The New Macroeconometrics: A Bayesian Approach*

Jesús Fernández-Villaverde

University of Pennsylvania, NBER, and CEPR

Pablo Guerrón-Quintana

North Carolina State University

Juan F. Rubio-Ramírez

Duke University and Federal Reserve Bank of Atlanta

January 19, 2009

*Corresponding author: Juan F. Rubio-Ramírez, 213 Social Sciences, Duke University, Durham, NC 27708, USA. E-mail: Juan.Rubio-Ramirez@duke.edu. We thank Alex Groves for his help with preparing the data used in the estimation. Beyond the usual disclaimer, we must note that any views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Finally, we also thank the NSF for financial support.

Abstract

This chapter studies the dynamics of the U.S. economy over the last fifty years via the Bayesian analysis of dynamic stochastic general equilibrium (DSGE) models. Our application is of particular interest because modern macroeconomics is centered around the construction of DSGE models. We formulate and estimate a benchmark DSGE model that captures well the dynamics of the data. This model can be easily applied to policy analysis at public institutions, such as central banks, private organizations, and businesses. We explain how to solve the model and how to evaluate the likelihood using filtering theory. We also discuss the role of priors and how pre-sample information is key for a successful understanding of macro data. Our results document a fair amount of real and nominal rigidities in the U.S. economy. We finish by pointing out lines for future research.

1. Introduction

This chapter studies the dynamics of the U.S. economy over the last fifty years via the Bayesian analysis of dynamic stochastic general equilibrium (DSGE) models. Our data are aggregate, quarterly economic variables and our approach combines macroeconomics (the study of aggregate economic variables like output or inflation) with econometrics (the application of formal statistical tools to economics). This is an example application in what is often called the *New Macroeconometrics.*

Our application is of particular interest because modern macroeconomics is centered around the construction of DSGE models. Economists rely on DSGE models to organize and clarify their thinking, to measure the importance of economic phenomena, and to provide policy prescriptions. DSGE models start by specifying a number of economic agents, typically households, firms, a government, and often a foreign sector, and embodying them with behavioral assumptions, commonly the maximization of an objective function like utility or profits.

The inelegant name of DSGE comes from what happens next. First, economists postulate some sources of shocks to the model: shocks to productivity, to preferences, to taxes, to monetary policy, etc. Those shocks are the "stochastic" part that will drive the model. Second, economists study how agents make their decisions over time as a response to these shocks (hence "dynamic"). Finally, economists focus on the investigation of aggregate outcomes, thus "general equilibrium."

Before we discuss our application, it is important to emphasize that economists employ the word "equilibrium" in a particular and precise technical sense, which is different from daily use or from its meaning in other sciences. We call equilibrium a situation where a) the agents in the model follow a concrete behavioral assumption (usually, but not necessarily, maximization of either their utility or of profits) and b) the decisions of the agents are consistent with each other (for example, the number of units of the good sold must be equal to the number of units of the good bought). Nothing in the definition of equilibrium implies that prices or allocations of goods are constant over time (they may be increasing, decreasing, or fluctuating in rather arbitrary ways), or that the economy is at a rest point, the sense in which equilibrium is commonly used in other fields, particularly in the natural sciences. Many misunderstandings about the accomplishments and potentialities of the concept of equilibrium in economics come about because of a failure to appreciate the subtle difference in language across fields.

General equilibrium has a long tradition in economics, being already implicit in the work of the French Physiocrats of the 18^{th} century, such as Cantillon, Turgot, or Quesnay, or in Adam Smith's magnum opus, *The Wealth of Nations*. However, general equilibrium did not take a front seat in economics until 1874 when Léon Walras published a pioneering book, *Elements of Pure Economics*, in which he laid down the foundations of modern equilibrium theory.

For many decades after Walras, general equilibrium theory focused on static models. This was not out of a lack of appreciation by economists of the importance of dynamics, which were always at the core of practical concerns, but because of the absence of proper methodological instruments to investigate economic change.

Fortunately, the development of recursive techniques in applied mathematics (dynamic programming, Kalman filtering, and optimal control theory, among others) during the 1950s and 1960s provided the tools that economists required for the formal analysis of dynamics. These new methods were first applied to the search for microfoundations of macroeconomics during the 1960s, which attempted to move from *ad hoc* descriptions of aggregate behavior motivated by empirical regularities to descriptions based on first principles of individual decision-making. Later, in the 1970s, recursive techniques opened the door to the rational expectations revolution, started by John Muth (1961) and Robert Lucas (1972), who highlighted the importance of specifying the expectations of the agents in the model in a way that was consistent with their behavior.

This research culminated in 1982 with an immensely influential paper by Finn Kydland and Edward Prescott, *Time to Build and Aggregate Fluctuations*.[1] This article presented the first modern DSGE model, on this propitious occasion, one tailored to explaining the fluctuations of the U.S. economy. Even if most of the material in the paper was already present in other articles written in the previous years by leading economists like Robert Lucas, Thomas Sargent, Christopher Sims, or Neil Wallace, the genius of Kydland and Prescott was to mix the existing ingredients in such a path-breaking recipe that they redirected the attention of the profession to DSGE modelling.

Kydland and Prescott also opened several methodological discussions about how to best construct and evaluate dynamic economic models. The conversation that concerns us in this chapter was about how to take the models to the data, how to assess their behavior, and, potentially, how to compare competing theories. Kydland and Prescott were skeptical about the abilities of formal statistics to provide a useful framework for these three tasks. Instead, they proposed to "calibrate" DSGE models, i.e., to select parameter values by matching some moments of the data and by borrowing from microeconomic evidence, and to judge models by their ability to reproduce properties of the data that had not been employed for calibration (for example, calibration often exploits the first moments of the data to determine parameter

[1] The impact of this work was recognized in 2004 when Kydland and Prescott received the Nobel Prize in economics for this paper and some previous work on time inconsistency in 1977.

values, while the evaluation of the model is done by looking at second moments).

In our previous paragraph, we introduced the idea of "parameters of the model." Before we proceed, it is worthwhile to discuss further what these parameters are. In any DSGE economy, there are functions, like the utility function, the production function, etc. that describe the preferences, technologies, and information sets of the agents. These functions are indexed by a set of parameters. One of the attractive features of DSGE models in the eyes of many economists is that these parameters carry a clear behavioral interpretation: they are directly linked to a feature of the environment about which we can tell an economic history. For instance, in the utility function, we have a discount factor, which tells us how patient the households are, and a risk aversion, which tells us how much households dislike uncertainty. These behavioral parameters (also called "deep parameters" or, more ambiguously, "structural parameters") are of interest because they are invariant to interventions, including shocks by nature or, more important, changes in economic policy.[2]

Kydland and Prescott's calibration became popular because, back in the early 1980s, researchers could not estimate the behavioral parameters of the models in an efficient way. The bottleneck was how to evaluate the likelihood of the model. The equilibrium dynamics of DSGE economies cannot be computed analytically. Instead, we need to resort to numerical approximations. One key issue is, then, how to go from this numerical solution to the likelihood of the model. Nowadays, most researchers apply filtering theory to accomplish this goal. Three decades ago, economists were less familiar with filters and faced the speed constraints of existing computers. Only limited estimation exercises were possible and even those were performed at a considerable cost. Furthermore, as recalled by Sargent (2005), the early experiments on estimation suggested that the prototype DSGE models from the early 1980s were so far away from fitting the data that applying statistical methods to them was not worth the time and effort. Consequently, for over a decade, little work was done on the estimation of DSGE models.[3]

---

[2]Structural parameters stand in opposition to reduced-form parameters, which are those that index empirical models estimated with a weaker link with economic theory. Many economists distrust the evaluation of economic policy undertaken with reduced parameters because they suspect that the historical relations between variables will not be invariant to the changes in economic policy that we are investigating. This insight is known as the *Lucas' critique*, after Robert Lucas, who forcefully emphasized this point in a renown 1976 article (Lucas, 1976).

[3]There was one partial exception. Lars Peter Hansen proposed the use of the Generalized Method of Moments, or GMM (Hansen, 1982), to estimate the parameters of the model by minimizing a quadratic form built from the difference between moments implied by the model and moments in the data. The GMM spread quickly, especially in finance, because of the simplicity of its implementation and because it was the generalization of well-known techniques in econometrics such as instrumental variables OLS. However, the GMM does not deliver all the powerful results implied by likelihood methods and it often has a disappointing performance in small samples.

Again, advances in mathematical tools and in economic theory rapidly changed the landscape. From the perspective of macroeconomics, the streamlined DSGE models of the 1980s begot much richer models in the 1990s. One remarkably successful extension was the introduction of nominal and real rigidities, i.e., the conception that agents cannot immediately adjust to changes in the economic environment. In particular, many of the new DSGE models focused on studying the consequences of limitations in how frequently or how easily agents can changes prices and wages (prices and wages are "sticky"). Since the spirit of these models seemed to capture the tradition of Keynesian economics, they quickly became known as New Keynesian DSGE models (Woodford, 2003). By capturing these nominal and real rigidities, New Keynesian DSGE models began to have a fighting chance at explaining the dynamics of the data, hence suggesting that formally estimating them could be of interest. But this desire for formal estimation would have led to naught if better statistical/numerical methods had not become widely available.

First, economists found out how to approximate the equilibrium dynamics of DSGE models much faster and more accurately. Since estimating DSGE models typically involves solving for the equilibrium dynamics thousands of times, each for a different combination of parameter values, this advance was crucial. Second, economists learned how to run the Kalman filter to evaluate the likelihood of the model implied by the linearized equilibrium dynamics. More recently, economists have also learned how to apply sequential Monte Carlo (SMC) methods to evaluate the likelihood of DSGE models where equilibrium dynamics is nonlinear and/or the shocks in the model are not Gaussian. Third, the popularization of Markov chain Monte Carlo (McMc) algorithms has facilitated the task of exploring and characterizing the likelihood of DSGE models.

Those advances have resulted in an explosion of the estimation of DSGE models, both at the academic level (An and Schorfheide, 2006) and, more remarkable still, at the institutional level, where a growing number of policy-making institutions are estimating DSGE models for policy analysis and forecasting (see appendix A for a partial list). Furthermore, there is growing evidence of the good forecasting record of DSGE models, even if we compare them with the judgmental predictions that staff economists prepare within these policy-making institutions relying on real-time data (Edge, Kiley, and Laforte, 2008).

One of the features of the New Macroeconometrics that the readers of this volume will find exciting is that the (overwhelming?) majority of it is done from an explicitly Bayesian perspective. There are several reasons for this. First, priors are a natural device for economists. To begin with, they use them extensively in game theory or in learning models. More to the point, many economists feel that they have reasonably concrete beliefs about plausible values for most of the behavioral parameters, beliefs built perhaps from years of experience with models and their properties, perhaps from introspection (how risk adverse am I?), perhaps from well-educated economic intuition. Priors offer researchers a flexible way to introduce these beliefs as pre-sample information.

Second, DSGE models are indexed by a relatively large number of parameters (from around 10 to around 60 or 70 depending on the size of the model) while data are sparse (in the best case scenario, the U.S. economy, we have 216 quarters of data from 1954 to 2007).[4] Under this low ratio data/parameters, the desirable small sample properties of Bayesian methods and the inclusion of pre-sample information are peculiarly attractive.

Third, and possibly as a consequence of the argument above, the likelihoods of DSGE models tend to be wild, with dozens of maxima and minima that defeat the best optimization algorithms. McMc are a robust and simple procedure to get around, or at least alleviate, these problems (so much so, that some econometricians who still prefer a classical approach to inference recommend the application of McMc algorithms by adopting a "pseudo-bayesian" stand; see Chernozhukov and Hong, 2003). Our own experience has been that models that we could estimate in a few days using McMc without an unusual effort turned out to be extremely difficult to estimate using maximum likelihood. Even the most pragmatic of economists, with the lowest possible interest in axiomatic foundations of inference, finds this feasibility argument rather compelling.

Fourth, the Bayesian approach deals in a transparent way with misspecification and identification problems, which are pervasive in the estimation of DSGE models (Canova and Sala, 2006). After all, a DSGE model is a very stylized and simplified view of the economy that focuses only on the most important mechanisms at play. Hence, the model is false by construction and we need to keep this notion constantly in view, something that the Bayesian way of thinking has an easier time with than frequentist approaches.

Finally, Bayesian analysis has a direct link with decision theory. The connection is particularly relevant for DSGE models since they are used for policy analysis. Many of the relevant policy decisions require an explicit consideration of uncertainty and asymmetric risk assessments. Most economists, for example, read the empirical evidence as suggesting that the costs of a 1% deflation are considerably bigger than the costs of 1% inflation. Hence, a central bank with a price stability goal (such as the Federal Reserve System in the U.S. or the European Central Bank) faces a radically asymmetric loss function with respect to deviations from the inflation target.

[4]Data before 1954 are less reliable and the structure of the economy back then was sufficiently different as to render the estimation of the same model problematic. Elsewhere (Fernández-Villaverde and Rubio-Ramírez, 2008), we have argued that a similar argument holds for data even before 1981. For most countries outside the U.S., the situation is much worse, since we do not have reliable quarterly data until the 1970s or even the 1980s.

To illustrate this lengthy discussion, we will present a benchmark New Keynesian DSGE model, we will estimate it with U.S. data, and we will discuss our results. We will close by outlining three active areas of research in the estimation of DSGE models that highlight some of the challenges that we see ahead of us for the next few years.

2. A Benchmark New Keynesian Model

Due to space constraints, we cannot present a benchmark New Keynesian DSGE model in all its details. Instead, we will just outline its main features to provide a flavor of what the model is about and what it can and cannot deliver. In particular, we will omit many discussions of why are we doing things the way we do. We will ask the reader to trust that many of our choices are not arbitrary but are the outcome of many years of discussion in the literature. The interested reader can access the whole description of the model at a complementary technical appendix posted at `www.econ.upenn.edu/~jesusfv/benchmark_DSGE.pdf`. The model we select embodies what is considered the current standard of New Keynesian macroeconomics (see Christiano, Eichenbaum, and Evans, 2005) and it is extremely close to the models being employed by several central banks as inputs for policy decisions.

The agents in the model will include a) households that consume, save, and supply labor to a labor "packer," b) a labor "packer" that puts together the labor supplied by different households into an homogeneous labor unit, c) intermediate good producers, who produce goods using capital and aggregated labor, d) a final good producer that mixes all the intermediate goods into a final good that households consume or invest in, and e) a government that sets monetary policy through open market operations. We present in turn each type of agents.

2.1. Households

There is a continuum of infinitely lived households in the economy indexed by j. The households maximize their utility function, which is separable in consumption, c_{jt}, real money balances (nominal money, mo_{jt}, divided by the aggregate price level, p_t), and hours worked, l_{jt}:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t e^{d_t} \left\{ \log\left(c_{jt} - hc_{jt-1}\right) + \upsilon \log\left(\frac{mo_{jt}}{p_t}\right) - e^{\varphi_t}\psi \frac{l_{jt}^{1+\vartheta}}{1+\vartheta} \right\}$$

where β is the discount factor, h controls habit persistence, ϑ is the inverse of the Frisch labor supply elasticity (how much labor supply changes when the wage changes while keeping consumption constant), d_t is a shock to intertemporal preference that follows the AR(1) process:

$$d_t = \rho_d d_{t-1} + e^{\sigma_d} \varepsilon_{d,t} \text{ where } \varepsilon_{d,t} \sim \mathcal{N}(0,1),$$

and φ_t is a labor supply shock that also follows an AR(1) process:

$$\varphi_t = \rho_\varphi \varphi_{t-1} + e^{\sigma_\varphi} \varepsilon_{\varphi,t} \text{ where } \varepsilon_{\varphi,t} \sim \mathcal{N}(0,1).$$

These two shocks, $\varepsilon_{d,t}$ and $\varepsilon_{\varphi,t}$, are equal across all agents.

Households trade on assets contingent on idiosyncratic and aggregate events. By a_{jt+1} we indicate the amount of those securities that pay one unit of consumption at time $t+1$ in event $\omega_{j,t+1,t}$ purchased by household j at time t at real price $q_{jt+1,t}$. These one-period securities are sufficient to span the whole set of financial contracts regardless of their duration. Households also hold an amount b_{jt} of government bonds that pay a nominal gross interest rate of R_t and invest a quantity x_t of the final good. Thus, the $j-th$ household's budget constraint is given by:

$$\begin{aligned} & c_{jt} + x_{jt} + \frac{mo_{jt}}{p_t} + \frac{b_{jt+1}}{p_t} + \int q_{jt+1,t} a_{jt+1} d\omega_{j,t+1,t} \\ = & \; w_{jt} l_{jt} + \left(r_t u_{jt} - \mu_t^{-1} \Phi \left[u_{jt} \right] \right) k_{jt-1} + \frac{mo_{jt-1}}{p_t} + R_{t-1} \frac{b_{jt}}{p_t} + a_{jt} + T_t + \digamma_t \end{aligned} \tag{1}$$

where w_{jt} is the real wage, r_t the real rental price of capital, $u_{jt} > 0$ the intensity of use of capital, $\mu_t^{-1}\Phi[u_{jt}]$ is the physical cost of u_{jt} in resource terms (where $\Phi[u] = \Phi_1(u-1) + \frac{\Phi_2}{2}(u-1)^2$ and $\Phi_1, \Phi_2 \geq 0$), μ_t is an investment-specific technological shock that we will describe momentarily, T_t is a lump-sum transfer from the government, and $\digamma_t$ is the household share of the profits of the firms in the economy.

Given a depreciation rate δ and a quadratic adjustment cost function

$$V\left[\frac{x_t}{x_{t-1}}\right] = \frac{\kappa}{2}\left(\frac{x_t}{x_{t-1}} - \Lambda_x\right)^2,$$

where $\kappa \geq 0$ and Λ_x is the long-run growth of investment, capital evolves as:

$$k_{jt} = (1-\delta) k_{jt-1} + \mu_t \left(1 - V\left[\frac{x_{jt}}{x_{jt-1}}\right]\right) x_{jt}.$$

Note that in this equation there appears an investment-specific technological shock μ_t that also follows an autoregressive process:

$$\mu_t = \mu_{t-1} \exp\left(\Lambda_\mu + z_{\mu,t}\right) \text{ where } z_{\mu,t} = \sigma_\mu \varepsilon_{\mu,t} \text{ and } \varepsilon_{\mu,t} \sim \mathcal{N}(0,1).$$

Thus, the problem of the household is a Lagrangian function formed by the utility function, the law of motion for capital, and the budget constraint. The first order conditions of this problem with respect to c_{jt}, b_{jt}, u_{jt}, k_{jt}, and x_{jt} are standard (although messy), and we refer the reader to the online appendix for a detailed exposition.

The first order condition with respect to labor and wages is more involved because, as we explained in the introduction, households will face rigidities in changing their wages.[5] Each household supplies a slightly different type of labor service (for example, some households write chapters on Bayesian estimation of DSGE models and some write chapters on hierarchical models) that is aggregated by a labor "packer" (in our example, an editor) into an homogeneous labor unit ("Bayesian research") according to the function:

$$l_t^d = \left(\int_0^1 l_{jt}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} \tag{2}$$

where η controls the elasticity of substitution among different types of labor and l_t^d is the aggregate labor demand. The technical role that the labor "packer" plays in the model is to allow for the existence of many different types of labor, and hence for the possibility of different wages, while keeping the heterogeneity of agents at a tractable level.

The labor "packer" takes all wages w_{jt} of the differentiated labor and the wage of the homogeneous labor unit w_t as given and maximizes profits subject to the production function (2). Thus, the first order conditions of the labor "packer" are:

$$l_{jt} = \left(\frac{w_{jt}}{w_t} \right)^{-\eta} l_t^d \qquad \forall j \tag{3}$$

We assume that there is free entry into the labor packing business. After a few substitutions, we get an expression for the wage w_t as a function of the different wages w_{jt}:

$$w_t = \left(\int_0^1 w_{jt}^{1-\eta} dj \right)^{\frac{1}{1-\eta}}.$$

The way in which households set their wages is through a device called Calvo pricing (for the economist Guillermo Calvo, who proposed this formulation in 1983). In each period, the household has a probability $1-\theta_w$ of being able to change its wages. Otherwise, it can only

[5] In this type of model, households set up their wages and firms decide how much labor to hire. We could also have specified a model where wages are posted by firms and households decide how much to work. For several technical reasons, our modelling choice is easier to handle. If the reader has problems seeing a household setting its wage, she can think of the case of an individual contractor posting a wage for hour worked, or a union negotiating a contract in favor of its members.

partially index its wages by a fraction $\chi_w \in [0,1]$ of past inflation. Therefore, if a household has not been able to change its wage for τ periods, its real wage is

$$\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi_w}}{\Pi_{t+s}} w_{jt},$$

the original wage times the indexation divided by the accumulated inflation. This probability $1-\theta_w$ represents a streamlined version of a more subtle explanation of wage rigidities (based on issues like contracts costs, Caplin and Leahy, 1991, or information limitations, Sims, 2002), which we do not flesh out entirely in the model to keep tractability.[6]

The average duration of a wage decision in this economy will be $1/(1-\theta_w)$, although all wages change period by period because of the presence of indexation (in that sense the economy would look surprisingly flexible to a naïve observer!). Since a suitable law of large numbers holds for this economy, the probability of changing wages, $1-\theta_w$, will also be equal to the fraction of households reoptimizing their wages in each period.

Calvo pricing implies three results. First, the form of the utility function that we selected and the presence of complete asset markets deliver the remarkable result that all households that choose their wage in this period will choose exactly the same wage, w_t^* (also, the consumption, investment, and Lagrangian multiplier λ_t of the budget constraint of all households will be the same and we drop the subindex j when no confusion arises).

Second, in every period, a fraction $1-\theta_w$ of households set w_t^* as their wage, while the remaining fraction θ_w partially index their price by past inflation. Consequently, the real wage index evolves as:

$$w_t^{1-\eta} = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t} \right)^{1-\eta} w_{t-1}^{1-\eta} + (1-\theta_w)\, w_t^{*1-\eta}.$$

Third, after a fair amount of algebra, we can show that the evolution of w_t^* and an auxiliary variable f_t are governed by two recursive equations:

$$f_t = \frac{\eta-1}{\eta} \left(w_t^*\right)^{1-\eta} \lambda_t w_t^{\eta} l_t^d + \beta\theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}} \right)^{1-\eta} \left(\frac{w_{t+1}^*}{w_t^*} \right)^{\eta-1} f_{t+1}$$

and:

$$f_t = \psi d_t \varphi_t \left(\frac{w_t}{w_t^*} \right)^{\eta(1+\vartheta)} \left(l_t^d\right)^{1+\vartheta} + \beta\theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}} \right)^{-\eta(1+\vartheta)} \left(\frac{w_{t+1}^*}{w_t^*} \right)^{\eta(1+\vartheta)} f_{t+1}.$$

[6] There is, however, a discussion in the literature regarding the potential shortcomings of Calvo pricing. See, for example, Dotsey, King, and Wolman (1999), for a paper that prefers to be explicit about the problem faced by agents when changing prices.

2.2. The Final Good Producer

The final good producer aggregates all intermediate goods into a final good with the following production function:

$$y_t^d = \left(\int_0^1 y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}. \tag{4}$$

where ε controls the elasticity of substitution.

The role of the final good producer is to allow for the existence of many firms with differentiated products while keeping a simple structure in the heterogeneity of products and prices. Also, the final good producer takes as given all intermediate goods prices p_{ti} and the final good price p_t, which implies a demand function for each good

$$y_{it} = \left(\frac{p_{it}}{p_t} \right)^{-\varepsilon} y_t^d \qquad \forall i,$$

where y_t^d is the aggregate demand and by a zero profit condition:

$$p_t = \left(\int_0^1 p_{it}^{1-\varepsilon} di \right)^{\frac{1}{1-\varepsilon}}.$$

2.3. Intermediate Good Producers

There is a continuum of intermediate good producers. Each intermediate good producer i has access to a technology described by a production function of the form $y_{it} = A_t k_{it-1}^{\alpha} \left(l_{it}^d \right)^{1-\alpha} - \phi z_t$ where k_{it-1} is the capital rented by the firm, l_{it}^d is the amount of the homogeneous labor input rented by the firm, the parameter ϕ corresponds to the fixed cost of production, and where A_t follows a unit root process in logs, $A_t = A_{t-1} \exp\left(\Lambda_A + z_{A,t}\right)$ where $z_{A,t} = \sigma_A \varepsilon_{A,t}$ and $\varepsilon_{A,t} \sim \mathcal{N}(0,1)$.

The fixed cost ϕ is scaled by the variable $z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$ to guarantee that economic profits are roughly equal to zero. The variable z_t evolves over time as $z_t = z_{t-1} \exp\left(\Lambda_z + z_{z,t}\right)$ where $z_{z,t} = \frac{z_{A,t} + \alpha z_{\mu,t}}{1-\alpha}$ and $\Lambda_z = \frac{\Lambda_A + \alpha \Lambda_\mu}{1-\alpha}$. We can also prove that Λ_z is the mean growth rate of the economy.

Intermediate good producers solve a two-stage problem. First, given w_t and r_t, they rent l_{it}^d and k_{it-1} to minimize real costs of production, which implies a marginal cost of:

$$mc_t = \left(\frac{1}{1-\alpha} \right)^{1-\alpha} \left(\frac{1}{\alpha} \right)^{\alpha} \frac{w_t^{1-\alpha} r_t^{\alpha}}{A_t}$$

This marginal cost does not depend on i: all firms receive the same shocks and rent inputs at the same price.

Second, intermediate good producers choose the price that maximizes discounted real profits under the same Calvo pricing scheme as households. The only difference is that now the probability of changing prices is given by $1-\theta_p$ and the partial indexation by the parameter $\chi \in [0,1]$. We will call p_t^* the price that intermediate good producers select when they are allowed to optimize their prices at time t.

Again, after a fair amount of manipulation, we find that the price index evolves as:

$$p_t^{1-\varepsilon} = \theta_p \left(\Pi_{t-1}^{\chi}\right)^{1-\varepsilon} p_{t-1}^{1-\varepsilon} + (1-\theta_p)\, p_t^{*1-\varepsilon}$$

and that p_t^* is determined by two recursive equations in the auxiliary variable g_t^1 and g_t^2:

$$g_t^1 = \lambda_t mc_t y_t^d + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{-\varepsilon} g_{t+1}^1$$

$$g_t^2 = \lambda_t \Pi_t^* y_t^d + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{1-\varepsilon} \left(\frac{\Pi_t^*}{\Pi_{t+1}^*}\right) g_{t+1}^2$$

where $\varepsilon g_t^1 = (\varepsilon - 1) g_t^2$ and $\Pi_t^* = p_t^*/p_t$.

2.4. The Government

The government plays an extremely limited role in this economy. It sets the nominal interest rates according to the following policy rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_\Pi} \left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\Lambda_z}\right)^{\gamma_y}\right)^{1-\gamma_R} e^{m_t} \qquad (5)$$

The policy is implemented through open market operations that are financed with lump-sum transfers T_t to ensure that the government budget is balanced period by period. The variable Π is the target level of inflation, R the steady-state gross return of capital, and Λ_z, as mentioned above, is the average gross growth rate of y_t^d (these last two variables are determined by the model, not by the government). The term m_t is a random shock to monetary policy that follows $m_t = \sigma_m \varepsilon_{mt}$ where ε_{mt} is distributed according to $\mathcal{N}(0,1)$.

The intuition behind the policy rules, motivated on a large body of empirical literature (Orphanides, 2002), is that a good way to describe the behavior of the monetary authority is to think about central banks as setting up the (short-term) interest rate as a function of the past interest rate, R_{t-1}, the deviation of inflation from a target, and the deviation of output growth from the long-run average growth rate.

2.5. Aggregation

With some additional work, we can sum up the behavior of all agents in the economy to find expressions for the remaining aggregate variables, including aggregate demand:

$$y_t^d = c_t + x_t + \mu_t^{-1}\Phi\left[u_t\right]k_{t-1}$$

aggregate supply:

$$y_t^s = \frac{A_t\left(u_t k_{t-1}\right)^{\alpha}\left(l_t^d\right)^{1-\alpha} - \phi z_t}{v_t^p}$$

where:

$$v_t^p = \int_0^1 \left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} di$$

and aggregate labor supply $l_t^d = l_t / v_t^w$ where

$$v_t^w = \int_0^1 \left(\frac{w_{jt}}{w_t}\right)^{-\eta} dj$$

The terms v_t^p and v_t^w represent the loss of efficiency induced by price and wage dispersion. Since all households and firms are symmetrical, a social planner would like to set every price and wage at the same level. However, the nominal rigidities prevent this socially optimal solution. By the properties of the index under Calvo's pricing, v_t^p and v_t^w evolve as:

$$\begin{aligned} v_t^p &= \theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + \left(1-\theta_p\right)\Pi_t^{*-\varepsilon} \\ v_t^w &= \theta_w\left(\frac{w_{t-1}}{w_t}\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{-\eta} v_{t-1}^w + \left(1-\theta_w\right)\left(\Pi_t^{w*}\right)^{-\eta}. \end{aligned}$$

2.6. Equilibrium

The equilibrium of this economy is characterized by the first order conditions of the household, the first order conditions of the firms, the policy rule of the government, and the aggregate variables that we just presented, with a total of 21 equations (the web appendix has the list of all of these 21 equations). Formally, we can think about them as a set of nonlinear functional equations where the unknowns are the functions that determine the evolution over time of the variables of the model. The variables that enter as arguments of these functions are the 19 state variables that determine the situation of the system at any point in time.

A simple inspection of the system reveals that it does not have an analytical solution. Instead, we need to resort to some type of numerical approximation to solve it. The literature on how to do so is enormous, and we do not review it here (see, instead, the comprehensive text-book by Judd, 1998). In previous work (Aruoba, Fernández-Villaverde, and Rubio-Ramírez, 2006), we convinced ourselves that a nice compromise between speed and accuracy in the solution of systems of this sort can be achieved by the use of perturbation techniques.

The foundation of perturbation methods is to substitute the original system by a suitable variation of it that allows for an analytically (or at least numerically) exact solution. In our particular case, this variation of the problem is to set the standard deviations of all of the model's shocks to zero. Under this condition, the economy converges to a balanced growth path, where all the variables grow at a constant (but potentially different) rate. Then, we use this solution and a functional version of the implicit function theorem to compute the leading coefficients of the Taylor expansion of the exact solution.

Traditionally, the literature has focused on the first order approximation to the solution because it generates a linear representation of the model whose likelihood is easily evaluated using the Kalman filter. Since this chapter is a brief introduction to the New Macroeconometrics, we will follow that convention. However, we could also easily compute a second order approximation. Higher order approximations have the advantage of being more accurate and of capturing precautionary behavior among agents induced by variances. We have found in our own research that in many problems (but not in all!), those second order terms are important.

Further elaborating this point, note that we end up working with an approximation of the solution of the model and not with the solution itself raises an interesting technical issue. Even if we were able to evaluate the likelihood implied by that solution, we would not be evaluating the likelihood of the exact solution of the model but the likelihood implied by the approximated solution of the model. Both objects may be quite different and some care is required when we proceed to perform inference (Fernández-Villaverde, Rubio-Ramírez, and Santos, 2006).

However, before digging into the details of the solution method, and since we have technological progress, we need to rescale nearly all variables to make them stationary. To do so, define $\widetilde{c}_t = \frac{c_t}{z_t}$, $\widetilde{\lambda}_t = \lambda_t z_t$, $\widetilde{r}_t = r_t \mu_t$, $\widetilde{q}_t = q_t \mu_t$, $\widetilde{x}_t = \frac{x_t}{z_t}$, $\widetilde{w}_t = \frac{w_t}{z_t}$, $\widetilde{w}_t^* = \frac{w_t^*}{z_t}$, $\widetilde{k}_t = \frac{k_t}{z_t \mu_t}$, and $\widetilde{y}_t^d = \frac{y_t^d}{z_t}$. Also note that $\Lambda_c = \Lambda_x = \Lambda_w = \Lambda_{w^*} = \Lambda_{y^d} = \Lambda_z$. Furthermore, we normalize $u = 1$ in the steady state by setting $\Phi_1 = \widetilde{r}$, eliminating it as a free parameter.

For each variable var_t, we define $\widehat{var}_t = var_t - var$, as the deviation with respect to the steady state. Then, the states of the model $\overline{S}_t$ are given by:

$$\overline{S}_t = \begin{pmatrix} \widehat{\Pi}_{t-1}, \widehat{\widetilde{w}}_{t-1}, \widehat{g}^1_{t-1}, \widehat{g}^2_{t-1}, \widehat{\widetilde{k}}_{t-1}, \widehat{R}_{t-1}, \widehat{\widetilde{y}}^d_{t-1}, \widehat{\widetilde{c}}_{t-1}, \widehat{v}^p_{t-1}, \widehat{v}^w_{t-1}, \\ \widehat{\widetilde{q}}_{t-1}, \widehat{\widetilde{f}}_{t-1}, \widehat{\widetilde{x}}_{t-1}, \widehat{\widetilde{\lambda}}_{t-1}, \widehat{\widetilde{z}}_{t-1}, z_{\mu,t-1}, \widehat{d}_{t-1}, \widehat{\varphi}_{t-1}, z_{A,t-1} \end{pmatrix}',$$

and the exogenous shocks are $\varepsilon_t = (\varepsilon_{\mu,t}, \varepsilon_{d,t}, \varepsilon_{\varphi,t}, \varepsilon_{A,t}, \varepsilon_{m,t})'$.

From the output of the perturbation, we build the law of motion for the states:

$$\overline{S}_{t+1} = \Psi_{s1}\left(\overline{S}_t', \varepsilon_t'\right)' \tag{6}$$

where Ψ_{s1} is a 19×24 matrix, and the law of motion for the observables:

$$\mathbb{Y}_t = \Psi_{o1}\left(S_t', \varepsilon_t'\right)' \tag{7}$$

where $S_t = \left(\overline{S}_t', \overline{S}_{t-1}', \varepsilon_{t-1}'\right)$ and Ψ_{o1} is a 5×48 matrix. We include lagged states in our observation equation, because some of our data will appear in first differences. Equivalently, we could have added lagged observations to the states to accomplish the same objective. Also, note that the function (7) includes a constant that captures the mean of the observables as implied by the equilibrium of the model.

Our observables are the first differences of the relative price of investment, output, real wages, inflation, and the federal funds rate, or in our notation:

$$\mathbb{Y}_t = \left(\triangle \log \mu_t^{-1}, \triangle \log y_t, \triangle \log w_t, \log \Pi_t, \log R_t\right)'.$$

While the law of motion for states is unique (or at least belonging to an equivalence class of representations), the observation equation depends on what the econometrician actually observes or chooses to observe. Since little is known about how those choices affect our estimation, we have selected the time series that we find particularly informative for our purposes (see Guerrón-Quintana, 2008, for a detailed discussion).

For observables, we can only select a number of series less than or equal to the number of shocks in the model. Otherwise, the model will be stochastically singular, i.e., we could write the extra observables as *deterministic* functions of the other observables and the likelihood would be $-\infty$. In the jargon of macroeconomics, these functions are part of the equilibrium cross-equation restrictions.

A popular way around the problem of stochastic singularity has been to assume that the observables come with measurement error. This assumption delivers, by itself, one shock per observable. There are three justifications for measurement error. First, statistical agencies make mistakes. Measuring U.S. output or wages is a daunting task, undertaken with extremely limited resources by different government agencies. Despite their remarkable efforts, statistical agencies can only provide us with an estimate of the series we need. Second, there are differences between the definitions of variables in the theory and in the data, some caused by divergent methodological choices (see, for instance, the discussion in appendix B about how to move from nominal output into real output), some caused by the limitations of the data the statistical agency can collect. Finally, measurement errors may account for parts of the dynamic of the data that are not captured by the model.

On the negative side, including measurement error complicates identification and faces the risk that the data ends up being explained by this measurement error and not by the model itself. In our own research we have estimated DSGE models with and without measurement error, according to the circumstances of the problem. In this chapter, since we have a rich DSGE economy with many sources of uncertainty, we assume that our data come without measurement error. This makes the exercise more transparent and easier to follow.

2.7. The Likelihood Function

Equations (6) and (7) plus the definition of S_t are the state space representation of a dynamic model. It is well understood that we can use this state representation to find the likelihood $\mathcal{L}(\mathbb{Y}_{1:T};\Psi)$ of our DSGE model, where we have stacked all parameter values in:

$$\Psi = \left\{\beta, h, \upsilon, \vartheta, \delta, \eta, \varepsilon, \alpha, \phi, \theta_w, \chi_w, \theta_p, \chi_p, \Phi_2, \gamma_R, \gamma_y, \gamma_\Pi, \Pi, \Lambda_\mu, \Lambda_A, \rho_d, \rho_\varphi, \sigma_\mu, \sigma_d, \sigma_A, \sigma_m, \sigma_\varphi\right\}.$$

and where $\mathbb{Y}_{1:T}$ is the sequence of data from period 1 to T

To perform this evaluation of the likelihood function given some parameter values, we start by factorizing the likelihood function as:

$$\mathcal{L}(\mathbb{Y}_{1:T};\Psi) = \prod_{t=1}^{T} \mathcal{L}(\mathbb{Y}_t|\mathbb{Y}_{1:T-1};\Psi)$$

Then:

$$\mathcal{L}(\mathbb{Y}_{1:T};\Psi) = \int \mathcal{L}(\mathbb{Y}_1|S_0;\Psi)\, dS_0 \prod_{t=2}^{T} \int \mathcal{L}(\mathbb{Y}_t|S_t;\Psi)\, p(S_t|\mathbb{Y}_{1:T-1};\Psi)\, dS_t \tag{8}$$

If we know S_t, computing $\mathcal{L}(\mathbb{Y}_t|S_t;\Psi)$ is conceptually simple since, conditional on S_t, the measurement equation (7) is a change of variables from ε_t to $\mathbb{Y}_t$. Similarly, if we know S_0, it is easy to compute $\mathcal{L}(\mathbb{Y}_1|S_0;\Psi)$ with the help of (6) and (7). Thus, all that remains to evaluate the likelihood is to find the sequence $\{p(S_t|\mathbb{Y}_{1:t-1};\Psi)\}_{t=1}^{T}$ and the initial distribution $p(S_0;\Psi)$.

This second task is comparatively simple. There are two procedures for doing so. First, in the case where we can characterize the ergodic distribution, for example, in linearized models, we can draw directly from it (this procedure takes advantage of the stationarity of the model achieved by the rescaling of variables). In our model, we just need to set the states to zero (remember, they are expressed in differences with respect to the steady state) and specify their initial variance-covariance matrix. Second, if we cannot characterize the ergodic distribution, we can simulate the model for a large number of periods (250 for a model like ours may suffice). Under certain regularity conditions, the values of the states at period 250 are a draw from the ergodic distribution of states. Consequently, the only remaining barrier is to characterize the sequence of conditional distributions $\{p(S_t|\mathbb{Y}_{1:T-1};\Psi)\}_{t=1}^{T}$ and to compute the integrals in (8).

Since a) we have performed a first order perturbation of the equilibrium equations of the problem and generated a linear state space representation, and b) our shocks are normally distributed, the sequence $\{p(S_t|\mathbb{Y}_{1:T-1};\Psi)\}_{t=1}^{T}$ is composed of conditionally Gaussian distributions. Hence, we can use the Kalman filter, find $\{p(S_t|\mathbb{Y}_{1:T-1};\Psi)\}_{t=1}^{T}$, and evaluate the likelihood in a quick and efficient way. A version of the Kalman filter applicable to our setup is described in appendix B. The Kalman filter is extremely efficient: it takes less than one second to run for each combination of parameter values.

However, in many cases of interest, we may want to perform a higher order perturbation or the shocks to the model may not be normally distributed. Then, the components of $\{p(S_t|\mathbb{Y}_{1:T-1};\Psi)\}_{t=1}^{T}$ would not be conditionally Gaussian. In fact, in general, it will not belong to any known parametric family of density functions. In all of those cases, we can track the desired sequence $\{p(S_t|\mathbb{Y}_{1:T-1};\Psi)\}_{t=1}^{T}$ using SMC methods (see Appendix A).

Once we have the likelihood of the model $\mathcal{L}(\mathbb{Y}_{1:T};\Psi)$, we combine it with a prior density for the parameters $p(\Psi)$ to form a posterior

$$p(\Psi|\,\mathbb{Y}_{1:T}) \propto \mathcal{L}(\mathbb{Y}_{1:T};\Psi)\,p(\Psi).$$

Since the posterior is also difficult to characterize (we do not even have an expression for the likelihood, only an *evaluation*), we generate draws from it using a random walk Metropolis-Hastings algorithm. The scaling matrix of the proposal density will be selected to generate the appropriate acceptance ratio of proposals (Roberts, Gelman, and Gilks, 1997).

We can use the resulting empirical distribution to obtain point estimates, variances, etc., and to build the marginal likelihood that is a fundamental component in the Bayesian comparison of models (see Fernández-Villaverde and Rubio-Ramírez, 2004, for an application of the comparison of dynamic models in economics). An important advantage of having draws from the posterior is that we can also compute other objects of interest, like the posterior of the welfare cost of business cycle fluctuations or the posterior of the responses of the economy to an unanticipated shock. These objects are often of more interest to economists than the parameter themselves.

3. Empirical Analysis

Now we are ready to take our model to the data and obtain point estimates and posterior distributions. Again, because of space limitations, our analysis should be read more as an example of the type of exercises that can be implemented than as an exhaustive investigation of the empirical properties of the model. Furthermore, once we have parameter posteriors, we can undertake many additional exercises of interest, like forecasting (where is the economy going?), counterfactual analysis (what would have happen if?), policy design (what is the best way to respond to shocks?), and welfare analysis (what is the cost of business cycle fluctuations?).

We estimate our New Keynesian model using five time series for the U.S. economy: 1) the relative price of investment with respect to the price of consumption, 2) real output per capita growth, 3) real wages per capita, 4) the consumer price index, and 5) the federal funds rate (the interest rate at which banks lend balances at the Federal Reserve System to each other, usually overnight). Our sample goes from 1959Q1 - 2007Q1. Appendix B explains how we construct these series.

Before proceeding further, we specify priors for the parameters. We will have two sets of priors that will provide us with two sets of posteriors. First, we adopt flat priors for all parameters except a few that we fix at some predetermined values. These flat priors are modified only by imposing boundary constraints to make the priors proper and to rule out parameter values that are incompatible with the model (i.e., a negative value for a variance or Calvo parameters outside the unit interval). Second, we use a set of priors nearly identical to the one proposed by Smets and Wouters (2007) in a highly influential study in which they estimate a similar DSGE model to ours.

Our first exercise with flat priors is motivated by the observation that, with this prior, the posterior is proportional to the likelihood function.[7] Consequently, our Bayesian results can be interpreted as a classical exercise where the mode of the likelihood function (the point estimate under an absolute value loss function for estimation) is the maximum likelihood estimate. Moreover, a researcher who prefers alternative priors can always reweight the draws from the posterior using importance sampling to encompass his favorite priors. We do not argue that our flat priors are uninformative. After a reparameterization of the model, a flat prior may become highly curved. Instead, we have found in our research that an estimation with flat priors is a good first step to learn about the amount of information carried by the likelihood and a natural benchmark against which to compare results obtained

[7] There is a small qualifier: the bounded support of some of the priors. We can eliminate this small difference by thinking about those bounds as frontiers of admissible parameter values in a classical perspective.

with nonflat priors. The main disadvantage of this first exercise is that, in the absence of further information, it is extremely difficult to get the McMc to converge and characterize the posterior accurately. The number of local modes is so high that the chain wanders away for the longest time without settling into the ergodic distribution.

In comparison, our second exercise, with Smets and Wouters' (2007) priors, is closer to the exercises performed at central banks and other policy institutions. The priors bring a considerable amount of additional information that allows us to achieve much more reasonable estimates. However, the use of more aggressive priors requires a careful preparation by the researcher, and in the case of institutional models, a candid conversation with the principal regarding the assumptions about parameters she feels comfortable with.

3.1. Flat Priors

Now we present results from our exercise with flat priors. Before reporting results, we fix some parameter values at the values reported in Table 1. In a Bayesian context, we can think about fixing parameter values as having Dirac priors over them.

Table 1: Fixed Parameters

β	h	ψ	ϑ	δ	α	ε	η	κ	ϕ	Φ_2
0.99	0.9	9	1.35	0.015	0.30	8	8	30	0	0.001

Macroeconomists fix some parameters for different reasons. One reason is related to the amount of data: being too ambitious in estimating all the parameters in the model with limited data sometimes delivers posteriors that are too spread for useful interpretation and policy analysis. A related reason is that some parameters are poorly identified with macro data (for example, the elasticity of output with respect to capital), while, at the same time, we have good sources of information from micro data not used in the estimation. It seems, then, reasonable to set these parameters at the conventional values determined by microeconometricians. In our first exercise with flat priors, the number of parameters that we need to fix is relatively high, 11. Otherwise, the likelihood has too many local peaks, and it turns out to be extremely difficult to design a successful McMc.

We briefly describe some of these parameter values in Table 1. The discount factor, β, is fixed at 0.99, a number close to one, to match the low risk free rate observed in the U.S. economy (the gross risk free rate depends on the inverse of β and the growth rate of the economy). Habit persistence, $h = 0.9$, arises because we want to match the evidence of a sluggish adjustment of consumption decisions to shocks. The parameter ϑ pins down a Frisch elasticity of 0.74, a value well within the range of the findings of microeconometrics

(Browning, Hansen, and Heckman, 1999). Our choice of δ, the depreciation rate, matches the capital-output ratio in the data. The elasticity of capital to output, $\alpha = 0.3$, is informed by the share of national income that goes to capital. The elasticity of substitution across intermediate goods, ε, accounts for the microeconomic evidence on the average mark-up of U.S. firms, which is estimated to be around 10-15 percent (the mark-up in our model is approximately $\varepsilon/(\varepsilon-1)$). By symmetry, we pick the same value for η. We take a high value for the adjustment cost κ, 30, to dampen the fluctuations of investment in the model and get them closer to the behavior of investment in the data. We set ϕ to zero, since we do not have information on pure profits by firms. Luckily, since we do not have entry and exit of firms in the model, the parameter is nearly irrelevant for equilibrium dynamics. Finally, the parameter controlling money demand υ does not affect the dynamics of the model because the monetary authority will supply as much money as required to implement the nominal interest rate determined by the policy rule.

We summarize the posterior distribution of the parameters in Table 2, where we report the median of each parameter and its 5 and 95 percentile values and in Figure 1, where we plot the histograms of the draws of each parameter. The results are based on a chain of 75,000 draws, initialized after a considerable amount of search for good initial starting values, and a 30 percent acceptance rate. We performed standard analysis of convergence of the chain to ensure that the results are accurate.

Table 2: Median Estimated Parameters (5 and 95 percentile in parentheses)

θ_p	χ	θ_w	χ_w	γ_R	γ_y	γ_π	Π	
0.72 [0.68,0.77]	0.94 [0.81,0.99]	0.62 [0.56,0.70]	0.92 [0.73,0.99]	0.87 [0.83,0.90]	0.99 [0.59,1.73]	2.65 [2.10,3.67]	1.012 [1.011,1.013]	
ρ_d	ρ_φ	σ_A	σ_d	σ_φ	σ_μ	σ_e	Λ_μ	Λ_A
0.14 [0.02,0.27]	0.91 [0.86,0.93]	-4.45 [-4.61,-4.29]	-2.53 [-2.64,-2.43]	-2.29 [-2.59,-1.74]	-5.49 [-5.57,-5.40]	-5.97 [-6.09,-5.84]	$3e-3$ [0.002,0.004]	$1e-4$ [0.00,0.0003]

[FIGURE 1 HERE]

We highlight a few results. The Calvo parameter for prices, θ_p, is closely estimated to be around 0.72. This indicates that firms revise their pricing decisions around 3.5 quarters on average, a number that seems close to a natural benchmark of a yearly pricing cycle in many firms complemented with some firms that revise their prices more often. The indexation level, χ, of 0.94, has a higher degree of uncertainty but also suggests that firms respond quickly to inflation. High indexation generates a high level of inflation inertia, and with it, a bigger role for monetary policy. The Calvo parameter for wages, θ_w, implies wage decisions every 2.6 quarters, also with high indexation.

The parameters of the policy rule show that the monetary authority smooths to a large degree the evolution of the federal funds rate, γ_R is equal to 0.87, cares about the growth level of the economy, γ_y is equal to 0.99, and finally, is rather aggressive against inflation, γ_π is equal to 2.65. This last parameter value is important, because it implies that the monetary authority respects the "Taylor principle" that requires nominal interest rates to rise more than inflation rises. Otherwise, the real interest rate falls, inflation rises, and we generate bad policy outcomes due to the indeterminacy of equilibria (Lubick and Schorfheide, 2004). The inflation target of 1 percent per quarter is higher than the stated goals of the Federal Reserve but consistent with the behavior of inflation in the sample.

The standard deviations of the five shocks in the economy show the importance of the preference shocks and the higher importance of the neutral technological shock in comparison with the investment-specific technological shock in accounting for fluctuations in the economy. In comparison, the posterior clearly indicated that the *average* growth of the economy is driven by investment-specific technological change, whose mean, Λ_μ, is an order of magnitude bigger than neutral technological change, Λ_A. These findings are similar to the ones obtained using a calibration approach by Greenwood, Herkowitz, and Krusell (1997).

3.2. Smets and Wouters' (2007) Priors

In our second estimation exercise, we incorporate priors for the parameters of the model based on the work of Smets and Wouters (2007). First, as we did before, we fix some parameter values. In this case, since the priors help to achieve identification, we can free 6 out of the 11 parameters fixed in the first exercise. The remaining 5 fixed parameters are those still difficult to identify with aggregate data and are summarized in Table 3. Note that, to maintain comparability with Smets and Wouters' paper, we increase the elasticities of substitution, ε and η, to 10, an increase that has a minuscule effect on the dynamics of the data.

Table 3: Fixed Parameters

δ	ε	η	ϕ	Φ_2
0.025	10	10	0	0.001

We modify several of the priors of Smets and Wouters to induce better identification and to accommodate some of the differences between their model and ours. Instead of a detailed discussion of the reasons behind each prior, it is better to highlight that the priors are common in the literature and are centered around values that are in the middle of the usual range of estimates, using both macro and micro data, and diverse econometric methodologies. We summarize our priors in Table 4.

Table 4: Priors					
$100\left(\beta^{-1}-1\right)$	h	ψ	θ_p	χ	θ_w
$Ga(0.25, 0.1)$	$Be(0.7, 0.1)$	$N(9, 3)$	$Be(0.5, 0.1)$	$Be(0.5, 0.15)$	$Be(0.5, 0.1)$
χ_w	γ_R	γ_y	γ_π	$100(\Pi-1)$	ϑ
$Be(0.5, 0.1)$	$Be(0.75, 0.1)$	$N(0.12, 0.05)$	$N(1.5, 0.125)$	$Ga(0.95, 0.1)$	$N(1, 0.25)$
κ	α	ρ_d	ρ_φ	$\exp(\sigma_A)$	$\exp(\sigma_d)$
$N(4, 1.5)$	$N(0.3, 0.025)$	$Be(0.5, 0.2)$	$Be(0.5, 0.2)$	$IG(0.1, 2)$	$IG(0.1, 2)$
$\exp(\sigma_\varphi)$	$\exp(\sigma_\mu)$	$\exp(\sigma_e)$	$100\Lambda_\mu$	$100\Lambda_A$	
$IG(0.1, 2)$	$IG(0.1, 2)$	$IG(0.1, 2)$	$N(0.34, 0.1)$	$N(0.178, 0.075)$	

We generate 75,000 draws from the posterior using our Metropolis-Hastings, also after an exhaustive search for good initial parameters of the chain. In this case, the search is notably easier than in the first exercise, showing the usefulness of the prior in achieving identification and a good behavior of the simulation. Table 5 reports the posterior medians, and the 5 and 95 percentile values of the 23 estimated parameters of the model, while Figure 2 plots the histograms of each parameter.

Table 5: Median Estimated Parameters (5 and 95 percentile in parentheses)					
β	h	ψ	ϑ	κ	α
0.998 [0.997,0.999]	0.97 [0.95,0.98]	8.92 [4.09,13.84]	1.17 [0.74,1.61]	9.51 [7.47,11.39]	0.21 [0.17,0.26]
θ_p	χ	θ_w	χ_w	γ_R	γ_y
0.82 [0.78,0.87]	0.63 [0.46,0.79]	0.68 [0.62,0.73]	0.62 [0.44,0.79]	0.77 [0.74,0.81]	0.19 [0.13,0.27]
γ_π	Π	ρ_d	ρ_φ	σ_A	σ_d
1.29 [1.02,1.47]	1.010 [1.008,1.011]	0.12 [0.04,0.22]	0.93 [0.89,0.96]	-3.97 [-4.17,-3.78]	-1.51 [-1.82,-1.11]
σ_φ	σ_μ	σ_e	Λ_μ	Λ_A	
-2.36 [-2.76,-1.74]	-5.43 [-5.52,-5.35]	-5.85 [-5.94,-5.74]	$3.4e-3$ [0.003,0.004]	$2.8e-3$ [0.002,0.004]	

[FIGURE 2 HERE]

The value of the discount factor, β, goes very close to 1. This result is typical in DSGE models. The growth rate of the economy and the log utility function for consumption induce a high risk free real interest rate even without any discounting. This result is difficult to reconcile with an observed low average real interest rate in the U.S. Hence, the likelihood wants to push β as close as possible to 1 to avoid an even worse fit of the data. We find a quite high level of habit, 0.97, which diverges from other findings in the literature that hover around 0.7. The Frisch elasticity of labor supply of 0.85 (1/1.17) is similar to the one we fixed in the first estimation exercise, a reassuring result since DSGE models have often been criticized for relying on implausible high Frisch elasticities.

The adjustment cost of investment, κ, is high, 9.51, although lower than the value we fixed in the first exercise. Since our specification of adjustment costs was parsimonious, this result hints at the importance of further research on the nature of investment plans by firms. The parameter α is centered around 0.2. This result, which coincides with the findings of Smets and Wouters, is lower than in other estimates, but it is hard to interpret because the presence of monopolistic competition complicates the mapping between this parameter and observed income shares in the national income and product accounts.

The Calvo parameter for price adjustment, θ_p, is 0.82 (an average five-quarter pricing cycle) and the indexation level χ is 0.63, although the posterior for this parameter is quite spread. The Calvo parameter for wage adjustment, θ_w, is 0.68 (wage decisions are revised every three quarters), while the indexation, χ_w, is 0.62. We see how the data push us considerably away from the prior. There is information to be learned from the observations. However, at the same time, the median of the posterior of the nominal rigidities parameters is relatively different from the median in the case with flat priors. This shows how more informative priors *do* have an impact on our inference, in particular in parameter values of key importance for policy analysis as θ_p and θ_w. We do not necessarily judge this impact in a negative way. If the prior is bringing useful information into the table (for example, from micro data on individual firms' price changes, as in Bils and Klenow, 2004), this is precisely what we want. Nevertheless, the researcher and the final user of the model must be aware of this fact.

The parameters of the policy rule $\left\{\gamma_R, \gamma_\Pi, \gamma_y, \Pi\right\}$ are standard, with the only significant difference that now the Fed is less responsive to inflation, with the coefficient $\gamma_\pi = 1.29$. However, this value still respects the Taylor principle (even the value at the 5 percentile does). The Fed smooths the interest rate over time (γ_R is estimated to be 0.79) and responds actively to inflation (γ_R is 1.25) and weakly to the output growth gap (γ_y is 0.19). We estimate that the Fed has a target for quarterly inflation of 1 percent.

The growth rates of the investment-specific technological change, Λ_μ, and of the neutral technology, Λ_A, are roughly equal. The estimated average growth rate of the economy in per capita terms, $(\Lambda_A + \alpha\Lambda_\mu) / (1-\alpha)$ is 0.43 percent per quarter, or 1.7 percent annually, roughly the historical mean in the U.S. in the period 1865-2007.

4. Lines of Further Research

While the previous pages offered a snapshot of where the frontier of the research is, we want to conclude by highlighting three avenues for further research. Our enumeration is only a small sampler from the large menu of items to be explored in the New Macroeconometrics. We could easily write a whole chapter just signaling areas of investigation where entrepreneurial economists can spend decades and yet leave most of the territory uncharted.

First, we are particularly interested in the exploration of "exotic preferences" that go beyond the standard expected utility function used in this chapter (Backus, Routledge, and Zin, 2005). There is much evidence that expected utility functions have problems accounting for many aspects of the behavior of economic agents, in particular those aspects determining asset prices. For example, expected utility cannot explain the time inconsistencies that we repeatedly see at the individual level (how long did you stick with your last diet?). Recently, Bansal and Yaron (2004) and Hansen and Sargent (2007) have shown that many of the asset pricing puzzles can be accounted for if we depart from expected utility. However, these authors have used little formal econometrics and we do not know which of the different alternatives to expected utility will explain the data better. Therefore, the estimation of DSGE models with "exotic preferences" seems an area where the interaction between empirical work and theory is particularly promising.

Second, we would like to relax some of the tight parametric restrictions of the model, preferably within a Bayesian framework. DSGE models require many auxiliary parametric assumptions that are not central to the theory. Unfortunately, many of these parametric assumptions do not have a sound empirical foundation. But picking the wrong parametric form may have horrible consequences for the estimation of dynamic models. These concerns motivate the study of how to estimate DSGE models combining parametric and nonparametric components and, hence, conducting inference that is more robust to auxiliary assumptions.

Finally, a most exciting frontier is the integration of microeconomic heterogeneity within estimated DSGE models. James Heckman and many other econometricians have shown beyond reasonable doubt that individual heterogeneity is the defining feature of micro data (see Browning, Hansen, and Heckman, 1999). Our macro models need to move away from the basic representative agent paradigm and include richer configurations. Of course, this raises the difficult challenge of how to effectively estimate these economies, since the computation of the equilibrium dynamics of the model is a challenge by itself. However, we are optimistic: advances in computational power and methods have allowed us to do things that were unthinkable a decade ago. We do not see any strong reasons why estimating DSGE models with heterogeneous agents will not be any different in a few years.

5. Appendix A: Broader Context and Background

5.1. Policy-Making Institutions and DSGE Models

We mentioned in the main text that numerous policy institutions are using estimated DSGE models as an important input for their decisions. Without being exhaustive, we mention the Federal Reserve Board (Erceg, Guerrieri, and Gust, 2006), the European Central Bank (Christoffel, Coenen, and Warne, 2007), the Bank of Canada (Murchison and Rennison, 2006), the Bank of England (Harrison *et al.*, 2005), the Bank of Sweden (Adolfson *et al.*, 2005), the Bank of Finland (Kilponen and Ripatti, 2006 and Kortelainen, 2002), and the Bank of Spain (Andrés, Burriel, and Estrada, 2006), among several others.

5.2. Sequential Monte Carlo Methods and DSGE Models

The solution of DSGE models can be written in terms of a state space representation. Often, that state space representation is nonlinear and/or non-Gaussian. There are many possible reasons for this. For example, we may want to capture issues, such as asymmetries, threshold effects, big shocks, or policy regime switching that are approximated very poorly (or not at all) by a linear solution. Also, one of the key issues in macroeconomics is the time-varying volatility in time series. McConnell and Pérez-Quirós (2000) and Kim and Nelson (1998) presented extremely definitive evidence that the U.S. economy had been more stable in the 1980s and 1990s than before. Fernández-Villaverde and Rubio-Ramírez (2007) and Justiniano and Primiceri (2008) demonstrated that DSGE models, properly augmented with non-Gaussian shocks, could account for this evidence.

Nonlinear and/or non-Gaussian state space representations complicate the filtering problem. All the relevant conditional distributions of states became nonnormal and, except for a few cases, we cannot resort to analytic methods to track them. Similarly, the curse of dimensionality of numerical integration precludes the use of quadrature methods to compute the relevant integrals of filtering (those that appear in the Chapman-Kolmogorov formula and in Bayes' theorem).

Fortunately, during the 1990s, a new set of tools was developed to handle this problem. This set of tools has become to be known as sequential Monte Carlo (SMC) methods because they use simulations period by period in the sample. The interested reader can see a general introduction to SMC methods in Arulampalam *et al.* (2002), a collection of background material and applications in Doucet, de Freitas, and Gordon (2001), and the first applications in macroeconomics in Fernández-Villaverde and Rubio-Ramírez (2005 and 2007). The appendix in that last paper also offers references to alternative approaches.

The simplest SMC method is the Particle filter, which can easily be applied to estimate our

DSGE model. This filter replaces the conditional distribution of states $\{p(S_t|\mathbb{Y}_{1:T-1};\Psi)\}_{t=1}^{T}$ by an empirical distribution of N draws $\left\{\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}\right\}_{t=1}^{T}$ from the sequence $\{p(S_t|\mathbb{Y}_{1:T-1};\Psi)\}_{t=1}^{T}$. These draws are generated by simulation. Then, by a trivial application of the law of large numbers:

$$\mathcal{L}(\mathbb{Y}_{1:T};\Psi) \simeq \frac{1}{N}\sum_{i=1}^{N}\mathcal{L}\left(\mathbb{Y}_1|s_{0|0}^{i};\Psi\right)\prod_{t=2}^{T}\frac{1}{N}\sum_{i=1}^{N}\mathcal{L}\left(\mathbb{Y}_t|s_{t|t-1}^{i};\Psi\right)$$

where the subindex tracks the conditioning set (i.e., $t|t-1$ means a draw at moment t conditional on information until $t-1$) and lower cases are realizations of a random variable.

The problem is then to draw from the conditional distributions $\{p(S_t|\mathbb{Y}_{1:T-1};\Psi)\}_{t=1}^{T}$. Rubin (1988) first noticed that such drawing can be done efficiently by an application of sequential sampling:

Proposition 1. *Let* $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ *be a draw from* $p(S_t|\mathbb{Y}_{1:T-1};\Psi)$*. Let the sequence* $\{\widetilde{s}_t^{i}\}_{i=1}^{N}$ *be a draw with replacement from* $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ *where the resampling probability is given by*

$$q_t^i = \frac{\mathcal{L}\left(\mathbb{Y}_t|s_{t|t-1}^{i};\Psi\right)}{\sum_{i=1}^{N}\mathcal{L}\left(\mathbb{Y}_t|s_{t|t-1}^{i};\Psi\right)},$$

Then $\{\widetilde{s}_t^{i}\}_{i=1}^{N}$ *is a draw from* $p(S_t|\mathbb{Y}_{1:T};\Psi)$*.*

Proposition 1 recursively uses a draw $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ from $p(S_t|\mathbb{Y}_{1:T-1};\Psi)$ to draw $\left\{s_{t|t}^{i}\right\}_{i=1}^{N}$ from $p(S_t|\mathbb{Y}_{1:T};\Psi)$. But this is nothing more than the update of our estimate of S_t to add the information on y_t that Bayes' theorem is asking for. Resampling ensures that this update is done in an efficient way.

Once we have $\left\{s_{t|t}^{i}\right\}_{i=1}^{N}$, we draw N vectors of the 5 exogenous innovations in the DSGE model (the two shocks to productivity, the two shocks to technology, and the monetary policy shock) from the corresponding distributions and apply the law of motion for states to generate $\left\{s_{t+1|t}^{i}\right\}_{i=1}^{N}$. This step, known as forecast, puts us back at the beginning of proposition 1, but with the difference that we have moved forward one period in our conditioning, implementing in that way the Chapman-Kolmogorov equation. By going through the sample repeating these steps, we complete the evaluation of the likelihood function. Künsch (2005) provides general conditions for the consistency of this estimator of the likelihood function and for a central limit theorem to apply.

The following pseudo-code summarizes the description of the algorithm:

Step 0, Initialization: Set $t \rightsquigarrow 1$. Sample N values $\left\{s_{0|0}^{i}\right\}_{i=1}^{N}$ from $p\left(S_0;\Psi\right)$.

Step 1, Prediction: Sample N values $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ using $\left\{s_{t-1|t-1}^{i}\right\}_{i=1}^{N}$, the law of motion for states and the distribution of shocks ε_t.

Step 2, Filtering: Assign to each draw $\left(s_{t|t-1}^{i}\right)$ the weight ω_t^i in Proposition 1.

Step 3, Sampling: Sample N times with replacement from $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ using the probabilities $\{q_t^i\}_{i=1}^{N}$. Call each draw $\left(s_{t|t}^{i}\right)$. If $t < T$ set $t \rightsquigarrow t+1$ and go to step 1. Otherwise stop.

6. Appendix B: Model and Computation

6.1. Computation of the Model

A feature of the New Macroeconometrics that some readers from outside economics may find less familiar is that the researcher does not specify some functional forms to be directly estimated. Instead, the economist postulates an environment with different agents, technology, preferences, information, and shocks. Then, she concentrates on investigating the equilibrium dynamics of this environment and how to map the dynamics into observables. In other words, economists are not satisfied with describing behavior, they want to explain it from first principles. Therefore, a necessary first step is to solve for the equilibrium of the model given arbitrary parameter values. Once we have that equilibrium, we can build the associated estimating function (the likelihood, some moments, etc.) and apply data to perform inference. Thus, not only is finding the equilibrium of the model of key importance, but doing it rapidly and accurately, since we may need to do it for many different combinations of parameter values (for example, in an McMc simulation, we need to solve the model for each draw of parameter values).

As we described in the main text, our solution algorithm for the model relies on the perturbation of the equations that characterize the equilibrium of the model given some fixed parameter values. The first step of the perturbation is to take partial derivatives of these equations with respect to the states and control variables. This step, even if conceptually simple, is rather cumbersome and requires an inordinate amount of algebra.

Our favorite approach for solving this step is to write a `Mathematica` program that generates the required analytic derivatives (which are in the range of several thousands) and that writes automatic Fortran 95 code with the corresponding analytic expressions. This step is crucial because, once we have paid the fixed cost of taking the analytic derivatives (which takes several hours on a good desktop computer), solving the equilibrium dynamics for a new combination of parameter values takes less than a second, since now we only need to solve a numerical system in Fortran. The solution algorithm will be nested inside a Metropolis-Hastings. For each proposed parameter value in the chain, we will solve the model, find the equilibrium dynamics, use those dynamics to find the likelihood using the Kalman filter, and then accept or reject the proposal.

6.2. Kalman Filter

The implementation of the Kalman filter to evaluate the likelihood function in a DSGE model like ours follows closely that of Stengel (1994). We start by writing the first order linear approximation to the solution of the model in a standard state space representation:

$$\begin{aligned} s_t &= As_{t-1} + B\varepsilon_t && (9) \\ y_t &= Cs_t + D\varepsilon_t && (10) \end{aligned}$$

where s_t are the states of the model, y_t are the observables, and $\varepsilon_t \sim \mathbf{N}(0, I)$ is the vector of innovations to the model.

We introduce some definitions. Let $x_{t|t-1} = \mathbb{E}(x_t|Y_{t-1})$ and $x_{t|t} = \mathbb{E}(x_t|Y_t)$ where $Y_t = \{y_1, y_2, ..., y_t\}$. Also, we have $P_{t-1|t-1} = \mathbb{E}\left(s_{t-1} - s_{t-1|t-1}\right)\left(s_{t-1} - s_{t-1|t-1}\right)'$ and $P_{t|t-1} = \mathbb{E}\left(s_{t-1} - s_{t|t-1}\right)\left(s_{t-1} - s_{t|t-1}\right)'$.With this notation, the one-step-ahead forecast error is $\eta_t = y_t - Cx_{t|t-1}$.

We forecast the evolution of states:

$$s_{t|t-1} = Ax_{t-1|t-1} \tag{11}$$

Since the possible presence of correlation in the innovations does not change the nature of the filter, it is still the case that

$$s_{t|t} = s_{t|t-1} + K\eta_t, \tag{12}$$

where K is the Kalman gain at time t. Define variance of forecast as $V_y = CP_{t|t-1}C' + DD'$.

Then, the conditional likelihood is just:

$$\log \mathcal{L} = -\frac{n}{2}\log 2\pi - \frac{1}{2}\log\det\left(V_y\right) - \frac{1}{2}\eta_t V_y^{-1}\eta_t$$

The last step is to update our estimates of the states. Define residuals $\xi_{t|t-1} = s_t - s_{t|t-1}$

and $\xi_{t|t} = s_t - s_{t|t}$. Subtracting equation (11) from equation (9)

$$\begin{aligned} s_t - s_{t|s-1} &= A\left(s_{t-1} - s_{t-1|t-1}\right) + Bw_t, \\ \xi_{t|t-1} &= A\xi_{t-1|t-1} + Bw_t \end{aligned} \tag{13}$$

Now subtract equation (12) from equation (9)

$$\begin{aligned} s_t - s_{t|t} &= s_t - s_{t|t-1} - K\left[Cx_t + Dw_t - Cx_{t|t-1}\right] \\ \xi_{t|t} &= \xi_{t|t-1} - K\left[C\xi_{t|t-1} + Dw_t\right]. \end{aligned} \tag{14}$$

Note $P_{t|t-1}$ can be written as

$$\begin{aligned} P_{t|t-1} &= \mathbb{E}\xi_{t|t-1}\xi'_{t|t-1}, \\ &= \mathbb{E}\left(A\xi_{t-1|t-1} + Bw_t\right)\left(A\xi_{t-1|t-1} + Bw_t\right)' \\ &= AP_{t-1|t-1}A' + BB'. \end{aligned} \tag{15}$$

For $P_{t|t}$ we have:

$$\begin{aligned} P_{t|t} &= \mathbb{E}\xi_{t|t}\xi'_{t|t} \\ &= \left(I - KC\right)P_{t|t-1}\left(I - C'K'\right) + KDD'K' - KDB' \\ &\quad -BD'K' + KCBD'K' + KDB'C'K'. \end{aligned} \tag{16}$$

The optimal gain minimizes $P_{t|t}$:

$$K_{opt} = \left[P_{t|t-1}C' + BD'\right]\left[Vy + CBD' + DB'C'\right]^{-1}$$

and, consequently, the updating equations are:

$$\begin{aligned} P_{t|t} &= P_{t|t-1} - K_{opt}\left[DB' + CP_{t|t-1}\right], \\ s_{t|t} &= s_{t|t-1} + K_{opt}\eta_t \end{aligned}$$

and we close the iterations.

6.3. Construction of Data

When we estimate the model, we need to make the series provided by the national and income product accounts (NIPA) consistent with the definition of variables in the theory. The main adjustment that we undertake is to express both real output and real gross investment in consumption units. Our DSGE model implies that there is a numeraire in terms of which all the other prices need to be quoted. We pick consumption as the numeraire. The NIPA, in comparison, uses an index of all prices to transform nominal GDP and investment into real values. In the presence of changing relative prices, such as the ones we have seen in the U.S. over the last several decades with the fall in the relative price of capital, NIPA's procedure biases the valuation of different series in real terms.

We map theory into data by computing our own series of real output and real investment. To do so, we use the relative price of investment, defined as the ratio of an investment deflator and a deflator for consumption. The denominator is easily derived from the deflators of nondurable goods and services reported in the NIPA. It is more complicated to obtain the numerator because, historically, NIPA investment deflators were poorly constructed. Instead, we rely on the investment deflator computed by Fisher (2006), a series that unfortunately ends early in 2000Q4. Following Fisher's methodology, we have extended the series to 2007Q1.

For the real output per capita series, we first define nominal output as nominal consumption plus nominal gross investment. We define nominal consumption as the sum of personal consumption expenditures on nondurable goods and services. We define nominal gross investment as the sum of personal consumption expenditures on durable goods, private residential investment, and nonresidential fixed investment. Per capita nominal output is equal to the ratio between our nominal output series and the civilian noninstitutional population between 16 and 65. To obtain per capita values, we divide the previous series by the civilian noninstitutional population between 16 and 65. Finally, real wages are defined as compensation per hour in the nonfarm business sector divided by the CPI deflator.

References

[1] Adolfson, M., S. Laséen, J. Lindé, and M. Villani (2005). "Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through." *Sveriges Riksbank Working Paper Series* 179.

[2] An, S. and F. Schorfheide (2006). "Bayesian Analysis of DSGE Models." *Econometric Reviews* 26, 113-172.

[3] Andrés, J., P. Burriel, and A. Estrada (2006). "BEMOD: a DSGE Model for the Spanish Economy and the Rest of the Euro Area." *Documento de Trabajo del Banco de España* 0631.

[4] Aruoba, S.B., J. Fernández-Villaverde, and J. Rubio-Ramírez (2006). "Comparing Solution Methods for Dynamic Equilibrium Economies." *Journal of Economic Dynamics and Control* 30, 2477-2508.

[5] Arulampalam, A.S., S. Maskell, N. Gordon, and T. Clapp (2002). "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking." *IEEE Transactions on Signal Processing* 50, 174-188.

[6] Backus, D.K., B.R. Routledge, and S.E. Zin (2005). "Exotic Preferences for Macroeconomists." *NBER Macroeconomics Annual* 2004, 319-390.

[7] Bansal, R. and A. Yaron (2004). "Risks For The Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance* 59, 1481-1509

[8] Bils, M. and P. Klenow (2004). "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy* 112, 947-985.

[9] Browning, M., L.P. Hansen, and J.J. Heckman (1999). "Micro Data and General Equilibrium Models." in J.B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics* 1, 543-633. Elsevier.

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

[11] Canova, F., and L. Sala (2006). "Back to Square One: Identification Issues in DSGE Models." *Mimeo*, Pompeu Fabra University.

[12] Caplin, A. and J. Leahy (1991). "State-Dependent Pricing and the Dynamics of Money and Output." *Quarterly Journal of Economics*, 106, 683-708.

[13] Chernozhukov, V., and H. Hong (2003). "A MCMC approach to classical estimation." *Journal of Econometrics* 115(2), 293–346.

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

[15] Christoffel, K, G. Coenen, and A. Warne (2007). "The New Area-Wide Model of the Euro Area: Specification and First Estimation Results." *Mimeo*, European Central Bank.

[16] Dotsey, M., R.G. King, and A. Wolman (1999). "State Dependent Pricing and the General Equilibrium Dynamics of Money and Output." *Quarterly Journal of Economics* 655-690.

[17] Doucet, A., N. de Freitas, and N. Gordon (2001). *Sequential Monte Carlo Methods in Practice*. Springer Verlag.

[18] Edge, R., M. Kiley, and J.P. Laforte (2008). "A Comparison of Forecast Performance Between Federal Reserve Staff Forecasts, Simple Reduced-Form Models, and a DSGE Model." *Mimeo*, Federal Reserve Board.

[19] Erceg, C.J., L. Guerrieri, and C. Gust (2006). "SIGMA: A New Open Economy Model for Policy Analysis." *International Journal of Central Banking* 2, 1-50.

[20] Fernández-Villaverde, J. and J. Rubio-Ramírez (2004). "Comparing Dynamic Equilibrium Models to Data: a Bayesian Approach." *Journal of Econometrics* 123*,* 153-187*.*

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

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

[23] Fernandez-Villaverde, J. and J. Rubio-Ramirez (2008). "How Structural are Structural Parameters?" *NBER Macroeconomics Annual* 2007, 83-137.

[24] Fernández-Villaverde, J., J. Rubio-Ramírez, and M.S. Santos (2006). "Convergence Properties of the Likelihood of Computed Dynamic Models." *Econometrica* 74, 93-119.

[25] Fisher, J., 2006. "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks." *Journal of Political Economy* 114, 413-52.

[26] Greenwood, J, Z. Herkowitz, and P. Krusell (1997). "Long-Run Implications of Investment-Specific Technological Change." *American Economic Review* 87, 342-362.

[27] Guerron-Quintana P. (2008). "What you match does matter: The Effects of Data on DSGE Estimation." *Working Paper*, North Carolina State University.

[28] Judd, K.L. (1998). *Numerical Methods in Economics.* MIT Press.

[29] Justiniano A. and G.E. Primiceri (2008). "The Time Varying Volatility of Macroeconomic Fluctuations." *American Economic Review* 98, 604-641.

[30] Hansen, L.P. (1982). "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50, 1029-54.

[31] Hansen, L.P. and T. Sargent (2007). *Robustness*. Princeton University Press, Princeton.

[32] Harrison, R., K. Nikolov, M. Quinn, G. Ramsey, A. Scott, and R. Thomas (2005). *The Bank of England Quarterly Model.* The Bank of England.

[33] Kim, C. and C.R. Nelson (1998). "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle". *Review of Economics and Statistics* 81, 608-616.

[34] Kilponen, J. and A. Ripatti (2006). "Introduction to AINO." *Mimeo*, Bank of Findland.

[35] Kortelainen, Mika (2002). "EDGE: a Model of the Euro Area with Applications to Monetary Policy.", Bank of Finland Studies E:23.

[36] Künsch, H.R. (2005). "Recursive Monte Carlo Filters: Algorithms and Theoretical Analysis." *Annals of Statistics* 33, 1983-2021.

[37] Kydland, F., and E.C. Prescott (1982). "Time to Build and Aggregate Fluctuations." *Econometrica* 50, 1345–1370.

[38] Lubick, T. and F. Schorfheide (2004). "Testing for Indeterminacy: An Application to U.S. Monetary Policy." *American Economic Review* 94, 190-217.

[39] Lucas, R. (1972). "Expectations and the Neutrality of Money" *Journal of Economic Theory* 4, 103-124.

[40] Lucas, R. (1976). "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy* 1, 19–46.

[41] McConnell, M.M. and G. Pérez-Quirós (2000). "Output Fluctuations in the United States: What Has Changed Since the Early 1980's?" *American Economic Review* 90, 1464-1476.

[42] Murchison, S. and A. Rennison (2006). "ToTEM: The Bank of Canada's New Quarterly Projection Model." *Bank of Canada Technical Report* 97.

[43] Muth, J. F. (1961). "Rational Expectations and the Theory of Price Movements." *Econometrica* 29, 315-335.

[44] Orphanides, A. (2002). "Monetary Policy Rules and the Great Inflation." *American Economic Review* 92, 115-120.

[45] Roberts, G., A. Gelman, and W. Gilks (1997). "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorthims." *Annals of Applied Probability* 7, 110-120.

[46] Rubin, D.B. (1988). "Using the SIR Algorithm to Simulate Posterior Distributions". in J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith (eds.), *Bayesian Statistics 3*, 395-402, Oxford University Press.

[47] Sargent, T.J. (2005). "An Interview with Thomas J. Sargent by George W. Evans and Seppo Honkapohja." *Macroeconomic Dynamics* 9, 561-583.

[48] Sims, C.A. (2002). "Implications of Rational Inattention." *Mimeo*, Princeton University.

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

[50] Smith, A. (1776). *An Inquiry into the Nature and Cause of the Wealth of Nations.* Published by Clanderon Press (1979), Oxford, Glasgow Edition of the Works and Correspondence of Adam Smith, 2.

[51] Stengel, R.F. (1994). "Optimal Control and Estimation." Dover Publications, New York.

[52] Walras, L. (1874). Elements of Pure Economics. Republished by Harvard University Press (1954), Cambridge.

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

Figure 1: Posterior Distribution, Flat Priors

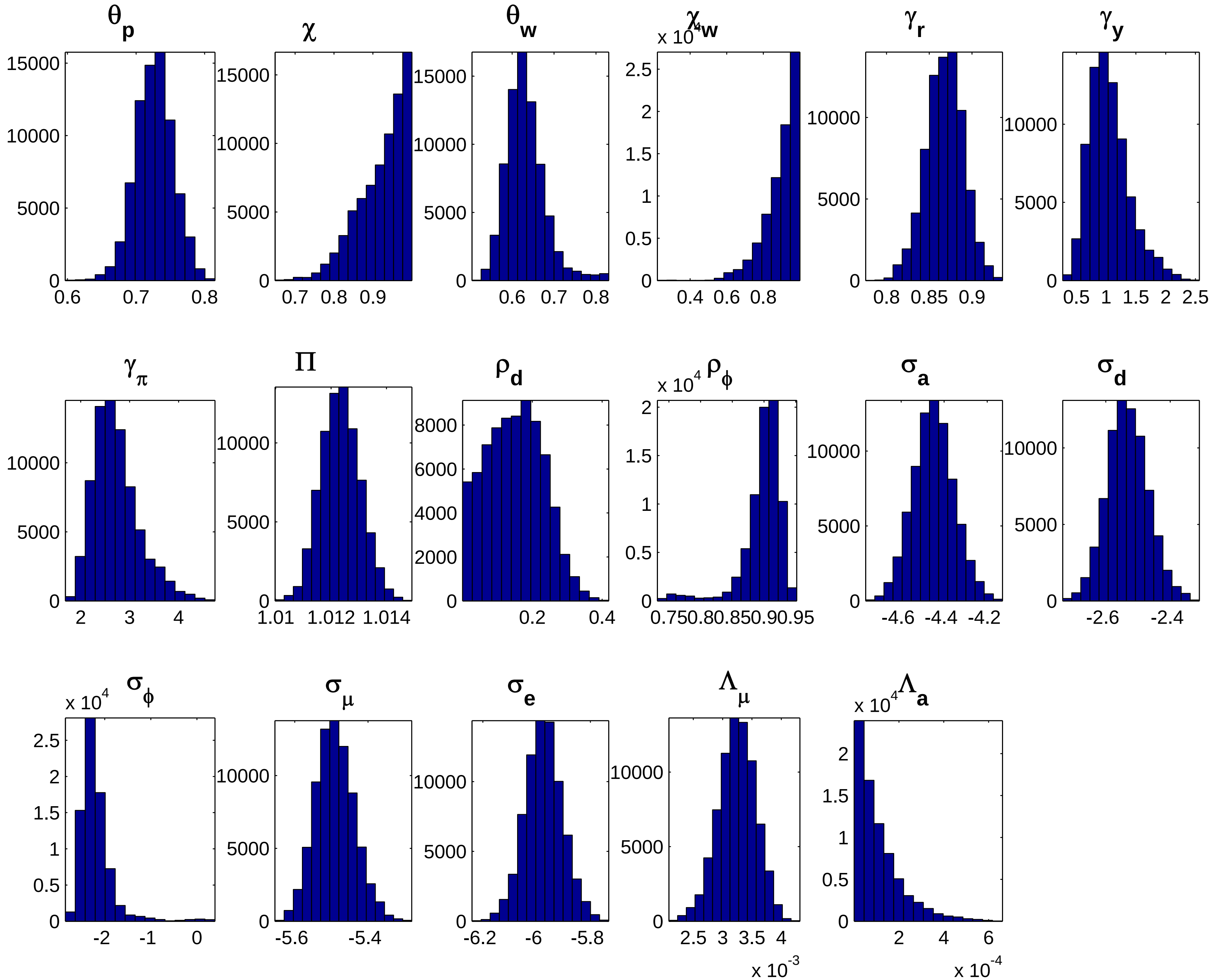

Figure 2: Posterior Distribution, Smets-Wouters Priors