WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 12-4
BAYESIAN ESTIMATION OF DSGE MODELS

Pablo A. Guerrón-Quintana
Federal Reserve Bank of Philadelphia

James M. Nason
Federal Reserve Bank of Philadelphia

February 2012

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Bayesian Estimation of DSGE Models*

Pablo A. Guerrón-Quintana†

and

James M. Nason‡

February 2, 2012

Abstract

We survey Bayesian methods for estimating dynamic stochastic general equilibrium (DSGE) models in this article. We focus on New Keynesian (NK)DSGE models because of the interest shown in this class of models by economists in academic and policy-making institutions. This interest stems from the ability of this class of DSGE model to transmit real, nominal, and fiscal and monetary policy shocks into endogenous fluctuations at business cycle frequencies. Intuition about these propagation mechanisms is developed by reviewing the structure of a canonical NKDSGE model. Estimation and evaluation of the NKDSGE model rests on being able to detrend its optimality and equilibrium conditions, to construct a linear approximation of the model, to solve for its linear approximate decision rules, and to map from this solution into a state space model to generate Kalman filter projections. The likelihood of the linear approximate NKDSGE model is based on these projections. The projections and likelihood are useful inputs into the Metropolis-Hastings Markov chain Monte Carlo simulator that we employ to produce Bayesian estimates of the NKDSGE model. We discuss an algorithm that implements this simulator. This algorithm involves choosing priors of the NKDSGE model parameters and fixing initial conditions to start the simulator. The output of the simulator is posterior estimates of two NKDSGE models, which are summarized and compared to results in the existing literature. Given the posterior distributions, the NKDSGE models are evaluated with tools that determine which is most favored by the data. We also give a short history of DSGE model estimation as well as pointing to issues that are at the frontier of this research.

JEL Classification Numbers: $C32$, $E10$, $E32$.

Key Words: dynamic stochastic general equilibrium; Bayesian; Metropolis-Hastings; Markov chain Monte Carlo; Kalman filter; likelihood.

† *e-mail*: Pablo.Guerron@phil.frb.org, *voice*: (215) 574–3813, *address*: Research Department, Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia, PA 19106.

‡ *e-mail*: jim.nason@phil.frb.org, *voice*: (215) 574–3463, *address*: Research Department, Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia, PA 19106.

* This article was prepared for the *Handbook of Empirical Methods in Macroeconomics*, Michael Thornton and Nigar Hashimzade editors, to be published by Edward Elgar Publishing Ltd., in the *Handbooks of Research Methods and Applications* series. The views herein are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This paper is available at http://www.philadelphiafed.org/research-and-data/publications/working-papers/ free of charge.

1 Introduction

Macroeconomists have made substantial investments in Bayesian time series during the last 30 years. One reason is that Bayesian methods afford researchers the chance to estimate and evaluate a wide variety of macro models that frequentist econometrics often find challenging. Bayesian vector autoregressions (BVARs) represent an early return on this research project manifested, for example, by Doan, Litterman, and Sims (1984). They show that BVARs are useful forecasting tools.[1] More recent work focuses on developing Bayesian methods capable of estimating time-varying parameter (TVP) VARs, associated with Cogley and Sargent (2005) and Primiceri (2005), and Markov-switching (MS) VARs initiated by Sims and Zha (2006).[2] The complexity of TVP- and MS-VARs underline the efforts macroeconomists have put into developing useful Bayesian time series tools.[3]

Bayesian times series methods are also attractive for macroeconomists studying dynamic stochastic general equilibrium (DSGE) models. Although DSGE models can be estimated using classical optimization methods, macroeconomists often prefer to use Bayesian tools for these tasks. One reason is that advances in Bayesian theory are providing an expanding array of tools that researchers can employ to estimate and evaluate DSGE models. The popularity of the Bayesian approach is also explained by the increasing computational power available to estimate and evaluate medium- to large-scale DSGE models using Markov chain Monte Carlo (MCMC) simulators. These DSGE models can pose identification problems for frequentist estimators that no amount of data or computing power can overcome.

Macroeconomists are also drawn to the estimation and evaluation framework Bayesians have created because DSGE models are often seen as abstractions of actual economies. A frequentist econometrician might say that DSGE models are misspecified versions of the true model. This is not consistent with the beliefs often held about DSGE models. These beliefs are animated by the well known mantra that "all models are false." Since Bayesians eschew the existence of a true model, employing Bayesian methods to study DSGE models dovetails with the views held by many macroeconomists.

This chapter presents an overview of Bayesian time series methods that have been developed to estimate and evaluate linearized DSGE models.[4] We aim to bring the reader to the point where her priors and DSGE model can, subsequent to linearization, meet the data to be estimated and evaluated using Bayesian methods. The reader may wonder why this chapter puts aside nonlinear estimation of DSGE models. Since these methods represent the frontier, which is being pushed out at an extraordinary rate, a review of Bayesian nonlinear estimation of DSGE models waits for more consensus about the merits of the different approaches.[5]

We describe procedures for estimating a medium-scale New Keynesian (NK) DSGE model in this chapter. The NKDSGE model is a descendant of ones analyzed by Smets and Wouters (2003) and Christiano, Eichenbaum and Evans (2005). As those authors do, we estimate a linearized approximation of the NKDSGE. The linearization is grounded in the stochastically

[1] L. Kilian gives a progress report on BVARs in this handbook.

[2] This volume has surveys of MS models by J-Y. Pitarakis and TVP models by A. Hall and O. Boldea.

[3] L. Bauwens and D. Korobilis provide a chapter on Bayesian methods for macroeconomists in this handbook.

[4] Fernández-Villaverde, et al (2009) and Schorfheide (2011) review Bayesian estimation of DSGE models, while Canova (2007) and DeJong and Dave (2007) give textbook treatments of the subject.

[5] An and Schorfheide (2007), Fernández-Villaverde and Rubio-Ramírez (2007), Fernández-Villaverde, et al (2010), Aruoba, et al (2011), and Liu, Waggoner, and Zha (2011) propose different nonlinear estimators of DSGE models.

detrended optimality and equilibrium conditions because the growth rate of the technology shock is stationary. These optimality and equilibrium conditions yield a solution that is cast in state space form, which is the starting point for the Kalman filter. Since the Kalman filter generates predictions and updates of the state vector of the linearized NKDSGE model, we have a platform for computing its likelihood. This likelihood is used by Bayesian MCMC simulators to produce posterior distributions of NKDSGE model parameters given actual data and prior beliefs about these parameters. Posterior distributions represent confidence in an NKDSGE model conditional on the evidence provided by its likelihood. Marginal likelihoods are used to evaluate which member of a suite of NKDSGE models is most favored by the data.

A brief history of DSGE model estimation is presented in the next section. Our purpose is to give a framework for understanding the interaction between the need to connect macro theory to current data and the development of tools to achieve that task. Section 3 outlines the DSGE model we study. The NKDSGE model is prepared for estimation in section 4. This is followed by a discussion of Bayesian methods to estimate the linear approximate solution of the NKDSGE model described in section 5. Results appear in section 6. Section 7 concludes.

2 A Brief History of DSGE Model Estimation

Efforts to estimate and evaluate DSGE models using Bayesian methods began in ernest in the late 1990s. Previously, macroeconomists used classical optimization methods to estimate DSGE models. This section reviews these frequentist approaches to estimate DSGE models, covers the transition from frequentist to Bayesian methods, and ends by mentioning several issues at the frontier of Bayesian estimation of DSGE models.

Non-Bayesians have used maximum likelihood (ML), generalized method of moments (GMM), and indirect inference (II) to estimate DSGE models. These estimators rely on classical optimization either of a log likelihood function or of a GMM criterion.[6]

Early examples of frequentist ML estimation of DSGE models are Altuğ (1989) and Bencivenga (1992). They apply classical optimization routines to the log likelihood of the restricted finite-order vector autoregressive-moving average (VARMA) implied by the linear approximate solutions of their real business cycle (RBC) models. The restrictions arise because the VARMA lag polynomials are nonlinear functions of the DSGE model parameters.

A restricted VARMA engages an ML estimator that differs from the approach of Sargent (1989). He maps the linear solution of permanent income (PI) models with a serially correlated endowment shock into likelihoods that are built on Kalman filter innovations of the observed data and the associated covariance matrix. Sargent assumes that the data are ridden with measurement errors, which evolve as independent first-order autoregressions, AR(1)s.[7] This aids in identification because serially correlated measurement errors add restrictions to the VARMA implied by the PI model solution. An extension of Sargent's approach is Ireland (2001). He replaces the independent AR(1) measurement errors with an unrestricted VAR(1); see Curdia and Reis (2011) for a Bayesian version of this method. Besides measurement error, this VAR(1) inherits the sample data dynamics left unexplained by the RBC model that Ireland studies.

[6]This handbook has chapters on frequentist ML (GMM) DSGE model estimation by M. Fukač (F. Ruge-Murcia).

[7]Assuming sample data suffers from classical measurement error helps Altuğ identify the Kydland and Prescott (1982) RBC model. Bencivenga achieves the same objective with AR(1) taste shocks in an RBC model.

The tools of classical optimization are also useful for GMM estimation of DSGE models. Christiano and Eichenbaum (1992) construct GMM estimates of a subset of the parameters of their RBC model using its steady state conditions and the relevant shock processes as moments. Since the moment conditions are outnumbered by RBC model parameters, only a subset of these parameters are identified by GMM.

Identification also matters for ML estimation of DSGE models. For example, Altuğ, Bencivenga, and Ireland only identify a subset of RBC model parameters after pre-setting or calibrating several other parameters. Analysis by Hall (1996) suggests a reason for this practice. He shows that whether ML or GMM is being used, these estimators are relying on the same sample and theoretical information about first moments to identify DSGE model parameters. Although ML is a full information estimator, which engages all the moment conditions expressed by the DSGE model, GMM and ML rely on the same first moment information for identification. This suggests that problems identifying DSGE models are similar whether ML or GMM is the estimator of choice; see Fernández-Villaverde, et al (2009) for more discussion of these issues.

The frequentist assumption of a true model binds the identification problem to the issue of DSGE model misspecification. The question is whether any parameters of a DSGE model can be identified when it is misspecified. For example, frequentist ML loses its appeal when models are known to be misspecified.[8] Thus, it seems that no amount of data or computing power will solve problems related to the identification and misspecification of DSGE models.

A frequentist response to these problems is II. The first application of II to DSGE models is Smith (1993). He and Gourieroux, Monfort, and Renault (1993) note that II yields an estimator and specification tests whose asymptotic properties are standard even though the true likelihood of the DSGE model is not known.[9] The II estimator minimizes a GMM-like criterion in the distance between a vector of theoretical and sample moments. These moments are readily observed in the actual data and predicted by the DSGE model. Estimating DSGE model parameters is "indirect" because the objective of the GMM-like criterion is to match moments not related directly to the structure of the DSGE model.[10] Theoretical moments are produced by simulating synthetic data from the solution of the DSGE model. A classical optimizer moves the theoretical moments closer to the sample moments by updating the DSGE model parameters holding the structural shock innovations fixed.[11]

Dridi, Guay, and Renault (2007) extend the II estimator by acknowledging that the DSGE model is false. They argue that the purpose of dividing the vector of DSGE model parameters, Θ, into the parameters of interest, Θ_1, and the remaining nuisance or pseudo-parameters, Θ_2, is to separate the part of a DSGE model having economic content from the misspecified part. Thus, Θ_1 represents the part of a DSGE model that is economically relevant for the moments it aims to match. However, Θ_2 cannot be ignored because it is integral to the DSGE model. Fixing Θ_2 or calibrating it with sample information contributes to identifying Θ_1, but without

[8] White (1982) develops quasi-ML for misspecified models, but its consistency needs a strong set of assumptions.

[9] Gregory and Smith (1990, 1991) anticipate the II approach to DSGE model estimation and evaluation.

[10] Also, II can estimate DSGE model parameters by minimizing the distance between the likelihoods of an auxiliary model generated using actual and simulated samples. Simulated quasi-ML yields an asymptotically less efficient estimator because the likelihood of the auxiliary model differs from that of the DSGE model; see Smith (1993).

[11] Christiano et al (2005) estimate an NKDSGE model by matching its predicted impulse responses to those of an SVAR. This approach to moment matching is in the class of II estimators. See Canova and Sala (2009) for a discussion of the identification problem facing this estimator and Hall, et al (2012) for an optimal impulse response matching estimator of DSGE models.

polluting it with the misspecification of the DSGE model encapsulated by Θ_2. This insight is the basis for Dridi, Guay, and Renault (DGR) to construct an asymptotic distribution of Θ_1 that accounts for misspecification of the DSGE model. The sampling theory is useful for tests of the degree of misspecification of the DSGE model and to gauge its ability to match the data.

Whether identification of DSGE models is a problem for Bayesians is not clear. For many Bayesians all that is needed for identification is a well posed prior.[12] Poirier (1998) points out that this position has potential costs in that prior and posterior distributions can be equivalent if the data are uninformative. This problem differs from identification problems frequentists face. Identification of a model is a problem that arises in population for a frequentist estimator, while for a Bayesian the source of the equivalence is data interacting with the prior. Nonetheless, Poirier provides analysis suggesting that Θ be split into those parameters for which the data are informative, Θ_1, given the priors from those, Θ_2, for which this is not possible.

Bayesians avoid having to assume there exists a true or correctly specified DSGE model because of the likelihood principle (LP). The LP is a foundation of Bayesian statistics and says that all evidence about a DSGE model is contained in its likelihood conditional on the data; see Berger and Wolpert (1988). Since the data's probabilistic assessment of a DSGE model is summarized by its likelihood, the likelihoods of a suite of DSGE models possess the evidence needed to judge which "best" fit the data. Thus, Bayesian likelihood-based evaluation is consistent with the view that there is no true DSGE model because, for example, this class of models is afflicted with incurable misspecification.

There exist several Bayesian approaches to estimate DSGE models. Most of these methods are fully invested in the LP, which implies likelihood-based estimation. The goal of Bayesian estimation is construction of the posterior distribution, $\mathcal{P}(\Theta|\mathcal{Y}_T)$, of DSGE model parameters conditional on sample data $\mathcal{Y}_T$ of length T. Bayesian estimation exploits the fact that the posterior distribution equals the DSGE model likelihood, $\mathcal{L}(\mathcal{Y}_T|\Theta)$, multiplied by the econometrician's priors on the DSGE model parameters, $\mathcal{P}(\Theta)$, up to a factor of proportionality

$$\mathcal{P}\Big(\Theta\Big|\mathcal{Y}_T\Big) \;\propto\; \mathcal{L}\Big(\mathcal{Y}_T\Big|\Theta\Big)\mathcal{P}\Big(\Theta\Big). \tag{1}$$

Bayesian estimation of DSGE models is confronted by posterior distributions too complicated to evaluate analytically. The complication arises because the mapping from a DSGE model to its $\mathcal{L}(\mathcal{Y}_T|\Theta)$ is nonlinear in Θ, which suggests using simulation to approximate $\mathcal{P}(\Theta|\mathcal{Y}_T)$.

Among the earliest examples of Bayesian likelihood-based estimation of a DSGE model is DeJong, Ingram, and Whiteman (2000a, b). They engage importance sampling to compute posterior distributions of functions of Θ, $\mathcal{G}(\Theta)$.[13] Importance sampling relies on a finite number N of $\mathcal{IID}$ random draws from an arbitrary density $\mathcal{D}(\Theta)$ to approximate $\mathcal{G}(\Theta)$. The approximation is computed with weights that smooth $\mathcal{G}(\Theta)$. The weights, $\mathcal{W}(\Theta_i)$, $i = 1, \ldots, N$, smooth the approximation by giving less (greater) mass to posterior draws of $\mathcal{G}(\Theta_i)$ that occur frequently (infrequently).[14] One drawback of importance sampling is that it is often unreliable when Θ has large dimension. Another is that there is little guidance about updating $\mathcal{P}(\Theta|\mathcal{Y}_t)$, and therefore $\mathcal{G}(\Theta)$, from one draw of $\mathcal{D}(\Theta)$ to the next, given $\mathcal{P}(\Theta)$.

[12] This is a proper prior that is independent of the data and has a density that integrates to one.

[13] The objective is to approximate $\mathbf{E}\{\mathcal{G}(\Theta)\} = \int \mathcal{G}(\Theta)\mathcal{P}(\Theta|\mathcal{Y}_t)d\Theta \;/\; \int \mathcal{P}(\Theta|\mathcal{Y}_t)d\Theta$.

[14] Given N draws from $\mathcal{D}(\Theta)$, $\mathbf{E}\{\mathcal{G}(\Theta)\}$ is approximated as $\overline{\mathcal{G}}_N = \sum_{i=1}^{N}\mathcal{W}(\Theta_i)\mathcal{G}(\Theta_i) \;/\; \sum_{i=1}^{N}\mathcal{W}(\Theta_i)$, where the weights, $\mathcal{W}(\Theta_i)$, equal $\mathcal{P}(\Theta_i|\mathcal{Y}_t) \;/\; \mathcal{D}(\Theta_i)$.

Otrok (2001) reports estimates of a DSGE model grounded on the Metropolis-Hasting (MH) algorithm. This is, perhaps, the first instance of MH-MCMC simulation applied to DSGE model estimation. The MH algorithm proposes to update Θ using a multivariate random walk, but first an initial draw of Θ from $\mathcal{P}(\Theta)$ is needed. The initial Θ is updated by adding to it draws from a distribution of "shock innovations." The decision to keep the initial Θ or to move to the updated Θ depends on whether the latter increases $\mathcal{L}(\mathcal{Y}_t|\Theta)$. This process is repeated by sampling from the multivariate random walk to update Θ.

The MH-MCMC simulator is often preferred to importance sampling methods to estimate DSGE models. One reason is that the MH algorithm places less structure on the MCMC simulator. Thus, a wide class of time series models can be estimated by MH-MCMC simulation. Also MH-MCMC simulators generate serial correlation in the posterior distribution, which induces good asymptotic properties, especially compared to importance samplers. These properties reduce the computational burden of updating the prior. Another useful feature of MH-MCMC simulation is that its flexibility lessens the demands imposed by high dimensional Θ. We postpone further discussion of the MH-MCMC simulator to section 5.3.

Bayesian estimation of NKDSGE models leans heavily on MH-MCMC simulation. Smets and Wouter (2003, 2007), Del Negro and Schorfheide (2004), and Del Negro, Schorfheide, Smets and Wouter (2007) estimate NKDSGE models similar to the one we estimate below. Open economy NKDSGE models are estimated using MH-MCMC simulators by, among others, Adolfson, Laséen, Lindé, and Villani (2007), Lubik and Schorfheide (2007), Kano (2009), Justiniano and Preston (2010), Rabanal and Tuesta (2010), and Guerrón-Quintana (2010b). Evidence of the wide applicability of the MH-MCMC algorithm is its applications to NKDSGE models with labor market search by Sala, Söderström, and Trigari (2008), with fiscal and monetary policy interactions by Leeper, Plante, and Traum (2010), and that compare sticky price monetary transmission to monetary search frictions by Aruoba and Schorfheide (2011).

Formal Bayesian evaluation of estimated DSGE models relies on Bayes factors or posterior odds ratios. The Bayes factor is

$$\mathcal{B}_{j,s|\mathcal{Y}_T} = \frac{\mathcal{L}\Big(\mathcal{Y}_T \,\Big|\, \Theta_j, \mathcal{M}_j\Big)}{\mathcal{L}\Big(\mathcal{Y}_T \,\Big|\, \Theta_s, \mathcal{M}_s\Big)}, \tag{2}$$

which measures the odds the data prefer DSGE model j, $\mathcal{M}_j$ (with parameter vector Θ_j), over $\mathcal{M}_s$.[15] Multiply $\mathcal{B}_{j,s|\mathcal{Y}_T}$ by the prior odds to find the posterior odds ratio, which as the name suggests is $\mathcal{R}_{j,s|\mathcal{Y}_T} = \mathcal{B}_{j,s|\mathcal{Y}_T}\mathcal{P}(\Theta_j)/\mathcal{P}(\Theta_s)$. Put another way, the log of the Bayes factor is the log of the posterior odds of $\mathcal{M}_j$ compared to $\mathcal{M}_s$ net of the log of the prior odds of these DSGE models. Geweke (1999, 2005) and Fernández-Villaverde and Rubio-Ramírez (2004) discuss the foundations of Bayesian evaluation of DSGE models, while Rabanal and Rubio-Ramírez (2005) calculate Bayes factors to gauge the fit of several NKDSGE models.

There are other Bayesian approaches to DSGE model evaluation. Schorfheide (2000) estimates DSGE models using the MH-MCMC simulator as well as a richly parameterized structural BVAR, which serves as a "reference" model. The fit of the DSGE and reference models to the data is judged within a Bayesian decision problem using a few selected moments under symmetric and asymmetric loss functions. The moments are structural IRFs that have

[15] In general, Bayes factor involves the ratio of marginal likelihoods of $\mathcal{M}_j$ and $\mathcal{M}_s$. The marginal likelihood integrates out Θ_j from $\mathcal{L}(\mathcal{Y}_T|\Theta_j, \mathcal{M}_j)$; see Geweke (2005).

economic meaning within the context of the DSGE models. Problems of DSGE model misspecification are sidestepped in this non-LP-based Bayesian evaluation process because, according to Schorfheide, the moments on which the DSGE models are evaluated are identified by the structural BVAR. He also argues that this approach yields valid DSGE model evaluation when no DSGE model fits the model well, which is not true of the Bayes factor; also see Geweke (2010). This argument is similar to arguments DGR make for parsimony (*i.e.*, do not rely on all the moments inherent in the likelihood), when selecting moments to bind the DSGE model to the data for II estimation.[16] DGR are guided to choose moments most economically meaningful for the DSGE model, which is a frequentist analogue to Schorfheide's Bayesian approach.

Another interesting approach to these issues is Guerrón-Quintana (2010a). He confronts a NKDSGE model with different sets of observed aggregate variables to ask which data set is most informative for estimating DSGE model parameters. Fixing the NKDSGE models and changing the observed data rules out using the posterior odds ratio to conduct model evaluation. Instead, Guerrón-Quintana engages impulse response functions and out-of-sample forecast exercises to choose among the competing data sets. These evaluation tools reveal that the posterior of a DSGE model is affected by the composition and size of the information sets used in Bayesian MH-MCMC estimation, which is a signal of misspecification.

Identification of DSGE models has become a research frontier for Bayesian econometrics. We mention briefly several here. One approach is Müller (2010). He constructs statistics that unwind the relative contributions of the prior and the likelihood to the posterior. These statistics measure the "identification strength" of DSGE model parameters with respect to a specific prior. Koop, Pesaran, and Smith (2011) describe two methods that depend on computing conditional and marginal posterior distributions for checking identification of DSGE models. Another useful approach is found in Guerrón-Quintana, Inoue, and Kilian (2010). When DSGE models are weakly identified (*i.e.*, Bayesian posterior distribution cannot be viewed as frequentist confidence sets), they advocate inverting the Bayes factor to construct confidence intervals with good small sample properties. We return to these issues at the end of this chapter.

3 A Canonical New Keynesian DSGE Model

This section builds a canonical NKDSGE model inspired by the recent literature. The specification of this NKDSGE model is similar to those estimated by Del Negro, Schorfheide, Smets, and Wouters (2007), Smets and Wouters (2007) and Del Negro and Schorfheide (2008), who in turn build on Smets and Wouters (2003) and Christiano, et al (2005).[17] The main features of the NKDSGE model are (a) the economy grows along a stochastic path, (b) prices and wages are assumed to be sticky à la Calvo, (c) preferences display internal habit formation in consumption, (d) investment is costly, and (e) there are five exogenous shocks. There are shocks to the monopoly power of the final good firm, the disutility of work, government spending and a shock to the growth rate of labor neutral total factor productivity (TFP). All of these shocks are stationary AR(1)s. The fifth is a monetary policy shock embedded in a Taylor rule.

[16] Kim (2002), Chernozhukov and Hong (2003), and Sims (2007) give Bayesian treatments of GMM and other limited information estimators.

[17] See the chapter in this handbook by P. Levine for a plethora of DSGE model specifications.

3.1 Firms

There is a continuum of monopolistically competitive firms indexed by $j \in [0,1]$. A firm produces an intermediate good using capital services, $k_{j,t}$, and labor services, $L_{j,t}$, which are rented in perfectly competitive markets. The production function of firm j is given by

$$Y_{j,t} = k_{j,t}^{\alpha}\left(Z_t L_{j,t}\right)^{1-\alpha} - \kappa Z_t, \quad \alpha \in \left(0,1\right), \ \kappa > 0, \tag{3}$$

where Z_t is labor neutral TFP common to all firms. The term κZ_t is removed from the output of firm j to guarantee that steady state profits are zero as well as to generate the period-by-period fixed cost needed to support monopolistic competition among intermediate goods firms. We assume that the growth rate of the TFP shock, $z_t = \ln\left(Z_t/Z_{t-1}\right)$, is an AR(1) process

$$z_t = \left(1-\rho_z\right)\gamma + \rho_z z_{t-1} + \sigma_z \epsilon_{z,t}.$$

This AR(1) is stationary around the deterministic TFP growth rate γ (> 0) because $|\rho_z| < 1$ and the innovation of z_t is time invariant and homoskedastic, $\epsilon_{z,t} \sim \mathcal{NID}(0,1)$ with $\sigma_z > 0$.[18]

Firm j chooses its price $P_{j,t}$ to maximize the present value of profits subject to the restriction that changes in their prices are time dependent. This form of price stickiness is called Calvo pricing; see Yun (1996). At each date t, a fraction of the unit mass of firms are able to update their price to its optimal level. The remaining firms update their prices by a fraction of the economy-wide lagged inflation rate, π_{t-1}. Inflation is defined as the growth rate of the aggregate price level, $\pi_t = P_t/P_{t-1} - 1$. We posit that firms are able to revise their prices at the exogenous probability $1-\zeta_p$ every date t, while a firm not re-optimizing its price updates according to the rule: $P_{j,t} = \left(\pi^*\right)^{1-\iota_p}\left(\pi_{t-1}\right)^{\iota_p} P_{j,t-1}$, where π^* is steady state inflation and $\iota_p \in [0,1]$. This has firms indexing (the log) of their prices to inflation to a weighted average of steady state inflation and lagged inflation, according to the weight ι_p, in periods when re-optimization is not allowed.

There is a competitive firm that produces the final good using intermediate goods aggregated using the technology

$$Y_t = \left[\int_0^1 Y_{j,t}^{1/\left(1+\lambda_{f,t}\right)} dj\right]^{1+\lambda_{f,t}},$$

where $\lambda_{f,t}$ is the time-varying degree of monopoly power (*i.e.*, the stochastic price elasticity is $[1+\lambda_{f,t}]/\lambda_{f,t}$). This monopoly power evolves according to the AR(1) process

$$\ln\lambda_{f,t} = \left(1-\rho_{\lambda_f}\right)\ln\lambda_f + \rho_{\lambda_f}\ln\lambda_{f,t-1} + \sigma_{\lambda_f}\epsilon_{\lambda,t},$$

where $|\rho_{\lambda_f}| < 1$, $\lambda_f, \ \sigma_{\lambda_f} > 0$, and $\epsilon_{\lambda,t} \sim \mathcal{NID}(0,1)$.

[18]A strictly positive deterministic growth term γ is also needed to have a well-defined steady state around which we can linearize and solve the NKDSGE model.

3.2 Households

The economy is populated by a continuum of households indexed by address $i \in [0, 1]$. Household i derives utility over "net" consumption and the disutility of work.[19] This relationship is summarized by the period utility function

$$\mathcal{U}\Big(C_{i,t},\, C_{i,t-1},\, L_{i,t}; \phi_t\Big) \;=\; \ln\Big(C_{i,t} - hC_{i,t-1}\Big) \;-\; \phi_t \frac{L_{i,t}^{1+\nu_l}}{1+\nu_l}, \tag{4}$$

where $C_{i,t}$ and $L_{i,t}$ are consumption and labor supply of household i, ν_l is the inverse of the Frisch labor supply elasticity, and ϕ_t is an exogenous and stochastic preference shifter. Period utility receives the flow of $C_{i,t}$ net of a fraction h of $C_{i,t-1}$, which is the habit in consumption displayed by preferences. Consumption habit is internal to households and governed by the preference parameter $h \in (0, 1)$. The preference shifter follows the AR(1) process

$$\ln \phi_t \;=\; \Big(1 - \rho_\phi\Big) \ln \phi + \rho_\phi \ln \phi_{t-1} + \sigma_\phi \epsilon_{\phi,t},$$

with $|\rho_\phi| < 1$, $\sigma_\phi > 0$, and $\varepsilon_{\phi,t} \sim \mathcal{NID}(0, 1)$.

Households are infinitely-lived. For household i, this means that it maximizes the expected present discounted value of period utility

$$\mathbb{E}_0^i \sum_{t=0}^{\infty} \beta^t \mathcal{U}\Big(C_{i,t},\, C_{i,t-1},\, L_{i,t}; \phi_t\Big), \quad \beta \in (0, 1), \tag{5}$$

subject to the budget constraint

$$P_t C_{i,t} + P_t \Big[I_{i,t} + a(u_{i,t})\overline{K}_{i,t}\Big] + B_{i,t+1} \;=\; R_t^K u_{i,t}\overline{K}_{i,t} + W_{i,t}L_{i,t} + R_{t-1}B_{i,t} + A_{i,t} + \Pi_t + T_{i,t}, \tag{6}$$

and the law of motion of capital

$$\overline{K}_{i,t+1} \;=\; (1-\delta)\overline{K}_{i,t} + I_{i,t}\left[1 - \Gamma\left(\frac{I_{i,t}}{I_{i,t-1}}\right)\right], \quad \delta \in (0, 1), \tag{7}$$

over uncertain streams of consumption, labor supply, capital intensity, $u_{i,t}$, investment, $I_{i,t}$, capital, $\overline{K}_{i,t+1}$, and 1-period government bonds, $B_{i,t+1}$. Here $\mathbb{E}_t^i$ is the expectation operator conditional on the information set available to household i at time t; $a(\cdot)$ is the cost (in units of the consumption good) household i generates when working $\overline{K}_{i,t+1}$ at intensity $u_{i,t}$; R_t^K is the nominal rental rate of capital; $W_{i,t}$ is the nominal wage household i charges for hiring out $L_{i,t}$; R_{t-1} is the gross nominal interest rate paid on $B_{i,t}$; $A_{i,t}$ captures net payments from complete markets; Π_t corresponds to profits from intermediate goods producers; $T_{i,t}$ corresponds to lump-sum transfers from the government to household i; and $\Gamma(\cdot)$ is a function reflecting costs associated with adjusting the flow $I_{i,t}$ into $\overline{K}_{i,t+1}$. The function $\Gamma(\cdot)$ is assumed to be increasing and convex satisfying $\Gamma(\gamma^*) = \Gamma'(\gamma^*) = 0$ and $\Gamma''(\gamma^*) > 0$, where $\gamma^* \equiv \exp(\gamma)$. Also note that $\overline{K}_t \equiv \int \overline{K}_{i,t} di$ is the aggregate stock of capital. Given $u_{i,t}$ is a choice variable for household i, the nominal return on capital is $R_t^K u_{i,t}\overline{K}_{i,t}$ gross of the real cost $a(u_{i,t})$. The cost function $a(\cdot)$ satisfies the restrictions $a(1) = 0$, $a'(1) > 0$, and $a''(1) > 0$.

[19]Agents in the economy are given access to complete insurance markets. This assumption is needed to eliminate wealth differentials arising from wage heterogeneity.

3.3 Staggered Nominal Wage Setting

Erceg et al. (2000) introduce Calvo staggered nominal wage setting into an NKDSGE model. We adopt their approach. Assume that household i is a monopolistic supplier of a differentiated labor service, $L_{i,t}$. Households sell these labor services to a firm that aggregates labor and sells it to final firms. This firm aggregates household labor services using the technology

$$L_t = \left[\int_0^1 L_{i,t}^{1/(1+\lambda_W)} dj \right]^{1+\lambda_W}, \quad 0 < \lambda_W < \infty$$

where the nominal wage elasticity is $(1 + \lambda_W)/\lambda_W$.

The role of this firm is to sell aggregate labor services, L_t, to intermediate goods firms in a perfectly competitive market at the aggregate nominal wage, W_t. The relationship between L_t, $L_{i,t}$, $W_{i,t}$, and W_t is given by

$$L_{i,t} = \left[\frac{W_{i,t}}{W_t} \right]^{-(1+\lambda_W)/\lambda_W} L_t.$$

We assume, as Erceg et al. (2000) did to induce wage sluggishness, that household i is allowed to reset its nominal wage in a similar manner to the approach that intermediate goods firms are forced to use to update the prices of their output. Calvo staggered nominal wage setting permits households to re-optimize their labor market decisions at the fixed exogenous probability $1-\zeta_W$ during each date t. Households not allowed to reset their nominal wages optimally employ the rule $W_{i,t} = (\pi^*\gamma^*)^{1-\iota_W} (\pi_{t-1} \exp(z_{t-1}))^{\iota_W} W_{i,t-1}$ to update, where $\iota_W \in [0, 1]$. This rule indexes (the log) of those nominal wages not being set optimally to a weighted average of steady state inflation grossed up by the deterministic growth rate and lagged inflation grossed up by lagged TFP growth, where ι_W determines the weights.

3.4 The Government

As often in the new Keynesian literature, we assume a cashless economy; see Woodford (2003). The monetary authority sets the short-term interest rate according to the Taylor rule used in Del Negro et al. (2007) and Del Negro and Schorfheide (2008)

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*}\right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi^*}\right)^{\psi_1} \left(\frac{Y_t}{Y_t^{\mathcal{T}}}\right)^{\psi_2} \right]^{1-\rho_R} \exp\Big(\sigma_R \epsilon_{r,t}\Big), \tag{8}$$

where R^* (> 0) corresponds to the steady state gross nominal interest rate, steady state inflation is π^*, $Y_t^{\mathcal{T}}$ denotes the target level of output, $\epsilon_{r,t}$ is a random shock to the systematic component of monetary policy, which is distributed $\mathcal{NID}(0, 1)$, and σ_r (> 0) is the size of the monetary shock. The Taylor rule has the central bank systematically smoothing its policy rate by ρ_R as well as responding to deviations of π_t from its steady state π^*, and of Y_t from its target $Y_t^{\mathcal{T}}$.

Finally, we assume that government spending is a time-varying fraction of output, $G_t = (1 - 1/g_t)\, Y_t$. The fraction is driven by the shock g_t, which follows the AR(1) process

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

where $|\rho_g| < 1$, g^*, $\sigma_g > 0$, and $\epsilon_{g,t} \sim \mathcal{NID}(0,1)$. Although taxes and 1-period bonds are notionally used to finance G_t, the government inhabits a Ricardian world such that along the equilibrium path 1-period bonds are in zero net supply, $B_t = 0$, at all dates t. This forces aggregate lump sum taxes, T_t, always to equal G_t (*i.e.*, the primary surplus, $T_t - G_t$, is zero).

4 Preparing the NKDSGE Model for Estimation

The scale of the NKDSGE model suggests that it does not admit a closed-form solution. Hence, we rely on linearization to obtain an approximate solution. The procedure consists of computing a first-order approximation of the NKDSGE model around its non-stochastic steady state.[20]

4.1 Stochastic Detrending

The productivity shock Z_t is non-stationary (*i.e.*, has a unit root). Since its growth rate, z_t, is stationary, the NKDSGE model grows along a stochastic path. We induce stationarity in the NKDSGE model by dividing the levels of trending real variables Y_t, C_t, I_t, and $\overline{K}_t$ by Z_t. This is the detrending step, where for example $y_t = Y_t/Z_t$. The nominal wage W_t also needs to be detrended after dividing it by the price level to obtain the detrended real wage, $w_t = W_t/(P_t Z_t)$. To transform the nominal rental rate of capital into the real rate, divide by P_t, $r_t^k = R_t^k/P_t$.

4.2 Linearization

We engage a first-order Taylor or linear approximation to solve the NKDSGE model. The linear approximation is applied to the levels of the variables found in the nonlinear optimality and equilibrium conditions of the NKDSGE model.[21] The first step is to detrend the optimality and equilibrium conditions. Consider the production function (3), which after detrending becomes

$$y_{j,t} = k_{j,t}^{\alpha} L_{j,t}^{1-\alpha} - \kappa.$$

We avoid excessive notation by representing the original and detrended levels of capital in firm j with k_j. Denote $\tilde{y}_{j,t}$ as the deviation of output from its steady state, $\tilde{y}_{j,t} = y_{j,t} - y_j$. Taking a linear approximation of the previous expression gives

$$\tilde{y}_{j,t} = \alpha \tilde{k}_{j,t} + (1-\alpha)\tilde{L}_{j,t}.$$

The approach is easily extended to the remaining equilibrium and optimality conditions. Del Negro and Schorfheide (2008) present the complete set of linearized optimality and equilibrium conditions of the NKDSGE model.

[20] A first-order approximation is sufficient for many macroeconomic applications. Otherwise, see Fernandez-Villaverde et al. (2010, 2011) for tools to solve and estimate DSGE models with higher-order approximations.

[21] First-order approximations can also linearize many variables in logs rather than in levels

4.3 Solution

Once the model has been detrended and linearized, the collection of its equilibrium conditions can be cast as an expectational stochastic difference equation

$$\mathbb{E}_t\Big\{\mathcal{F}\Big(N_{t+1},\ N_t,\ X_{t+1},\ X_t\Big)\Big\} \;=\; 0, \tag{9}$$

where X_t and N_t are vectors of predetermined (states) and non-predetermined (controls) variables, respectively. These vectors include

$$X_t \;\equiv\; \Big[\tilde{y}_{t-1}\ \ \tilde{c}_{t-1}\ \ \tilde{i}_{t-1}\ \ \tilde{k}_t\ \ \tilde{w}_{t-1}\ \ \tilde{R}_{t-1}\ \ \tilde{\pi}_{t-1}\ \ \tilde{z}_t\ \ \tilde{g}_t\ \ \tilde{\phi}_t\ \ \tilde{\lambda}_{f,t}\Big]'$$

and

$$N_t \;\equiv\; \Big[\tilde{y}_t\ \ \tilde{c}_t\ \ \tilde{i}_t\ \ \tilde{l}_t\ \ \tilde{r}^k_t\ \ \tilde{u}_t\ \ \tilde{w}_t\ \ \tilde{\pi}_t, \tilde{R}_t\Big]'.$$

whose elements are deviations from their steady state values. Hence, finding the solution of the model is tantamount to solving the system of linear stochastic difference equations (9). We rely on a suite of programs developed by Stephanie Schmitt-Grohe and Martin Uribe to solve for the linear approximate equilibrium decision rules of the state variables of the NKDSGE model.[22] The solution of the NKDSGE model takes the form

$$\begin{aligned} X_t &= \Pi X_{t-1} + \Phi \xi_t \\ N_t &= \Psi X_t, \end{aligned} \tag{10}$$

where the first system of equations is the linear approximate equilibrium decision rules of the state variables, the second set maps from the state variables to the control variables, Π, Φ, and Ψ are matrices that are nonlinear functions of the structural parameters of the NKDSGE model, and ξ_t is the vector of structural innovations, $\Big[\epsilon_{z,t}\ \ \epsilon_{\lambda,t}\ \ \epsilon_{\phi,t}\ \ \epsilon_{r,t}\ \ \epsilon_{g,t}\Big]'$.

5 Bayesian Estimation of the NKDSGE Model

This section presents the tools needed to generate Bayesian estimates of the linear approximate NKDSGE model of the previous section. Bayesian estimation employs the Kalman filter to construct the likelihood of the NKDSGE model. Next, priors for the NKDSGE model are reported because the likelihood multiplied by the prior is proportional to the posterior according to expression (1). We end this section by reviewing several details of the MH-MCMC simulator.

[22] These programs are available at http://www.columbia.edu/ mu2166/2nd_order.htm. Other examples of widely used software to solve DSGE models are found in the Dynare and Iris software packages. This handbook includes reviews of Dynare and Iris by J. Madeira and J. Beneš, respectively.

5.1 The Kalman Filter and the Likelihood

A key step in Bayesian MH-MCMC estimation of a linearized NKDSGE model is evaluation of its likelihood. A convenient tool to evaluate the likelihood of linear models is the Kalman filter. The Kalman filter generates projections or forecasts of the state of the linear approximate solution (10) of the NKDSGE model given an information set of observed macro time series. Forecasts of these observables are also produced by the Kalman filter. The Kalman filter is useful for evaluating the likelihood of a linearized NKDSGE model because the forecasts are optimal within the class of all linear models. When shock innovations and the initial state of the NKDSGE model are assumed to be Gaussian (*i.e.*, normally distributed), the Kalman filter renders forecasts that are optimal against all data-generating processes of the states and observables. Another implication is that at date t the observables are normally distributed with mean and variance that are functions of forecasts of the state of the linearized NKDSGE model and lagged observables. Thus, the Kalman filter provides the building blocks of the likelihood of a linear approximate NKDSGE model.

We describe the link between the solution of the linearized NKDSGE model with the Kalman filter.[23] Define the expanded vector of states as $\mathbb{S}_t = \left[N_t' \;\; X_t'\right]'$. Using this definition, the state space representation of the NKDSGE model consists of the system of state equations

$$\text{(11.1)} \qquad \mathbb{S}_t \; = \; \mathbb{F}\mathbb{S}_{t-1} + \mathbb{Q}\xi_t, \quad \xi_t \; \sim \mathcal{NID}\left(\mathbf{0}, \mathbf{I}_m\right),$$

and the system of observation equations

$$\text{(11.2)} \qquad \mathbb{Y}_t \; = \; \mathbb{M} + \mathbb{H}\mathbb{S}_t + \xi_{u,t}, \quad \xi_{u,t} \; \sim \mathcal{NID}\left(\mathbf{0}, \Sigma_u\right).$$

Here, $\mathbb{Y}_t$ corresponds to the vector of observables at time t; $\mathbb{F}$ and $\mathbb{Q}$ are functions of the matrices Π, Φ, and Ψ; the matrix $\mathbb{H}$, which contains zeros and ones, relates the model's definitions with the data; $\mathbb{M}$ is a vector required to match the means of the observed data; and $\xi_{u,t}$ is a vector of measurement errors. Assume the vector of observables and the vector of states have dimensions m and n, respectively. Also, define $\mathbb{S}_{t|t-1}$ as the conditional forecast or expectation of $\mathbb{S}_t$ given $\{\mathbb{S}_1, \ldots, \mathbb{S}_{t-1}\}$, or $\mathbb{S}_{t|t-1} \equiv \mathbf{E}\left[\mathbb{S}_t | \mathbb{S}_1, \ldots, \mathbb{S}_{t-1}\right]$. Its mean square error or covariance matrix is $P_{t|t-1} \equiv \mathbf{E}\Big[\left(\mathbb{S}_t - \mathbb{S}_{t-1}\right)\left(\mathbb{S}_t - \mathbb{S}_{t-1}\right)'\Big]$.

The likelihood of the linearized NKDSGE model is built up by generating forecasts from the state space system (11.1) and (11.2) period-by-period

$$\text{(12)} \qquad \mathcal{L}\left(\mathcal{Y}_T \middle| \Theta\right) \; = \; \prod_{t=1}^{T} \mathcal{L}\left(\mathbb{Y}_t \middle| \mathcal{Y}_{t-1}, \Theta\right),$$

where $\mathcal{L}\left(\mathbb{Y}_t | \mathcal{Y}_{t-1}, \Theta\right)$ is the likelihood conditional on the information available up to date $t-1$ and to be clear $\mathcal{Y}_{t-1} \equiv \{\mathbb{Y}_0, \ldots, \mathbb{Y}_{t-1}\}$. The Kalman filter computes this likelihood using the following steps:

[23] See Anderson and Moore (2005) for more information on linear filtering and Harvey (1989) for details on the Kalman filter and likelihood-based estimation.

1. Set $\mathbb{S}_{1|0} = 0$ and $P_{1|0} = \mathbb{F}P_{0|0}\,\mathbb{F} + \mathcal{Q}', \quad \mathcal{Q}' = \mathbb{Q}\mathbb{Q}'$.[24]
2. Compute $\mathbb{Y}_{1|0} = \mathbb{H}'\mathbb{S}_{1|0} = 0,\ \Omega_{1|0} = E\Big(\Big[\mathbb{Y}_1 - \mathbb{Y}_{1|0}\Big]\Big[\mathbb{Y}_1 - \mathbb{Y}_{1|0}\Big]'\Big) = \mathbb{H}'\,P_{1|0}\,\mathbb{H} + \Sigma_u$.
3. The predictions made in Steps 1 and 2 produce the date 1 likelihood:

$$\mathcal{L}\Big(\mathbb{Y}_1\Big|\Theta\Big) \;=\; (2\pi)^{-m/2}\left|\Omega_{1|0}^{-1}\right|^{1/2}\exp\left[-\frac{1}{2}\left(\mathbb{Y}_1'\Omega_{1|0}^{-1}\mathbb{Y}_1\right)\right].$$

4. Next, update the date 1 forecasts:

$$\mathbb{S}_{1|1} \;=\; \mathbb{S}_{1|0} + P_{1|0}\mathbb{H}\Omega_{1|0}^{-1}\Big(\mathbb{Y}_1 - \mathbb{Y}_{1|0}\Big),$$
$$P_{1|1} \;=\; P_{1|0} - P_{1|0}\mathbb{H}\Omega_{1|0}^{-1}\mathbb{H}'P_{1|0}.$$

5. Repeat steps 2, 3, and 4 to generate Kalman filter predictions of $\mathbb{S}_t$ and $\mathbb{Y}_t$:

$$\mathbb{S}_{t|t-1} \;=\; \mathbb{F}\mathbb{S}_{t-1}P_{t|t-1},$$

$$P_{t|t-1} \;=\; \mathbb{F}P_{t-1|t-1}\mathbb{F}' + \mathbb{Q}',$$

$$\mathbb{Y}_{t|t-1} \;=\; \mathbb{H}'\mathbb{S}_{t|t-1},$$

$$\Omega_{t|t-1} \;=\; \mathbf{E}\Big[(\mathbb{Y}_t - \mathbb{Y}_{t|t-1})\,(\mathbb{Y}_t - \mathbb{Y}_{t|t-1})'\Big] \;=\; \mathbb{H}'P_{t|t-1}\mathbb{H} + \Sigma_u,$$

the likelihood,

$$\mathcal{L}\Big(\mathbb{Y}_t\Big|\mathcal{Y}_{t-1},\Theta\Big) \;=\; (2\pi)^{-m/2}\left|\Omega_{t|t-1}^{-1}\right|^{1/2}\exp\left[-\frac{1}{2}\Big(\mathbb{Y}_t - \mathbb{Y}_{t|t-1}\Big)'\Omega_{t|t-1}^{-1}\Big(\mathbb{Y}_t - \mathbb{Y}_{t|t-1}\Big)\right]$$

and the updates of the state vector and its mean square error matrix

$$\mathbb{S}_{t|t} \;=\; \mathbb{S}_{t|t-1} + P_{t|t-1}\mathbb{H}\Omega_{t|t-1}^{-1}\Big(\mathbb{Y}_t - \mathbb{Y}_{t|t-1}\Big),$$
$$P_{t|t} \;=\; P_{t|t-1} - P_{t|t-1}\mathbb{H}\Omega_{t|t-1}^{-1}\mathbb{H}'P_{t|t-1}.$$

for $t = 2, \ldots, T$.

The likelihoods, $\mathcal{L}(\mathbb{Y}_1|\Theta)$, $\mathcal{L}(\mathbb{Y}_2|\mathcal{Y}_1,\Theta)$, $\mathcal{L}(\mathbb{Y}_3|\mathcal{Y}_2,\Theta), \ldots, \mathcal{L}(\mathbb{Y}_{T-1}|\mathcal{Y}_{T-2},\Theta)$, and $\mathcal{L}(\mathbb{Y}_T|\mathcal{Y}_{T-1},\Theta)$, computed at Steps 2 and 5 are used to build up the likelihood function (12) of the linearized NKDSGE model.

[24] Let $\Sigma_{\mathbb{S}}$ be the unconditional covariance matrix of $\mathbb{S}$. The state equations (11.1) imply $\Sigma_{\mathbb{S}} = \mathbb{F}\Sigma_{\mathbb{S}}\mathbb{F}' + \mathcal{Q}'$. Its solution is $\text{vec}(\Sigma_{\mathbb{S}}) = [\mathbf{I}_n - \mathbb{F}\otimes\mathbb{F}]^{-1}\,\text{vec}(\mathcal{Q}')$, where $\text{vec}(\mathbb{A}\mathbb{B}\mathbb{C}) = (\mathbb{C}'\otimes\mathbb{A})\,\text{vec}(\mathbb{B})$, which in turn sets $P_{0|0} = \text{vec}(\Sigma_{\mathbb{S}})$.

5.2 Priors

Our priors are borrowed from Del Negro and Schorfheide (2008). They construct priors by separating the NKDSGE model parameters into three sets. Their first set consists of those parameters that define the steady state of the NKDSGE model; see table 2 of Del Negro and Schorfheide (2008, p. 1201). The steady state, which as Hall (1996) shows ties the steady state of the NKDSGE model to the unconditional first moments of $\mathcal{Y}_T$, has no effect on the mechanism that endogenously propagates exogenous shocks. This mechanism relies on preferences, technologies, and market structure. The parameters of these primitives of the NKDSGE model are included in the second set of priors. Along with technology, preference, and market structure parameters, Del Negro and Schorfheide add parameters of the Taylor rule (8) to this set; see the agnostic sticky price and wage priors of tables 1 and 2 of Del Negro and Schorfheide (2008, pp. 1200–1201). The third set of parameters consist of AR1 coefficients and standard deviations of the exogenous shocks; see table 3 of Del Negro and Schorfheide (2008, p. 1201).

We divide the parameter vector Θ into two parts to start. The 25×1 column vector

$$\Theta_1 = \Big[\zeta_p \; \pi^* \; \iota_p \; h \; \nu_l \; a'' \; \Gamma'' \; \lambda_W \; \zeta_W \; \iota_W \; R^* \; \rho_R \; \psi_1 \; \psi_2 \; \gamma \; \lambda_f \; \rho_z \; \rho_\phi \; \rho_{\lambda_f} \; \rho_g \; \sigma_z \; \sigma_\phi \; \sigma_{\lambda_f} \; \sigma_g \; \sigma_R\Big]',$$

contains the parameters of economic interest, which are to be estimated, in the order in which they appear in section 3. Under the Del Negro and Schorfheide (2008) prior rubric, the elements of Θ_1 are grouped into the steady state parameter vector

$$\Theta_{1,ss} = \Big[\pi^* \;\; \gamma \;\; \lambda_f \;\; \lambda_W \;\; R^*\Big]',$$

the parameters tied to endogenous propagation in the NKDSGE model

$$\Theta_{1,prop} = \Big[\zeta_p \;\; \iota_p \;\; h \;\; \nu_l \;\; a'' \;\; \Gamma'' \;\; \zeta_W \;\; \iota_W \;\; \rho_R \;\; \psi_1 \;\; \psi_2\Big]',$$

and

$$\Theta_{1,exog} = \Big[\rho_z \;\; \rho_\phi \;\; \rho_{\lambda_f} \;\; \rho_g \;\; \sigma_z \;\; \sigma_\phi \;\; \sigma_{\lambda_f} \;\; \sigma_g \;\; \sigma_R\Big]'.$$

contains the slope coefficients and standard deviations of the exogenous AR(1) shocks that are the source of fluctuations in the NKDSGE model.

Table 1 lists priors for $\Theta_{1,ss}$, $\Theta_{1,prop}$, and $\Theta_{1,exog}$. We draw priors for Θ_1 from normal, beta, gamma, and inverse gamma distributions; see Del Negro and Schorfheide (2008) for details. The priors are summarized by the distribution from which we draw, the parameters of the distribution, and implied 95 percent probability intervals.

Our choices reflect, in part, a desire to elicit priors on Θ_1 that are easy to understand. For example, π^* is endowed with a normally distributed prior. Its mean is 4.3 percent, which is less than twice its standard deviation giving a 95 percent probability interval running from nearly -1 percent to more than 9 percent. Thus, the prior reveals the extent of the uncertainty that surrounds steady state inflation.

The beta distribution is useful because it restricts priors on NKDSGE model parameters to the open unit interval. This motivates drawing the sticky price and wage parameter, ζ_p, ι_p, ζ_W, and ι_W, the consumption habit parameters, h, and the AR1 parameters, ρ_R, ρ_z, ρ_ϕ, ρ_{λ_f}, and ρ_g, from the beta distribution. The means and standard deviations of the priors display our uncertainty about these NKDSGE model parameters. For example, the prior on h indicates less uncertainty about it than is placed on the priors for ζ_p, ι_p, ζ_W, and ι_W (*i.e.*, the ratio of the mean to the standard deviation of the priors of these parameters is less than three, while the same ratio for the prior of h is 14). This gives larger intervals on which to draw the sticky price and wage parameters than on h. Also, the prior 95 percent probability interval of h is in the range that Kano and Nason (2010) show to be relevant for consumption habit to generate business cycle fluctuations in similar NKDSGE models.

The AR1 coefficients also rely on the beta distribution for priors. The prior on ρ_R suggests a 95 percent probability interval of draws that range from 0.22 to 0.73. At the upper end of this range, the Taylor rule is smoothing the policy rate R_t. This interval has the same length but is shifted to the left for ρ_z, which endows the technology growth prior with less persistence. The taste, monopoly power, and government spending shocks exhibit more persistence with AR1 coefficients priors lying between 0.5 and 0.95.

The gamma distribution is applied to NKDSGE model parameters that only require priors that rule out non-negative draws or impose a lower bound. The former restriction describes the use of the gamma distribution for priors on the goods and labor market monopoly power parameters, λ_f and λ_W, the capital utilization parameter, a'', and the Taylor rule parameter on output, ψ_2. A lower bound is placed on the prior of the deterministic growth of technology, γ, the mean policy rate R^*, the labor supply parameter, ν_l, the investment cost parameter, Γ'', and the Taylor rule parameter on inflation, ψ_1. The prior on ψ_1 is set to obey the Taylor principle that R_t rises by more than the increase in π_t net of π^*. This contrasts with the prior on ψ_2 that suggests a smaller response of R_t to the output gap, $Y_t - Y_t^{\tau}$, but this response is non-zero.

The priors on the standard deviations of the exogenous shocks are drawn from inverse-gamma distributions. This distribution has support on an open interval that excludes zero and is unbounded. This allows σ_z, σ_{λ_f}, σ_g, and σ_R to have priors with 95 percent probability intervals with lower bounds near zero and large upper bounds. These priors show the uncertainty held about these elements of the exogenous shock processes of the NKDSGE model. The same is true for the prior on σ_ϕ, but its scale parameter has a 95 percent probability interval that exhibits more uncertainty as it is shifted to the right especially for the upper bound.

The remaining parameters are necessary to solve the linearized NKDSGE model but are problematic for estimation. The fixed or calibrated parameters are collected into

$$\Theta_2 = \begin{bmatrix} \alpha & \delta & g^* & \mathcal{L}_{\mathcal{A}} & \kappa \end{bmatrix}'.$$

The calibration of Θ_2 results in

$$\begin{bmatrix} \alpha & \delta & g^* & \mathcal{L}_{\mathcal{A}} & \kappa \end{bmatrix}' = \begin{bmatrix} 0.33 & 0.025 & 0.22 & 1.0 & 0.0 \end{bmatrix}'.$$

Although these values are standard choices in the DSGE literature, some clarification is in order. As in Del Negro and Schorfheide (2008), our parametrization imposes the constraint that firms make zero profits in the steady state. We also assume that households work one unit of time in steady state. This assumption implies that the parameter ϕ, the mean of the taste shock ϕ_t, is endogenously determined by the optimality conditions in the model. This restriction on steady state hours worked in the NKDSGE model differs from the sample mean of hours worked. We deal with this mismatch by augmenting the measurement equation in the state space representation with a constant or "add-factor" that forces the theoretical mean of hours worked to match the sample mean; see Del Negro and Schorfheide (2008, p. 1197). This amounts to adding $\mathcal{L}_{\mathcal{A}}$ to the log likelihood of the linearized NKDSGE model

$$\ln \mathcal{L}\left(\mathcal{Y}_T \middle| \Theta_1; \Theta_2\right) + \ln \mathcal{L}_{\mathcal{A}}.$$

Also, rather than imposing priors on the great ratios, C^*/Y^*, $I^*/\overline{K}^*$, $\overline{K}^*/Y^*$, and G^*/Y^*, we fix the capital share, α, the depreciation rate, δ, and the share of government expenditure, g^*. This follows well established practices that pre-date Bayesian estimation of NKDSGE models.

5.3 Useful Information about the MH-MCMC Simulator

The posterior distribution of the NKDSGE model parameters in Θ_1 is characterized using the MH-MCMC algorithm. The MH-MCMC algorithm is started up with an initial Θ_1. This parameter vector is passed to the Kalman filter routines described in section 5.1 to obtain an estimate of $\mathcal{L}(\mathcal{Y}_T|\Theta_1;\Theta_2)$. Next, the initial Θ_1 is updated according to the MH random walk law of motion. Inputing the proposed update of Θ_1 into the Kalman filter produces a second estimate of the likelihood of the linear approximate NKDSGE model. The MH decision rule determines whether the initial or proposed update of Θ_1 and the associated likelihood is carried forward to the next step of the MH algorithm. Given this choice, the next step of the MH algorithm is to obtain a new proposed update of Θ_1 using the random walk law of motion and to generate an estimate of the likelihood at these estimates. This likelihood is compared to the likelihood carried over from the previous MH step using the MH decision rule to select the likelihood and Θ_1 for the next MH step. This process is repeated $\mathcal{H}$ times to generate the posterior of the linear approximate NKDSGE model, $\mathcal{P}(\Theta_1|\mathcal{Y}_T;\Theta_2)$.

We summarize this description of the MH-MCMC algorithm with

1. Label the vector of NKDSGE model parameters chosen to initialize the MH algorithm $\widehat{\Theta}_{1,0}$.

2. Pass $\widehat{\Theta}_{1,0}$ to the Kalman filter routines described in section 5.2 to generate an initial estimate of the likelihood of the linear approximate NKDSGE model, $\mathcal{L}\left(\mathcal{Y}_T \middle| \widehat{\Theta}_{1,0}; \Theta_2\right)$.

3. A proposed update of $\widehat{\Theta}_{1,0}$ is $\Theta_{1,1}$ which is generated using the MH random walk law of motion, $\Theta_{1,1} = \widehat{\Theta}_{1,0} + \varpi\vartheta\varepsilon_1,\ \varepsilon_1 \sim \mathcal{NID}\left(\mathbf{0}_d, \mathbf{I}_d\right)$, where ϖ is a scalar that controls the size of the "jump" of the proposed MH random walk update, ϑ is the Cholesky decomposition of the covariance matrix of Θ_1, and d (= 25) is the dimension of Θ_1. Obtain $\mathcal{L}\left(\mathcal{Y}_T \middle| \Theta_{1,1}; \Theta_2\right)$ by running the Kalman filter using $\Theta_{1,1}$ as input.

4. The MH algorithm employs a two-stage procedure to decide whether to keep the initial $\widehat{\Theta}_{1,0}$ or move to the updated proposal $\Theta_{1,1}$. First, calculate

$$\omega_1 \;=\; \min\left\{ \frac{\mathcal{L}\left(\mathcal{Y}_T \,\middle|\, \Theta_{1,1};\, \Theta_2\right) \mathcal{P}\left(\Theta_{1,1}\right)}{\mathcal{L}\left(\mathcal{Y}_T \,\middle|\, \widehat{\Theta}_{1,0};\, \Theta_2\right) \mathcal{P}\left(\widehat{\Theta}_{1,0}\right)},\, 1 \right\},$$

   where, for example, $\mathcal{P}\left(\Theta_{1,1}\right)$ is the prior at $\Theta_{1,1}$. The second stage begins by drawing a uniform random variable $\varphi_1 \sim U(0,1)$ to set $\widehat{\Theta}_{1,1} = \Theta_{1,1}$ and the counter $\wp = 1$ if $\varphi_1 \leq \omega_1$, otherwise $\widehat{\Theta}_{1,1} = \widehat{\Theta}_{1,0}$ and $\wp = 0$.

5. Repeat steps 3 and 4 for $\ell = 2, 3, \ldots, \mathcal{H}$ using the MH random walk law of motion

$$\Theta_{1,\ell} \;=\; \widehat{\Theta}_{1,\ell-1} + \varpi \vartheta \varepsilon_\ell, \quad \varepsilon_\ell \sim \mathcal{NID}\left(\mathbf{0}_{d\times 1}, \mathbf{I}_d\right), \tag{13}$$

   and drawing the uniform random variable $\varphi_\ell \sim U(0,1)$ to test against

$$\omega_\ell \;=\; \min\left\{ \frac{\mathcal{L}\left(\mathcal{Y}_T \,\middle|\, \Theta_{1,\ell};\, \Theta_2\right) \mathcal{P}\left(\Theta_{1,\ell}\right)}{\mathcal{L}\left(\mathcal{Y}_T \,\middle|\, \widehat{\Theta}_{1,\ell-1};\, \Theta_2\right) \mathcal{P}\left(\widehat{\Theta}_{1,\ell-1}\right)},\, 1 \right\},$$

   for equating $\widehat{\Theta}_{1,\ell}$ to either $\Theta_{1,\ell}$ or $\widehat{\Theta}_{1,\ell-1}$. The latter implies that the counter is updated according to $\wp = \wp + 0$, while the former has $\wp = \wp + 1$.

Steps 1–5 of the MH-MCMC algorithm produce the posterior, $\mathcal{P}(\widehat{\Theta}_1 | \mathcal{Y}_T;\, \Theta_2)$, of the linear approximate NKDSGE model by drawing from $\left\{\widehat{\Theta}_{1,\ell}\right\}_{\ell=1}^{\mathcal{H}}$. Note that in Steps 4 and 5 the decision to accept the updated proposal, $\varphi_\ell \leq \omega_\ell$, is akin to moving to a higher point on the likelihood surface.

There are several more issues that have to be resolved to run the MH-MCMC algorithm to create $\mathcal{P}(\widehat{\Theta}_1 | \mathcal{Y}_T;\, \Theta_2)$. Among these are obtaining an $\widehat{\Theta}_{1,0}$ to initialize the MH-MCMC, computing ϑ, determining $\mathcal{H}$, fixing ϖ to achieve the optimal acceptance rate for the proposal $\Theta_{1,\ell}$ of $\wp/\mathcal{H}$, and checking that the MH-MCMC simulator has converged.[25]

Step 1 of the MH-MCMC algorithm leaves open the procedure for setting $\widehat{\Theta}_{1,0}$. We employ classical optimization methods and an MH-MCMC "burn-in" stage to obtain $\widehat{\Theta}_{1,0}$. First, a classical optimizer is applied repeatedly to the likelihood of the linear approximate NKDSGE model with initial conditions found by sampling 100 times from $\mathcal{P}(\Theta_1)$.[26] These estimates yield the mode of the posterior distribution of Θ_1 that we identify as initial conditions for a "burn-in" stage of the MH-MCMC algorithm. The point of this burn-in of the MH-MCMC algorithm is to

[25] Gelman et al (2004, pp 305–307) discuss rules for the MH-MCMC simulator that improve the efficiency of the law of motion (13) to give acceptance rates that are optimal.

[26] Chris Sims is responsible for the optimizer software that we use. The optimizer is csminwel and available at http://sims.princeton.edu/yftp/optimize/.

remove dependence of $\mathcal{P}(\widehat{\Theta}_1|\mathcal{Y}_T;\Theta_2)$ on the initial condition $\widehat{\Theta}_{1,0}$. Drawing $\widehat{\Theta}_{1,0}$ from a distribution that resembles $\mathcal{P}(\widehat{\Theta}_1|\mathcal{Y}_T;\Theta_2)$ eliminates this dependence. Next, 10,000 MH steps are run with $\varpi = 1$ and $\vartheta = \mathbf{I}_d$ to complete the burn-in stage. The final MH step of the burn-in gives $\widehat{\Theta}_{1,0}$ to initialize the $\mathcal{H}$ steps of the final stage of the MH-MCMC algorithm. The 10,000 estimates of Θ_1 generated during the MH burn-in steps are used to construct an empirical estimate of the covariance matrix $\vartheta\vartheta'$. The Cholesky decomposition of this covariance matrix is the source of ϑ needed for the MH law of motion (13).

The scale of the "jump" from $\Theta_{1,\ell}$ to $\widehat{\Theta}_{1,\ell-1}$ determines the speed at which the proposals $\Theta_{1,\ell}$ converge to $\mathcal{P}(\widehat{\Theta}_1|\mathcal{Y}_T;\Theta_2)$ within the MH-MCMC simulator. The speed of convergence is sensitive to ϖ as well as to $\mathcal{H}$. The number of steps of the final stage of the MH-MCMC simulator has to be sufficient to allow for convergence. We obtain $\mathcal{H} = 300{,}000$ draws from the posterior $\mathcal{P}(\widehat{\Theta}_1|\mathcal{Y}_T;\Theta_2)$, but note that for larger and richer NKDSGE models the total number of draws is often many times larger. Nonetheless, the choice of the scalar ϖ is key for controlling the speed of convergence of the MH-MCMC. Although Gelman et al (2004) recommend that greatest efficiency of the MH law of motion (13) is found with $\varpi = 2.4/\sqrt{d}$, we set ϖ to drive the acceptance rate $\wp/\mathcal{H} \in [0.23, 0.30]$.[27]

It is standard practice to test to check the convergence of the MH-MCMC simulator, besides requiring $\wp/\mathcal{H}$ to 0.23. Information about convergence of the MH-MCMC simulator is provided by the $\widehat{R}$ statistic of Gelman et al. (2004, pp. 269–297). This statistic compares the variances of the elements within the sequence of $\left\{\widehat{\Theta}_{1,\ell}\right\}_{\ell=1}^{\mathcal{M}}$ to the variance across several sequences produced by the MH-MCMC simulator given different initial conditions. These different initial conditions are produced using the same methods already described with one exception. The initial condition for the burn-in stage of the MH-MCMC algorithm is typically set at the next largest mode of the posterior distribution obtained by applying the classical optimizer to the likelihood of the linear approximate NKSDSGE model. This process is often repeated three to five times. Gelman et al. (2004) suggest that $\widehat{R} < 1.1$ for each element of $\widehat{\Theta}_1$. If not, across the posteriors of the MH-MCMC chains there is excessive variation relative to the variance within the sequences. When $\widehat{R}$ is large, Gelman et al propose increasing $\mathcal{H}$ until convergence is achieved as witnessed by $\widehat{R} < 1.1$.[28]

6 Results

This section describes the data and reports the results of estimating the linear approximate NKDSGE model using the Bayesian procedures of the previous section.

6.1 Data

We follow Del Negro and Schorfheide (2008) in estimating the NKDSGE model given five aggregate U.S. variables. The observables are per capita output growth, per capita hours worked,

[27] This involves an iterative process of running the MH-MCMC simulator to calibrate ϖ to reach the desired acceptance rate.

[28] Geweke (2005) advocates a convergence test examining the serial correlation within the sequence of each element of $\widehat{\Theta}_{1,\ell}$, $\ell = 1, \ldots, \mathcal{H}$.

labor share, inflation, and the nominal interest rate on the $1982Q1$–$2009Q4$ sample. Thus, Bayesian estimates of the NKDSGE model parameters are conditional on the information set

$$\mathbb{Y}_t = \left[400\Delta \ln Y_t \;\; 100 \ln L_t \;\; 100 \ln \frac{W_t L_t}{P_t Y_t} \;\; 400\pi_t \;\; 400 \ln R_t\right]',$$

where Δ is the first difference operator. Per capita output growth, labor share, inflation, and the nominal interest rate are multiplied to obtain data that are annualized, which is consistent with the measurement of per capita hours worked, and in percentages. Real GDP is divided by population (16 years and older) to create per capita output. Hours worked is a series constructed by Del Negro and Schorfheide (2008) that we extend several more quarters. They interpolate annual observations on aggregate hours worked in the U.S. into the quarterly frequency using the growth rate of an index of hours of all persons in the nonfarm business sector. Labor share equals the ratio of total compensation of employees to nominal GDP. Inflation is equated to the (chained) GDP deflator. The effective federal funds rate defines the nominal interest rate.[29]

6.2 Posterior Estimates

Table 2 contains summary statistics of the posterior distributions of two NKDSGE models. We include posterior medians, modes, and 95 percent probability intervals of the NKDSGE model parameters in table 2. Estimates of the NKDSGE model labeled $\mathcal{M}_1$ are grounded in the priors that appear in table 1 and discussed in section 5.2. We also estimate an NKDSGE model that fixes ι_p at zero, which defines the weights on π^* and π_{t-1} in the indexation rule used by firms unable to update their prices at any date t. This NKDSGE model is labeled $\mathcal{M}_2$. The motivation for estimating $\mathcal{M}_2$ is that table 6 of Del Negro and Schorfheide (2008, p. 1206) has 90 percent probability intervals for ι_p with a lower bound of zero for all but one of their priors.

We obtain similar estimates for $\Theta_{1,prop}$ across $\mathcal{M}_1$ and $\mathcal{M}_2$ as listed in the middle panel of table 2, except for ι_p. The posterior distributions of these models indicate substantial consumption habit, $h \in (0.73, 0.87)$, a large Frisch labor supply elasticity, $\nu_l^{-1} \in (0.56, 1.39)$, costly capital utilization, $a'' \in (0.11, 0.46)$, investment costs of adjustments, $\Gamma'' \in (6.9, 14.2)$, sticky prices, $\zeta_p \in (0.58, 0.74)$, nominal wage indexation, $\iota_W \in (0.22, 0.82)$, and interest rate smoothing by a monetary authority, $\rho_R \in (0.74, 0.82)$, that satisfies the Taylor principle, $\psi_1 \in (2.14, 2.90)$. These estimates show which elements of the NKDSGE models interact endogenously to replicate fluctuations found $\mathcal{Y}_T$. These estimates are also in the range often found in the existing literature; for example, see Negro and Schorfheide (2008).

Sticky nominal wages, price indexation, and the monetary authority's response to deviations of output from its target appear to matter less for generating endogenous propagation in the NKDSGE models. The 95 percent probability interval of ι_p has a lower bound of 0.006 in the posterior distribution of $\mathcal{M}_1$. For $\mathcal{M}_1$ and $\mathcal{M}_2$, the estimates of ζ_W and ψ_2 are also relatively

[29] The data are available at http://research.stlouisfed.org/fred2/. This website, which is maintained by the Federal Reserve Bank of St. Louis, contains data produced by the Bureau of Economic Analysis (BEA), the Bureau of Labor Statistics (BLS), and the Board of Governors of the Federal Reserve System (BofG). The BEA compiles real GDP, annual aggregate hours worked, total compensation of employees, nominal GDP, and the chained GDP deflator. The BLS provides the population series and the index of hours of all persons in the nonfarm business sector. The effective federal funds rate is collected by the BofG.

small. Thus, sticky prices and nominal wage indexation, not sticky nominal wages and price indexation, matter for endogenous propagation in $\mathcal{M}_1$ and $\mathcal{M}_2$ given $\mathcal{Y}_T$ and our priors.

Estimates of $\Theta_{1,exog}$ show that exogenous propagation matters for creating fluctuations in the posterior distributions of $\mathcal{M}_1$ and $\mathcal{M}_2$. The bottom panel of table 2 shows that the taste shock ϕ_t, the goods market monopoly power shock λ_f, and the government spending shock g_t are persistent. In $\mathcal{M}_1$ and $\mathcal{M}_2$, the half-life of a structural innovation to these shocks are about seven quarters for λ_f and 11 quarters for ϕ_t and g_t at the medians and modes of ρ_{λ_f}, ρ_ϕ, and ρ_g, respectively.[30] The NKDSGE $\mathcal{M}_1$ and $\mathcal{M}_2$ yield estimates of ρ_z that signal much less persistence. Estimates of $\rho_z \in (0.23, 0.47)$ surround estimates of the unconditional first-order autocorrelation coefficient of U.S. output growth; see Cogley and Nason (1995). Further, $\mathcal{M}_1$ and $\mathcal{M}_2$ produce posterior distributions in which the lower end of the 95 percent probability intervals of ρ_z suggests little or no persistence in z_t.

Exogenous shock volatility contributes to $\mathcal{M}_1$ and $\mathcal{M}_2$ replicating variation in $\mathcal{Y}_T$. The scale parameters σ_ϕ and σ_{λ_f} matter most for this aspect of the fit of the NKDSGE models. Estimates of these elements of $\Theta_{1,exog}$ are 2.5 to more than 9 times larger than estimates of σ_z and σ_g. When σ_R is included in this comparison, it reveals that exogenous variation in monetary policy matters less for $\mathcal{M}_1$ and $\mathcal{M}_2$ to explain variation in $\mathcal{Y}_T$. Thus, $\mathcal{M}_1$ and $\mathcal{M}_2$ attribute the sources of business cycle fluctuations more to taste and goods market monopoly power shocks than to TFP growth, government spending, or monetary policy shocks.

The top panel of table 2 displays estimates of $\Theta_{1,ss}$ that are nearly identical across $\mathcal{M}_1$ and $\mathcal{M}_2$. These estimates indicate that the posterior distributions of these NKDSGE models place a 95 percent probability that steady state inflation in the U.S. was as low as 2.1 percent and just a little more than 3.5 percent.[31] There is greater precision in the posterior estimates of γ. Deterministic TFP growth is estimated to range from 2.85 to 3 percent per annum with a 95 percent probability interval according to $\mathcal{M}_1$ and $\mathcal{M}_2$. In contrast, the 95 percent probability intervals of R^* are shifted slightly to the left of the ones shown for π^*. These estimates suggest the NKDSGE models $\mathcal{M}_1$ and $\mathcal{M}_2$ predict steady state real interest rates near zero.

Del Negro and Schorfheide (2008) report estimates of price and wage stickiness that often differ from those of $\mathcal{M}_1$ and $\mathcal{M}_2$. For example, the middle panel of table 2 shows that the median degree of price stickiness yields a frequency (*i.e.*, $1/[1 - \zeta_p]$) at which the firms of $\mathcal{M}_1$ and $\mathcal{M}_2$ change prices about once every two to three quarters. Del Negro and Schorfheide obtain estimates of ζ_p that imply an almost identical frequency of price changes for only three of the six priors they use. Notably, when they adopt priors with greater price stickiness, posterior estimates have firms changing prices as infrequently as once every 10 quarters on average.

Nominal wages exhibit less rigidity in the posterior distributions of $\mathcal{M}_1$ and $\mathcal{M}_2$. The 95 percent probability intervals of ζ_W range from 0.07 to 0.27. This indicates that the households of $\mathcal{M}_1$ and $\mathcal{M}_2$ change their nominal wages no more than every other quarter. However, at the posterior median and modes of ι_W, those households unable to optimally adjust their nominal wages depend in about equal parts on π^*, γ^*, π_{t-1}, and z_{t-1} when updating to $W_{i,t-1}$. In comparison, the posterior of $\mathcal{M}_1$ shows that firms unable to reset their prices optimally rely almost entirely on π_{t-1} and not π^* when updating because the 95 percent probability interval of $\iota_p \in (0.0, 0.2)$. The lower end of this interval is near the restriction imposed on ι_p by $\mathcal{M}_2$.

[30]The half-life estimates are computed as $\ln 0.5/\ln \rho_s$, $s = \phi$, λ_f, and g.

[31]The posterior probability interval differs from a frequentist confidence band. The latter holds the relevant parameter fixed and depends only on data, while the former is conditional on the model, priors, and data.

The marginal likelihoods of $\mathcal{M}_1$ and $\mathcal{M}_2$ give evidence about which NKDSGE model is preferred by $\mathcal{Y}_T$. The Bayes factor (2) is employed to gauge the relative merits of $\mathcal{M}_1$ and $\mathcal{M}_2$. We adopt methods described in Geweke (1999, 2005) to integrate or marginalize $\widehat{\Theta}_1$ out of $\mathcal{L}(\mathcal{Y}_T|\widehat{\Theta}_1, \mathcal{M}_j; \Theta_2)$, $j = 1, 2$; also see and Chib and Jeliazkov (2001).[32] The top of table 2 lists the log marginal likelihoods of $\mathcal{M}_1$ and $\mathcal{M}_2$. The Bayes factor of the marginal likelihoods of $\mathcal{M}_1$ and $\mathcal{M}_2$ is 2.23. According to Jeffreys (1998), a Bayes factor of this size shows that $\mathcal{Y}_T$'s preference for $\mathcal{M}_2$ over $\mathcal{M}_1$ is "barely worth mentioning."[33] Thus, the marginal likelihoods of $\mathcal{M}_1$ and $\mathcal{M}_2$ provide evidence that, although $\mathcal{Y}_T$ support $\iota_p = 0$, the evidence in favor of this restriction is not sufficient for an econometrician with the priors displayed in table 1 to ignore $\mathcal{M}_1$, say, for conducting policy analysis.

7 Conclusion

This article surveys Bayesian methods for estimating NKDSGE models with the goal of raising the use of these empirical tools. We give an outline of an NKDSGE model to develop intuition about the mechanisms it has to transmit exogenous shocks into endogenous business cycle fluctuations. Studying the sources and causes of these propagation mechanisms requires us to review the operations needed to detrend its optimality and equilibrium conditions, a technique to construct a linear approximation of the model, a strategy to solve for its linear approximate decision rules, and the mapping from this solution into a state space model that can produce Kalman filter projections and the likelihood of the linear approximate NKDSGE model. The projections and likelihood are useful inputs into the MH-MCMC simulator. Since the source of Bayesian estimates of the NKDSGE model is the MH-MCMC simulator, we present an algorithm that implements it. This algorithm relies on our priors of the NKDSGE model parameters and setting initial conditions for the simulator. We employ the simulator to generate posterior distributions of two NKDSGE models. These posterior distributions yield summary statistics of the Bayesian estimates of the NKDSGE model parameters that are compared to results in the extant literature. These posterior distributions are needed as well to address the question of which NKDSGE models is most favored by the data. We also provide a short history of DSGE model estimation as well as pointing to issues that are at the frontier of this research.

We describe Bayesian methods in this article that are valuable because DSGE models are useful tools for understanding the sources and causes of business cycles and for conducting policy evaluation. This article supplies empirical exercises in which NKDSGE models are estimated and evaluated using data and priors that are standard in the published literature. Thus, it is no surprise that our estimates of the NKDSGE models resemble estimates found in the published literature. Although comforting, the similarity in estimates raises questions about whether the data are truly informative about the NKDSGE models or if posterior distributions of the NKDSGE models are dominated by our priors. Also, little is known about the impact of misspecification on the relationship between data, priors, and posterior distributions of NKDSGE models. We hope this article acts as a foundation supporting future research on these issues.

[32] Geweke advises computing the marginal likelihood with a harmonic mean estimator along with several refinements that he proposes. Useful instructions for computing marginal likelihoods along these lines are provided by Fernández-Villaverde and Rubio-Ramírez (2004, pp. 169–170).

[33] The Bayes factor needs to exceed odds of 3 to 1 before there is "substantial" evidence against $\mathcal{M}_2$.

References

Adolfson, M., S. Laséen, J. Lindé, and M. Villani, (2007), 'Bayesian estimation of an open economy DSGE model with incomplete pass-through,' *Journal of International Economics*, *72*(2), 481–511.

Altuğ, S. (1989), 'Time-to-Build and Aggregate Fluctuations: Some New Evidence,' *International Economic Review*, **30**(4), 889–920.

An, S., F. Schorfheide (2007), 'Bayesian Analysis of DSGE Models,' *Econometric Reviews*, **26**(2–4), 113–172.

Anderson, B., J. Moore (2005), *Optimal Filtering*, Dover Publications.

Aruoba, B., L. Bocola, F. Schorfheide (2011), 'A New Class of Nonlinear Time Series Models for the Evaluation of DSGE Models,' manuscript, Department of Economics, University of Pennslyvania.

Aruoba, B., F. Schorfheide (2011), 'Sticky prices versus monetary frictions: An estimation of policy trade-offs,' *American Economic Journal: Macroeconomics*, **3**(1), 60–90.

Bencivenga, V.R. (1992), 'An Econometric Study of Hours and Output Variation with Preference Shocks,' *International Economic Review*, **33**(2), 449–471.

Berger, J.O., R.L. Wolpert (1988), *The Likelihood Principle, Second Edition*, Hayward, CA: Institute of Mathematical Statistics.

Canova, F. (2007), *Methods for Applied Macroeconomic Research*, Princeton, NJ: Princeton University Press.

Canova, F., L. Sala (2009), 'Back to Square One: Identification Issues in DSGE Models,' *Journal of Monetary Economics*, **56**(4), 431–449.

Chernozhukov, V., H. Hong (2003), 'An MCMC Approach to Classical Estimation,' *Journal of Econometrics*, *115*(2), 293–346.

Chib, S., I. Jeliazkov (2001), 'Marginal Likelihood from the Metropolis-Hastings Output,' *Journal of the American Statistical Association*, **96**, 270–281.

Christiano, L., M. Eichenbaum (1992), 'Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations,' *Ameican Economic Review*, *82*(3), 430–450.

Christiano, L., M. Eichenbaum, C. Evans (2005), 'Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,' *Journal of Political Economy*, *113*(1), 1–45.

Cogley, T., J.M. Nason (1995), 'Output Dynamics in Real-Business-Cycle Models,' *American Economic Review*, **85**(3), 492–511.

Cogley, T., T.J. Sargent (2005), 'Drifts and volatilities: Monetary policies and outcomes in the post WWII US,' *Review of Economic Dynamics*, **8**(2), 262–302.

Curdia, V., R. Reis (2011), Correlated Disturbances and U.S. Business Cycles,' manuscript, Department of Economics, Columbia University.

DeJong, D.N., C. Dave (2007), *Structural Macroeconometrics*, Princeton, NJ: Princeton University Press.

DeJong, D.N., B.F. Ingram, C.H. Whiteman (2000a), 'A Bayesian approach to dynamic macroeconomics,' *Journal of Econometrics*, **98**(2), 203–223.

DeJong, D.N., B.F. Ingram, C.H. Whiteman (2000b), 'Keynesian impulses versus solow residuals: Identifying sources of business cycle fluctuations,' *Journal of Applied Econometrics*, **15**(3), 311–329.

Del Negro, M., F. Schorfheide (2004), 'Priors from general equilibrium models for VARs,' *International Economic Review*, **45**(2), 643–673.

Del Negro, M., F. Schorfheide (2008), 'Forming Priors for DSGE models (and How It Affects the Assessment of Nominal Rigidities),' *Journal of Monetary Economics*, **55** (7), 1191–1208.

Del Negro, M., F. Schorfheide, F. Smets, R. Wouters (2007), 'On the Fit and Forecasting Performance of New Keynesian Models,' *Journal of Business and Economic Statistics*, **25** (2), 123–162.

Doan, T., R. Litterman, C.A. Sims, (1984). "Forecasting and conditional projections using a realistic prior distribution," *Econometric Reviews* , **3**(1), 1–100.

Dridi, R., A. Guay, E. Renault (2007), 'Indirect inference and calibration of dynamic stochastic general equilibrium models,' *Journal of Econometrics*, **136**(2), 397–430.

Erceg, C., D. Henderson, A. Levin (2000), 'Optimal Monetary Policy with Staggered Wage and Price Contracts,' *Journal of Monetary Economics*, **46**(2), 281–313.

Fernández-Villaverde, J., P.A. Guerrón-Quintana, J.F. Rubio-Ramírez (2009), 'The New Macroeconometrics: A Bayesian Approach,' *Handbook of Applied Bayesian Analysis*.

Fernández-Villaverde, J., P.A. Guerrón-Quintana, J.F. Rubio-Ramírez (2010), 'Fortune or Virture: Time-Variant Volatilities versus Parameter Drifting,' Federal Reserve Bank of Philadelphia working paper 10–14.

Fernández-Villaverde, J., P.A. Guerrón-Quintana, K. Kuester, and J.F. Rubio-Ramírez (2011), 'Fiscal Volatility Shocks and Economic Activity,' Federal Reserve Bank of Philadelphia working paper 11–32.

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

Fernández-Villaverde, J., J.F. Rubio-Ramírez (2007), 'Estimating Macroeconomic Models: A Likelihood Approach,' *Review of Economic Studies*, **74**(4), 1059–1087.

Gelman, Andrew, John B. Carlin, Hal S. Stern, Donald B. Rubin (2004), *Bayesian Data Analysis*, second edition, Chapman and Hall/CRC: Boca Raton, FL.

Geweke, John (1999) 'Simulation methods for model criticism and robustness analysis,' in James O. Berger, José M. Bernado, A. Philip Dawid, Adrian F.M. Smith (eds.), *Bayesian Statistics, Vol. 6*, Oxford, UK: Oxford University Press, pp. 275—299.

Geweke, John (2005), *Contemporary Bayesian Econometrics and Statistics*, Hoboken, NJ: John Wiley & Sons, Inc.

Geweke, John (2010), *Complete and Incomplete Econometric Models*, Princeton, NJ: Princeton University Press.

Gourieroux, C, A. Monfort, E. Renault (1993), 'Indirect Inference,' *Journal of Applied Econometrics*, **18**(S1), S85–S118.

Gregory, A.W., G.W. Smith (1990), 'Calibration as Estimation,' *Econometric Reviews*, **9**(1), 57–89.

Gregory, A.W., G.W. Smith (1991), 'Calibration as Testing: Inference in Simulated Macroeconomic Models,' *Journal of Business and Economic Statistics*, **9**(3), 297–303.

Guerrón-Quintana, P.A. (2010a), 'What You Match Does Matter: The Effects of Data on DSGE Estimation,' *Journal of Applied Econometrics*, **25**(5), 774–804.

Guerrón-Quintana, P.A. (2010b), 'Common Factors in Small Open Economies: Inference and Consequences,' Working Paper 10–04, Federal Reserve Bank of Philadelphia.

Guerrón-Quintana, P.A., A. Inoue, L. Kilian (2010), 'Frequentist inference in weakly identified DSGE models', manuscript, Research Department, Federal Reserve Bank of Philadelphia.

Hall, A.R., A. Inoue, J.M. Nason, B. Rossi (2012), 'Information Criteria for Impulse Response Function Matching Estimation of DSGE Models,' *Journal of Econometrics*, forthcoming.

Hall, G.J. (1996), 'Overtime, Effort, and the Propagation of Business Cycle Shocks,' *Journal of Monetary Economics*, **38**(1), 139–160.

Harvey, A.C. (1989), *Forecasting, Structural Time Series Models, and the Kalman Filter*, Cambridge University Press: Cambridge, England.

Ireland, P.N., (2001), 'Technology shocks and the business cycle: An empirical investigation,' *Journal of Economics Dynamics and Control*, **25**(5), 703–719.

Jeffreys, H. (1998), *The Theory of Probability, Third Edition*, Oxford University Press: Oxford, England.

Justiniano, A., B. Preston (2010), 'Monetary policy and uncertainty in an empirical small open-economy model,' *Journal of Applied Econometrics*, **25**(1), 93–128.

Kano, T. (2009), 'Habit formation and the present-value model of the current account: Yet another suspect,' *Journal of International Economics*, *78*(1), 72–85.

Kano, T., J.M. Nason (2010), 'Business Cycle Implications of Internal Consumption Habit for New Keynesian Models,' manuscript, Federal Reserve Bank of Philadelphia.

Kim, J-Y. (2002), 'Limited information likelihood and Bayesian analysis,' *Journal of Econometrics*, **107**(1-2), 175–193.

Koop, G., M.H. Pesaran, R.P. Smith (2011), 'On Identification of Bayesian DSGE Models,' manuscript, Department of Economics, Birkneck College, London, UK.

Kydland F.E., E.C. Prescott (1982), 'Time to Build and Aggregate Fluctuations,' *Econometrica*, **50**(6) 1345–1370.

Leeper, E.M., M. Plante., and N. Traum (2010), 'Dynamics of fiscal financing in the United States,' *Journal of Econometrics*, **156**(2), 304–321.

Liu, Z., D.F. Waggoner, T. Zha (2011), 'Sources of Macroeconomic Fluctuations: A Regime-Swicthing DSGE Approach,' *Quantitative Economics*, **2**(2), 251–301.

Lubik, T.A., F. Schorfheide (2007), 'Do central banks respond to exchange rate movements? A structural investigation,' *Journal of Monetary Economics*, **54**(4), 1069–1087.

Müller, U.K. (2010), 'Measuring Prior Sensitivity and Prior Informativeness in Large Bayesian Models,' manuscript, Department of Economics, Princeton University.

Otrok, C. (2001), 'On measuring the welfare cost of business cycles,' *Journal of Monetary Economics*, **47** (1), 61–92.

Poirier, D.J. (1998), 'Revising Beliefs in Nonidentified Models,' *Econometric Theory*, **14**(4), 483–509.

Primiceri, G., 2005. 'Time varying structural vector autoregressions and monetary policy,' *Review of Economic Studies*, **72**(3), 821–852.

Rabanal, P., J.F. Rubio-Ramírez (2005), 'Comparing New Keynesian Models of the Business Cycle: A Bayesian Approach,' *Journal of Monetary Economics*, **52**(6), 1151–1166.

Rabanal, P., V. Tuesta (2010), 'Euro-dollar real exchange rate dynamics in an estimated two-country model: An assessment,' *Journal of Economic Dynamics and Control*, **34**(4), 780–797.

Sala, L., U. Söderström, and A. Trigari (2008), 'Monetary policy under uncertainty in an estimated model with labor market frictions,' *Journal of Monetary Economics*, **55**(5), 983–1006.

Sargent, T.J. (1989), 'Two model of measurements and the investment accelerator,' *Journal of Political Economy*, **97**(2), 251–287.

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

Schorfheide, F. (2011), 'Estimation and Evaluation of DSGE Models: Progress and Challenges,' Working Paper 11–07, Federal Reserve Bank of Philadelphia.

Sims, C.A. (2007), 'Thinking about instrumental variables,' manuscript, Department of Economics, Princeton University.

Sims, C.A., T. Zha (2006), 'Were There Regime Switches in U.S. Monetary Policy?,' *American Economic Review*, **96**(1), 54–81.

Smets, F., R. Wouters (2003), 'An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area,' *Journal of the European Economic Association*, **1**(5), 1123–1175.

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

Smith, A.A. (1993), 'Estimating nonlinear time-series models using simulated vector autoregressions,' *Journal of Applied Econometrics*, **18**(S1), S63–S84.

White, H. (1982), 'Maximum likelihood estimation of mis-specified models,' *Econometrica* **50**(1), 1–25.

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

Yun, T. (1996), 'Nominal price rigidity, money supply endogeneity, and business cycles,' *Journal of Monetary Economics*, **37**(2), 345–370.

Table 1. Priors of NKDSGE Model Parameter

Steady State Parameters: $\Theta_{1,ss}$				
	Priors			Probability
	Distribution	A_1	A_2	intervals, 95%
π^*	Normal	4.30	2.50	[−0.600, 9.200]
γ	Gamma	1.65	1.00	[0.304, 3.651]
λ_f	Gamma	0.15	0.10	[0.022, 0.343]
λ_W	Gamma	0.15	0.10	[0.022, 0.343]
R^*	Gamma	1.50	1.00	[0.216, 3.430]

Endogenous Propagation Parameters: $\Theta_{1,prop}$				
	Priors			Probability
	Distribution	A_1	A_2	intervals, 95%
ζ_p	Beta	0.60	0.20	[0.284, 0.842]
ι_p	Beta	0.50	0.28	[0.132, 0.825]
h	Beta	0.70	0.05	[0.615, 0.767]
ν_l	Gamma	2.00	0.75	[0.520, 3.372]
a''	Gamma	0.20	0.10	[0.024, 0.388]
Γ''	Gamma	4.00	1.50	[1.623, 6.743]
ζ_W	Beta	0.60	0.20	[0.284, 0.842]
ι_W	Beta	0.50	0.28	[0.132, 0.825]
ρ_R	Beta	0.50	0.20	[0.229, 0.733]
ψ_1	Gamma	2.00	0.25	[1.540, 2.428]
ψ_2	Gamma	0.20	0.10	[0.024, 0.388]

Exogenous Propagation Parameters: $\Theta_{1,exog}$				
	Priors			Probability
	Distribution	A_1	A_2	intervals, 95%
ρ_z	Beta	0.40	0.25	[0.122, 0.674]
ρ_ϕ	Beta	0.75	0.15	[0.458, 0.950]
ρ_{λ_f}	Beta	0.75	0.15	[0.458, 0.950]
ρ_g	Beta	0.75	0.15	[0.458, 0.950]
σ_z	Inv-Gamma	0.30	4.00	[0.000, 7.601]
σ_ϕ	Inv-Gamma	3.00	4.00	[2.475, 28.899]
σ_{λ_f}	Inv-Gamma	0.20	4.00	[0.000, 6.044]
σ_g	Inv-Gamma	0.50	4.00	[0.002, 10.048]
σ_R	Inv-Gamma	0.20	4.00	[0.000, 6.044]

Columns headed A_1 and A_2 contain the means and standard deviations of the beta, gamma, and normal distributions. For the inverse-gamma distribution, A_1 and A_2 denote scale and shape coefficients.

Table 2. Summary of Posterior Distributions of the NKDSGE Models

Sample: 1982Q1–2009Q4

ln Marginal Likelihoods	
$\mathcal{M}_1 = -39.49$	$\mathcal{M}_2(\iota_p = 0) = -38.69$

Steady State Parameters: $\Theta_{1,ss}$

	Posterior medians	Posterior modes	Probability intervals, 95%	Posterior medians	Posterior modes	Probability intervals, 95%
π^*	2.822	2.831	[2.133, 3.635]	2.804	2.551	[2.116, 3.559]
γ	1.771	1.766	[1.206, 2.356]	1.773	2.109	[1.191, 2.409]
λ_f	0.178	0.178	[0.160, 0.216]	0.177	0.176	[0.160, 0.211]
λ_W	0.215	0.159	[0.086, 0.458]	0.225	0.274	[0.090, 0.473]
R^*	2.629	2.705	[2.014, 3.242]	2.622	1.915	[2.001, 3.229]

Endogenous Propagation Parameters: $\Theta_{1,prop}$

	Posterior medians	Posterior modes	Probability intervals, 95%	Posterior medians	Posterior modes	Probability intervals, 95%
ζ_p	0.656	0.653	[0.578, 0.734]	0.673	0.725	[0.600, 0.743]
ι_p	0.059	0.007	[0.006, 0.215]	NA	NA	NA
h	0.814	0.825	[0.729, 0.872]	0.816	0.830	[0.736, 0.873]
ν_l	1.157	1.003	[0.717, 1.773]	1.156	1.074	[0.720, 1.787]
a''	0.241	0.198	[0.112, 0.459]	0.238	0.249	[0.109, 0.462]
Γ''	10.05	10.14	[6.948, 13.88]	10.13	13.90	[7.029, 14.18]
ζ_W	0.153	0.113	[0.072, 0.270]	0.154	0.180	[0.076, 0.270]
ι_W	0.461	0.514	[0.228, 0.818]	0.467	0.427	[0.224, 0.803]
ρ_R	0.787	0.780	[0.742, 0.823]	0.784	0.780	[0.739, 0.822]
ψ_1	2.513	2.514	[2.161, 2.902]	2.503	2.356	[2.138, 2.897]
ψ_2	0.055	0.052	[0.025, 0.093]	0.053	0.078	[0.024, 0.093]

Exogenous Propagation Parameters: $\Theta_{1,exog}$

	Posterior medians	Posterior modes	Probability intervals, 95%	Posterior medians	Posterior modes	Probability intervals, 95%
ρ_z	0.256	0.226	[0.080, 0.454]	0.257	0.228	[0.077, 0.469]
ρ_ϕ	0.936	0.934	[0.875, 0.976]	0.936	0.963	[0.878, 0.977]
ρ_{λ_f}	0.915	0.921	[0.797, 0.974]	0.912	0.909	[0.804, 0.971]
ρ_g	0.944	0.944	[0.906, 0.975]	0.944	0.937	[0.906, 0.975]
σ_z	0.739	0.722	[0.659, 0.839]	0.741	0.732	[0.660, 0.846]
σ_ϕ	2.259	1.945	[1.741, 3.224]	2.239	2.133	[1.732, 3.154]
σ_{λ_f}	6.639	6.174	[4.987, 9.624]	6.798	8.474	[5.114, 9.822]
σ_g	0.772	0.757	[0.678, 0.889]	0.772	0.759	[0.679, 0.885]
σ_R	0.195	0.193	[0.170, 0.225]	0.196	0.203	[0.172, 0.226]