Rapid estimation of nonlinear DSGE models

Hall, Jamie

UNSW

11 September 2012

Rapid estimation of nonlinear DSGE models

Jamie Hall*
School of Economics
University of New South Wales
jamie1212@gmail.com

September, 2012
This version: November 9, 2012

Abstract

This article describes a new approximation method for dynamic stochastic general equilibrium (DSGE) models. The method allows nonlinear models to be estimated efficiently and relatively quickly with the fully-adapted particle filter, without using high-performance parallel computation. The article demonstrates the method by estimating, on US data, a nonlinear New Keynesian model with time-varying volatility.

Keywords: Particle filter, New Keynesian.

JEL codes: C1, E0

1 Introduction

Nonlinear models of the macroeconomy can include a variety of features that are off limits to linear models, such as time-varying volatility or policy regime switching. These models have been explored in simulation studies over recent years, but estimation remains rare.[1] This is largely because of the computational difficulties involved. Finding an accurate nonlinear solution to a dynamic stochastic general equilibrium (DSGE) model can be time-consuming, and in order to estimate the model we must solve it a large number of times.

*I would like to thank Robert Kohn, Gianni Amisano and Bingbing Dong for useful comments and discussions. I also wish to acknowledge partial support from ARC grants LP0774950 and DP0988579.

[1]Notable exceptions are Fernández-Villaverde and Rubio-Ramírez (2007), Amisano and Tristani (2010) and Doh (2011).

This paper describes a method for taking nonlinear DSGE models to the data. The method is based on local linearisation, that is, a linearisation of the model's policy function conditional on the current state. Conditional linearity implies that the model's prediction density has full support over the space of possible observations, which is not true of (for instance) a second-order approximation. While that issue causes no difficulty in simulation, it can be a critical issue in estimating the model, as described below. Additionally, a conditionally linear model with Gaussian shocks is fully adapted in the sense of Pitt and Shephard (1999), which makes estimation very efficient. (Note that while this paper assumes the structural shocks are Gaussian, mixtures of Gaussians are also fully adapted.)

This is not the only possible answer: one could also use a higher-order local approximation, or a global approximation. In Sections 1.1 and 1.2, I outline these alternatives and argue that a local linearisation is also worth considering. Section 2 describes the method in general terms, and Section 3 shows how it fits into a particle filtering framework. Since these discussions are fairly abstract, Section 4 provides a worked example based on the neoclassical growth model. Finally, in Section 5, I report some estimation results for a nonlinear New Keynesian model.

1.1 Why not use a second-order approximation?

Simulation studies of nonlinear DSGE models often use a second-order Taylor approximation to the model solution (Schmitt-Grohé and Uribe 2004, Kim, Kim, Schaumburg, and Sims 2008). That is to say, the law of motion of the model's state variables is approximated by an autoregression with linear terms, squares, and cross-products. Since this approximation has proven satisfactory for simulation studies, why not use it for estimation?

One reason is illustrated in Figure 1. This graph shows the possible values of the observed variable (growth in the price-dividend ratio) as a function of the structural consumption shock ν_t, in an asset pricing model with external habits (Campbell and Cochrane 1999). The precise details of the model are not important for this illustration; see Appendix A for a full explanation. The three lines in the graph correspond to three possible approximations: linear, quadratic, and quartic. The quartic approximation is fairly close to the exact solution, while the linear approximation is quite inaccurate. The important thing to notice is that the quadratic approximation bends back on itself for $\nu_t < 0.007$ (a little over one standard deviation away from zero). In other words, conditional on the given value of the state x_{t-1}, it is impossible for the second-order approximation to generate an observation lower than -6%. This is not a problem of misspecification, since all three approximations in the chart are calculated using the same parameters and the same model. Suppose that the model and parameters used to make this graph represent the true data generating process. Then the quartic approximation indicates that the exact solution would generate a value of $\Delta \log y_t$ lower than -6% with a reasonable probability. Thus, even though the second-order approximation uses the correct model and the correct parameters, it would estimate the probability of that observation to be zero (in the absence of a noise term in the observation equation).

FIGURE 1 ABOUT HERE

In general, quadratic functions can approximate a curved function accurately on a neighbourhood of the origin, but the location and size of this neighbourhood depend on the model's parameters, making estimation difficult. Additionally, as illustrated in Figure 1, the even nature of a quadratic function can make it useless for estimation, because it is not absolutely continuous with respect to the process that generates the observations, even when both the parameters and the model are correct.

It would be possible to continue using second-order approximations by acting as if the data series were observed with a large amount of noise. This would be worth pursuing if no alternatives were available. The next section briefly discusses some possible alternatives, and the rest of the paper presents another one.

1.2 Why not use better global methods?

Recent work on DSGE models has employed more exotic approximation methods, which promise far greater accuracy than the first- or second-order Taylor approximations can deliver. This class of methods includes Smolyak polynomials (Fernández-Villaverde, Gordon, Guerrón-Quintana, and Rubio-Ramírez 2012), as well as Chebyshev polynomials and other projection methods (Judd 1998, Heer and Maußner 2009).[2]

Although these methods offer greater sophistication and accuracy, it is still worth considering locally linear approximations, for two reasons. The first is speed. Those global methods can be somewhat time-consuming, even using modern high-performance computing. This makes them unappealing for use in estimation, where the solution to the model might be recalculated many thousands of times. The second reason for maintaining an interest in a locally linear approximation is that the latter allows us to use the fully adapted particle filter, whereas global methods may not. As discussed in Sections 3 and 5, the fully adapted particle filter can be much more efficient than the standard version. In this respect, the locally linear approximation also improves on the second-order approximations discussed above. It is possible to use a variation on the auxiliary particle filter specially tailored for second-order approximations (Hall, Pitt, and Kohn 2012), but this cannot attain the efficiency of a fully adapted filter.[3]

Intuitively, the sequential Monte Carlo framework of particle filtering relies on representing the model's likelihood function with the product of a series of conditional likelihoods:

$$p(y_{1:T}|\theta) = \prod_{t=1}^{T} p(y_t|\theta, y_{1:(t-1)})$$

Therefore, the approximate solution of the model at time t does not need to be unconditionally accurate for all time. It only needs to maintain its validity into time $(t+1)$, at which point it can be recalculated. Thus the locally linear approach can attain higher accuracy close to the current state, at the cost of a larger discrepancy in other areas of the state space—but those areas are unlikely to be reached in a single step.

[2]Specially tailored solutions for particular situations are also possible, such as that of Amisano and Tristani (2011) for heteroscedastic models.

[3]It is also possible to remove some of the nonlinearity in a second-order model by ignoring the interaction between shocks to volatility and structural shocks, since they are of third-order importance (Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez 2012). However, the same comments would apply to the simplified second-order model.

2 General setup

In general, an economic model with optimising agents and rational expectations can be written in the form

$$\mathbb{E}_t F(c_{t+1}, c_t, k_t, k_{t-1}, z_t) = 0 \tag{1}$$

where $\mathbb{E}_t$ denotes an expectation conditional on date t information, c_t is a vector of choice variables (including forward-looking variables and jump variables), k_t is a vector of endogenous predetermined variables, and z_t is a vector of exogenous forcing variables. The vector-valued function F includes the law of motion of z_t, the equations determining k_t, and the equations that implicitly determine c_t as a function of k_{t-1} and z_t. I will write n_c for the number of components of c_t, and n_k and n_z for the length of k_t and z_t.

The goal is to find a corresponding expression for the model in state-space form, which can then be estimated. That is, we need functions of the form

$$c_t = \sigma(k_{t-1}, z_t) \tag{2}$$

$$k_t = \rho(k_{t-1}, z_t) \tag{3}$$

with the property that, when these values are used in F, equation (1) is satisfied. In general, these functions are not available in closed form, and need to be approximated.

2.1 Assumptions

To apply the method described in this paper, it is necessary to split the function F into its separate components, and make some assumptions about their structure. First, I assume that the law of motion of z_t can be written in log-linear form:

$$\log z_t = d + T_t \log z_{t-1} + R_t \epsilon_t \tag{4}$$

Here T_t and R_t are coefficient matrices which are assumed to be fixed, conditional on the vector of structural parameters θ and the values of z_{t-1} and k_{t-1}. The time subscripts indicate that the matrices are not assumed to be functions of θ alone. For example, time-varying volatility can be accommodated by including functions of z_{t-1} and/or k_{t-1} in the components of R_t. I assume that ϵ_t is a multivariate standard normal distribution: $\epsilon_t \sim N(0, I)$.[4]

To describe the equations determining the endogenous variables, we need the following notation: for each variable x_t, define $\widetilde{x}$ as the log-deviation of x_t from its nonstochastic steady state $\overline{X}$, i.e.

$$x_t = \overline{X} \exp(\widetilde{x}_t)$$

The second assumption I make is that each endogenous predetermined variable can be written as a function g of last period's predetermined variables, the current period's exogenous processes, and the choice variables:

$$k_t = g(\widetilde{k}_{t-1}, \widetilde{z}_t, \widetilde{c}_t) \tag{5}$$

The structure of a model will usually make the function g available in a natural way. If not, it could be calculated implicitly or approximated.

The final assumption concerns the equations that characterise the choice variables. In general, they can be written in the form

$$\mathbb{E}_t \left\{ f(\widetilde{k}_{t-1}, \widetilde{z}_t, \widetilde{c}_t, \widetilde{k}_t, \widetilde{z}_{t+1}, \widetilde{c}_{t+1}) \right\} = 0 \tag{6}$$

where f is some $\mathbb{R}^{n_c}$-valued function with the indicated arguments, and $\mathbb{E}_t$ denotes expectation conditional on date-t information. I assume that f can be re-expressed, approximately if necessary, to give an exponential function of the vector c_t as a weighted sum of N_c exponential terms:

[4] The methods described below can be generalised to allow each component of ϵ_t to be a finite mixture of distributions from the exponential family.

$$A_0 \circ \exp\left(D_0 \widetilde{c}_t\right) = \mathbb{E}_t \sum_{j=1}^{N_c} A_j \circ \exp\Big(B_j \widetilde{k}_{t-1} + C_j \widetilde{z}_t + D_j \widetilde{c}_t + E_j \widetilde{k}_t + F_j \widetilde{z}_{t+1} + G_j \widetilde{c}_{t+1} \Big) \tag{7}$$

Here, $\circ$ denotes the Hadamard product (elementwise multiplication) and the exponential is understood to operate elementwise. Thus the coefficient matrix D_0 is $n_c \times n_c$ dimensional, the vector A_i is $n_c \times 1$ dimensional, the matrix B_i is $n_k \times n_k$ dimensional, and similarly for the others. While the vector $\widetilde{c}_t$ can appear on both the left and right hand sides of equation (7), I assume that it is written so that the left-hand coefficient matrix D_0 is full rank. In the case of jump variables, which are determined by intratemporal optimality conditions only, the coefficients on the $(t+1)$-dated variables in (7) will be zero.

The approximation in equation (7) differs from the log-linearisation frequently used in macroeconomics in two ways. First, the approximation is not calculated at the nonstochastic steady state, but rather conditional on the values of k_{t-1} and z_t. Second, while the form of the equation may appear similar to the exponential of a loglinear structural model, the right-hand side of the equation is a weighted sum of exponential terms, rather than a single exponential of a sum. In other words, if the function f consists of several terms, then each term is loglinearised separately, conditional on k_{t-1} and z_t.

In many cases, the approximation in equation (7) will in fact be exact. This is true of all the equations in the examples below. If this type of exactness is not possible, the approximation can be calculated by log-differentiating the function f.

2.2 The approximate solution

In general, the solution of the model (1) involves expressing the choice variables c_t as determined by a policy function σ:

$$c_t = \sigma(z_t, k_{t-1}) \tag{8}$$

The function σ, which is intractable in general, is characterised by the fact that it satisfies (1) when substituted for c:

$$\mathbb{E}_t F(\sigma(z_{t+1}, k_t), \sigma(z_t, k_{t-1}), k_t, k_{t-1}, z_t) = 0$$

Now, suppose we approximate the unknown σ by expressing the vector of choice variables c_t as a log-affine function of $\widetilde{z}_t$ and $\widetilde{k}_{t-1}$:

$$c_t \approx \overline{C} \exp\left(\xi_t + \Omega_t \widetilde{z}_t + \Phi_t \widetilde{k}_{t-1}\right) \tag{9}$$

$$c_{t+1} \approx \overline{C} \exp\left(\xi_t + \Omega_t \widetilde{z}_{t+1} + \Phi_t \widetilde{k}_t\right) \tag{10}$$

where the vector ξ_t and the matrices Ω_t and Φ_t are functions of θ, z_t and k_{t-1}, corresponding to a first-order approximation of the policy function σ at those values. The rest of this section will describe how they can be calculated. I assume that unique initial values for Ω and Φ are available by log-linearising the model around its nonstochastic steady state using standard methods (Klein 2000, Sims 2001). In other words, I take it for granted that these matrices exist, that it is feasible to compute them, and that they are unique. The first contribution of this paper is to describe an efficient method for updating these approximations.

Suppose we use equations (9) and (10) to approximate equation (7). We can then substitute it into the right-hand side of equation (7).

$$\begin{aligned} A_0 \circ \exp\left(D_0 \xi_t + D_0 \Omega_t \widetilde{z}_t + D_0 \Phi_t \widetilde{k}_t\right) \approx \mathbb{E}_t \sum_{j=1}^{N_c} A_j \circ \exp\Big(& B_j \widetilde{k}_{t-1} + C_j \widetilde{z}_t + D_j \xi_t + D_j \Omega_t \widetilde{z}_t \\ & + D_j \Phi_t \widetilde{k}_{t-1} + E_j \widetilde{k}_t + F_j \widetilde{z}_{t+1} + G_j \xi_t + G_j \Omega_t \widetilde{z}_{t+1} + G_j \Phi_t \widetilde{k}_t\Big) \end{aligned} \tag{11}$$

We can then use the law of motion for z, equation (4), to replace $\widetilde{z}_{t+1}$ in (11):

$$
\begin{aligned}
A_0 \circ \exp\left(D_0\xi_t + D_0\Omega_t\widetilde{z}_t \right. & \left. + D_0\Phi_t\widetilde{k}_t\right) \\
& \approx \mathbb{E}_t \sum_{j=1}^{N_c} A_j \circ \exp\Big(B_j\widetilde{k}_{t-1} + C_j\widetilde{z}_t + D_j\xi_t + D_j\Omega_t\widetilde{z}_t + D_j\Phi_t\widetilde{k}_{t-1} \\
& \quad + E_j\widetilde{k}_t + G_j\xi_t \left[F_j + G_j\Omega_t\right] T_t\widetilde{z}_t + \left[F_j + G_j\Omega_t\right] R_t\epsilon_{t+1} + G_j\Phi_t\widetilde{k}_t\Big) \qquad (12)
\end{aligned}
$$

Conditional on date-t information, the only stochastic part in equation (12) is ϵ_{t+1}. So we can factor it as

$$
\begin{aligned}
A_0 \circ \exp\left(D_0\xi_t \right. & \left. + D_0\Omega_t\widetilde{z}_t + D_0\Phi_t\widetilde{k}_t\right) \\
& \approx \sum_{j=1}^{N_c} A_j \circ \exp\Big(\left[B_j + D_j\Phi_t\right]\widetilde{k}_{t-1} + \left[D_j + G_j\right]\xi_t \\
& \qquad + \left[C_j + D_j\Omega_t + F_jT_t + G_j\Omega_tT_t\right]\widetilde{z}_t + \left[E_j + G_j\Phi_t\right]\widetilde{k}_t\Big) \\
& \quad \circ \mathbb{E}_t \exp\left(\left[F_j + G_j\Omega_t\right] R_t\epsilon_{t+1}\right) \qquad (13)
\end{aligned}
$$

Each component of the expectation part is now in the form $\mathbb{E}(q'\epsilon_{t+1})$, where q' is a row of $\left[F_j + G_j\Omega_t\right] R_t$. In other words, it is equal to a value of the moment generating function of ϵ_{t+1}. Since ϵ is multivariate normal, the moment generating function is given by

$$
\mathbb{E}\exp\left(q'\epsilon\right) = \exp\left(\frac{1}{2}q'q\right) \qquad (14)
$$

Using (14) in (13), we obtain the following:

$$
\begin{aligned}
A_0 \circ \exp \big(D_0\xi_t \ &+ D_0\Omega_t\widetilde{z}_t + D_0\Phi_t\widetilde{k}_t\big) \\
&\approx \sum_{j=1}^{N_c} A_j \circ \exp \Big(\left[B_j + D_j\Phi_t\right] \widetilde{k}_{t-1} + \left[D_j + G_j\right] \xi_t \\
&+ \left[C_j + D_j\Omega_t + F_jT_t + G_j\Omega_tT_t\right] \widetilde{z}_t + \left[E_j + G_j\Phi_t\right] \widetilde{k}_t \\
&\quad + \frac{1}{2} K \left[F_j + G_j\Omega_t\right] R_tR_t' \left[F_j + G_j\Omega_t\right]' \Big)
\end{aligned}
\tag{15}
$$

where K is the $n_c \times n_c^2$ matrix that selects the diagonal elements of an $n_c \times n_c$ matrix.

The coefficient matrices T_t and R_t are assumed to be known at the start of time t, while the other coefficient matrices are determined by the economic equations defining the model (and by the value of the structural parameter vector θ). Taking the values of k_t as given, the only free parameters are in the c_t-approximation.

The values of of the coefficients in equation (9) are characterised by the fact that it is a first-order approximation of the policy function σ. We can therefore identify Ω and Φ by log-differentiating equation (15), and approximate ξ_t by ensuring that (15) holds with equality (to within a tolerable accuracy).

For notational convenience, let the vectors S_j be defined as

$$
\begin{aligned}
S_j = (A_j/A_0) \circ \exp \Big(&\left[B_j + D_j\Phi\right] \widetilde{k}_{t-1} + \left[D_j + G_j\right] \xi_t + \left[C_j + D_j\Omega + F_jT_t + G_j\Omega T_t\right] \widetilde{z}_t \\
&+ \left[E_j + G_j\Phi\right] \widetilde{k}_t + \frac{1}{2} K \left[F_j + G_j\Omega_t\right] R_tR_t' \left[F_j + G_j\Omega_t\right]' \Big)
\end{aligned}
\tag{16}
$$

That is, we rewrite equation (15) in the form $\exp(D_0\widetilde{c}_t) = \sum_{j=1}^{N_c} S_j$.

Log differentiating then gives the following expressions.

$$
\Omega_t = D_0^{-1} \sum_{j=1}^{N_c} \frac{S_j}{\sum_{m=1}^{N_c} S_m} \left[C_j + D_j\Omega_t + F_jT_t + G_j\Omega_tT_t + (E_j + G_j\Phi_t) \left(\frac{\partial \widetilde{k}_t}{\partial \widetilde{z}_t'} + \frac{\partial \widetilde{k}_t}{\partial \widetilde{c}_t'} \Omega_t \right) \right] \tag{17}
$$

$$\Phi_t = D_0^{-1} \sum_{j=1}^{N_c} \frac{S_j}{\sum_{m=1}^{N_c} S_m} \left[B_j + D_j \Phi_t + (E_j + G_j \Phi_t) \left(\frac{\partial \widetilde{k}_t}{\partial \widetilde{k}'_{t-1}} + \frac{\partial \widetilde{k}_t}{\partial \widetilde{c}'_t} \Phi_t \right) \right] \tag{18}$$

Finally, the vector ξ_t can be calculated from

$$\xi = D_0^{-1} \log \left(\sum_{j=1}^{N_c} S_j \right) - \Omega z_t - \Phi k_{t-1} \tag{19}$$

These three equations define a continuous self-map on the elements of the coefficient matrices. In computations, it is possible to update Ω, Φ and ξ by iterating the last three equations until convergence is achieved. Note that although the first two equations have some similarity to the forward-looking linear structures taken as inputs by the algorithms of Klein (2000) and Sims (2001), those methods are not applicable here, since the S_j factors are functions of the reduced-form solutions, and the final terms in equations (17) and (18) are quadratic functions of Ω and Φ.

In practice, I used Anderson acceleration to speed up the convergence of the fixed-point iterations (Anderson 1965). See Appendix B for more details.

2.3 Constraints on predetermined variables

It is often of interest to consider constraints on the endogenous predetermined variables of the form

$$C_1 \widetilde{k}_t \geq C_2 \tag{20}$$

In principle, a constrained solution to (11) can be obtained using the methods described here. The moment generating function part of equation (13) can be calculated using a generalisation of equation (14), for Gaussian random variables subject to inequality constraints (Tallis 1965).

However, when I carried this out on a model with a zero lower bound on the nominal interest rate, the fixed point iterations tended to diverge. This may be related to the unusual properties the zero lower bound: its effects are exceptionally sensitive to parameter values (Dong 2012), and it can be consistent with multiple equilibria (Braun, Körber, and Waki 2012). A solution may be possible within this framework, but specially tailored global methods are also available (Adjemian and Juillard 2010, Fernández-Villaverde, Gordon, Guerrón-Quintana, and Rubio-Ramírez 2012). I leave this analysis to future research.

3 Estimation method

While any nonlinear state-space algorithm could apply to the approximation described in Section 2, its locally linear quality makes it particularly well suited to the fully-adapted particle filter (Pitt and Shephard 1999). The standard particle filter approximates the posterior filtering density $p(x_t|y_t, x_{t-1})$ by sampling from the transition density $p(x_t|x_{t-1})$ and then reweighting by $p(y_t|x_t)$. The fully-adapted version samples directly from $p(x_t|y_t, x_{t-1})$. This is only possible if $p(x_t|x_{t-1})$ and $p(y_t|x_t)$ are conjugate in x. When it is possible, it can be several orders of magnitude more efficient than the standard particle filter, particularly when the observations y_t are highly informative about the underlying state x_t (Pitt, Silva, Giordani, and Kohn 2012). In this section I describe how to modify the approximation algorithm to take advantage of those features.

The locally linear approximation of $\widetilde{c}_t$ and $\widetilde{k}_t$ gives a system in the following form:

$$\widetilde{c}_t = \xi_t + \Omega_t \widetilde{z}_t + \Phi_t \widetilde{k}_{t-1} \tag{9}$$

$$\widetilde{k}_t = g\left(\widetilde{c}_t, \widetilde{z}_t, \widetilde{k}_{t-1}\right) \tag{5}$$

$$\widetilde{z}_t = T_t \widetilde{z}_{t-1} + R_t \epsilon_t \tag{4}$$

The coefficient matrices T_t and R_t are assumed to be given (conditional on last period's information), but the others depend on $\widetilde{z}_t$. Since changes in these coefficients in a single period are likely to be small, the previous period's values can serve as approximations to them. Additionally, we can take a linear approximation of $g()$ conditional on $\widetilde{z}_{t-1}$ and $\widetilde{k}_{t-1}$. This gives the following system:

$$\widehat{c}_t \approx \xi_{t-1} + \Omega_{t-1}\widetilde{z}_t + \Phi_{t-1}\widetilde{k}_{t-1} \tag{21}$$

$$\widehat{k}_t \approx \kappa_{t-1} + J_{t-1}\left(\widetilde{z}_t - \widetilde{z}_{t-1}\right) \tag{22}$$

where $\kappa_{t-1} = g\left(\varsigma_{t-1}, \widetilde{z}_{t-1}, \widetilde{k}_{t-1}\right)$, $\varsigma_{t-1} = \xi_{t-1} + \Omega_{t-1}\widetilde{z}_{t-1} + \Phi_{t-1}\widetilde{k}_{t-1}$, and $J = \left(\frac{\partial g}{\partial c'}\Omega_{t-1} + \frac{\partial g}{\partial z'}\right)$.

Suppose that the observation vector y_t consists of elements of $\widetilde{c}_t$ and $\widetilde{k}_t$ selected by projection matrices Z^c and Z^k, that is,

$$y_t = \begin{bmatrix} Z^c\widetilde{c}_t \\ Z^k\widetilde{k}_t \end{bmatrix}$$

Using the approximations just derived gives

$$y_t = \underbrace{\begin{bmatrix} Z^c\xi_{t-1} + Z^c\Phi_{t-1}\widetilde{k}_{t-1} \\ Z^k\kappa_{t-1} - Z^kJ_{t-1}z_{t-1} \end{bmatrix}}_{\mathbf{d_t}} + \underbrace{\begin{bmatrix} Z^c\Omega_{t-1} \\ Z^kJ_{t-1} \end{bmatrix}}_{\mathbf{Z_t}}\widetilde{z}_t \tag{23}$$

which can be used as the observation equation for a one-step Kalman filter, using the law of motion for $\widetilde{z}_t$ as the state equation (Harvey 1991). Thus the updated estimate of $\widetilde{z}_t$, given last period's state and the observation y_t, is

$$\widehat{z}_t = T_t\widetilde{z}_{t-1} + R_tR_t'\mathbf{Z_t}'\left(\mathbf{Z_t}R_tR_t'\mathbf{Z_t}' + H\right)^{-1}\left(y_t - \mathbf{Z_t}T_t\widetilde{z}_{t-1} - \mathbf{d_t}\right) \tag{24}$$

where H is the covariance matrix of measurement noise in y_t. The covariance of this estimate is given by

$$P_t = R_tR_t'\left[I - \mathbf{Z}_t'\left(\mathbf{Z_t}R_tR_t'\mathbf{Z_t}' + H\right)^{-1}\mathbf{Z}_tR_tR_t'\right] \tag{25}$$

The one-step-ahead forecast error and forecast covariance are given by

$$v_t = y_t - \mathbf{Z_t}T_t\widetilde{z}_{t-1} - \mathbf{d_t} \tag{26}$$

and

$$F_t = \left(\mathbf{Z_t} R_t R_t' \mathbf{Z_t}' + H\right) \tag{27}$$

Based on an estimate $\widehat{z}_t$, the implied values of the endogenous variables are given by (21) and (22). When H is small, so that observations are very informative about some elements of $\widetilde{c}_t$ and $\widetilde{k}_t$, the resulting values of $\widehat{c}_t$ and $\widehat{k}_t$ will match the observed y_t, but will not in general be equal to the values of $\widetilde{c}_t$ and $\widetilde{k}_t$ given by (9) and (5)—that is, after the coefficient matrices ξ, Ω and Φ are updated conditional on $\widetilde{z}_t = \widehat{z}_t$. However, this inaccuracy will be of second-order importance. To see this, note that

$$\begin{aligned}
\widetilde{c}_t(\widehat{z}_t) - \widehat{c}_t &= \xi(\widehat{z}_t) - \xi_{t-1} + \left[\Omega(\widehat{z}_t) - \Omega_{t-1}\right] \widehat{z}_t + \left[\Phi(\widehat{z}_t) - \Phi_{t-1}\right] \widetilde{k}_{t-1} \\
&\approx \frac{\partial \xi}{\partial z'}(\widehat{z}_t - \widetilde{z}_{t-1}) + o\left(\widetilde{z}^2\right)
\end{aligned}$$

Since the Taylor approximation of the policy function σ has an error of magnitude $o\left(\Delta \widetilde{z}^2\right)$, the change $\frac{\partial \xi}{\partial z'} \Delta z$ will be of a similar order of magnitude.

To recapitulate, these considerations suggest that a fully adapted particle filter can be implemented using the following steps at each time t:

1. Begin with values for $\widetilde{z}_{t-1}^{(m)}$, $\widetilde{k}_{t-1}^{(m)}$, $\xi_{t-1}^{(m)}$, $\Omega_{t-1}^{(m)}$ and $\Phi_{t-1}^{(m)}$ for particles indexed by $m = 1, \ldots, M$. Denote this information set as $\mathcal{F}_{t-1}^{(m)}$.

2. Calculate $v_t^{(m)}$ and $F_t^{(m)}$ from (26) and (27) for each particle, then resample the particles using weights given by $p(y_t|\mathcal{F}_t^{(m)}) = N(v_t^{(m)}, F_{t-1}^{(m)})$.

3. Calculate $\widehat{z}_t$ and P_t from (24) and (25), then draw $\widetilde{z}_t^{(m)} \sim N(\widehat{z}_t^{(m)}, P_t^{(m)})$.

4. Calculate ξ_t, Ω_t and Φ_t using equations (17), (18) and (19) from Section 2. Update $\widetilde{k}_t$ via equation (5).

An unbiased estimate of the likelihood is given by $\ell = \prod_{t=1}^{T} \frac{1}{M} \sum_{m=1}^{M} p(y_t|\mathcal{F}_{t-1}^{(m)})$. For proofs of its unbiasedness, as well as other properties, see Del Moral (2004) and Pitt, Silva, Giordani, and Kohn (2012).

4 Example 1: Neoclassical growth

In this section, I consider a basic neoclassical growth model. I choose this model because it is a useful and simple benchmark for solving and estimating DSGEs, used for example by Schmitt-Grohé and Uribe (2004) and Gomme and Klein (2011). The model is based on the decisions of a representative household, which chooses between consumption c_t and investment in next period's capital stock k_t.

I begin by showing how the model is written in the general framework of Section 2. I then evaluate the accuracy of the resulting approxmimations.

4.1 The model

The household's goal is to maximise discounted lifetime utility, given by

$$U = \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma}$$

subject to a feasibility constraint

$$c_t + k_t = A_t k_{t-1}^{\alpha} + (1-\delta) k_{t-1} \ , \tag{28}$$

where $\delta \in [0,1]$, and a productivity shock

$$\log A_t = \rho \log A_{t-1} + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2) \ . \tag{29}$$

where $\rho \in (0,1)$. The solution of the model consists of equations (28) and (29) plus a consumption Euler equation,

$$c_t^{-\gamma} = \beta \mathbb{E}_t \left\{ c_{t+1}^{-\gamma} \left[\alpha A_{t+1} k_t^{\alpha-1} + 1 - \delta \right] \right\} \tag{30}$$

4.2 Approximation

In terms of the notation in Section 2, here the choice variable is c_t, the endogenous predetermined variable is k_t, and the exogenous forcing variable is A_t. Equation (29) is exactly in the form of equation (4); thus $d = 0$, $T_t = T = \rho$, and $R_t = R = \sigma_\epsilon$. Note that the nonstochastic steady state of A is 1, so that $\widetilde{a}_t = \log A_t$. Equation (28) corresponds to equation (5):

$$\overline{K}\exp(\widetilde{k}_t) = \overline{K}^{\alpha}\exp(\widetilde{a}_t + \alpha\widetilde{k}_{t-1}) + (1-\delta)\overline{K}\exp(\widetilde{k}_{t-1}) - \overline{C}\exp(\widetilde{c}_t) \tag{31}$$

Equation (30) corresponds to (7):

$$\begin{aligned}\exp\left(-\gamma\widetilde{c}_t\right) = \mathbb{E}\beta\alpha\left(\overline{K}\right)^{(\alpha-1)}\exp\left[-\gamma\widetilde{c}_{t+1} + (\alpha-1)\widetilde{k}_t + \widetilde{a}_{t+1}\right] \\ + \mathbb{E}\beta(1-\delta)\exp\left(-\gamma\widetilde{c}_{t+1}\right)\end{aligned} \tag{32}$$

Thus, in the notation of equation (7), $A_0 = 1$, $A_1 = \beta\alpha\left(\overline{K}\right)^{(\alpha-1)}$, $A_2 = \beta(1-\delta)$, $B_1 = B_2 = 0$, $C_1 = C_2 = 0$, $D_0 = -\gamma$,$D_1 = D_2 = 0$, $E_1 = (\alpha-1)$, $E_2 = 0$, $F_1 = 1$, $F_2 = 0$, $G_1 = G_2 = -\gamma$.

Using the Gaussian mgf (14) and the linear approximations (9) and (10) then gives

$$\begin{aligned}&\exp\left(-\gamma\xi_t - \gamma\Omega_t\widetilde{a}_t - \gamma\Phi_t\widetilde{k}_{t-1}\right) = \\ &\quad\beta\alpha\left(\overline{K}\right)^{(\alpha-1)}\exp\left[-\gamma\xi_t + (1-\gamma\Omega_t)\rho\widetilde{a}_t + (\alpha-1-\gamma\Phi_t)\widetilde{k}_t + \frac{\sigma^2\left(1-\gamma\Omega_t\right)^2}{2}\right] \\ &\quad + \beta(1-\delta)\exp\left(-\gamma\xi_t - \gamma\Omega_t\rho\widetilde{a}_t - \gamma\Phi_t\widetilde{k}_t + \frac{\sigma^2\gamma^2\Omega_t^2}{2}\right)\end{aligned} \tag{33}$$

which is equal to equation (15).

The constraint is

$$\begin{aligned}\widetilde{k}_t = \log\Big[&\overline{K}^{(\alpha-1)}\exp(\widetilde{a}_t + \alpha\widetilde{k}_{t-1}) + (1-\delta)\exp(\widetilde{k}_{t-1}) \\ &-\frac{\overline{C}}{\overline{K}}\exp(\xi_t + \Omega_t\widetilde{a}_t + \Phi_t\widetilde{k}_{t-1})\Big]\end{aligned} \tag{34}$$

4.3 Results

How accurate is the approximation described in Section 4.2? I performed a numerical analysis to compare it to some other benchmarks. I chose standard values for most of the parameters: $\beta = 0.96$, $\alpha = 1/3$, $\rho = 0.9$, $\delta = 0.05$, and $\sigma_\epsilon = 0.02$. For the risk-aversion parameter γ, I chose the rather low value of 0.5, to place more emphasis on the nonlinear character of the model. If γ is large, then the model is almost log-affine and accurate approximation is less important. I made this choice because the intention is to demonstrate a method that can be usefully applied to more complex models with stronger nonlinear features.

Since this model has only one shock, it is possible to calculate the expectation in equation (30) numerically, given a policy function $\widetilde{c}_t = \sigma(\widetilde{z}_t, \widetilde{k}_{t-1})$, and therefore to estimate the policy function's Euler error. (This procedure is of course not feasible for larger models.) I used a 50×50-point grid on $z_t \in [-0.2, 0.2]$ and $\widetilde{k}_{t-1} = [-0.6, 0.6]$ to do so.

I applied this method to calculate the Euler errors for three different methods: the locally linear approximation proposed in this paper, a loglinearisation, and a second-order perturbation. The loglinearisation can be calculated using standard methods (Klein 2000, Sims 2001). The second-order perturbation is described in Schmitt-Grohé and Uribe (2004), Kim, Kim, Schaumburg, and Sims (2008) and Gomme and Klein (2011).

Figure 2 shows the squared Euler errors for each approximation method, on a logarithmic scale. The quadratic approximation is comparable to the locally linear one at certain points of the grid close to the origin. However, the quadratic approximation does not maintain this accuracy in the tails of the state variables. The upper left and lower right panels of the Figure are evaluated at values of $\widetilde{k}_{t-1}$ around 5 standard deviations away from the steady state. They illustrate one advantage of the local linearisation method: while these values of the state variable would rarely be encountered in a simulation, they would be more common in estimation, since the level of the nonstochastic steady state is a function of the structural parameters. Thus, while searching through the parameter space using a given series of real observations, it is advantageous to maintain accuracy throughout the state space.

The *average* Euler errors are summarised in Table 1. The average squared errors, in the first column of results, were estimated with respect to the stationary distributions of z and k. The $\|\eta\|_\infty$ estimates were calculated with the state variables restricted to a distance of three standard deviations from their steady states.

FIGURE 2 ABOUT HERE

TABLE 1 ABOUT HERE

5 Example 2: A basic New Keynesian model

In this section, I consider a baseline New Keynesian macro model. There are many possible variations on the basic structure. I use a simplified version of the model based on Amisano and Tristani (2010) and Fernández-Villaverde, Gordon, Guerrón-Quintana, and Rubio-Ramírez (2012).[5] It includes a representative household with a utility function separable in consumption c_t and hours worked l_t; a continuum of profit-maximising goods producers in monopolistic competition, with sticky prices; and a government sector that sets the nominal interest rate through a Taylor rule. Investment is not modelled, with the capital stock instead being taken as fixed.

5.1 The model

The household's objective is to maximise utility, given by

$$U_t = \sum_{h=0}^{\infty} \beta^h \mathrm{e}^{b_{t+h-1}} \left(\frac{C_{t+h}^{1-\gamma}}{1-\gamma} - \mathrm{e}^{e_{t+h}} \frac{1}{1+\phi} l_{t+h}^{1+\phi} \right) \tag{35}$$

where C_t is consumption, l_t is hours worked, b_t is an exogenous disturbance representing demand-side shocks, and e_t is an exogenous disturbance representing labour-supply shocks. The maximisation is subject to a budget constraint

$$P_t C_t + B_t = W_t l_t + R_{t-1} B_{t-1} \tag{36}$$

[5] See also Woodford (2003) and Galí (2008) for further details and background.

The household's income is derived from a nominal wage W_t and a gross return R_{t-1} paid on risk-free nominal bonds B_{t-1}. This can be spent on consumption and on saving for next period. P_t is the level of the consumer price index in period t. To keep the model simple, I assume that taxes and transfers offset any profits accruing to the household.

The resulting first-order conditions of the household are

$$\lambda_t = \mathbb{E}_t \beta e^{b_t} \lambda_{t+1} \frac{R_t}{\Pi_{t+1}} \tag{37}$$

$$C_t^{-\gamma} = \lambda_t \tag{38}$$

$$\frac{W_t}{P_t} \lambda_t = e^{e_t} l_t^{\phi} \tag{39}$$

Equation (37) is an intertemporal consumption Euler equation, connecting the marginal utility of consumption λ_t with its expected value next period, deflated by a time preference factor $\beta \in (0,1)$ and the expected real interest rate. Π_{t+1} is the ratio of the consumer price index in period $t+1$ and period t.

Equation (39) is an intratemporal optimality condition, equating the marginal disutility of labour (l_t^{ϕ}) to the marginal benefit of increased consumption ($\frac{W_t}{P_t}\lambda_t$).

There is a continuum of firms in a monopolistically competitive market. The production function of the representative firm is

$$Y_t(i) = A_t l_t(i)^{\alpha} \tag{40}$$

Here A_t is a technology shock and $l_t(i)$ is the amount of labour hired by firm i. It follows that real marginal cost for the representative firm is given by

$$MC_t = \frac{W_t/P_t}{A_t} \tag{41}$$

I make the standard assumption that firms are able to change their price with fixed probability θ_p each period (Calvo 1983). The firm's problem is to choose a price in order to maximise expected profits subject to this constraint, and subject to a constant elasticity of demand θ. Let the auxiliary variables $G_{1,t}$ and $G_{2,t}$ equal the present values of marginal cost and marginal revenue. The profit maximisation conditions are then given by

$$\theta G_{1,t} = (\theta - 1) G_{2,t} \tag{42}$$

$$G_{1,t} = \lambda_t M C_t Y_t + \beta \mathrm{e}^{b_t} \theta_p \mathbb{E}_t \Pi_{t+1}^{\theta} G_{1,t+1} \tag{43}$$

$$G_{2,t} = \Pi_t^{\star} \left(\lambda_t Y_t + \beta \mathrm{e}^{b_t} \theta_p \mathbb{E}_t \frac{\Pi_{t+1}^{\theta-1}}{\Pi_{t+1}^{\star}} G_{2,t+1} \right) \tag{44}$$

Aggregate CPI inflation is given by

$$1 = \theta_p \Pi_t^{\theta-1} + (1 - \theta_p)(\Pi_t^{\star})^{1-\theta} \tag{45}$$

If the model is loglinearised around its nonstochastic steady state, with steady-state inflation $\log \Pi = 0$, then equations (42) to (45) collapse into the familiar New Keynesian Phillips curve.

The market clearing conditions are

$$Y_t = C_t \tag{46}$$

$$B_t = 0 \tag{47}$$

Interest rate policy is set according to

$$R_t = R^{1-\rho_r} R_{t-1}^{\rho_r} \left[\left(\frac{\Pi_t}{\Pi} \right)^{\phi_\pi} \left(\frac{y_t}{y} \right)^{\phi_y} \right]^{1-\rho_r} m_t \eta_{r,t} \tag{48}$$

This is a standard Taylor Rule for monetary policy. The exogenous process m_t is a persistent monetary-policy shock (similar to a time-varying inflation target) and $\eta_{r,t}$ is a transitory one.

The model is closed by specifying the laws of motion for the exogenous processes. These are assumed to be as follows:

$$A_t = A^{1-\rho_A} A_t^{\rho_a} \exp(\sigma_{a,t} \epsilon_{a,t}) \tag{49}$$

$$b_t = \rho_b b_{t-1} + \sigma_{b,t} \epsilon_{b,t} \tag{50}$$

$$m_t = \rho_m m_{t-1} + \sigma_{m,t} \epsilon_{m,t} \tag{51}$$

$$e_t = \rho_e e_{t-1} + \sigma_{e,t}\epsilon_{e,t} \tag{52}$$

$$\eta_{r,t} = \sigma_{r,t}\epsilon_{r,t} \tag{53}$$

Finally, I assume that each volatility $\sigma_{i,t}$ is determined by an independent GARCH process:

$$\sigma^2_{i,t} = \alpha^2_i + \beta_i\epsilon^2_{i,t-1} + \gamma_i\sigma^2_{i,t-1} \tag{54}$$

It is convenient to model time-varying volatility with a GARCH process, because GARCH is fully adapted (Pitt, Silva, Giordani, and Kohn 2010). In other words, the coefficients R_t in equation (4) are completely determined, conditional on $\widetilde{z}_{t-1}$. Therefore, the method described in Section 3 can be used without modification.

5.2 Approximation

It is straightforward to express this model in the form used in Section 2. As in the growth model example, the expressions in equations (7) and (5) are exact in this case. While the algebra involved in these expressions is not difficult, it is not as compact as in the previous example, so is reported in Appendix C. In brief, the choice variables of the model are $\widetilde{c}_t = (\widetilde{\pi}_t, \widetilde{\pi}^\star_t, \widetilde{G}_{1,t}, \widetilde{G}_{2,t}, \widetilde{\lambda}_t, \widetilde{l}_t, \widetilde{MC}_t, \widetilde{y}_t)$, the predetermined variables are $\widetilde{k}_t = (\widetilde{r}_t)$, the exogenous variables are $\widetilde{z}_t = (\widetilde{b}_t, \widetilde{a}_t, \widetilde{m}_t, \widehat{e}_t, \eta_{r,t}, \sigma^2_{b,t}, \sigma^2_{a,t}, \sigma^2_{m,t}, \sigma^2_{e,t}, \sigma^2_{r,t})$, and the shocks are $\epsilon_t = (\epsilon_{a,t}, \epsilon_{b,t}, \epsilon_{m,t}, \epsilon_{e,t}, \epsilon_{r,t})$.

For each set of parameter values considered, the locally linear approximation was recalculated at each time period. The iterations of the Anderson method (Appendix B) were continued until all coefficients were stable to at least four decimal places. Because the coefficients were very similar across the particle swarm at each time period, I updated the approximation only once each period, at the previous period's mean value of $\widetilde{z}_{t-1}$ and $\widetilde{k}_{t-2}$—that is, the sample mean of the particle swarm after the resampling step. This produced loglikelihood estimates that were indistinguishable from the case where the approximation was updated individually for each particle, and enabled a considerable improvement in speed. This simplification is also consistent with the analysis in Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2012), which shows that changes in volatility are of third-order importance for the determination of the choice variables. Note that each particle maintained separate estimates of the volatilities $\sigma_{i,t}$ at each time period.

5.3 Estimation

I used a fully-adapted particle filter, as described in Section 3, to estimate the model on US data, using 50 particles. For comparison, I estimated a loglinearised version of the model using the Kalman filter. I chose the number of particles for the nonlinear estimation by taking repeated estimates of the loglikelihood at the estimated mode of the linear approximation. With 50 particles, the estimates of the loglikelihood had a standard deviation of around 1, which is optimal for a Metropolis Hastings run (Pitt, Silva, Giordani, and Kohn 2012).

In the linear case, the parameters of the model are β, θ, θ_p, ϕ, ϕ_π, ϕ_y, Π, ρ_r, ρ_b, σ_b, α, γ, ρ_a, σ_a, ρ_e, σ_e, σ_r, ρ_m, and σ_m, a total of 19 free parameters.[6] For the nonlinear model, one β_i and one γ_i GARCH coefficient for each exogenous shock adds ten additional parameters. I chose independent prior distributions for these parameters based on those used in Smets and Wouters (2003), Smets and Wouters (2007), and Amisano and Tristani (2010). They are summarised in Table C.

The observable variables are $\log y_t$, $\log \Pi_t$ and R_t. I used quarterly data from 1984Q4 to 2011Q4 in the FRED (Federal Reserve Economic Data) database from the Federal Reserve Bank of St. Louis. For $\log y_t$, I used quarterly chain-volume GDP (code GDPC1); for $\log \Pi_t$, quarterly CPI inflation for all urban consumers, excluding energy (code CPILFESL_PCH); and for R_t, the Federal Funds rate (code FEDFUNDS).[7] Following Smets and Wouters (2003) and Amisano and Tristani (2010), I substracted a loglinear trend from the GDP series prior to estimation. The other two series were estimated relative to the steady state implied by each draw of the parameter vector. All data series were seasonally adjusted prior to estimation.

I assumed that the data series were observed with a small amount of noise. Specifically, each observation was assumed to be affected by a mean-zero iid Gaussian shock with variance 10^{-8}. (This is negligible compared to the size of the structural shocks; it is simply a convenient device

[6] The steady-state level of technology, A, was calibrated to 1.

[7] I used the ex-energy CPI series because the rapid change in energy prices in late 2008 is difficult to account for in a simple DSGE model.

for avoiding stochastic singularity.)

For both the linear and nonlinear estimation, the parameters were estimated using an adaptive random walk Metropolis Hastings algorithm (Haario, Saksman, and Tamminen 2001). The MCMC chains for both models were initialised at the estimated mode of the linear model. For the non-linear model, I initialised all the GARCH parameters at 0.1. In both cases, I took 100,000 draws, discarding the first 50,000 as a burn-in, with adaptation beginning after 1000 draws. I initialised the proposal covariance matrix with the diagonal of the estimated Hessian at the linear mode, with small positive values for the GARCH parameters.

The code for both estimation methods was written predominantly in MATLAB, with some parts of both methods written in C++. The loglinear approximation at the model's steady state was carried out using Dynare (Adjemian, Bastani, Karamé, Juillard, Maih, Mihoubi, Perendia, Ratto, and Villemot 2012). I ran the code on an Intel Core i7-880 workstation.

TABLE 2 ABOUT HERE

5.4 Results

Using the locally linear approximation method, the estimation run took roughly 20 times longer than in the linear case (Table 3). This is in the vicinity of the fastest possible time; with 50 particles, we must perform a one-step Kalman filter calculation 50 times per observation. On a multi-core computer, the symmetry of the particle filter algorithm allows many operations to be vectorised, reducing the required time. (MATLAB implemented this automatically in low-level computations, without explicit parallelism in the estimation algorithm.) The additional overhead, which is mainly due to having to update the nonlinear approximation, could perhaps be reduced by redesigning the model, for instance by substituting out some of the jump variables. I chose not to do that in this case, in order to demonstrate that this paper's nonlinear approximation method is feasible to use on a reasonably sized multivariate model.

TABLE 3 ABOUT HERE

Turning to the results, we see that the nonlinear model fits the data considerably better (Table 4). Its marginal logposterior is around 15 points higher than the linear model's, indicating that a Bayes Factor ratio would decisively prefer it.[8] However, the evidence is more mixed once we penalise the nonlinear version for its higher number of free parameters. The Akaike Information Criterion (Akaike 1974) and the Bayesian Information Criterion (Schwarz 1978) both favour the linear model, while the Deviance Information Criterion (Spiegelhalter, Best, Carlin, and Van Der Linde 2002) suggests that the nonlinear model fits better.

TABLES 4 AND 5 ABOUT HERE

The posterior estimates for individual parameters are somewhat different for the linear and nonlinear models (Table C). The nonlinear model's estimates are more precise for some components, but more diffuse for others. For many parameters, the estimated linear and nonlinear posterior distributions do not overlap to any significant degree, notably in the case of the risk aversion coefficient γ and the inflation response coefficient ϕ_π.

6 Conclusion

A locally linear approximation of a DSGE model has two features that can make it useful in estimation. Because it can be recalculated at each value of the latent state vector, it maintains a high degree of accuracy throughout the possible state space. And, because it is conditionally linear, it fits naturally into the framework of fully-adapted particle filtering. A broad range of nonlinear models can be included in that framework: for instance, models with time varying structural parameters, GARCH errors, or mixture-of-normal errors. Thus, the method described in this paper provides a practical method for nonlinear applied models with those characteristics to be brought to the data.

[8] I estimated the marginal logposteriors of both models using the method of Gelfand and Dey (1994), as implemented by Geweke (1999).

A The model used for Figure 1

Figure 1 is based on the model described in Campbell and Cochrane (1999). The model assumes that the representative agent's consumption process is

$$\Delta \log C_t = g + \nu_t \tag{55}$$

where $\nu \sim N(0, \sigma^2)$. The agent's utility function is given by

$$U_t = E_t \sum_{t=0}^{\infty} \beta^t \frac{(C_t - X_t)^{1-\gamma}}{1-\gamma} \tag{56}$$

where X_t is the (external) habit stock, interpreted as the minimum level of consumption required to maintain a well-defined utility (i.e., the household must ensure that $C_t > X_t$). The surplus consumption ratio S_t is defined by

$$S_t = \frac{C_t - X_t}{C_t}$$

and $\widetilde{s}_t = \log S_t - \log \overline{S}$ is the deviation of $\log S_t$ from its mean $\overline{S}$.

The law of motion of $\widetilde{s}_t$ is assumed to be

$$\widetilde{s}_t = \phi \widetilde{s}_{t-1} + \left(\frac{1}{\overline{S}} \sqrt{1 - 2\widetilde{s}_{t-1}} - 1 \right) \nu_t \tag{57}$$

where the disturbance ν_t is the same as the consumption innovation in equation (55), and the steady-state level of S_t is given by

$$\overline{S} = \sigma \sqrt{\frac{\gamma}{1-\phi}} \tag{58}$$

On that basis, it can be shown that the equilibrium price-dividend ratio of a financial asset satisfies

$$\frac{P_t}{D_t} = \beta_t \mathbb{E}_t \left[\exp\left[\gamma(\widetilde{s}_t - \widetilde{s}_{t+1}) + (1-\gamma)(g + \nu_{t+1})\right] \left(1 + \frac{P_{t+1}}{D_{t+1}}\right) \right] \tag{59}$$

Figure 1 shows the growth rate $\Delta \log \frac{P_t}{D_t}$ as a function of ν_t, conditional on $\widehat{s}_{t-1} = -.45$, with the level $\log \frac{P_t}{D_t}$ approximated by Taylor series of different orders. The value $\widehat{s}_{t-1} = -.45$ is well below the steady state, but has a reasonable chance of being observed. (Close to the steady state, the first- and second-order approximations will work much better, by construction.) The parameter values used are $r^f = 0.0025$, $g = .00444$, $\sigma = .00555$, $\gamma = 2.372$, $\phi = 0.97$.

B Computation of ξ, Ω and Φ

I computed the locally linear approximations using Anderson acceleration applied to the fixed-point equations (17), (18) and (19). Briefly, Anderson acceleration is similar to Newton's method for root-finding; the Jacobian is estimated by regressing previous fixed-point residuals on the corresponding fixed-point outputs. The algorithm provides quadratic convergence, like Newton's method, but is often more robust.

Write $x_k = (\xi_k, \operatorname{vec}\Omega_k, \operatorname{vec}\Phi_k)$ for the k^{th} estimate of the fixed point, g_{k+1} for the output of the fixed-point equations applied to x_k, and let $f_{k+1} = g_{k+1} - x_k$ be the residual. Let F_{k+1} be the matrix with columns given by $f_i - f_{k+1}$, for $i = (k - m_k), \ldots, k$; the number of columns, m_k, is given by $\min(k, m)$, where m is a control parameter of the algorithm. After some experimentation, I chose $m = 5$. Let G_{k+1} be the similar-sized matrix with columns given by g_{k-m_k-1} to g_{k+1}, and X_{k+1} consist of x_{k-m_k-1} to x_{k+1}.

The next iterate is then given by solving the least-squares problem

$$\vartheta = \left(X'_{k+1} X_{k+1}\right)^{-1} X'_{k+1} f_k \tag{60}$$

then setting $\alpha = (\frac{1}{1+\sum\vartheta}, \frac{\vartheta_1}{1+\sum\vartheta}, \ldots)$, and calculating the next iterate as

$$x_{k+1} = \beta G_{k+1}\alpha + (1-\beta) X_{k+1}\alpha$$

where $\beta \in (0, 1]$ controls the speed of adjustment. After experimenting with different values, I used $\beta = 1$.

There are various methods for computing equation (60)—see Fang and Saad (2009) and Walker and Ni (2011) for analysis and discussion. I used the QR decomposition. Following those two sources, I restarted the iterations whenever the condition number of the R matrix exceeded 10^5 (indicating that the accuracy of the least-squares solution had degraded) and whenever $\|f_k\| > 0.2\|f_{k-1}\|$.

C New Keynesian example

The New Keynesian model used in Section 5 is characterised by 8 equations for the choice variables: (42), (43), (44), (38), (37), (45), (40) and (39); one equation for the endogenous predetermined variable (48); and three equations for the exogenous shocks (49), (50) and (51).

Using the notation of Section 2, the components of the state vector are

$$\widetilde{c}_t = \begin{bmatrix} \widetilde{g}_{1,t} \\ \widetilde{g}_{2,t} \\ \widetilde{\pi}^\star_t \\ \widetilde{\lambda}_t \\ \widetilde{y}_t \\ \widetilde{\pi}_t \\ \widetilde{l}_t \\ \widetilde{mc}_t \end{bmatrix} \qquad \widetilde{k}_t = [\widetilde{r}_t] \qquad \widetilde{z}_t = \begin{bmatrix} \widetilde{a}_t \\ \widetilde{b}_t \\ \widetilde{m}_t \\ \widetilde{e}_t \\ \eta_{r,t} \end{bmatrix}$$

(To economise on space, I omit the volatilities $\sigma^2_{i,t}$ from the $\widetilde{z}_t$ vector.) The model's coefficient matrices are

$$A_0 = \begin{bmatrix} (\theta-1)G_2 \\ G_1 \\ \frac{G_1}{\Pi^\star} \\ Y^{-\gamma} \\ 1 \\ \theta_p \Pi^{\theta-1} \\ l \\ MC \end{bmatrix} \quad D_0 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\gamma & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (\theta-1) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} \theta G_1 \\ \lambda MCY \\ \lambda Y \\ \lambda \\ 1 \\ -(1-\theta_p)(\Pi^\star)^{1-\theta} \\ Y^{\frac{1}{\alpha}} \\ l^\phi/\lambda \end{bmatrix} \quad B_1 = \mathbf{0}_{8\times 1} \quad C_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{\alpha} & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$D_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-\theta) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\alpha} & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & \phi & 0 \end{bmatrix} \quad E_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad F_1 = \mathbf{0}_{8\times 5}$$

1

$$G_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 \\ \beta\theta_p\left(\Pi^{\theta}\right)G1 \\ \beta\theta_p\left(\Pi^{\theta-1}\right)\frac{G2}{\Pi^{\star}} \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad B_2 = \mathbf{0}_{8\times 1} \quad C_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad D_2 = \mathbf{0}_{8\times 8} \quad E_2 = \mathbf{0}_{8\times 1}$$

2

$$F_2 = \mathbf{0}_{8\times 5} \quad G_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & \theta & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & (\theta-1) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Finally, the steady state of the model is:

$$R = \frac{\Pi}{\beta} \qquad \Pi^\star = \left(\frac{1-\theta_p \Pi^{\theta-1}}{1-\theta_p}\right)^{\frac{1}{1-\theta}}$$

$$Y = \left[\Pi^\star \frac{(\theta-1)}{\theta} \frac{1-\beta\theta_p\Pi^\theta}{(1-\beta\theta_p\Pi^{\theta-1})}^{\frac{1}{\frac{\phi}{\alpha}+\gamma}}\right]$$

$$\lambda = Y^{-\gamma} \qquad l = Y^{\frac{1}{\alpha}} \qquad MC = Y^{\frac{\phi}{\alpha}+\gamma}$$

$$G_1 = \frac{Y^{1+\frac{\phi}{\alpha}}}{1-\beta\theta_p\Pi^\theta} \qquad G_2 = \frac{\theta}{\theta-1} G_1$$

References

Adjemian, S., H. Bastani, F. Karamé, M. Juillard, J. Maih, F. Mihoubi, G. Perendia, M. Ratto, and S. Villemot (2012): "Dynare: Reference Manual, Version 4," Dynare Working Paper 1, CEPREMAP.

Adjemian, S., and M. Juillard (2010): "Dealing with ZLB in DSGE models: An application to the Japanese economy," Discussion Paper 258, Economic and Social Research Institute, Tokyo.

Akaike, H. (1974): "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, 19(6), 716 – 723.

Amisano, G., and O. Tristani (2010): "Euro area inflation persistence in an estimated nonlinear DSGE model," *Journal of Economic Dynamics and Control*, 34(10), 1837–1858.

——— (2011): "Exact Likelihood Computation for Nonlinear DSGE Models with Heteroskedastic Innovations," Working Paper 1341, ECB.

Anderson, D. G. (1965): "Iterative Procedures for Nonlinear Integral Equations," *J. ACM*, 12(4), 547–560.

Braun, R. A., L. M. Körber, and Y. Waki (2012): "Some Unpleasant Properties of Log-Linearized Solutions When the Nominal Rate Is Zero," Working Paper 2012-5a, Federal Reserve Bank of Atlanta.

Calvo, G. A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, 12(3), 383–398.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107(2), 205–251, ArticleType: research-article / Full publication date: April 1999 / Copyright © 1999 The University of Chicago Press.

Del Moral, P. (2004): *Feynman-Kac formulae: genealogical and interacting particle systems with applications*. Springer, New York.

Doh, T. (2011): "Yield curve in an estimated nonlinear macro model," *Journal of Economic Dynamics and Control*, 35(8), 1229–1244.

Dong, B. (2012): "Mystery at the zero lower bound," Mimeo, University of Virginia.

Fang, H.-r., and Y. Saad (2009): "Two classes of multisecant methods for nonlinear acceleration," *Numerical Linear Algebra with Applications*, 16(3), 197–221.

Fernández-Villaverde, J., G. Gordon, P. A. Guerrón-Quintana, and J. Rubio-Ramírez (2012): "Nonlinear Adventures at the Zero Lower Bound," *National Bureau of Economic Research Working Paper Series*, No. 18058.

Fernández-Villaverde, J., P. A. Guerrón-Quintana, and J. Rubio-Ramírez (2012): "Estimating Dynamic Equilibrium Models with Stochastic Volatility," Working Paper 18399, National Bureau of Economic Research.

Fernández-Villaverde, J., and J. F. Rubio-Ramírez (2007): "How Structural Are Structural Parameters?," *National Bureau of Economic Research Working Paper Series*, No. 13166, published as Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez. "How Structural Are Structural

Parameters?," in Daron Acemoglu, Kenneth Rogoff and Michael Woodford, editors, "NBER Macroeconomics Annual 2007, Volume 22" University of Chicago Press (2008).

Galí, J. (2008): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press.

Gelfand, A. E., and D. K. Dey (1994): "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society. Series B (Methodological)*, 56(3), 501–514, ArticleType: research-article / Full publication date: 1994 / Copyright © 1994 Royal Statistical Society.

Geweke, J. (1999): "Using simulation methods for Bayesian econometric models: inference, development,and communication," *Econometric Reviews*, 18(1), 1–73.

Gomme, P., and P. Klein (2011): "Second-order approximation of dynamic models without the use of tensors," *Journal of Economic Dynamics and Control*, 35(4), 604–615.

Haario, H., E. Saksman, and J. Tamminen (2001): "An Adaptive Metropolis Algorithm," *Bernoulli*, 7(2), 223–242, ArticleType: research-article / Full publication date: Apr., 2001 / Copyright © 2001 International Statistical Institute (ISI) and Bernoulli Society for Mathematical Statistics and Probability.

Hall, J., M. K. Pitt, and R. Kohn (2012): "Bayesian inference for nonlinear structural time series models," *arXiv:1209.0253*.

Harvey, A. C. (1991): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Heer, B., and A. Maussner (2009): *Dynamic General Equilibrium Modeling: Computational Methods and Applications*. Springer.

Judd, K. L. (1998): *Numerical Methods in Economics*. MIT Press.

Kim, J., S. Kim, E. Schaumburg, and C. A. Sims (2008): "Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models," *Journal of Economic Dynamics and Control*, 32(11), 3397–3414.

Klein, P. (2000): "Using the generalized Schur form to solve a multivariate linear rational expectations model," *Journal of Economic Dynamics and Control*, 24(10), 1405–1423.

Pitt, M., R. Silva, P. Giordani, and R. Kohn (2010): "Auxiliary Particle filtering within adaptive Metropolis-Hastings Sampling," *arXiv:1006.1914*.

Pitt, M. K., and N. Shephard (1999): "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94(446), 590–599, ArticleType: research-article / Full publication date: Jun., 1999 / Copyright © 1999 American Statistical Association.

Pitt, M. K., R. Silva, P. Giordani, and R. Kohn (2012): "On some properties of Markov chain Monte Carlo simulation methods based on the particle filter," Mimeo, UNSW.

Schmitt-Grohé, S., and M. 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.

Schwarz, G. (1978): "Estimating the Dimension of a Model," *The Annals of Statistics*, 6(2), 461–464, Mathematical Reviews number (MathSciNet): MR468014; Zentralblatt MATH identifier: 0379.62005.

Sims, C. A. (2001): "Solving linear rational expectations models," *Computational Economics*, 20(1), 1—20.

Smets, F., and R. Wouters (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1(5), 1123–1175, ArticleType: research-article / Full publication date: Sep., 2003 / Copyright © 2003 The MIT Press.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde (2002): "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583–639.

Tallis, G. M. (1965): "Plane Truncation in Normal Populations," *Journal of the Royal Statistical Society. Series B (Methodological)*, 27(2), 301–307, ArticleType: research-article / Full publication date: 1965 / Copyright © 1965 Royal Statistical Society.

Walker, H. F., and P. Ni (2011): "Anderson Acceleration for Fixed-Point Iterations," *SIAM J. Numer. Anal.*, 49(4), 1715–1735.

Woodford, M. (2003): *Interest and prices: Foundations of a theory of monetary policy.* Princeton University Press, Princeton, NJ.

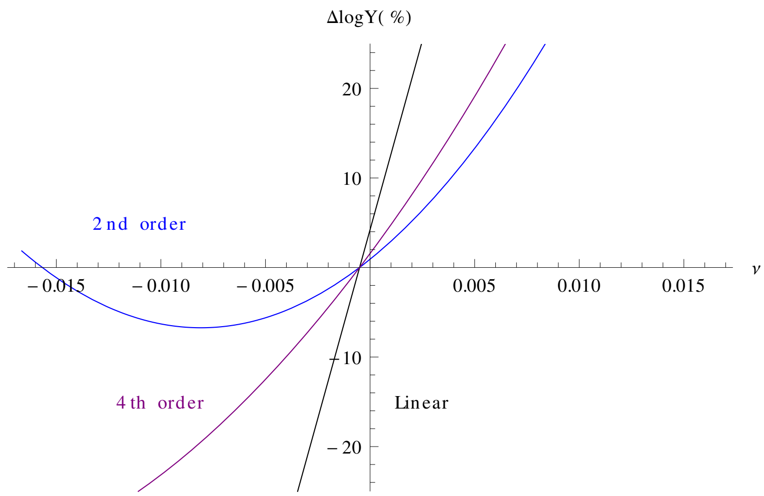

Figure 1: $\log y_t$ as a function of ν_t for a given x_{t-1} and θ: linear, quadratic, and quartic approximations. See Appendix A for details.

Method	$\mathbb{E}\eta^2$	$\|\eta\|_\infty$
Loglinear	4.8×10^{-8}	2.3×10^{-3}
Second-order	9.2×10^{-10}	1.4×10^{-4}
Locally linear	1.1×10^{-13}	6.7×10^{-7}

Table 1: Average Euler errors for different approximation methods

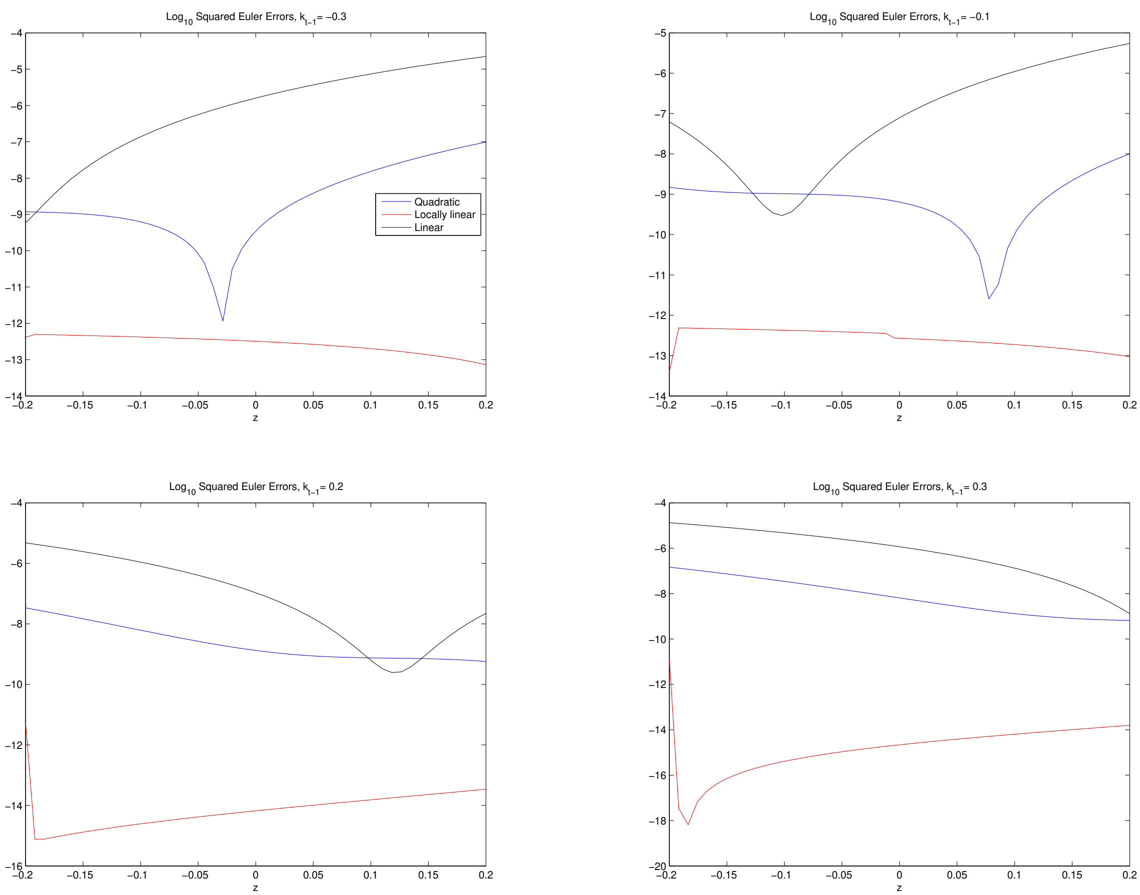

Figure 2: Squared Euler errors for the growth model (log scale), shown as a function of z_t, calculated for various values of k_{t-1}, using linear (black), quadratic (blue) and locally linear (red) approximation methods.

Parameter	Interpretation	Distribution	Mean	Std. dev
Demand-side parameters				
β	Discount factor	Beta	0.99	0.01
γ	Consumption elasticity	Gamma	1	0.2
ρ_b	Demand shock persistence	Beta	0.8	0.05
$\mathbb{E}\sigma_b$	Demand shock marginal volatility	Gamma	0.01	0.005
β_b	Demand shock GARCH parameter 1	Beta	0.2	0.2
γ_b	Demand shock GARCH parameter 2	Beta	0.2	0.2
Supply-side parameters				
ϕ	Labour supply elasticity	Gamma	1	0.5
α	Labour share of output	Beta	0.7	0.1
θ_p	Price stickiness	Beta	0.7	0.05
θ	Degree of imperfect competition	Gamma	7	2
ρ_a	TFP persistence	Beta	0.8	0.05
$\mathbb{E}\sigma_a$	TFP shock marginal volatility	Gamma	0.01	0.005
β_a	TFP shock GARCH parameter 1	Beta	0.2	0.2
γ_a	TFP shock GARCH parameter 2	Beta	0.2	0.2
ρ_e	Supply shock persistence	Beta	0.6	0.2
$\mathbb{E}\sigma_e$	Supply shock marginal volatility	Gamma	0.01	0.005
β_e	Supply shock GARCH parameter 1	Beta	0.2	0.2
γ_e	Supply shock GARCH parameter 2	Beta	0.2	0.2
Policy parameters				
ρ_r	Interest rate smoothing	Beta	0.5	0.2
ϕ_π	Central bank inflation response	Gamma	1.2	0.2
ϕ_y	Central bank output response	Gamma	0.5	0.2
Π	Steady-state gross inflation	Normal	1.005	0.0005
ρ_m	Policy shock persistence	Beta	0.6	0.2
$\mathbb{E}\sigma_m$	Policy shock marginal volatility	Gamma	0.01	0.005
β_m	Policy shock GARCH parameter 1	Beta	0.2	0.2
γ_m	Policy shock GARCH parameter 2	Beta	0.2	0.2
$\mathbb{E}\sigma_r$	Interest-rate shock volatility	Gamma	0.01	0.005
β_r	Policy shock GARCH parameter 1	Beta	0.2	0.2
γ_r	Policy shock GARCH parameter 2	Beta	0.2	0.2

Note: In the nonlinear approximation, each pair of GARCH parameters β_i and γ_i was restricted to the stable region, $0 \leq \beta_i + \gamma_i < 1$.

Table 2: Prior distributions for New Keynesian model

Method	Total time	Time per 100 draws
Linear	25 min	1.5 s
Nonlinear	9.5 hrs	34 s

Table 3: Estimation time for New Keynesian model

Method	Marginal logposterior	AIC	BIC	DIC
Loglinear	1564.8	3240.7	3168.2	-3246.7
Locally linear	1580.3	3299.1	3188.4	-3359.1

Table 4: Estimation results for New Keynesian model

Parameter	Linear		Nonlinear	
	Mean	90% CI	Mean	90% CI
Demand-side parameters				
β	0.992	[0.99,0.99]	0.996	[0.99,1]
γ	1.83	[1.5,2.2]	5.3	[4,6.8]
ρ_b	0.854	[0.83,0.88]	0.96	[0.91,1]
$\mathbb{E}\sigma_b$	0.00273	[0.0022,0.0032]	0.00355	[0.0022,0.0049]
β_b	—	—	0.161	[0.028,0.27]
γ_b	—	—	0.437	[0.27,0.58]
Supply-side parameters				
ϕ	0.669	[0.23,1.2]	0.147	[0.013,0.41]
α	0.703	[0.54,0.84]	0.796	[0.63,0.95]
θ_p	0.786	[0.71,0.86]	0.874	[0.83,0.91]
θ	7.09	[5,10]	1.9	[1.1,4.1]
ρ_a	0.794	[0.75,0.83]	0.511	[0.42,0.6]
$\mathbb{E}\sigma_a$	0.00196	[0.00079,0.0037]	0.00262	[0.00034,0.0058]
β_a	—	—	0.0321	[0.00029,0.11]
γ_a	—	—	0.0726	[0.00078,0.25]
ρ_e	0.988	[0.97,1]	0.995	[0.99,1]
$\mathbb{E}\sigma_e$	0.0065	[0.0044,0.0088]	0.0151	[0.011,0.018]
β_e	—	—	0.32	[0.19,0.5]
γ_e	—	—	0.0551	[0.00028,0.13]
Policy parameters				
ρ_r	0.902	[0.81,0.97]	0.991	[0.97,1]
ϕ_π	1.59	[1.3,1.9]	3.03	[2.6,3.7]
ϕ_y	0.279	[0.15,0.45]	0.204	[0.047,0.37]
Π	1.005	[1.004,1.006]	1.0098	[1.006,1.013]
ρ_m	0.923	[0.88,0.96]	0.978	[0.95,1]
$\mathbb{E}\sigma_m$	0.00244	[0.0018,0.0031]	0.0061	[0.0049,0.0075]
β_m	—	—	0.144	[0.018,0.27]
γ_m	—	—	0.225	[0.01,0.45]
$\mathbb{E}\sigma_r$	0.000673	[0.00039,0.00095]	0.00204	[0.0016,0.0025]
β_r	—	—	0.648	[0.41,0.85]
γ_r	—	—	0.179	[0.0038,0.44]

Table 5: Posterior estimates for New Keynesian model