Volume Title: NBER Macroeconomics Annual 2007, Volume 22

Volume Author/Editor: Daron Acemoglu, Kenneth Rogoff and Michael Woodford, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 978-0-226-00202-6

Volume URL: http://www.nber.org/books/acem07-1

Conference Date: March 30-31, 2007

Publication Date: June 2008

Chapter Title: How Structural Are Structural Parameters?

Chapter Author: Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez

Chapter URL: http://www.nber.org/chapters/c4087

Chapter pages in book: (p. 83 - 137)

2

How Structural Are Structural Parameters?

Jesús Fernández-Villaverde, *University of Pennsylvania, NBER, and CEPR*
Juan F. Rubio-Ramírez, *Duke University and Federal Reserve Bank of Atlanta*

1 Introduction

This paper studies the following problem: how stable over time are the so-called "structural parameters" of dynamic stochastic general equilibrium (DSGE) models? To answer this question, we estimate a medium-scale DSGE model with real and nominal rigidities, using U.S. data. In our model, we allow for parameter drifting and rational expectations of the agents with respect to this drift. We document that there is strong evidence that parameters change within our sample. In particular, we illustrate variations in the parameters describing the monetary policy reaction function and in the parameters characterizing the pricing behavior of firms and households. Moreover, we show how the movements in the pricing parameters are correlated with inflation. Thus, our results cast doubts on the empirical relevance of Calvo models.

Our findings are important because DSGE models are at the core of modern macroeconomics. They promise to be a laboratory that researchers can employ to match theory with reality, to design economic policy, and to evaluate welfare. The allure of DSGE models has captured the imagination of many, inside and outside academia. In universities, a multitude of economists implement DSGE models in their rich varieties and fashions. More remarkable still, a burgeoning number of policy-making institutions are estimating DSGE models for policy analysis and forecasting. 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 Sweden (Adolfson, Laséen, Lindé, and Villani 2005), and the Bank of Spain (Andrés, Burriel and Estrada 2006) are at the leading edge of the tide, but a dozen other institutions are jumping on the bandwagon. In addition, economists are accumulating experience with the good forecasting record of DSGE models, even when compared with judgmental predictions from staff economists (Christoffel, Coenen, and Warne, 2007).

At the center of DSGE models we have the "structural parameters" that define the preferences and technology of the economy. We call these parameters "structural" in the sense of Hurwicz (1962): they are invariant to interventions, including shocks by nature. The structural character of the parameters is responsible for much of the appeal of DSGE models. Since the parameters are fully interpretable from the perspective of economic theory and invariant to policy interventions, DSGE models avoid the Lucas critique and can be used to quantitatively evaluate policy.

Our point of departure is that, at least at some level, it is hard to believe that the "structural parameters" of DSGE models are really structural, given the class of interventions we are interested in for policy analysis. Let us think, for instance, about technology. Most DSGE models specify a stable production function, perhaps subject to productivity growth. Except in a few papers (Young 2004), the features of the technology, like the elasticity of output to capital, are constant over time. But this constant elasticity is untenable in a world where technological change is purposeful. We can expect that changes in relative input prices will induce changes in the new technologies developed and that those may translate into different elasticities of output to inputs. Similar arguments can be made along nearly every dimension of a modern DSGE model.

The previous argument is not sufficient to dismiss the practice of estimating DSGE models with constant parameter values. Simplifying assumptions, like stable parameters, are required to make progress in economics. However, as soon as we realize the possible changing nature of structural parameters, we weaken the justifications for inference exercises underlying the program of DSGE modeling. The separation between what is "structural" and what is reduced form becomes much more ambiguous.[1]

The possibility but not the necessity of parameter drifting motivates the main question of this paper: how much evidence of parameter drifting in DSGE models is in the data? If the answer is that we find much support for drifting (where the metric to decide "much" needs to be discussed), we would have to reevaluate the usefulness of our estimation exercises, or at least modify them to account for parameter variation.

Moreover, parameter drifting may also be interpreted as a sign of model misspecification and, possibly, as a guide for improving our models. If the answer is negative—that is, if we find little parameter drifting—we would increase our confidence in DSGE models as a procedure to tackle relevant policy discussions.

Beyond addressing our substantive question, this paper also develops new tools for the estimation of dynamic equilibrium models with parameter drifting. We show how the combination of perturbation methods and the particle filter allows the efficient estimation of this class of economies. Indeed, all the required computations can be performed on an average personal computer in a reasonable amount of time. We hope that those tools may be put to good use in other applications, not necessarily in general equilibrium, that involve time-varying parameters in essential ways.

Our main results are as follows. First, we offer compelling proof of changing parameters in the Fed's behavior. Monetary policy became appreciably more aggressive in its stand against inflation after Volcker's appointment. This agrees with Clarida, Galí, and Gertler (2000), Lubick and Schorfheide (2004), Boivin (2006), and Rabanal (2007). Our contribution is to rederive the result within a model where agents understand and act upon the fact that monetary policy changes over time.

Second, we expose the instability of the parameters controlling the level of nominal rigidity and indexation of prices and wages. Those changes are strongly correlated with changes in inflation in an intuitive way: lower rigidities correlate with higher inflation and higher rigidities with lower inflation. Our finding suggests that a more thorough treatment of nominal rigidities, possibly through state-dependent pricing models, may yield a high payoff.

We want to be up front about the shortcomings of our exercise. First and foremost, we face the limitations of the data. With 184 quarterly observations of the U.S. economy, there is a tight bound on how much we can learn from the data (Ploberger and Phillips 2003, frame the problem of empirical limits for time series models precisely in terms of information bounds). The main consequence of the limitations of the short sample size is relatively imprecise estimates.

The second limitation, forcefully emphasized by Sims (2001), is that we do not allow for changing volatilities in the innovations of the model, which is itself a particular form of parameter drift. If the innovations in the U.S. data are heteroskedastic (as we report in Fernández-Villaverde and Rubio-Ramírez 2007), the estimation may attempt to pick up the changing variance by spurious changes in the structural parameters. At the same time, Cogley and Sargent (2005) defend that there is still variation in the parameters of a vector autoregression (VAR), even after controlling for heteroskedasticity. We are currently working on an extension of the model with both parameter drifting and changing volatilities.

We build upon an illustrious tradition of estimating models with parameter drifting. One classic reference is Cooley and Prescott (1976), where the authors studied the estimation of regression parameters that are subject to permanent and transitory shocks. Unfortunately, the techniques in this tradition are within the context of the Cowles Commission's framework and, hence, are of little direct application to our investigation.

Our paper is also linked with a growing body of research that shows signs of parameter drifting on dynamic models. Since the estimation of this class of models is a new undertaking, the evidence is scattered. One relevant literature estimates VARs with time-varying parameters and/or stochastic volatility. Examples include Uhlig (1997), Bernanke and Mihov (1998), Cogley and Sargent (2005), Primiceri (2005), and Sims and Zha (2006). The consensus emerging from these papers is that there is evidence of time variation in the parameters of a VAR, although there is a dispute about whether the variation comes from changes in the autoregressive components or from stochastic volatility. This evidence, however, is only suggestive, since a DSGE model with constant parameters may be compatible with a time-varying VAR (Cogley and Sbordone 2006).

A second group of studies has estimated equilibrium models with parameter variation, but it has been less ambitious in the extent of the fluctuations studied. Fernández-Villaverde and Rubio-Ramírez (2007) and Justiniano and Primiceri (2005) demonstrate the importance of stochastic volatility to account for U.S. data using a DSGE model. King (2006) works with a simple real business cycle (RBC) economy with parameter drift in four parameters. However, his approach relies on particular properties of his model and it is too cumbersome to be of general applicability. Canova (2005) estimates a small-scale New Keynesian model with parameter drifting but without the agents being aware of these changes in the parameters. He uncovers important movements in the parameters that enter into the Phillips curve and the Euler equations. Boivin (2006) estimates a parameter-drifting Taylor rule with real-time data. He corroborates previous findings of changes in the rule coefficients obtained with final data. Benati (2006), elaborating on an argument by Woodford

(2006), questions the indexation mechanisms introduced in New Keynesian models and shows that they are not structural to changes in monetary policy rules. Oliner, Rudebusch, and Sichel (1996) find unstable parameters even investment models with more intricate representations of capital spending than those found in current DSGE models. Owyang and Ramey (2004) estimate regime-switching models of monetary policy and identify the evolving preferences of the monetary authority through their interaction with the structural parameters.

There are also numerous papers that tell us about parameter drifting, albeit in an indirect way. A common practice when estimating models has been to divide the sample into two periods, usually before and after 1979, and argue that there are significant differences in the inference results. One celebrated representative of this method is Clarida, Galí, and Gertler (2000), a paper we will discuss later.

Finally, a literature that has connections with our analysis is the one that deals with DSGE models with a Markov-switching process in different aspects of the environment, like monetary or fiscal policy (Davig and Leeper 2006a and 2006b, Chung, Davig, and Leeper 2006, and Farmer, Waggoner, and Zha 2006). The stated motivation of these papers is that Markov switches help us understand the dynamics of the economy better. So far, none of these papers has produced an estimated model.

The rest of the article is organized as follows. First, in section 2, we discuss different ways to think about parameter drifting in dynamic equilibrium models. In section 3, we develop two simple examples of parameter drift that motivate our investigation. Section 4 defines a medium-scale model of the U.S. economy and discusses how to introduce this model to the data. Section 5 introduces parameter drifting and explains how to adapt the approach in section 4 to handle this situation. We report our results in section 6. Section 7 concludes. An appendix provides the interested reader with technical details.

2 Parameter Drifting and Dynamic Equilibrium Models

There are at least three ways to think about parameter drifting in an estimated DSGE model. The simplest approach, which we call the *pure econometric interpretation*, is to consider parameter drifting as a convenient phenomenon to fit the data better or as the consequence of a capricious nature that agents in the model neither understand nor forecast. Despite its simplicity, this interpretation violates the spirit of rational expectations: not having free parameters that the researcher can play with. Consequently, we will not investigate this case further.

The second way to think about parameter drifting is as a characteristic of the environment that the agents understand and act upon. Let us come back to our example of the production function. Imagine that the aggregate technology is given by a Cobb-Douglas function $Y_t = AK_t^{\alpha_t} L_t^{1-\alpha_t}$, where output Y_t is produced with capital K_t and labor L_t given a technology level A and share parameter α_t. The only difference with the standard environment is that α_t is indexed by time (neither the realism nor the empirical justification of our example is crucial for the argument, although we could argue in favor of both features). Let us also assume that α_t evolves over time as a random walk with reflecting boundaries at 0 and 1, to ensure that the production function satisfies the usual properties. We could imagine that such drift comes about because the new technologies developed have a random requirement of capital. The solution of the agents' problems are decision rules that have as one of their arguments the current α_t. Why? First, because α_t determines current prices. Second, because α_t helps to forecast future values α_{t+j} and hence to predict future prices. This interpretation is our favorite one, and it will frame our reading of the results in section 6.

The final perspective about parameter drifting is as a telltale of model misspecification. This point, raised by Cooley (1971) and Rosenberg (1968), is particularly cogent when estimating DSGE models. These models are complex constructions. To make them useful for policy purposes, researchers add many mechanisms that affect the dynamics of the economy: sticky prices and wages, adjustment costs, and so on. In addition, DSGE models require right parametric assumptions for the utility function, production function, adjustment costs, distribution of shocks, and so forth. If we seriously misspecified the model along at least one dimension, parameter drifting may appear as the only possibility left to the model to fit the data. Our example in section 3 illustrates this point in detail. We will exploit this possibility in our empirical results and assess how the drift in the parameters determining the degree of nominal rigidity in the economy implies that time-dependent models of pricing decisions may be flawed.

3 Two Examples

In this section, we present two simple examples that generate parameter drifting in estimated DSGE models. We have chosen the examples to illustrate our points as clearly as possible, and not based on their relevance or plausibility. However, the examples are not far-fetched: they deal with recurrent themes in the literature and are linked (albeit we do not explore this connection to its fullest) to relevant features of the economy.

3.1 *Parameter Drift as a Consequence of Changing Policies*

The first example deals with the changes in the elasticity of monetary policy to different variables. It is common to postulate that the monetary authority uses open market operations to set the short-run nominal interest rate R_t according to a Taylor rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R}\left[\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_\Pi}\left(\frac{y_t}{\hat{y}_t}\right)^{\gamma_y}\right]^{1-\gamma_R}\exp(\sigma_m \varepsilon_{mt})$$

The variable Π represents the target levels of inflation of the monetary authority, R the steady-state gross return of capital, y_t is output, and $\hat{y}_t$ a measure of target output. The term ε_{mt} is a random shock distributed according to $\mathcal{N}(0, 1)$.

In an influential contribution, Clarida, Galí, and Gertler (2000) attracted the attention of economists to changes in the elasticity parameter γ_Π before and after Volcker's appointment as Fed chairman in 1979. They document, with a slightly different specification of the Taylor rule, that γ_Π more than doubles after 1979. This finding has been corroborated in many studies and found resilient to modifications in the empirical specification (Lubick and Schorfheide 2004). The division of the sample between the time before and after 1979 has also been exploited by Boivin and Giannoni (2006), who find that the point estimates of the structural parameters also substantially vary between the two periods.

Changes in the policy coefficients are one particular example of parameter drift. They can be the consequence of the shifting priorities of the policymakers or, as emphasized by Sargent (1999), of changes in the perception of the effectiveness of monetary policy. Once we recognize that there is evidence of the parameter γ_Π drifting over time, it is natural to assume that agents are aware of the changes and act upon them. Such an environment may capture some of the insights of Sims (1980) about the difference between a change in policy regime (in our Taylor rule, a change in the way the interest rate is determined) and the evolution of the policy within one regime, which could be represented in our context as the drift of the parameters of the rule.

3.2 *Parameter Drift as a Telltale of Model Misspecification*

Our second example revisits several of the themes in Browning, Hansen, and Heckman (1999). We explore the consequences for inference of an econometrician estimating a model with infinitely lived agents when the data are actually generated by an overlapping-generations model. We show how our estimate of the discount factor will be a function of the true discount factor, the elasticity of output to capital, and the changing age distribution of the population. This example is relevant because variations in the age structure of the U.S. population have been continuous due to shifts in fertility and mortality.

3.2.1 An Artificial World We begin by creating a simple artificial world. In each period t, there are two generations of households alive, young and old. Each household maximizes the life utility

$$\log c_t^t + \beta \mathbb{E}_t \log c_{t+1}^t$$

where the superindex denotes that the household was born in period t, the subindex the period in which it consumes, and $\mathbb{E}_t$ is the conditional expectations operator. The discount factor, β, captures the preference for current consumption. We pick a log utility function to simplify the algebra that follows.

Households work when young and get a wage w_t for a unit of time that they supply inelastically. Households live off their savings when they are old. The period budget constraints are $c_t^t + s_t = w_t$ and $c_{t+1}^t = R_{t+1}s_t$, where s_t is the household savings and R_{t+1} the gross return on capital. From the first order condition of households, we have that $c_t^t = [1/(1+\beta)]w_t$ and $c_t^t = [\beta/(1+\beta)]w_t$.

In each period, a number n_t of new households is born. For the moment, we will assume only that l_t is the realization of some random process. Nothing of substance for our argument is lost by assuming that the size of the new generation is exogenous.

The production side of the economy is defined by a Cobb-Douglas function, $y_t = k_t^\alpha l_t^{1-\alpha}$, where k_t is the total amount of capital in the economy and l_t the total amount of labor. If we assume total depreciation in the economy, again to simplify the algebra, and impose the condition $l_t = n_t$, we get by competitive pricing $w_t = (1-\alpha)k_t^\alpha l_t^{n-\alpha}$ and $R_t = \alpha k_t^{\alpha-1} n_t^{1-\alpha}$.

All that remains is some accounting. Total consumption in the economy in period t, C_t, is equal to the consumption of the old generation plus the consumption of the young generation. The old consume all of their income, which is equal to the capital income of the economy, $R_t k_t = \alpha k_t^{\alpha} n_t^{1-\alpha}$. The young consume a fraction $[1/(1 + \beta)]$ of their income, which is equal to the labor income of the economy $w_t l_t = (1 - \alpha) k_t^{\alpha} n_t^{1-\alpha}$. Then total consumption is:

$$C_t = \frac{1 + \alpha\beta}{1 + \beta} k_t^{\alpha} n_t^{1-\alpha}$$

By the aggregate resource constraint, investment (or, equivalently, capital in period $t + 1$) is

$$I_t = k_{t+1} = \frac{(1 - \alpha)\beta}{1 + \alpha\beta} C_t$$

Finally, we find per capita consumption c_t^{pc} as:

$$c_t^{pc} = \frac{C_t}{n_t + n_{t-1}}$$

3.2.2 An Econometrician Let us now suppose that we have an econometrician who aims to estimate a model with a representative infinitely lived agent and T observations generated from our economy. To do so, the econometrician postulates that the agent has a utility function:

$$\max_{(c_t^{pc})} \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t \left[\prod_{i=0}^{t} (1 + \gamma_t) \right] \log c_t^{pc}$$

where γ_t is the (random) growth rate of the population between periods $t - 1$ and t:

$$1 + \gamma_t = \frac{n_t + n_{t-1}}{n_{t-1} + n_{t-2}}$$

and $\gamma_0 = 0$. This utility function is the same as in the canonical presentation of the RBC model in Cooley and Prescott (1995) except that the growth rate of the population is stochastic instead of constant. The production side of the economy is the same as before, $y_t = k_t^{\alpha} l_t^{1-\alpha}$. Thus, the only difference between the artificial world we have created and the model the econometrician estimates is that, instead of having two generations alive in each moment, the econometrician estimates a model with a representative agent.

What are the consequences on the estimated parameters? Imagine that the econometrician knows α and that the depreciation factor is 1. Then, a simple procedure to estimate the only remaining unknown parameter in the world, the discount factor β, is to build the population moment:

$$\frac{1}{c_t^{pc}} = \beta \mathbb{E}_t (1 + \gamma_{t+1}) \frac{R_{t+1}}{c_{t+1}^{pc}}$$

and substitute the expectation by the sample mean:

$$\hat{\beta}_T = \frac{\frac{1}{T-1} \sum_{t=0}^{T-1} \frac{1}{c_t^{pc}}}{\frac{1}{T-1} \sum_{t=0}^{T-1} (1 + \gamma_{t+1}) \frac{R_{t+1}}{c_{t+1}^{pc}}}$$

We study how this expression evolves over time. First, note that, by substituting the expressions found before, we get:

$$(1 + \gamma_{t+1}) \frac{R_{t+1}}{c_{t+1}^{pc}} = \frac{(n_{t+1} + n_t)^2}{n_t + n_{t-1}} \frac{\alpha}{1-\alpha} \frac{1+\beta}{\beta} \frac{1}{C_t}$$

Then:

$$\hat{\beta}_T = \beta \frac{1-\alpha}{\alpha} \frac{1}{1+\beta} \frac{\sum_{t=0}^{T-1} (n_t + n_{t-1}) \frac{1}{C_t}}{\sum_{t=0}^{T-1} \frac{(n_{t+1} + n_t)^2}{n_t + n_{t-1}} \frac{1}{C_t}}$$

We want to work on the previous expression. First, we substitute aggregate consumption for its value in terms of capital and labor:

$$\hat{\beta}_T = \beta \frac{1}{1+\beta} \frac{1-\alpha}{\alpha} \frac{\sum_{t=0}^{T-1} (n_t + n_{t-1}) \frac{1}{k_t^{\alpha} n_t^{1-\alpha}}}{\sum_{t=0}^{T-1} \frac{(n_{t+1} + n_t)^2}{n_t + n_{t-1}} \frac{1}{k_t^{\alpha} n_t^{1-\alpha}}}$$

The only remaining endogenous element in this equation is k_t. To eliminate it, we recursively substitute k_{t-i} to find:

$$k_t = \left[\frac{(1-\alpha)\beta}{1+\alpha\beta} n_{t-1}^{1-\alpha} \prod_{i=1}^{t-1} \left(\frac{(1-\alpha)\beta}{1+\alpha\beta} n_{t-1-i}^{1-\alpha} \right)^{\alpha i} \right] k_0^{\alpha t}$$

Then:

$$\hat{\beta}_T = \beta \frac{1}{1+\beta} \frac{1-\alpha}{\alpha} \frac{\sum_{t=0}^{T-1} \frac{n_t + n_{t-1}}{n_t^{1-\alpha}} \left\{ \left[n_{t-1}^{1-\alpha} \prod_{i=1}^{t-1} \left(\frac{(1-\alpha)\beta}{1+\alpha\beta} n_{t-1-i}^{1-\alpha} \right)^{\alpha i} \right] k_0^{\alpha t} \right\}^{-\alpha}}{\sum_{t=0}^{T-1} \frac{(n_{t+1} + n_t)^2}{(n_t + n_{t-1}) n_t^{1-\alpha}} \left\{ \left[n_{t-1}^{1-\alpha} \prod_{i=1}^{t-1} \left(\frac{(1-\alpha)\beta}{1+\alpha\beta} n_{t-1-i}^{1-\alpha} \right)^{\alpha i} \right] k_0^{\alpha t} \right\}^{-\alpha}}$$

which delivers a $\hat{\beta}_T$, which is biased and drifts over time according to the evolution of the population. This expression is composed of three parts. First, the true parameter, β, second the deterministic bias,

$$\frac{1}{1+\beta}\frac{1-\alpha}{\alpha}$$

and finally the term involving the n_ts and k_0, which fluctuates over time.

Without further structure on population growth over time, it is difficult to say much about $\hat{\beta}_T$. In the simple case where $\gamma_t = \gamma$ is constant, as $T \to \infty$, the only factor dominating is:

$$\hat{\beta}_T \simeq \beta\frac{1}{1+\beta}\frac{1-\alpha}{\alpha}(1+\gamma)^{-2} \tag{1}$$

To explore the behavior of $\hat{\beta}_T$ in the general case where γ_t varies, we simulate the model and estimate the parameter recursively with data from an economy with $\alpha = 0.3$ and $\beta = 0.96$. The growth rates of population are 2, 4, 3, 1, 2, and 5 percent each for 50 periods (i.e., for period 1 to 50, growth rate is 2 percent, for period 51 to 100, the growth rate is 4 percent and so forth). We plot our results in figure 2.1, where we see the evolution over time of $\hat{\beta}_T$ and how it inherits the properties of γ_t. To facilitate comparison with (1), we superimpose the value of (1) that would

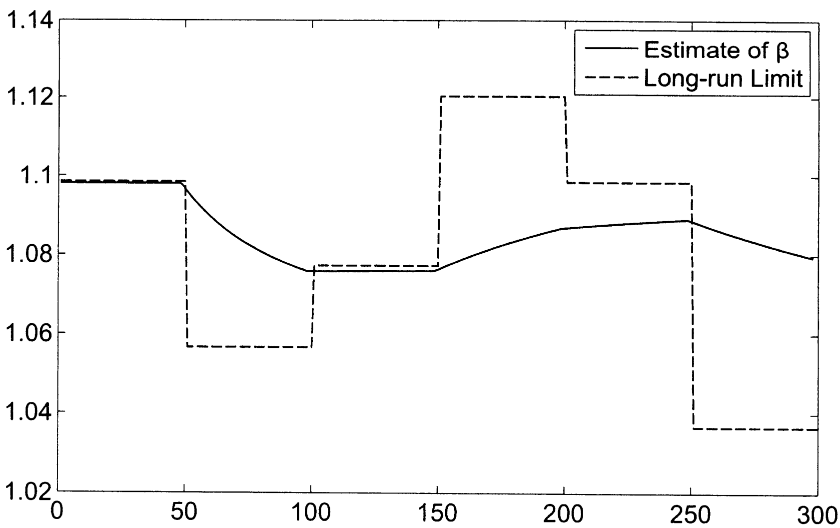

Figure 2.1
Estimate of β versus long-run limit

be implied if the growth rate in a period stayed constant over time. The graph shows how $\hat{\beta}_T$ converges to (1) within each block of 50 periods.

4 The Baseline Model

We will structure our investigation around a baseline New Keynesian business cycle model. We pick this model because it is the paradigmatic representative of the DSGE economies estimated by practitioners. Since (Fernández-Villaverde 2005) we have gone on the record on other occasions, criticizing the problems of this framework, we do not feel obliged to repeat those shortcomings here. Suffice it to say as a motivation that given the level of interest by policymaking institutions in this model, it is difficult to see a more appropriate vessel for our exploration.

The New Keynesian model is well known (see the book-length description in Woodford 2003). Consequently, we will be brief in our presentation and will omit some of the technical aspects. On the other hand, for concreteness, we need to discuss the model at a certain level of detail. The interested reader can access the entire description of the model at a complementary technical appendix posted at www.econ.upenn.edu/~jesusfv/benchmark_DSGE.pdf. In this section, to clarify our ideas, we will introduce the model without changes in the parameters. In section 5, we will introduce the parameter change over time.

4.1 *Households*

The basic structure of the economy is as follows. A representative household consumes, saves, holds real money balances, supplies labor, and sets its own wages subject to a demand curve and Calvo's pricing. The final output is manufactured by a competitive final-good producer, which uses as inputs a continuum of intermediate goods manufactured by monopolistic competitors. The intermediate-good producers rent capital and labor to manufacture their good. Also, the intermediate-good producers face the constraint that they can only change prices following a Calvo's rule. Finally, there is a monetary authority that fixes the one-period nominal interest rate through open market operations with public debt. Long-run growth is induced by the presence of two unit roots, one in the level of neutral technology and one in the investment-specific technology. These stochastic trends will allow us to estimate the model with the raw, undetrended data.

We have a continuum of households in the economy indexed by j. The households maximize the following lifetime utility function, which is separable in consumption, c_{jt}, real money balances, m_{jt}/p_t, and hours worked, l_{jt}:

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

where β is the discount factor, h controls habit persistence, ϑ is the inverse of Frisch labor supply elasticity, d_t is a shock to intertemporal preference with the law of motion:

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

and φ_t is a labor supply shock with the law of motion:

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

Households trade on the whole set of Arrow-Debreu securities, contingent on idiosyncratic and aggregate events. Our notation α_{jt+1} indicates the amount of those securities that pay one unit of consumption in event $w_{j,t+1,t}$ purchased by household j at time t at (real) price $q_{jt+1,t}$. To save on notation, we drop the explicit dependence on the event. Households also hold an amount, b_{jt}, of government bonds that pay a nominal gross interest rate of R_t and invest x_t. Then, the j – th household's budget constraint is:

$$c_{jt} + x_{jt} + \frac{m_{jt}}{p_t} + \frac{b_{jt+1}}{p_t} + \int q_{jt+1,t}\alpha_{jt+1} dw_{j,t+1,t}$$

$$= w_{jt} l_{jt} + [r_t u_{jt} - u_t^{-1}\Phi(u_{jt})]k_{jt-1} + \frac{m_{jt-1}}{p_t} + R_{t-1}\frac{b_{jt}}{p_t} + a_{jt} + T_t + \mathbb{F}_t$$

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, μ_t is an investment-specific technological shock (to be described momentarily), T_t is a lump-sum transfer, and $\mathbb{F}_t$ is the profits of the firms in the economy. We assume that $\Phi(1) = 0$, Φ' and $\Phi'' > 0$.

Investment x_{jt} induces a law of motion for capital:

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

where δ is the depreciation rate and $V(\cdot)$ is a quadratic adjustment cost function such that $V(\Lambda_x) = 0$, where Λ_x is the growth rate of investment along the balance growth path. Note that we index capital by the time its level is decided. The investment-specific technological shock follows an autoregressive process:

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

The first order conditions with respect to c_{jt}, b_{jt}, u_{jt}, k_{jt}, and x_{jt} are:

$$d_t(c_{jt} - hc_{jt-1})^{-1} - b\beta\mathbb{E}_t d_{t+1}(c_{jt+1} - hc_{jt})^{-1} = \lambda_{jt},$$

$$\lambda_{jt} = \beta\mathbb{E}_t\left(\lambda_{jt+1}\frac{R_t}{\Pi_{t+1}}\right),$$

$$r_t = \mu_t^{-1}\Phi'(u_{jt}),$$

$$q_{jt} = \beta\mathbb{E}_t\left\{\frac{\lambda_{jt+1}}{\lambda_{jt}}[(1-\delta)q_{jt+1} + r_{t+1}u_{jt+1} - \mu_{t+1}^{-1}\Phi(u_{jt+1})]\right\}, \text{ and}$$

$$1 = q_{jt}\mu_t\left[1 - V\left(\frac{x_{jt}}{x_{jt-1}}\right) - V'\left(\frac{x_{jt}}{x_{jt-1}}\right)\frac{x_{jt}}{x_{jt-1}}\right]$$
$$+ \beta\mathbb{E}q_{jt+1}\mu_{t+1}\frac{\lambda_{jt+1}}{\lambda_{jt}}V'\left(\frac{x_{jt+1}}{x_{jt}}\right)\left(\frac{x_{jt+1}}{x_{jt}}\right)^2,$$

where λ_{jt} is the Lagrangian multiplier associated with the budget constraint and q_{jt} is the marginal Tobin's Q, the Lagrangian multiplier associated with the investment adjustment constraint normalized by λ_{jt}.

The first order condition with respect to labor and wages is more involved. The labor employed by intermediate-good producers is supplied by a representative, competitive firm that hires the labor supplied by each household j. The labor supplier aggregates the differentiated labor of households with the production function:

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

where η controls the elasticity of substitution among different types of labor and l_t^d is the aggregate labor demand.

The labor "packer" maximizes profits subject to the production function (2), taking as given all differentiated labor wages w_{jt} and the wage w_t. From this maximization problem we get:

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

Then, to find the aggregated wage, we again use the zero profit condition $w_t l_t^d = \int_0^1 w_{jt} l_{jt} dj$ to deliver:

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

Households set their wages following a Calvo's setting. In each period, a fraction $1 - \theta_w$ of households reoptimize their wages. All other households can only partially index their wages by past inflation. Indexation is controlled by the parameter $\chi_w \in (0, 1)$. This implies that if the household cannot change its wage for τ periods, its normalized wage after τ periods is $\Pi_{s=1}^{\tau} (\Pi_{t+s-1}^{\chi_w} / \Pi_{t+s}) w_{jt}$.

Since we assume complete markets and separable utility in labor (see Erceg, Henderson, and Levin, 2000), we will concentrate on a symmetric equilibrium where $c_{jt} = c_t$, $u_{jt} = u_t$, $k_{jt-1} = k_t$, $x_{jt} = x_t$, $\lambda_{jt} = \lambda_t$, $q_{jt} = q_t$, and $w_{jt}^* = w_t^*$. In anticipation of that equilibrium, and after a fair amount of manipulation, we arrive at the recursive equations:

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

and:

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

that determine the evolution of wages.

Then, 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:

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

4.2 *The Final-Good Producer*

There is one final good produced using intermediate goods with the following production function:

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

where ε controls the elasticity of substitution.

Final-good producers are perfectly competitive and maximize profits subject to the production function (4), taking as given all intermediate goods' prices p_{it} and the final good price p_t. Repeating the same steps as for wages, we obtain the demand functions for each intermediate good:

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

where y_t^d is the aggregate demand and the zero profit condition $p_t y_t^d = \int_0^1 p_{it} y_{it} di$ to deliver:

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

4.3 *Intermediate-Good Producers*

There is a continuum of intermediate-good producers. Each intermediate-good producer i has access to a technology represented by a production function:

$$y_{it} = A_t k_{it-1}^{\alpha} (l_{it}^d)^{1-\alpha} - \phi z_t$$

where k_{it-1} is the capital rented by the firm, l_{it}^d is the amount of the "packed" labor input rented by the firm, the parameter ϕ corresponds to the fixed cost of production, and where A_t follows:

$$A_t = A_{t-1} \exp(\Lambda_A + z_{A,t}) \text{ where } z_{A,t} = \sigma_A \varepsilon_{A,t} \text{ and } \varepsilon_{A,t} \sim \mathcal{N}(0, 1).$$

The fixed cost ϕ is scaled by the variable $z_t = A_t^{1/(1-\alpha)} \mu_t^{\alpha/(1-\alpha)}$. We can think of z_t as a weighted index of the two technology levels A_t and μ_t, where the weight is the share of capital in the production function. The product ϕz_t guarantees that economic profits are roughly equal to zero in the steady state. Also, we rule out the entry and exit of intermediate-good producers. Note that z_t evolves over time as $z_t = z_{t-1} \exp(\Lambda_z + z_{z,t})$ where $z_{z,t} = (z_{A,t} + \alpha z_{\mu,t})/(1-\alpha)$ and $\Lambda_z = (\Lambda_A + \alpha\Lambda_\mu)/(1-\alpha)$. We will see in the following 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} in perfectly competitive factor markets in order to minimize real costs, which implies a marginal cost of:

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

The 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 pricing scheme as households. In each period, a fraction $1-\theta_p$ of firms reoptimize their prices. All other firms can only index their prices by past inflation. Indexation is controlled by the parameter $\chi \in (0,1)$, where $\chi = 0$ is no indexation and $\chi = 1$ is total indexation.

The problem of the firms is then:

$$\max_{p_{it}} \mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_p)^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} \left[\left(\prod_{s=1}^{\tau} \Pi^{\chi}_{t+s-1} \frac{p_{it}}{p_{t+\tau}} - mc_{t+\tau} \right) y_{it+\tau} \right]$$

subject to

$$y_{it+\tau} = \left(\prod_{s=1}^{\tau} \Pi^{\chi}_{t+s-1} \frac{p_{it}}{p_{t+\tau}} \right)^{-\varepsilon} y^d_{t+\tau},$$

where the marginal value of a dollar to the household is treated as exogenous by the firm. Since there are complete markets in securities, this marginal value is constant across households and, consequently, $\lambda_{t+\tau}/\lambda_t$ is the correct valuation on future profits.

We write the solution of the problem in terms of two recursive equations in g^1_t and g^2_t:

$$g^1_t = \lambda_t mc_t y^d_t + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi^{\chi}_t}{\Pi_{t+1}} \right)^{-\varepsilon} g^1_{t+1}$$

$$g^2_t = \lambda_t \Pi^*_t y^d_t + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi^{\chi}_t}{\Pi_{t+1}} \right)^{1-\varepsilon} \left(\frac{\Pi^*_t}{\Pi^*_{t+1}} \right) g^2_{t+1}$$

where $\varepsilon g^1_t = (\varepsilon - 1) g^2_t$ and $\Pi^*_t = p^*_t / p_t$.

Given Calvo's pricing, the price index evolves:

$$p_t^{1-\varepsilon} = \theta_p (\Pi^{\chi}_{t-1})^{1-\varepsilon} p_{t-1}^{1-\varepsilon} + (1-\theta_p) p_t^{*1-\varepsilon}$$

or, dividing by $p_t^{1-\varepsilon}$,

$$1 = \theta_p \left(\frac{\Pi^{\chi}_{t-1}}{\Pi_t} \right)^{1-\varepsilon} + (1-\theta_p) \prod_t^{*1-\varepsilon}$$

4.4 *The Government*

The government sets the nominal interest rates according to the Taylor rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma R}\left[\left(\frac{\Pi_t}{\Pi}\right)^{\gamma\Pi}\left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\Lambda_y d}\right)^{\gamma y}\right]^{1-\gamma_R}\exp(m_t) \tag{5}$$

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 represents the target levels of inflation (equal to inflation in the steady state), R is the steady-state gross return of capital, and $\Lambda_y d$ the steady-state gross growth rate of y_t^d. With a bit of abuse of language, we will refer to the term $(y_t^d/y_{t-1}^d)/\Lambda_{y^d}$ as the *growth gap*. 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 $\mathcal{N}(0, 1)$. We introduce the previous period interest rate, R_t, to match the smooth profile of the interest rate over time observed in the United States.

4.5 *Aggregation*

First, we begin with the aggregate demand:

$$y_t^d = c_t + x_t + \mu_t^{-1}\,\Phi(u_t)k_{t-1}.$$

Then, using the production function for intermediate-good producers, the fact that all the firms pick the same capital-labor ratio, and market clearing in the output and input markets, we find the aggregate demand must be equal to aggregate supply:

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

where

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

is the aggregate loss of efficiency induced by price dispersion. By the properties of the index under Calvo's pricing:

$$v_t^p = \theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + (1-\theta_p)\Pi_t^{*-\varepsilon}.$$

Finally, we integrate labor demand over all households j to obtain:

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

where l_t is the aggregate labor supply of households. Hence if we define

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

we get:

$$l_t^d = \frac{1}{\upsilon_t^w} l_t$$

and:

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

4.6 *Equilibrium*

A definition of equilibrium in this economy is standard and the equations that characterize it are determined by the first order conditions of the household, the first order conditions of the firms, the Taylor rule of the government, and market clearing.

To undertake our quantitative analysis, we must approximate the equilibrium dynamics of the economy. Ours is a large model (even the version without parameter drifting has 19 state variables). Moreover, we will need to solve the model repeatedly during our estimation process. We have argued elsewhere (Fernández-Villaverde, Rubio-Ramírez, and Santos 2006) that there is much to be gained from a nonlinear estimation of the model, both in terms of accuracy and in terms of identification. This is particularly true if we want to allow the agents in the economy to ensure themselves against future changes in the parameters of the model. Hence, we require a nonlinear solution method that is fast and accurate. In previous work (Aruoba, Fernández-Villaverde, and Rubio-Ramírez 2006), we have found that a second order perturbation around the deterministic steady state of the model fulfills the previous desiderata.

But before solving the model, we clear up some technical issues. First, because of technological change, most of the variables are growing in average. To achieve the right accuracy in the computation, we make the variables stationary and solve the model in the transformed variables. Hence, we define $\tilde{c}_t = c_t/z_t$, $\tilde{\lambda}_t = \lambda_t z_t$, $\tilde{r}_t = r_t \mu_t$, $\tilde{q}_t = q_t \mu_t$, $\tilde{x}_t = x_t/z_t$, $\tilde{w}_t = w_t/z_t$, $\tilde{w}_t^* = w_t^*/z_t$, $\tilde{k}_t = k_t/z_t \mu_t$, and $\tilde{y}_t^d = y_t^d/z_t$. Also note that $\Lambda_c = \Lambda_x = \Lambda_w = \Lambda_{w^*} = \Lambda_{y^d} = \Lambda_z$. Second, we choose functional forms for $\Phi(\cdot)$ and $V(\cdot)$. For $\Phi(u)$ we pick $\Phi(u) = \Phi_1(u - 1) + (\Phi_2/2)(u - 1)^2$. We normalize

$u = 1$ in the steady state. Hence, $\tilde{r} = \Phi'(1) = \Phi_1$ and $\Phi(1) = 0$. The investment adjustment cost function is $V(x_t/x_{t-1}) = (\kappa/2)[(x_t/x_{t-1}) - \Lambda_x]^2$. Then, along the balanced growth path, $V(\Lambda_x) = V'(\Lambda_x) = 0$.

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

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

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

As a first step, we parameterize the matrix of variances-covariances of the exogenous shocks as $\overline{\Omega}(\chi) = \chi\Omega$, where $\overline{\Omega}(1) = \Omega$ is a diagonal matrix. However, nothing really depends on that assumption, and we could handle an arbitrary matrix of variances-covariances. Then, we take a perturbation solution around the deterministic steady state of the model, that is, $\chi = 0$.

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

$$\overline{S}_{t+1} = \Psi_{s1}(\overline{S}'_t, \varepsilon'_t)' + \frac{1}{2}(\overline{S}'_t, \varepsilon'_t)\Psi_{s2}(\overline{S}'_t, \varepsilon'_t)' + \Psi_{s3}, \tag{6}$$

where Ψ_{s1} is a 1×24 vector and Ψ_{s2} is a 24×24 matrix. The term $\Psi_{s1}(\overline{S}'_t, \varepsilon'_t)'$ constitutes the linear solution of the model, $(\overline{S}'_t, \varepsilon'_t)\Psi_{s2}(\overline{S}'_t, \varepsilon'_t)'$ is the quadratic component, and Ψ_{s3} is a 1×24 vector of constants added by the second order approximation that corrects for precautionary behavior. Some of the entries of the matrices Ψ_{si} will be zero.

From the same output, we find the law of motion for the observables

$$\mathcal{Y}^T = (\Delta \log \mu_t^{-1}, \Delta \log y_t, \Delta \log l_t, \log \Pi_t, \log R_t)'.$$

Now, define $S_t = (\overline{S}'_t, \overline{S}'_{t-1}, \varepsilon'_{t-1})$. We keep track of the past states, $\overline{S}'_{t-1}$, because some of the observables in the following measurement equation will appear in first differences. Then, we write to the observation equation:

$$\mathcal{Y}^T = \Psi_{o1}(S'_t, \varepsilon'_t)' + \frac{1}{2}(S'_t, \varepsilon'_t)\Psi_{o2}(S'_t, \varepsilon'_t)' + \Psi_{o3} \tag{7}$$

where Ψ_{o1} and $\Psi_{o3} 1 \times 48$ vectors and Ψ_{o2} is a 48×48 matrix.

While the law of motion for states is unique (or at least equivalent to a class of different states, all of which have the same implications for the dynamics of the model), the observation equation depends on what we assume the researcher actually observes. In our case, we have chosen the first differences of the relative price of investment, output, hours, inflation, and the federal funds rate. Unfortunately, we do not know much about the right choice of observables and how they may affect our estimation results (for one of the few articles on this topic, see Boivin and Giannoni 2006).

4.7 *The Likelihood Function*

Equations (6) and (7) constitute the state space representation of our model. One convenient property of this representation is that we can exploit it to evaluate the likelihood of a DSGE model, an otherwise challenging task. The likelihood, $\mathcal{L}(\mathbb{Y}^T; \Psi)$, is the probability that the model assigns to a sequence of realizations of the observable $\mathbb{Y}^T$ given parameter values:

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

Note that Φ_1 is not included in Ψ because it is a function of the other parameters in the economy to ensure that $\tilde{r} = \Phi_1$. With $\mathcal{L}(\mathbb{Y}^T; \Psi)$, we can estimate Ψ by maximizing the likelihood or by combining it with a prior density for the model parameter to form a posterior distribution.

How do we evaluate the likelihood $\mathcal{L}(\mathbb{Y}^T; \Psi)$? Given the Markov structure of our state space representation, we begin by factorizing the likelihood function as:

$$\mathcal{L}(\mathbb{Y}^T; \Psi) = \prod_{t=1}^{T} \mathcal{L}(\mathbb{Y}^T \mid \mathbb{Y}^{t-1}; \Psi).$$

Then, conditioning on the states:

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

If we know S_t, computing $\mathcal{L}(\mathbb{Y}_t \mid S_t; \Psi)$ is relatively easy. Conditional on S_t, the measurement equation (7) is a change of variables from ε_t to $\mathcal{Y}^T$. Hence, we can apply the change-of-variable formula to evaluate the required probabilities. Similarly, if we know S_0, we can employ (6) and the measurement equation (7) to compute $\mathcal{L}(\mathbb{Y}_t \mid S_0; \Psi)$. Consequently, knowledge of the sequence $(p(S_t \mid \mathbb{Y}^{t-1}; \Psi))_{t=1}^{T}$ and of $p(S_0; \Psi)$ allows us to find $\mathcal{L}(\mathbb{Y}^T; \Psi)$. Evaluating (or at least drawing from) $p(S_0; \Psi)$ is usually straightforward, although often costly (Santos and Peralta-Alva 2005). The difficulty is to characterize the sequence of conditional distributions $[p(S_t \mid \mathbb{Y}^{t-1}; \Psi)]_{t=1}^T$ and to compute the integrals in (8).

An algorithm for doing so (but not the only one; see the technical appendix to Fernández-Villaverde and Rubio-Ramírez 2007, for alternatives and references) is to use a simulation technique known as the *particle filter.* Fernández-Villaverde and Rubio-Ramírez (2005 and 2007) have shown that the particle filter can be successfully applied to the estimation of nonlinear and/or nonnormal DSGE models. The particle filter is a sequential Monte Carlo method that replaces the $[p(S_t \mid \mathbb{Y}^{t-1}; \Psi)]_{t=1}^T$ by an empirical distribution of draws generated by simulation. The bit of magic in the particle filter is that the simulation is generated through a procedure known as sequential importance resampling (SIR). Sequential importance resampling guarantees that the Monte Carlo method achieves sufficient accuracy in a reasonable amount of time, something that cannot be achieved without resampling (Arulampalam, Maskell, Gordon, and Clapp 2002). The appendix describes in further detail the working of the particle filter.

4.8 A Bayesian Approach

We will confront our model with the data using Bayesian methods. The Bayesian paradigm is a powerful and flexible perspective for the estimation of DSGE models (see the survey by An and Schorfheide 2006). First, Bayesian analysis is a coherent approach to inference based on a clear set of axioms. Second, the Bayesian approach handles in a natural way misspecification and lack of identification, both serious concerns in the estimation of DSGE models (Canova and Sala 2006). Moreover, it has desirable small-sample and asymptotic properties, even when evaluated by classical criteria (Fernández-Villaverde and Rubio-Ramírez 2004). Third, priors are a flexible procedure to introduce presample information and to reduce the dimensionality problem associated with the number of parameters. This property will be especially attractive in our application, since parameter drifting will increase the practical number of dimensions of our model.

The Bayesian approach combines the likelihood of the model $\mathcal{L}(\mathbb{Y}^T; \Psi)$ with a prior density for the parameters $p(\Psi)$ to form a posterior

$$p(\Psi \mid \mathbb{Y}^T) \propto \mathcal{L}(\mathbb{Y}^T; \Psi) p(\Psi).$$

The posterior summarizes the uncertainty regarding the parameters, and it can be used for point estimation. For example, under a quadratic loss function, our point estimates will be the mean of the posterior.

Since the posterior is also difficult to characterize, we generate draws from it using a Metropolis-Hastings algorithm. We use the resulting empirical distribution to obtain point estimates, variances, and so on. We describe this algorithm in the appendix.

5 Parameter Drifting

Now we are ready to deal with parameter drifting. Since the extension to other cases of parameter variation is rather straightforward, we present only one example of drift within our model.

Motivated by the first example in section 3, we will investigate the situation where the Taylor rule is specified as:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_{Rt}} \left[\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_{\Pi t}} \left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\Lambda_{y^d}}\right)^{\gamma_{yt}}\right]^{1-\gamma_{Rt}} \exp(m_t). \tag{9}$$

Note the difference with the specification in (5): in the new equation the elasticities of the response of the interest rate $(\gamma_{Rt}, \gamma_{\Pi t}, \gamma_{yt})$ are indexed by time.

We will postulate that the parameters follow an autoregressive model (AR[1]) in logs to ensure that the parameter is positive:

$$\log \gamma_{Rt} = \min[(1 - \rho_R)\log \overline{\gamma}_R + \rho_R \log \gamma_{Rt-1} + \varepsilon_{Rt}, 0] \tag{10}$$

$$\log \gamma_{\Pi t} = (1 - \rho_\Pi) \log \overline{\gamma}_\Pi + \rho_\Pi \log \gamma_{\Pi t-1} + \varepsilon_{\Pi t} \tag{11}$$

$$\log \gamma_{yt} = (1 - \rho_y) \log \overline{\gamma}_y + \rho_y \log \gamma_{yt-1} + \varepsilon_{yt} \tag{12}$$

where $[\varepsilon_{Rt}, \varepsilon_{\Pi t}, \varepsilon_{yt}]$ are i.i.d. normal shocks and Q is a 3×3 matrix of covariances.[2] We allow for arbitrary correlation in the innovations, since it is plausible that the reasons why the monetary authority becomes more (less) responsive to inflation are the same reasons it will become less (more) responsive to the growth gap. Also, we could generalize the changes in parameters by allowing changes in Π or in the variance of m_t (R and Λ_{y^d} are not chosen by the monetary authority but they are implied by the other parameters of the model and by Π). Finally, we impose the stability condition that the smoothing coefficient γ_{Rt} must be less than 1 in levels (or less than 0 in logs).

Our specification of parameter drift emphasizes the continuity of the change process, in opposition to the discrete changes in the parameters captured by a Markov-switching process (see Davig and Leeper 2006a and 2006b). We do not have a strong prior preference for one version or the other. We prefer our specification because it is parsimonious and easy to handle, and it captures phenomena such as the Fed's gradual learning about the behavior of the economy.

According to our favorite interpretation of parameter drifting, we will assume that agents understand that policy evolves over time following equations (10)–(12). Consequently, they react to it and make their decisions based on the current values of γ_t and on the fact that γ_t will evolve over time.

The drift of the parameters implies that the economy will travel through zones where the Taylor principle is not satisfied. However, this may not necessarily mean that the equilibrium is not unique. In the context of Markov-regime changes in the coefficients of the Taylor rule, Davig and Leeper (2006b) have developed what they call the generalized Taylor principle. Davig and Leeper argue that a unique equilibrium survives if the Taylor rule is sufficiently active when the economy is in the active policy regime or if the expected length of time the economy will be in the nonactive policy regime is sufficiently small. To keep this paper focused, we will not dwell on generating results equivalent to Davig and Leeper's in our environment. Suffice it to say that one further advantage of the Bayesian approach is that we can handle restrictions on the parameter drifting with the use of the priors. For example, we can implement a reflecting boundary on (10) by putting a zero prior on the possibility of violating that boundary. Also, in our empirical analysis, we estimate $\overline{\gamma}_\Pi$ as being bigger than 1. This suggests that the Taylor principle will be satisfied, at least on average.

Our formulation of parameter drifting has one important drawback: we do not model explicitly why the parameters change over time. In section 3, we discuss that changes in the policy parameters could be a reflection of changing political priorities or evolving perceptions about the effectiveness of policy. A more complete model would include explicit mechanisms through which we discipline the movement of the parameters over time. Many of those mechanisms can be incorporated into our framework, since we are rather flexible with the type of functional forms for the parameter drift that we can handle.

The model in section 4 carries on except with the modification of (9) and the fact that all the conditional expectations now incorporate the process (10). Thus, the states of the model with parameter drifting are:

$$\overline{S}_t = \begin{pmatrix} \hat{\Pi}_{t-1}, \hat{\tilde{w}}_{t-1}, \hat{g}^1_{t-1}, \hat{g}^2_{t-1}, \hat{\tilde{k}}_{t-1}, \hat{R}_{t-1}, \hat{\tilde{y}}^d_{t-1}, \hat{\tilde{c}}_{t-1}, \hat{\upsilon}^p_{t-1}, \hat{\upsilon}^w_{t-1}, \\ \hat{\tilde{q}}_{t-1}, \hat{\tilde{f}}_{t-1}, \hat{\tilde{x}}_{-1}, \hat{\tilde{\lambda}}_{t-1}, \hat{\tilde{z}}_{t-1}, z_{\mu,t-1}, \hat{d}_{t-1}, \hat{\varphi}_{t-1}, z_{A,t-1}; \gamma_{Rt-1}, \gamma_{\Pi t-1}, \gamma_{yt} \end{pmatrix}',$$

where we have included γ_{Rt}, $\gamma_{\Pi t}$, and γ_{yt} as three additional states. We will follow the convention of separating drifting parameters from the other states with a semicolon (;) since they are an object of interest by themselves. Similarly, we apply the particle filter to evaluate the likelihood of the model and the Metropolis-Hasting algorithm to simulate from the posterior.

6 Empirical Analysis

This section presents our empirical analysis. First, we report the point estimates of the model when we keep all parameters fixed over the sample. This estimation sets a natural benchmark for the rest of the study. Second, we discuss the results of an exercise where we allow the parameters of the Taylor rule of the monetary authority to change over time. Third, we analyze the evolution of the parameters that control the level of price and wage rigidities. In the interest of space we select these two exercises as particularly illustrative of the procedure we propose. However, we could have performed many other exercises within the framework of our methodology.

We estimate the model using five time series for the United States: (1) the relative price of investment with respect to the price of consumption, (2) real output per capita growth, (3) hours worked per capita, (4) the CPI, and (5) the federal funds rate. Our sample goes from 1955:Q1 to 2000:Q4. We stop our sample at the end of 2000 because of the absence of good information on the relative price of investment after that time. To make the observed series compatible with the model, we compute both real output and real gross investment in consumption units. For that purpose, we use the relative price of investment defined as the ratio of an investment deflator and a deflator for consumption. The consumption deflator is constructed from the deflators of nondurable goods and services reported in the national income and product accounts (NIPA). Since the NIPA investment deflators are poorly measured, we rely on the investment deflator constructed by Fisher (2006), a series that ends at 2000:Q4. The appendix provides further information on the construction of the data.

6.1 Point Estimation

Before reporting results, we specify priors for the model's parameters. We adopt flat priors for all parameters. We impose boundary constraints only to make the priors proper and to rule out parameter values that are either incompatible with the model (i.e., a negative value for a variance, Calvo parameters outside the unit interval) or implausible (the response to inflation in the Taylor rule being bigger than 100). The looseness of such constraints is shown by the fact that the simulations performed in the following never travel even close to the bounds. Also, we fix four parameters, (υ, ϕ, Φ_2, δ). The parameter controlling money demand υ is irrelevant for equilibrium dynamics because the government will supply as much money as required to implement the nominal interest rate determined by the Taylor rule. We fix the parameter ϕ to zero, since we do not have information on pure profits by firms (in the absence of entry/exit of firms, there are no serious implications for equilibrium dynamics). The parameter of the investment adjustment cost, Φ_2, is set to 0.001, and depreciation, δ, to 0.0149 because they are difficult to identify. Our choice of δ matches the capital-output ratio in the data (remember that in our model we have both physical depreciation, controlled by δ, and economic depreciation, induced by the change in the relative price of capital).

Our choice of flat priors is motivated by the observation that, with this prior, the posterior is proportional to the likelihood function.[3] 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 more informative priors can always reweight the draws from the posterior to accommodate his favorite priors (Geweke 1989).[4] We repeated our estimation with an informative prior without finding important differences in the results.

Table 2.1 summarizes our results by reporting the mean and the standard deviation of the posterior.[5] Most of our point estimates coincide with the typical findings of other estimation exercises and the standard deviations are small. Hence, we comment only on a few of them. We have a high degree of habit persistence—h is 0.88—and we have a Frisch elasticity of labor supply of 0.74 (1/1.36), well within the bounds of findings in the recent microeconomic literature (Browning, Hansen, and Heckman 1999). The estimates of elasticities of substitution ε and η are around 8, implying average markups of around 14 percent.

Table 2.1
Point estimates

Parameter	Point Estimate	S.D.	Parameter	Point Estimate	S.D.	Parameter	Point Estimate	S.D.
β	0.999	0.001	γ_R	0.790	0.012	ρ_d	0.951	0.006
h	0.877	0.009	γ_y	0.190	0.056	ρ_ϑ	0.942	0.015
ψ	8.942	0.045	γ_Π	1.260	0.075	σ_μ	0.101	0.006
ϑ	1.359	0.004	Π	1.008	3.6e-4	σ_A	0.007	0.002
κ	7.679	0.600	α	0.255	0.011	σ_m	0.003	8.4e-5
θ_w	0.451	0.0923	ε	7.957	0.1593	σ_d	0.060	0.003
χ_w	0.849	0.1231	η	7.965	0.2984	σ_φ	0.070	0.011
θ_p	0.907	0.012	Λ_μ	0.010	2.86e-4			
χ_p	0.151	0.100	Λ_A	0.0005	4.57e-4			

The Calvo parameter for price adjustment, θ_p, is a relatively high 0.91, while the indexation level, χ_p, is 0.15. It is tempting to compare our estimates with the microeconomic evidence on the average duration of prices (Bils and Klenow 2004, or Nakamura and Steinsson 2006). However, the comparison is difficult because we have partial indexation: prices change every quarter for all producers, a fraction θ_p because producers reoptimize and a fraction $1 - \theta_p$ because of indexation. The Calvo parameter for wage adjustment, θ_w, is 0.45, while the indexation, χ_w, is 0.85. Our point estimates imply stronger nominal rigidities in price than in wages, in line with Rabanal and Rubio-Ramírez (2005) or Galí and Rabanal (2004) but diverging from Smets and Wouters (2003), who have much more informative priors.

The policy parameters (γ_R, γ_Π, γ_y, Π) are quite standard. 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 0.78 percent (or around 3 percent yearly).

The growth rates of the investment-specific technological change, Λ_μ, and of the neutral technology, Λ_A, imply that most of the growth in the U.S. economy (83 percent) is induced by improvements in the capital-producing technology. This result corroborates the importance of modelling biased technological change for understanding growth and fluctuations that Greenwood, Herkowitz, and Krusell (1997 and 2000) have so forcefully defended. The estimated long-run growth rate of the economy, $(\Lambda_A + \alpha\Lambda_\mu)/(1 - \alpha)$ is 0.4 percent per quarter, or 1.6 percent annually, roughly the observed mean in the sample. Also, the standard deviation σ_μ is much higher than σ_A.

Our estimation serves different roles. First, it validates our model as a promising laboratory for our exercises with parameter drifting. Since in the benchmark case we obtain results compatible with the literature and with the basic growth properties of the U.S. economy, we know that the results with parameter drifting will indeed come from that feature of the estimation. Second, we use our point estimates to initialize the parameters in the exercises with parameter drifting.

In the next two subsections, we will report our findings when we allow one parameter to vary at a time. We do this for convenience. First, allowing several parameters to move simultaneously makes the computation and estimation of the model much more costly. Second, the information in the sample is limited and it is difficult to obtain stable estimates otherwise. Third, especially in our second exercise, our objective is not so much to have the richest possible model to fit the data well but to show that as soon as parameters are allowed to change over time, strong signs of misspecification appear. We will continue the exploration of joint moves of parameters in the near future.

6.2 *Evolution of Policy Parameters*

Our first exercise studies the evolution of the policy parameters in the Taylor rule. This investigation evaluates how much evidence there is in the data of a changing monetary policy over time. As we discussed in section 3, the literature has extensively debated the topic (e.g., Clarida, Galí, and Gertler 2000, Cogney and Sargent 2001, Lubick and Schorfheide 2004, Sims and Zha 2006, Boivin 2006). However, the empirical methods applied so far are unsatisfactory because they rely either on divisions of the sample that do not let the agents in the model forecast the changes in policy or on the estimation of reduced forms.

Arguably the most interesting parameter is $\gamma_{\Pi t-1}$, since this parameter controls how aggressively the monetary authority responds to inflation. In addition, $\gamma_{\Pi t-1}$ is intimately linked with the issue of multiplicity of equilibria and the possibility of monetary policy being a source of instability. Figure 2.2 plots our point estimate of the evolution of $\gamma_{\Pi t-1}$ over time plus the two standard deviations interval to gauge the uncertainty present in the estimation. We report the smoothed values of $\gamma_{\Pi t-1}$ using the whole sample (Godsill, Doucet, and West 2004). We find it convenient, for expositional purposes, to eliminate some of the quarter-to-

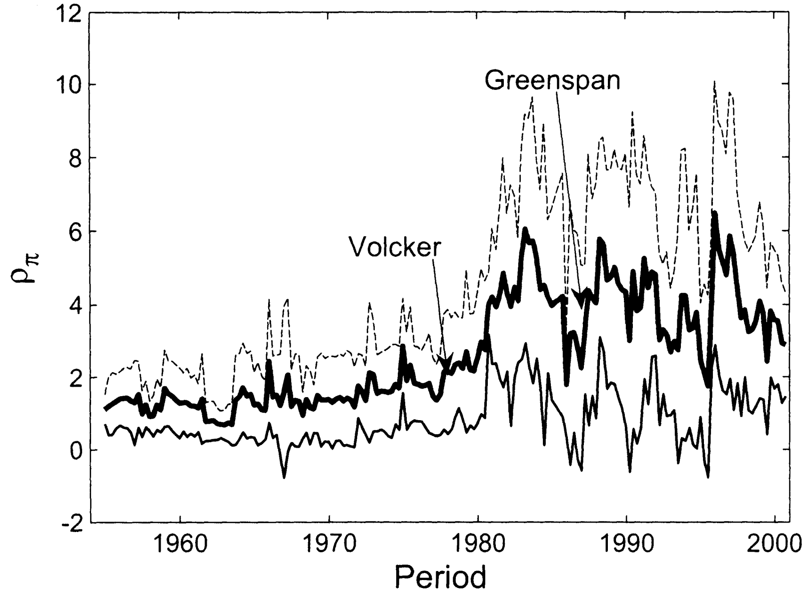

Figure 2.2
Evolution of Response to Inflation

quarter variation of the parameter. To accomplish this goal, in figure 2.3 we graph the trend of the change of the parameter where we compute the trend using a Hodrick-Prescott filter. We emphasize that this trend is only a device to read the graph more clearly and lacks a formal statistical interpretation.

In both figures 2.2 and 2.3, $\gamma_{\Pi t-1}$ starts low, slightly above 1 during the 1950s, 1960s, and early 1970s, with periods when it was even below 1. However, in the mid-1970s, and especially after Volcker's appointment as chairman of the Board of Governors, $\gamma_{\Pi t-1}$ soared. The response to inflation reached its peak in the early 1980s, where it was as high as 6 in one quarter. After that maximization, $\gamma_{\Pi t-1}$ slowly decreases during the 1990s, perhaps reflecting the Fed's more permissive attitude to accommodate the strong productivity growth associated with the Internet boom.

Since our model has parameter drifting, it is not straightforward to compare these numbers with estimates obtained in fixed-parameter models. However, we clearly confirm the findings of Clarida, Galí, and Gertler (2000), Lubick and Schorfheide (2004), and Boivin (2006)—that monetary policy has become much more active in the last 25 years. Our

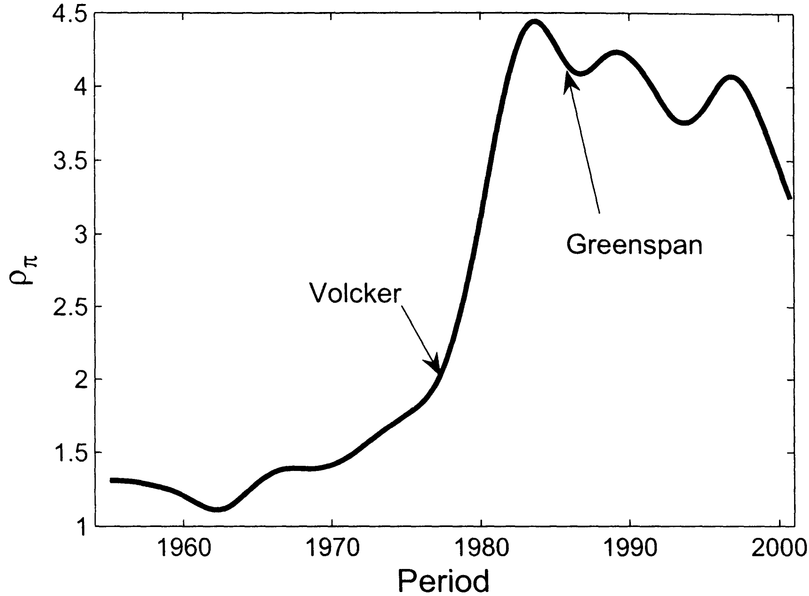

Figure 2.3
HP-Trend Evolution of Response to Inflation

finding is also consistent with the results of figure 12 in Cogley and Sargent (2001), where they trace the evolution of the activism coefficient as measured by a parameter-drifting VAR.

Another parameter of importance is the inflation target of the monetary authority, Π. Histories like those in Taylor (1998), Sargent (1999), or Primiceri (2006) explain that the inflation target may have changed over time as a reflection of the Fed's varying beliefs about the trade-off between unemployment and inflation. Figure 2.4 plots the evolution of the target over time plus the two standard deviation interval. From the start of the sample until the early 1970s and, later, for the 1990s, Π hovers around 1.004, or, in annual terms, around 1.6 percent. This number is close to the informal target or comfort zone that, according to many commentators, describes the Fed's behavior. During the intermediate years, the inflation target increases, reflecting perhaps the views the Fed had about the possibility of exploiting the Phillips curve or illustrating the information lags regarding the changing features of the economy emphasized by Orphanides (2002). We find intriguing the similarity of figure 2.4 to Romer and Romer's (2002) hypothesis, based on narrative accounts and internal Greenbook forecasts of the Fed, that monetary

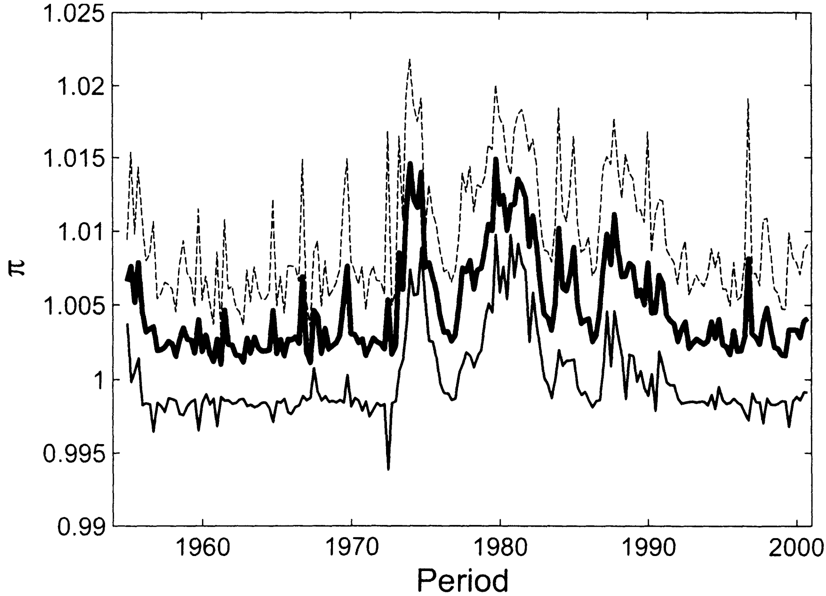

Figure 2.4
Evolution of inflation target

policy in the United States has gone through a long cycle of moderation, aggressiveness, and renewed temperance.

Our estimates of the evolution of the inflation target provide a reality check on our procedure. In figure 2.5, we plot the inflation target versus realized inflation. If the estimation is working properly, part of the variation in the inflation target needs to be accounted for, in a purely mechanical fashion, by changes in inflation. That is precisely what we observe: as inflation increases and then falls during the late 1960s and 1970s, the target inflation estimated goes up and down.

Note, however, that the inflation target fluctuates roughly between 40 and 50 percent less than inflation. Particularly during the 1970s, the inflation target is well below actual inflation. This difference is accounted in two ways. First, by the form of our Taylor rule. We assume that one input into the rule is the growth gap between the growth of output y_t^d/y_{t-1}^d and the long-run growth rate of the economy Λ_{y^d}. The 1970s were years of very low growth in comparison with Λ_{y^d}.[6] Thus, our model interprets the behavior of the Fed as lowering the interest rates as a response to low output growth in exchange for higher inflation. Second, our model backs up large negative technology shocks in the 1970s that

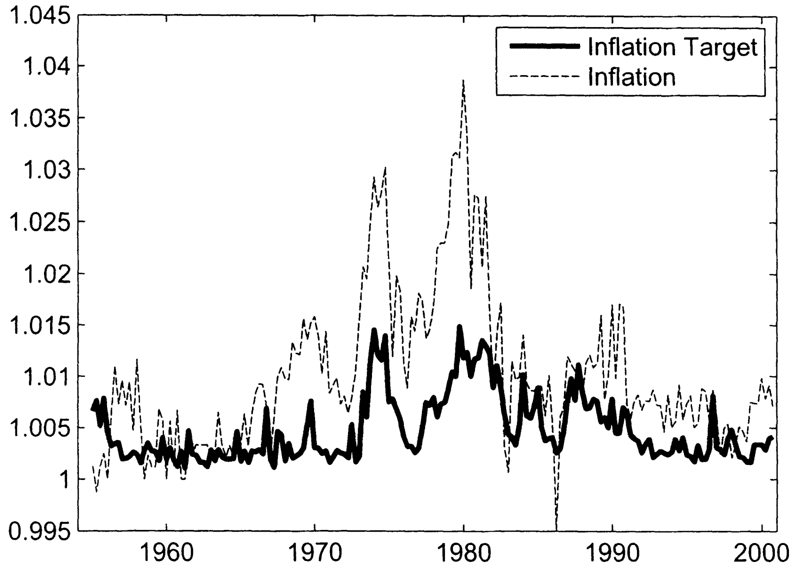

Figure 2.5
Inflation Target versus Inflation

push inflation above the target level. Hence, an alternative way to think about this result is that our model suggests that the big rise in inflation during the 1970s had less to do with changes in the inflation target than with a series of unfavorable aggregate shocks.

We summarize our results. First, the Fed's response toward inflation became more aggressive in the late 1970s and early 1980s and has stayed high since then, with perhaps a small decline. Second, the inflation target was relaxed in the 1970s, but not enough to account for the high inflation of that decade. We trust our results not only because they come from the estimation of a coherent DSGE model, but also because they are consistent with the findings of the existing literature that uses alternative estimation procedures, with narrative accounts of monetary policy, and with the reality check explained previously.

6.3 Evolution of Price and Wage Rigidities

A key set of parameters in our model are those determining the extent of price and wage rigidities, $(\theta_p, \chi_p, \theta_w, \chi_w)$. These four parameters generate the nominal rigidity in the economy required to match the impulse response functions documented by VARs (Christiano, Eichenbaum, and Evans 2005).

Given their importance in the model, it is unfortunate that these parameters have only a tenuous link with microeconomic foundations. Even if the Calvo adjustment probabilities are the reduced form of a convex adjustment cost model, the environment that produces this reduced form has changed over the years in our sample. We have gone from periods of high inflation and low response of the monetary authority to rising prices to periods of much lower inflation and a more aggressive attitude by the Fed toward inflation. Moreover, the U.S. economy has experienced a notable level of deregulation, increasing competition in internal markets from international trade, and lower unionization rates. The justification of the indexation parameters or their relation to the Calvo adjustment probabilities is even less clear. Why do agents index their prices and wages? And if they do, to which quantity? Past inflation? Current inflation? Steady-state inflation? Wage inflation? Consequently, it is natural to examine the possibility that the parameters (θ_p, χ_p, θ_w, χ_w) drift over time, both as a measure of how strong nominal rigidities have been in each different moment and as a tool to assess the extent of possible misspecification of the model along this dimension.

As in the case of policy parameters, we specify an AR(1) as the law of motion for the parameters:

$$\log \theta_{pt} = \min[(1 - \rho_{\theta_p})\log \overline{\theta}_p + \rho_{\theta_p} \log \theta_{pt-1} + \varepsilon_{\theta_{pt}}, 0]$$

$$\log \chi_{pt} = \min[(1 - \rho_{\chi_p})\log \overline{\chi}_p + \rho_{\chi} \log \chi_{pt-1} + \varepsilon_{\chi_{pt}}, 0]$$

$$\log \theta_{wt} = \min[(1 - \rho_{\theta_w})\log \overline{\theta}_w + \rho_{\theta_w} \log \theta_{wt-1} + \varepsilon_{\theta_{wt}}, 0]$$

$$\log \chi_{wt} = \min[(1 - \rho_{\theta_w})\log \overline{\chi}_w + \rho_{\chi} \log \chi_{wt-1}\, \varepsilon_{\chi_{wt}}, 0]$$

where ($\varepsilon_{\theta_{pt}}$, $\varepsilon_{\chi_{pt}}$, $\varepsilon_{\theta_{wt}}$, $\varepsilon_{\chi_{wt}}$) are i.i.d. normal shocks and where we take the minimum of the value of the parameter induced by the autoregressive component and 0 to be sure that the two standard deviation interval and figure 2.6 its HP-trend (again, with the HP-trend of the CPI superimposed). Indexation evolves in an opposite way to price duration: it starts low in the 1950s and 1960s but raises very strongly during the late 1960s. Then, it drops dramatically in the mid-1970s and stays low over the next 20 years (except for a temporary increase in the early 1980s). In the last part of the sample, during the 1990s, χ_{pt} steadily drops. The drop in indexation in the second half of the 1970s may be accounted for by firms

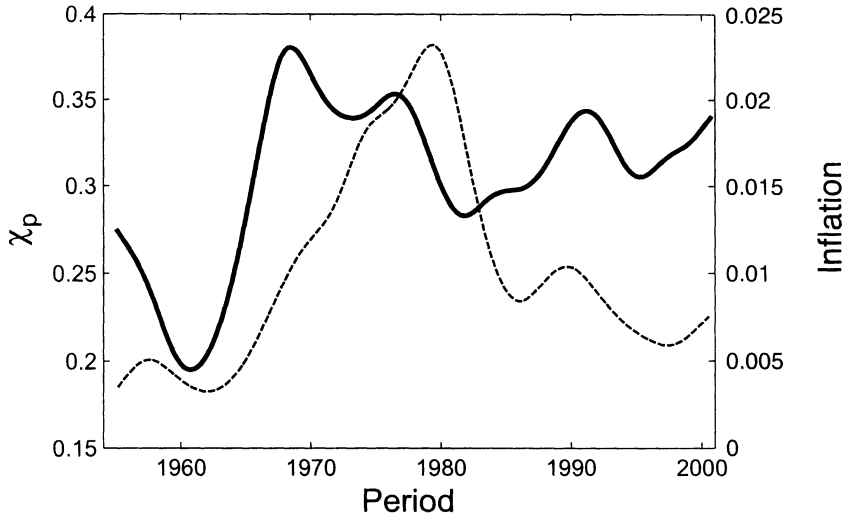

Figure 2.6
HP-Trend Price Indexation vs. HP-Trend Inflation

switching to more often optimal price adjustments and less automatic pricing rules. Firms were perhaps induced by the volatile inflation of those years, which made partial indexation a costly option. Mechanically, our estimation finds less indexation because inflation is less persistent in the 1970s.

We find it illuminating to combine the evolution of the Calvo parameter θ_{pt} and of indexation χ_{pt}. We do so in figure 2.7 (for their levels) and in figure 2.8 (for their HP-trends). The comparison of both parameters shows that periods of high price rigidities are also periods of low indexation. The converse is true as well, except for the mid-1970s. This result points out that adding indexation as an ad hoc procedure to increase the level of inflation inertia may hide important dynamics in price adjustments.

We repeat our two experiments for wages. Figure 2.9 (in levels) and figure 2.10 (in HP-trends, with inflation superimposed) plot the evolution of the average duration of the spell before workers reoptimize wages, $1(/1 - \theta_{wt})$, in quarter terms. In this case the evidence is more difficult to interpret, with a big spike in the second half of the 1980s, which is probably due to sampling uncertainty. However, it is still the case that during the 1970s, as inflation went up, wage rigidity went down, and as inflation was tamed in the early 1980s, wages again became more rigid.

Figures 2.11 and 2.12 draw the evolution of wage indexation. Here, in

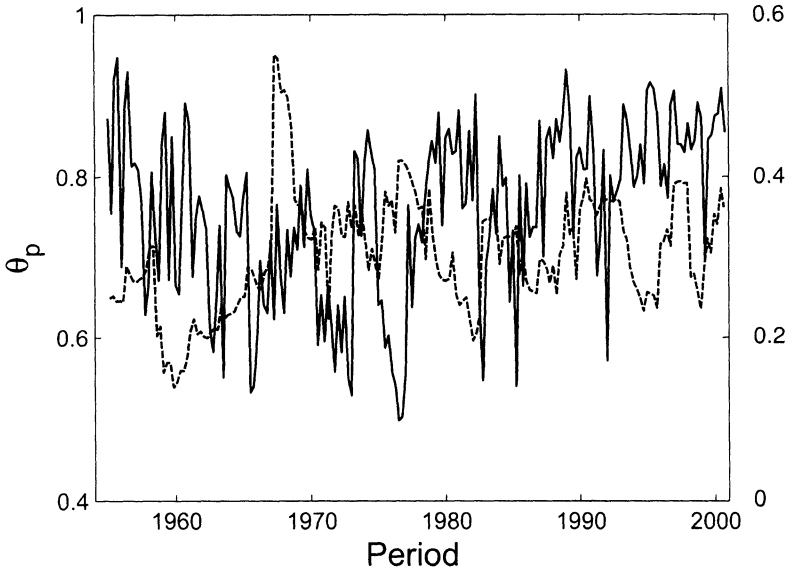

Figure 2.7
Price Rigidity vs. Indexation

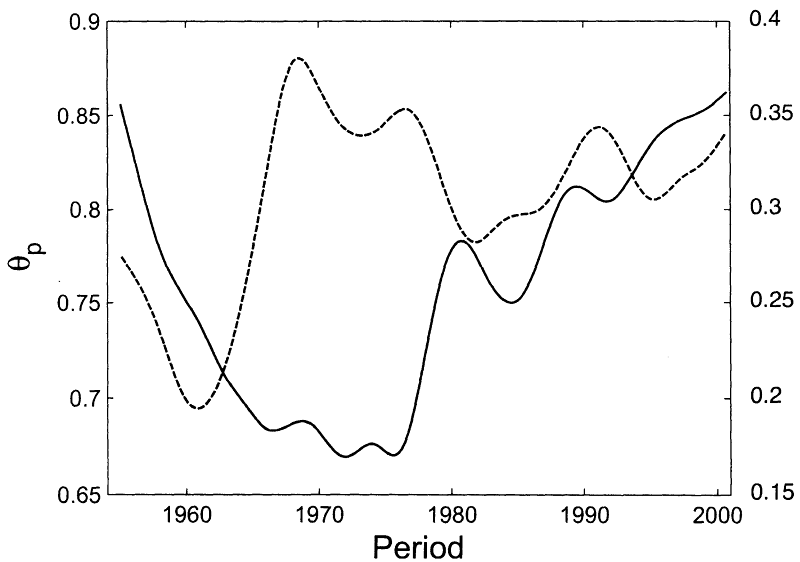

Figure 2.8
HP-Trend Price Rigidity vs. HP-Trend Indexation

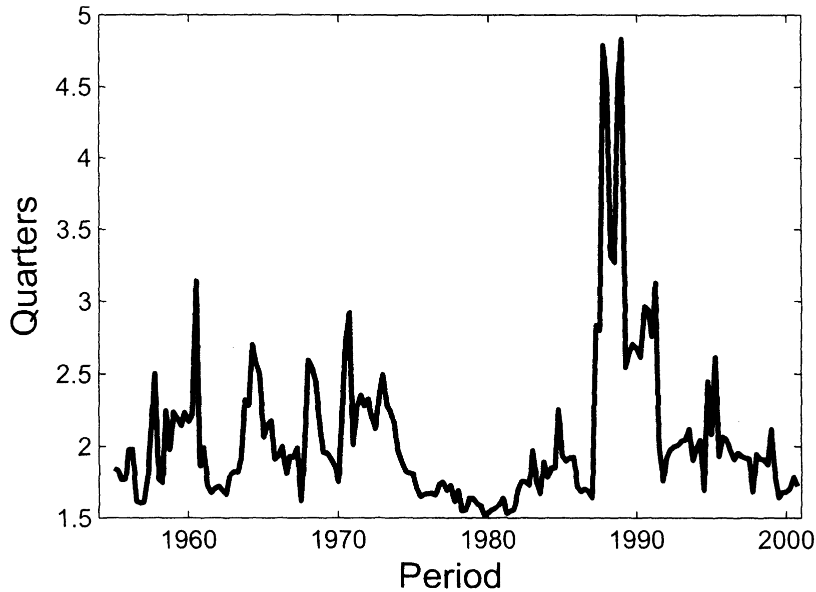

Figure 2.9
Average Wage Duration

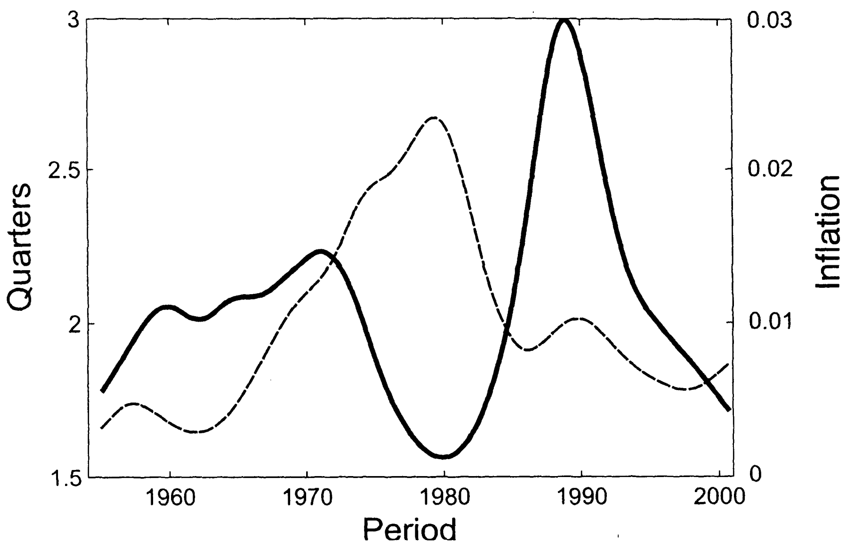

Figure 2.10
HP-Trend Wage Rigidity versus HP-Trend Inflation

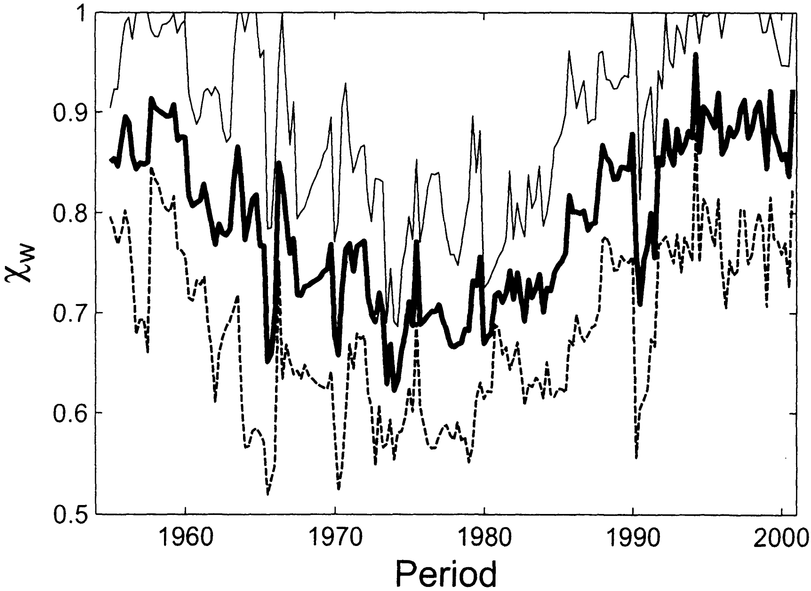

Figure 2.11
Wage Indexation

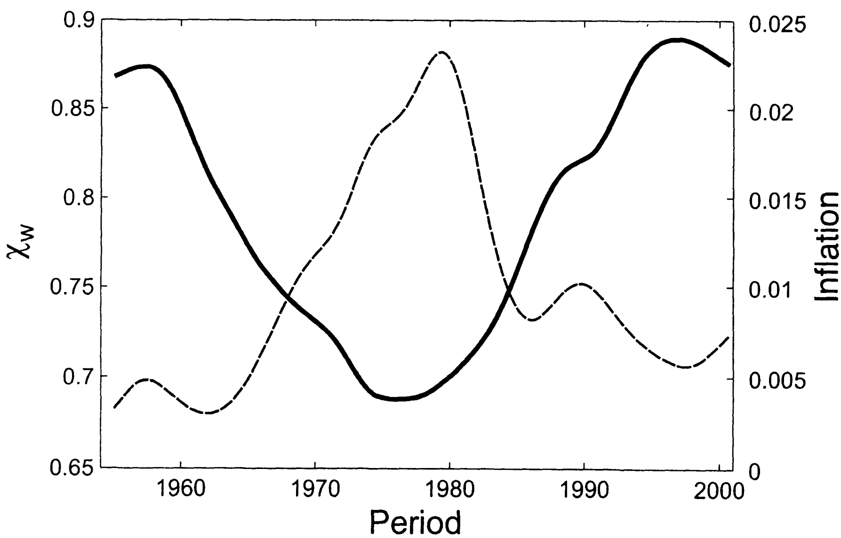

Figure 2.12
HP-Trend Wage Indexation versus HP-Trend Inflation

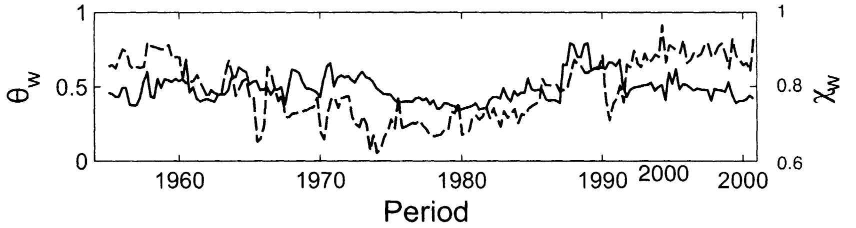

Figure 2.13
Wage Rigidity vs. Indexation

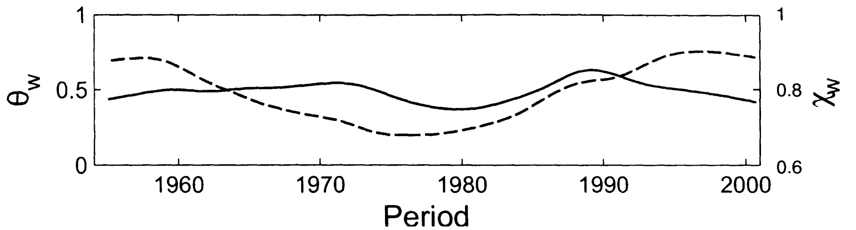

Figure 2.14
HP-Trend Wage Rigidity vs. HP-Trend Indexation

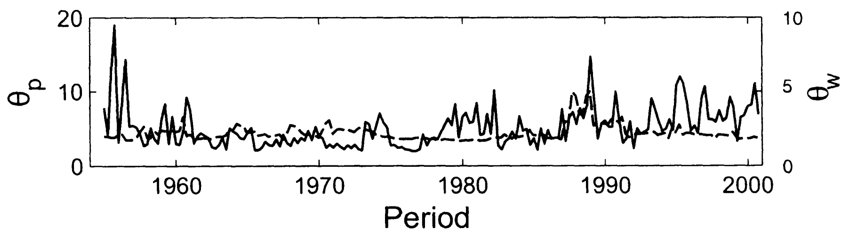

Figure 2.15
Price vs. Wage Rigidity

comparison, the clarity of the result is embarrassing: wage indexation is nearly the perfect mirror of inflation. As we did for prices, we interpret this finding as the natural consequence of workers switching to more often wage reoptimizations that make indexation less of an interesting rule in times of high inflation.[8] Less wage indexation is what the model needs to capture the higher volatility of inflation in the data.

For completeness, we finish our graphical display with figures 2.13 to 2.18, where we plot the evolution of the different parameters controlling nominal rigidities against other. Because of space constraints, we refrain from further discussion of the plots. However, the reader can appreciate

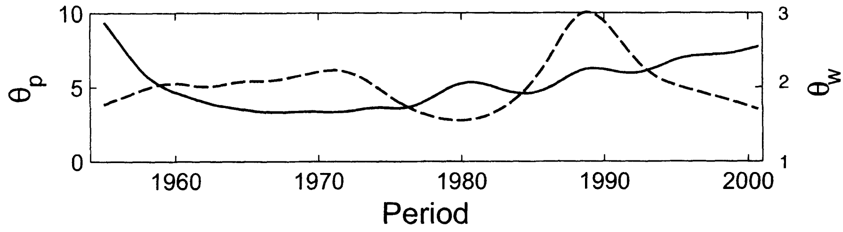

Figure 2.16
HP-Trend Price vs. HP-Trend Wage Rigidity

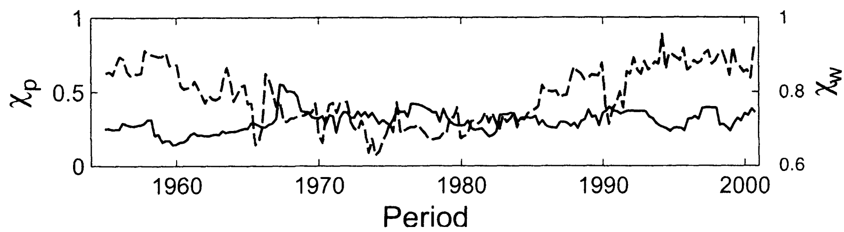

Figure 2.17
Price Indexation vs. Wage Indexation

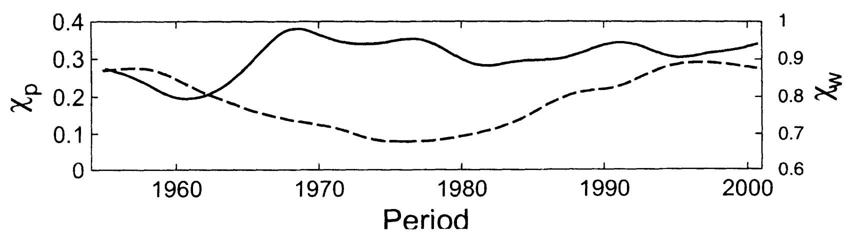

Figure 2.18
HP-Trend Price vs. HP-Trend Wage Indexation

that the similarity in the evolution of the parameters over time solidifies our confidence that we are uncovering a systematic pattern of relationships between nominal rigidities and inflation.

We consider our findings to be strong proof of the changing nature of the nominal rigidities in the economy and of a strong indication of model misspecification along the dimension of price and wage adjustment. Calvo's price adjustment cannot capture the evolution of the parameters are less than 1 in levels (they will always be more than 0 because we are taking logs).

We first report the experiment where we let θ_{pt}, the Calvo parameter of

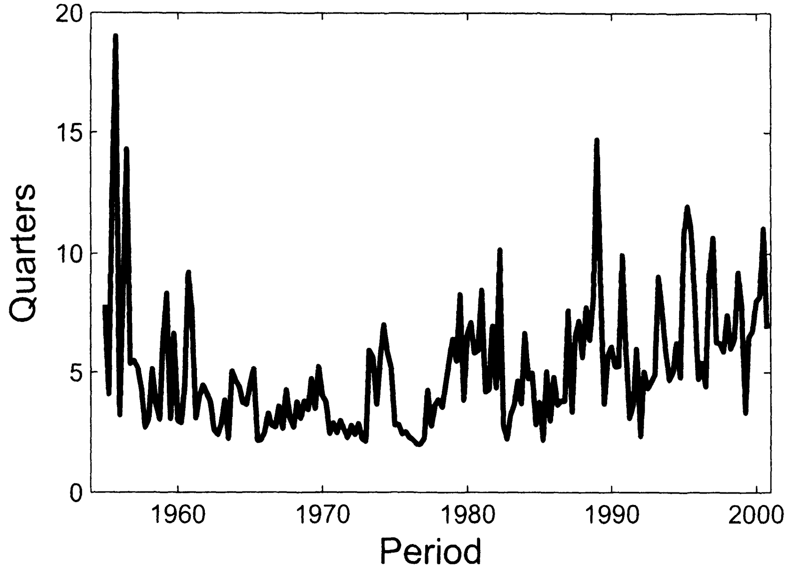

Figure 2.19
Average Price Duration

price changes, evolve over time. We find it more informative (and more directly comparable to the micro evidence) to report the average duration of the spell before the producers reoptimize, $1(/1 - \theta_{pt})$, in quarter terms. Figure 2.19 plots that duration while figure 2.20 plots the Hodrick and Prescott (HP)-trend and, for comparison purposes, the HP-trend of the consumer price index (CPI). In this figure, as in all the rest of the figures of the paper where we plot two different variables, we follow the convention that the continuous line represents the parameter on the left y-axis and the discontinuous one the parameter on the right y-axis.[7]

Figures 2.19 and 2.20 reveal a clear pattern: average duration was high in the late 1950s, dropped quickly in the 1960s, and only started to pick up in the late 1970s, continuing with an upward trend until today. Interestingly enough, the changes in the average duration of the spell before the producers reoptimize are strongly correlated with changes in inflation. In figure 2.20 we see how times of increasing trend inflation (late 1960s, 1970s) are times of falling average duration and vice versa: how times of decreasing trend inflation (the 1980s and the 1990s) are times of increasing average duration.

Our second experiment regarding price rigidities is with χ_{pt}, the parameter that controls price indexation. Figure 2.21 plots the evolution of

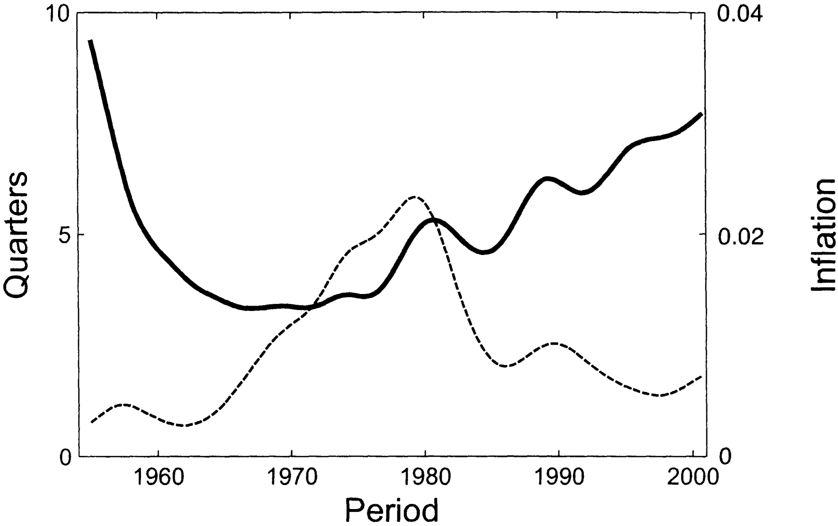

Figure 2.20
HP-Trend Price Rigidity vs. HP-Trend Inflation

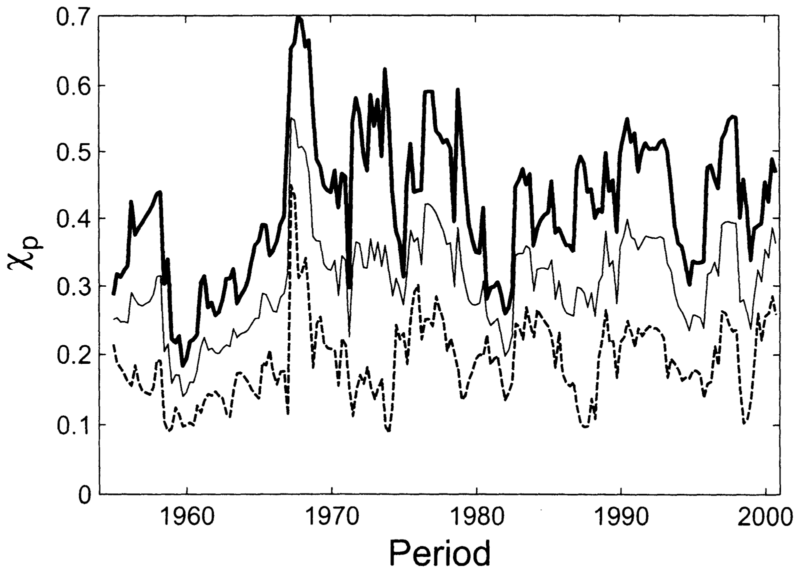

Figure 2.21
Price Indexation

the parameter over the sample plus the fundamentals that determine the pricing decisions of firms and households. Our results underscore that this problem is relevant empirically. Also, they suggest that the evidence in Klenow and Kryvtsov (2005)—that the intensive margin of price changes accounts for 95 percent of the monthly variance of inflator—may be a product of the sample period (1988–2003), where the low level of inflation limits identification because it eliminates the source of variation of the data. Indeed, in our figures 2.7 and 2.8, if we look at the period 1988–2000, we observe less variation in the pricing parameters.

There are at least two possible sources for this misspecification of the pricing mechanism of the model that could rationalize our findings. First, time-varying price and wage rigidity parameters may be revealing a problem of omitted variables. For example, a change in the probability of price adjustment translates into a different slope of the (implicit) Phillips curve in our model and thus, into a variation of inflation. However, in the data, there are other shocks that affect inflation, like the price of energy, the price of commodities, or exchange rate fluctuations. Since we do not include these shocks, we may be capturing the changing influence of these sources of inflation through variations in the Calvo parameters.[9]

The second source of misspecification may be the time-dependent structure of pricing (either à la Calvo as in the model we have presented or à la Taylor). Thus, we can read our results as favoring models of state-dependent pricing (Caballero and Engel 1999, Caplin and Leahy 1991 and 1997), since those have an endogenously changing duration of prices and wages. The extra analytical difficulty implied by state-dependent models (Dotsey, King, and Wolman 1999) may be a price we are forced to pay. Another strand of the literature that may consider our results interesting is the one that deals with sticky information (Mankiw and Reiss 2002, and Sims 2002). Higher inflation increases the incentives to gather information and, hence, it is likely to imply more frequent price and wage adjustments.

Finally, our findings have relevant implications for optimal policy design. First, if we interpret the evolution of parameters like θ_{pt} as exogenously given, it may be something that the monetary authority may condition its behavior on (we do not enter into a discussion of how it would estimate them in real time, we only raise this as a theoretical possibility). Second, if we real our results as showing that the measured amount of price rigidities are endogenous to monetary policy, optimal design becomes tougher than in the basic Ramsey exercises.

7 Conclusion

How structural are the structural parameters of DSGE models? Less so than we often claim. Our analysis indicates that there are large variations in the estimated values of several of the key parameters of a benchmark medium-scale macroeconomic model during our sample period.

We document changes in the response of the monetary authority to inflation and in the inflation target that confirm previous findings by other researchers. In particular, we report a move by the Fed toward a much more aggressive stand against raising prices in the late 1970s. Also, we find that changes in the inflation target account, at most, for 40 to 50 percent of the increase in inflation in the 1970s. Our results are remarkable because they are derived in a context where agents understand that policy evolves over time and respond to that evolution.

We uncover that the parameters controlling nominal rigidities drift in a substantial way, and more important, are strongly correlated with inflation. These findings cast serious doubts on the usefulness of models based on Calvo pricing and invite deeper investigations of state-dependent pricing models.

We do not want our work to be interpreted as a sweeping criticism of the estimation of DSGE models, because it is not. The literature has made impressive progress over the last years and has contributed much to improving our understanding of aggregate fluctuations and the effects of economic policy. We ourselves have been engaged in this research agenda and plan to continue doing so. We hope, instead, that our paper will be read as an invitation to further estimation of DSGE models with parameter drifting. This avenue is promising, both as a mechanism for incorporating richer dynamics and as a diagnostic tool for detecting gross misspecifications.

In fact, as our discussants have rightly pointed out, much remains to be done. We have only scratched the surface of the problem of estimating DSGE models with parameter drifting. We have not explored the model when we have different sources of variations in the parameters at the same time or when there is stochastic volatility in the shocks. Also, we have not studied the consequences of drifting parameters for the dynamics of the business cycle or for the impulse-response functions of the model. Finally, we have not evaluated different specifications of parameter drift or analyzed the possible reasons for parameter drifting in detail.

Our skepticism about the structural nature of most structural parameters is not a call to perform reduced-form exercises. Along with Tom Sargent and Mark Watson (Fernández-Villaverde and Rubio-Ramírez 2007), we have singled out some of the problems of estimating reduced-form models. But there are many other papers emphasizing the weaknesses of reduced-form inference. The fundamental point is that every empirical procedure has strengths and limitations. As Hurwicz (1962) warned us many years ago, just because we name something "structural," we should not believe we have taken the theoretical high ground.

Acknowledgments

Corresponding author: Juan F. Rubio-Ramírez, 213 Social Sciences, Duke University, Durham, NC 27708, USA. E-mail: Juan.Rubio-Ramírez@duke.edu. We thank the editors, Daron Acemoglu, Kenneth Rogoff, and Mike Woodford; our two discussants, Tim Cogley and Frank Schorfheide; Pau Rabanal, Garey Ramey, Stephanie Schmitt-Grohé, and Martín Uribe, and participants at the NBER Macroeconomics Annual conference for comments. 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.

Endnotes

1. Indeed, Hurwicz (1962) himself emphasized the contingency of the definition of structural parameters: "the *concept* of structure *is relative to the domain of modifications anticipated*"; "If two individuals differ with regard to modifications they are willing to consider, they will probably differ with regard to the relations accepted as structural," and "this relativity of the concept of structure is due to the fact that it represents *not a property of the material system under observation,* but rather a property of the anticipations of those asking for predictions concerning the state of the system" (p. 238; italics in the original).

2. The autoregressive coefficients (ρ_R, ρ_{Π_t}, ρ_y) and the matrix Q become in this formulation the new structural parameters. We are also skeptical about their true structural nature, but to avoid the infinite regression problem, we will ignore our doubts for the moment.

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

4. We do not argue that our flat priors are uninformative. After a reparameterization of the model, a flat prior may become highly curved. Moreover, if we wanted to use the model for other purposes like forecasting or to compare it with, for example, a VAR, we would need to elicit our priors more carefully.

5. A word of caution here: the estimates of the standard deviation with the particle filter are relatively unstable (Fernández-Villaverde and Rubio-Ramírez 2007, and DeJong, Dharmarajan, Liesenfeld, and Richard 2007). Computational constraints preclude us from running a simulation sufficiently long to fully avoid this problem.

6. This observation may have motivated a model where Λ_{yd} changes over time, but such models are, as argued by Bansal and Yaron (2004), quite difficult to estimate in small samples.

7. We do not plot the standard deviations interval for the average price duration (nor later for the average wage duration) because the transformation $1(/1-\theta_{pt})$ generates implausibly large upper bounds as soon as the simulation of θ_{pt} travels close to 1. The standard deviations interval for θ_{pt} show, however, that the parameter itself is estimated without too much uncertainty.

8. During the early 1970s, there was a raise in the prevalence of cost-of-living allowance (COLA) escalators in collective bargaining agreements (Hendricks and Kahn 1985). This observation could be used to undermine our result. However, even at their peak, COLAs only covered 6 million workers, a small percentage of the labor force. Moreover, it is difficult to map COLAs from the 1970s into our model, since they had many contingent rules that make them quite different from the naive indexation rules that we use. In fact, it could even be possible to think about a state-contingent COLA as an implicit form of reoptimization.

9. Similarly, part of the variation in the Calvo parameters may be accounted for by markup shocks, which play an important role in models like Smets and Wouters' (2003). However, it is difficult to see which type of markup shocks will have the level of persistence that we observe in the movements of the Calvo parameters that we estimate.

References

Adolfson, M., S. Laséen, J. Lindé, and M. Villani. 2005. Bayesian estimation of an open economy DSGE model with incomplete pass-through. Working Paper Series 179. Stockholm: Sveriges Riksbank (Central Bank of Sweden).

An, S., and F. Schorfheide. 2006. Forthcoming. Bayesian analysis of DSGE models. *Econometric Reviews.*

Andrés, J., P. Burriel, and A. Estrada. 2006. BEMOD: A DSGE model for the Spanish economy and the rest of the Euro area. Working Paper no. 631. Madrid: Banco de España.

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* 50:174–88.

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.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59:1481–1509.

Benati, L. 2006. Investigating inflation persistence across monetary regimes. Unpublished paper. Frankfurt am Main, Germany: European Central Bank.

Bernanke, B. S., and I. Mihov. 1998. Measuring monetary policy. *Quarterly Journal of Economics* 113:869–902.

Bils, M., and P. Klenov. 2004. Some evidence on the importance of sticky prices. *Journal of Political Economy* 112:947–85.

Boivin, J. 2006. Has U.S. monetary policy changed? Evidence from drifting coefficients and real-time data. *Journal of Money, Credit and Banking* 38 (5): 1149–73.

Boivin, J., and M. Giannoni. 2006. DSGE models in a data-rich environment. NBER Working Paper no. 12272. Cambridge, MA: National Bureau of Economic Research.

Browning, M., L. P. Hansen, and J. Heckman. 1999. Micro data and general equilibrium models. In *Handbook of Macroeconomics*, volume 1, chapter 8, ed. J. B. Taylor and M. Woodford, 543–633. The Netherlands: Elsevier.

Caballero, R., and E. Engel. 1999. Heterogeneity and output fluctuations in a dynamic menu-costs economy. *Review of Economic Studies* 60:95–119.

Canova, F. 2005. Monetary policy and the evolution of the U.S. economy. Centre for Economics Policy Research (CEPR) Working Paper no. 5467. London: CEPR.

Canova, F., and L. Sala. 2006. Back to square one: Identification issues in DSGE models. ECB Working Paper no. 583. Frankfurt am Main, Germany: European Central Bank.

Caplin, A., and J. Leahy. 1991. State-dependent pricing and the dynamics of money and output. *Quarterly Journal of Economics* 106:683–708.

———. 1997. Aggregation and optimization with state-dependent pricing. *Econometrica* 65:601–26.

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.

Christoffel, K., G. Coenen, and A. Warne. 2007. The new area-wide model of the Euro area: Specification and first estimation results. *Mimeo.* Frankfurt am Main, Germany: European Central Bank.

———. 2007. Conditional and unconditional forecasting with the new area-wide model of the Euro area. European Central Bank. Unpublished paper.

Clarida, R., J. Galí, and M. Gertler. 2000. Monetary policy rules and macroeconomic stability: Evidence and some theory. *Quarterly Journal of Economics* 115:147–80.

Cogley, T., and T. J. Sargent. 2001. Evolving post-World War II U.S. inflation dynamics. In *NBER Macroeconomics Annual 2001*, ed. B. Bernanke and K. Rogoff, 331–73. Cambridge, MA: MIT Press.

———. 2005. Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S. *Review of Economic Dynamics* 8:262–302.

Cogley, T., and A. Sbordone. 2006. Trend inflation and inflation persistence in the new Keynesian Phillips curve. FRB of New York Staff Report no. 270. New York: Federal Reserve Bank of New York.

Cooley, T. 1971. Estimation in the presence of sequential parameter variation. Unpublished PhD diss., University of Pennsylvania.

Cooley, T. F., and E. C. Prescott. 1976. Estimation in the presence of stochastic parameter variation. *Econometrica* 44:167–84.

———. 1995. Economic growth and business cycles. In *Frontiers of business cycle research,* ed. T. F. Cooley, 1–38. Princeton, NJ: Princeton University Press.

Davig, T., and E. Leeper. 2006a. Forthcoming. Endogenous monetary policy regime change. *NBER International Seminar on Macroeconomics.*

———. 2006b. Forthcoming. Generalizing the Taylor Principle. *American Economic Review.*

Davig, T., E. Leeper, and H. Chung. 2007. Monetary and fiscal policy switching. *Journal of Money, Credit, and Banking* 39 (4): 809–42.

DeJong, D. N., H. Dharmarajan, R. Liesenfeld, and J. Richard. 2007. An efficient filtering approach to likelihood approximation for state-space representations. Working Paper no. 321. University of Pittsburgh.

Del Moral, P., and J. Jacod. 2002. The Monte-Carlod method for filtering with discrete time observations. central limit theorems. In *Numerical Methods and Stochastics,* ed. T. J. Lyons and T. S. Salisbury, 29–54. New York: American Mathematical Society.

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* 114 (2): 655–90.

Doucet, A., N. de Freitas, and N. Gordon. 2001. *Sequential Monte Carlo methods in practice.* New York: Springer Verlag.

Erceg, C. J., L. Guerrieri, and Christopher Gust. 2006. SIGMA: A new open economy model for policy analysis. *International Journal of Central Banking* 2:1–50.

Erceg, C. J., D. Henderson, and A. Levin. 2000. Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics* 46:281–313.

Farmer, R., D. Waggoner, and T. Zha. 2006. Indeterminacy in a forward looking regime switching model. FRB of Atlanta Working Paper no. 2006-19. New York: Federal Reserve Bank.

Fernández-Villaverde, J. 2005. Comments on S. Schmitte-Grohé and M. Uribe, Optimal fiscal and monetary policy in a medium-scale macroeconomic model. In *NBER Macoeconomics Annual 2005,* ed. M. Gertler and K. Rogoff, 427–44. Cambridge, MA: MIT Press.

Fernández-Villaverde, J., and J. Rubio-Ramírez. 2004. Comparing dynamic equilibrium models to data: A Bayesian approach. *Journal of Econometrics* 12:153–87.

———. 2005. Estimating dynamic equilibrium economies: Linear versus nonlinear likelihood. *Journal of Applied Econometrics* 20:891–910.

———. 2007. Forthcoming. Estimating macroeconomic models: A likelihood approach. *Review of Economic Studies.*

Fernández-Villaverde, J., J. Rubio-Ramírez, and M. S. Santos. 2006. Convergence properties of the likelihood of computed dynamic models. *Econometrica* 74:93–119.

Fernández-Villaverde, J., J. Rubio-Ramírez, T. J. Sargent, and M. Watson. 2006. A.B.C.'s (and D)'s for understanding VARs. *American Economic Review.* Forthcoming.

Fisher, J. 2006. The dynamic effects of neural and investment-specific technology shocks. *Journal of Political Economy* 114:413–51.

Galí, J., and P. Rabanal. 2004. Technology shocks and aggregate fluctuations: How well does the RBC model fit postwar U.S. data? In *NBER Macroeconomics Annual 2004*, 225–88. Cambridge: MIT Press.

Geweke, J. 1989. Bayesian inference in econometric models using Monte Carlo integration. *Econometrica* 24:1037–99.

Godsill, S. J., A. Doucet, and M. West. 2004. Monte Carlo smoothing for nonlinear time series. *Journal of the American Statistical Association* 99:156–68.

Greenwood, J., Z. Herkowitz, and P. Krusell. 1997. Long-run implications of investment-specific technological change. *American Economic Review* 87:342–62.

Greenwood, J., Z. Herkowitz, and P. Krusell. 2000. The role of investment-specific technological change in the business cycle. *European Economic Review* 44:91–115.

Hendricks, W. E., and L. Khan. 1985. *Wage indexation in the United States: Cola or uncola?* Cambridge, MA: Ballinger.

Hurwicz, L. 1962. On the structural form of independent systems. In *Logic, methodology, and philosophy of science*, ed. E. Nagel, P. Suppes, and A. Tarski, 232–39. Stanford, CA: Stanford University Press.

Justiniano, A., and G. E. Primiceri. 2005. The time varying volatility of macroeconomic fluctuations. 2006 Meeting: Paper no. 353. New York: Society for Economic Dynamics.

King, T. B. 2006. Dynamic equilibrium models with time-varying structural parameters. Mimeograph. St. Louis, MO: Washington University.

Klenow, P., and O. Kryvstov. 2007. State-dependent or time-dependent pricing: Does it matter for recent U.S. inflation? Unpublished paper. Stanford, CA: Stanford University.

Künsch, H. R. 2005. Recursive Monte Carlo filters: Algorithms and theoretical analysis. *Annals of Statistics* 33:1983–2021.

Lubick, T., and F. Schorfheide. 2004. Testing for indeterminacy: An application to U.S. monetary policy. *American Economic Review* 94:190–217.

Mankiw, G. N., and R. Reiss. 2002. Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics* 117:1295–1328.

McFadden, D. L. 1989. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* 57:995–1026.

Murchison, S., and A. Rennison. 2006. ToTEM: The Bank of Canada's new quarterly projection model. Bank of Canada Technical Report no. 97. Ottawa: Bank of Canada.

Nakamura, E., and J. Steinsson. 2007. Forthcoming. Five facts about prices: A reevaluation of menu cost models. *Quarterly Journal of Economics.*

Oliner, S., G. Rudebusch, and D. Sichel. 1996. The Lucas critique revisited: Assessing the stability of empirical Euler equations for investment. *Journal of Applied Econometrics* 70:291–316.

Orphanides, A. 2002. Monetary policy rules and the great inflation. *American Economic Review* 92:115–20.

Owyang, M. T., and G. Ramey. 2004. Regime switching and monetary policy measurement. *Journal of Monetary Economics* 51:1577–97.

Pakes, A., and D. Pollard. 1989. Simulation and the asymptotics of optimization estimators. *Econometrica* 57:1027–57.

Pitt, M. K., and N. Shephard. 1999. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association* 94:590–99.

Ploberger, W., and P. C. B. Phillips. 2003. Empirical limits for time series econometric models. *Econometrica* 71:627–73.

Primiceri, G. E. 2005. Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies* 72:821–52.

———. 2006. Why inflation rose and fell: Policymakers' beliefs and U.S. postwar stabilization policy. *Quarterly Review of Economics* 121:867–901.

Rabanal, P. 2007. Does inflation increase after a monetary policy tightening? answers based on an estimated DSGE model. *Journal of Economic Dynamics and Control* 31: 906–37.

Rabanal, P., and J. F. Rubio-Ramírez. 2005. Comparing new Keynesian models of the business cycle: A Bayesian approach. *Journal of Monetary Economics* 52:1151–66.

Romer, C., and D. Romer. 2002. The evolution of economic understanding and postwar stabilization policy. In *Rethinking Stabilization Policy,* 11–78. Kansas City: Federal Reserve Bank of Kansas City.

Roberts, G., A. Gelman, and W. Gilks. 1997. Weak convergence and optimal scaling of random walk metropolis algorithms. *Annals of Applied Probability* 7:110–20.

Rosenberg, B. 1968. Varying parameter estimation. Unpublished PhD diss., Harvard University.

Rubin, D. B. 1988. Using the SIR algorithm to simulate posterior distributions. In *Bayesian Statistics 3,* ed. J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, 395–401. Oxford: Oxford University Press.

Santos, M. S., and A. Peralta-Alva. 2005. Accuracy of simulations for stochastic dynamic models. *Econometrica* 73:1939–76.

Sargent, T. J. 1999. *The conquest of American inflation.* Princeton, NJ: Princeton University Press.

Sims, C. A. Macroeconomics and reality. *Econometrica* 48 (1): 1–48.

———. 2001. Discussion of evolving post-World War II U.S. inflation dynamics by T. Cogney and T. J. Sargent. *NBER Macroeconomics Annual 2001,* 373–79.

———. 2003. Implications of rational inattention. *Journal of Monetary Economics* 50 (3): 665–90.

Sims, C. A., and T. Zha. 2006. Were there regime switches in U.S. monetary policy? *American Economic Review* 96:54–81.

Smets, F., and R. Wouters. 2003. An estimated stochastic dynamic general equilibrium model of the Euro area. *Journal of the European Economic Association* 1:1123–75.

Taylor, J. 1998. Monetary policy guidelines for unemployment and inflation stability. In *Inflation, unemployment, and monetary policy,* ed. R. M. Solown and J. B. Taylor, 29–34. Cambridge: MIT Press.

Uhlig, H. 1997. Bayesian vector autoregressions with stochastic volatility. *Econometrica* 65:59–73.

Woodford, M. 2006. Forthcoming. Interpreting inflation persistence: Comments on the Conference on Quantitative Evidence on Price Discrimination. *Journal of Money, Credit and Banking.*

Young, A. T. 2004. Labor's share fluctuations, biased technical change, and the business cycle. *Review of Economic Dynamics* 7:916–31.

8. Appendix

This appendix offers further details about the technical aspects of the paper. First, we discuss some general computational aspects and elaborate on the solution of the model. Second, we describe the particle filter that evaluates the likelihood function of the model. Third, we comment on the estimation procedure. Fourth, we close with the details of the construction of the data.

8.1 Computation of the Model

The most important feature of the algorithm to be described below to solve and estimate the model is that it can be implemented on a good desktop computer. We coded all programs for the perturbation of the model and the particle filter in Fortran 95 and compiled them in Intel Visual Fortran 9.1 to run on Windows-based machines (except some Mathematica programs to generate analytic derivatives). We use a Xeon Processor 5160 EMT64 at 3.00 GHz with 16 GB of RAM.

The solution of the model is challenging because we have 19 state variables plus the drifting parameters that we allow in each empirical exercise. Moreover, we need to recompute the solution of the model for each new set of parameter values in the estimation. The only non-linear procedure that accomplishes this computation in a reasonable amount of time is perturbation (Aruoba, Fernández-Villaverde, and Rubio-Ramírez 2006). We implement our solution by perturbing the equilibrium conditions of the rescaled version of the model (i.e., the one where we have already eliminated the two unit roots) around the deterministic steady state. This means that the solution is locally accurate regardless of the level of technology in the economy. Also, note that the steady state will depend on the level of inflation targeted by the monetary authority.

We use Mathematica to compute the analytical derivatives and to generate Fortran 95 code with the corresponding analytical expression. Then, we load that output into a Fortran 95 code that evaluates the solution of the model for each parameter value as implied by the Metropolis-Hastings algorithm to be described below. The solution will have the form:

$$(\overline{S}'_{t+1}, J'_t)' = \Gamma_{s1}(\overline{S}'_t, \varepsilon'_t)' + \frac{1}{2}(\overline{S}'_t, \varepsilon'_t)\Gamma_{s2}(\overline{S}'_t, \varepsilon'_t)' + \Gamma_{s3} \tag{13}$$

where, recalling our notation, $\overline{S}_t$ are the states, ε_t are the shocks, J_t is a vector of variables of interest in the model that are not states, and the Γ_{si}'s are matrices of the right size. With (13), and by selecting the appropriate rows, we build the state space representation:

$$\overline{S}_{t+1} = \Psi_{s1}(\overline{S}'_t, \varepsilon'_t)' + \frac{1}{2}(\overline{S}'_t, \varepsilon'_t)\Psi_{s2}(\overline{S}'_t, \varepsilon'_t)' + \Psi_{s3}$$

$$\mathcal{Y}^T = \Psi_{o1}(S'_t, \varepsilon'_t)' + \frac{1}{2}(S'_t, \varepsilon'_t)\Psi_{o2}(S'_t, \varepsilon'_t)' + \Psi_{o3}$$

where $S_t = (\overline{S}'_t, \overline{S}'_{t-1}, \varepsilon'_{t-1})$ and $\mathcal{Y}^T = (\Delta \log \mu_t^{-1}, \Delta \log y_t, \Delta \log l_t, \log \Pi_t, \log R_t)'$.

8.2 *Description of the Particle Filter*

We provide now a short description of the particle filter. We will deliberately focus on the intuition of the procedure and we will gloss over many technical issues that are relevant for a successful application of the filter. We direct the interested reader to Fernández-Villaverde and Rubio-Ramírez (2007), where we discuss most of those issues in detail, and the articles in Doucet, de Freitas, and Gordon (2001), which present improved sequential Monte Carlo algorithms, like Pitt and Shephard's (1999) auxiliary particle filter.

As we described in the main text, given the Markov structure of our state space representation, we can factorize the likelihood function as:

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

and obtain the factorization:

$$\mathcal{L}(\mathbb{Y}^T;\Psi) = \int \mathcal{L}(\mathbb{Y}_1 \mid S_0;\Psi)dS_0 \prod_{t=2}^{T} \int \mathcal{L}(\mathbb{Y}_t \mid S_t;\Psi)p(S_t \mid \mathbb{Y}^{t-1};\Psi)dS_t \tag{14}$$

Consequently, if we had the sequence $(p(S_t \mid \mathbb{Y}^{t-1};\Psi))_{t=1}^T$ and $p(S_0;\Psi)$, we could evaluate the likelihood of the model. Santos and Peralta-Alva (2005) show conditions under which we can draw the numerical solution of the model to approximate $p(S_0;\Psi)$. The two difficulties of evaluation (14) are then to characterize the sequence of conditional distributions$[p(S_t \mid \mathbb{Y}^{t-1};\Psi)]_{t=1}^T$ and to compute the different integrals in the expression.

The particle filter begins from the observation that, if somehow we can get N draws of the form $[(s^i_{t|t-1})_{i=1}^N]_{t=1}^T$ from the sequence $(p(S_t \mid \mathbb{Y}^{t-1};\Psi))_{t=1}^T$, we can appeal to a law of large numbers and substitute the integrals with a mean of the conditional likelihoods evaluated in the empirical draws:

$$\mathcal{L}(\mathbb{Y}^T;\Psi) \simeq \frac{1}{N}\sum_{i=1}^{N} \mathcal{L}(\mathbb{Y}_1 \mid s^i_{0|0};\Psi)\sum_{t=2}^{T}\frac{1}{N}\sum_{i=1}^{N}\mathcal{L}(\mathbb{Y}_t \mid s^i_{t|t-1};\Psi)$$

where our notation for the draws indicates in the subindex the conditioning set (i.e., $t \mid t-1$ means a draw at moment t conditional on information until $t-1$) and the superindex denotes the index of the draw. The intuition of the procedure is that we substitute the exact but unknown sequence $[p(S_t \mid \mathbb{Y}^{t-1};\Psi)]_{t=1}^T$ by its empirical counterpart.

How do we draw from $[p(S_t \mid \mathbb{Y}^{t-1};\Psi)]_{t=1}^T$? The second key idea of the particle filter is that we can extend importance sampling (Geweke 1989) to a sequential environment. The following proposition, due in its original form to Rubin (1988), formalizes the idea:

Proposition 1. *Let $[s^i_{t|t-1}]_{i=1}^N$ be a draw from $p(S_t \mid \mathbb{Y}^{t-1};\Psi)$. Let the sequence $[\tilde{s}^i_t]_{i=1}^N$ be a draw with replacement from $[s^i_{t|t-1}]_{i=1}^N$ where the resampling probability is given by*

$$q^i_t = \frac{\mathcal{L}(\mathbb{Y}_t \mid s^i_{t|t-1};\Psi)}{\sum_{i=1}^N \mathcal{L}(\mathbb{Y}_t \mid s^i_{t|t-1};\Psi)},$$

Then $(\tilde{s}^i_t)_{i=1}^N$ is a draw from $p(S_t \mid \mathbb{Y}^t;\Psi)$.

The proposition 1 shows how to recursively use a draw $(s^i_{t|t-1})_{i=1}^N$ from $p(S_t \mid \mathbb{Y}^{t-1};\Psi)$ to get a draw $(s^i_{t|t})_{i=1}^N$ from $p(S_t \mid \mathbb{Y}^t;\Psi)$. This result is crucial.

It allows us to incorporate the information in $\mathbb{Y}_t$ to change our current estimate of S_t. This is why this step is known in filtering theory as update (the discerning reader has probably already realized that this update is nothing more than an application of Bayes' theorem).

The resampling step is key for the success of the filter. A naive extension of Monte Carlo techniques will just draw a whole sequence of $[(s^i_{t|t-1})^N_{i=1}]^T_{t=1}$ without stopping period by period to resample according to proposition 1. Unfortunately, this naive scheme diverges. The reason is that all the sequences become arbitrarily far away from the true sequence of states, which is a zero measure set and the sequence that is closer to the true states dominates all the remaining ones in weight. A simple simulation shows that the degeneracy appears even after very few steps.

Given $(s^i_{t|t})^N_{i=1}$, we draw N exogenous shocks, something quite simple, since the shocks in our model $\varepsilon^i_{t+1} = (\varepsilon^i_{\mu,t+1}, \varepsilon^i_{d,t+1}, \varepsilon^i_{\varphi,t+1}, \varepsilon^i_{A,t+1}, \varepsilon^i_{m,t+1})'$ are normally distributed. Then, we apply the law of motion for states that relates the $s^i_{t|t}$ and the shocks ε^i_{t+1} to generate $(s^i_{t+1|t})^N_{i=1}$. This step, known as forecast, put us back at the beginning of proposition 1, but with the difference that we have moved forward one period in our conditioning.

The following pseudocode summarizes the description of the algorithm:

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

Step 1, Prediction: Sample N values $(s^i_{t|t-1})^N_{i=1}$ using $(s^i_{t-1|t-1})^N_{i=1}$, the law of motion for states and the distribution of shocks ε_t.

Step 2, Filtering: Assign to each draw $(s^i_{t|t-1})$ the weight q^i_t in proposition 1.

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

With the output of the algorithm, we just substitute into our formula

$$\mathscr{L}(\mathbb{Y}^T; \Psi) \simeq \frac{1}{N}\sum_{i=1}^{N} \mathscr{L}(\mathbb{Y}_1 \,|\, s^i_{0|0}; \Psi) \prod_{t=2}^{T} \frac{1}{N}\sum_{i=1}^{N} \mathscr{L}(\mathbb{Y}_t \,|\, s^i_{t|t-1}; \Psi) \tag{15}$$

and get an estimate of the likelihood of the model. Del Moral and Jacod (2002) and Künsch (2005) show weak conditions under which the right-hand side of the previous equation is a consistent estimator of $\mathscr{L}(\mathbb{Y}^T; \Psi)$ and a central limit theorem applies.

8.3 Estimation Procedure

We mention in the main part of the text that the posterior of the model

$$p(\Psi \mid \mathbb{Y}^T) \propto \mathcal{L}(\mathbb{Y}^T; \Psi)p(\Psi)$$

is difficult, if not impossible, to characterize. However, we can draw from it and build its empirical counterpart using a Metropolis-Hastings algorithm. The algorithm is as follows:

Step 0, Initialization: Set $i \rightsquigarrow 0$ and an initial Ψ_i. Solve the model for Ψ_i and build the state space representation. Evaluate prior $p(\Psi_i)$ and approximate $\mathcal{L}(\mathbb{Y}^T; \Psi)$ with (15). Set $i \rightsquigarrow i + 1$.

Step 1, Proposal draw: Get a draw Ψ_i^* from a proposal density $q(\gamma_{i-1}, \gamma_i^*)$.

Step 2, Solving the Model: Solve the model for Ψ_i^* and build the new state space representation.

Step 3, Evaluating the proposal: Evaluate $p(\Psi_i^*)$ and $\mathcal{L}(\mathbb{Y}^T; \Psi_i^*)$ with (15).

Step 4, Accept/Reject: Draw $\chi_i \sim U(0, 1)$. If $\chi_i \leq [\mathcal{L}(\mathbb{Y}^T; \Psi_i^*)p(\Psi_i^*)q(\Psi_{i-1}, \Psi_i^*)]/[\mathcal{L}(\mathbb{Y}^T; \Psi_{i-1})p(\Psi_{i-1})q(\Psi_i^*, \Psi_{i-1})]$ set $\Psi_i = \Psi_i^*$, otherwise $\Psi_i = \Psi_{i-1}$.

Step 5, Iteration: If $i < M$, set $i \rightsquigarrow i + 1$ and go to step 1. Otherwise stop.

This algorithm requires us to specify a proposal density $q(\cdot, \cdot)$. We follow the standard practice and choose a random walk proposal, $\Psi_i^* = \Psi_{i-1} + \kappa_i$, $\kappa_i \sim \mathcal{N}(0, \Sigma_\kappa)$, where Σ_κ is a scaling matrix. This matrix is selected to get the appropriate acceptance ratio of proposals (Roberts, Gelman, and Gilks 1997).

To reduce the "chatter" of the problem, we will keep the innovations in the particle filter (i.e., the draws from the exogenous shock distributions and the resampling probabilities) constant across different passes of the Metropolis-Hastings algorithm. As pointed out by McFadden (1989) and Pakes and Pollard (1989), this is required to achieve stochastic equicontinuity, and even if the condition is not strictly necessary in a Bayesian framework, it reduces the numerical variance of the procedure.

8.4 Construction of Data

As we mention in the text, we compute both real output and real gross investment in consumption units to make the observed series compatible with the model. We define the relative price of investment as the ratio of the investment deflator and the deflator for consumption. The consumption deflator is constructed from the deflators of nondurable goods and services reported in the NIPA. Since the NIPA investment deflators are poorly measured, we use the investment deflator constructed by Fisher (2006). 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, national defense consumption expenditures, federal nondefense consumption expenditures, and state and local government consumption expenditures. We define nominal gross investment as the sum of personal consumption expenditures on durable goods, national defense gross investment, federal government nondefense gross investment, state and local government gross investment, private nonresidential fixed investment, and private residential fixed investment. Per capita nominal output is defined as the ratio between our nominal output series and the civilian noninstitutional population between 16 and 65. Since we need to measure real output per capita in consumption units, we deflate the series by the consumption deflator. For the real gross investment per capita series, we divide our above mentioned nominal gross investment series by the civilian noninstitutional population between 16 and 65 and the consumption deflator. Finally, the hours worked per capita series is constructed with the index of total number of hours worked in the business sector and the civilian noninstitutional population between 16 and 65. Since our model implies that hours worked per capita are between 0 and 1, we normalize the observed series of hours worked per capita such that it is, on average, 0.33.