# Bayesian Estimation of DSGE Models: Lessons from Second-Order Approximations

Sungbae An*

*Singapore Management University*

May 2007

**Abstract**

This paper investigates a general procedure to estimate second-order approximations to a DSGE model and compares the performance with the widely used estimation technique for a log-linearized economy on a version of new Keynesian monetary model. It is done in the context of posterior distributions, welfare cost, and impulse response analysis. Our findings include the followings. First, we find that all the results of An and Schorfheide (2007) are confirmed with U.S. data. With the nonlinear estimation we can identify parameters that are neglected previously; the marginal data density evaluation shows that data support the nonlinear estimation procedure; and parameter estimates that are related to nondeterministic steady states are quite different from the linear estimates. Second, the estimated welfare differentials are more aggressive for the second-order approximations, that is, the posterior welfare differentials from the linear estimation may underestimate the welfare cost resulted from changes in the monetary policy. Third, the second-order approximation unveils quite different dynamics which are neglected in a log-linearized economy.

*JEL Classification:* C11, C32, C51, C52, E52
*Keywords:* Bayesian Analysis, DSGE Models, Second-order Approximation, Particle Filter

---

*School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903. E-mail: sungbae@smu.edu.sg. I am grateful to Frank Schorfheide for his guidance and advice. Francis X. Diebold and Jesús Fernández-Villaverde provided many helpful comments, as did seminar participants at University of Pennsylvania, University of Rochester, Hong Kong University of Science and Technology, Singapore Management University, Bank of Canada, Board of Governors of the Federal Reserve System, and Indiana University. All errors are mine.

# 1 Introduction

It is by now widespread practice to estimate log-linearized dynamic stochastic general equilibrium (DSGE) models with Bayesian methods. While log-linear approximations are sufficiently accurate in many applications, the recent literature on the policy analysis shows that the welfare comparisons based on first-order approximations can be misleading. Kim and Kim (2003) have first pointed out that second-order perturbation methods lead to more accurate approximations of welfare measures. For example, when the second-order approximation to the welfare function involves linear terms in addition to quadratic ones, which is frequent for widely used models such as Smets and Wouters (2003) and Christiano, Eichenbaum, and Evans (2005), the log-linearization will not yield the same level of welfare as that from the second-order approximation.

Since the seminal paper by Rotemberg and Woodford (1997), the utility-based metric has played a central role in the evaluation of the monetary policy of dynamic stochastic general equilibrium (DSGE) models. It has several advantages over the alternative approach with ad hoc loss functions to assess policies. Woodford (2003, Chapter 6) shows that an ad hoc loss function can be rationalized as a welfare metric, that is, evaluations of a second-order approximation of the utility function depend only on the quadratic terms of the expansion as long as the steady state of the model is in the first best equilibrium. However, this approach has been confined to very simplified economies that satisfy the following: (1) the government holds the balanced budget at all times by means of lump-sum taxation, that is, the fiscal policy is non-distorting and passive, (2) the government has access to a subsidy which offsets the steady-state effects of monopolistic distortions, and (3) there is no capital accumulation in the economy.

With an advent of the computational algorithm developed by Schmitt-Grohé and Uribe (2006b), Kim, Kim, Schaumburg, and Sims (2005), and Swanson, Anderson, and Levin (2005), second-order accurate approximations to the model and the welfare metric are available, and they lead to the analysis of the optimal policy rules. See for example, Laforte (2004), Levin and López-Salido (2004), Levin, Onatski, Williams, and Williams (2006), and Schmitt-Grohé and Uribe (2006a).

Despite the rising interest in the second-order perturbation methods, the likelihood-based estimation of the second-order approximated economy has not been executed because of the computational difficulties. The likelihood cannot be constructed with the Kalman filter that is readily available for the log-linearized economy.[1] Instead, researchers have taken the short cut that they

[1]The Kalman filter is widely used to evaluate the likelihood of the models with linear policy functions. See, for example, Sargent (1989) and McGrattan (1994) for the maximum likelihood estimation, Schorfheide (2000), Otrok (2001), DeJong, Ingram, and Whiteman (2000), and Smets and Wouters (2003) for the Bayesian estimation.

first estimate the log-linearized economy and plug the estimates into the second-order approximated economy. See, for example, Laforte (2004) and Levin, Onatski, Williams, and Williams (2006).

Can this practice be misleading? To answer this question, we consider a particle filter in constructing the likelihood of the second-order approximated DSGE models. Because the second-order approximation results in a nonlinear and non-Gaussian state space representation, numerical methods are exploited to integrate out the latent states.[2] Gordon, Salmond, and Smith (1993) and Kitagawa (1996) are the early contributors to the literature.[3] In economics, the particle filter has been applied to analyze the stochastic volatility models by Pitt and Shephard (1999) and Kim, Shephard, and Chib (1998). In the series of paper, Fernández-Villaverde and Rubio-Ramírez (2004, 2005, 2006) first introduce the particle filter to investigate the nonlinear nature of DSGE models.

An and Schorfheide (2007) apply our method to a version of the new Keynesian monetary model in line with Woodford (2003), which is widely used to compare monetary policy rules in terms of the welfare loss criterion based on the agent's utility. They study the performance of the procedure with artificial data which are generated based on the second-order approximated model economy with the calibration that matches real U.S. data, and linear (log-linearization with the Kalman filter) and quadratic (second-order approximation with the particle filter) estimation methods are applied to it. They found that (a) the quadratic estimation method can identify more structural parameters that are unidentifiable under the linear method, even though the identifications are weak; (b) the nonlinear estimation provides a better fit of the model as measured by the marginal data density; and (c) the posterior constructed from the second-order approximation tends to attain true values for the steady state parameters. The latter two findings are in line with those from Fernández-Villaverde and Rubio-Ramírez (2005).

In this paper, we apply this nonlinear approach to real U.S. data and investigate its effect on welfare evaluations. We further investigate the dynamics of the model with exogenous shock processes. Since our approximated model is nonlinear, the generalized impulse responses proposed by Koop, Pesaran, and Potter (1996) are considered. Our findings include the followings. First, we find that all the results of An and Schorfheide (2007) are confirmed with U.S. data. With the nonlinear estimation we can identify parameters that are neglected previously; the marginal data density evaluation shows that data support the nonlinear estimation procedure; and parameter estimates that are related to nondeterministic steady states are quite different from the linear estimates. Second, the estimated welfare differentials are more aggressive for the second-order approximations, that is, the posterior welfare differentials from the linear estimation may underestimate the welfare

[2]For that reason, the particle filter is also known as the sequential Monte Carlo filter.

[3]See Arulampalam, Maskell, Gordon, and Clapp (2002) for an excellent survey.

cost resulted from changes in the monetary policy. Third, the second-order approximation unveils quite different dynamics which are neglected in a log-linearized economy.

The remainder of the paper is organized as follows. Section 2 contains a brief description of the DSGE model that we are using. Section 3 discusses various solution methods to approximate DSGE model equilibria. Section 4 introduces the estimation procedure including the state space representation, the particle filter, and the specification of priors. Section 5 gives the empirical results using U.S. data. Section 6 concludes.

# 2 The Model

The model economy consists of a firm producing a final good, a continuum of firms producing intermediate goods, households, and a monetary as well as a fiscal authority. This model has become a benchmark specification for an analysis of monetary policies and is analyzed in detail, for instance, in Woodford (2003). We abstract wage rigidities and capital accumulation. More elaborate versions of the model can be found in Smets and Wouters (2003) and Christiano, Eichenbaum, and Evans (2005).

## 2.1 The Agents and Their Decision Problems

The perfectly competitive final good producing firm combines a continuum of intermediate goods indexed by $j \in [0, 1]$ using the technology

$$Y_t = \left( \int_0^1 Y_t(j)^{1-\nu} dj \right)^{\frac{1}{1-\nu}}. \tag{1}$$

Here $1/\nu > 1$ represents the elasticity of demand for each intermediate good. The firm takes input prices $P_t(j)$ and output prices $P_t$ as given. Profit maximization implies that the demand for intermediate goods is

$$Y_t(j) = \left( \frac{P_t(j)}{P_t} \right)^{-1/\nu} Y_t. \tag{2}$$

The relationship between intermediate goods prices and the price of the final good is

$$P_t = \left( \int_0^1 P_t(j)^{\frac{\nu-1}{\nu}} dj \right)^{\frac{\nu}{\nu-1}}. \tag{3}$$

Intermediate good $j$ is produced by a monopolist who has access to the following linear production technology:

$$Y_t(j) = A_t N_t(j), \tag{4}$$

where $A_t$ is an exogenous productivity process that is common to all firms and $N_t(j)$ is the labor input of firm $j$. Labor is hired in a perfectly competitive factor market at the real wage $W_t$. Firms face nominal rigidities in terms of quadratic price adjustment costs

$$AC_t(j) = \frac{\varphi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - \pi\right)^2 Y_t(j), \tag{5}$$

where $\varphi$ governs the price stickiness in the economy and $\pi$ is the steady state inflation rate associated with the final good. Firm $j$ chooses its labor input $N_t(j)$ and the price $P_t(j)$ to maximize the present value of the future profits

$$\mathbf{E}_t\left[\sum_{s=0}^{\infty}\beta^s Q_{t+s}\left(\frac{P_{t+s}(j)}{P_{t+s}}Y_{t+s}(j) - W_{t+s}N_{t+s}(j) - AC_{t+s}(j)\right)\right]. \tag{6}$$

Here, $Q_{t+s}$ is the marginal value of a unit of the consumption good to the household, which is treated as exogenous by the firm.

The representative household derives utility from real money balances $M_t/P_t$ and consumption $C_t$ relative to a habit stock. We assume that the habit stock is given by the level of technology $A_t$. This assumption assures that the economy evolves along a balanced growth path even if the utility function is additively separable in consumption, real money balances, and leisure. The household derives disutility from hours worked $H_t$ and maximizes

$$\mathbf{E}_t\left[\sum_{s=0}^{\infty}\beta^s\left(\frac{(C_{t+s}/A_{t+s})^{1-\tau} - 1}{1-\tau} + \chi_M \ln\left(\frac{M_{t+s}}{P_{t+s}}\right) - \chi_H H_{t+s}\right)\right], \tag{7}$$

where $\beta$ is the discount factor, $1/\tau$ is the intertemporal elasticity of substitution, and $\chi_M$ and $\chi_H$ are the scale factors that determine the steady state real money balances and the hours worked. The household supplies perfectly elastic labor services to the firms taking the real wage $W_t$ as given. The household has access to a domestic bond market where nominal government bonds are traded that pay (gross) interest $R_t$. Furthermore, it receives aggregate residual real profits $D_t$ from the firms and has to pay lump-sum taxes $T_t$. Thus, the households' budget constraint is of the form

$$P_t C_t + B_t + M_t - M_{t-1} + T_t \quad = \quad P_t W_t H_t + R_{t-1} B_{t-1} + P_t D_t \tag{8}$$

The usual transversality condition on asset accumulation applies which rules out the Ponzi schemes.

The monetary policy is described by an interest rate feedback rule of the form

$$R_t = R_t^{*\,1-\rho_R} R_{t-1}^{\rho_R} e^{\epsilon_{R,t}}, \tag{9}$$

where $\epsilon_{R,t}$ is a monetary policy shock and $R_t^*$ is the (nominal) target rate. We consider an output gap rule as a specification for $R_t^*$, in which the central bank reacts to inflation and deviations of output from potential output:

$$R_t^* = r\pi^* \left(\frac{\pi_t}{\pi^*}\right)^{\psi_1} \left(\frac{Y_t}{Y_t^*}\right)^{\psi_2} \tag{10}$$

Here $r$ is the steady state real interest rate, $\pi_t$ is the gross inflation rate defined as $\pi_t = P_t/P_{t-1}$, and $\pi^*$ is the target inflation rate, which in equilibrium coincides with the steady state inflation rate. $Y_t^*$ in (10) is the level of the output that would prevail in the absence of nominal rigidities.

The fiscal authority consumes a fraction $\zeta_t$ of the aggregate output $Y_t$, where $\zeta_t \in [0, 1]$ follows an exogenous process. The government levies a lump-sum tax (subsidy) to finance any shortfalls in the government revenues (or to rebate any surplus). The government's budget constraint is given by

$$P_t G_t + R_{t-1} B_{t-1} \quad = \quad T_t + B_t + M_t - M_{t-1}. \tag{11}$$

## 2.2 Exogenous Processes

The model economy is perturbed by three exogenous processes. Aggregate productivity evolves according to

$$\ln A_t = \ln \gamma + \ln A_{t-1} + \ln z_t, \quad \text{where} \quad \ln z_t = \rho_z \ln z_{t-1} + \epsilon_{z,t}. \tag{12}$$

Define $g_t = 1/(1-\zeta_t)$. We assume that

$$\ln g_t = (1-\rho_g) \ln g^* + \rho_g \ln g_{t-1} + \epsilon_{g,t}. \tag{13}$$

Finally, the monetary policy shock $\epsilon_{R,t}$ is assumed to be serially uncorrelated. The three innovations are independent of each other at all leads and lags and are normally distributed with means zero and standard deviations $\sigma_z$, $\sigma_g$, $\sigma_R$, respectively.

## 2.3 Equilibrium Relationships

We consider the symmetric equilibrium in which all intermediate goods producing firms make identical choices so that the $j$ subscript can be omitted. The market clearing conditions are given by

$$Y_t = C_t + G_t + AC_t \quad \text{and} \quad H_t = N_t. \tag{14}$$

It can be shown that output, consumption, interest rates, and inflation have to satisfy the following optimality conditions

$$\begin{aligned}
1 &= \beta \mathbf{E}_t \left[ \left( \frac{C_{t+1}/A_{t+1}}{C_t/A_t} \right)^{-\tau} \frac{A_t}{A_{t+1}} \frac{R_t}{\pi_{t+1}} \right] && (15)\\
1 &= \frac{1}{\nu} \left[ 1 - \chi_H \left( \frac{C_t}{A_t} \right)^{\tau} \right] + \varphi(\pi_t - \pi) \left[ \left( 1 - \frac{1}{2\nu} \right) \pi_t + \frac{\pi}{2\nu} \right] && (16)\\
&\quad - \varphi \beta \mathbf{E}_t \left[ \left( \frac{C_{t+1}/A_{t+1}}{C_t/A_t} \right)^{-\tau} \frac{Y_{t+1}/A_{t+1}}{Y_t/A_t} (\pi_{t+1} - \pi) \pi_{t+1} \right].
\end{aligned}$$

In the absence of nominal rigidities ($\varphi = 0$), the aggregate output is given by

$$Y_t^* = \left( \frac{1-\nu}{\chi_H} \right)^{1/\tau} A_t g_t, \tag{17}$$

which is the target level of output that appears in the output gap rule specification.

Since the non-stationary technology process $A_t$ induces a stochastic trend in output and consumption, it is convenient to express the model in terms of detrended variables $c_t = C_t/A_t$ and $y_t = Y_t/A_t$. The model economy has a unique steady state in terms of the detrended variables that is attained if the innovations $\epsilon_{R,t}$, $\epsilon_{g,t}$, and $\epsilon_{z,t}$ are zero at all times. The steady state inflation $\pi$ equals the target rate $\pi^*$ and

$$r = \frac{\gamma}{\beta}, \quad R = r\pi^*, \quad c = \left( \frac{1-\nu}{\chi_H} \right)^{1/\tau}, \quad \text{and} \quad y = g \left( \frac{1-\nu}{\chi_H} \right)^{1/\tau}. \tag{18}$$

Let $\hat{x}_t = \ln(x_t/x)$ denote the percentage deviation of a variable $x_t$ from its steady state $x$. Then the model can be expressed as

$$\begin{aligned}
1 &= \mathbf{E}_t \left[ e^{-\tau \hat{c}_{t+1} + \tau \hat{c}_t + \hat{R}_t - \hat{z}_{t+1} - \hat{\pi}_{t+1}} \right] && (19)\\
\frac{1-\nu}{\nu \varphi \pi^2} \left( e^{\tau \hat{c}_t} - 1 \right) &= \left( e^{\hat{\pi}_t} - 1 \right) \left[ \left( 1 - \frac{1}{2\nu} \right) e^{\hat{\pi}_t} + \frac{1}{2\nu} \right] && (20)\\
&\quad - \beta\, \mathbf{E}_t \left[ \left( e^{\hat{\pi}_{t+1}} - 1 \right) e^{-\tau \hat{c}_{t+1} + \tau \hat{c}_t + \hat{y}_{t+1} - \hat{y}_t + \hat{\pi}_{t+1}} \right] \\
e^{\hat{c}_t - \hat{y}_t} &= e^{-\hat{g}_t} - \frac{\varphi \pi^2 g}{2} \left( e^{\hat{\pi}_t} - 1 \right)^2 && (21)\\
\hat{R}_t &= \rho_R \hat{R}_{t-1} + (1-\rho_R)\psi_1 \hat{\pi}_t + (1-\rho_R)\psi_2 \left( \hat{y}_t - \hat{g}_t \right) + \epsilon_{R,t} && (22)\\
\hat{g}_t &= \rho_g \hat{g}_{t-1} + \epsilon_{g,t} && (23)\\
\hat{z}_t &= \rho_z \hat{z}_{t-1} + \epsilon_{z,t}. && (24)
\end{aligned}$$

# 3 Solution Methods

We begin by presenting the canonical form for our first- and second-order approximated models. Also, the welfare measure is defined as an infinite sum of the household's discounted utility and we show how to get the second-order approximation to the level of welfare in line with Schmitt-Grohé and Uribe (2004), and its relation to the ad hoc loss function. Finally, the impulse responses for nonlinear processes are considered to emphasize the differences in dynamics between the first- and the second-order approximated models.

## 3.1 Model Solutions

Equations (19) to (24) form a nonlinear rational expectations system in the variables $\hat{y}_t$, $\hat{c}_t$, $\hat{\pi}_t$, $\hat{R}_t$, $\hat{g}_t$, and $\hat{z}_t$ that are driven by the vector of innovations $\epsilon_t = [\epsilon_{R,t}, \epsilon_{g,t}, \epsilon_{z,t}]'$. This rational expectations system has to be solved before the DSGE model can be estimated. Define

$$s_t = [\hat{y}_t, \hat{c}_t, \hat{\pi}_t, \hat{R}_t, \epsilon_{R,t}, \hat{g}_t, \hat{z}_t]'.$$

The solution of the rational expectations system takes the form

$$s_t = \Phi(s_{t-1}, \epsilon_t; \theta). \tag{25}$$

From an econometric perspective, $s_t$ can be viewed as a (partially latent) state vector in a state space representation and (25) is the state transition equation.

A variety of numerical techniques are available to solve rational expectations systems. In the context of likelihood-based DSGE model estimation, linear approximation methods are very popular because they lead to a state-space representation of the DSGE model that can be analyzed with the Kalman filter. A linearization of Equations (19) to (21) yields

$$\hat{c}_t = \mathbf{E}_t\left[\hat{c}_{t+1}\right] - \frac{1}{\tau}\left(\hat{R}_t - \mathbf{E}\left[\hat{\pi}_{t+1}\right] - \mathbf{E}\left[\hat{z}_{t+1}\right]\right) \tag{26}$$

$$\hat{\pi}_t = \beta\mathbf{E}_t[\hat{\pi}_{t+1}] + \kappa\hat{c}_t \tag{27}$$

$$\hat{y}_t = \hat{c}_t + \hat{g}_t, \tag{28}$$

where

$$\kappa = \tau\frac{1-\nu}{\nu\pi^2\varphi}. \tag{29}$$

Equations (26) to (28) combined with (22) to (24) form a linear rational expectations system in $s_t$ for which several solution algorithms are available, for instance, Blanchard and Kahn (1980), Binder and Pesaran (1997), King and Watson (1998), Uhlig (1999), Anderson (2000), Klein (2000), Christiano (2002), and Sims (2002). Depending on the parameterization of the DSGE model, there are three possibilities: (1) no stable rational expectations solution exists, (2) the stable solution is unique (determinacy), or (3) there are multiple stable solutions (indeterminacy). We will focus on the case of the determinacy and restrict the parameter space accordingly. The resulting law of motion for the $j$'th element of $s_t$ takes the form

$$s_{j,t} = \sum_{i=1}^{J} \Phi_{j,i}^{(s)} s_{i,t-1} + \sum_{l=1}^{n} \Phi_{j,l}^{(\epsilon)} \epsilon_{l,t}, \quad j = 1, \ldots, J. \tag{30}$$

Here the coefficients $\Phi_{j,i}^{(s)}$ and $\Phi_{j,l}^{(\epsilon)}$ are functions of the structural parameters of the DSGE model.

While in many applications first-order approximations are sufficient, higher-order refinements are an active area of research. For instance, if the goal of the analysis is to compare welfare across policies that do not have first-order effects on the model's steady state or to study asset pricing implications of DSGE models, a more accurate solution may be necessary. See for instance, Kim and Kim (2003) or Woodford (2003).

A second-order accurate solution to the DSGE model can be obtained from the second-order perturbation of the equilibrium conditions (19) to (24). Algorithms to construct such solutions have been developed by Judd (1998), Collard and Juillard (2001), Jin and Judd (2002), Schmitt-Grohé and Uribe (2004), Kim, Kim, Schaumburg, and Sims (2005), Swanson, Anderson, and Levin (2005), and Klein (2005). The resulting state transition equation can be expressed as

$$\begin{aligned} s_{j,t} \;=\;& \Phi_j^{(0)} + \sum_{i=1}^{J} \Phi_{j,i}^{(s)} s_{i,t-1} + \sum_{l=1}^{n} \Phi_{j,l}^{(\epsilon)} \epsilon_{l,t} + \sum_{i=1}^{J} \sum_{l=1}^{J} \Phi_{j,il}^{(ss)} s_{i,t-1} s_{l,t-1} \\ &+ \sum_{i=1}^{J} \sum_{l=1}^{n} \Phi_{j,il}^{(s\epsilon)} s_{i,t-1} \epsilon_{l,t} + \sum_{i=1}^{n} \sum_{l=1}^{n} \Phi_{j,il}^{(\epsilon\epsilon)} \epsilon_{i,t} \epsilon_{l,t}. \end{aligned} \tag{31}$$

As before, the coefficients are functions of the parameters of the DSGE model. For the subsequent analysis we use Klein's (2000) procedure to compute a first-order accurate solution of the DSGE model and Schmitt-Grohé and Uribe's (2004)'s algorithm to obtain a second-order accurate solution.

Noting that a second-order approximation is an extension of a first-order approximation, i.e., linear terms in (30) and (31) match exactly, a higher-order perturbation is readily available and based on the same method with the higher-order Taylor expansion. While the perturbation methods approximate policy functions locally, there are several global approximation schemes including the projection methods such as the finite element method and the Chebyshev polynomial method on the spectral domain. Judd (1998) covers various solution methods, and Taylor and Uhlig (1990) and Den Haan and Marcet (1994) compare the accuracy of alternative solution methods with a first-order perturbation method. Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) find that there exist compelling advantages to change some computations currently undertaken with a linear method to a second-order approximation, while higher-order perturbations display a much superior performance.

### 3.2 Welfare Analysis

The welfare associated with a particular monetary policy can be measured once the model parameter $\theta$ is determined. Following Schmitt-Grohé and Uribe (2006a) we use the conditional expectation of the household's lifetime utility as of time zero as our welfare measure. We further assume that all the variables are in their deterministic steady states at time zero. Let $C_t^a$ and $H_t^a$ follow an equilibrium path of consumption and hours of labor under a particular policy (given parameter $\theta$), then our welfare measure is given by

$$W_0 = \mathbf{E}_0 \left[ \sum_{t=0}^{\infty} \beta^t \left( \frac{(C_t^a/A_t)^{1-\tau} - 1}{1-\tau} - \chi_H H_t^a \right) \right]. \quad (32)$$

As mentioned earlier, the log-linear approximation to (32) around the steady state does not depend on the policy parameters, $\psi_1$ and $\psi_2$, and hence a higher-order approximation is necessary to compare welfare across different policies.

Several approaches have been proposed to evaluate $W_t$ in the literature. Woodford (2003) shows that linear approximation to the policy functions is sufficient to accurately approximate welfare up to the second order if the second-order approximation to the welfare function contains quadratic terms only. In cases the above condition does not hold, Kim, Kim, Schaumburg, and Sims (2005) and Laforte (2004) use second-order accurate solutions to carry out welfare calculations. A recent advance in computational algorithms, developed by Schmitt-Grohé and Uribe (2004), has delivered a simple way to evaluate any higher-order approximation of welfare. It exploits the recursive nature of (32) and is made possible by augmenting the state space with $W_t$. Noting that the welfare measure is the infinite sum of the discounted utilities and that $H_t = y_t$ in equilibrium we have

$$\begin{aligned} W_t &= \mathbf{E}_t \left[ \sum_{s=0}^{\infty} \beta^s \left( \frac{c_{t+s}^{1-\tau} - 1}{1-\tau} - \chi_H y_{t+s} \right) \right] \\ &= \mathbf{E}_t \left[ \sum_{s=0}^{\infty} \beta^s \left( \frac{c_{t+s}^{1-\tau} - 1}{1-\tau} \right) \right] - \chi_H \mathbf{E}_t \left[ \sum_{s=0}^{\infty} \beta^s y_{t+s} \right] \\ &= U_t - \chi_H V_t. \end{aligned}$$

Moreover, recursive representations of value functions, $U_t$ and $V_t$, are

$$\begin{aligned} U_t &= \frac{c_t^{1-\tau}-1}{1-\tau} + \beta \mathbf{E}_t U_{t+1}, \\ V_t &= y_t + \beta \mathbf{E}_t V_{t+1}. \end{aligned}$$

Let $u_t = e^{U_t}$ and $v_t = e^{V_t}$. Then the second-order accurate welfare approximations can be obtained by adding the following equations into the system (19) – (24)

$$\hat{u}_t = \frac{c^{1-\tau}}{1-\tau}\left(e^{(1-\tau)\hat{c}_t} - 1\right) + \beta \mathbf{E}_t \hat{u}_{t+1}, \quad (33)$$

$$\hat{v}_t = y\left(e^{\hat{y}_t} - 1\right) + \beta \mathbf{E}_t \hat{v}_{t+1} \quad (34)$$

and using the identity $W_t = W + \hat{u}_t - \chi_H \hat{v}_t$ where $W$ denotes the steady-state level of welfare.

Suppose that we have two sets of parameters describing the economy, reference $\theta^R$ and alternative $\theta^A$, and different levels of welfare are implied by each set of parameters. More often than not, the level of welfare itself has little economic meaning, and we are interested in the fraction of consumption stream from alternative to be added (or subtracted) to achieve the reference level of welfare, $W_t^R$. That is, we measure the welfare cost $\lambda$ satisfying

$$\begin{aligned} W_t^R &= \mathbf{E}_t\left[\sum_{s=0}^{\infty}\left(\beta^A\right)^s\left(\frac{\left((1+\lambda)C_{t+s}^A/A_{t+s}^A\right)^{1-\tau^A}-1}{1-\tau^A} - \chi_H^A H_{t+s}^A\right)\right] \\ &= (1+\lambda)^{1-\tau^A} U_t^A + \frac{1}{1-\beta^A}\frac{(1+\lambda)^{1-\tau^A}-1}{1-\tau^A} - \chi_H^A V_t^A \end{aligned}$$

and therefore,

$$\lambda = \left[\frac{1+(1-\beta^A)(1-\tau^A)(W_t^R+\chi_H^A V_t^A)}{1+(1-\beta^A)(1-\tau^A)(W_t^A+\chi_H^A V_t^A)}\right]^{\frac{1}{1-\tau^A}} - 1.$$

### 3.3 Impulse Response Analysis

The impulse response has been a key instrument for exploring the dynamics of DSGE and VAR models. It traces out the response pattern of a model to exogenous shocks. While the impulse response is straightforward for the linearized version of the economy, Gallant, Rossi, and Tauchen (1993) and Koop, Pesaran, and Potter (1996) have pointed out that nonlinear impulse responses essentially differ from linear ones in several aspects. First, the effect of the current shock is neither symmetric nor scalable. In linear models, the opposite shock with the same magnitude has exactly the opposite effect on the response, and moreover, a shock with double magnitude has twice as much effect. Second, nonlinear impulse responses are path dependent, while linear impulse responses are independent of history of the process. They are affected not only by current shock but also by the consecutive shocks.

Following Koop, Pesaran, and Potter (1996), the generalized impulse response (GI) is defined as

$$GI(n) \quad = \quad \mathbf{E}\left[x_{t+n}|\epsilon_t, \Omega_{t-1}\right] - \mathbf{E}\left[x_{t+n}|\Omega_{t-1}\right],$$

in which the consecutive shocks are integrated out. Rewriting (30) as

$$s_t \quad = \quad \Phi^{(s)} s_{t-1} + \Phi^{(\epsilon)} \epsilon_t \tag{35}$$

it immediately follows that

$$GI^L(n) \quad = \quad \left(\Phi^{(s)}\right)^n \Phi^{(\epsilon)} \epsilon_t$$

which coincides with the impulse response when $\epsilon_t$ takes a particular value, usually the unit standard deviation. For the second-order approximation we rewrite the state transition equation (31) as follows. Decompose $s_t = [x_t', z_t']'$, where $z_t$ (not to be confused with the technology growth process in Section 2) is composed of the exogenous processes $\epsilon_{R,t}$, $\hat{g}_t$, and $\hat{z}_t$, which follows a linear autoregressive structure

$$z_t \quad = \quad \Psi^{(z)} z_{t-1} + \Psi^{(\epsilon)} \epsilon_t. \tag{36}$$

The law of motion for the endogenous state variables $x_j$ can be expressed as

$$x_{j,t} = \Psi_j^{(0)} + \Psi_j^{(1)} w_t + w_t' \Psi_j^{(2)} w_t \tag{37}$$

where $w_t = [x_{t-1}', z_t']'$. With conformable partition of $\Psi_j^{(1)}$ and $\Psi_j^{(2)}$ and (36), we can write (37) as

$$\begin{aligned} x_{j,t} \quad = \quad & \Psi_j^{(0)} + \Psi_{j1}^{(1)} x_{t-1} + \Psi_{j2}^{(1)} \Psi^{(z)} z_{t-1} + \Psi_{j2}^{(1)} \Psi^{(\epsilon)} \epsilon_t \\ & + \operatorname{tr}\left(\Psi_{j11}^{(2)} x_{t-1} x_{t-1}'\right) + 2\operatorname{tr}\left(\Psi^{(z)\prime} \Psi_{j21}^{(2)} x_{t-1} z_{t-1}'\right) + \operatorname{tr}\left(\Psi^{(z)\prime} \Psi_{j22}^{(2)} \Psi^{(z)} z_{t-1} z_{t-1}'\right) \\ & + 2\operatorname{tr}\left(\Psi^{(\epsilon)\prime} \Psi_{j21}^{(2)} x_{t-1} \epsilon_t'\right) + \operatorname{tr}\left(\Psi^{(\epsilon)\prime} \Psi_{j22}^{(2)} \Psi^{(z)} z_{t-1} \epsilon_t'\right) + \operatorname{tr}\left(\Psi^{(\epsilon)\prime} \Psi_{j22}^{(2)} \Psi^{(\epsilon)} \epsilon_t \epsilon_t'\right). \end{aligned}$$

Noting that we start from the steady state of the economy, $s_{t-k} = 0, k = 1, 2, \ldots$, the generalized impulse response at the initial horizon is

$$GI_j^Q(0) \quad = \quad \Psi_{j2}^{(1)} \Psi^{(\epsilon)} \epsilon_t + \operatorname{tr}\left(\Psi^{(\epsilon)\prime} \Psi_{j22}^{(2)} \Psi^{(\epsilon)} (\epsilon_t \epsilon_t' - I)\right),$$

from which it is obvious that the impulse response for the second-order approximated economy is neither symmetric nor scalable even at the initial horizon.

# 4 Estimation Methods

We focus on Bayesian estimations of DSGE models, which have three main characteristics. First, unlike a GMM estimation based on equilibrium relationships such as the consumption Euler equation (19), the price setting equation of the intermediate goods producing firms (20), or the monetary policy rule (22), the Bayesian analysis is system-based and fits the solved DSGE model to a vector of aggregate time series. Second, the estimation is based on the likelihood function generated by the DSGE model rather than, for instance, the discrepancy between the DSGE model responses and the VAR impulse responses. Third, prior distributions can be used to incorporate additional information into the parameter estimation.

In what follows, we specify the measurement equations to complete the state space representation, of which the likelihood can be evaluated with two kinds of filters. While the Kalman filter is used for the linear transition (35), the particle filter for the quadratic transitions (36) and (37) is considered to construct the likelihood and explained in detail. With appropriate prior distributions, the posterior draws are to be generated using the Metropolis-Hastings algorithm.

## 4.1 Measurement Equations

Our model is completed by defining a set of measurement equations that relate the elements of $s_t$ to a set of observables. We assume that we have observations on the quarter-to-quarter per capita GDP growth rates (YGR), the annualized quarter-to-quarter inflation rates (INFL), and the annualized nominal interest rates (INT) in percentages:

$$\begin{aligned} \mathrm{YGR}_t &= \gamma^{(Q)} + 400 \times (\hat{y}_t - \hat{y}_{t-1} + \hat{z}_t) + u_{y,t} \\ \mathrm{INFL}_t &= \pi^{(A)} + 400 \times \hat{\pi}_t + u_{\pi,t} \\ \mathrm{INT}_t &= \pi^{(A)} + r^{(A)} + 4 \times \gamma^{(Q)} + 400 \times \hat{R}_t + u_{r,t}. \end{aligned} \tag{38}$$

with measurement errors whose standard deviations are $\sigma_y$, $\sigma_\pi$, and $\sigma_r$, respectively. The parameters $\gamma^{(Q)}$, $\pi^{(A)}$, and $r^{(A)}$ are related to the steady states of the model economy as follows

$$\gamma = e^{\gamma^{(Q)}/100}, \quad \beta = \frac{1}{e^{r^{(A)}/400}}, \quad \pi = e^{\pi^{(A)}/400}.$$

The structural parameters are collected in the vector $\theta$. Since in the first-order approximation the parameters $\nu$ and $\varphi$ are not separately identifiable, we express the model in terms of $\kappa$, defined in (29). Let

$$\theta = \left[\tau, \kappa, \psi_1, \psi_2, \rho_R, \rho_g, \rho_z, r^{(A)}, \pi^{(A)}, \gamma^{(Q)}, \sigma_R, \sigma_g, \sigma_z\right]'.$$

For the quadratic approximation, the composite parameter $\kappa$ will be replaced by either $(\varphi, \nu)$ or alternatively by $(\kappa, \nu)$. Moreover, $\theta$ will be augmented with the steady state consumption-output ratio $c/y$ which is equal to $1 - 1/g$.

## 4.2 Posterior Computations and the Likelihoods

The inference is based on the likelihood, $\mathcal{L}(\theta|Y)$. However, the evaluation of it is not trivial. When the transition (25) is approximated linearly with a Gaussian error as in (35), the Kalman filter is applied to deliver the likelihood.

For the second-order approximated transition (36) and (37), however, the Kalman filter cannot be used to compute the likelihood. The problem arises not only from the nonlinear nature of the state variables but also from the non-Gaussian error forms. To see this, rearrange (31) as

$$\begin{aligned} s_{j,t} &= \sum_{i=1}^{n}\sum_{l=1}^{n}\Phi_{j,il}^{(\epsilon\epsilon)}\epsilon_{i,t}\epsilon_{l,t} + \sum_{l=1}^{n}\left(\Phi_{j,l}^{(\epsilon)} + \sum_{i=1}^{J}\Phi_{j,il}^{(s\epsilon)}s_{i,t-1}\right)\epsilon_{l,t} \\ &\quad + \left(\Phi_{j}^{(0)} + \sum_{i=1}^{J}\Phi_{j,i}^{(s)}s_{i,t-1} + \sum_{i=1}^{J}\sum_{l=1}^{J}\Phi_{j,il}^{(ss)}s_{i,t-1}s_{l,t-1}\right) \end{aligned}$$

which implies that conditional $s_t$ follows a multivariate version of non-central chi-squared distribution. The extended Kalman filter (EKF) that is frequently used for the nonlinear state space model would show poor performance because a fundamental flaw of the EKF is that the densities of the various random variables are no longer normal after undergoing their respective nonlinear transformations. The EKF is simply an ad hoc state estimator that only approximates the optimality of the Bayes rule by linearization. Tanizaki (1996) investigates the performance of various deterministic filters including EKF and finds Monte Carlo evidence that EKF delivers poor performance when applied to real economic applications.

The evaluation of the likelihood can be implemented with the particle filter, also known as the sequential Monte Carlo filter. Early contributors to the particle filter include Gordon, Salmond, and Smith (1993) and Kitagawa (1996). In economics, the particle filter has been applied to analyze stochastic volatility models as in Pitt and Shephard (1999) and Kim, Shephard, and Chib (1998). Recently Fernández-Villaverde and Rubio-Ramírez (2004) use the filter to construct the likelihood for a DSGE model solved with a projection method. We follow their approach and use the particle filter for our DSGE model solved with a second-order perturbation method. A brief description of the procedure follows.

**Initialization** Draw $N$ particles $\left\{s_0^i\right\}_{i=1}^N$ from the initial distribution $p(s_0|\,\theta)$. By induction, in period $t$, we start with the particles $\left\{s_{t-1}^i\right\}_{i=1}^N$ which approximate $p\left(s_{t-1}|Y^{t-1},\theta\right)$.

**Prediction** Draw one-step ahead forecasted particles $\left\{\tilde{s}_t^i\right\}_{i=1}^N$ from $p\left(s_t|Y^{t-1},\theta\right)$. Note that

$$p\left(s_t|Y^{t-1},\theta\right) = \int p\left(s_t|s_{t-1},\theta\right)\, p\left(s_{t-1}|Y^{t-1},\theta\right)\, ds_{t-1} \approx \frac{1}{N}\sum_{i=1}^{N} p\left(s_t|s_{t-1}^i,\theta\right).$$

Hence, one can draw $N$ particles from $p\left(s_t|Y^{t-1},\theta\right)$ by generating one draw from $p\left(s_t|s_{t-1}^i,\theta\right)$ for each $i$.

**Filtering** The goal is to approximate

$$p\left(s_t|Y^t,\theta\right) = \frac{p\left(y_t|s_t,\theta\right) p\left(s_t|Y^{t-1},\theta\right)}{p\left(y_t|Y^{t-1},\theta\right)}, \tag{39}$$

which amounts to updating the probability weights assigned to the particles $\left\{\tilde{s}_t^i\right\}_{i=1}^N$. We begin by computing the unnormalized importance weights $\tilde{\pi}_t^i = p\left(y_t|\tilde{s}_t^i,\theta\right)$. The denominator in (39) can be approximated by

$$p\left(y_t|Y^{t-1},\theta\right) = \int p\left(y_t|s_t,\theta\right)\, p\left(s_t|Y^{t-1},\theta\right)\, ds_t \;\approx\; \frac{1}{N}\sum_{i=1}^{N}\tilde{\pi}_t^i. \tag{40}$$

Now define the normalized weights

$$\pi_t^i = \frac{\tilde{\pi}_t^i}{\sum_{j=1}^N \tilde{\pi}_t^j}$$

and note that the importance sampler $\left\{\tilde{s}_t^i, \pi_t^i\right\}_{i=1}^N$ approximates the updated density $p\left(s_t|\,Y^t,\theta\right)$. The Sequential Importance Sampler (SIS) is a Monte Carlo method which utilizes this aspect, but it is well documented to suffer from a degeneracy problem, where after a few iterations, all but one particle will have negligible weight.

**Resampling** We now generate a new set of particles $\left\{s_t^i\right\}_{i=1}^N$ by resampling with replacement $N$ times from an approximate discrete representation of $p\left(s_t|\,y^t,\theta\right)$ given by $\left\{\tilde{s}_t^i, \pi_t^i\right\}_{i=1}^N$ so that

$$\Pr\left(s_t^i = \tilde{s}_t^i\right) = \pi_t^i, \qquad i = 1,\ldots,N.$$

The resulting sample is in fact an iid sample from the discretized density of $p\left(s_t|\,y^t,\theta\right)$, and hence is equally weighted.

**Likelihood Evaluation** According to (40) the log likelihood function can be approximated as follows

$$\ln \mathcal{L}(\theta|Y) = \ln p\left(y_1|\,\theta\right) + \sum_{t=2}^{T} \ln p\left(y_t|Y^{t-1},\theta\right) \approx \frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{N}\tilde{\pi}_t^i.$$

### 4.3 Configuration of the Particle Filter

There are several issues on the practical implementation of the particle filter. First, we need a scheme to draw from the initial state distribution, $p(s_0|\theta)$. In the linear case, it is straightforward to calculate the unconditional distribution of $s_t$ associated with the vector autoregressive representation (30). For the second-order approximation we exploit (36) and (37). We generate $x_0$ using (37) by drawing $z_0$ from its unconditional distribution using (36) and setting $x_{-1} = x$, where $x$ is the steady state of $x_t$.

Second, we have to choose the number of particles. For a good approximation of the prediction error distribution, it is desirable to have many particles, especially enough particles to capture the tails of $p(s_t|Y^t)$. Moreover, the number of particles affects the performance of the resampling algorithm. If the number of particles is too small the resampling will not work well. In our implementation, the stratified resampling scheme proposed by Kitagawa (1996) is applied. It is optimal in terms of variance in the class of unbiased resampling schemes. Fernández-Villaverde and Rubio-Ramírez (2006) show by simulation that the gain from more particles is very marginal after 20,000 particles in their specification, and it motivates our choice of 40,000 particles. If the measurement errors in the conditional distribution $p(y_t|s_t)$ are small, more particles are needed to obtain an accurate approximation of the likelihood function and to ensure that the posterior weights $\pi_t^i$ do not assign probability one to a single particle. In our application, the standard deviations of the measurement errors are fixed around 20 percents of those of observations.

## 5 Results

We begin the section by explaining the data used and the prior specifications, and the results from the posterior follows. The same prior as in the artificial data is used.

## 5.1 The Data

Now we apply our procedure to estimate our model with U.S. quarterly data. Quarterly observations of per capita output growth rate, inflation and interest rate from the third quarter of 1987 to the fourth quarter 2002 are used. Most of data are extracted from FRED2 database maintained by Federal Reserve Bank of St. Louis. Per capita output growth is based on the real GDP series (GDPC96) and the working age population series, which refers to population between age 16 and 64 and is constructed as follows. Age 16 years and over (PAN17) and age 65 years and over (PAN19) are extracted from the DRI-Global Insight, and hence, the difference between the two makes the working age population. The updates for these population estimates after 1995 are available from the US Census Bureau. Since the estimates between 1990 and 1995 in the DRI-Global Insight are postcensal estimates based on the 1990 census, they are superseded by the intercensal estimates that are based on the 1990 and 2000 census. After 2000, postcensal estimates are used.[4] Annual data are converted to quarterly frequency using a quadratic interpolation. Per capita output growth is defined as $100 * (\log(GDP_t/POP_t) - \log(GDP_{t-1}/POP_{t-1}))$. Inflation is based on the consumer price index (CPIAUCSL) from FRED2 in monthly frequency. Quarterly frequency series are obtained by arithmetic averaging. Inflation is defined as $400*(\log(CPI_t/CPI_{t-1}))$. The nominal interest rate is the effective Federal fund rate (FEDFUNDS) extracted from FRED2. The monthly series is converted into quarterly frequency by arithmetic averaging.

## 5.2 Priors

The choice of priors plays an important role in the estimation of DSGE models. They might down-weigh regions of the parameter space that are at odds with observations not contained in the estimation sample. They might also add curvature to a likelihood function that is (nearly) flat in some dimensions of the parameter space and therefore strongly influence the shape of the posterior distribution. While, in principle, priors can be gleaned from personal introspection to reflect strongly held beliefs about the validity of economic theories, in practice most priors are chosen based on some observations.

Table 1 lists the marginal prior distributions for the structural parameters of the DSGE model. The inverse elasticity of intertemporal substitution, $\tau$, has a gamma prior centered around 2 and with standard deviation 0.5, which reflects the range frequently found in the literature. For example,

[4]For 1990-2000 the following series are used: [Intercensal estimates of the united states resident plus armed forces overseas population by age and sex, 1990-2000: all months]. For 2000 and after [Monthly postcensal resident population plus armed forces overseas, by single year of age, sex, race, and Hispanic origin] are used.

in the business cycle literature $\tau$ is calibrated as low as 1 (e.g. King, Plosser, and Rebelo 1988) and as high as 3 (e.g. Rotemberg and Woodford 1992). The prior for the parameter that governs price stickiness, $\kappa$, is chosen based on the micro-evidence on price setting behavior provided, for instance, in Bils and Klenow (2004). The priors for the coefficients in the monetary policy rule, $\psi_1$ and $\psi_2$, are loosely centered around values typically associated with the Taylor rule, 1.5 and 0.5. For the steady state consumption-output ratio, $c/y$, a beta prior with mean 0.85 and standard deviation 0.1 is used.

For convenience, it is typically assumed that all parameters are *a priori* independent. In applications in which the independence assumption is unreasonable, one could derive parameter transformations such as steady state ratios, autocorrelations, or relative volatilities, and specify independent priors on the transformed parameters, which induce dependent priors for the original parameters. As mentioned before, rational expectations models can have multiple equilibria. While this may be of an independent interest, we do not pursue this direction in this paper. Hence, the prior distribution is truncated at the boundary of the determinacy region. The distribution specified in Table 1 places about 2 percents of its mass on the parameter values that imply indeterminacy.

### 5.3 The Posterior

We now compare the posterior distributions obtained from our linear and nonlinear analysis. In both cases we use the same prior distribution which is reported in Table 1. We use the Metropolis-Hastings algorithm to generate draws from the posterior distributions associated with the linear and the quadratic approximations of the DSGE model. We first compute the posterior mode for the linear specification with the additional parameters, $\nu$ and $c/y$, fixed at their true values. After that, we evaluate the Hessian to be used in the Metropolis algorithm at the (linear) mode without fixing $\nu$ and $c/y$, so that the inverse Hessian reflects the prior variance of the additional parameters. This Hessian is used for both the linear and the nonlinear analysis. We use scaling factors 0.4 for the linear estimation and 0.25 for the quadratic estimation to target an acceptance rate of about 35 percents. The Markov chains are initialized in the neighborhood of the linear posterior mode in both approaches. The following findings are in line with those from Fernández-Villaverde and Rubio-Ramírez (2005, 2006) and An and Schorfheide (2007).

Figure 1 and Figure 2 depict the draws from prior, linear posterior, and quadratic posterior distribution. 60,000 draws from each distribution are generated, the first 10,000 draws are discarded for the convergence of Markov chain, and every 250th draw is plotted. As expected, the draws appear to be more concentrated as we move from prior to posterior. However for most of our parameters, plots from linear and quadratic posterior do not show much difference. This finding can be confirmed in Table 3 where we report the means and 90 percent probability intervals for prior, linear and quadratic posterior distributions. Differences are pronounced with policy parameters on output. The posterior estimates of $\psi_2$ using the quadratic approach show that $\psi_2$ is distributed more compactly around significantly smaller mean than those from the linear estimation. That is, the posterior means (the inter-90 percentile ranges) are 0.514 (0.769) and 0.479 (0.693) for linear and quadratic estimations, respectively. The same can be observed for the steady state parameters $r^{(A)}$, $\pi^{(A)}$, and $\gamma^{(Q)}$. For these parameters the difference between linear and quadratic estimates are quite huge which is mainly resulted from the adjustment factor for the uncertainty neglected in the linear estimation procedure. As mentioned earlier the steady state parameters in general can have different meanings in linear and quadratic approximated DSGE models and the bias correction for the stochastic steady state works in the quadratic approximation. This difference can make the monetary authority to misperceive the state of the economy and hence to perform an over (or under) reaction to the output gap and/or inflation, which can be seen in the policy parameter estimates. These parameters are also more concentrated around their means for the quadratic posteriors. Another finding is that the newly identified parameters, especially $\nu$, have effects on the quadratic posterior distribution, even though the identification is weak. Estimates of $\nu$ for quadratic posterior is different from those for prior and linear posterior, in both mean and 90% interval, which implies that there is some information contained in quadratic likelihood. Moreover, the correlation between $\kappa$ and $\nu$ can be found in the quadratic posterior distribution.

Table 2 reports the log marginal data densities for our linear and quadratic estimation. Geweke's (1999) modified harmonic mean estimator is used to compute the marginal data density. The result shows that the quadratic estimation fits the data better than the linear estimation. The difference is 0.5 in logarithm (and 1.65 in level) in favor of quadratic estimation, which does not fit into Jeffreys' (1961) criterion for decisive evidence. Considering that the baseline particle filter shows inefficiency in terms of the size of particle, that is, a large number of particles fails to survive because of the lack of information on current observations, this evidence in favor of quadratic estimation can be improved by increasing the size of particle.

### 5.4 Welfare Analysis

Now we compare the performance of our linear and quadratic estimations in terms of the welfare variations from changes in monetary policy. We consider welfare differentials relative to a baseline policy using the welfare measure discussed in Section 3, the welfare cost in terms of consumption.

Figure 3 shows welfare cost differentials relative to a baseline policy $\psi_1 = 1.5$ and $\psi_2 = 0.5$ (vertical dashed line). The welfare cost differentials at the posterior means are depicted (solid line), that is, all the non-policy parameters are fixed at their posterior mean values and we calculate the changes in welfare cost as a function of one of the policy parameters, $\psi_1$ in upper and $\psi_2$ in lower panels, for the linear (left) and the quadratic (right) posterior. The shaded area represents 90 percent intervals. Non-policy parameters are fixed at each posterior draw and the differentials are calculated as a function of each policy parameter.

The first thing we notice from Figure 3 is that, as expected, welfare cost increases rapidly as the inflation responsiveness parameter, $\psi_1$, tends to the indeterminacy region. This finding is in line with Del Negro and Schorfheide (2005). Especially for the quadratic posterior this tendency is pronounced, which implies that we may lose critical information on welfare cost originated from changes in monetary policy when we stick to the linear posterior as is widely used in the literature, that is, to estimate a new Keynesian DSGE model with a log-linear approximation and the Kalman filter, and then to perform a policy analysis based on a second-order approximated model. Even though there are noticeable differences in welfare costs constructed from linear and quadratic posteriors, changes in the policy in the opposite direction where the welfare of the economy can be improved are not effective in terms of magnitude for both $\psi_1$ and $\psi_2$. As pointed in Schmitt-Grohé and Uribe (2006a), these negligible effects of the monetary policy based on either linear or quadratic estimations can depend on the abstract nature of our model, for example, absence of capital accumulation and money demand, and a passive fiscal policy. Still, the finding strongly supports the advantage of the estimation based on second-order approximations in the sense that linear procedure can underestimate the effect of monetary policy on the welfare of the economy.

### 5.5 Impulse Response Analysis

Considering nonlinearities associated with a second-order approximated DSGE model, the impulse responses developed for linear processes are not enough to analyze the dynamics of our model economy. Figure 4 shows the impulse responses of standard practice, that is, those for log-linearized economy. The impulse responses at the posterior mean (solid) are plotted alongside of 90 percent intervals evaluated with the quadratic posterior draws (shaded).[5] Most of the responses are in line with the standard results in business cycles literature. As expected from our model specification, the inflation and interest rate do not respond to government spending shocks in our model.

As mentioned in Section 3 , impulse responses for linear processes have several special properties that can not be shared with nonlinear processes: symmetry, scalability, and path-dependency. To deal with these problems when we extend the impulse response analysis to our second-order

[5]The intervals constructed from linear posterior draws are identical qualitatively and hence we do not report.

approximated DSGE model, we rather employ the generalized impulse responses proposed by Koop, Pesaran, and Potter (1996). In our application, past histories are fixed at the steady states and future shocks are integrated out by Monte Carlo to accommodate path-dependency. Moreover, due to lack of symmetry and scalability we simulate various initial shocks and trace out the response distributions along the horizon rather than fix them to a unit standard deviation that is commonly used for linear processes.

Figure 5 shows the generalized impulse responses for our second-order approximated model at the quadratic posterior mean. On the top-left panel, response distributions of output to monetary policy shocks show less dispersion as the horizon gets longer. The tendency that the response distribution becomes more concentrated around zero implies the mean reversion of output to monetary shocks, and it is consistent with the impulse responses in Figure 4. Noting that the technology shock has a persistent effect on output in Figure 4, we can expect that the response distributions will not show the mean reversion property easily and it is confirmed in the bottom-left panel of Figure 5 that they become slightly more disperse over the horizon. It is notable that the magnitudes of the traditional impulse responses to unit shocks coincide with the standard deviations of generalized impulse responses for linear processes. The technology and government spending shocks have persistent effects on output and interest rate, which is implied by the estimated persistence of shocks. While the traditional impulse responses do not reflect higher-order effects of the government spending shocks to inflation and interest rate, persistent effects on interest rate can be shown through a generalized impulse response analysis. Their responses in Figure 4 reflect only the first-order effects from the shock and are flat, that is, the inflation and interest rate do not respond to government spending shocks. However, Figure 5 tells different story. If the inflation and interest rate were not affected by government spending shocks, the response distributions in Figure 5 would have been degenerate from the initial horizon with zero the standard deviations.

# 6 Conclusion

We have presented a general procedure to estimate second-order approximations to a DSGE model and compare the performance to the widely used estimation technique for a log-linearized economy on a version of new Keynesian monetary model. It is done in the context of posterior distributions, welfare cost, and impulse response analysis.

Our findings include the followings. First, we find that all the results of An and Schorfheide (2007) are confirmed with U.S. data. With the nonlinear estimation we can identify parameters that are neglected previously; the marginal data density evaluation shows that data support the nonlinear estimation procedure; and parameter estimates that are related to nondeterministic steady states are quite different from the linear estimates. Second, the estimated welfare differentials are more aggressive for the second-order approximations, that is, the posterior welfare differentials from the linear estimation may underestimate the welfare cost resulted from changes in the monetary policy. Third, the second-order approximation unveils quite different dynamics which are neglected in a log-linearized economy. With advance of computational algorithm and computing power, it is interesting to expand our analysis to more realistic models such as Smets and Wouters (2003) and Christiano, Eichenbaum, and Evans (2005) in future research.

## References

An, Sungbae, and Frank Schorfheide (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26(2-4), 113–172.

Anderson, Gary (2000): "A Reliable and Computationally Efficient Algorithm for Imposing the Saddle Point Property in Dynamic Models," *Manuscript*, Federal Reserve Board of Governors.

Arulampalam, M. Sanjeev, Simon Maskell, Neil Gordon, and Tim Clapp (2002): "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," *IEEE Transactions on Signal Processing*, 50(2), 173–188.

Aruoba, S. Borağan, Jesús Fernández-Villaverde, and Juan F. Rubio-Ramírez (2006): "Comparing Solution Methods for Dynamic Equilibrium Economies," *Journal of Economic Dynamics and Control*, 30(12), 2477–2508.

Bils, Mark, and Peter Klenow (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112(5), 947–985.

Binder, Michael, and M. Hashem Pesaran (1997): "Multivariate Linear Rational Expectations Models: Characterization of the Nature of the Solutions and Their Fully Recursive Computation," *Econometric Theory*, 13(6), 877–888.

Blanchard, Olivier Jean, and Charles M. Kahn (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48(5), 1305–1312.

Christiano, Lawrence J. (2002): "Solving Dynamic Equilibrium Models by a Methods of Undetermined Coefficients," *Computational Economics*, 20(1-2), 21–55.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Collard, Fabrice, and Michel Juillard (2001): "Accuracy of Stochastic Perturbation Methods: The Case of Asset Pricing Models," *Journal of Economic Dynamics and Control*, 25(6-7), 979–999.

DeJong, David N., Beth F. Ingram, and Charles H. Whiteman (2000): "A Bayesian Approach to Dynamic Macroeconomics," *Journal of Econometrics*, 98(2), 203–223.

Del Negro, Marco, and Frank Schorfheide (2005): "Policy Predictions if the Model Does Not Fit," *Journal of European Economic Association*, 3(2-3), 434–443.

Den Haan, Wouter J., and Albert Marcet (1994): "Accuracy in Simulation," *Review of Economic Studies*, 61(1), 3–17.

Fernández-Villaverde, Jesús, and Juan F. Rubio-Ramírez (2004): "Comparing Dynamic Equilibrium Models to Data: A Bayesian Approach," *Journal of Econometrics*, 123(1), 153–187.

Fernández-Villaverde, Jesús, and Juan F. Rubio-Ramírez (2005): "Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood," *Journal of Applied Econometrics*, 20(7), 891–910.

Fernández-Villaverde, Jesús, and Juan F. Rubio-Ramírez (2006): "Estimating Dynamic Equilibrium Economies: A Likelihood Approach," *Review of Economic Studies*, forthcoming.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen (1993): "Nonlinear Dynamic Structures," *Econometrica*, 61(4), 871–907.

Geweke, John (1999): "Computational Experiments and Reality," *Manuscript*, University of Iowa.

Gordon, Neil J., David John Salmond, and Adrian F.M. Smith (1993): "Novel Approach to Nonlinear and Non-Gaussian Bayesian State Estimation," *IEE Proceedings-F*, 140(2), 107–113.

Jeffreys, Harold (1961): *Theory of Probability.* Oxford University Press.

Jin, He-Hui, and Kenneth L. Judd (2002): "Perturbation Methods for General Dynamic Stochastic Models," *Manuscript*, Hoover Institute.

Judd, Kenneth L. (1998): *Numerical Methods in Economics.* Cambridge, MA: MIT Press.

Kim, Jinill, and Sunghyun Henry Kim (2003): "Spurious Welfare Reversal in International Business Cycle Models," *Journal of International Economics*, 60(2), 471–500.

Kim, Jinill, Sunghyun Henry Kim, Ernst Schaumburg, and Christopher A. Sims (2005): "Calculating and Using Second Order Accurate Solutions of discrete Time Dynamic Equilibirum Models," *Manuscript*, Princeton University.

Kim, Sangjoon, Neil Shephard, and Siddhartha Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65(3), 361–393.

King, Robert G., Charles I. Plosser, and Sergio T. Rebelo (1988): "Production, Growth and Business Cycles: I. The Basic Neoclssical Model," *Journal of Monetary Economics*, 21(2-3), 195–232.

King, Robert G., and Mark W. Watson (1998): "The Solution of Singluar Linear Difference Systems under Rational Expectations," *International Economic Review*, 39(4), 1015–1026.

Kitagawa, Genshiro (1996): "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models," *Journal of Computational and Graphical Statistics*, 5(1), 1–25.

Klein, Paul (2000): "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control*, 24(10), 1405–1423.

Klein, Paul (2005): "Second-Order Approximation of Dynamic Models without the Use of Tensors," *Manuscript*, University of Western Ontario.

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter (1996): "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics*, 74(1), 119–147.

Laforte, Jean-Philippe (2004): "Comparing Monetary Policy Rules in an Estimated Equilibrium Model of the US Economy," *Manuscript*, Federal Reserve Board of Governors.

Levin, Andrew T., and J. David López-Salido (2004): "Optimal Monetary Policy with Endogenous Capital Accumulation," *Manuscript*, Federal Reserve Board of Governors.

Levin, Andrew T., Alexei Onatski, John C. Williams, and Noah Williams (2006): "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models," in *NBER Macroeconomics Annual 2005*, ed. by Mark Gertler, and Kenneth Rogoff. Cambridge, MA: MIT Press.

McGrattan, Ellen R. (1994): "The Macroeconomic Effects of Distortionary Taxation," *Journal of Monetary Economics*, 33(3), 573–601.

Otrok, Christopher (2001): "On Measuring the Welfare Cost of Business Cycles," *Journal of Monetary Economics*, 47(1), 61–92.

Pitt, Micheal K., and Neil Shephard (1999): "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94(446), 590–599.

Rotemberg, Julio J., and Michael D. Woodford (1992): "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity," *Journal of Political Economy*, 100(6), 1153–1207.

Rotemberg, Julio J., and Michael D. Woodford (1997): "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," in *NBER Macroeconomics Annual 1997*, ed. by Ben S. Bernanke, and Julio J. Rotemberg. Cambridge, MA: MIT Press, 297–346.

Sargent, Thomas J. (1989): "Two Models of Measurements and the Investment Accelerator," *Journal of Political Economy*, 97(2), 251–287.

Schmitt-Grohé, Stephanie, and Martín Uribe (2004): "Solving Dynamic General Equilibrium Models Using a Second-order Approximation to the Policy Function," *Journal of Economic Dynamics and Control*, 28(4), 755–775.

Schmitt-Grohé, Stephanie, and Martín Uribe (2006a): "Optimal Fiscal and Monetary Policy in a Medium-Scale Macroeconomic Model," in *NBER Macroeconomics Annual 2005*, ed. by Mark Gertler, and Kenneth Rogoff. Cambridge, MA: MIT Press, 383–425.

Schmitt-Grohé, Stephanie, and Martín Uribe (2006b): "Optimal Simple and Implementable Monetary and Fiscal Rules," *Journal of Monetary Economics*, forthcoming.

Schorfheide, Frank (2000): "Loss Function-Based Evaluation of DSGE Models," *Journal of Applied Econometrics*, 15(6), 645–670.

Sims, Christopher A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1-2), 1–20.

Smets, Frank, and Raf Wouters (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of European Economic Association*, 1(5), 1123–1175.

Swanson, Eric, Gary Anderson, and Andrew Levin (2005): "Higher-Order Perturbation Solutions to Dynamic, Discrete-Time Rational Expectations Models," *Manuscript*, Federal Reserve Bank of San Fransisco.

Tanizaki, Hisashi (1996): *Nonlinear Filters: Estimation and Application*. Springer-Verlag.

Taylor, John B., and Harald Uhlig (1990): "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods," *Journal of Business and Economic Statistics*, 8(1), 1–17.

Uhlig, Harald (1999): "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily," in *Computational Methods for the Study of Dynamic Economies*, ed. by Ramon Marimón, and Andrew Scott. Oxford, U.K.: Oxford University Press, 30–61.

Woodford, Michael (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.

Table 1: PRIOR DISTRIBUTION

| Name | Domain | Density | Para (1) | Para (2) |
|---|---|---|---|---|
| $\tau$ | $\mathbf{R}^+$ | Gamma | 2.000 | 0.500 |
| $\nu$ | $[0,1)$ | Beta | 0.500 | 0.200 |
| $\kappa$ | $\mathbf{R}^+$ | Gamma | 0.250 | 0.100 |
| $c/y$ | $[0,1)$ | Beta | 0.850 | 0.100 |
| $\psi_1$ | $\mathbf{R}^+$ | Gamma | 1.500 | 0.250 |
| $\psi_2$ | $\mathbf{R}^+$ | Gamma | 0.500 | 0.250 |
| $\rho_R$ | $[0,1)$ | Beta | 0.500 | 0.200 |
| $\rho_g$ | $[0,1)$ | Beta | 0.750 | 0.100 |
| $\rho_z$ | $[0,1)$ | Beta | 0.660 | 0.150 |
| $r^{(A)}$ | $\mathbf{R}^+$ | Gamma | 1.000 | 0.500 |
| $\pi^{(A)}$ | $\mathbf{R}^+$ | Gamma | 3.000 | 2.000 |
| $\gamma^{(Q)}$ | $\mathbf{R}$ | Normal | 0.400 | 0.200 |
| $100\sigma_R$ | $\mathbf{R}^+$ | InvGamma | 0.400 | 4.000 |
| $100\sigma_g$ | $\mathbf{R}^+$ | InvGamma | 0.200 | 4.000 |
| $100\sigma_z$ | $\mathbf{R}^+$ | InvGamma | 0.200 | 4.000 |

*Notes:* Para (1) and Para (2) list the means and the standard deviations for Beta, Gamma, and Normal distributions; the upper and lower bound of the support for the Uniform distribution; $s$ and $\nu$ for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu,s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$. The effective prior is truncated at the boundary of the determinacy region.

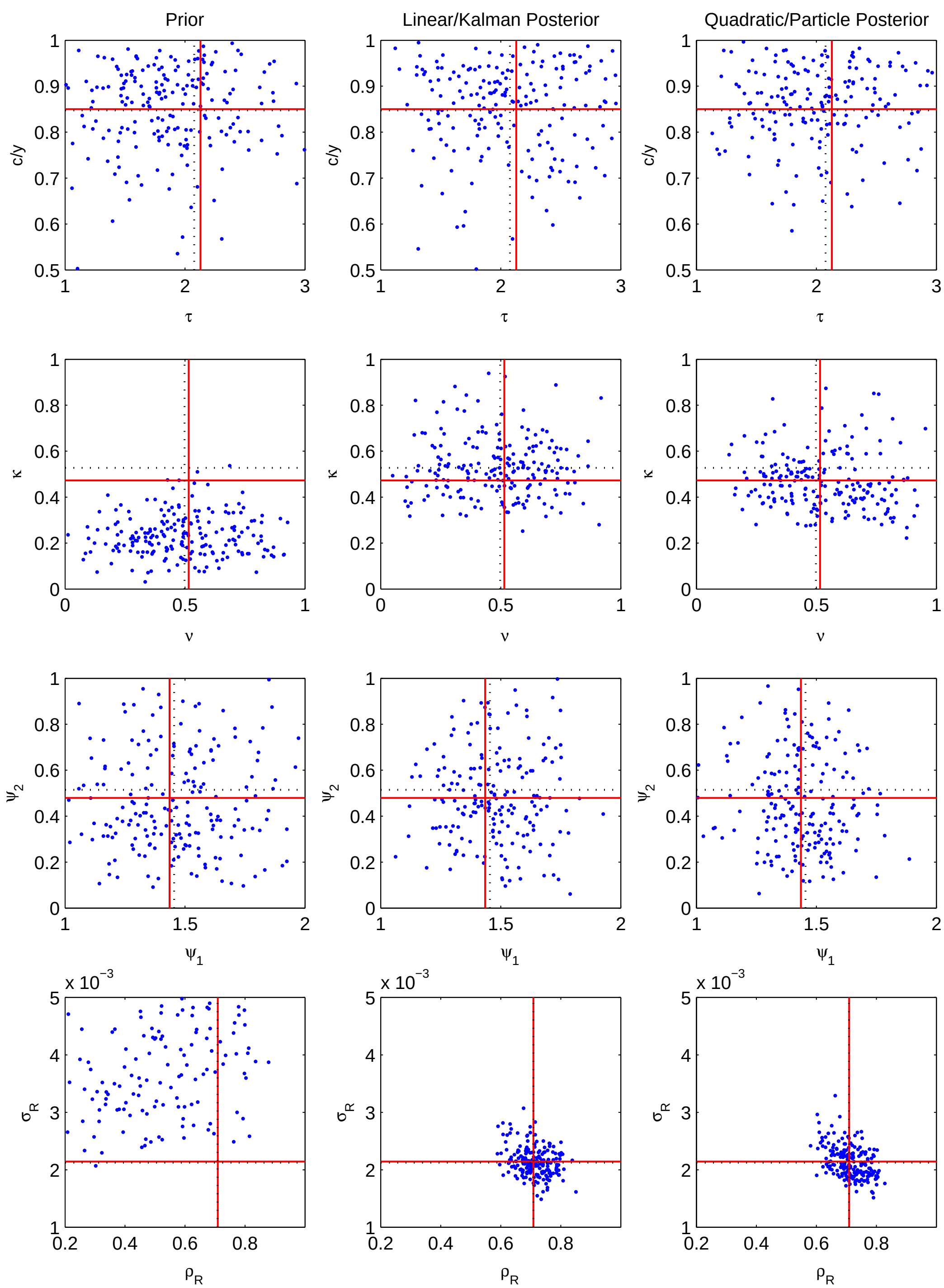
Table 2: LOG MARGINAL DATA DENSITIES: U.S. DATA

| Linear/Kalman | Quadratic/Particle |
|---|---|
| −249.180 | −248.678 |

Table 3: Posterior Estimates: U.S. Data

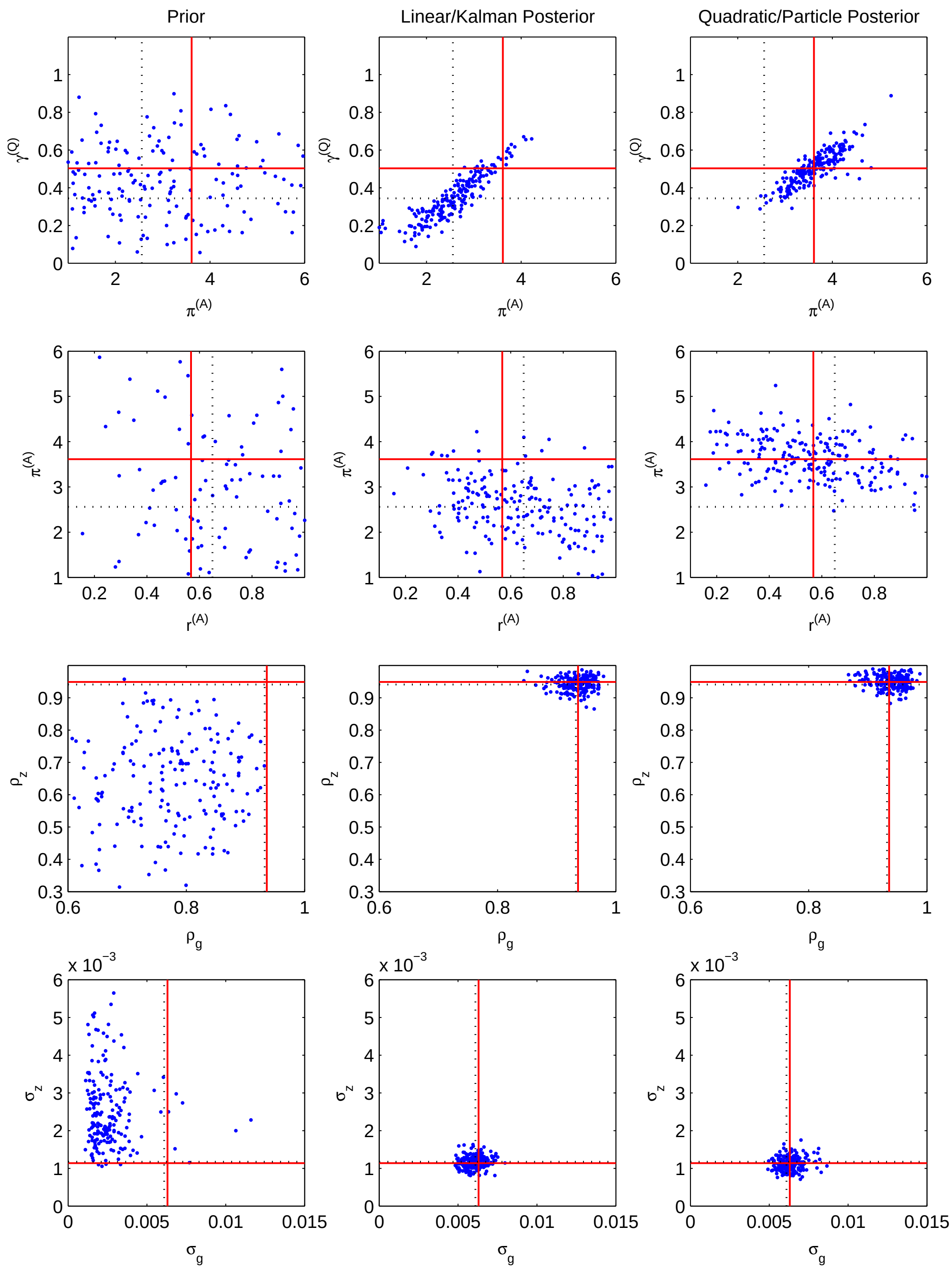
| | Prior | | Posterior | | | |
|---|---|---|---|---|---|---|
| | | | Linear/Kalman | | Quadratic/Particle | |
| | Mean | 90% Interval | Mean | 90% Interval | Mean | 90% Interval |
| $\tau$ | 2.004 | [1.202, 2.814] | 2.077 | [1.268, 2.917] | 2.129 | [1.292, 3.030] |
| $\nu$ | 0.499 | [0.168, 0.824] | 0.498 | [0.175, 0.837] | 0.515 | [0.200, 0.846] |
| $\kappa$ | 0.250 | [0.091, 0.400] | 0.527 | [0.315, 0.733] | 0.472 | [0.279, 0.669] |
| $c/y$ | 0.850 | [0.707, 0.993] | 0.847 | [0.700, 0.991] | 0.850 | [0.712, 0.993] |
| $\psi_1$ | 1.507 | [1.090, 1.906] | 1.455 | [1.179, 1.717] | 1.436 | [1.170, 1.723] |
| $\psi_2$ | 0.501 | [0.119, 0.876] | 0.514 | [0.129, 0.898] | 0.479 | [0.131, 0.824] |
| $\rho_R$ | 0.501 | [0.174, 0.829] | 0.709 | [0.621, 0.796] | 0.709 | [0.629, 0.802] |
| $\rho_g$ | 0.750 | [0.596, 0.913] | 0.932 | [0.893, 0.974] | 0.936 | [0.898, 0.976] |
| $\rho_z$ | 0.662 | [0.418, 0.900] | 0.941 | [0.906, 0.976] | 0.949 | [0.915, 0.984] |
| $r^{(A)}$ | 0.998 | [0.231, 1.728] | 0.649 | [0.271, 1.011] | 0.568 | [0.200, 0.889] |
| $\pi^{(A)}$ | 2.991 | [0.228, 5.761] | 2.558 | [1.537, 3.575] | 3.612 | [2.885, 4.341] |
| $\gamma^{(Q)}$ | 0.401 | [0.068, 0.723] | 0.344 | [0.147, 0.554] | 0.503 | [0.366, 0.637] |
| $100\sigma_R$ | 0.499 | [0.209, 0.784] | 0.212 | [0.170, 0.253] | 0.214 | [0.173, 0.259] |
| $100\sigma_g$ | 0.252 | [0.108, 0.400] | 0.610 | [0.510, 0.704] | 0.630 | [0.520, 0.724] |
| $100\sigma_z$ | 0.251 | [0.105, 0.395] | 0.117 | [0.089, 0.144] | 0.114 | [0.087, 0.143] |

Figure 1: Posterior Draws: U.S. Data I



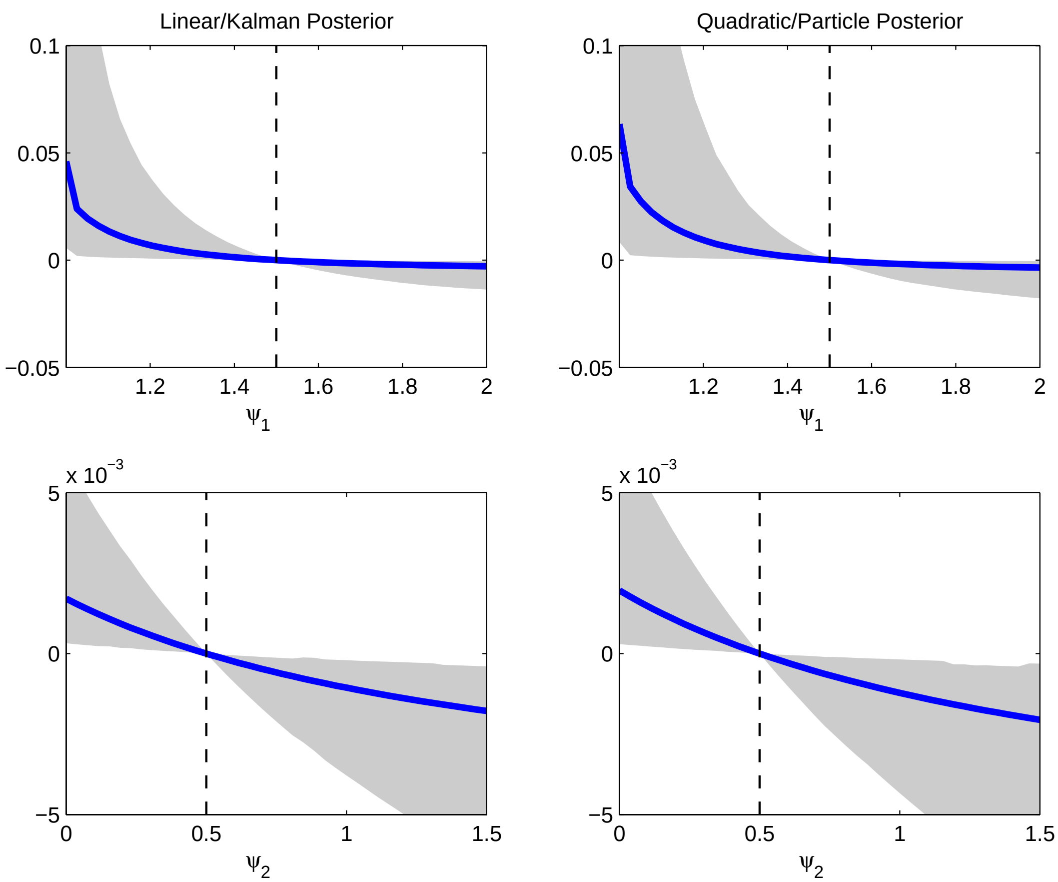

*Notes:* 60,000 draws from the prior and posterior distributions. First 10,000 draws are discarded for the convergence of Markov chain. Every 250th draw is plotted. Quadratic (solid) and linear (dotted) posterior means are shown. 40,000 particles are used.

Figure 2: Posterior Draws: U.S. Data II



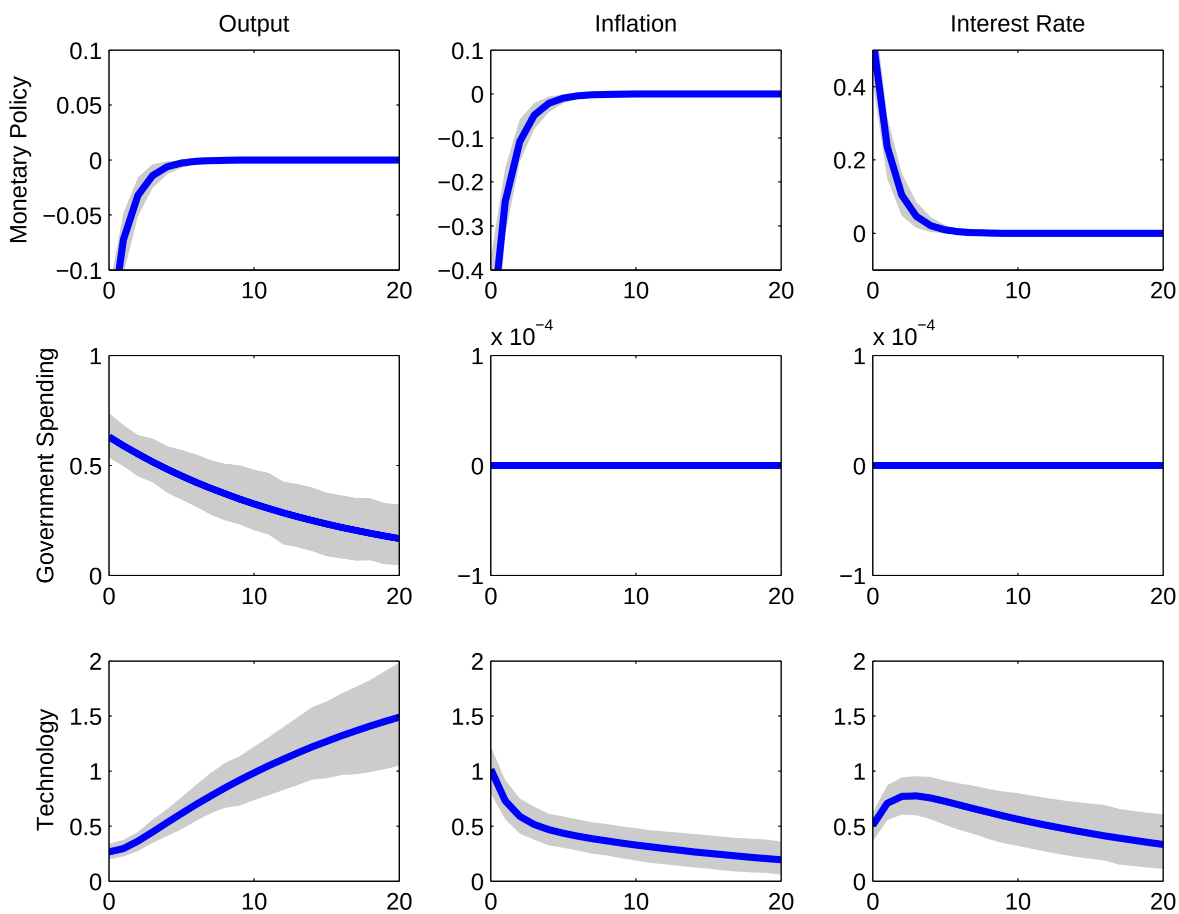

*Notes:* 60,000 draws from the prior and posterior distributions. First 10,000 draws are discarded for the convergence of Markov chain. Every 250th draw is plotted. Quadratic (solid) and linear (dotted) posterior means are shown. 40,000 particles are used.

Figure 3: Welfare Cost Differentials: U.S. Data



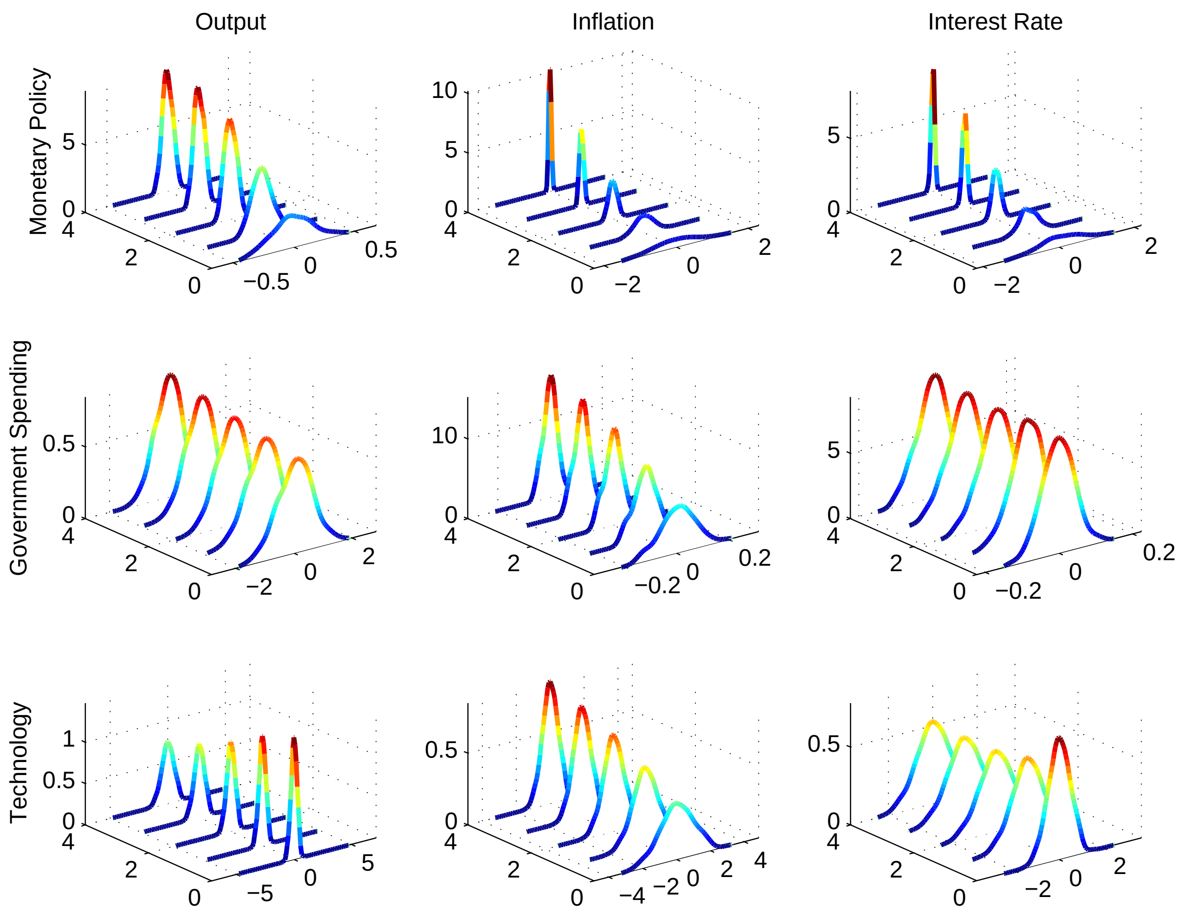

*Notes:* Welfare cost differentials at linear and quadratic posterior means (solid) relative to baseline policy rule $\psi_1 = 1.5$ and $\psi_2 = 0.5$ (dashed). 90% intervals are evaluated using posterior draws (shaded). Negative differentials signify an improvement relative to baseline rule.

Figure 4: IMPULSE RESPONSES: U.S. DATA




*Notes:* Impulse responses based on log-linearized DSGE model. Drawn at quadratic posterior mean (solid) with 90% intervals evaluated using posterior draws (shaded).

Figure 5: Generalized Impulse Responses: U.S. Data




*Notes:* Generalized impulse responses based on second-order approximated DSGE model. Drawn at quadratic posterior mean (solid) from horizon 0 to 4.