Measuring Monetary Policy when the Nominal Short-Term Interest Rate is Zero: A Dynamic Stochastic General Equilibrium Approach*

Tomiyuki Kitamura †

Research and Statistics Department, Bank of Japan

PRELIMINARY AND INCOMPLETE

December 2009

Abstract

This paper proposes an empirical dynamic stochastic general equilibrium (DSGE) framework to measure monetary policy when the nominal short-term interest rate is zero. The framework assumes that there exists a shadow rate which represents the monetary policy stance of a central bank. When the shadow rate is positive, it is observed as the policy rate of the central bank. However, when it is negative, it deviates from the policy rate which remains at its lower-bound of zero. It is not the policy rate but the shadow rate that affects the economy when the two rates deviate. With this framework, standard DSGE models can be fitted to data using a version of particle filter, and the historical movements of the shadow rate can be estimated. As an application, we estimate a small New Keynesian model using Japanese data. The results suggest that the shadow rate was well below zero during the period of zero policy rate in the 2000s.

Keywords: Monetary policy; Dynamic stochastic general equilibrium; Zero lower-bound of nominal interest rate; Particle filter; Shadow rate

*The views expressed in the paper are those of the author and are not reflective of those of the Bank of Japan.

†E-mail address: tomiyuki.kitamura@boj.or.jp

1 Introduction

Many central banks today use short-term nominal interest rates as their main policy instruments. A potentially serious challenge caused by the use of nominal interest rate, however, is the zero lower-bound of interest rate. Once the policy rate becomes zero, monetary policy can become ineffective as long as the central bank does not do anything other than the control of current policy rate.

Given the low inflation environment in many developed countries in recent decades, the challenge caused by the zero lower-bound of interest rate has become not of mere academic curiosity, but of practical importance to the central banks. A recent example of countries that actually tackled such a situation is, of course, Japan. Japanese economy has experienced a prolonged stagnation in the 1990s, and as a result of the deflationary pressures, Bank of Japan (BoJ) lowered its policy rate to the lower bound of zero twice in recent 15 years.

However, one can argue that the view that monetary policy loses its effects on the economy once the policy rate reaches zero is too naive. Recent theoretical studies show that, even when the policy rate has reached the lower-bound of zero, monetary policy can stimulate the economy by committing to keeping the policy rate at zero longer than usual (e.g., Reifschneider and Williams, 2000). Indeed, as summarized in the survey by Ugai (2006), empirical studies on the Japanese experience suggest that policy measures of BoJ during the periods of zero interest rate actually had at least some effects on the economy.

Motivated by such arguments, this paper proposes an empirical framework that can be used to measure the degree of monetary policy accomodation during the periods of zero interest rate. In so doing, we assume that there exists a 'shadow rate' that represents the policy stance of the central bank. While this shadow rate coincides with the policy rate in normal circumstances where it is positive, when it becomes negative, they deviate from each other. It is not the policy rate but the shadow rate that affects the economy when the two rates deviate, i.e., when the policy rate has hit its lower-bound of zero.

Once we assume the presence of the shadow rate, standard dynamic stochastic general equilibrium (DSGE) models can be fitted to data in a Bayesian framework. Normally, estimating a DSGE model is not a very difficult task because the likelihood function of the model can be easily evaluated using the Kalman filter. In our setup, however, since the relationship between the shadow rate and the policy rate is non-linear, a direct application of the Kalman filter is not possible. To overcome this problem, we employ a version of particle filter.[1] The particle filter allows us to approximate the distribution of the shadow rate, and using the approximated distribution, we can estimate the model parameters. Finally, with the parameter estimates, we can estimate the unobserved path of the shadow rate.

[1]For an overview of particle filters, see Arulampalam et al. (2002). As an application to macroeconometrics, Fernandez-Villaverde and Rubio-Ramirez (2005) also employ a particle filter to estimate a DSGE model, but in a very different context than that of the present paper.

The concept of shadow rate is not new in the literature (e.g. Black, 1995). Iwata and Wu (2006) estimate a reduced-form model to evaluate the monetary policy effects during the periods of zero interest rate in Japan, assuming the presence of the shadow rate. Also, Ichiue and Ueno (2007, 2008) introduced a shadow rate in estimating a yield-curve model using Japanese data. As we will see in Section 2, the distinctions between our approach and theirs are two-fold. First, none of these studies employs a DSGE framework like ours. Second, these studies do not fully take into account the possible effects of changes in the shadow rate on the economy. In particular, the framework of Iwata and Wu (2006) assume that the shadow rate affects the economy only contemporaneously, and thus imposes a restriction that the lags of the shadow rate do not affect current variables including the current shadow rate. But such a restriction does not generally hold in a rational-expectations model. For example, when monetary policy represented by the shadow rate has some inertia, the shadow rate depends on its own lags.

As an application of the framework, we fit a small New Keynesian DSGE model to Japanese data whose sample period contains the experience of zero interest rate. The results indicate the usefulness of our approach. The estimated shadow rate is well below zero during the periods of zero interest rate, especially in the period in the 2000s. Overall, the estimated parameters of the model seem to be reasonable compared with those of the existing studies in the literature.

The paper is organized as follows. Section 2 presents our DSGE framework that assumes the presence of a shadow rate. Since the framework can be applied to a broader class of DSGE models than the small model estimated in a later section in this paper, we describe it based on a generic state-space representation of a DSGE model with shadow rate. Section 3 describes a version of particle filter which we employ to deal with the nonlinearity arising from the zero lower-bound of interest rate. Section 4 discusses the estimation strategy based on the particle filter. As an application of our proposed estimation method, Section 5 fits a small New Keynesian model with the shadow rate to Japanese data. Section 6 concludes. Appendix includes the derivations of some equations, relevant conditional densities, and complete descriptions of the algorithms used to implement the estimation strategy of the paper.

2 A DSGE framework with shadow rate

In this section, we sketch the DSGE framework in which a shadow rate is present. First we present a generic state-space representation of a DSGE model with shadow rate, and point out the nonlinearity caused by the zero lower-bound of interest rate. Then we compare our framework with those of the existing literature on the zero lower-bound of interest rate.

2.1 State-space representation

Let us consider an economy consisting of agents (including a government, if necessary) and a central bank. We assume that the policy stance of the central bank is represented by a shadow rate, r_t^*. The central bank has a full control power over the shadow rate r_t^*. This shadow rate is the rate that affects the economy, and thus, enters in the objective functions and the constraints of the agents in the economy.

When the shadow rate is positive, it coincides with the short-term nominal interest rate, r_t, which the central bank uses as its policy instrument in normal circumstances. When the shadow rate is negative, however, the two rates deviate, and the short-term rate r_t stays at zero. Thus, the short-term rate and the shadow rate are related by the following observation equation:

$$r_t = \max \{0, r_t^*\} . \tag{1}$$

Suppose that the optimization problems of the agents are specified and solved, that the policies of the central bank and the government are specified, and that the equilibrium is defined. By log-linearizing the resulting behavioral equations, policy functions, and the resource constraints around the steady state, a DSGE model of the economy can be expressed as a system of expectational difference equations:

$$AE_t (S_{t+1}) = BS_t + C + D\varepsilon_t \tag{2}$$

where S_t is the $m \times 1$ state vector containing variables (possibly including expectational variables and lagged variables) in the model, and $\varepsilon_t \sim N(0, \Sigma)$ is an $n \times 1$ vector of structural shocks.

We assume that the shadow rate r_t^* can be written as

$$r_t^* = \beta_r + \Lambda_r S_t \tag{3}$$

for some scalar β_r and some $1 \times m$ vector Λ_r.[2]

Given the parameter values, the system of expectational difference equations in (2) can be solved with standard methods (e.g., Klein, 2000, and Sims, 2002). Assuming the determinacy of the system, the rational-expectations solution of (2) can be written as

$$S_t = \alpha + \Pi S_{t-1} + \Psi \varepsilon_t \tag{4}$$

where α, Π and Ψ are nonlinear functions of the structural parameters contained in A, B, C

[2] When the variables in the model are all expressed in terms of the deviation from the steady state value, β_r is the steady state value of the shadow rate. When the units are different between r_t and the interest rate in the model, Λ_r contains a unit transformation coefficient (which is, for example, 4 if the transformation is from quarter-on-quarter basis to year-on-year basis).

and D.[3]

Combining (4) with an observation equation which relates the observed data to the state vector S_t, we have a state-space representation of the model. Let $Y_t \equiv (x_t', r_t)'$ be a observation vector containing r_t as well as other observed data series x_t. Then the observation equation can be written as

$$Y_t = F(\beta + \Lambda S_t) \equiv \begin{bmatrix} \beta_x + \Lambda_x S_t \\ \max\{0, \beta_r + \Lambda_r S_t\} \end{bmatrix} \tag{5}$$

where $\beta \equiv (\beta_x', \beta_r)'$ is a vector of constants, Λ_x is a matrix, and $\Lambda \equiv (\Lambda_x', \Lambda_r')'$.[4] We assume (i) there are as many observed series as structural shocks, i.e., Y_t is an $n \times 1$ vector, and (ii) $\Lambda\Psi$ is invertible. The implication of these assumptions will be discussed in the next section.

The function F in the observation equation (5) is nonlinear. When such nonlinearity is absent, estimation of the system and the state vector S_t is not a difficult task. The likelihood of the system consisting of (4) and a linear observation equation can be easily evaluated via the Kalman filter (see, e.g., Hamilton, 1994). Then the parameters can be estimated by maximum likelihood or Bayesian method (see, e.g., An and Schorfheide, 2007). With the parameter estimates, the state vector S_t can be estimated by the Kalman smoothing.

However, with the nonlinear observation equation (5), the standard Kalman filter is not applicable. In the next section we will describe the difficulties caused by the nonlinearity in detail and how a version of particle filter can be used to deal with the problem.

2.2 Comparison with the literature

To illustrate the novelty of our approach to deal with the nonlinearity caused by the zero lower-bound, let us compare our framework with those of the previous studies.

First, compared to the frameworks of most theoretical studies on the zero lower-bound of interest rate, our framework does not suffer from the lack of analytical solution method for nonlinear difference equation. In the theoretical studies, the presence of the shadow rate is not assumed, and the short-term rate is the rate that affects the economy. In other words, the shadow rate is assumed to always coincide with the short-term rate, even when the short-term rate is zero. These studies normally assume that the short-term rate evolves according to an interest rate rule with the zero lower-bound constraint. In our notation, the rule can be written as

$$r_t = r_t^* = \max\{0, c + \phi S_t\},$$

[3] We do not consider sunspot shocks. Lubik and Schordheide (2004) show how a DSGE model without any nonlinearity can be estimated allowing for the possibility of indeterminacy. Although it is beyond the scope of the current paper, whether the estimation strategy proposed here can be extended to allow for the possibility of indeterminacy is an interesting question.

[4] As is evident in the equation, we do not consider any measurement error in this paper.

where c is a constant and ϕ is an $1 \times m$ vector. Typically, the equation is based on a Taylor-type rule with interest rate smoothing, and thus ϕS_t is a sum of the inflation term, the output gap term, and the term of the lagged short-term rate. Note that the above nonlinear equation implies that the dynamics of r_t^* are affected by the nonlinearity caused by the zero lower-bound constraint. Thus, the system of the structural equations can not be written in a linear form. Currently, no analytical solution method for such a nonlinear system is available.[5] As a result, the estimation techniques used for linear DSGE models can not be applied to such models. In contrast, the structural equation (2) in our framework is linear. Thus, in our case, the model can be solved by the standard solution methods. As a result, it can also be estimated as long as the nonlinearity of the observation equation is taken care of with particle filtering, as we will show later.

We do not claim that our approach is the ideal solution to deal with the nonlinearity caused by the zero lower-bound. One might argue that the concept of the shadow rate has little theoretical foundations. However, a benefit from assuming the presence of the shadow rate can be, in our opinion, large enough to pay the cost of losing the theoretical coherence of the framework. The benefit is that we can estimate a DSGE model and the policy stance during the periods of zero interest rate.

Next, compared to the frameworks of the empirical studies which assume the presence of the shadow rate, there are two major improvements in our framework. First, our framework is based on a fully-structural DSGE model while the previous studies are based on reduced-form models.[6] Second, our framework allows for the possibility that the lags of the shadow rate affect the current variables in the system. To see this point, let us look at the framework of Iwata and Wu (2006). They employ a reduced-form system

$$A_0 Y_t^* = A(L) Y_t + \mu + u_t$$

where $Y_t^* \equiv (x_t', r_t^*)'$, A_0 is a matrix, $A(L) \equiv A_1 L - \cdots - A_p L^p$, L is the lag operator, μ is a vector of constants, and u_t is a vector of shocks. Note that the shadow rate r_t^* is included not in the right-hand side but only in the left-hand side. This implies that the shadow rate can affect the variables in the economy only contemporaneously, and that the lags of the shadow rate can not have any effect on the current variables. In particular, the shadow rate can not depend on its own lags, and thus monetary policy inertia is assumed to be lost once the shadow rate becomes negative. In contrast, the state-transition equation (4) in our framework allows for the effects of the lags of the shadow rate on the current

[5] There are some methods which can handle certain types of nonlinearities in state transition equation. An example is the Markovian Jump Linear Quadratic approach proposed by Svensson and Williams (2007, 2008). However, such methods have not been applied to the problem of zero bound of interest rate.

[6] In fact, our framework does not have to be based on a DSGE model, and can be based on a reduced-form model. Without considering the underlying DSGE model (2), one can base the whole analysis on the reduced-form equation (4) treating α, Π and Ψ instead of A, B, C and D as the parameters to be estimated.

variables. Regarding this point, Ichiue and Ueno (2007, 2008) also allow for the dependence of the current shadow rate on its own lags. Their framework, however, is based on a partial-equilibrium yield curve model with the shadow rate. Therefore, unlike us, they ignore the effects of shadow rate on the other variables in the economy.

3 The filtering method

We propose to employ a version of particle filter to deal with the nonlinearity in the observation equation (5). Particle filters are sequential Monte Carlo filtering methods which are applicable to any state-space model. The key idea of the methods is to represent the conditional density function of the state vector by a set of random samples ("particles") with associated weights. It is shown in the literature that as the number of samples increases, this approximate representation approaches the true conditional density.

In this section, we consider the situation in which all the parameter values are given. The estimation of the parameters will be discussed in the next section. The focus of this section is on how the distribution of the shadow rate can be approximated with particles. We also assume that, when the particle filter is invoked, the initial value of S_0 is given.

Below, to motivate the introduction of the particle filter, we first consider the filtering problem in a general form. Then, we describe the specific version of particle filter which we implement for our framework. We also discuss the implications of the assumptions which we made on the system for our implementation of the particle filter.

3.1 The filtering problem

Consider the filtering problem for the system of (4) and (5). Let us denote the history of observations by $Y^t \equiv \{Y_s\}_{s=0}^t$. Filtering problem is to compute $f\left(S_t|Y^t\right)$, the density of the state vector S_t conditional on the history Y^t. Appendix A shows that the following recursive formula for $f\left(S_t|Y^t\right)$ holds:

$$f\left(S_t|Y^t\right) = \int f\left(S_t|S_{t-1}, Y_t\right) f\left(S_{t-1}|Y^t\right) dS_{t-1}, \tag{6}$$

where

$$f\left(S_{t-1}|Y^t\right) = \frac{f\left(Y_t|S_{t-1}\right)}{\int f\left(Y_t|S'_{t-1}\right) f\left(S'_{t-1}|Y^{t-1}\right) dS'_{t-1}} f\left(S_{t-1}|Y^{t-1}\right). \tag{7}$$

This formula is different from the standard recursive formula of filtering problem often used in the literature.[7] We find this version more helpful to clarify the idea of our particle filtering method.

[7]The standard formula often used in the literature involves $f\left(Y_t|S_t\right)$ and $f\left(S_t|S_{t-1}\right)$ instead of $f\left(Y_t|S_{t-1}\right)$ and $f\left(S_t|S_{t-1}, Y_t\right)$ in our formula. See, e.g., Equation (4) of Doucet et al. (2000).

According to (6) and (7), observing the new data Y_t, $f(S_t|Y^t)$ can be updated from $f(S_{t-1}|Y^{t-1})$ using two densities. The first is $f(Y_t|S_{t-1})$, the probability of observing Y_t when S_{t-1} is the true state in $t-1$. Since this probability represents the plausibility of S_{t-1} as the true state in $t-1$ given the new information Y_t, reweighting the old density $f(S_{t-1}|Y^{t-1})$ with them as in (7) yields $f(S_{t-1}|Y^t)$, the new density of S_{t-1} updated with the new information. The second is $f(S_t|S_{t-1}, Y_t)$, the probability of S_t being the next state when S_{t-1} is the old state. The product of this probability and the new density of S_{t-1} gives the probability that the state evolved from S_{t-1} to S_t. Integrating this over all possible starting states S_{t-1} yields the density of the current state, $f(S_t|Y^t)$.

When there was no nonlinearity and shocks were all Gaussian, there would be no need to employ any approximation method. In that case, $f(S_t|S_{t-1}, Y_t)$ and $f(Y_t|S_{t-1})$ would be Gaussian, and $f(S_t|Y^t)$ would take the same simple Gaussian form as $f(S_{t-1}|Y^{t-1})$. This is the case to which the Kalman filter can be applied. In such a case, we could easily compute the exact distribution $f(S_t|Y^t)$, which can be completely characterized by only two parameters, mean and variance.

However, things are not that easy when we have the nonlinearity in the observation equation in (5). Once r_t^* becomes negative, $f(S_t|Y^t)$ takes a different form than $f(S_{t-1}|Y^{t-1})$, and no simple analytical solution like the Kalman filter is available. This is because $f(S_t|S_{t-1}, Y_t)$ and $f(Y_t|S_{t-1})$ involve nonstandard distributions, as we show in Appendix B. As (6) says, the density of S_t conditional on the history of the observations is an integral of the product of the two nonstandard densities over S_{t-1}. The nonnormalities of the two densities make it impossible to express the resulting integral in (6) in any analytically simple form. Therefore, some approximation must be employed to conduct filtering based on (6) in this case.

3.2 Particle filtering

Particle filters approximate the conditional density $f(S_t|Y^t)$ with a set of particles $\{S_t^i\}_{i=1}^N$ and the associated weights $\{w_t^i\}_{i=1}^N$. Our version of particle filter belongs to a class called Sequential Importance Sampling (SIS) filter. In this class of particle filters, the particles are sequentially sampled from a density called importance density, and the weights are also sequentially updated.

Within the class of SIS filters, it has been shown in the literature to be optimal to sample S_t^i and update w_t^i given S_{t-1}^i and w_{t-1}^i by

$$S_t^i \sim f\left(S_t|S_{t-1}^i, Y_t\right), \tag{8}$$

$$w_t^i \propto w_{t-1}^i f\left(Y_t|S_{t-1}^i\right). \tag{9}$$

These are optimal in the sense that they minimize the variance of the weights $\{w_t^i\}_{i=1}^N$

conditional on $\{S_t^i\}_{i=1}^N$ and Y^t.[8]

Note that the two conditional densities in (8) and (9) are the same as those in (6) and (7). In fact, the particle filter based on (8) and (9) simply implements the idea embodied in the recursive formula (6) and (7). Suppose that we have particles and weights $\left\{S_{t-1}^i, w_{t-1}^i\right\}_{i=1}^N$ approximating $f\left(S_{t-1}|Y^{t-1}\right)$. Observing Y_t, we can first update the weight assigned to each particle of the previous period using $f\left(Y_t|S_{t-1}^i\right)$, the plausibility of S_{t-1}^i as the true state in $t-1$ given the new information Y_t. Also, for each particle of the previous period, we can generate particles of the current period using $f\left(S_t|S_{t-1}^i, Y_t\right)$, the probability of S_t being the next state when S_{t-1}^i is the old state. The resulting particles and weights $\{S_t^i, w_t^i\}_{i=1}^N$ approximates $f\left(S_t|Y^t\right)$.

We use the optimal importance density function (8) and the optimal weight updating equation (9) for our particle filter. In many cases considered in the literature, the use of this optimal density function and the associated weights is not feasible. Specifically, there are two reasons why the optimal procedure is infeasible in many cases. First, it is usually very difficult to sample from $f\left(S_t|S_{t-1}^i, Y_t\right)$. Second, it is often computationally hard to evaluate $f\left(Y_t|S_{t-1}^i\right)$. In our particular case, however, these two common difficulties are not present. First, as Appendix B shows, $f\left(S_t|S_{t-1}^i, Y_t\right)$ is a density of a variable which is a linear function of a truncated normal variable. Therefore, sampling from the density can be done by drawing from a truncated normal density and plugging the draw into the linear function. Second, as Appendix B also shows, although $f\left(Y_t|S_{t-1}^i\right)$ is not normal when $r_t = 0$, it is a product of normal density and a probability which can be calculated from a cumulative normal density. Thus, it can easily be evaluated.

Starting from initial particles $\{S_0^i\}_{i=1}^N$ and weights $\{w_0^i\}_{i=1}^N$, the particle filter sequentially approximate $f\left(S_t|Y^t\right)$ by sampling new particles according to (8) and updating weights according to (9). The complete description of the algorithm is given in Appendix C.

3.3 The role of the assumptions

Since the state vector in general contains multiple variables, its conditional density $f\left(S_t|Y^t\right)$ is a multivariate density. Therefore, in general, we may have to approximate a multivariate density in implementing the particle filter.

However, approximating a multivariate density is notoriously a difficult task. As the number of dimensions increases, the number of particles required for a given precision of approximation exponentially increases. Then, the amount of time needed for the computation easily explodes, rendering the approximation method computationally infeasible.

[8]A large variance of the weights means that there are a lot of particles to which very small weights are assigned. Such particles contain less information on the density to be approximated than particles with larger weights. Therefore, the efficiency of the approximation deteriorates as the variance of the weights increases. See, e.g., Arulampalam et al. (2002).

The two assumptions which we imposed on the system in Section 2, and the assumption made in this section that S_0 is given, are intended to avoid this problem. With the two assumptions, namely, (i) there are as many observed series as structural shocks, and (ii) $\Lambda\Psi$ is invertible, we effectively limit our attention to the case in which an approximation of the one-dimensional density of r_t^* is sufficient to approximate the density of entire S_t. This is because under the assumptions there is a mapping from r_t^* to S_t given Y_t and S_{t-1}, as we show in Appendix B. Intuitively, when $\Lambda\Psi$ is invertible, the structural shocks ε_t can be computed from r_t^* and x_t given S_{t-1}. Then, with the structural shocks ε_t and given S_{t-1}, the entire state vector S_t can be backed out from the state-transition equation (4). Note that the argument holds for all $t \geq 1$ under the assumption that S_0 is given.

Note that the assumptions do not hold for some of the typical DSGE models in the literature such as Smets and Wouters (2003). In Smets and Wouters (2003), the number of shocks exceeds the number of observed data series. Intuitively, in that model some of the shocks are inherently attached to the natural output, and the natural output is treated as an unobservable state variable in the state-space representation. When such an additional unobservable variable other than the shadow rate was present in our framework, the approximation must have been done for a multivariate density. However, as we noted above, the computation could be infeasible in such a case.[9]

4 Estimation strategies

Now we proceed to the estimation step of the system given by (4) and (5). The parameter values are not known and now to be estimated. We denote by θ all the structural parameters contained in matrices A, B, C, and D in (2).

We assume that the sample period for the estimation is chosen such that the short-term rate is not zero in the initial period. This assumption guarantees that the unconditional distribution of the observed data in the initial period is a normal distribution. Then, as we will show below, the period likelihood function for the initial period can be easily evaluated.

In the previous section, we assumed that S_0 is given when the particle filter is invoked. When S_0 can be computed from Y_0, the assumption can be easily satisfied. Even when some elements of S_0 can not be computed from Y_0, we can treat the unknown elements in S_0 as additional parameters, and add them to θ.

[9]For the case with other unobservable variables than the shadow rate, we suggest that a Gibbs sampler can be applied in estimating the model using the particle filter presented here. Note that, if the path of the shadow rate is known, the system becomes standard and the DSGE model can be easily estimated. Therefore, starting from an initial path of the shadow rate, we can draw parameters and the paths of unobservable variables other than the shadow rate from the posterior distribution conditional on the path of the shadow rate. Then, given the draws of the path of the unobservables and parameter values, we can apply the particle filtering method to draw a path of the shadow rate. One can iterate these steps to sample parameters and paths of the unobservables from the unconditional posterior distribution.

We first describe how the likelihood of the model for a set of the parameter values can be computed with the particles and weights generated by the particle filter. Then, we describe the strategy for estimating the parameters. We also describe the particle smoothing method, by which we can estimate the path of the shadow rate given the parameter estimates.

4.1 Likelihood evaluation

For a set of parameter values, the particle filter generates the particles and associated weights $\left\{ \{S_t^i\}_{i=1}^N, \{w_t^i\}_{i=1}^N \right\}_{t=0}^T$ which approximate the conditional densities of the state vector. With them, we can evaluate the likelihood as follows.

First, the likelihood function can be written as a product of the period likelihood functions $f(Y_t|Y^{t-1})$

$$f\left(Y^T|\theta\right) = \prod_{t=0}^{T} f\left(Y_t|Y^{t-1}\right) \tag{11}$$

with $f(Y_0|Y^{-1}) \equiv f(Y_0)$. We now make explicit the dependence of the likelihood on the parameter values by putting θ in the expression on the left hand side.

For $t > 0$, the period likelihood function $f(Y_t|Y^{t-1})$ can be written as

$$f\left(Y_t|Y^{t-1}\right) = \int f\left(Y_t|S_{t-1}\right) f\left(S_{t-1}|Y^{t-1}\right) dS_{t-1},$$

where we used the fact that $f(Y_t|S_{t-1}, Y^{t-1}) = f(Y_t|S_{t-1})$. We approximate this using the particles $\left\{S_{t-1}^i\right\}_{i=1}^N$ and weights $\left\{w_{t-1}^i\right\}_{i=1}^N$ approximating $f(S_{t-1}|Y^{t-1})$ as

$$f\left(Y_t|Y^{t-1}\right) \approx \sum_{i=1}^{N} w_{t-1}^i f\left(Y_t|S_{t-1}^i\right).$$

For $t = 0$, we use the unconditional distribution of Y_0. Since the short-term rate in this period is not zero by assumption, we have $r_0^* = r_0$, and thus (5) implies that $Y_0 = \beta + \Lambda S_0$. Since (4) implies that the unconditional distribution of S_0 is normal with mean $(I - \Pi)^{-1}\alpha$ and variance $\sum_{i=0}^{\infty} \Pi^j \Psi \Sigma \Psi' (\Pi^j)'$, the unconditional distribution of Y_0 is normal with mean $\beta + \Lambda (I - \Pi)^{-1}\alpha$ and variance $\sum_{i=0}^{\infty} \Lambda \Pi^j \Psi \Sigma \Psi' (\Pi^j)' \Lambda'$.

4.2 Parameter estimation

Maximum likelihood estimation can be done by simply maximizing the likelihood function (11) with respect to θ. However, in the current literature, maximum likelihood estimation of DSGE models is not the common choice of the estimation method. As An and Schorfheide (2007) point out, one reason for this is that the estimates of structural parameters obtained by maximum likelihood methods are often at odds with the parameter values suggested by other evidence such as the results from micro studies. This reflects that the likelihood function often peaks in regions of the parameter space that are inconsistent with the parameter values suggested by other evidence.

The common approach employed in the literature of estimation of DSGE models is a Bayesian approach. In the Bayesian approach, a prior distribution $p(\theta)$ is placed on parameters. Then by Bayes' Theorem the posterior distribution of θ can be constructed from the prior distribution and the likelihood function as

$$p\left(\theta|Y^T\right) = \frac{f\left(Y^T|\theta\right)p(\theta)}{\int f\left(Y^T|\theta\right)p(\theta)\,d\theta}$$

It is usually impossible to analytically compute this posterior distribution. Thus, instead, draws from the posterior distribution are generated using Markov Chain Monte Carlo (MCMC) methods, and inference on the parameters are made based on these posterior draws.

In this paper, following An and Schorfheide (2007), we employ the Random-Walk Metropolis algorithm as the MCMC method to generate posterior draws. The details of the algorithm are given in Appendix C.

4.3 Smoothing

Given parameter values, smoothing is a step to compute the densities of state vectors S_t at each period given the entire data sample Y^T, that is, $f\left(S_t|Y^T\right)$ for all t. When the particle filter is employed for filtering, smoothing can be done using the particles and weights generated by the particle filter. Let us assume that we now have the parameter estimates obtained from the estimation step described above, and that using the parameter estimates we have invoked the particle filter to get the particles and weights $\left\{\{S_t^i, w_t^i\}_{i=1}^N\right\}_{t=1}^T$.

To conduct smoothing with the particles and weights, we use backward simulation following Godsill, Doucet, and West (2004). As we describe below, the idea of the backward simulation is straightforward. Our primary interest is on the path of shadow rate, and therefore let us focus here not on smoothing the entire state vector but on smoothing only the path of shadow rate. The details of the algorithm, which conducts smoothing for the entire state vector, is given in Appendix C.

First note that plugging each particle S_t^i into (3) yields the corresponding shadow rate $r_t^{*,i}$. Thus $\left\{r_t^{*,i}, w_t^i\right\}_{i=1}^N$ approximates the conditional density $f\left(r_t^*|Y^t\right)$. With these particles of the shadow rate and their weights, the smoothing procedure approximates $f\left(r_t^*|Y^T\right)$ by generating lots of draws from the density. To get one such draw, we start from the last period T and move backward to the initial period 0. First, for the last period T, $f\left(r_T^*|Y^T\right)$ is exactly the conditional density which the particles and weights $\left\{r_T^{*,i}, w_T^i\right\}_{i=1}^N$ approximate. Thus, we simply choose the sample of the shadow rate $\tilde{r}_T^*$ from $\left\{r_T^{*,i}\right\}_{i=1}^N$ with probability $\{w_T^i\}_{i=1}^N$.

For $t < T$, given the chosen shadow rate $\tilde{r}^*_{t+1}$, we can reevaluate the weight assigned to each particle with the information that the shadow rate in the next period is $\tilde{r}^*_{t+1}$. This can be done by reweighting each weight $\{w^i_t\}^N_{i=1}$ with the probability that $\tilde{r}^*_{t+1}$ actually realizes in the next period when the current state is S^i_t. Then, we choose $\tilde{r}^*_t$ from $\left\{r^{*,i}_t\right\}^N_{i=1}$ with the reweighted probabilities. The steps proceed recursively backward, and we obtain a sample path of the shadow rate $\{\tilde{r}^*_t\}^T_{t=0}$. By generating a desired number of paths of the shadow rate, we have the approximation of the density $f\left(r^*_t|Y^T\right)$ represented by those paths.

5 An example: small DSGE model with Japanese data

In this section, we fit a small DSGE model to Japanese data as an application of the estimation method presented above.

5.1 A small New Keynesian model

The small model which we fit to Japanese data is a typical DSGE model in the literature. It basically consists of three equations. As usual, the equations have all been log-linearized, and the variables are in terms of deviations from the deterministic steady state, except that the output gap is in terms of deviation from the flexible-price equilibrium. The definitions of the parameters are summarized in Table 1.

The first equation is the forward-looking IS curve with internal habits:[10]

$$y_t = c^y_1 y_{t-1} + c^y_2 E_t y_{t+1} - c^y_3 E_t y_{t+2} - c^y_4 \left(r^*_t - E_t \pi_{t+1}\right) + \varepsilon^y_t, \tag{12}$$

where y_t is the output gap, π_t is the inflation rate, and ε^n_t is the shock to the natural rate whose distribution is i.i.d.normal, $c^y_1 \equiv \chi/\zeta$, $c^y_2 \equiv \left(1 + \beta\chi + \beta\chi^2\right)/\zeta$, $c^y_3 \equiv \beta\chi/\zeta$, $c^y_4 \equiv \left(1 - \beta\chi\right)\left(1 - \chi\right)/\left(\sigma_c\zeta\right)$, $\zeta \equiv 1 + \chi + \beta\chi^2$. The discount factor β is a function of $\bar{r}$, the steady state annual real interest rate (in percentage point), and $\beta = 1/\exp\left(\bar{r}/400\right)$.

The second equation is the hybrid New Keynesian Phillips curve:[11]

$$\pi_t = c^\pi_1 E_t \pi_{t+1} + c^\pi_2 mc_t + c^\pi_3 \pi_{t-1} + \varepsilon^\pi_t, \tag{13}$$

[10] This IS curve can be derived from the following household utility function:

$$E_0 \sum_{t=0}^\infty \beta^t \left[\left(1 - \sigma_c\right)^{-1} \left(C_t - \chi C_{t-1}\right)^{1-\sigma_c} - \chi_h \left(1 + \sigma_h\right)^{-1} N_t^{1+\sigma_h}\right]$$

subject to an appropriate budget constraint with the shadow rate. Here, C_t is consumption level, N_t is labor supplied.

[11] The Phillips curve is derived in the Calvo-style sticky-price setup with backward indexation. The underlying production function is assumed to be Cobb-Duglous with capital being fixed. Labor market is assumed to be competitive.

where ε_t^π is the price shock whose distribution is i.i.d.normal, $c_1^\pi \equiv \beta/\left(1+\beta\gamma_p\right)$, $c_2^\pi \equiv \kappa/\left\{\left(1+\beta\gamma_p\right)\left(1+\theta\left(1-\alpha\right)/\alpha\right)\right\}$, $\kappa \equiv \left(1-\xi_p\right)\left(1-\beta\xi_p\right)/\xi_p$, $c_3^\pi \equiv \gamma_p/\left(1+\beta\gamma_p\right)$ and mc_t is the real marginal cost defined by

$$mc_t \equiv c_4^\pi y_t + c_5^\pi \left\{\left(1+\beta\chi^2\right)\left(y_t - \varepsilon_t^y\right) - \chi\left(\beta E_t y_{t+1} + y_{t-1} + \varepsilon_t^y\right)\right\},$$

with $c_4^\pi \equiv \left(\sigma_h + 1 - \alpha\right)/\alpha$ and $c_5^\pi \equiv \sigma_c/\left\{\left(1-\chi\right)\left(1-\beta\chi\right)\right\}$. The parameters θ and α are the steady state price elasticity of demand and the cost share of labor input in the Cobb-Douglous production function, respectively. Following Ichiue et al. (2008), we calibrate these two parameters and set $\theta = 6$ and $\alpha = 0.63$.

The third equation is the Taylor-type monetary policy rule with inertia:

$$r_t^* = \phi_r r_{t-1}^* + \left(1-\phi_r\right)\left\{\phi_y y_t + \phi_\pi \pi_t\right\} + \varepsilon_t^r. \tag{14}$$

where ε_t^r is the monetary policy shock whose distribution is i.i.d.normal.

Note that the interest rate that enters the structural equations (12)-(14) is not the short-term interest rate r_t but the shadow rate r_t^*. The two rates are related by (1).

Let $S_t \equiv \left(y_{t-1}, \pi_{t-1}, r_{t-1}^*, y_t, \pi_t, r_t^*, E_t y_{t+1}\right)'$ and $\varepsilon_t \equiv \left(\varepsilon_t^y, \varepsilon_t^\pi, \varepsilon_t^r\right)'$. Then the structural equations (12)-(14) can be cast into a system of expectational difference equations (2). Within the parameter region of determinacy, the solution can be written in the form of (4). Letting $\bar{\pi}$ be the steady state inflation rate, the observed data series are related to the model variables through the following observation equation.

$$\begin{pmatrix} GAP_t \\ INF_t \\ R_t \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{\pi} \\ \bar{r} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 100 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 400 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 400 & 0 \end{pmatrix} S_t.$$

where GAP_t is the data of GDP gap, INF_t is the data of annual inflation rate, R_t is the data of short-term annual interest rate. All the data are in percentage point.

Finally, S_0 can be computed from Y_0 in this case. $y_{-1}, \pi_{-1}, r_{-1}^*, y_0, \pi_0, r_0^*$ are computed from the observation equation with Y_{-1} and Y_0. As for $E_0 y_1$, it can be computed by solving $e_{Ey}' S_0 = e_y'\left(\alpha + \Pi S_0\right)$, where e_{Ey} and e_y are selection vector whose values are zeros except for the elements corresponding to $E_t y_{t+1}$ and y_t, respectively, being one.

5.2 Data and implementation

We use three quarterly Japanese time series as observable data series: the CPI inflation rate for INF_t, GDP gap constructed by the method described in Hara et al. (2006) for GAP_t, and the overnight call rate for R_t. All the variables except the call rate are seasonally adjusted. The sample period is 1981:1Q to 2008:3Q. We treat a value of short-term interest rate less than 5 basis point as zero. As a result, there are two periods of zero interest rate: 1999:2Q-2000:2Q and 2001:2Q-2006:2Q.

We set the number of particles to 10,000. While the approximation error becomes smaller as the number of the particles increases, the computation time significantly increases. In order to make the computation more efficient, we implemented the particle filter in Fortran code, but still there is considerable computational burden in running the particle filter. The number of 10,000 is adopted taking into account this tradeoff.

As for the implementation of the Random-Walk Metropolis algorithm, the covariance matrix of the proposal density, $\tilde{\Sigma}$, is set to 0.5 times the Hessian of the posterior density of an auxiliary model evaluated at the posterior mode. The auxiliary model is the same as the above model except that it ignores the nonlinearity caused by the zero lower-bound of interest rate. More precisely, it assumes that the shadow rate is always equal to the short-term rate.[12]

The initial draw is drawn from $N_P\left(\bar{\theta}, 2\tilde{\Sigma}\right)$, where $\bar{\theta}$ is the posterior mode obtained by maximizing $\ln f\left(Y^T|\theta\right) + \ln p\left(\theta\right)$. We repeat the steps of the Random-Walk Metropolis algorithm 1,000,000 times, and throw away the first half of the draws. The resulting 500,000 draws are used to approximate the posterior density.

5.3 Results

Table 1 shows the prior specifications and the parameter values at the posterior mean. Note that the steady state real rate at the posterior mean is lower than its prior mean, 2.00, which is about the average of the ex-post real rate during the sample period. This result obtains, of course, because we allowed the shadow rate to decline below zero.

Figure 1 shows the prior and posterior densities of the structural parameters. In many cases, the posterior density is narrower than the prior density, suggesting that the observations actually added useful information for estimating the parameter values. Also, all the posterior densities are single-peaked, suggesting that a severe identification problem is not present in this estimation.

Figure 2 shows the impulse responses to structural shocks. Most of them seem to be reasonable. In particular, the responses of output gap and inflation rate to a monetary policy shock are hump-shaped, and the response of inflation is delayed, attaining its peak in about 2 years.

[12]Note that in principle $\tilde{\Sigma}$ can be any positive definite matrix for the algorithm to correctly generate draws from the posterior density. Having said that, usually $\tilde{\Sigma}$ is set to some constant times the Hessian of the posterior density of the model being estimated, evaluated at the posterior mode. In our case, however, the Hessian of the model being estimated can not be computed with a high precision because the likelihood is computed with approximation. As a result, when the Hessian of the model being estimated is used, the speed of convergence of the algorithm becomes very slow. We avoid this problem by using the Hessian of the auxiliary model, which is computed with much higher precision because no approximation by the particle filter is necessary for this model.

Figure 3 shows the smoothed path of the shadow rate at the posterior mean, together with the output gap and the inflation rate. The path of shadow rate is the mean of 10,000 paths generated by the method described in Section 4. The 90% confidence interval is also constructed as the 5- and 95- percentiles of the 10,000 paths. As is clear from the figure, the shadow rate declines to a level below zero during the periods of zero interest rate. Especially, the decline of the shadow rate is large in the second zero-interest-rate period. This is consistent with the fact that in that period BoJ conducted a larger number of 'unconventional' policy measures, which are detailed in Ugai (2006). The shadow rate hits the bottom of -2.15% in the 4th quarter of 2002, when the output gap and the inflation rate are also near their bottoms.

To compare the estimated path of the shadow rate with those of the previous studies, Figure 4 and 5 show the paths of the shadow rate estimated in Iwata and Wu (2006) and Ichiue and Ueno (2008). These are taken from their original papers. First look at the estimated path of Iwata and Wu (2006). Their path ("Implied Call Rate") looks similar to ours at least until 2001, which is the end of their sample period. However, since our results are based on the sample period ending 2008:3Q, which includes additional 4 and a half years of period of zero interest rate, one should not draw a conclusion from this comparison. Next, let us look at the estimated path of Ichiue and Ueno (2008). Their sample period extends to the beginning of 2006, and thus a more direct comparison can be made. Overall, their path is far below our path for the most of the periods of zero interest rate. This might reflect the fact that their approach does not take into account the effects of the shadow rate on the economy.[13] On the other hand, note also that the movements of the two paths share some common features. Both paths hit the bottoms around 2003, and quickly rise till about 2004. Such similarities are interesting given that the two paths are estimated by totally different approaches.

6 Conclusion

This paper proposed an empirical DSGE framework to measure monetary policy when the nominal short-term interest rate is zero. In so doing, we introduced the shadow rate representing the monetary policy stance of a central bank into the standard DSGE framework. By the introduction of the shadow rate, the nonlinearity caused by the zero lower-bound does not affect the linearity of the state-transition equation, and only the observation equation becomes nonlinear. We argued that Bayesian estimation of such a model can be done by employing a version of particle filter. As an application of the method, we fitted a small

[13]On the other hand, the results of Ichiue and Ueno (2007) suggest that the large decline of the estimated shadow rate in Figure 5 is due to the assumption of the constant equilibrium interest rate in their yield curve model. In Ichiue and Ueno (2007), they extend their framework to allow for shifts in the equilibrium rate. The bottom of the estimated path using the extended model is around -0.8%, and thus it is now above ours.

New Keynesian model to Japanese data. The results suggest that the shadow rate was well below zero during the period of zero interest rate in the 2000s.

There are many directions for extensions of the current work, and here we suggest two of them. First, the small New Keynesian model which we fitted to Japanese data is very likely to be too simple a model to capture the dynamics of Japanese economy. Employing richer DSGE models to estimate the shadow rate using the framework of this paper is a promising direction. Second, as we pointed out in Section 3, the framework presented in this paper can not directly be applied to some models with unobserved natural output including Smets and Wouters (2003). However, the development of the natural output during the periods of zero interest rate is of great interest especially when we evaluate the policies conducted during those periods. Extending the framework for such models is another direction.

A Derivation of (6)

$$
\begin{aligned}
f\left(S_t|Y^t\right) &= \frac{f\left(S_t, Y_t|Y^{t-1}\right)}{f\left(Y_t|Y^{t-1}\right)} \\
&= \frac{\int f\left(S_t, Y_t, S_{t-1}|Y^{t-1}\right) dS_{t-1}}{\int f\left(Y_t, S'_{t-1}|Y^{t-1}\right) dS'_{t-1}} \\
&= \frac{\int f\left(S_t|Y_t, S_{t-1}, Y^{t-1}\right) f\left(Y_t|S_{t-1}, Y^{t-1}\right) f\left(S_{t-1}|Y^{t-1}\right) dS_{t-1}}{\int f\left(Y_t|S'_{t-1}, Y^{t-1}\right) f\left(S'_{t-1}|Y^{t-1}\right) dS'_{t-1}} \\
&= \int f\left(S_t|Y_t, S_{t-1}\right) \frac{f\left(Y_t|S_{t-1}\right) f\left(S_{t-1}|Y^{t-1}\right)}{\int f\left(Y_t|S'_{t-1}\right) f\left(S'_{t-1}|Y^{t-1}\right) dS'_{t-1}} dS_{t-1}.
\end{aligned}
$$

In the final line, we used the fact that

$$
\begin{aligned}
f\left(S_t|S_{t-1}, Y_t, Y^{t-1}\right) &= f\left(S_t|S_{t-1}, Y_t\right), \\
f\left(Y_t|S_{t-1}, Y^{t-1}\right) &= f\left(Y_t|S_{t-1}\right).
\end{aligned}
$$

These relationships hold because, if we know S_{t-1}, the past history Y^{t-1} does not add any information to infer about S_t and Y_t.

B Conditional densities

Let us fix the notation first. We denote by $Y_t^* \equiv (x_t', r_t^*)'$ the vector of data which would be observed if the shadow rate was observable. Then we have $Y_t^* = \beta + \Lambda S_t$, and thus (4) implies that

$$
Y_t^* = \gamma + \Phi S_{t-1} + e_t, \tag{B.1}
$$

where $\gamma \equiv \beta + \Lambda\alpha$, $\Phi \equiv \Lambda\Pi$, $e_t \equiv \Lambda\Psi\varepsilon_t$, $e_t \sim N(0, \Omega)$, and $\Omega \equiv \Lambda\Psi\Sigma\Psi'\Lambda'$. We partition e_t, γ, Φ and Ω as

$$
e_t \equiv \begin{bmatrix} e_{x,t} \\ e_{r,t} \end{bmatrix}, \quad \gamma \equiv \begin{bmatrix} \gamma_x \\ \gamma_r \end{bmatrix}, \quad \Phi \equiv \begin{bmatrix} \Phi_x \\ \Phi_r \end{bmatrix}, \quad \Omega \equiv \begin{bmatrix} \Omega_{xx} & \Omega_{xr} \\ \Omega_{rx} & \Omega_{rr} \end{bmatrix},
$$

with the number of rows of $e_{x,t}$, γ_x, Φ_x and Ω_{xx} being equal to that of x_t.

Note that under the assumptions that Y_t is an $n \times 1$ vector and that $\Lambda\Psi$ is invertible, there is a mapping from r_t^* to S_t given S_{t-1} and Y_t. That is, combining $\varepsilon_t = (\Lambda\Psi)^{-1} e_t$ and (B.1) and then substituting the result into (4), we have

$$
S_t = \alpha + \Pi S_{t-1} + \Psi (\Lambda\Psi)^{-1} \left(Y_t^* - (\gamma + \Phi S_{t-1})\right). \tag{B.2}
$$

As we pointed out in Section 3, this mapping enables us to approximate the density of S_t, which is in general multivariate, by approximating one-dimensional density.

Below, we denote by $N_P(x;\mu,\Sigma)$ and $N_C(x;\mu,\Sigma)$ the normal density and its cumulative version at x with mean μ and variance Σ. Also, we denote by $N_P^{Tr}(x;c,\mu,\sigma^2)$ the truncated normal density at x with mean μ, variance σ^2, and the truncation from above at c.

[SUGOU: Below, $f(Y_t|S_{t-1})$ should come first because it is referred before $f(S_t|S_{t-1},Y_t)$ on p.7??]

B.1 $f(S_t|S_{t-1},Y_t)$

(B.1) implies that the density of Y_t^* conditional on S_{t-1} is

$$f(Y_t^*|S_{t-1}) = N_P(Y_t^*;\gamma+\Phi S_{t-1},\Omega). \tag{B.3}$$

Then, the density of r_t^* conditional on x_t is given by

$$f(r_t^*|S_{t-1},x_t) = N_P\left(r_t^*;\gamma_r+\Phi_r S_{t-1}+\rho_{rx}e_{x,t},\sigma_{r|x}^2\right), \tag{B.4}$$

where $\rho_{rx}\equiv\Omega_{rx}\Omega_{xx}^{-1}$, $\sigma_{r|x}^2\equiv\Omega_{rr}-\Omega_{rx}\Omega_{xx}^{-1}\Omega_{xr}$.

When $r_t>0$, we know for sure that $Y_t^*=Y_t$ and thus the value of S_t can be computed by (B.2) without any uncertainty. In this case, the density becomes degenerate. When $r_t=0$, since $r_t=0$ if and only if $r_t^*\leq 0$, we have $f(r_t^*|S_{t-1},Y_t)=f(r_t^*|S_{t-1},x_t|r_t^*\leq 0)$. Therefore, $f(r_t^*|S_{t-1},Y_t)$ is a truncated normal density given by

$$f(r_t^*|S_{t-1},Y_t) = \begin{cases} \text{degenerate at } r_t^*=r_t & \text{when } r_t>0 \\ N_P^{Tr}\left(r_t^*;0,\gamma_r+\Phi_r S_{t-1}+\rho_{rx}e_{x,t},\sigma_{r|x}^2\right) & \text{when } r_t=0 \end{cases} \tag{B.5}$$

Then the conditional density $f(S_t|S_{t-1},Y_t)$ is implicitly given by (B.2), the mapping from r_t^* to S_t.

B.2 $f(Y_t|S_{t-1})$

First note that $f(Y_t|S_{t-1})$ is a joint density of x_t and r_t. Therefore, it can be written as the product of the density of r_t conditional on x_t and the density of x_t, i.e.,

$$f(Y_t|S_{t-1}) = f(r_t|S_{t-1},x_t)\,f(x_t|S_{t-1}).$$

As for the first term, note that r_t is a censored normal variable with the latent variable r_t^*. Since the density of r_t^* conditional on x_t is given by (B.4), we have

$$f\left(r_t|S_{t-1},x_t\right)=\begin{cases} N_P\left(r_t;\gamma_r+\Phi_r S_{t-1}+\rho_{rx}e_{x,t},\sigma^2_{r|x}\right) & \text{if } r_t>0\\ N_C\left(0;\gamma_r+\Phi_r S_{t-1}+\rho_{rx}e_{x,t},\sigma^2_{r|x}\right) & \text{if } r_t=0\end{cases}$$

As for the second term, (B.1) implies that x_t has the conditional density

$$f\left(x_t|S_{t-1}\right)=N_P\left(x_t;\gamma_x+\Phi_x S_{t-1},\Omega_{xx}\right).$$

Multiplying the two and noting that the resulting density for $r_t>0$ is simply the joint normal density of Y_t, we have

$$f\left(Y_t|S_{t-1}\right)=\begin{cases} N_P\left(Y_t;\gamma+\Phi S_{t-1},\Omega\right) & \text{if } r_t>0\\ N_C\left(0;\gamma_r+\Phi_r S_{t-1}+\rho_{rx}e_{x,t},\sigma^2_{r|x}\right)N_P\left(x_t;\gamma_x+\Phi_x S_{t-1},\Omega_{xx}\right) & \text{if } r_t=0\end{cases} \tag{B.6}$$

C Algorithms

Algorithm 1 (Particle filter for shadow rate)**.**

1. *Set initial particles and weights as* $S_0^i=S_0$*,* $w_0^i=1/N$ *for all* $i=1,\ldots,N$*.*
2. *For* $t=1$ *to* T*:*
   (a) *Draw* $\left\{r_t^{*,i}\right\}_{i=1}^N$ *as follows.*
      - *When* $r_t>0$*, set* $r_t^{*,i}=r_t$*.*
      - *When* $r_t=0$*, draw* $\left\{r_t^{*,i}\right\}_{i=1}^N$ *from (B.5).*

   (b) *Set* $S_t^i=\alpha+\Pi S_{t-1}^i+\Psi\left(\Lambda\Psi\right)^{-1}\left(Y_t^{*,i}-\left(\gamma+\Phi S_{t-1}^i\right)\right)$ *with* $Y_t^{*,i}=(x_t',r_t^{*,i})'$ *for all* $i=1,\ldots,N$*.*

   (c) *Update weights by* $w_t^i\propto w_{t-1}^i f\left(Y_t|S_{t-1}^i\right)$ *with (B.6).*
3. $\left\{\{S_t^i,w_t^i\}_{i=1}^N\right\}_{t=1}^T$ *is an approximation of the sequence of the conditional distributions of the state vector,* $\{f\left(S_t|Y^t\right)\}_{t=1}^T$*.*

Algorithm 2 (Random-Walk Metropolis algorithm)**.**

1. *Set the followings:*
   - $\tilde{\Sigma}$*, the covariance matrix of the proposal distribution*
   - $\theta^{(0)}$*, the initial draw*
   - n_{sim}*, the number of draws to be generated*

- n_{burn}, *the number of initial draws to be discarded*

2. *For* $s = 1$ *to* n_{sim}*:*

   - *Draw* ϑ *from the proposal distribution* $N_P\left(\theta^{(s-1)}, \tilde{\Sigma}\right)$.
   - *Set* $\chi^{(s)} = \min\left\{1, \frac{f\left(Y^T|\vartheta\right)p(\vartheta)}{f\left(Y^T|\theta^{(s-1)}\right)p\left(\theta^{(s-1)}\right)}\right\}$.
   - *Set* $\theta^{(s)}$ *as follows:*
     - *Set* $\theta^{(s)} = \vartheta$ *with probability* $\chi^{(s)}$.
     - *Otherwise, set* $\theta^{(s)} = \theta^{(s-1)}$.

3. $\left\{\theta^{(s)}\right\}_{s=n_{burn}+1}^{n_{sim}}$ *is an approximate realization from* $p\left(\theta|Y^T\right)$.

Algorithm 3 (Particle smoothing)**.**

1. *Smooth the shadow rate by moving backward as follows:*

   (a) *Choose* $\tilde{r}^*_T = r^{*,i}_T$ *with probability* w^i_T *and set* $\tilde{Y}^*_T \equiv (x'_T, \tilde{r}^*_T)'$.

   (b) *For* $t = T-1$ *to* 1*:*

   - *For each* $i = 1, \ldots, N$, *compute* $w^i_{t|t+1} \propto w^i_t f(\tilde{Y}^*_{t+1}|S^i_t)$ *with (B.3).*
   - *Choose* $\tilde{r}^*_t = r^i_t$ *with probability* $w^i_{t|t+1}$ *and set* $\tilde{Y}^*_t = (x'_t, \tilde{r}^*_t)'$.

2. *Recover the entire state vector by moving forward as follows:*

   (a) *Set* $\tilde{S}_0 = S_0$.

   (b) *For* $t = 1$ *to* T*:*

   - *Set* $\tilde{S}_t = \alpha + \Pi\tilde{S}_{t-1} + \Psi\left(\Lambda\Psi\right)^{-1}\left(\tilde{Y}^*_t - \gamma - \Phi\tilde{S}_{t-1}\right)$.

3. $\left\{\tilde{S}_t\right\}_{t=0}^{T}$ *is an approximate realization from* $f(S^T|Y^T)$.

References

[1] An, S., and F. Schorfheide, 2007. Bayesian analysis of DSGE models. Econometric Reviews 26(2-4), 113-172.

[2] Arulampalam, M. S., S. Maskell, N. Gordon, and T. Clapp, 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Transactions on Signal Processing 50(2), 174-188.

[3] Black, F., 1995. Interest rates as options. Journal of Finance 50, 1371-1376.

[4] Doucet, A., S. Godsill, and C. Andrieu, 2000. On sequential Monte Carlo sampling methods for Bayesian filtering. Statistics and Computing 10, 197-208.

[5] Fernandez-Villaverde, J. and J. F. Rubio-Ramirez, 2005. Estimating dynamic equilibrium economies: Linear and nonlinear likelihood. Journal of Applied Econometrics 20, 891-910.

[6] Godsill, S. J., A. Doucet, and M. West, 2004. Monte Carlo smoothing for nonlinear time series. Journal of the American Statistical Association 99(465), 156-168.

[7] Hamilton, J. D., 1994. Time Series Analysis. Princeton, New Jersey: Princeton University Press.

[8] Hara, N., N. Hirakata, Y. Inomata, S. Ito, T. Kawamoto, T. Kurozumi, M. Minegishi, and I. Takagawa, 2006. The new estimates of output gap and potential growth rate. Bank of Japan Review 2006-E-3.

[9] Ichiue, H., T. Kurozmi, and T. Sunakawa, 2008. Inflation dynamics and labor adjustments in Japan: A Bayesian DSGE approach. BOJ Working Paper Series No. 08-E-9.

[10] Ichiue, H. and Y. Ueno, 2007. Equilibrium interest rate and the yield curve in a low interest rate environment. BOJ Working Paper Series No.07-E-18.

[11] Ichiue, H. and Y. Ueno, 2008. Monetary policy and the yield curves at zero interest. Mimeo (available at http://hichiue.googlepages.com/home).

[12] Iwata, S. and S. Wu, 2006. Estimating monetary policy effects when interest rates are close to zero. Journal of Monetary Economics 53, 1395-1408.

[13] Klein, P., 2000. Using the generalized Schur form to solve a multivariate linear rational expectations model. Journal of Economic Dynamics & Control 24, 1405-1423.

[14] Lubik, T. A. and F. Schorfheide, 2004. Testing for indeterminacy: An application to U.S. monetary policy. American Economic Review 94(1), 190-217.

[15] Reifschneider, D., and J. C. Williams, 2000. Three lessons for monetary policy in a low inflation era. Journal of Money, Credit, and Banking 32(4), 936-966.

[16] Sims, C. A., 2002. Solving linear rational expectations models. Computational Economics 20, 1-20.

[17] Smets, F. and R. Wouters, 2003. An estimated dynamic stochastic general equilibrium model of the Euro area. Journal of European Economic Association 1(5), 1123-1175.

[18] Svensson, L. E. O., and N. Williams, 2007. Monetary policy with model uncertainty: distribution forecast targeting. Mimeo.

[19] Svensson, L. E. O., and N. Williams, 2008. Optimal monetary policy under uncertainty: a Markov jump-linear-quadratic approach. Federal Reserve Bank of St. Louis Review 90(4), 275-293.

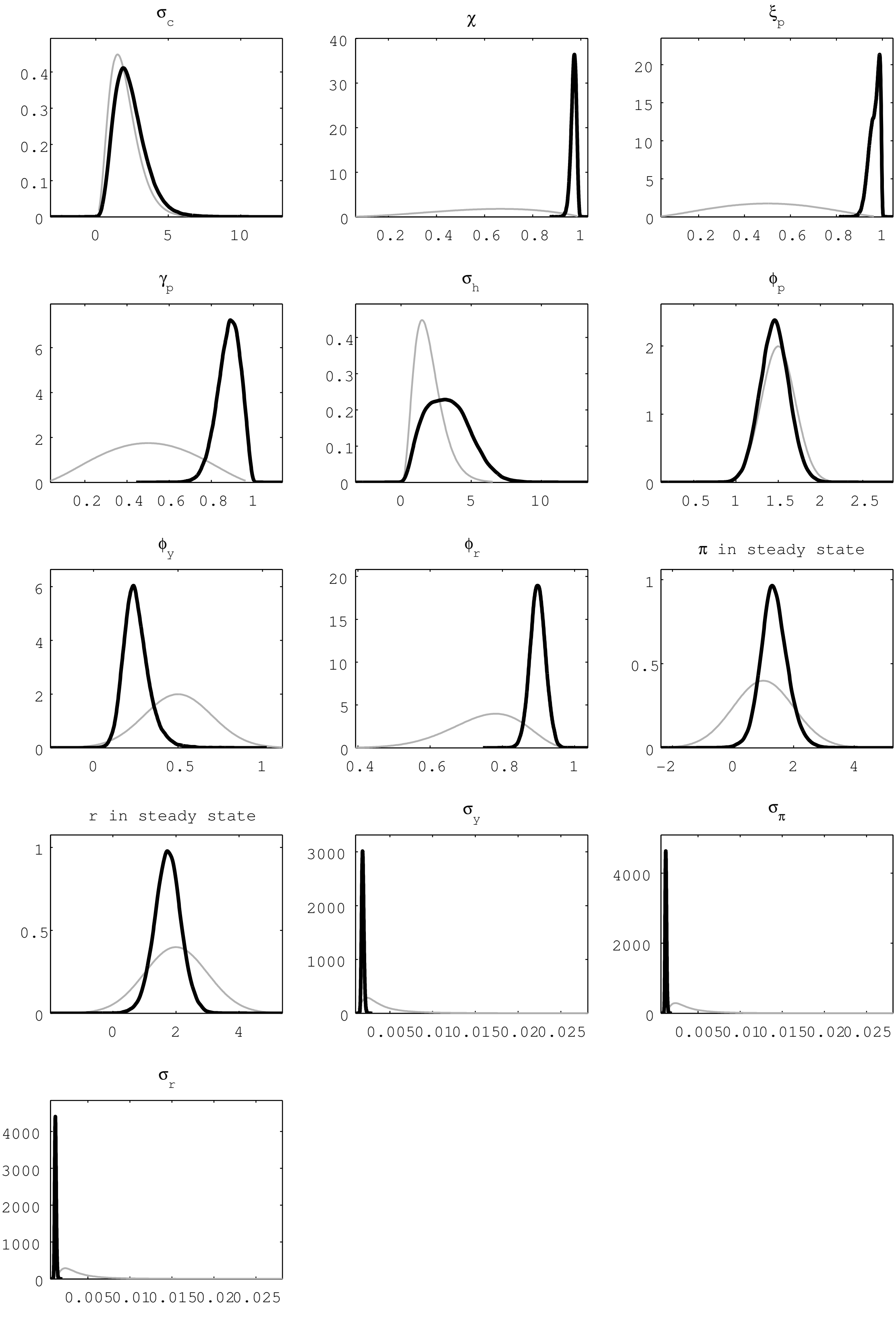
[20] Ugai, H., 2006. Effects of the quantitative easing policy: A survey of empirical analysis. BOJ Working Paper Series No.06-E-10.

Table 1. Prior specifications and parameter values at the posterior mean

Parameter		Prior		
		Type	Mean	90% interval
σ_c	Relative risk aversion	Gamma	2.00	[0.68,3.88]
χ	Habit persistence	Beta	0.60	[0.25,0.90]
ξ_p	Price fixity	Beta	0.50	[0.17,0.83]
γ_p	Price indexation	Beta	0.50	[0.17,0.83]
σ_h	Inverse of Frisch elasticity	Gamma	2.00	[0.68,3.88]
ϕ_π	Monetary policy reaction to inflation	Normal	1.50	[1.17,1.83]
ϕ_y	Monetary policy reaction to output gap	Normal	0.50	[0.17,0.83]
ϕ_r	Monetary policy inertia	Beta	0.75	[0.57,0.90]
$\bar{\pi}$	Steady state inflation	Normal	1.00	[-0.64,2.64]
$\bar{r}$	Steady state real rate	Normal	2.00	[0.36,3.64]
σ_y	S.d. of natural rate shock	Inverse gamma	0.005	[0.001,0.014]
σ_π	S.d. of price shock	Inverse gamma	0.005	[0.001,0.014]
σ_r	S.d. of monetary policy shock	Inverse gamma	0.005	[0.001,0.014]

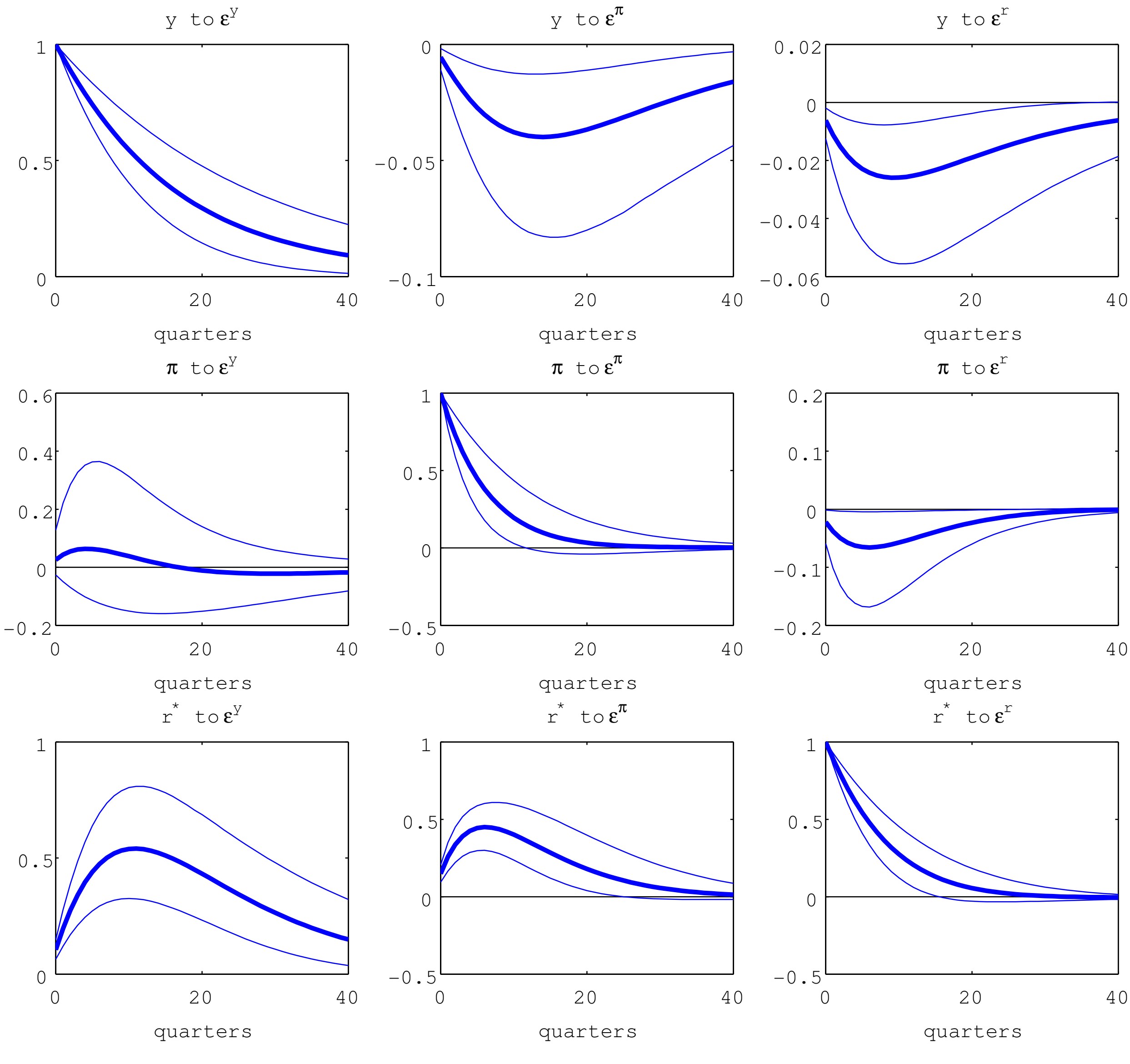
Parameter		Posterior	
		Mean	90% interval
σ_c	Relative risk aversion	2.378	[0.714,3.976]
χ	Habit persistence	0.970	[0.952,0.988]
ξ_p	Price fixity	0.966	[0.933,0.997]
γ_p	Price indexation	0.882	[0.798,0.971]
σ_h	Inverse of Frisch elasticity	3.324	[0.729,5.583]
ϕ_π	Monetary policy reaction to inflation	1.454	[1.170,1.733]
ϕ_y	Monetary policy reaction to output gap	0.253	[0.133,0.372]
ϕ_r	Monetary policy inertia	0.898	[0.864,0.933]
$\bar{\pi}$	Steady state inflation	1.355	[0.627,2.076]
$\bar{r}$	Steady state real rate	1.731	[1.041,2.420]
σ_y	S.d. of natural rate shock	0.0018	[0.0016,0.0021]
σ_π	S.d. of price shock	0.0012	[0.0010,0.0013]
σ_r	S.d. of monetary policy shock	0.0012	[0.0010,0.0013]

Figure 1: Prior and posterior densities



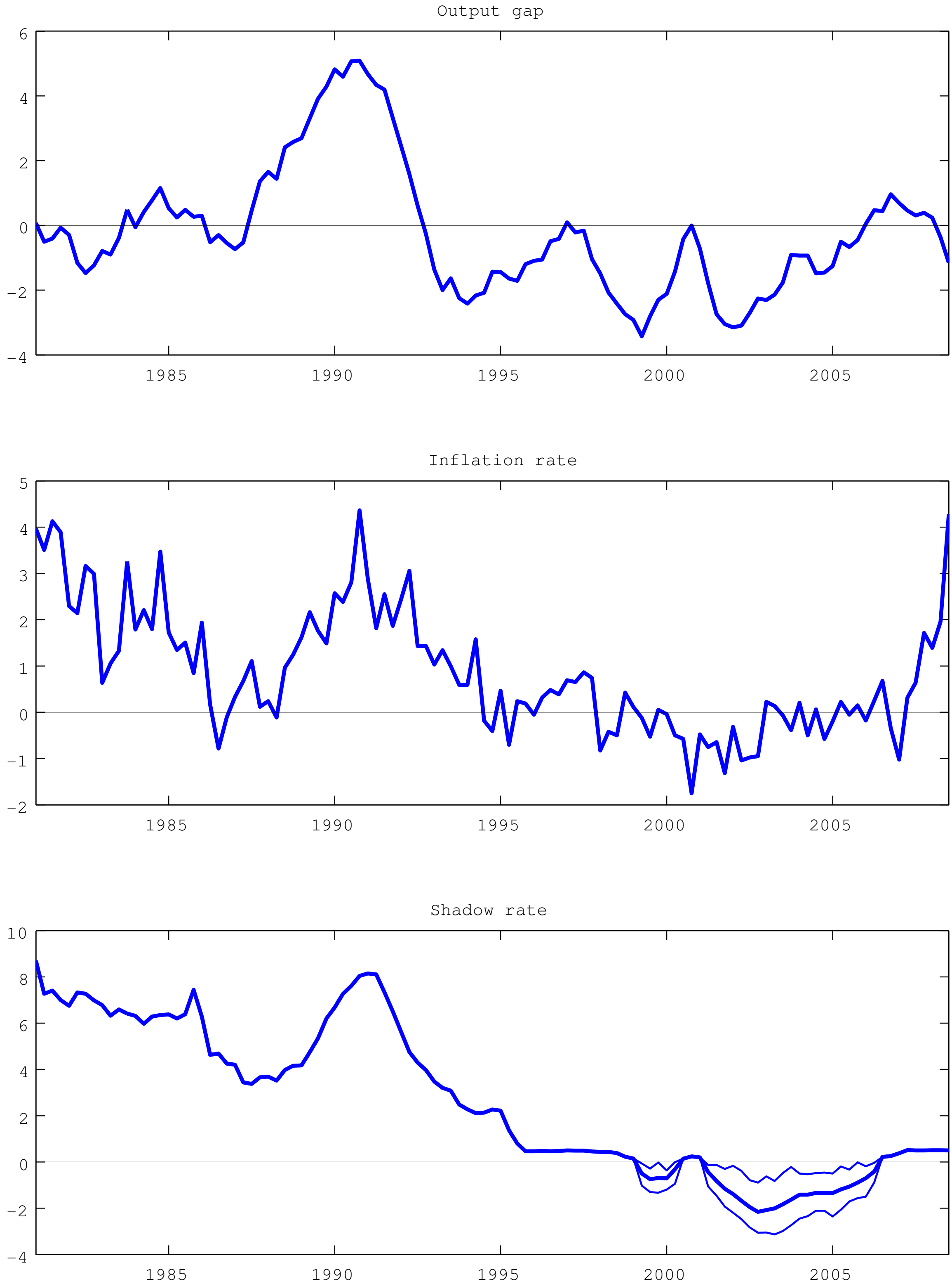
Note: The figure plots the marginal posterior distributions (black lines) against their marginal prior distributions (gray lines).

Figure 2: Impulse responses.



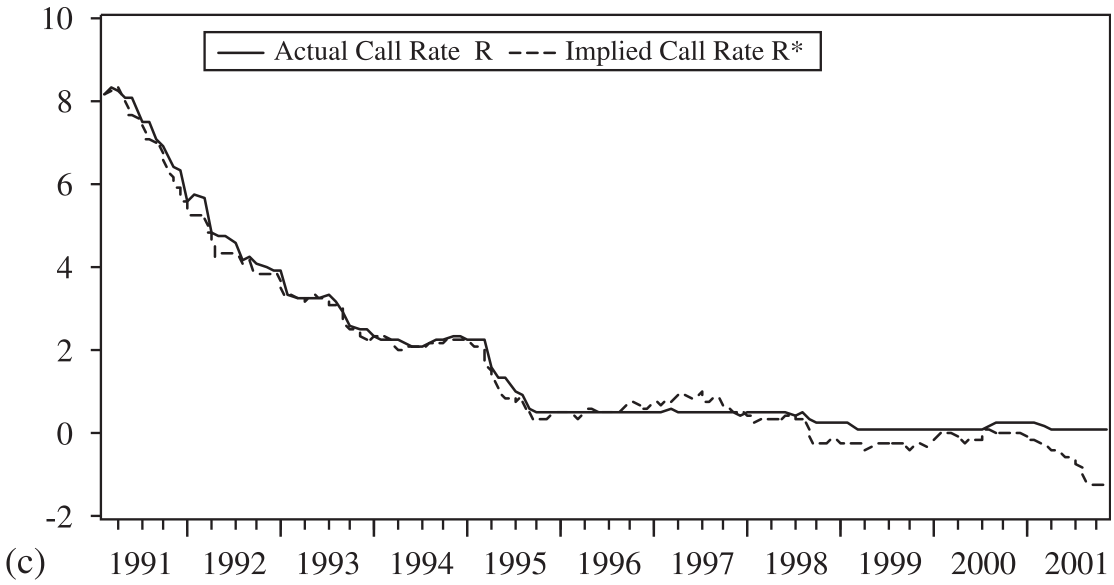
Note: The 90% confidence intervals are shown with thin lines.

Figure 3: Output gap, inflation rate, and shadow rate in Japan



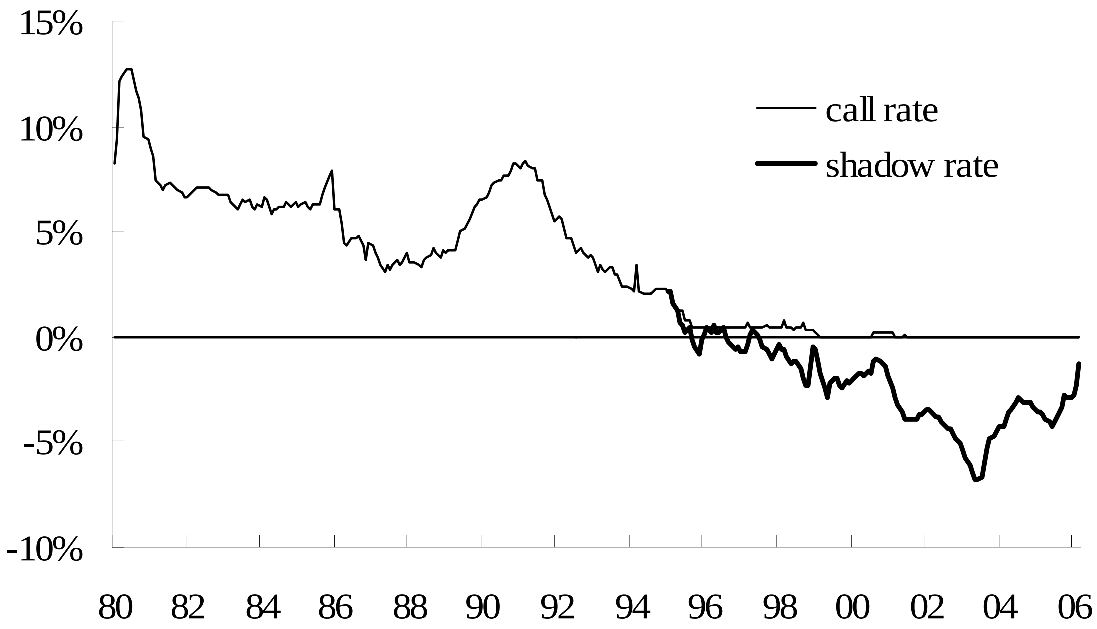
Note: The path of the shadow rate is the mean of 10,000 paths generated by the method described in Section 4. The 90% confidence interval is also shown.

Figure 4: The estimated shadow rate of Iwata and Wu (2006)



Note: The figure is taken from Iwata and Wu (2006).

Figure 5: The estimated shadow rate of Ichiue and Ueno (2008)



Note: The figure is taken from Ichiue and Ueno (2008).