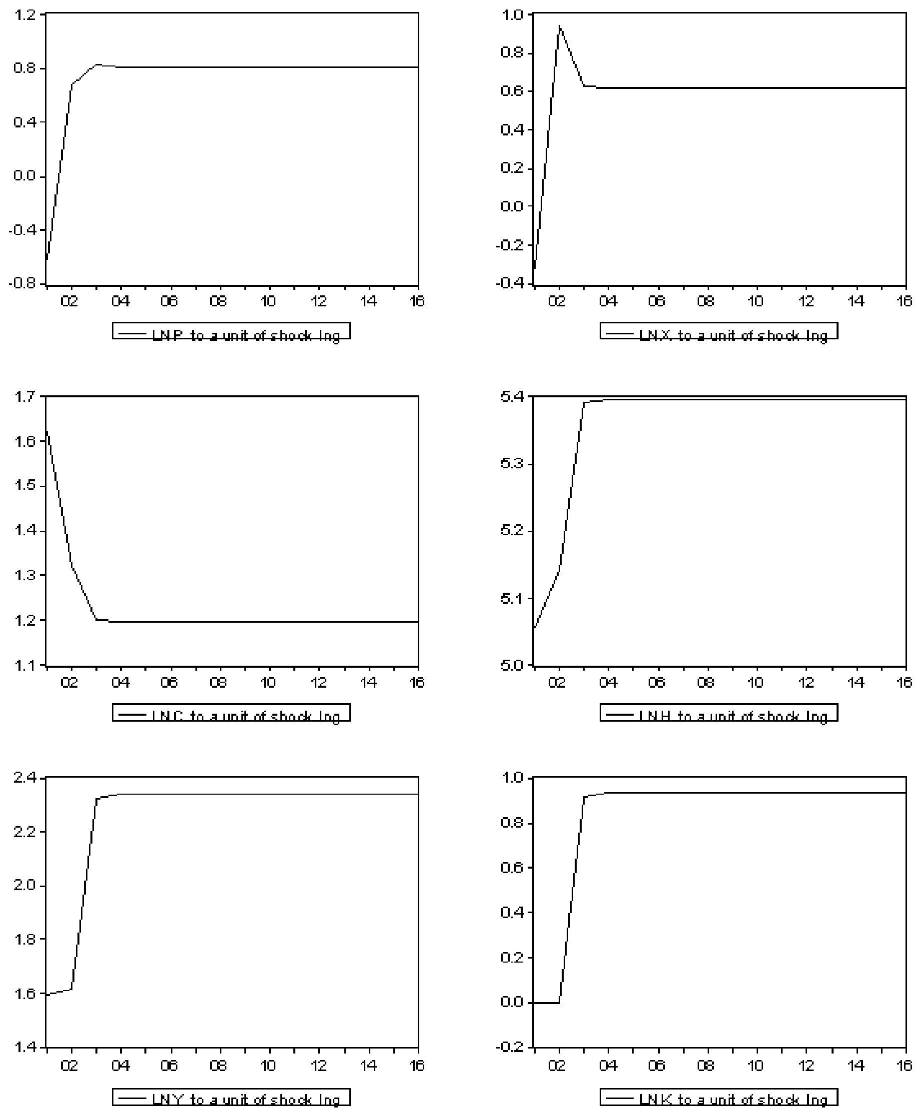

Figure 4: The innovation responses of the levels of the variables to a unit monetary shock over a 15 periods horizon. The unit shock happens in period 2.

5 Summary and Conclusion

Our 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 cash-in-advance model are not penalized, it can account for about 90% of the information of the empirical model in the nonsingular direction after changing stochastic process to an AR(2) process, where it accounts for only 47% under the AR(1) assumption. This suggests that the assumptions about the stochastic structure are critical for the performance of the theoretical model.

Second, the Cooley and Hansen model does not do well in explaining the observed data if the possible model misspecification or missing variables are considered. One evidence is that the estimated capital depreciation rate and capital share are much higher than the usual calibrated values. Despite that the parameters can be restricted to an area close to the calibrated values, we choose to leave the parameters free and let the data speaks. The much higher values implies that some important features of the economy are missing from the model although the failure is not so immediately evident.

Third, the main limitation of the econometric approaches used in this study is that only a subset of the deep parameters can be estimated. Further analysis that requires all deep parameters, such as the analysis of welfare cost of inflation, is not applicable here. Although the coefficients in the decision rules provide some links between the coefficients in the estimated model and the deep parameters, we did not use them to solve the deep parameters due to the complexity of those links.

Fourth, computing the conditional reality bound requires a close approximation of the data-generating process, which is obtained by finding an econometric model that is best fit for the data. However, in some dynamic stochastic general equilibrium models, the DGP may contain some unobserved variables, such as the labor effort in the model of Burnside, Eichenbaum, and Rebelo (1993). In that case our approach may not be applicable .

Appendix

Derivation of equation (23):

$$\begin{aligned}
\ln H &= \frac{1}{1-\rho}\ln\left(X+\frac{1}{\hat{p}}\right)+\frac{z}{1-\rho}+\frac{\rho}{1-\rho}\ln K, \\
&\simeq h_1+h_2\ln X+h_3\ln\hat{p}+\frac{z}{1-\rho}+\frac{\rho}{1-\rho}\ln K, \\
&\simeq [h_1+h_2d_{21}+h_3d_{11}]+[h_2d_{22}+h_3d_{12}+1/(1-\rho)]\,z \\
&\quad +[h_2d_{23}+h_3d_{13}]\ln g+[h_2d_{24}+h_3d_{14}+1/(1-\rho)]\ln K, \\
&= \alpha_1+\alpha_2z_t+\alpha_3\ln g_t+\alpha_4\ln K.
\end{aligned}$$

The second line is obtained by applying the first-order Taylor's expansion in the neighborhood of the steady state. Here the h_i's are functions of the deterministic steady-state values $\widetilde{X}^*$, $\tilde{p}^*$, and the coefficient ρ. The third line results from replacing $\ln\hat{p}$ and $\ln X$ by (20) and (21). $\alpha_1=h_1+h_2d_{21}+h_3d_{11}$, $\alpha_2=h_2d_{22}+h_3d_{12}+1/(1-\rho)$, $\alpha_3=h_2d_{23}+h_3d_{13}$, $\alpha_4=h_2d_{24}+h_3d_{14}+1/(1-\rho)$.

Derivation of equation (24):

$$\begin{aligned}
\ln Y &= z+\rho\ln K+(1-\rho)\ln H, \\
&\simeq (1-\rho)\,[h_1+h_2d_{21}+h_3d_{11}] \\
&\quad +\{1+(1-\rho)\,[h_2d_{22}+h_3d_{12}+1/(1-\rho)]\}\,z \\
&\quad +\{(1-\rho)\,[h_2d_{23}+h_3d_{13}]\}\ln g \\
&\quad +\{\rho+(1-\rho)\,[h_2d_{24}+h_3d_{14}+1/(1-\rho)]\}\ln K, \\
&= \beta_1+\beta_2z_t+\beta_3\ln g_t+\beta_4\ln K,
\end{aligned}$$

where $\beta_1=(1-\rho)\,[h_1+h_2d_{21}+h_3d_{11}]$, $\beta_2=\{1+(1-\rho)\,[h_2d_{22}+h_3d_{12}+1/(1-\rho)]\}$, $\beta_3=\{(1-\rho)\,[h_2d_{23}+h_3d_{13}]\}$, $\beta_4=\{\rho+(1-\rho)\,[h_2d_{24}+h_3d_{14}+1/(1-\rho)]\}$.

Derivation of equation (25):

$$\begin{aligned}
\tilde{K}' &\simeq (1-\delta)\tilde{K}+\frac{\exp(\widetilde{X}^*)}{\exp(\widetilde{K}'^*)}\tilde{X} \\
&\quad -\left[(1-\tilde{K}'^*)-(1-\delta)(1-\tilde{K}^*)+\frac{\exp(\widetilde{X}^*)}{\exp(\widetilde{K}'^*)}(1-\tilde{X}^*)\right], \\
&\simeq k_1+k_2\ln K+k_3\ln X, \\
&\simeq k_1+k_3d_{21}+k_3d_{22}z+k_3d_{23}\ln g+(k_2+k_3d_{24})\ln K, \\
&= \gamma_1+\gamma_2z_t+\gamma_3\ln g_t+\gamma_4\ln K_t.
\end{aligned}$$

where $k_1 = \delta [2 - (1 + \delta) \ln K^*]$, $k_2 = (1 - \delta)$, $k_3 = \delta$, $\gamma_1 = k_1 + k_3 d_{21}$, $\gamma_2 = k_3 d_{22}$, $\gamma_3 = k_3 d_{23}$, $\gamma_4 = k_2 + k_3 d_{24}$.

Links of the vectors in (28)-(30) to the parameters:

$$
\begin{aligned}
\delta_0 &= (d_{11} + 1, \ln \delta, -d_{11}, h_1 + h_2 \ln \delta + h_3 d_{11}, (1 - \rho)(h_1 + h_2 \ln \delta + h_3 d_{11}))^{\mathrm{T}}, \\
\delta_1 &= (d_{12}, d_{22}, -d_{12}, h_2 d_{22} + h_3 d_{12} + 1/(1 - \rho), 2 + (1 - \rho)(h_2 d_{22} + h_3 d_{12}))^{\mathrm{T}}, \\
\delta_2 &= (d_{13} + 1, d_{23}, -d_{13}, h_2 d_{23} + h_3 d_{13}, (1 - \rho)(h_2 d_{23} + h_3 d_{13}))^{\mathrm{T}}, \\
\delta_3 &= (1, 0, 0, 0, 0)^{\mathrm{T}}, \\
\delta_4 &= (d_{14}, 1, -d_{14}, h_2 + h_3 d_{14} + \rho/(1 - \rho), 2\rho + (1 - \rho)(h_2 + h_3 d_{14}))^{\mathrm{T}}.
\end{aligned}
$$

References

Bierens, H. J., 2006, Econometric Analysis of Linearized Singular Dynamic Stochastic General Equilibrium Models", Journal of Econometrics 236, 595-627.

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.

Burnside, C. M., M. Eichenbaum, and S. Rebelo, 1993, Labor hoarding and the business cycle, Journal of Political Economy 101, 245-273.

Christiano, L. J., 1990, Solving the stochastic growth model by linear-quadratic approximation and by value-function iteration, Journal of Business and Economic Statistics 8, 23-26.

Christiano, L.J. and M. Eichenbaum, 1992, Current business cycle theories and aggregate labor market fluctuations, American Economic Review 82, 430-50.

Cooley, T. F. and G. D. Hansen, 1989, The inflation tax in a real business cycle, American Economic Review 79, 733-748.

Del Negro and Schorfheide F. ,2004, Priors from general equilibrium models for VARs, International Economic Review 45, 643-673.

Del Negro, M., Schorfheide F., Smets F. and R. Wouters, 2006, On the fit and Forecasting Performance of New-Keynesian Models, ECB working paper 491.

DeJong, D. N., B. F. Ingram, and C. H. Whiteman, 1996, A bayesian approach to calibration, Journal of Business and Economic Statistics 14, 1-9.

DeJong, D. N., B. F. Ingram, and C. H. Whiteman, 2000, A bayesian approach to dynamic macroeconomics, Journal of Econometrics 98, 203-223.

Geweke, J. , 1999, Computational experiments and reality, Department of Economics, University of Minnesota and Federal Reserve Bank of Minneapolis, working paper.

Hannan, E. J., and B. G. Quinn, 1979, The determination of the order of an autoregression, Journal of the Royal Statistical Society, Series B 41, 190-195.

Hansen, G. D. ,1985, Indivisible labor and the business cycle, Journal of Monetary Economics 16, 309-327.

Hansen, G.D. and E.C. Prescott, 1995, Recursive methods for computing equilibria of business cycle models, in: T.C. Cooley, eds., Frontiers of business cycle research, (Princeton University Press, Princeton, NJ.) 39-64.

Ireland, P. N., 2004, A method for taking models to the data, Journal of Economic Dynamics and Control 28, 1205-1226.

Johansen, S., 1995, Likelihood-based inference in cointegration vector autoregressive models, Oxford University Press, Oxford, U.K.

Johansen, S., 1988, Statistical analysis of cointegration vectors, Journal of Economic Dynamics and Control 12, 231-254.

King, R. G., C. I. Plosser, and S. T. Rebelo, 2001, Production, growth and business cycles: Technical appendix, Computational Economics.

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.

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.

Kydland, F. E., and E. C. Prescott, 1982, Time to build and aggregate fluctuations, Econometrica 50, 1345-1370.

Lutkepohl, H., 1993, Introduction to multiple time series analysis (Springer-Verlag, Berlin).

McGrattan, E. R., 1990, Solving the stochastic growth model by linear-quadratic approximation, Journal of Business and Statistics 8, 41-44.

McGrattan, E. R., R. Rogerson and R. Wright, 1997, An equilibrium model of the business cycle with household production and fiscal policy, International Economic Review 38, 267-290.

Schorfheide, F., 2000, Loss function-based evaluation of DSGE models, Journal of Applied Econometrics 15, 645-670.

Schwarz, G., 1978, Estimating the dimension of a model, Annals of Statistics 6, 461-464.