Evaluation of Bayesian Nonlinear DSGE Models*

Kuo-Hsuan Chin†

Abstract

We apply a Bayesian approach to estimate a small-scale New Keynesian Dynamic Stochastic General Equilibrium (DSGE) model on the basis of first-order and second-order approximation techniques. In particular, the likelihood function for the linear and nonlinear model approximation are constructed via the Kalman and particle filters respectively. We evaluate the performance of the linear/nonlinear model approximation in terms of the log marginal data density and out-of-sample forecasting exercise. With the use of U.S. macroeconomic data, spanning from 1960 to 2017 and crossing the so-called "Great Inflation", "Great Moderation" and "Great Recession" periods in order, we find the nonlinear model approximation provides the best fit to the full sample of data. Moreover, by considering the transitions of Great Inflation to Great Moderation (GI-GM) and Great Moderation to Great Recession (GM-GR) periods as out-of-sample evaluation periods, we find, in terms of the forecast unbiasedness test, that the nonlinear DSGE model generates unbiased forecasts of inflation and interest rates but biased forecasts of real GDP growth rate in most cases. Lastly, in the forecast competition, the nonlinear DSGE model is dominant in forecasting all variables of interest in both transition periods except for the inflation forecasts evaluated in the GI-GM transition period.

JEL classification: C15, C53, E37.

Keywords: Bayesian Approach, Nonlinear New Keynesian DSGE, Particle Filter.

*I gratefully acknowledges the financial support from the Ministry of Science and Technology (MOST), R.O.C. (Project Number: MOST107-2410-H-035-004).

†Associate Professor, Department of Economics, Feng Chia University, No.100, Wenhwa Rd., Seatwen Dist., Taichung City, Taiwan 40724, R.O.C. (khchin@fcu.edu.tw)

1 Introduction

Over the past 60 years, the U.S. economy experienced in sequence the so-called "Great Inflation" (GI), "Great Moderation" (GM) and "Great Recession" (GR) periods, named according to the extent of the volatility of macroeconomic variables, the economic activity in particular. Compares with the other two periods, the volatility of the output growth has moderated dramatically in the GM period.[1] Most macroeconomic researchers have a broad consensus on the structural change in the transition from the GI to GM periods, and it spurred the consideration of the nonlinear nature of the economy.[23]

In the past two decades, dynamic stochastic general equilibrium (DSGE) models, which are constructed with a micro-foundation and rational expectation, have been commonly used as instruments for macroeconomic policy analysis and forecasting over long periods of data, crossing the above-mentioned regimes. Those models, for the most part, neglect nonlinearities by simply linearizing a set of nonlinear equilibrium conditions around the steady states and then obtaining the model solution afterward via linear solution techniques such as those proposed in Blandchard and Kahn (1980), Klein (2000), and Sims (2002). Aruoba, Fernández-Villaverde and Rubio-Ramírez (2006) state that the approximated errors generated with linear solution techniques (hereafter, the linear DSGE model) are much larger than those generated with nonlinear solution techniques (hereafter, the nonlinear DSGE model) in a typical DSGE model.[4] Those errors are brought into the construction of the likelihood function of the model and then result in inaccurate likelihood-based estimates and worse fit of the model to the data.

On these grounds, we study the time-varying fitness of the nonlinear DSGE model to the U.S. macroeconomic data over the past 60 years. Technically, we first solve a small-

[1] Kim and Nelson (1999) show a substantial decline in the volatility of U.S. real GDP growth rate uisng a Bayesian Markov-switching model, and the estimated break date that they found is the first quarter of the year 1984.

[2] Some studies find the causes of the transition in these regimes are possibly the variation in the size of the shocks. Under the framework of a multivariate Markov-switching model with time-varying parameters, Sims and Zha (2006) find the exogenous shocks which occurred in the GI period are much more volatile than those in the GM period. Another class of explanations about the causes of the transition in these regimes emphasizes the conduct of the monetary policy. Clarida, Gali and Gertler (2000) and Cogley and Sargent (2005) have argued that U.S. monetary policy is conducted in a more aggressive way in the GM period. However, Stock and Watson (2003) state that the improved moneatry policy cannot explain a structural decline in the volatility of the output growth, and their statements are supported by a variety of modern macroeconometric models, ranging from the reduced-form vector autoregression (VAR) to the structural DSGE models.

[3] There exists a debate on whether the "Great Moderation" (GM) period is ending or not. Using the Markov regime-switching models, Canarella, Fang, Miller and Pollard (2010) argue in favor of a transition to a new era of output growth volatility in the U.S. and U.K. However, Charles, Darné and Ferrara (2018) find no evidence for the end of the GM period in a GARCH-type model with U.S. and international data.

[4] By comparing different solution techniques for the DSGE models, Aruoba, Fernández-Villaverde and Rubio-Ramírez (2006) state that using linear solution techniques, log-linearization in particular, produces the worst fit of the model to the data.

scale New Keynesian model with both linear and nonlinear solution techniques, and the likelihood functions of the model are then constructed via the Kalman and particle filters respectively. Next, we apply a Bayesian approach to combine the likelihood function of the model with the prior density of model parameter to form the posterior density, leading to the approximation of the marginal data density, used to measure the fitness of the model. We repeat the above-mentioned process recursively to study and compare the time-varying fitness of the linear and nonlinear DSGE models. Moreover, we conduct the out-of-sample forecasting exercises in both models over the transition periods of Great Inflation to Great Moderation (GI-GM) and Great Moderation to Great Recession (GM-GR), and evaluate their absolute and relative performance of forecasting under a sequence of forecast unbiasedness and Diebold-Mariano (DM) tests.

Regarding the related studies, An (2005) states that Bayesian estimation of second-order approximations to a New Keynesian DSGE model improves the model fit and the identification of the structural parameters. Fernández-Villaverde and Rubio-Ramírez (2005) apply both linear and particle filters to a stochastic growth model, and they find the model estimated by the means of the particle filter yields a better fit, evaluated on the basis of marginal likelihoods, to both simulated and real data.[5] The present study differs from the previous studies in that we consider the transition regimes of GI-GM and GM-GR in linear and nonlinear DSGE models when evaluating fitness to the data and the forecasting performance in terms of a variety of statistical tests.

Several results are found in this paper. First, we find, in terms of the marginal data density, that the nonlinear DSGE model has better fit to the whole sample. In particular, the time-varying model fit of the nonlinear DSGE model improves over the transition GI-GM period. In contrast, the model fit of the linear DSGE model gets worse over time in the transition GI-GM period. The results imply that the use of the nonlinear solution technique helps to capture a structural decline in the volatility of the macroeconomic variables over the GI-GM period. Second, compared to the linear DSGE model in most cases, the nonlinear DSGE model helps to generate unbiased forecasts of inflation and interest rates over the transition periods of GI-GM and GM-GR. Third, the nonlinear DSGE model overwhelmingly outperforms the linear DSGE model on forecasting real output growth, inflation and interest rates over two transition periods, particulaly the transition GM-GR period.

The reminder of the paper is organized as follows. Section 2 introduces a small-scale New Keynesian DSGE model. The estimation methodology is introduced in Section 3, and we particularly focus on the particle filter, used for approximating the likelihood function of the nonlinear DSGE model, and the Bayesian approach. In Section 4, we

[5] Fernández-Villaverde and Rubio-Ramírez (2005) state that although there is a small difference between the point estimates obtained respectively by linear and particle filters, it leads to important effects on the statistical moments of the model.

describe the source of data and present the empirical results. Section 5 concludes.

2 New Keynesian DSGE Model

The New Keynesian DSGE model used in this paper is based on Ireland (2004). Specifically, the closed economy simply consists of homogeneous households, homogeneous finished-goods producers, a continuum of differentiated intermediate-goods producers indexed by $i \in [0,1]$, and a central bank. The households are demanders in the goods market and provide labor service to the intermediate-goods producers. During each period $t = 0, 1, 2, ...$, the representative household buys the amount of consumption bundle (C_t) from the final-goods producers and provides the labor supply (h_t) to the intermediate-goods producers to maximize the following expected utility function[6]

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[a_t \ln (C_t) - \left(\frac{1}{\eta} \right) h_t^{\eta} \right] \tag{2.1}$$

subject to the budget constraint of

$$\frac{B_{t-1} + T_t + W_t h_t + D_t}{P_t} \geq C_t + \frac{\frac{B_t}{r_t}}{P_t} \tag{2.2}$$

where β $(0 < \beta < 1)$ is the discount factor, η $(\eta \geq 1)$ measures the elasticity of labor supply, and a_t is the preference shock, following the stationary first-order autoregressive process

$$\ln (a_t) = (1 - \rho_a) \ln (a) + \rho_a \ln (a_{t-1}) + \varepsilon_{at} \tag{2.3}$$

where $0 \leq \rho_a < 1$, $a\,(a > 0)$ is the steady state of a_t, and ε_{at} is the serially uncorrelated innovation and normally distributed with zero mean and standard deviation σ_a. In addition, B_t, T_t, W_t, D_t, P_t and r_t denote the nominal bonds, lump-sum taxes, wage, dividend, price and interest rate respectively.

The representative finished-goods producer aggregates the intermediate goods, $Y_t(i)$, $i \in [0,1]$, into the final good, Y_t, according to the constant-returns-to-scale (CRTS) technology characterized by

$$\left[\int_0^1 Y_t (i)^{\frac{(\theta_t - 1)}{\theta_t}} di \right]^{\frac{\theta_t}{\theta_t - 1}} \geq Y_t \tag{2.4}$$

where θ_t is the price markup shock, following the stationary first-order autoregressive process

$$\ln (\theta_t) = (1 - \rho_\theta) \ln (\theta) + \rho_\theta \ln (\theta_{t-1}) + \varepsilon_{\theta t} \tag{2.5}$$

[6]Utility is additively separable in consumption and hours worked. Given this additive separability, the logarithmic specification for preferences over consumption is necessary, as shown by King, Plosser, and Rebelo (1988), for the model to be consistent with balanced growth.

where $0 \leq \rho_\theta < 1$, $\theta \, (\theta > 0)$ is the steady state of θ_t, and $\varepsilon_{\theta t}$ is the serially uncorrelated innovation and normally distributed with a mean of zero and standard deviation σ_θ. Each final-goods producer sells the aggregate goods in a perfectly competitive market to the consumers with the price, P_t, and its profit maximization problem is characterized by

$$P_t \left[\int_0^1 Y_t \left(i \right)^{\frac{(\theta_t - 1)}{\theta_t}} di \right]^{\frac{\theta_t}{\theta_t - 1}} - \int_0^1 P_t \left(i \right) Y_t \left(i \right) di \tag{2.6}$$

where $P_t \left(i \right)$, the price charged by the intermediate-goods producer i, is related to the aggregate price index via the zero-profit condition shown as follows

$$P_t = \left[\int_0^1 P_t \left(i \right)^{1-\theta_t} di \right]^{\frac{1}{1-\theta_t}}, \quad t = 0, 1, 2, ... \tag{2.7}$$

Each intermediate-goods producer, indexed by i, uses the labor (h_t) hired from the households to produce a differentiate good, $Y_t \left(i \right)$, according to the CRTS technology described by

$$h_t \left(i \right) \geq Y_t \left(i \right) \tag{2.8}$$

where the labor is the only input used in production. The intermediate-goods producers sell the products to the final-goods producers in the monopolistically competitive market, and the prices, $P_t \left(i \right)$, they set are subject to the price rigidity. In this paper, we use the price-rigidity framework, proposed by Rotemberg (1982), in which the intermediate-goods producers face a quadratic cost of adjusting the nominal price between periods, measured in terms of the final goods, shown as

$$\frac{\phi}{2} \left[\frac{P_t \left(i \right)}{\pi P_{t-1} \left(i \right)} - 1 \right]^2 Y_t \tag{2.9}$$

where $\phi \geq 0$ measures the degree of the price rigidity and $\pi > 1$ is the gross steady-state inflation rate. During each period $t = 0, 1, 2, ...$, each intermediate-goods producer chooses the optimal price to maximize its profit as

$$\frac{P_0}{\Lambda_0} E_0 \sum_{t=0}^{\infty} \beta^t \Lambda_t \left\{ \left[\frac{P_t \left(i \right)}{P_t} \right]^{1-\theta_t} Y_t - \left[\frac{P_t \left(i \right)}{P_t} \right]^{-\theta_t} \left(\frac{W_t}{P_t} \right) Y_t - \frac{\phi}{2} \left[\frac{P_t \left(i \right)}{\pi P_{t-1} \left(i \right)} - 1 \right]^2 Y_t \right\}, \tag{2.10}$$

where $\Lambda_t = \frac{a_t}{C_t}$ measures the marginal utility value to the representative household of each additional unit of real profits.

Finally, the central bank follows a simplified Taylor rule to conduct the monetary policy as

$$r_t = \pi_t^{\rho_\pi} Y_t^{\rho_Y} \exp \left(\varepsilon_{mt} \right) \tag{2.11}$$

where the linearized rule implies that the central bank raises or lowers the short-term nominal interest rate in response to the deviation of inflation rate and output from their steady-state levels, and the parameters ρ_π and ρ_Y denote the degree of the responses of the interest rate to the deviation of the inflation rate and output respectively. The monetary policy shock, ε_{mt}, is the serially uncorrelated innovation, and is normally distributed with a mean of zero and standard deviation of σ_m.

Based on the optimization problems of the economic agents, a set of nonlinear equilibrium conditions are derived, and the steady-state conditions are also obtained as a by-product. In addition to log-linearizing those equations around their steady states, we apply a nonlinear solution technique, a second-order perturbation approach in particular, to solve the model. The nonlinear solution of the model will be written in a nonlinear and non-normal state-space representation for approximating the likelihood function of the model via a particle filter.

3 Particle Markov Chain Monte Carlo Approach

In order to apply the Bayesian approach to the nonlinear DSGE model, several steps are needed for the posterior inference. First, the nonlinear solution of the model is augmented with a measurement equation, which relates the observed variables to the unobserved state variables in a nonlinear state-space form representation. Second, a particle filter is applied to approximate the likelihood function of the model. Third, by specifying the prior distribution of the underlying structural parameters (the priors), the Bayesian estimates can be obtained by combining the prior and the likelihood function via Markov Chain Monte Carlo (MCMC) simulation techniques.[7] Below we explain each step in detail.

In step 1, a nonlinear state-space form representation is written as follows,

$$s_t = f\left(s_{t-1}, \varepsilon_t; \Theta\right) \tag{3.1}$$

$$y_t = g\left(s_t; \Theta\right) + e_t \tag{3.2}$$

where s_t denotes the vector of finite state (predetermined) variables, assumed to be a first-order Markov process, in a DSGE model, y_t is the vector of observed variables, including probably predetermined and non-predetermined variables, both f and g, not explicitly defined, denote the set of state and measurement equations, Θ is the set of unknown parameters, and ε_t and e_t respectively denote the set of structural shocks and

[7]Adjemian and Karame (2016) propose tools for solving and estimating nonlinear DSGE models via the particle filter in Dynare. Please refer to their technical paper.

additive measurement errors.[8][9] In order to obtain the likelihood function of the nonlinear DSGE model, one needs a recursive algorithm that updates the state variables and determines the observed variables from the state variables.[10] In step 2, we use the sequential Monte Carlo (particle filter) approach, particularly the sequential importance sampling tehcnique, to reach the goal.[11]

Technically, the sample likelihood function of the model is the joint density of the data sample $y_{1:T} = \{y_t\}_{t=1}^{T}$ conditional on the parameters Θ, which can be written as

$$p\left(y_{1:T} \mid \Theta\right) = p\left(y_1 \mid s_0; \Theta\right) p\left(s_0 \mid \Theta\right) \overbrace{\Pi_{t=2}^{T} p\left(y_t \mid y_{1:t-1}; \Theta\right)}^{(*)} \tag{3.3}$$

where the product of the conditional densities, denoted as (*), can be written with the use of the law of conditional probability and Bayes' rule as

$$p\left(y_t \mid y_{1:t-1}; \Theta\right) = \int \frac{p\left(y_{1:t} \mid s_t; \Theta\right) p\left(s_t \mid \Theta\right) q\left(s_t \mid y_{1:t}; \Theta\right) p\left(s_{t-1} \mid y_{1:t-1}; \Theta\right)}{q\left(s_t \mid y_{1:t}; \Theta\right) p\left(y_{1:t-1} \mid s_{t-1}; \Theta\right) p\left(s_{t-1} \mid \Theta\right)} ds_t \tag{3.4}$$

where $q\left(s_t \mid y_{1:t}; \Theta\right)$ is the proposal (importance) density, an easily sampled one, and both normalized and unnormalized weights, $\tilde{w}_t\left(s_t\right)$ and $\hat{w}_t\left(s_t\right)$, are defined as $\frac{p(s_t \mid y_{1:t}; \Theta)}{q(s_t \mid y_{1:t}; \Theta)}$ and $\frac{p(y_{1:t} \mid s_t; \Theta) p(s_t; \Theta)}{q(s_t \mid y_{1:t}; \Theta)}$ respectively. According to Bayes' rule, $\tilde{w}_t\left(s_t\right) \propto \hat{w}_t\left(s_t\right)$. The sequential importance sampling technique can be implemented by choosing a proposal density with the following property

$$q\left(s_t \mid y_{1:t}; \Theta\right) = q\left(s_t \mid s_{t-1}, y_t; \Theta\right) q\left(s_{t-1} \mid y_{1:t-1}; \Theta\right) \tag{3.5}$$

where the proposal density proceeds in a recursive way. By incorporating Equation (3.5) with the definition of the unnormalized weights, we obtain

$$\hat{w}_t\left(s_t\right) = \hat{w}_{t-1}\left(s_{t-1}\right) \frac{p\left(y_t \mid s_t; \Theta\right) p\left(s_t \mid s_{t-1}; \Theta\right)}{q\left(s_t \mid s_{t-1}, y_t; \Theta\right)} \tag{3.6}$$

$$\hat{w}_t\left(s_t\right) \propto \tilde{w}_{t-1}\left(s_{t-1}\right) \frac{p\left(y_t \mid s_t; \Theta\right) p\left(s_t \mid s_{t-1}; \Theta\right)}{q\left(s_t \mid s_{t-1}, y_t; \Theta\right)} \tag{3.7}$$

where the current particle weight depends on its past weight, the so-called incremental weight. Equations (3.5) and (3.7) play key roles in obtaining the likelihood function in a sequential MCMC approach. In particular, by taking both equations into Equation (3.4),

[8] We assume that the observations are conditionally independent for simplicity, $p\left(y_t \mid y_{1:t-1}, s_{0:t}\right) = p\left(y_t \mid s_t\right)$, and this results in the expression of Equation (3.2).

[9] A measurement error is required to be associated with each observed variable. In other words, it is required to have as many measurement errors as observed variables to estimate a model with a nonlinear filter.

[10] Given that both structural shock and measurement error follow the Gaussian densities in a linearized DSGE model, the traditional Kalman filter can be applied to obtain the likelihood function analytically.

[11] Creal (2012) provides an excellent survey of sequential Monte Carlo techniques.

we obtain

$$p\left(y_t \mid y_{1:t-1};\Theta\right)=\int \tilde{w}_{t-1}\left(s_{t-1}\right)\frac{p\left(y_t \mid s_t;\Theta\right)p\left(s_t \mid s_{t-1};\Theta\right)}{q\left(s_t \mid s_{t-1},y_t;\Theta\right)}q\left(s_t \mid y_{1:t};\Theta\right)ds_t \tag{3.8}$$

By taking draws (particles) from the easy-to-sample proposal density, $q\left(s_t \mid y_{1:t};\Theta\right)$, one can approximate the conditional likelihood function. In sum, we summarize the previous discussion about the particle filter as follows: given $\{s_0, w_0\}$ is known, for $t = 1, 2, ..., T$ and $i = 1, 2, ..., N$,

1. Draw $\left\{\widetilde{s}_t^{(i)}\right\}_{i=1:N}$ from $q\left(s_t \mid s_{t-1}^{(i)}, y_t;\Theta\right)$.
2. Compute the importance weights: $\hat{w}_t^{(i)} \propto \tilde{w}_{t-1}^{(i)}\frac{p\left(y_t|\widetilde{s}_t^{(i)};\Theta\right)p\left(\widetilde{s}_t^{(i)}|s_{t-1}^{(i)};\Theta\right)}{q\left(\widetilde{s}_t^{(i)}|s_{t-1}^{(i)},y_t;\Theta\right)}$.
3. Obtain the normalized weights: $\tilde{w}_t^{(i)} = \frac{\hat{w}_t^{(i)}}{\sum_{j=1}^N \hat{w}_t^{(j)}}$.
4. Resampling: if $\frac{1}{\sum_{i=1}^N\left[\tilde{w}_t^{(i)}\right]^2}$ is less than the chosen value, resample the particles and replace $\left\{\widetilde{s}_t^{(i)}, \widetilde{w}_t^{(i)}\right\}$ with $\left\{s_t^{(i)}, w_t^{(i)} = \frac{1}{N}\right\}$. That is, we discard the particles with low weights and replace particles with high weights on interesting regions of the density using a fixed amount of particles.

We consider the resampling step to avoid the degeneracy problem, in which the weights degenerate as t increases, resulting in all but one of the particles having extremely small weights. According to Equations (3.3) and (3.8), one can approximate both full and conditional likelihood functions via numerical integration as

$$p\left(y_t \mid y_{1:t-1};\Theta\right)\approx\sum_{i=1}^N \tilde{w}_{t-1}^{(i)}\frac{p\left(y_t \mid \widetilde{s}_t^{(i)};\Theta\right)p\left(\widetilde{s}_t^{(i)} \mid s_{t-1}^{(i)};\Theta\right)}{q\left(\widetilde{s}_t^{(i)} \mid s_{t-1}^{(i)},y_t;\Theta\right)} \tag{3.9}$$

$$p\left(y_{1:T} \mid \Theta\right)=p\left(y_1 \mid s_0;\Theta\right)p\left(s_0 \mid \Theta\right)\Pi_{t=2}^T\left\{\sum_{i=1}^N \tilde{w}_{t-1}^{(i)}\frac{p\left(y_t \mid \widetilde{s}_t^{(i)};\Theta\right)p\left(\widetilde{s}_t^{(i)} \mid s_{t-1}^{(i)};\Theta\right)}{q\left(\widetilde{s}_t^{(i)} \mid s_{t-1}^{(i)},y_t;\Theta\right)}\right\} \tag{3.10}$$

where the likelihood function depends on the densities for the initial states (s_0). In practical implementation, we choose the ergodic distribution of the states by considering a first-order approximation around the steady-state of the DSGE model.

Regarding the choice of the proposal density, we impose the Gaussian assumption to choose the one incorporating the current information on observed variables. Specifically, the model's nonlinear equations are still preserved, and the approximation concerns the distribution of state variables, which will be assumed to be Gaussian. The common idea is to use the Kalman filter updating step to build a posterior distribution for state variables, by including the current information on observables y_t. The Kalman filter updating step will be used to form the proposal $q\left(s_t \mid s_{t-1}, y_t; \Theta\right)$ for state variables at time t in particle filters.[12] Specifically, the approach uses the posterior distribution provided by the nonlinear Kalman filter as a proposal. Technically, $\left\{\widetilde{s}_t^{(i)}\right\}_{i=1:N}$ are sampled from the posterior distribution $N\left(s_t; \overline{s}_{t|t}, P_{s_{t|t}}\right)$ obtained with the nonlinear Kalman filter previously presented. The current states distribution can be obtained with a Monte Carlo approach. The associated weights $\left\{\widetilde{w}_t^{(i)}\right\}_{i=1:N}$ are simply given by

$$\hat{w}_t^{(i)} \propto \tilde{w}_{t-1}^{(i)} \frac{p\left(y_t \mid \widetilde{s}_t^{(i)}; \Theta\right) p\left(\widetilde{s}_t^{(i)} \mid \widetilde{s}_{t-1}^{(i)}; \Theta\right)}{p\left(\widetilde{s}_t^{(i)} \mid \widetilde{s}_{t-1}^{(i)}; \Theta\right)} = \frac{1}{N} \frac{p\left(y_t \mid \widetilde{s}_t^{(i)}; \Theta\right) N\left(\widetilde{s}_t^{(i)}; \overline{s}_{t|t-1}, P_{s_{t|t-1}}\right)}{N\left(\widetilde{s}_t^{(i)}; \overline{s}_{t|t}, P_{s_{t|t}}\right)} \tag{3.11}$$

where the transition density of states $p\left(\widetilde{s}_t \mid s_{t-1}; \Theta\right)$ is approximated by the prior distribution $N\left(\widetilde{s}_t^{(i)}; \overline{s}_{t|t-1}, P_{s_{t|t-1}}\right)$, provided by the nonlinear Kalman filter, and the normalized weight is set to $\frac{1}{N}$. Since we have made a Gaussian assumption, we only have to track the first two moments of the states' distribution.

In step 3, we use a Bayesian approach to connect the prior belief about the parameters, Θ, with the data information coming from the likelihood function of the model. Technically, the posterior density of the parameters can be expressed as the product of the sample likelihood $p\left(y_{1:T}; \Theta\right)$ and the prior density of parameters $p\left(\Theta\right)$:

$$p\left(\Theta \mid y_{1:T}\right) \propto p\left(y_{1:T}; \Theta\right) p\left(\Theta\right) \tag{3.12}$$

where in the case of the linear and nonlinear models, the sample likelihood is approximated through the Kalman and particle filters respectively. In an MCMC framework like the random-walk Metropolis-Hastings (RWMH) algorithm which we apply for the linear model, the random-walk Particle Marginal Metropolis-Hastings (PMMH) algorithm is used in a nonlinear model to approximate the posterior density of parameters. Specifically, a candidate is drawn according to the following random-walk process:

$$\Theta_j^* = \Theta_{j-1} + \epsilon_j \tag{3.13}$$

where $\epsilon_j \sim N\left[0, \gamma_{RW} V\left(\Theta_0\right)\right]$ and γ_{RW} is set in order to obtain an acceptance ratio around

[12] Compared to the standard particle filter, this approach, the so-called marginal particle filter, drastically reduces the computational burden and is particularly easy to implement. Specifically, we only rely on the evolution of the mean and variance of the Gaussian density, approximated by weighted particles.

25%. The posterior distribution can be approximated according to the acceptation rule:

$$\Theta_j = \begin{cases} \Theta_j^* & if\ U_{[0,1]} \leq \min\left\{1, \frac{p\left(\Theta_j^*|y_{1:T}\right)}{p(\Theta_{j-1}|y_{1:T})}\right\} \\ \Theta_{j-1} & otherwise \end{cases} \tag{3.14}$$

4 Empirical Results

The linear and nonlinear DSGE models are estimated based on quarterly data, downloaded from the official website of the Federal Reserve Bank at St. Louis, for the U.S. economy, including inflation, real GDP, and the federal funds rates. The whole sample spans from 1948:Q1 to 2017:Q4, covering the so-called "Great Inflation", "Great Moderation" and "Great Recession" periods, respectively. In particular, I divide the part of the whole sample into two overlapping subsamples, the periods of 1960Q1-1986Q4 and 1983Q1-2010Q4, with roughly equal size for an out-of-sample forecasting exercise. The forecast evaluation period is chosen as 1979:Q1-1986:Q4 in subsample 1 and 2003:Q1-2010:Q4 in subsample 2, covering the transition regimes of GI-GM and GM-GR.

The time-varying fitness of the model to the data is presented in Table 1. In particular, the nonlinear DSGE model has better model fit in the whole sample period of 1960Q1-2017Q4, crossing three regimes of GI, GM and GR. Most important of all, the fitness of the nonlinear DSGE model improves over time in the transition period of GI-GM, but the fitness of the linear model gets worse over time in the same transition period. Accordingly, the results imply the importance of the nonlinear DSGE model for successfully capturing the nonlinear nature of the economy, resulting from the structural decline of the volatility of the macroeconomic variables. However, the time-varying fitness of the nonlinear DSGE model is always inferior to the linear model in the transition period of GM-GR.

Regarding the ability of absolute forecast, I rely on the forecast unbiasedness test, conducting the hypothesis testing on the coefficient of α on the basis of the following simple regression,

$$\hat{\varepsilon}^f_{t+h} = \alpha + \beta y^f_{t+h} + \eta_{t+h} \tag{4.1}$$

where $\hat{\varepsilon}^f_{t+h}$ is the h-step-ahead estimated forecast error, y^f_{t+h} represents the h-step-ahead forecast and the null hypothesis is formed as $\alpha = 0$, indicating the forecast unbiasedness. A sequence of forecast unbiasedness tests is conducted in a rolling way to capture models' absolute forecasting ability, and the results are reported in Tables 2.A., 2.B. and 2.C. For the values shown on the vertical axis of the graphs, I report the p-value of the t-statistic. The graphs show that the nonlinear DSGE model generates unbiased forecasts of inflation and interest rates but biased forecasts of the real GDP growth rate in most cases.

Table 1. Log Marginal Data Density

GI-GM Transition Period			GM-GR Transition Period		
Sample Period	NL-DSGE	L-DSGE	Sample Period	NL-DSGE	L-DSGE
1960Q1-1978Q4	-677.93	-244.29	1983Q1-2002Q4	-889.47	-561.98
1960Q1-1979Q1	-563.56	-242.97	1983Q1-2003Q1	-843.46	-328.30
1960Q1-1979Q2	-538.53	-246.11	1983Q1-2003Q2	-796.63	-326.72
1960Q1-1979Q3	-535.91	-256.47	1983Q1-2003Q3	-787.76	-329.11
1960Q1-1979Q4	-528.15	-271.03	1983Q1-2003Q4	-756.57	-334.70
1960Q1-1980Q1	-484.98	-279.76	1983Q1-2004Q1	-731.47	-337.41
1960Q1-1980Q2	-483.63	-280.76	1983Q1-2004Q2	-731.68	-338.96
1960Q1-1980Q3	-484.12	-295.98	1983Q1-2004Q3	-731.61	-339.53
1960Q1-1980Q4	-483.61	-306.85	1983Q1-2004Q4	-731.45	-340.25
1960Q1-1981Q1	-483.92	-338.49	1983Q1-2005Q1	-732.01	-337.62
1960Q1-1981Q2	-484.13	-365.98	1983Q1-2005Q2	-731.58	-339.88
1960Q1-1981Q3	-483.91	-379.08	1983Q1-2005Q3	-731.76	-339.31
1960Q1-1981Q4	-483.29	-399.99	1983Q1-2005Q4	-731.67	-339.17
1960Q1-1982Q1	-483.91	-426.05	1983Q1-2006Q1	-731.61	-339.71
1960Q1-1982Q2	-484.17	-434.04	1983Q1-2006Q2	-731.34	-340.10
1960Q1-1982Q3	-483.91	-441.69	1983Q1-2006Q3	-731.56	-346.42
1960Q1-1982Q4	-483.87	-451.45	1983Q1-2006Q4	-731.63	-355.88
1960Q1-1983Q1	-483.72	-463.20	1983Q1-2007Q1	-731.39	-358.59
1960Q1-1983Q2	-483.52	-471.38	1983Q1-2007Q2	-731.67	-362.73
1960Q1-1983Q3	-483.65	-484.43	1983Q1-2007Q3	-731.60	-363.25
1960Q1-1983Q4	-483.57	-492.80	1983Q1-2007Q4	-731.38	-363.72
1960Q1-1984Q1	-483.50	-507.38	1983Q1-2008Q1	-731.67	-364.11
1960Q1-1984Q2	-483.86	-526.63	1983Q1-2008Q2	-731.80	-365.91
1960Q1-1984Q3	-484.14	-541.28	1983Q1-2008Q3	-731.80	-373.53
1960Q1-1984Q4	-484.65	-546.10	1983Q1-2008Q4	-731.40	-375.70
1960Q1-1985Q1	-483.89	-557.37	1983Q1-2009Q1	-731.31	-388.74
1960Q1-1985Q2	-484.43	-568.18	1983Q1-2009Q2	-731.48	-393.42
1960Q1-1985Q3	-483.86	-579.34	1983Q1-2009Q3	-731.69	-397.69
1960Q1-1985Q4	-483.64	-591.33	1983Q1-2009Q4	-731.65	-401.06
1960Q1-1986Q1	-484.36	-601.75	1983Q1-2010Q1	-731.66	-405.20
1960Q1-1986Q2	-483.54	-610.03	1983Q1-2010Q2	-731.74	-408.32
1960Q1-1986Q3	-483.64	-616.02	1983Q1-2010Q3	-731.70	-411.53
Full Sample: 1960Q1-2017Q4					
Nonlinear DSGE: -2689.22			Linear DSGE: -2883.74		

Footnote: "NL-DSGE" and "L-DSGE" are abbreviations of "Nonlinear DSGE" and "Linear DSGE".

Table 2.A. Sequence of Forecast Unbiasedness Test (Real GDP Growth Rate)

GI-GM Transition Period			GM-GR Transition Period		
Rolling Period	NL-DSGE	L-DSGE	Rolling Period	NL-DSGE	L-DSGE
1979Q1-1982Q4	0.0003	0.0000	2003Q1-2006Q4	0.0342	0.0568
1979Q2-1983Q1	0.0002	0.0000	2003Q2-2007Q1	0.0156	0.0851
1979Q3-1983Q2	0.0000	0.0000	2003Q3-2007Q2	0.0118	0.1353
1979Q4-1983Q3	0.0000	0.0000	2003Q4-2007Q3	0.0101	0.0790
1980Q1-1983Q4	0.0000	0.0000	2004Q1-2007Q4	0.0014	0.1841
1980Q2-1984Q1	0.0000	0.0000	2004Q2-2008Q1	0.0005	0.8782
1980Q3-1984Q2	0.0000	0.0002	2004Q3-2008Q2	0.0010	0.8863
1980Q4-1984Q3	0.0000	0.0000	2004Q4-2008Q3	0.0003	0.4078
1981Q1-1984Q4	0.0000	0.0000	2005Q1-2008Q4	0.0000	0.0772
1981Q2-1985Q1	0.0000	0.0000	2005Q2-2009Q1	0.0004	0.0359
1981Q3-1985Q2	0.0000	0.0000	2005Q3-2009Q2	0.0055	0.0932
1981Q4-1985Q3	0.0000	0.0000	2005Q4-2009Q3	0.0173	0.2077
1982Q1-1985Q4	0.0000	0.0000	2006Q1-2009Q4	0.0728	0.5896
1982Q2-1986Q1	0.0000	0.0004	2006Q2-2010Q1	0.1040	0.7705
1982Q3-1986Q2	0.0000	0.0000	2006Q3-2010Q2	0.2924	0.7161
1982Q4-1986Q3	0.0000	0.0000	2006Q4-2010Q3	0.5291	0.3379
1983Q1-1986Q4	0.0000	0.0001	2007Q1-2010Q4	0.7804	0.1647

Footnote: the values listed in Table are p-values; forecasts are one-step-ahead forecasts.

Table 2.B. Sequence of Forecast Unbiasedness Test (Inflation Rate)

GI-GM Transition Period			GM-GR Transition Period		
Rolling Period	NL-DSGE	L-DSGE	Rolling Period	NL-DSGE	L-DSGE
1979Q1-1982Q4	0.7194	0.0043	2003Q1-2006Q4	0.4528	0.0584
1979Q2-1983Q1	0.6110	0.0001	2003Q2-2007Q1	0.9942	0.0475
1979Q3-1983Q2	0.2035	0.0000	2003Q3-2007Q2	0.8983	0.1495
1979Q4-1983Q3	0.1393	0.0000	2003Q4-2007Q3	0.3259	0.0385
1980Q1-1983Q4	0.0809	0.0000	2004Q1-2007Q4	0.0906	0.0275
1980Q2-1984Q1	0.1016	0.0000	2004Q2-2008Q1	0.1683	0.0351
1980Q3-1984Q2	0.0795	0.0000	2004Q3-2008Q2	0.1620	0.0411
1980Q4-1984Q3	0.0373	0.0000	2004Q4-2008Q3	0.3084	0.1077
1981Q1-1984Q4	0.0180	0.0000	2005Q1-2008Q4	0.1367	0.0385
1981Q2-1985Q1	0.1402	0.0008	2005Q2-2009Q1	0.1251	0.0450
1981Q3-1985Q2	0.1238	0.0001	2005Q3-2009Q2	0.0174	0.0049
1981Q4-1985Q3	0.0998	0.0000	2005Q4-2009Q3	0.0090	0.0056
1982Q1-1985Q4	0.3153	0.0003	2006Q1-2009Q4	0.0105	0.0152
1982Q2-1986Q1	0.1680	0.0000	2006Q2-2010Q1	0.0105	0.0269
1982Q3-1986Q2	0.0074	0.0000	2006Q3-2010Q2	0.0141	0.0402
1982Q4-1986Q3	0.0257	0.0000	2006Q4-2010Q3	0.0152	0.0418
1983Q1-1986Q4	0.0936	0.0000	2007Q1-2010Q4	0.0040	0.0179

Footnote: the values listed in Table are p-values; forecasts are one-step-ahead forecasts.

Table 2.C. Sequence of Forecast Unbiasedness Test (Interest Rate)

GI-GM Transition Period			GM-GR Transition Period		
Rolling Period	NL-DSGE	L-DSGE	Rolling Period	NL-DSGE	L-DSGE
1979Q1-1982Q4	0.0164	0.1452	2003Q1-2006Q4	0.8968	0.8377
1979Q2-1983Q1	0.0586	0.0808	2003Q2-2007Q1	0.9321	0.3391
1979Q3-1983Q2	0.2035	0.0330	2003Q3-2007Q2	0.9706	0.2270
1979Q4-1983Q3	0.4457	0.0239	2003Q4-2007Q3	0.7243	0.0352
1980Q1-1983Q4	0.6995	0.0140	2004Q1-2007Q4	0.1641	0.0236
1980Q2-1984Q1	0.8737	0.0126	2004Q2-2008Q1	0.2747	0.0725
1980Q3-1984Q2	0.9014	0.0081	2004Q3-2008Q2	0.2722	0.1028
1980Q4-1984Q3	0.6228	0.0015	2004Q4-2008Q3	0.5208	0.0655
1981Q1-1984Q4	0.5865	0.0029	2005Q1-2008Q4	0.5111	0.0144
1981Q2-1985Q1	0.7605	0.0536	2005Q2-2009Q1	0.9610	0.0005
1981Q3-1985Q2	0.5252	0.0224	2005Q3-2009Q2	0.5793	0.0004
1981Q4-1985Q3	0.1473	0.0036	2005Q4-2009Q3	0.4975	0.0006
1982Q1-1985Q4	0.3824	0.0159	2006Q1-2009Q4	0.7167	0.0038
1982Q2-1986Q1	0.2682	0.0027	2006Q2-2010Q1	0.7712	0.0051
1982Q3-1986Q2	0.0330	0.0000	2006Q3-2010Q2	0.9537	0.0114
1982Q4-1986Q3	0.2726	0.0000	2006Q4-2010Q3	0.9434	0.0123
1983Q1-1986Q4	0.9841	0.0003	2007Q1-2010Q4	0.9242	0.0229

Footnote: the values listed in Table are p-values; forecasts are one-step-ahead forecasts.

However, most forecasts generated from the linear DSGE model are biased. Results of the relative forecasting performance over time on the basis of the fluctuation test proposed by Giacomini and Rossi (2010) are also reported. In practice, the Diebold and Mariano (1995) test statistic can be computed in a rolling way to keep track of the models' relative forecasting performance. I report the results of the fluctuation test in Table 3. In the forecast competition, the nonlinear DSGE model is dominant for forecasting all variables of interest in both transition periods except the inflation forecasts in the transition period of GI-GM.

In sum, the results shown above demonstrate that when one uses the DSGE framework to conduct empirical research for the U.S. economy over a long time span, particularly crossing both the "Great Inflation" and "Great Moderation" periods, it is appropriate to choose the nonlinear DSGE model as the modelling tool since it has better model fitness and helps to generate more unbiased forecasts of the key macroecoonmic variables.

5 Conclusion

In the past two decades, DSGE models have played important roles in empirical macroeconomic research. Many macroeconomists use these models to study the optimal policy, analyze the short-run business cycle, and forecast the variables of interest. Those studies, for the most part, neglect the nonlinearities by simply considering the conventionally linearized DSGE model for the U.S. economy over a long period, covering the so-called "Great Inflation" (GI), "Great Moderation" (GM) and "Great Recession" (GR). Compare to linear models, it is argued here that the nonlinear DSGE model can capture the nonlinear nature of the U.S. economy over time, resulting from structural change in the volatility of macroeconomic variables, and has better fit of the model to the data. This conjecture is clarified by evaluating the nonlinear DSGE model in terms of the model fitness and out-of-sample forecasting exercise, with a particular focus on the transition periods of GI-GM and GM-GR.

Several results are found in this paper. First, the nonlinear DSGE model has the better model fit for sample data covering a long time span, including the GI, GM and GR regimes. Moreover, it is found that the fitness of the nonlinear DSGE model to data improves over time in the transition period of GI-GM, but the fitness of the linear model worsens over time in the same transition period. As a result, it is argued that the nonlinear DSGE model is superior to the linear model when one considers a dataset covering a long time span, at least including both the GI and GM periods. Second, the nonlinear DSGE model generates unbiased forecasts of inflation and interest rates but biased forecasts of the real GDP growth rate in most cases. However, most of the forecasts generated from the linear DSGE model are biased. Lastly, in the forecast competition,

Table 3. Fluctuation Test

GI-GM Transition Period				GM-GR Transition Period			
Rolling Period	Y	INF	INT	Rolling Period	Y	INF	INT
1979Q1-1982Q4	0.99	0.00	0.99	2003Q1-2006Q4	0.99	0.07	0.77
1979Q2-1983Q1	0.99	0.00	0.99	2003Q2-2007Q1	0.99	0.10	0.81
1979Q3-1983Q2	0.99	0.00	0.99	2003Q3-2007Q2	0.99	0.14	0.71
1979Q4-1983Q3	0.99	0.00	0.99	2003Q4-2007Q3	0.99	0.12	0.42
1980Q1-1983Q4	0.99	0.00	0.99	2004Q1-2007Q4	0.99	0.10	0.06
1980Q2-1984Q1	0.99	0.00	0.99	2004Q2-2008Q1	0.99	0.08	0.05
1980Q3-1984Q2	0.99	0.00	0.99	2004Q3-2008Q2	0.99	0.08	0.04
1980Q4-1984Q3	0.99	0.00	0.99	2004Q4-2008Q3	0.99	0.09	0.04
1981Q1-1984Q4	0.99	0.00	0.99	2005Q1-2008Q4	0.95	0.14	0.10
1981Q2-1985Q1	0.99	0.00	0.99	2005Q2-2009Q1	0.96	0.12	0.31
1981Q3-1985Q2	0.99	0.00	0.99	2005Q3-2009Q2	0.98	0.08	0.49
1981Q4-1985Q3	0.99	0.00	0.99	2005Q4-2009Q3	0.99	0.04	0.60
1982Q1-1985Q4	0.99	0.00	0.99	2006Q1-2009Q4	0.99	0.03	0.65
1982Q2-1986Q1	0.99	0.00	0.99	2006Q2-2010Q1	0.99	0.03	0.70
1982Q3-1986Q2	0.99	0.00	0.99	2006Q3-2010Q2	0.99	0.03	0.76
1982Q4-1986Q3	0.99	0.00	0.99	2006Q4-2010Q3	0.99	0.03	0.89
1983Q1-1986Q4	0.99	0.00	0.99	2007Q1-2010Q4	0.99	0.03	0.93

Footnote: the values listed in Table are p-values; "Y", "INF" and "INT" are the abbreviations of "Real GDP growth rate", "Inflation rate" and "Interest Rate".

the nonlinear DSGE model is dominant on forecasting all variables of interest in both transition periods, except the inflation forecasts in the transition period of GI-GM.

References

[1] Adjemian, S., H. Boustani, F. Karamé, M. Juillard, J. Maih, F. Mihoubi, G. Perendia, J. Pfeifer, M. Ratto, and S. Villemot. (2014). Dynare Reference Manual v.4.4.3. Technical report, Cepremap.

[2] An, Sungbae. (2005). "Bayesian Estimation of DSGE Models: Lessons from Second-Order Approximations." Manuscript, University of Pennsylvania.

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

[4] Blanchard, Olivier J., and Charles M. Kahn. (1980). "The Solution of Linear Difference Models under Rational Expectations." *Econometrica*, 48, 5, 1305-1311.

[5] Canarella, Giorgio, Wen-Shwo Fang, Stephen M. Miller, and Stephen K. Pollard. (2010). "Is the Great Moderation Ending? — UK and US Evidence." *Modern Economy*, 1, 17-42.

[6] Charles, Amélie, Olivier Darné, and Laurent Ferrara. (2018). "Does the Great Recession Imply the End of the Great Moderation? International Evidence." *Economic Inquiry*, 56, 2, 745-760.

[7] Clarida, Richard, Jordi Gali, and Mark Gertler. (2000). "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics*, 115, 1, 147-180.

[8] Cogley, Timothy, and Thomas J. Sargent. (2005). "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US." *Review of Economic Dynamics*, 8, 2, 262-302.

[9] Creal, Drew. (2012). "A Survey of Sequential Monte Carlo Methods for Economics and Finance." *Econometric Reviews*, 31, 3, 245-296.

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

[11] Fernández-Villaverde, Jesús, and Juan F. Rubio-Ramírez. (2007). "Estimating Macroeconomic Models: A Likelihood Approach." *Review of Economic Studies*, 74, 4, 1059-1087.

[12] Ireland, Peter N. (2004). "Technology Shocks in the New Keynesian Model." *Review of Economics and Statistics*, 86, 4, 923-936.

[13] Kim, Chang-Jin, and Charles R. Nelson. (1999). "Has the US Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle." *Review of Economics and Statistics*, 81, 4, 608-616.

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

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

[16] Sims, Christopher. A. (2002). "Solving Linear Rational Expectations Models." *Computational Economics*, 20, 1, 1-20.

[17] Sims, Christopher A., and Tao Zha. (2006). "Were There Regime Switches in US Monetary Policy?" *American Economic Review*, 96, 1, 54-81.

[18] Stock, James H., and Mark W. Watson. (2003). "Has the Business Cycle Changed?" *In Proceedings-Economic Policy Symposium-Jackson Hole*, 9-56.