Fortune or Virtue:

Time-variant Volatilities versus Parameter Drifting

in U.S. Data*

Jesús Fernández-Villaverde[†] Pablo Guerrón-Quintana[‡]

Juan F. Rubio-Ramírez[§]

February 4, 2010

*We thank Frank Schorfheide and Tao Zha for useful 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, the Federal Reserve Bank of Philadelphia, or the Federal Reserve System. Finally, we also thank the NSF for financial support.

[†]University of Pennsylvania, NBER, CEPR, and FEDEA, <jesusfv@econ.upenn.edu>.

[‡]Federal Reserve Bank of Philadelphia, <pablo.guerron@phil.frb.org>.

[§]Duke University, Federal Reserve Bank of Atlanta, and FEDEA, <juan.rubio-ramirez@duke.edu>.

Abstract

This paper compares the role of stochastic volatility versus changes in monetary policy rules in accounting for the time-varying volatility of U.S. aggregate data. Of special interest to us is understanding the sources of the great moderation of business cycle fluctuations that the U.S. economy experienced between 1984 and 2008. We build a medium-scale dynamic stochastic general equilibrium (DSGE) model with both stochastic volatility and parameter drifting in the Taylor rule. We estimate the model using U.S. data and Bayesian methods. Methodologically, we show how to confront such a rich model with the data by exploiting the structure of the high-order approximation to the decision rules that characterize the equilibrium of the economy. Our main substantive finding is that, even after controlling for stochastic volatility, there is evidence of a change in monetary policy rules during Volcker's tenure as Chairman of the Fed. At the same time, we document how good volatility shocks played an important role in the great performance of the economy during the 1990s.

Keywords: DSGE Models, Stochastic Volatility, Parameter drifting, Bayesian methods.

JEL classification numbers: E10, E30, C11.

1. Introduction

This paper addresses one of the main open questions in empirical macroeconomics: what is the role of time-varying disturbance variances versus changes in monetary policy rules in accounting for the evolving volatility of U.S. aggregate data? This discussion is particularly relevant to understand the sources of the Great Moderation of business cycle fluctuations that the U.S. economy experienced between 1984 and 2007 and to forecast whether low volatility will return after the turbulences of 2008-2009.[1] To answer this question, we build a medium-scale dynamic stochastic general equilibrium (DSGE) model with both stochastic volatility in the disturbances that drive the dynamics of the economy and parameter drifting in the Taylor rule followed by the monetary authority. Then, we estimate the model using U.S. data and Bayesian methods. We use our results to run a battery of counterfactual exercises in which we build artificial histories of economies where some source of variation has been eliminated or modified in an illustrative manner.

The motivation for this investigation is transparent. Time-varying volatility tells a history built around the changing size of the variance of shocks that hit the economy. The Great Moderation is, then, a tale of fortune: for two decades and a half we were favored by fate in the form of small shocks. It is also a pessimistic perspective: we dwell in joy during periods of low volatility and we struggle through times of high volatility, but there is disappointingly little scope for the policy maker to battle the elements. Therefore, our current turbulences may be the opening stages of a era of large business cycle swings.

Parameter drifting constructs a radically divergent account: it believes that fundamental changes in the economy were the cause of higher stability. Some versions of the parameter drifting narrative emphasize technological change. Two commonly cited factors are better inventory control (McConnell and Pérez-Quirós, 2000, Ramey and Vine, 2006) or financial innovation (Dynan, Elmendorf, and Sichel, 2006 or Guerrón-Quintana, 2009a). Other versions of the parameter drift history, the most prominent of which is Clarida, Galí, and Gertler (2000), single out better monetary policy as the key for the reduce size of business cycle fluctuations. Thus, parameter drifting is a tale of virtue: thanks to either better technologies (inventory control, the move to services) or better policies, the economy is more stable than before. It is also an optimistic view. As long as we do not abandon new technologies or unlearn

[1] Kim and Nelson (1999), McConell and Pérez-Quirós (2000), and Blanchard and Simon (2001) were the first papers to point out that time-varying volatility was an important component of U.S. aggregate fluctuations. While Kim and Nelson and McConell and Pérez-Quirós emphasized a big change in volatility around 1984, Blanchard and Simon saw the Great Moderation as part of long-run trend towards lower volatility only momentarily interrupted during the 1970s. Stock and Watson (2002) undertake a thorough review of the evidence.

the lessons of monetary economics, we should expect the great moderation to continue, current maladies notwithstanding.

There is evidence in favor of different forms of parameter drifting. Besides the classic work by Clarida, Galí, and Gertler (2000), basic references in this line of research include Cogley and Sargent (2002), Lubick and Shorfheide (2004), Boivin and Giannoni (2006), or Canova (2009). More recently and most directly related with out investigation, Fernández-Villaverde and Rubio-Ramírez (2008) report compelling evidence of parameter drifting in the parameters that control the monetary policy rule and in the degree of nominal rigidities in a standard DSGE model.

But there has been another branch of the literature, which appeared largely as a response to Clarida, Galí, and Gertler (2000) and the other follow-up papers, that has presented a strong case in favor of time-varying volatility of disturbances. Perhaps the most influential paper in this tradition is Sims and Zha (2006). Relying on a Structural Vector Autorregresion (SVAR) with regime switching, Sims and Zha find that the model that best fits the data only has changes over time in the variances of structural disturbances and no variation in the monetary rule or in the private sector of the model. Even when they allow for policy regime changes, Sims and Zha do not find that the estimated changes are large enough to account for the evolution of observed volatility.[2]

Using similar approaches, other papers also find support for this view. Among others, we can cite Cogley and Sargent (2005), Primiceri (2005), and Canova and Gambetti (2009). In general, once time-varying volatility is allowed, SVARs find little support for the tale of virtue; fortune seems to be the preferred option.

But SVARs approaches face challenges of their own. Benati and Surico (2009) use data generated from a simple New Keynesian DSGE model to show how SVARs may misinterpret changes in policy by changes in variances because changes in policy also have implications for the volatility of endogenous variables (we can think about this argument as one instance of the Lucas' critique). They read their results as suggesting that existing SVAR evidence may be uninformative to the question at hand.

To avoid these problems we can follow a perspective more firmly grounded on explicit equilibrium models. First attempts in that line of work are Justiniano and Primiceri (2007) and Fernández-Villaverde and Rubio-Ramírez (2007) who estimate DSGE economies that incorporate the stochastic volatility of the shocks. Both papers show that such models fit the data much better than traditional economies with constant variance. However, neither of these two papers compares the stochastic volatility specification with one with changes in

[2]Furthermore, Sims and Zha (2006) reject single-equation approaches because they require the use of instruments, which the authors argue rely on implausible restriction assumptions and fragile identification.

policy rules.

The natural next step is, thus, to estimate a DSGE that can measure how much of the volatility change observed in the U.S. aggregate data can be attributed to either fortune or virtue. The project is challenging because, to get an econometrically satisfactory answer, we need to simultaneously allow for both stochastic volatility and parameter drifting. A "one-at-a-time" approach is fraught with peril. If we only allow one source of variation in the model, the likelihood may want to take advantage of this extra degree of flexibility to fit the data better. For example, if we have parameter drifting in nominal rigidities, a DSGE model with stochastic volatility may interpret this drift as time-varying volatility in mark-up shocks. Analogously, if we have time-varying volatility in technological shocks, a DSGE model with only parameter drifting may conclude, erroneously, that the parameters of the Taylor rule are changing.

Our contributions are both methodological and substantive. Methodologically, we show how to confront a rich DSGE model with the data by exploiting the structure of the high-order approximation to the decision rules that characterize the equilibrium of the economy. We prove a theorem, for a general class of DSGE models, that characterizes the structure of these decision rules. This theorem allows us to handily evaluate the likelihood function of the model. As an added bonus, this approach allows us to estimate the model without measurement errors in observables. One of the advantages of having stochastic volatility is that we multiply the number of random shocks to the model by two: for each exogenous stochastic process, we have a shock to level and a shock to volatility. We take advantage of this profusion of shocks to dispense from measurement errors.

Our main substantive finding is that, even after controlling for stochastic volatility (and there is a fair amount of it), there is evidence of change in monetary policy during the analyzed period. We establish this fact with two exercises. First, we obtain the smoothed series of the estimated response of the policy rule to inflation. These series shows a steep increase at the arrival of Volcker and a fast drop after the arrival of Greenspan. In fact, the response of monetary policy to inflation is back to the levels of Burns-Miller times by the early 1990s. Interestingly, the strong change in monetary policy during Volcker's tenure is consistent with one of the policy regimes identified in Sims and Zha (2006).

Second, we construct counterfactual histories feeding alternative policy rules to different periods of time. Our main finding is that had Volcker been the Chairman during Burns-Miller or Greenspan times, the Fed would have responded more aggressively to inflation at a low output cost. The intuition is that, the tougher stand of monetary policy (which is fully observed by the agents in our model) would have been enough to lower inflation and generate low interest rates. At the same time, we also document how good volatility shocks played an important role in the good performance of the economy during the 1990s. Greenspan was, indeed, a lucky Chairman.

Besides the papers cited before, there is a large literature on this topic and we cannot do justice here to it.[3] Instead, we selectively review some of them that we found particularly compelling. Galí and Gambetti (2009) present evidence that the Great Moderation in the U.S. was associated with large changes in comovements among variables, both conditional and unconditionally, and in impulse response functions (IRFs). The authors interpret this evidence as difficult to reconcile with a pure "good luck" history of the Great Moderation. Also, regime-switching models in policy rules such as those of Farmer, Waggoner, and Zha (2006a and b) or Bianchi (2009) provide with an extra degree of flexibility in modelling aggregate dynamics that is highly promising. More recently, Liu, Waggoner, and Zha (2009) show with a Markov-regime switching model that allows for changes in the policy rules that the switch from the dovish regime to the hawkish regime may not be the main source of substantial reductions in the volatilities of inflation and output. The reason is because, even if in the dovish regime is very different from the hawkish one, the expectation that we may switch from one regime to another dampens the effects on aggregate dynamics. However, this paper is not estimated and, therefore, it is more an intriguing hypothesis than an empirical investigation.

The rest of the paper is organized as follows. Section 2 describes the benchmark model that we will use for our exercise and section 3 defines the equilibrium in our economy and our approximated solution method. Section 4, the core of the methodological contribution, presents the description of how we evaluate the likelihood and our estimation method, including the theorem that characterizes the structure of the solution of high-order approximations to DSGE models with stochastic volatility. After describing the data and the estimation approach in section 5, the substantive results appear section 6, which reports our parameter estimates, section 7, with a discussion of the impulse-response functions of the model, section 8, which talks about model fit and smoothed estimates of shocks, volatilities, and policy parameters, and section 9, which constructs counterfactuals histories. Section 10 concludes and an appendix provides further details on some technical aspects of the paper (proof of the main theorem, computation, and construction of the data).

[3] Also, from the time-series perspective, other papers that estimate SVARs with time-varying parameters or stochastic volatility include Uhlig (1997), Bernanke and Mihov (1988), Cogley and Sargent (2005), and Primiceri (2005).

2. A Benchmark Model

We adopt as the benchmark economy for our empirical investigation what has become the standard New Keynesian DSGE model in the literature (Woodford, 2003). The model is based on Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007) and we have used it, without stochastic volatility, in Fernández-Villaverde and Rubio-Ramírez (2008) and in Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2009). The model has many strengths but also important weaknesses that fully acknowledge. Suffice it to say here that, since this model has been the base of much applied policy analysis by central banks,[4] it is the natural laboratory for this paper.

Since the model is well know, our presentation will be brief.[5]A continuum of households consume, save, hold real money balances, supply labor, and determine their own wages subject to a demand curve and nominal rigidities in the form of Calvo's pricing with partial indexation. The final good is produced by a representative firm that aggregates a continuum of intermediate goods produced by monopolistic competitive firms. These firms manufacture the intermediate good by renting the labor supplied and the capital accumulated by the households. Intermediate good producers also face nominal rigidities in the form of Calvo's pricing with partial indexation. The model is closed with a monetary authority that fixes the one-period nominal interest rate according to a Taylor policy rule through open market operations. In our specification, we introduce long-run growth through two unit roots, one in the neutral technology and one in the investment-specific technology. Stochastic volatility appears as changing the standard deviation of the five structural shocks to the model (two shocks to preferences, two shocks to technology, and one shock to monetary policy). Parameter drifting appears as changing values of the parameters in the Taylor policy rule.

2.1. Households

The economy is populated by a continuum of households indexed by j. The household j's preferences are representable by a lifetime utility function:

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

[4]Closely related models are used by the Federal Reserve Board (Edge, Killey, and Laforte, 2006), the European Central Bank (Christoffel, Coenen, and Warne, 2008) or the Bank of Sweden (Adolfson *et al.*, 2005).

[5]The interested reader can find the web document, `www.econ.upenn.edu/~jesusfv/benchmark_DSGE.pdf`, where we present the model without stochastic volatility or parameter drifting in careful detail.

which is separable in consumption, c_{jt}, real money balances, m_{jt}/p_t, and hours worked, l_{jt}. In our notation, $\mathbb{E}_0$ is the conditional expectation operator, β is the discount factor, h controls habit persistence, ϑ is the inverse of Frisch labor supply elasticity, d_t is a shifter to intertemporal preference that follows:

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

and φ_t is a labor supply shifter that evolves as:

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

These shifters are common to all household and provide flexibility for the equilibrium dynamics of the model to capture fluctuations in interest rates and changes in hours worked not accounted by variations in consumption levels.

The principal novelty of these preferences is that, for both shifters d_t and φ_t, the standard deviations, $\sigma_{d,t}$ and $\sigma_{\varphi,t}$, of their innovations, $\varepsilon_{d,t}$ and $\varepsilon_{\varphi,t}$, are indexed by time, that is, they stochastically move period by period according to the autoregressive processes:

$$\log \sigma_{d,t} = \left(1-\rho_{\sigma_d}\right)\log \sigma_d + \rho_{\sigma_d}\log \sigma_{d,t-1} + \eta_d u_{d,t} \text{ where } u_{d,t} \sim \mathcal{N}(0,1)$$

and

$$\log \sigma_{\varphi,t} = \left(1-\rho_{\sigma_\varphi}\right)\log \sigma_\varphi + \rho_{\sigma_\varphi}\log \sigma_{\varphi,t-1} + \eta_\varphi u_{\varphi,t} \text{ where } u_{\varphi,t} \sim \mathcal{N}(0,1).$$

Our specification for the law of motion of the standard deviations of the innovations is parsimonious and it only introduces four new parameters, ρ_{σ_d}, ρ_{σ_φ}, η_d, and η_φ. At the same time, it is surprisingly powerful in capturing some important features of the data (Shepard, 2005).

We can think about the shocks to preferences and to their stochastic volatility as reflecting the random evolution of more complicated phenomena. For example, stochastic volatility may appear as the consequence of changing demographic structures. An economy with older agents might be both less patient because of higher mortality risk (in our notation, a lower d_t) and less prone to reallocations in the labor force because of longer attachments to particular jobs (in our notation, a lower $\sigma_{\varphi,t}$).

We assume complete financial markets: households can trade a whole set of securities contingent on idiosyncratic and aggregate events. An amount of those securities, a_{jt+1}, which pay one unit of consumption in event $\omega_{j,t+1,t}$, is traded at time t at (real) unitary price $q_{jt+1,t}$. We drop the dependence on the event to ease the notational burden. In addition, households also hold b_{jt} government bonds that pay a nominal gross interest rate of R_t. Therefore, the

$j-th$ household's budget constraint is given by:

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

where x_t is investment, w_{jt} is the real wage, r_t the real rental price of capital, $u_{jt}>0$ the rate of use of capital, $\mu_t^{-1}\Phi\left[u_{jt}\right]$ is the depreciation cost of utilizing capital at rate u_{jt} in terms of the final good, μ_t is an investment-specific technological level, T_t is a lump-sum transfer, and $\digamma_t$ is the profits of the firms in the economy. We specify that

$$\Phi\left[u\right]=\Phi_1\left(u-1\right)+\frac{\Phi_2}{2}(u-1)^2$$

a form that satisfies the standard conditions that $\Phi\left[1\right]=0$, $\Phi'\left[\cdot\right]=0$, and $\Phi''\left[\cdot\right]>0$. This function carries the normalization that $u=1$ in the balanced growth path of the economy. Using the relevant first-order conditions, we can find $\Phi_1=\Phi'\left[1\right]=\widetilde{r}$ where $\widetilde{r}$ is the (rescaled) balanced growth path rental price of capital (determined by all the other parameters in the model). This will leave us with only one free parameter, Φ_2.

The capital accumulated by household j at the end of period t is given by:

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

where δ is the depreciation rate and $V\left[\cdot\right]$ is a quadratic adjustment cost function:

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

with adjustment parameter κ. This function is written in deviations with respect to the balanced growth rate of investment, Λ_x. Therefore, along the balanced growth path, $V\left[\Lambda_x\right]=V'\left[\Lambda_x\right]=0$.

Our third structural shock, the investment-specific technology level μ_t, follows a random walk in logs:

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

where Λ_μ is the drift of the process and $\varepsilon_{\mu,t}$ is the innovation to its growth rate (see Greenwood, Herkowitz, Krusell, 1997, for the classic motivation for this shock). In a similar way to the standard deviation of the innovations to the preference shocks, the standard deviation $\sigma_{\mu,t}$

evolves as an autoregressive process:

$$\log \sigma_{\mu,t} = \left(1 - \rho_{\sigma_\mu}\right) \log \sigma_\mu + \rho_{\sigma_\mu} \log \sigma_{\mu,t-1} + \eta_\mu u_{\mu,t} \text{ where } u_{\mu,t} \sim \mathcal{N}(0,1).$$

Again, we can think about stochastic volatility as a stand-in for a more detailed explanation of technological progress in capital production that we do not model explicitly.

We can define two Lagrangian multipliers, λ_{jt}, the multiplier associated with the budget constraint, and q_{jt} (the marginal Tobin's Q), the multiplier associated with the investment adjustment constraint normalized by λ_{jt}. Thus, the first order conditions of the household problem with respect to c_{jt}, b_{jt}, u_{jt}, k_{jt}, and x_{jt} can be written as:

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

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

$$r_t = \mu_t^{-1} \Phi' \left[u_{jt}\right],$$

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

and

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

We need more work to find the optimality condition with respect to labor and wages because of the presence of monopolistic competition and nominal rigidities. Each household j supplies a slightly different type of labor services l_{jt} that are aggregated by a "labor packer" into homogenous labor l_t^d with the production function:

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

that is rented to intermediate good producers at the wage w_t. The "labor packer" is perfectly competitive and it takes all differentiated labor wages w_{jt} and the wage w_t as given.

The first order conditions of the labor "packer" imply a demand function for labor:

$$l_{jt} = \left(\frac{w_{jt}}{w_t}\right)^{-\eta} l_t^d \qquad \forall j \tag{2}$$

and, together with a zero profit condition $w_t l_t^d = \int_0^1 w_{jt} l_{jt} dj$, an expression for the wage:

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

Households follow a Calvo pricing mechanism when they set their wages. At the start of every period, a randomly selected fraction $1 - \theta_w$ of households can reoptimize their wages (where, by an appropriate law of large numbers, individual probabilities and aggregate fractions are equal). All other households simply index their wages given past inflation with an indexation parameter $\chi_w \in [0, 1]$. Therefore, the real wage of a household j that has not changed wages for τ periods is:

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

Since we postulated above both complete financial markets for the households and separable utility in consumption, the marginal utilities of consumption are the same for all households. Thus, in equilibrium, $c_{jt} = c_t$, $u_{jt} = u_t$, $k_{jt-1} = k_t$, $x_{jt} = x_t$, $q_{jt} = q_t$, $\lambda_{jt} = \lambda_t$, and $w_{jt}^* = w_t^*$.

The last two equalities are the most relevant to simplify our analysis: they tell us that the shadow cost of consumption is equated across households and that all households that can reset their wages optimally will do it at the same level w_t^*. With these two results, and after several steps of algebra, we find that the evolution of wages is described by two 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

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

on the auxiliary variable f_t.

Taking advantage that, in every period, a fraction $1 - \theta_w$ of households set w_t^* as their wage and the remaining fraction θ_w partially index their price by past inflation, we can write the law of motion of real wage as:

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

2.2. The Final Good Producer

There is one final good producer that aggregates a continuum of intermediate goods according to the production function:

$$y_t^d = \left(\int_0^1 y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}} \tag{3}$$

where ε is the elasticity of substitution.

The final good producer is perfectly competitive and minimizes its costs subject to the production function (3) and taking as given all intermediate goods prices p_{ti} and the final good price p_t. The optimality conditions of this problem result in a demand functions for each intermediate good with the classic form:

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

where y_t^d is the aggregate demand and a price for the final good:

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

2.3. Intermediate Good Producers

Each of the intermediate goods is produced by a monopolistic competitor whose technology is given by a Cobb-Douglas production function with a fixed cost:

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

where k_{it-1} is the capital rented by the firm, l_{it}^d is the amount of the "packed" labor input rented by the firm, the parameter ϕ corresponds to the fixed cost of production, and A_t (our fourth structural shock) is neutral productivity that follows:

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

In this specification, Λ_A is the drift of the neutral technological level and $\varepsilon_{A,t}$ is the innovation to its growth rate.

The time-varying standard deviation of this innovation evolves stochastically following our, by now already familiar, specification:

$$\log \sigma_{A,t} = \left(1 - \rho_{\sigma_A}\right) \log \sigma_A + \rho_{\sigma_A} \log \sigma_{A,t-1} + \eta_A u_{A,t} \text{ where } u_{A,t} \sim \mathcal{N}(0,1).$$

The technology is translated by a fixed cost parameter ϕ and a scale variable $z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$. Given our definitions of neutral productivity A_t and investment-specific productivity μ_t:

$$\log z_t = \Lambda_z + \log z_{t-1} + z_{z,t}$$

where $\Lambda_z = \frac{\Lambda_A + \alpha \Lambda_\mu}{1-\alpha}$, $z_{z,t} = \frac{z_{A,t} + \alpha z_{\mu,t}}{1-\alpha}$, $z_{A,t} = \sigma_{A,t} \varepsilon_{A,t}$, and $z_{\mu,t} = \sigma_{\mu,t} \varepsilon_{\mu,t}$. We can think about z_t as the weighted level of technology in the economy, where the weight is given by the elasticity of output with respect to capital. The constant Λ_z is the average growth rate of the economy. The role of ϕ is to make economic profits roughly equal to zero. We scale it by z_t to keep the fixed costs constant in relative terms to the technology level. Finally, note that $z_{z,t}$ will also have a stochastic volatility structure product of the mixture of two processes with stochastic volatility themselves.

Intermediate good producers produce the quantity demanded of the good by renting l_{it}^d and k_{it-1} at prices w_t and r_t. Then, by minimization, we have 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 is constant for all firms and all production levels given A_t, w_t, and r_t.

The quantity sold of the good is determined by the demand function derived above. Given this demand function, the intermediate good producers set prices to maximize profits. However, when they do so, they follow the same Calvo pricing scheme as households. In each period, a fraction $1 - \theta_p$ of intermediate good producers reoptimize their prices. All other firms partially index their prices by past inflation with an indexation parameter $\chi \in [0, 1]$.

Therefore, prices are set to solve the problem:

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

subject to

$$y_{it+\tau} = \left(\prod_{s=1}^{\tau} \Pi_{t+s-1}^{\chi} \frac{p_{it}}{p_{t+\tau}} \right)^{-\varepsilon} y_{t+\tau}^d.$$

In this problem, future profits are discounted using the pricing kernel of the economy, $\beta^\tau \lambda_{t+\tau}/\lambda_t$, (which is the right valuation criteria from the perspective of the households) and the probability of the event "only indexation for τ periods", θ_p^τ.

The solution for the pricing problem of the firm has a recursive structure in two new auxiliary variables g_t^1 and g_t^2 that take the form:

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

and

$$\varepsilon g_t^1 = (\varepsilon - 1) g_t^2$$

where

$$\Pi_t^* = \frac{p_t^*}{p_t}$$

is the ratio between the optimal new price (common across all firms that can reset their prices) and the price of the final good.

With this structure, the price index follows:

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

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

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

2.4. The Monetary Authority

The model is closed by the presence of a monetary authority that sets the nominal interest rates through open market operations financed with lump-sum transfers T_t and a balanced budget. The monetary authority follows a modified Taylor rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_{\Pi,t}} \left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\exp\left(\Lambda_{y^d}\right)}\right)^{\gamma_{y,t}}\right)^{1-\gamma_R} e^{\sigma_{m,t}\varepsilon_{mt}}. \tag{4}$$

The first term on the right hand side, $\frac{R_{t-1}}{R}$, represents a desire for interest rate smoothing, expressed in terms of R, the balanced growth path nominal return of capital. The second term, $\frac{\Pi_t}{\Pi}$, a "inflation gap," responds to the deviation of inflation to its balanced growth path level Π. The third term,

$$\frac{\frac{y_t^d}{y_{t-1}^d}}{\exp\left(\Lambda_{y^d}\right)}$$

is a "growth gap:" the ratio between the growth rate of the economy and Λ_{y^d}, the balanced path gross growth rate of y_t^d. The term, ε_{mt}, is the innovation to the monetary policy shock and follows a $\mathcal{N}(0,1)$ process with a time-varying standard deviation, $\sigma_{m,t}$, that follows an autoregressive process:

$$\log \sigma_{m,t} = \left(1-\rho_{\sigma_m}\right) \log \sigma_m + \rho_{\sigma_m} \log \sigma_{m,t-1} + \eta_m u_{m,t}.$$

In this policy rule, we have two drifting parameters: the responses of the monetary authority, $\gamma_{\Pi,t}$ and $\gamma_{y,t}$, to the inflation and growth gap. The parameters drift over time in an autoregressive fashion:

$$\log \gamma_{\Pi,t} = \left(1-\rho_{\gamma_\Pi}\right) \log \gamma_\Pi + \rho_{\gamma_\Pi} \log \gamma_{\Pi,t-1} + \eta_\pi \varepsilon_{\pi,t} \text{ where } \varepsilon_{\pi,t} \sim \mathcal{N}(0,1)$$

and

$$\log \gamma_{y,t} = \left(1-\rho_{\gamma_y}\right) \log \gamma_y + \rho_{\gamma_y} \log \gamma_{y,t-1} + \eta_y \varepsilon_{y,t} \text{ where } \varepsilon_{y,t} \sim N(0,1).$$

We assume here that the agents perfectly observe the changes in monetary policy parameters. A more plausible scenario would involved some filtering in real-time by the agents who need to learn the stand of the monetary authority from observed decisions. A similar argument can be made for the values of the standard deviations of all the other shocks in the economy. But since this learning would further complicate what it is already a large model, we leave this extension for future work.

2.5. Aggregation

Aggregate demand is given by:

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

By relying on the observation that capital-labor ratio is constant across firms, we can derive that aggregate supply is:

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

where:

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

is the aggregate loss of efficiency induced by price dispersion of the intermediate good producers.

Market clearing requires that

$$y_t = y_t^d = y_t^s.$$

By the properties of 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, demanded labor is given by:

$$l_t^d = \frac{1}{v_t^w} \int_0^1 l_{jt} dj$$

where:

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

is the aggregate loss of labor input induced by wage dispersion among differentiated types of labor. Again, by Calvo's pricing, this inefficiency evolves as:

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

3. Equilibrium

We can characterize an equilibrium in our economy by piling all the first order conditions of the household and firms, the Taylor rule of the monetary authority, and market clearing.

This equilibrium is not stationary because we have two unit roots in the processes for technology. However, it is easy to circumvent this problem by rescaling the model $\widetilde{c}_t = \frac{c_t}{z_t}$, $\widetilde{\lambda}_t = \lambda_t z_t$, $\widetilde{r}_t = r_t \mu_t$, $\widetilde{q}_t = q_t \mu_t$, $\widetilde{x}_t = \frac{x_t}{z_t}$, $\widetilde{w}_t = \frac{w_t}{z_t}$, $\widetilde{w}_t^* = \frac{w_t^*}{z_t}$, $\widetilde{k}_t = \frac{k_t}{z_t \mu_t}$, and $\widetilde{y}_t^d = \frac{y_t^d}{z_t}$. The model is stationary in these transformed variables and therefore, along the balanced growth path:

$$\Lambda_c = \Lambda_x = \Lambda_w = \Lambda_{w^*} = \Lambda_{y^d} = \Lambda_z.$$

The equilibrium does not have a closed form solution and we need to resort to a numerical approximation to compute it. There are three challenges in this computation. The first challenge is that we have a large state vector S_t with 19 components:

$$S_t = \begin{pmatrix} \Lambda, \widehat{R}_{t-1}, \widehat{\widetilde{k}}_{t-1}, \widehat{v}_{t-1}^p, \widehat{v}_{t-1}^w, \widehat{\widetilde{w}}_{t-1}, \widehat{\widetilde{c}}_{t-1}, \widehat{\Pi}_{t-1}, \widehat{\widetilde{x}}_{t-1}, \widehat{\widetilde{y}}_{t-1}, \\ \widehat{d}_{t-1}, \widehat{\varphi}_{t-1}, \widehat{\gamma}_{\Pi,t-1}, \widehat{\gamma}_{y,t-1}, \widehat{\sigma}_{d,t-1}, \widehat{\sigma}_{\varphi,t-1}, \widehat{\sigma}_{\mu,t-1}, \widehat{\sigma}_{A,t-1}, \widehat{\sigma}_{m,t-1} \end{pmatrix}'$$

where we have expressed each variable var_t in terms of log deviation with respect to the steady state, $\widehat{var}_t = \log var_t - \log var$, and Λ is the perturbation parameter; with a vector $W_{1,t}$ of five innovations to exogenous shocks (two to preferences, two to technology, and one to monetary policy) and two innovations to the parameter drifts (one to the response to inflation and one to the response to output):

$$W_{1,t} = \left(\varepsilon_{d,t}, \varepsilon_{\varphi,t}, \varepsilon_{\mu,t}, \varepsilon_{A,t}, \varepsilon_{m,t}, \varepsilon_{\pi,t}, \varepsilon_{y,t}\right)';$$

and with a vector $W_{2,t}$ of five innovations to volatility shocks:

$$W_{2,t} = \left(u_{d,t}, u_{\varphi,t}, u_{\mu,t}, u_{A,t}, u_{m,t}\right)'.$$

The second challenge is that, since we have stochastic volatility, an inherently non-linear structure, standard linearization techniques cannot be applied. More pointedly, if we linearized the model, stochastic volatility would disappear from the scene because the solution of the model would be certainty equivalent. The third challenge is that, since we will need to compute the model for a large number of different parameter values in our estimation process, speed is of the outmost importance.

Fortunately, perturbation methods provide a nice solution to the computation of the model. Beyond being extremely fast, perturbation offers high levels of accuracy even relatively far away from the perturbation point (Aruoba, Fernández-Villaverde, and Rubio-Ramírez, 2006). Therefore, we perform a second order perturbation around the deterministic steady-state of the model. The quadratic terms of this approximation allows us to capture, to a large extent, the effects of time-varying volatility while keeping computational complexity at a reasonable level.

Define $\mathbb{S}_t = \left(S_t', W_{1,t}', W_{2,t}'\right)$, a vector that stacks states and shocks. The solution of the model is then given by a transition equation for states:

$$S_{t+1} = \begin{pmatrix} \Psi_{s1}^1 \mathbb{S}_t' \\ \vdots \\ \Psi_{s19}^1 \mathbb{S}_t' \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \mathbb{S}_t \Psi_{s1}^2 \mathbb{S}_t' \\ \vdots \\ \mathbb{S}_t \Psi_{s19}^2 \mathbb{S}_t' \end{pmatrix} \tag{5}$$

and an observation equation:

$$\mathbb{Y}_t = \mathbb{C} + \begin{pmatrix} \Psi_{o1}^1 \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right)' \\ \vdots \\ \Psi_{o5}^1 \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right)' \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right) \Psi_{o1}^2 \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right)' \\ \vdots \\ \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right) \Psi_{o5}^2 \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right)' \end{pmatrix} \tag{6}$$

where $\mathbb{Y}_t$ is a vector of observables for the econometrician (five in our case) and $\mathbb{C}$ is a vector of means of these observables. The lagged vector $\mathbb{S}_{t-1}$ appears in the equation because, as we will see momentarily, our observables have components in first differences.

In these equations, Ψ_{si}^{1} is a 1×31 vector and Ψ_{si}^{2} is a 31×31 matrix for $i = 1, \ldots, 19$. The first term is the linear solution of the model while the second term is the quadratic component of the solution. Similarly, Ψ_{oi}^{1} is a 1×62 vector and Ψ_{oi}^{2} a 62×62 matrix for $i = 1, \ldots, 5$ and the interpretation of each term is the same as before: the first term is the linear component and the second one the quadratic component of the solution.

It is important to emphasize that we are *not* assuming the presence of any measurement error. Although we consider measurement errors both plausible and relevant, in our exercise we want to measure how time-varying volatility and parameter drifting help accounting for the data. Consequently, we eliminate measurement errors to sharpen our analysis. This decision also helps to illustrate how DSGE models with stochastic volatility have a profusion of shocks that we can exploit in our estimation. We will elaborate on this last point below.

The transition equation (5) is unique (up to an equivalence class of representations) but the observable equation (6) is not because it depends on what we assume the econometrician can observe. Boivin and Giannoni (2006) and Guerrón-Quintana (2009b) discuss the consequences for inference of selecting different observables.

We pick as observables the first difference of the log of the relative price of investment, the log federal funds rate, log inflation, the first difference of log output, and the first difference of log real wages, or in our notation:

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

This implies that $\mathbb{C} = \left(\Lambda_\mu, \log R, \log \Pi, \Lambda_{y^d}, \Lambda_w\right)'$. Later, when we take the model to the data, we will let the likelihood pick these means, since Λ_μ, $\log R$, $\log \Pi$, Λ_{y^d}, and Λ_w depend on the structural parameters. We select these variables because they bring us information about aggregate behavior (output), the stand of monetary policy (the interest rate and inflation), and the different shocks (the relative price of investment about investment-specific technological change, the other four variables about technology and preference shocks) that we are concerned about.

The state space representation generated by the transition equation (5) and the measurement equation (6) has an interesting structure that we exploit to evaluate the likelihood of the model. In the next section, we present a general description of that structure and how it applies to our economy.

4. Stochastic Volatility and Evaluation of the Likelihood

In this section we explain how to evaluate the likelihood function of our model. If we allow $\mathbb{Y}_t^{data}$ to be the data counterpart of our observables, $\mathbb{Y}^{data,t} = \left(\mathbb{Y}_1^{data}, \dots, \mathbb{Y}_t^{data}\right)$ to be the history up to time t of our observables, $W_1^t = (W_{1,1}, \dots, W_{1,t})$ to be the history up to time t of the level shocks, and $W_2^t = (W_{2,1}, \dots, W_{2,t})$ to be the history up to time t of the volatility shocks, we can write the likelihood as:

$$p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \gamma\right) = \prod_{t=1}^{T} p\left(\mathbb{Y}_t = \mathbb{Y}_t^{data} | \mathbb{Y}^{data,t-1}; \gamma\right), \tag{7}$$

where

$$p\left(\mathbb{Y}_t = \mathbb{Y}_t^{data} | \mathbb{Y}^{data,t-1}; \gamma\right)$$
$$= \int\int\int p\left(\mathbb{Y}_t = \mathbb{Y}_t^{data} | W_1^t, W_2^{t-1}, S_0; \gamma\right) p\left(W_1^t, W_2^{t-1}, S_0 | \mathbb{Y}^{data,t-1}; \gamma\right) dW_1^t dW_2^{t-1} dS_0. \tag{8}$$

Computing this likelihood is a difficult problem. It cannot be evaluated exactly and deterministic integration problems are too slow for practical use (we have three integrals period per period over large dimensions). Instead, we use Sequential Monte Carlo method to obtain a numerical estimate of (8).[6] As shown in Fernández-Villaverde and Rubio-Ramírez (2007), conditional on having N draws of $\left\{w_1^{t,i}, w_2^{t-1,i}, s_0^i\right\}_{i=1}^N$ from the densities $p\left(W_1^t, W_2^{t-1}, S_0 | \mathbb{Y}^{data,t-1}; \gamma\right)$ (which we will explain later how we generate), a law of large numbers imply that the integral (8) can be approximated by:

$$p\left(\mathbb{Y}_t = \mathbb{Y}_t^{data} | \mathbb{Y}^{data,t-1}; \gamma\right) \simeq \frac{1}{N} \sum_{i=1}^{N} p\left(\mathbb{Y}_t = \mathbb{Y}_t^{data} | w_1^{t,i}, w_2^{t-1,i}, s_0^i; \gamma\right) \tag{9}$$

Hence, we need to evaluate:

$$p\left(\mathbb{Y}_t = \mathbb{Y}_t^{data} | w_1^{t,i}, w_2^{t-1,i}, s_0^i; \gamma\right) \tag{10}$$

for each draw. This evaluation step is crucial not only because it is a term in (9), but also because, in the Sequential Monte Carlo that we will implement, we need (10) to resample from the draws from $p\left(W_1^t, W_2^{t-1}, S_0 | \mathbb{Y}^{data,t-1}; \gamma\right)$ and, in that way, get draws from $p\left(W_1^{t+1}, W_2^t, S_0 | \mathbb{Y}^{data,t}; \gamma\right)$.

[6] This is not the only possible algorithm to do so, although it is a procedure that we have found useful in previous work. Alternives include DeJong *et al.* (2007), Kim, Shephard, and Chib (1998), Fiorentini, Sentana, and Shepard (2004), and Fermanian and Salanié (2004).

The measurement equation (6) implies that evaluating (10) involves solving the equation:

$$\mathbb{Y}_t^{data} = \mathbb{C} + \begin{pmatrix} \Psi_{o1}^1 \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right)' \\ \vdots \\ \Psi_{o5}^1 \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right)' \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right) \Psi_{o1}^2 \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right)' \\ \vdots \\ \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right) \Psi_{o5}^2 \left(\mathbb{S}_t, \mathbb{S}_{t-1}\right)' \end{pmatrix} \tag{11}$$

for $W_{2,t}$ given $\mathbb{Y}_t^{data}, w_1^{t,i}, w_2^{t-1,i}, s_0^i$. Since (11) is quadratic, we will have 2^5 different solutions to this equation. We are not aware of any accurate and efficient way to find these 2^5 different solutions. This problem would seem to prevent us from achieving our goal of evaluating the likelihood function of this model.

But, in this section, we show how considering stochastic volatility allows us to convert the above described quadratic problem into a linear and simpler one. In particular, we illustrate how, when stochastic volatility is present in the problem, equation (11) has only one solution. Moreover, that solution can be found by simply inverting a matrix. Thanks to this insight, the evaluation of the likelihood function becomes possible.[7]

The key of our approach is to note that, when stochastic volatility is considered, the optimal policies functions of many dynamic general equilibrium models share a particular pattern that we can exploit. To make this point more generally, we switch in the next few paragraphs to a somehow more abstract notation.

The set of equilibrium conditions of a wide variety of dynamic general equilibrium model (included the one described in the paper) can be written as:

$$\mathbb{E}_t f \left(\mathcal{Y}_{t+1}, \mathcal{Y}_t, \mathcal{S}_{t+1}, \mathcal{S}_t, \mathcal{Z}_{t+1}, \mathcal{Z}_t\right) = 0 \tag{12}$$

where $\mathbb{E}_t$ is the expectation operator conditional on information available at time t, $\mathcal{Y}_t = (\mathcal{Y}_{1,t}, \mathcal{Y}_{2,t}, \ldots, \mathcal{Y}_{k,t})$ is the vector of non-predetermined variables of size k, $\mathcal{S}_t = (\mathcal{S}_{1,t}, \mathcal{S}_{2,t}, \ldots, \mathcal{S}_{n,t})$ is the vector of endogenous predetermined variables of size n, $\mathcal{Z}_t = (\mathcal{Z}_{1,t}, \mathcal{Z}_{2,t}, \ldots, \mathcal{Z}_{m,t})$ is the vector of exogenous predetermined variables of size m (which, for simplicity, we often call structural or exogenous shocks), and f maps $\mathbb{R}^{2\times k+2\times n+2\times m}$ into $\mathbb{R}^{k+n+m}$.

We want to consider the case where the exogenous shocks follow a stochastic volatility

[7] Stochastic volatility may also help to circumvent a problem of some DSGE models: stochastic singularity. In general, we need at least as many shocks as observables for the likelihood function to be well defined. This requirement forces researchers to add extra shocks or measurement errors in situations where they might have not desired to do so. Stochastic volatility, by introducing an additional volatility shock for each structural shock, doubles the number of shocks in the model. Even if in our model, this is not necessary (we have five observables and five exogenous shocks), it is important to remember that the researcher may want to take advantage of this extra flexibility and either to augment the number of observables or to reduce the number of shocks.

process of the form:

$$\mathcal{Z}_{i,t+1} = \rho_i \mathcal{Z}_{i,t} + \Lambda \sigma_{i,t+1} \varepsilon_{i,t+1}$$

where the standard deviation of the innovations evolves as:

$$\log \sigma_{i,t+1} = \vartheta_i \log \sigma_{i,t} + \Lambda \eta_i u_{i,t+1}$$

for all $i = \{1, \ldots, m\}$. To ease notation, we are assuming that all exogenous shocks have stochastic volatility. It is straightforward yet cumbersome to generalize the notation to other cases (in fact, in our model, some of the shocks, the ones in the parameter drifting, do not have stochastic volatility).

The solution to the model given in equation (12) can be summarized by the following two equations, one describing the evolution of predetermined variables:

$$\mathcal{S}_{t+1} = h\left(\mathcal{S}_t, \mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t, \Lambda\right) \tag{13}$$

and one describing the evolution of non-predetermined ones:

$$\mathcal{Y}_t = g\left(\mathcal{S}_t, \mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t, \Lambda\right), \tag{14}$$

where $\Sigma_t = (\log \sigma_{1,t}, \log \sigma_{2,t}, \ldots, \log \sigma_{m,t})$, $\mathcal{E}_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \ldots, \varepsilon_{m,t})$, and $\mathcal{U}_t = (u_{1,t}, u_{2,t}, \ldots, u_{m,t})$ (this assumes that the volatility shocks are uncorrelated, a restriction that could be relaxed by the appropriate extension of the state space). To clarify notation, we can think of Σ_t as the volatility shocks, $\mathcal{E}_t$ are the innovations to the exogenous or structural shocks, and $\mathcal{U}_t$ innovations to volatility shocks. Sometimes, we will also call Σ_t the time-varying standard deviation of the innovations to the exogenous shocks.

We wish to find a second-order approximation of the functions $h(\cdot) : \mathbb{R}^{n+(4\times m)+1} \rightarrow \mathbb{R}^n$ and $g(\cdot) : \mathbb{R}^{n+(4\times m)+1} \rightarrow \mathbb{R}^k$ around the steady state, $\mathcal{S}_t = \mathcal{S}$ and $\Lambda = 0$. Therefore, we need to characterize the first and second order derivatives of the functions $h(\cdot)$ and $g(\cdot)$ evaluated at the steady state. The following theorem shows that the first partial derivatives of $h(\cdot)$ and $g(\cdot)$ with respect to any component of $\mathcal{U}_t$ and Σ_{t-1} evaluated at the steady is zero, that is, volatility shocks, $\log \sigma_{i,t}$, and their innovations, $u_{i,t}$, do not affect the linear component of the optimal decision rule of the agents for any $i = \{1, \ldots, m\}$ (the same occurs with the perturbation parameter Λ). A similar result has been already established by Schmitt-Grohé and Uribe (2004) for the homoskedastic shocks case. More importantly, the theorem shows (among other results) that the second partial derivative of $h(\cdot)$ and $g(\cdot)$ with respect to $u_{i,t}$ and any other variable but $\varepsilon_{i,t}$ is also zero for any $i = \{1, \ldots, m\}$.

Theorem 1. *Let us denote $[\Upsilon_\omega]^i_j$ as the derivative of the $i-th$ element of generic function Υ with respect to the $j-th$ element generic variable ω evaluated at the non-stochastic steady state (where we drop this index if ω is unidimensional). Then, for the dynamic equilibrium model specified in equation (12), we have that:*

$$\left[h_{\Sigma_{t-1}}\right]^{i_1}_j = \left[g_{\Sigma_{t-1}}\right]^{i_2}_j = [h_{\mathcal{U}_t}]^{i_1}_j = [g_{\mathcal{U}_t}]^{i_2}_j = [h_\Lambda]^{i_1} = [g_\Lambda]^{i_2} = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j \in \{1,\ldots,m\}$.

Furthermore, if we denote $[\Upsilon_{\omega\xi}]^i_{j_1,j_2}$ as the derivative of the $i-th$ element of generic function Υ with respect to the j_1-th element generic variable ω and the j_2-th element generic variable ξ evaluated at the non-stochastic steady state (where again we drop the index for unidimensional variables), we have that:

$$[h_{\Lambda,\mathcal{S}_t}]^{i_1}_j = [g_{\Lambda,\mathcal{S}_t}]^{i_2}_j = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j \in \{1,\ldots,n\}$,

$$\left[h_{\Lambda,\mathcal{Z}_{t-1}}\right]^{i_1}_j = \left[g_{\Lambda,\mathcal{Z}_{t-1}}\right]^{i_2}_j = \left[h_{\Lambda,\Sigma_{t-1}}\right]^{i_1}_j = \left[g_{\Lambda,\Sigma_{t-1}}\right]^{i_2}_j = [h_{\Lambda,\mathcal{E}_t}]^{i_1}_j = [g_{\Lambda,\mathcal{E}_t}]^{i_2}_j = [h_{\Lambda,\mathcal{U}_t}]^{i_1}_j = [g_{\Lambda,\mathcal{U}_t}]^{i_2}_j = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j \in \{1,\ldots,m\}$,

$$\left[h_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = [h_{\mathcal{S}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_2} = [g_{\mathcal{S}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, $j_1 \in \{1,\ldots,n\}$, and $j_2 \in \{1,\ldots,m\}$,

$$\left[h_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = \left[h_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = \left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = 0$$

and

$$\left[h_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_2}_{j_1,j_2} = \left[h_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} = \left[g_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_2}_{j_1,j_2} = [h_{\mathcal{U}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_2} = [g_{\mathcal{U}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j_1, j_2 \in \{1,\ldots,m\}$, and

$$\left[h_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = [h_{\mathcal{E}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_2} = [g_{\mathcal{E}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j_1, j_2 \in \{1,\ldots,m\}$ if $j_1 \neq j_2$.

Proof. See Appendix. ■

Since the statement of the theorem is long and cumbersome, it is useful to clarify it with a table where we put the second derivatives of $h(\cdot)$ and $g(\cdot)$ with respect to the different variables $(\mathcal{S}_t, \mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t, \Lambda)$. The way to read the table is as follows. Take an arbitrary entry, for instance entry (1,2), $\mathcal{S}_t \mathcal{Z}_{t-1} \neq 0$. In this entry, we state that the cross-derivatives of $h(\cdot)$ and $g(\cdot)$ with respect to $\mathcal{S}_t$ and $\mathcal{Z}_{t-1}$ are different from zero. Similarly, entry (3,3), $\Sigma_{t-1}\mathcal{U}_t = 0$, tells us that the cross-derivatives of $h(\cdot)$ and $g(\cdot)$ with respect to Σ_{t-1} and $\mathcal{U}_t$ are all zero. Entries (3,2) and (4,2) have a "*" to denote that the only cross-derivatives of those entries that are different from zero are those that correspond to the same index j (that is, the cross derivatives of each innovation to the exogenous shocks with respect to its own volatility shock and the cross derivatives of the innovation to the exogenous shocks to the innovation to its own volatility shock). The lower triangular part of the table is empty because of the symmetry of second derivatives.

Table 4.1: Second Derivatives

$\mathcal{S}_t\mathcal{S}_t \neq 0$	$\mathcal{S}_t\mathcal{Z}_{t-1} \neq 0$	$\mathcal{S}_t\Sigma_{t-1} = 0$	$\mathcal{S}_t\mathcal{E}_t \neq 0$	$\mathcal{S}_t\mathcal{U}_t = 0$	$\mathcal{S}_t\Lambda = 0$
$\mathcal{Z}_{t-1}\mathcal{Z}_{t-1} \neq 0$	$\mathcal{Z}_{t-1}\Sigma_{t-1} = 0$	$\mathcal{Z}_{t-1}\mathcal{E}_t \neq 0$	$\mathcal{Z}_{t-1}\mathcal{U}_t = 0$	$\mathcal{Z}_{t-1}\Lambda = 0$	
$\Sigma_{t-1}\Sigma_{t-1} = 0$	$\Sigma_{t-1}\mathcal{E}_t \neq 0^*$	$\Sigma_{t-1}\mathcal{U}_t = 0$	$\Sigma_{t-1}\Lambda = 0$		
$\mathcal{E}_t\mathcal{E}_t \neq 0$	$\mathcal{E}_t\mathcal{U}_t \neq 0^*$	$\mathcal{E}_t\Lambda = 0$			
$\mathcal{U}_t\mathcal{U}_t = 0$	$\mathcal{U}_t\Lambda = 0$				
$\Lambda\Lambda \neq 0$					

Table 4.1 tells us that, of the 21 possible sets of second derivatives, 12 are zero and 9 are not. The implications for the decision rules of agents and for the equilibrium function are striking. The perturbation parameter, Λ, will only have a coefficient different from zero in the term where it appears in a square of itself. This term is a constant that corrects for precautionary behavior induced by risk. Volatility shocks, Σ_{t-1}, appear with coefficients different from zero only in the term where they are multiplied by the exogenous shocks or the innovation to its own exogenous shock. Finally, innovations to the volatility shocks, $\mathcal{U}_t$, also appear with coefficients different from zero when they show up with the innovation to their own exogenous shock $\mathcal{E}_t$.

The main implication of Theorem 1 for our goal of evaluating the likelihood function is that, of the terms that complicate our work, only the ones associated with $[h_{\mathcal{E}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_1}$ and $[g_{\mathcal{E}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_1}$ are different from zero. As we will see in the next corollary, this result has an important yet rather direct implication for the structure of observable equation.

Corollary 2. *The second order approximation to the measurement equation (6) can be written as:*

$$\mathbb{Y}_t = \mathbb{C} + \begin{pmatrix} \Psi_{o1}^{1}\left(\mathbb{S}_t,\mathbb{S}_{t-1}\right)' \\ \vdots \\ \Psi_{o5}^{1}\left(\mathbb{S}_t,\mathbb{S}_{t-1}\right)' \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)\Psi_{o1}^{2,1}\left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)' \\ \vdots \\ \left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)\Psi_{o5}^{2,1}\left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)' \end{pmatrix} + \begin{pmatrix} W_{1,t}'\Psi_{o1}^{2,2} \\ \vdots \\ W_{1,t}'\Psi_{o5}^{2,2} \end{pmatrix} W_{2,t}$$

where $\Psi_{oi}^{2,1}$ denotes the cross derivative between elements of $\left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)$ and $\Psi_{oi}^{2,2}$ denotes the cross derivative between elements of $W_{1,t}$ and elements of $W_{2,t}$ for $i\in\{1,\dots,5\}$.

We are now ready to evaluate the likelihood function. Using corollary 2 and if we define:

$$\mathbb{A}_t\left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right) \equiv \mathbb{Y}_t^{data} - \mathbb{C} - \begin{pmatrix} \Psi_{o1}^{1} \\ \vdots \\ \Psi_{o5}^{1} \end{pmatrix}\left(\mathbb{S}_t,\mathbb{S}_{t-1}\right)' - \frac{1}{2}\begin{pmatrix} \left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)\Psi_{o1}^{2,1}\left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)' \\ \vdots \\ \left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)\Psi_{o5}^{2,1}\left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)' \end{pmatrix}'$$

and

$$\mathbb{B}_t\left(W_{1,t}'\right) \equiv \begin{pmatrix} W_{1,t}'\Psi_{o1}^{2,2} \\ \vdots \\ W_{1,t}'\Psi_{o5}^{2,2} \end{pmatrix}$$

we have that:

$$p\left(\mathbb{Y}_t = \mathbb{Y}_t^{data}|W_1^t,W_2^{t-1},S_0;\gamma\right) = \det\left(\mathbb{B}_t^{-1}\left(W_{1,t}'\right)\right)\Pr\left(W_{2,t} = \mathbb{B}_t^{-1}\left(W_{1,t}'\right)\mathbb{A}_t\left(S_t',W_{1,t}',\mathbb{S}_{t-1}\right)\right)$$

which is direct to evaluate given that we know $\mathbb{B}_t$ and the distribution of $W_{2,t}$.

With this expression, evaluated at N draws of $\left\{w_1^{t,i},w_2^{t-1,i},s_0^i\right\}_{i=1}^N$ from the densities $p\left(W_1^t,W_2^{t-1},S_0|\mathbb{Y}^{data,t-1};\gamma\right)$, the likelihood (8) can be approximated by:

$$p\left(\mathbb{Y}_t = \mathbb{Y}_t^{data}|\mathbb{Y}^{data,t-1}\right) \simeq \frac{1}{N}\sum_{i=1}^{N}\det\left(\mathbb{B}_t^{-1}\left(w_1^{t,i\prime}\right)\right)\Pr\left(w_2^{t,i} = \mathbb{B}_t^{-1}\left(w_1^{t,i\prime}\right)A_t\left(S_t',w_1^{t,i\prime},\mathbb{S}_{t-1}\right)\right) \tag{15}$$

In addition, we can find the importance weights for each draw:

$$q_t^i = \frac{\det\left(\mathbb{B}_t^{-1}\left(w_1^{t,i\prime}\right)\right)\Pr\left(w_2^{t-1,i} = \mathbb{B}_t^{-1}\left(w_1^{t,i\prime}\right)\mathbb{A}_t\left(S_t', w_1^{t,i\prime}, \mathbb{S}_{t-1}\right)\right)}{\frac{1}{N}\sum_{i=1}^{N}\det\left(\mathbb{B}_t^{-1}\left(w_1^{t,i\prime}\right)\right)\Pr\left(w_2^{t-1,i} = \mathbb{B}_t^{-1}\left(w_1^{t,i\prime}\right)\mathbb{A}_t\left(S_t', w_1^{t,i\prime}, \mathbb{S}_{t-1}\right)\right)} \tag{16}$$

that we will use momentarily to update our swarm of particles.

To generate the N draws of $\left\{w_1^{t,i}, w_2^{t-1,i}, s_0^i\right\}_{i=1}^N$, we rely on a Sequential Monte Carlo that proceeds as follows (see Fernández-Villaverde and Rubio-Ramírez, 2007, for details):

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

Step 1, Prediction: `Sample` N `values` $\left\{s_{t|t-1}^i\right\}_{i=1}^N$ `using` $\left\{s_{t-1|t-1}^i\right\}_{i=1}^N$, `the law of motion for states and the distribution of shocks` $\left\{w_1^t, w_2^{t-1}\right\}$.

Step 2, Filtering: `Assign to each draw` $\left(s_{t|t-1}^i\right)$ `the weight` q_t^i `in (16).`

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

Del Moral and Jacod (2002) and Künsch (2005) prove, under weak conditions, that this Sequential Monte Carlo delivers a consistent estimator of the likelihood function and that a central limit theorem applies.

5. Data and Estimation

We estimate our model using five time series for the U.S. economy: 1) the relative price of investment goods with respect to the price of consumption goods, 2) the federal funds rate, 3) real output per capita growth, 4) the consumer price index, and 5) real wages per capita. Our sample covers from 1959.Q1 to 2007.Q1, with 192 observations. Appendix B explains how we construct the series.

Once we have evaluated the likelihood as outlined before, we can either maximize it or combine it with a prior and rely on a Markov chain Monte Carlo algorithm to simulate from the posterior distribution. We follow the second route. However, we pick flat priors on a bounded support for all the parameters. The bounds are either natural economic restrictions (for instance, the Calvo and indexation parameters lie between 0 and 1) or are so wide that the likelihood assigns (numerically) zero probability to values outside them. Bounded flat priors induce a proper posterior, a convenient feature for our exercise.

We resort to flat prior for two reasons. First, to reduce the impact of presample information and show that our results arise mainly from the shape of the likelihood and not from the prior. Thus, the reader who wants to interpret our posterior modes as maximum likelihood point estimates can do so. Second, because as we learned in Fernández-Villaverde *et al.* (2010), eliciting priors for stochastic volatility is difficult since we deal with unfamiliar units, such as the variance of volatility shocks, about which we do not have clear beliefs.

Flat priors come, though, at a price: before proceeding to the estimation, we have to fix several parameters. We are dealing with a large model that suffers from weak identification along some dimensions. Our ability to learn from the data is sharpened if we avoid searching over these dimensions.

Table 5.1: Fixed Parameters

β	h	ψ	ϑ	δ	α	κ
0.99	0.9	8	1.17	0.025	0.21	9.5
ε	η	ϕ	Φ_2	ρ_{γ_Π}	ρ_{γ_y}	η_y
10	10	0	0.001	0.95	0	0

Table 5.1 lists the fixed parameters. Our guiding criterion in selecting them was to pick conventional values in the literature. The discount factor, $\beta = 0.99$, is a default choice, habit persistence $h = 0.9$, matches the observed sluggish response of consumption to shocks, the parameter controlling the level of labor supply, $\psi = 8$, captures the average amount of hours in the data, and the depreciation rate, $\delta = 0.025$, induces the appropriate capital-output ratio. The elasticities of substitution, $\varepsilon = \eta = 10$, deliver average mark-ups of around 10 percent, a common value in these models. We set the fixed cost of production, ϕ, to zero, since it is nearly irrelevant for the dynamics of the model and the cost of capital utilization, Φ_2, to a small number to introduce some curvature in this decision.

Three parameter values are borrowed from the point estimates from a similar model without stochastic volatility or parameter drifting presented in Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2009). The first is the inverse of the labor Frisch elasticity, $\vartheta = 1.17$. As argued by Rogerson and Wallenius (2009), this aggregate elasticity is compatible with microeconomic data once we allow for intensive and extensive margins on labor supply. The second is the coefficient of the intermediate goods production function, $\alpha = 0.21$. This value is lower than the common calibration of Cobb-Douglas production functions in real business cycle model because, in our environment, we have positive profits that also appear as capital income in the National Income and Product Accounts. The third value that we borrow is the adjustment cost, $\kappa = 9.5$, a number in line with other estimates from DSGE models (κ would be particularly hard to identify since investment is not one of our observables).

The autoregressive parameter of the evolution of the response to inflation, ρ_{γ_Π}, is set to 0.95. In preliminary estimations, we discovered that the likelihood pushed this parameter to 1. When this happened, and although the model was still in the determinacy region, the simulations became numerically unstable: after some shocks, the reaction of interest rates to inflation could be too tepid for too long. The last two parameters, ρ_{γ_y} and η_y, are equal to zero because, also in exploratory estimations, the likelihood favored values of $\eta_y \approx 0$. Thus, we decided to forget about them and make $\gamma_{y,t} = \gamma_y$.

To find the posterior, we proceed as follows. First, we define a grid of parameter values and check for the regions of high posterior density by evaluating the likelihood function in each point of the grid. This is a time-consuming procedure, but it ensures that we are searching in the right zone of the parameter space. Once we have identified the global maximum in the grid, we initialize a Random-Walk Metropolis-Hastings algorithm from this point. After an extensive fine-tuning of the algorithm, we draw 10,000 times from the chain.

6. Parameter Estimates

We examine now our parameter estimates. To ease the discussion, we group them in different tables, one for each set of parameters dealing with related aspects of the model. In all cases, we report the mode of the posterior and the standard deviation in parenthesis below (in the interest of space, we do not include the whole histograms of the posterior).

Table 6.1: Posterior, Parameters of Nominal Rigidities

θ_p	χ	θ_w	χ_w
0.8139 (0.0143)	0.6186 (0.024)	0.6869 (0.0432)	0.634 (0.0074)

Table 6.1 presents the results for the nominal rigidities parameters. Our estimates indicate an economy with substantial rigidities in prices, which are reoptimized once roughly every five quarters, and in wages, which are reoptimized approximately every three quarters. Moreover, since the standard deviations are small, there is enough information on the data about this result. At the same time, there is a fair degree of indexation, between 0.62-0.63, which brings a strong persistence of inflation. While it is tempting to compare our estimates with the evidence on the individual duration of prices as reviewed by Klenow and Malin (2009), in our model all prices and wages change every quarter. That is why, to a naive observer, our economy would look as one displaying tremendous price flexibility.

Table 6.2: Posterior, Parameters of the Stochastic Processes for Structural Shocks

ρ_d	ρ_φ	Λ_μ	Λ_A
0.1182 (0.0049)	0.9331 (0.0425)	0.0034 ($6.6e{-5}$)	0.0028 ($4.1e{-5}$)

Table 6.2 reports the findings for the parameters of the stochastic processes for the structural shocks. We estimate a low persistence of the intertemporal preference shock and a high persistence of the intratemporal one. The low estimate of ρ_d gets the quick variations in marginal utilities of consumption that match output growth and inflation fluctuations. The intratemporal shock is persistent to account for long-lived movements in hours worked. We estimate mean growth rates of technology of 0.0034 (neutral) and 0.0028 (investment-specific). Those numbers give us an average growth of the economy of 0.44 percent per quarter, or around 1.77 percent on an annual basis (0.46 and 1.86 percent on the data, respectively). Technology shocks, in our model, are deviations with respect to these drifts. Thus, we estimate that A_t falls in only 8 of the 192 quarters in our sample (which roughly corresponds to the percentage of quarters where measured TFP falls in the data), even if we estimate negative innovations to neutral technology in 103 quarters.

Table 6.3: Posterior, Parameters of the Stochastic Volatility Processes

$\log\sigma_d$	$\log\sigma_\varphi$	$\log\sigma_\mu$	$\log\sigma_A$	$\log\sigma_m$
-1.9834 (0.0726)	-2.4983 (0.0917)	-6.0283 (0.1278)	-3.9013 (0.0745)	-6.0004 (0.1471)
ρ_{σ_d}	ρ_{σ_φ}	ρ_{σ_μ}	ρ_{σ_a}	ρ_{σ_m}
0.9506 (0.0298)	0.1275 (0.0032)	0.7508 (0.035)	0.2411 (0.005)	0.855 (0.0231)
η_d	η_φ	η_μ	η_a	η_m
0.3246 (0.0083)	2.8549 (0.0669)	0.4716 (0.006)	0.7955 (0.013)	1.1034 (0.0185)

The estimates for the parameters of the stochastic volatility processes appear in table 6.3. In all cases, the ρ's and the η's are far away from zero: the likelihood strongly favors values where stochastic volatility plays an important role. The standard deviation of the innovations of the intertemporal preference shock and of the monetary policy shock are the most persistent, while the standard deviation of the innovation of the intratemporal preference shock is the least persistent. The standard deviation of the innovations of the volatility shock for the innovation of the intratemporal preference shock, $\eta_\varphi = 2.8549$, is large: the model asks for fast movements in marginal utilities of leisure to reproduce the hours data.

Table 6.4: Posterior, Policy Parameters

γ_R	$\log\gamma_y$	Π	$\log\gamma_\Pi$	η_π
0.7855 (0.0162)	-1.4034 (0.0498)	1.0005 (0.0043)	0.0441 (0.0005)	0.1479 (0.002)

In table 6.4., we have the estimates of the policy parameters. The autoregressive component of the federal funds rate is high, 0.7855, although somewhat smaller than in estimations without parameter drift. The value of γ_y (0.24 in levels) is similar to other results in the literature. The inflation along the balanced growth path is estimated to be 0.005 per quarter. This parameter requires some comment. Like for all the other endogenous variables, the law of motion of inflation is not centered around its value along the balanced growth path. Instead, it is moved away by the second order terms in the solution. For example, the constant in the second order associated with the squared value of the perturbation parameter is 0.0013. These effects allow the model to capture the level of inflation in the data. Also, they enormously complicate the introduction of a time-varying level of inflation along the balance growth path. The data speaks too softly about the difference between the evolution of this level and movements along the ergodic distribution of inflation.

Finally, the estimated value of γ_{Π} (1.045 in levels) guarantees local determinacy of the equilibrium. To see this, note that the relevant part of the solution of the model for local determinacy is the linear component. This component depends on γ_{Π}, the mean policy response, and not on the current value of $\gamma_{\Pi,t}$. While we cannot find an analytical expression for the determinacy region, numerical experiments show that, conditional on the other point estimates, values of γ_{Π} above 0.98 ensure uniqueness. Since the likelihood assigns probability zero to values of γ_{Π} lower than 1.01, well inside the determinacy region, multiplicity of local equilibria is not an issue here. The economic intuition is that local unicity survives even if $\gamma_{\Pi,t}$ temporarily violates the Taylor principle as long as there is reversion to the mean in the policy response and, thus, the agents have the expectation that $\gamma_{\Pi,t}$ will satisfy the Taylor principle on average (see, for a related result in models with Markov-switching regime changes, Davig and Leeper, 2006).

7. Impulse Response Functions

Before continuing the exploration of our results, we plot the impulse response functions (IRFs) generated by the model to a monetary policy shock. This exercise is a powerful reality-check. If the IRFs match the shapes and sizes of those gathered by time series methods such as SVARs, it will strengthen our belief in the rest of our results. Otherwise, we should at least understand where the differences come from.

Fortunately, the answer is positive: our model generates dynamic responses that are close to the ones from SVARs (see, for instance, Sims and Zha, 2006). Figure 7.1 plots the IRFs to three variables commonly discussed in monetary models: the federal funds rate, output growth, and inflation. Since we have a non-linear model, in all the figures in this section, we report the generalized IRFs starting from the mean of the ergodic distribution (Koop *et al.*, 1996). After a one standard-deviation shock to the federal funds rate, inflation goes down in a hump-shaped pattern for many quarters and output growth drops.

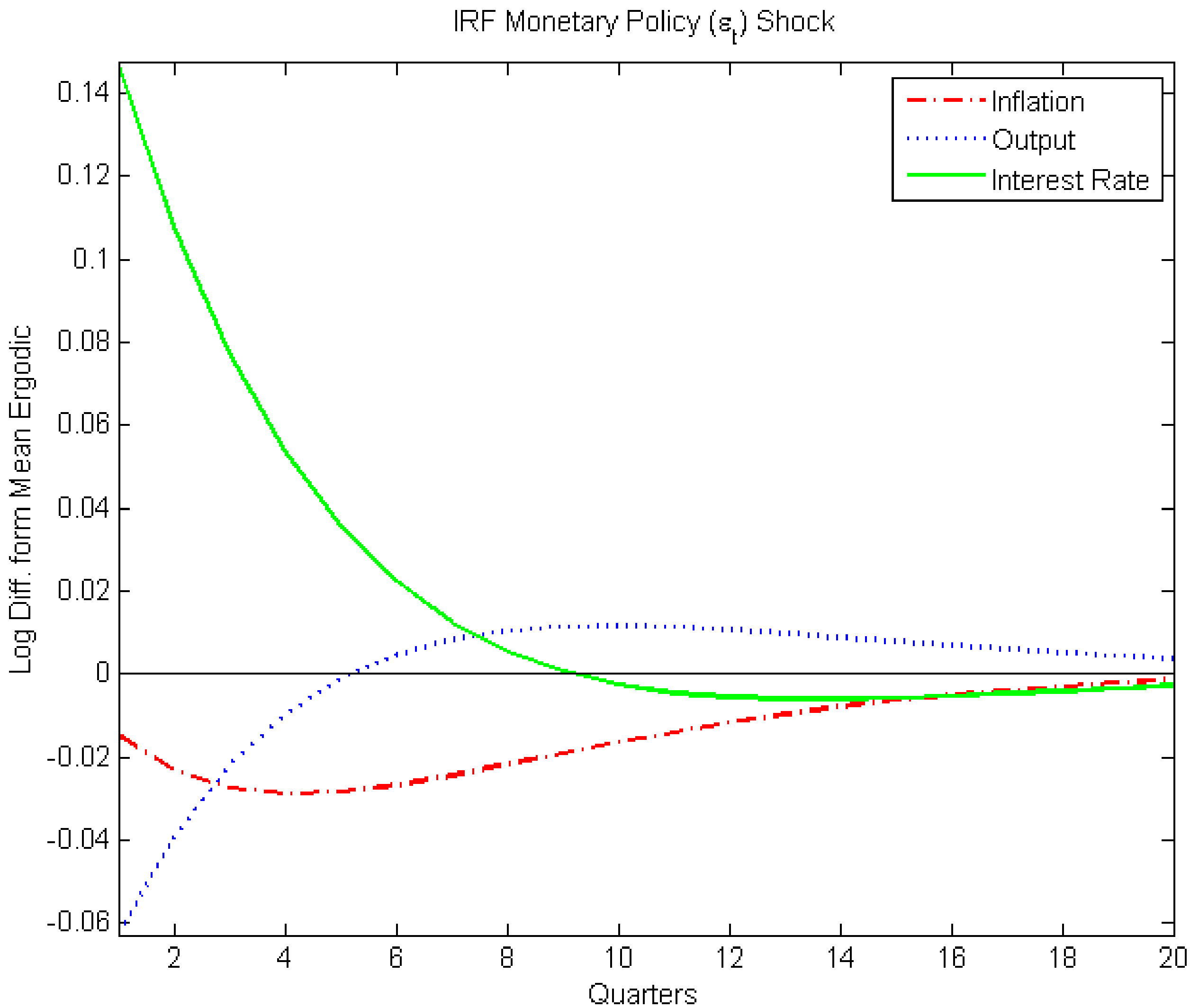


Figure 7.1: IRFs of inflation, output growth, and interest rates to a monetary policy (ε_t) shock. The responses are measured as log differences with respect to the mean of the ergodic distribution.

Figure 7.2 plots the IRFs after a one standard-deviation innovation to the monetary policy shock computed conditional on fixing $\gamma_{\Pi,t}$ to the estimated mean during the tenure of three different Chairmen of the Board: the combination Burns-Miller, Volcker, and Greenspan. This exercise tells us how the variation on the systematic component of monetary policy has affected the dynamics of aggregate variables. Furthermore, it allows a comparison with numerous similar exercises done in the literature with SVARs where the IRFs are estimated on different subsamples.

The most interesting difference is that the response of output under Volcker was the mildest: the energetic stand of monetary policy under his tenure stabilizes the economy. Inflation responds more moderately as well since the agents have the expectation that future shocks will be smoothed out more vigorously by the monetary authority. This also explains why the IRFs of the interest rate are nearly on top of each other for all three periods: while Volcker is more aggressive for any given level of inflation than Burns-Miller or Greenspan, his though policy lowers inflation deviations and hence moderates his actual response along the equilibrium path of the economy.

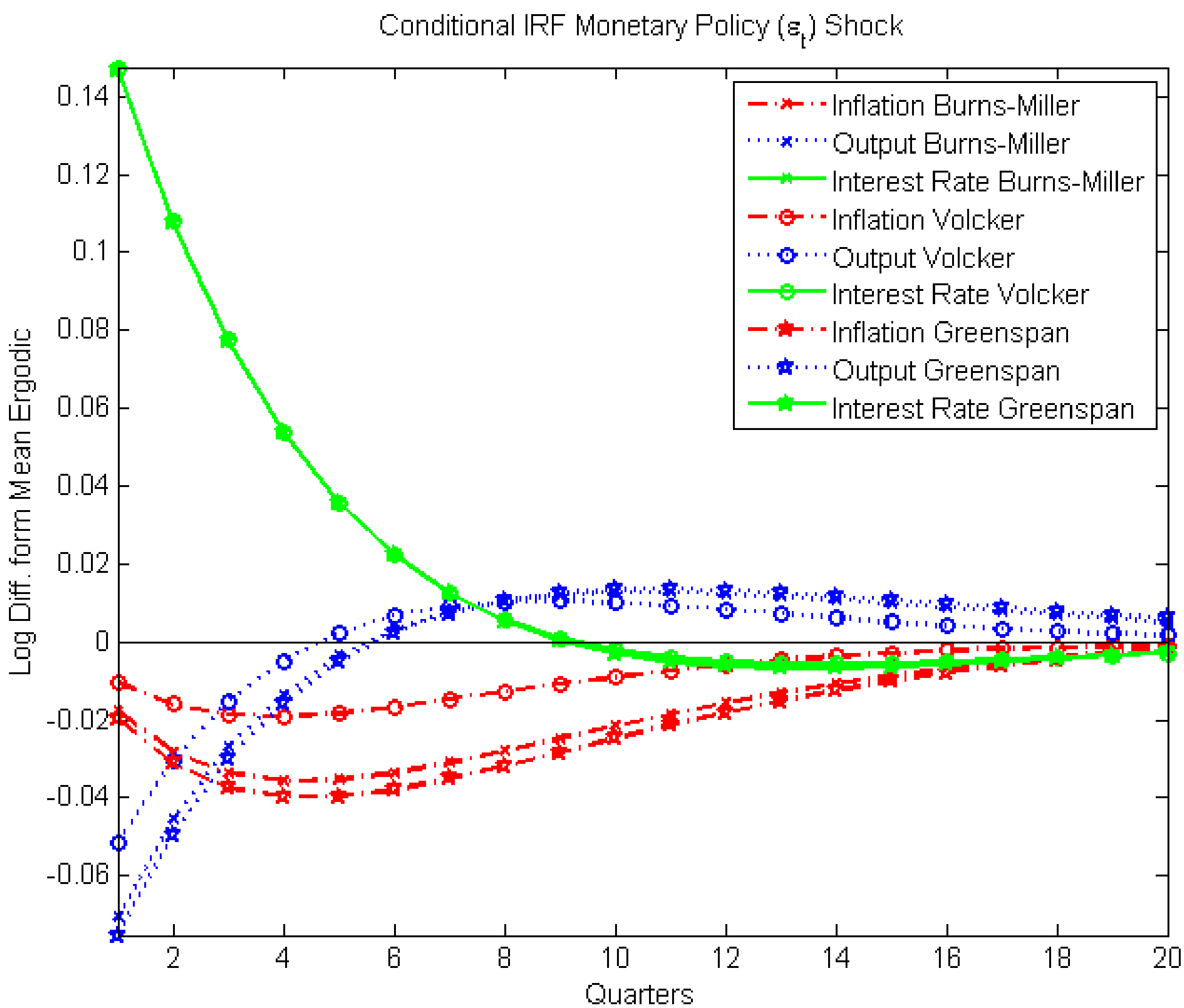


Figure 7.2: IRFs of inflation, output growth, and interest rates to a monetary policy (ε_t) shock conditional on the mean of the Taylor Rule parameters of each of the three Chairmen. The responses are measured as log differences with respect to the mean of the ergodic distribution.

For completeness, we also plot, in figure 7.3, the IRFs to each of the other four shocks in our model: the two preferences shocks (intertemporal and intratemporal) and the two technology shocks (investment-specific and neutral). The behavior of the model is standard. The intertemporal preference shock raises output growth and inflation because there is an increase in the desire for consumption in the current period. The intratemporal shock lowers output because labor becomes less attractive, driving up the marginal costs prices. The two supply shocks raise output growth and lower inflation by increasing productivity.

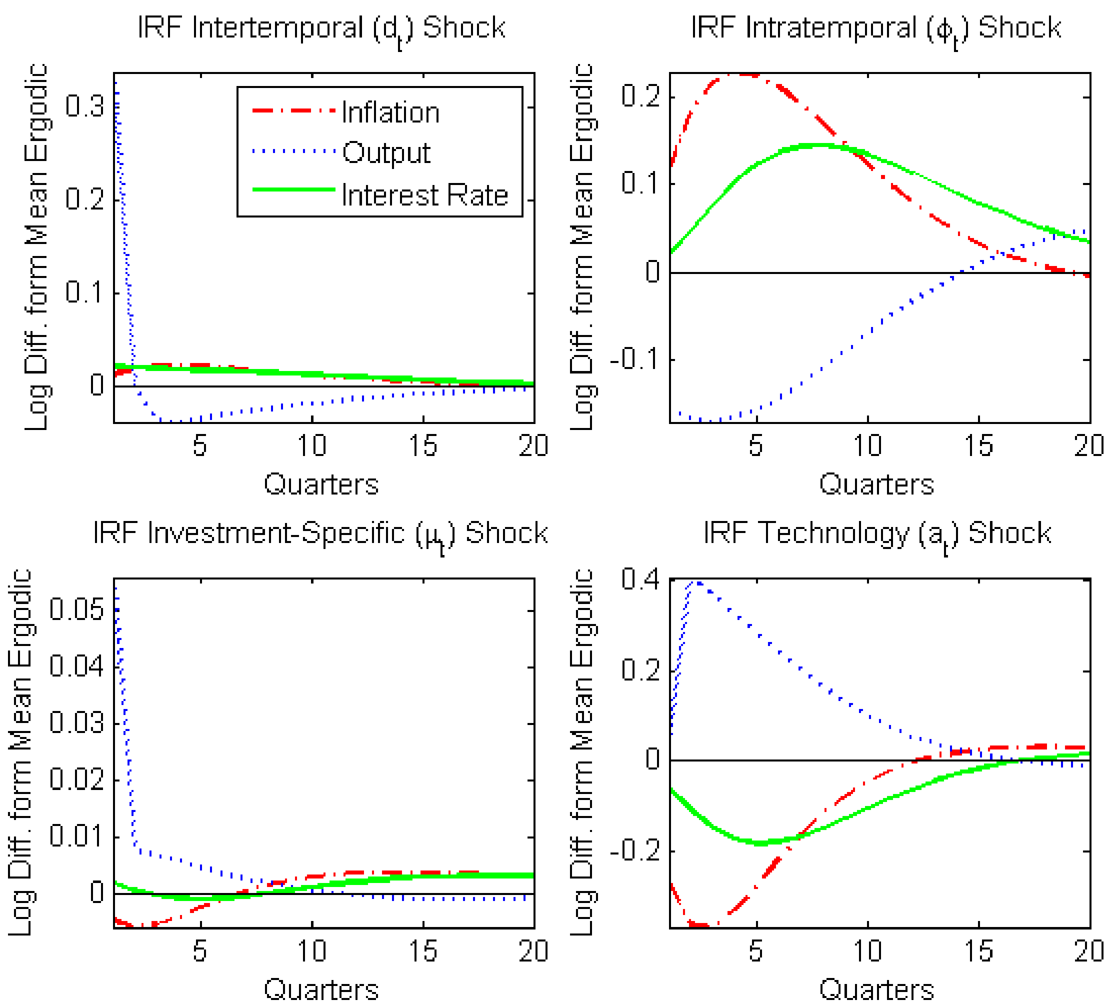


Figure 7.3: IRFs of inflation, output growth, and interest rates to an intertemporal demand (d_t) shock, an intratemporal demand (ϕ_t) shock, an investment-specific (μ_t) shock, and a neutral technology (a_t) shock. The responses are measured as log differences with respect to the mean of the ergodic distribution.

8. Model Fit

After the point estimates and the IRFs, our next step is to examine the fit of the model and how it compares with alternative specifications. In this way, we document one of the main findings of the paper: the data strongly supports the view that monetary policy has changed over time even after including stochastic volatility.

Given our Bayesian framework, a natural approach for model comparison is the computation of log marginal data densities (log MDD) and log Bayes factors. The log MDD of the version i of a model is defined as

$$\log p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; i\right) = \log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \gamma, i\right) p\left(\gamma; i\right) d\gamma, \tag{17}$$

where $p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \gamma, i\right)$ is the likelihood and $p(\gamma; i)$ is the prior for the parameters of the version i. The log Bayes factor between specifications i and j, a measure of the evidence in the data for one specification above the other, is

$$\log B_{i,j} = \log p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; i\right) - \log p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; j\right).$$

The Bayes factor is attractive because it automatically penalizes specifications with unneeded free parameters.

We compare the full model with stochastic volatility and parameter drifting (*drift*) with a version without parameter drifting (*no drift*). In this second case, we have two parameters less, ρ_{γ_Π} and η_π (but we still have the mean response $\log \gamma_\Pi$ of monetary policy to inflation deviations). To ease notation, we partition the parameter vector γ as $\gamma = \left(\widetilde{\gamma}, \rho_{\gamma_\Pi}, \eta_\pi\right)$ where $\widetilde{\gamma}$ is the vector of all the other parameters, common to the two versions of the model.

Given that 1) our priors are uniform, 2) independent of each other, 3) cover all the area where the likelihood is (numerically) positive, and 4) the priors on $\widetilde{\gamma}$ are common across the two specifications of the model, we can write

$$\begin{aligned}
& \log p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; drift\right) \\
= \ & \log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \gamma, drift\right) p\left(\widetilde{\gamma}\right) p\left(\rho_{\gamma_\Pi}\right) p\left(\eta_\pi\right) d\gamma = \\
= \ & \log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \gamma, drift\right) d\gamma + \log p\left(\widetilde{\gamma}\right) + \log p\left(\rho_{\gamma_\Pi}\right) + \log p\left(\eta_\pi\right),
\end{aligned}$$

where $\log p(\widetilde{\gamma})$, $\log p\left(\rho_{\gamma_\Pi}\right)$, and $\log p(\eta_\pi)$ are constants and

$$
\begin{aligned}
& \log p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \textit{no drift}\right) \\
= \; & \log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \widetilde{\gamma}, \textit{no drift}\right) p(\widetilde{\gamma})\, d\widetilde{\gamma} \\
= \; & \log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \widetilde{\gamma}, \textit{no drift}\right) d\widetilde{\gamma} + \log p(\widetilde{\gamma}).
\end{aligned}
$$

Thus

$$
\log B_{\textit{drift, no drift}} = \log p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \textit{drift}\right) - \log p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \textit{no drift}\right)
$$
$$
= \log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \gamma, \textit{drift}\right) d\widetilde{\gamma} - \log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \widetilde{\gamma}, \textit{no drift}\right) d\widetilde{\gamma} - \log p\left(\rho_{\gamma_\Pi}\right) - \log p(\eta_\pi).
$$

The first two terms, $\log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \gamma, \textit{drift}\right) d\gamma - \log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \gamma, \textit{no drift}\right) d\widetilde{\gamma}$, tells us how much better the version with parameter drift fits the data in comparison with the version with no drift. The last two terms, $\log p\left(\rho_{\gamma_\Pi}\right) + \log p(\eta_\pi)$, penalize for the presence of two extra parameters in the version with parameter drift.

We estimate the log MDDs following Geweke's (1998) harmonic mean method. This requires us to generate a new draw of the model for the specification with no parameter drift to compute $\log \int p\left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \widetilde{\gamma}, \textit{no drift}\right) d\widetilde{\gamma}$. After doing so, we find that

$$
\log B_{\textit{drift, no drift}} = 126.1331 + \log p\left(\rho_{\gamma_\Pi}\right) + \log p(\eta_\pi).
$$

This expression shows a potential problem of Bayes factors: by picking uniform priors for ρ_{γ_Π} and η_π spread out over a sufficiently large interval, we could overcome any difference in fit. But the prior for ρ_{γ_Π} is pinned down by our desire to keep that process stationary, which imposes natural bounds in $[-1, 1]$ and makes $\log p\left(\rho_{\gamma_\Pi}\right) = -0.6931$. Thus, there is only one degree of freedom left: our choice of $\log p(\eta_\pi)$.

Any sensible prior for η_π will only put mass in a relatively small interval: the point estimate is 0.1479, the standard deviation is 0.002, and the likelihood is numerically zero for values bigger than 0.2. Hence, we can safely impose that $\log p(\eta_\pi) > -1$ ($\log p(\eta_\pi) = -1$ would imply a uniform prior between 0 and 2.7183, a considerably wider support than any evidence in the data) and conclude that $\log B_{\textit{drift, no drift}} > 124.4400$. This is conventionally considered overwhelming evidence in favor of the model with parameter drift (Jeffreys, 1961, for instance, suggests that differences bigger than 5 are decisive).[8] Thus, even after controlling

[8] An alternative way to see this is that, to overcome the evidence in the data as recorded by the difference in loglikelihoods, $\log p(\eta_\pi)$ should be defined between 0 and $3.0054e+054$, clearly an absurd proposition.

for stochastic volatility, the data strongly prefers an specification of the model where monetary policy has changed over time even.

It has been noted, though, that the estimation of log MDDs is fraught with peril because of numerical instabilities in the evaluation of the integral (17) (see, among many, the discussion in Sims and Zha, 2006). This concern is particularly relevant in our case since we have a large model saddled with a burdensome computation. As a robustness analysis, we also computed the Bayesian Information Criterion (BIC) (Schwarz, 1978). The BIC, which avoids the need to handle the integral in (17), can be understood as an asymptotic approximation of the Bayes factor that also automatically penalizes for unneeded free parameters. The BIC of model i is defined:

$$BIC_i = -2 \ln p \left(\mathbb{Y}^T = \mathbb{Y}^{data,T}; \widehat{\gamma}, i \right) + k_i \ln n$$

where $\widehat{\gamma}$ is the maximum likelihood estimator (or, in our case given our flat priors, the mode of the posterior), k_i is the number of parameters, and n is the number of observations. Then, the BIC of the model with stochastic volatility and parameter drifting is:

$$BIC_{drift} = -2 * 3885 + 28 * \ln 192 = -7,622.8$$

If we eliminate parameter drifting and the parameters ρ_{γ_Π} and η_π associated with it (and, of course, with a new point estimate of the other parameters):

$$BIC_{no\ drift} = -2 * 3810.7 + 26 * \ln 192 = -7,484.7$$

The difference is, therefore, of over 138 log points, which is again overwhelming evidence in favor of the model with parameter drifting.

The comparison with the case without stochastic volatility is more difficult, since we are taking advantage of its presence to evaluate the likelihood. Fortunately, Justiniano and Primiceri (2007) and Fernández-Villaverde and Rubio-Ramírez (2007) estimate models similar to ours with and without stochastic volatility (in the first case, using only a first-order approximation to the decision rules of the agents and in the second with measurement errors). Both papers find that the fit of the model improves substantially when we include stochastic volatility. Finally, Fernández-Villaverde and Rubio-Ramírez (2008) compare a model with parameter drifting and no stochastic volatility with a model without parameter drifting and no stochastic volatility and report that parameter drifting is also strongly preferred by the likelihood.

9. Smoothed Shocks

We present now the smoothed estimates of the shocks, volatilities, and drifting parameters of the model. Figure 9.1 reports the level of the five shocks that drive the economy. To ease reading of the results, we color different vertical bars to represent each of the Chairmen periods at the Fed: the McChesney Martin's years from the start of our sample in 1959 to the appointment of Burns in February 1970 (white), the Burns-Miller era (light blue), the Volcker interlude from August 1979 to August 1987 (grey), the Greenspan times (orange), and Bernanke's tenure from February 2006 (yellow).

We see in the top left panel of figure 9.1 that the intertemporal shock d_t is particularly high in the 1970s. This increases the households' desire for current consumption. A higher aggregate demand triggers, in the model, the higher inflation observed in the data for those years and generates challenges for the monetary authority that we will discuss below. The shock has a dramatic drop in the second quarter of 1980. This is precisely the quarter where, in its increasingly desperate fight against inflation, the Carter administration invoked the Credit Control Act (started on March 14, 1980), which caused a large turmoil in financial markets and, most likely, distorted intertemporal choices of households (see the historical description in Shreft, 1990) reflected in a large innovation to d_t. The low level of d_t in the 1990s eases the inflationary pressures in the economy. This is the first piece of evidence that we present in favor of our interpretation that monetary policy faced a much tougher environment in the 1970s and early 1980s than in the 1990s.

The shock to the utility of leisure grows in the 1970s and falls in the 1980s to stabilize in the 1990s. The likelihood wants to track, in this way, the path of average hours worked: low in the 1970s, increasing in the 1980s, and stabilizing in the 1990s. Higher hours also lower the marginal cost of firms (wages fall relatively to the technological level). The reduction in marginal costs helped to reduce inflation during Greenspan's tenure.

The evolution of the investment-specific technology μ shows a clear drop after 1973 (when it is likely that energy-intensive capital goods suffered the consequences of the oil shocks in the form of economic obsolescence) and very positive realizations in the late 1990s. Our model interprets the sustained boom of those years as the consequence of strong improvements in investment technology. Later, we will come back to this point. In comparison, the neutral-technology shocks are stable since 1959 with only a few big shocks at the end of the sample.

Finally, the evolution of the monetary policy shock (panel (3,1)) reveals considerable fluctuations during Volcker's years (and again with large innovations in 1980). This is due both to the fast change in policy brought about by Volcker and to the fact that a Taylor rule might not fully capture the dynamics of monetary policy during a period where money growth targeting was attempted. Sims and Zha (2006) also find that the Volcker period appears as one with large disturbances of the policy rule and argue that the Taylor rule formalism can be a misleading perspective to understand policy during that time.

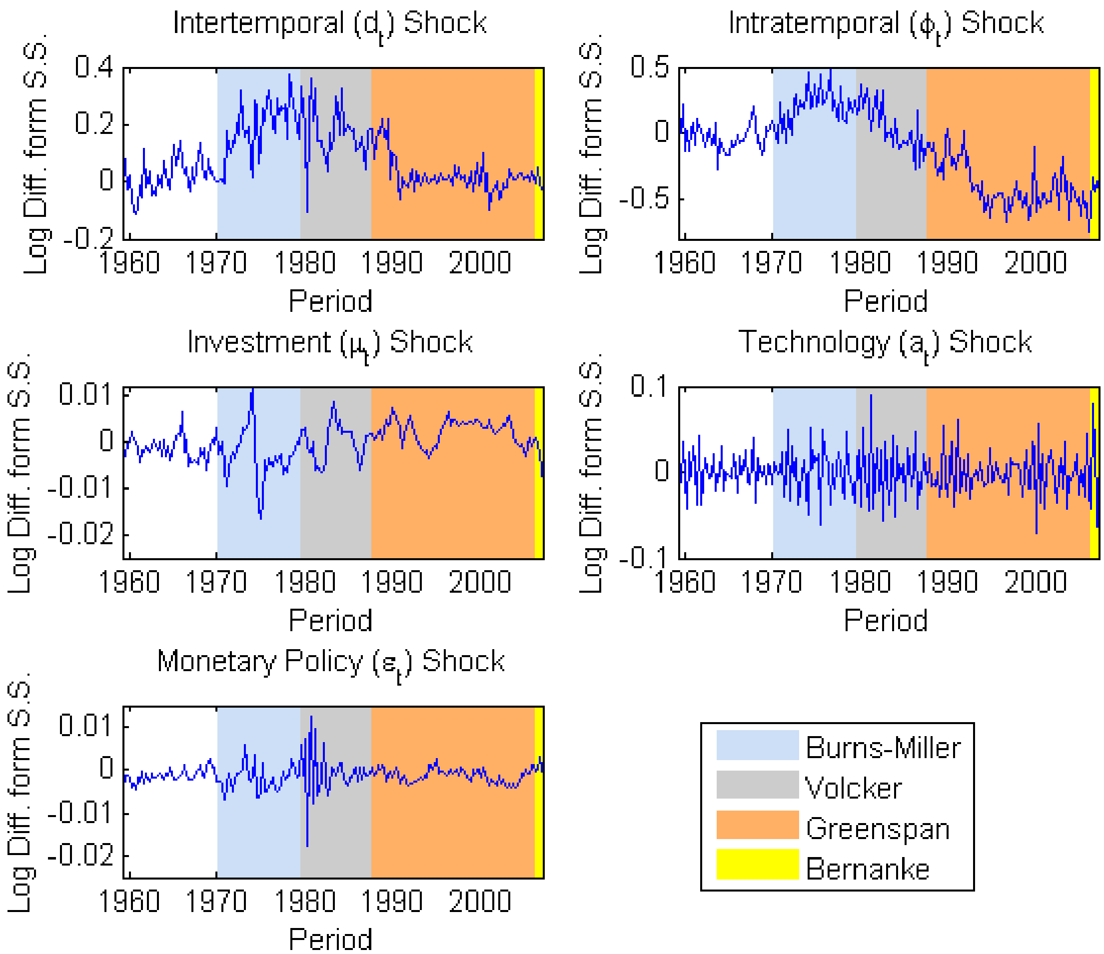


Figure 9.1: Smoothed intertemporal demand (d_t) shocks, intratemporal demand (ϕ_t) shocks, investment-specific (μ_t) shocks, technology (a_t) shocks, and monetary policy (ε_t) shocks.

As a way to gauge the level of uncertainty of our smoothed estimates, we plot in figure 6.2 the shock plus/minus two standard deviations. The lesson to get from this figure is that, in all the cases, the estimates are tight.

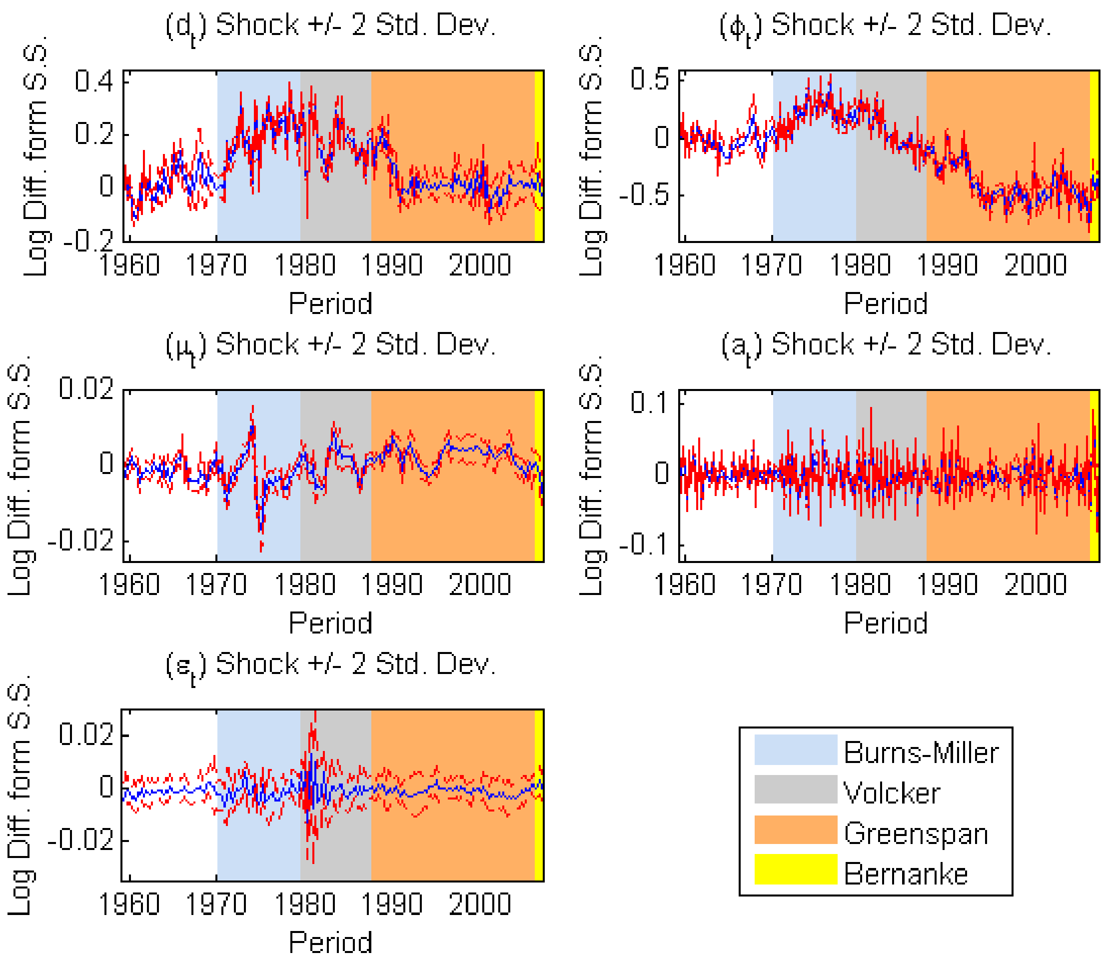


Figure 8.2: Smoothed intertemporal demand (d_t) shocks, intratemporal demand (ϕ_t) shocks, investment-specific (μ_t) shocks, technology (a_t) shocks, and monetary policy (ε_t) shocks.s +/- 2 Standard Deviations.

We move now, in figure 6.3, to plot the evolution of the five standard deviation levels and of the log of the response of monetary policy to inflation, all of them in log-deviations with respect to their estimated means to facilitate the interpretation of their movements. We see in this figure that the standard deviation of the intertemporal shock was particularly high in the 1970s and only went down slowly during the 1980s and early 1990s. After a through in the mid 1990s, the standard deviation recover somewhat to end up roughly at the level where it started at the beginning of the sample. In comparison the standard deviation of all the other shocks is relatively stable except, perhaps, for the big drop in the standard deviation of the monetary policy shock in the early 1980s. We delay the discussion of the evidence in panel (3,2) regarding monetary policy until a few paragraphs below.

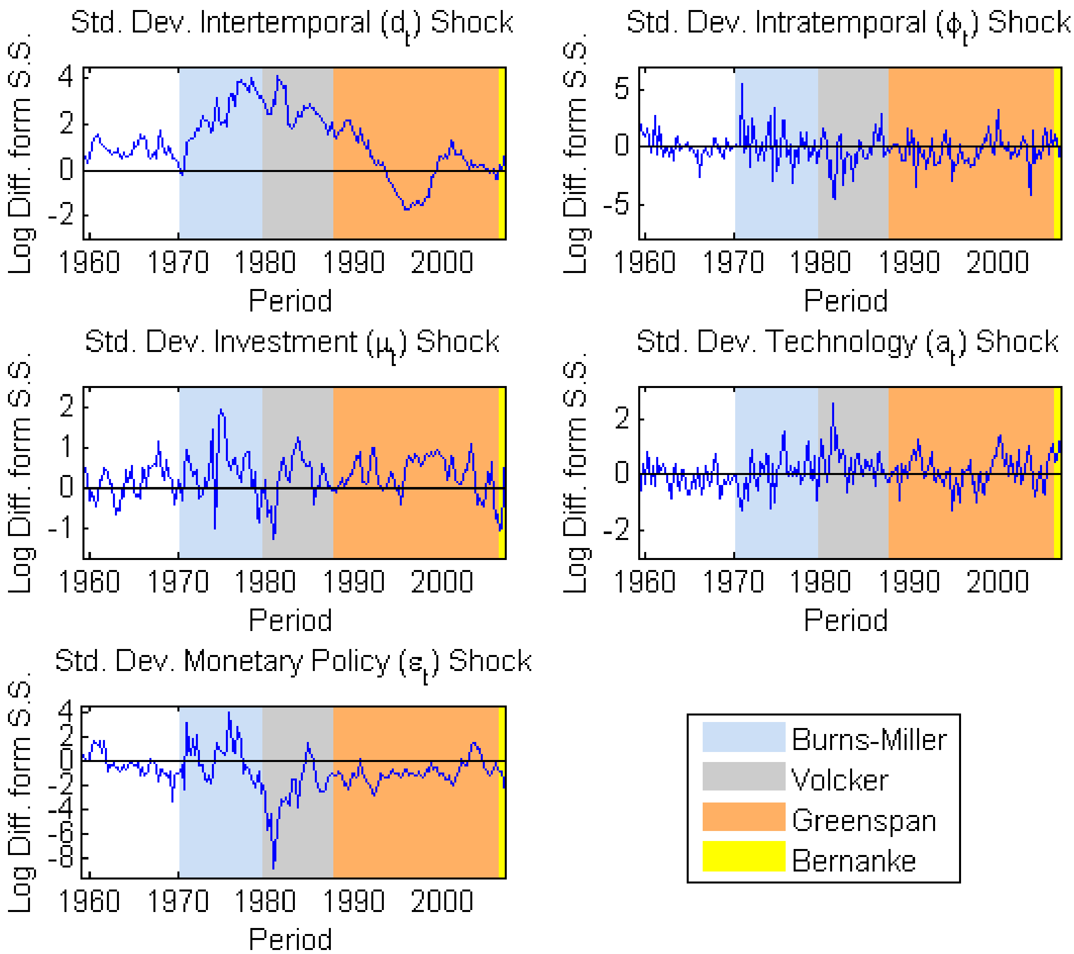


Figure 8.3: Smoothed standard deviation shocks to the intertemporal demand (d_t) shock, the intratemporal demand (ϕ_t) shock, the investment-specific (μ_t) shock, the technology (a_t) shock, and the monetary policy (ε_t) shock.

In figure 6.4, we plot the same results except that now we add two standard deviations to assess posterior uncertainty. The main lesson from this figure is that the big movements in the different series that we reported below can be ascertained with a reasonable degree of confidence.

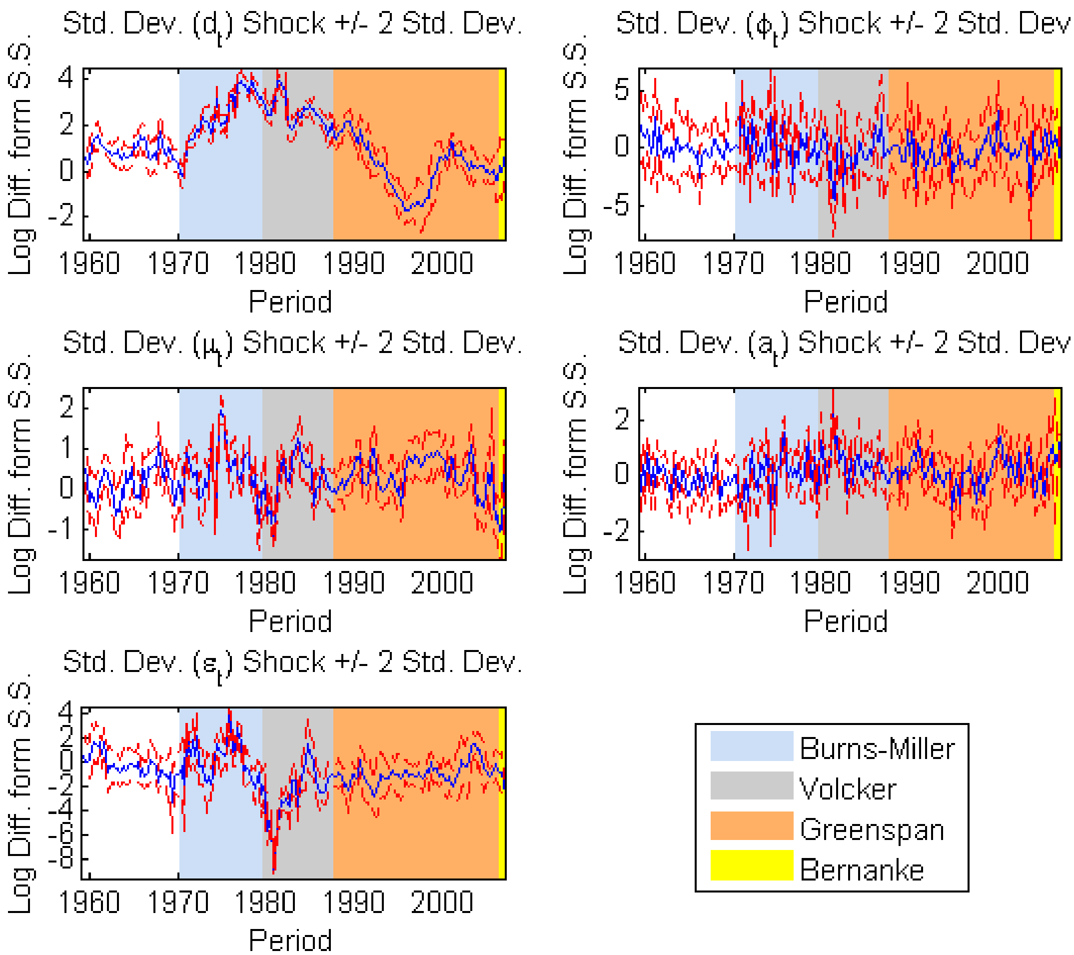


Figure 8.4: Smoothed standard deviation shocks to the intertemporal demand (d_t) shock, the intratemporal demand (ϕ_t) shock, the investment-specific (μ_t) shock, the technology (a_t) shock, and the monetary policy (ε_t) shock +/- 2 standard deviations.

Finally, in figure 6.5, we plot the evolution of the response of monetary that we already reported in panel (3,2) but in levels, a more natural units for economists, plus a two standard deviation interval. Figure 6.5 tells us an intriguing narrative. The parameter $\gamma_{\Pi,t}$ started the sample around its estimated mean, slightly over 1 and, then it grew more or less steadily during the 1960s until reaching a peak in early 1968. After that year, $\gamma_{\Pi,t}$ suffered a fast collapse that took it below 1971. To put this evolution in perspective, it is useful to remember that Burns was appointed Chairman at the Fed in February of 1970. The parameter stayed below 1 for all the decade, showing either that the monetary policy was not sufficiently active or that the Taylor rule is not a good description of the behavior of the Fed at the time. The arrival of Volcker is quickly picked by our smoothed estimates: $\gamma_{\Pi,t}$ increases to over 2 in just a few months and stays high during all the years of Volcker's tenure.

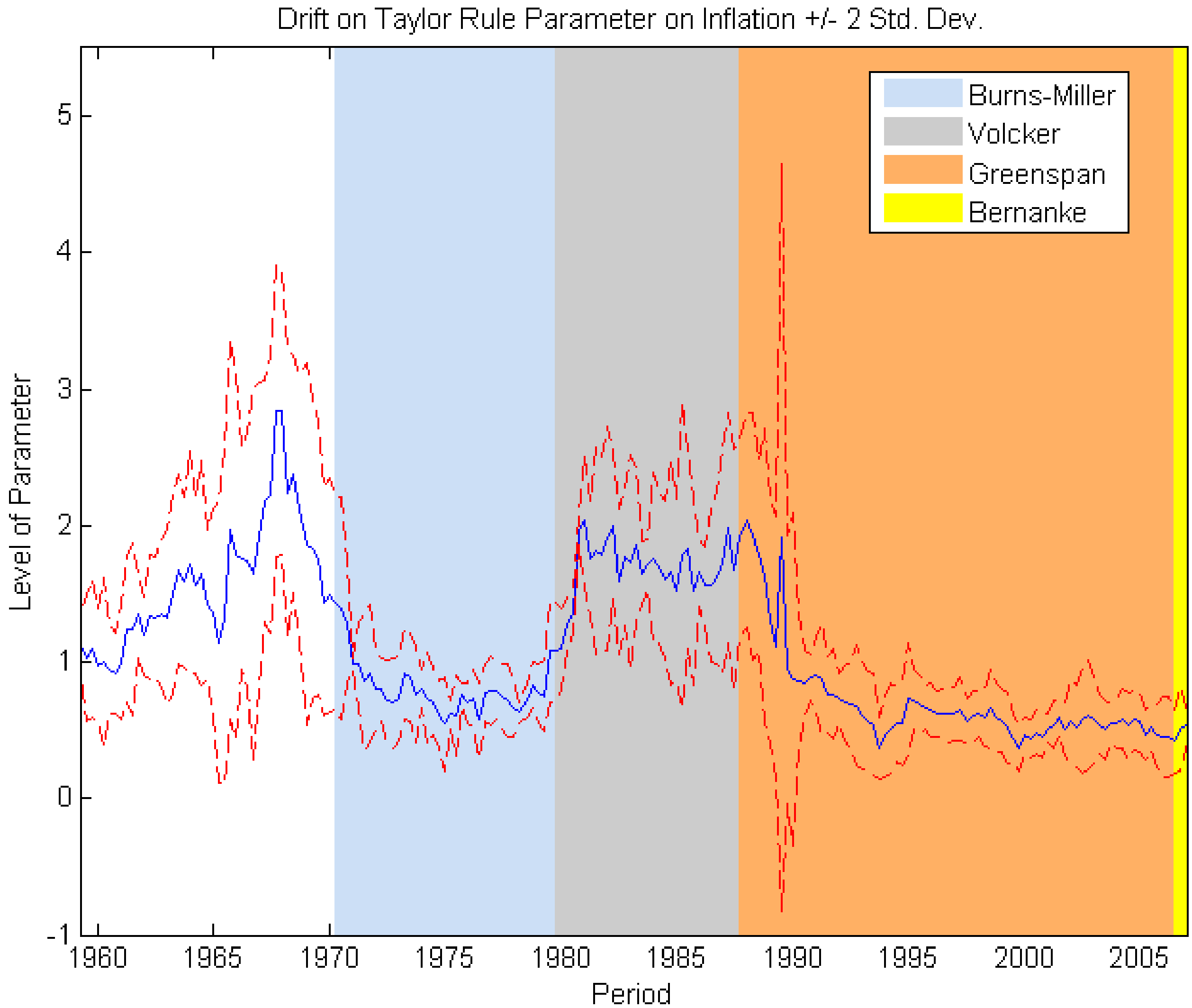


Figure 8.5: Smoothed path for the Taylor rule parameter on inflation +/- 2 standard deviations.

But, as quickly as $\gamma_{\Pi,t}$ rose when Volcker arrived, it went down again when he departed. Greenspan's tenure at the Fed meant that, by 1990, the response of the monetary authority to inflation was again below 1. During all the following years, $\gamma_{\Pi,t}$ was low, probably even below the values that it took during the Burns-Miller times. Moreover, our estimates of $\gamma_{\Pi,t}$ are relatively tight, suggesting that posterior uncertainty may not be the full explanation behind these movements. How can we account then for the good economic performance of the economy under Greenspan? While our smoothed shocks already hint at the reason, to satisfactorily answer this question, we move in the next section to build counterfactual histories.

10. Historical Counterfactuals

One important goal of our research is to quantify how much of the observed changes in the volatility of aggregate variables can be accounted for by changes in the standard deviations of shocks and how much by changes in policy. To accomplish this, we build a number of historical counterfactuals. Those are internally coherent exercises where we remove one source of variation at a time and we measure how the other variables behave with the remaining shocks. As long as our model is structural in the sense of Hurwicz (1962) (its structure is invariant to interventions, including shocks by nature such as the ones we are postulating), the model will provide an answer that is robust to Lucas' critique.

In this section, we will always plot the same three basic variables that we used in section 7: inflation, output, and the federal funds rate. Also, we will have vertical bars for the tenure of each Chairman of the fed, following the same coloring scheme as in section 9.

In our first historical counterfactual, we compute how the economy would have behaved in the absence of changes in volatility of the shocks, that is, if the volatility of the different shocks had been fixed at its historical mean. To do so, we back up the smoothed structural shocks as we did in section 9 and we feed them to the model, given our parameter point estimates and the historical mean of volatility, to generate series for inflation, output, and the federal funds rate. In figure 10.1, we compare this counterfactual history (blue line) with the observed one (red line).

Figure 10.1 tells us that volatility shocks mattered all across the sample. The run up for inflation would have been much slower in the late 1960s (with inflation actually being negative during the last years of Martin) without big differences in output growth or the federal funds rate (except at the very end of the sample). Prices would have not picked nearly as much during the first oil shock but output growth would have suffered. During Volcker's times, inflation would have fallen much faster with little output growth cost. Finally, in the 1990s, inflation would have been much more volatile, with a big increase in the middle of the decade. Similarly, during those years, output growth would have been much lower and the federal funds rate would have been much higher. This is another manifestation of how favorable the 1990s were for the policy makers.

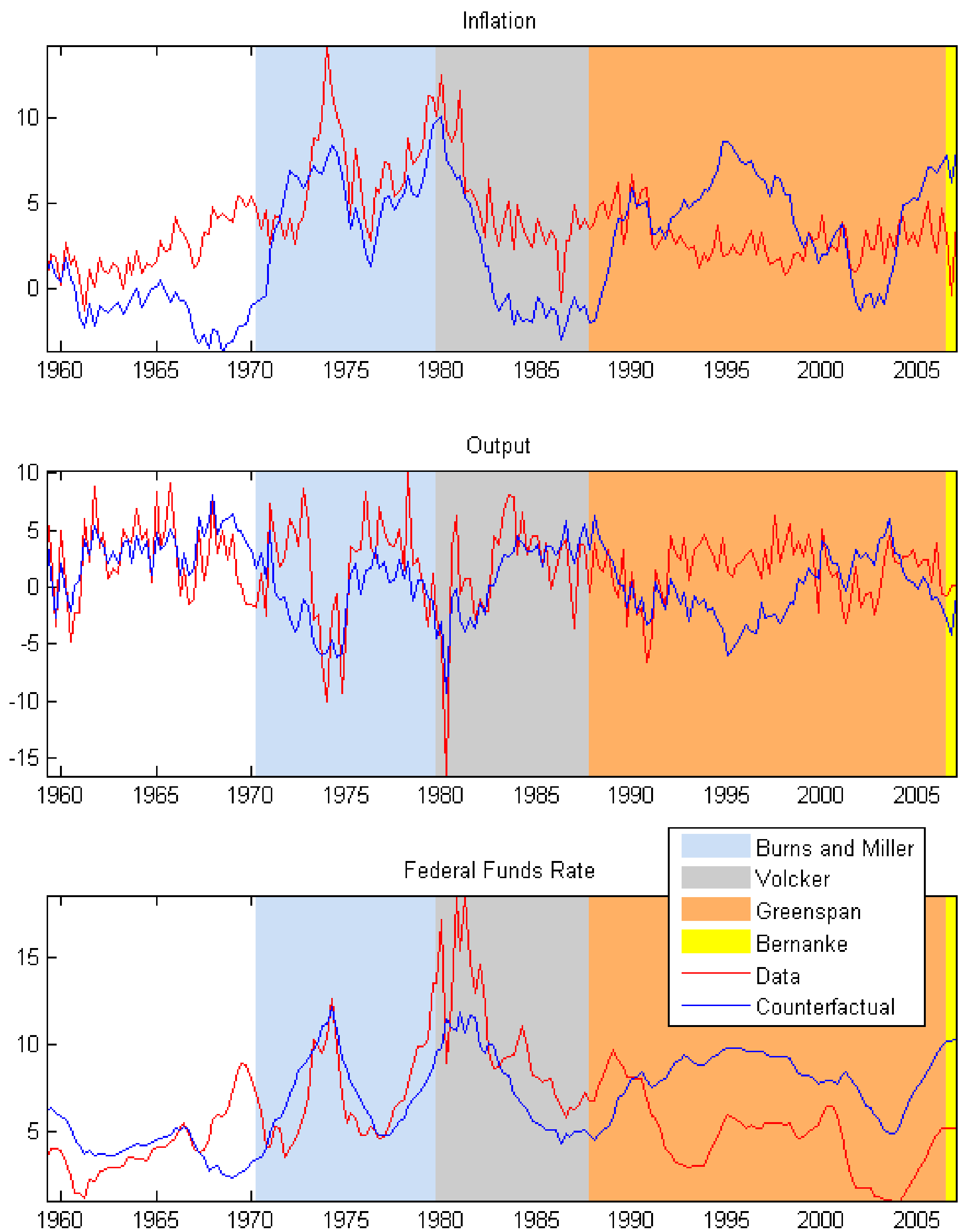


Figure 10.1: Counterfactual history with no changes in volatility.

In our second counterfactual we move one Chairman from his mandate to an alternative time period. For example, we appoint Greenspan as Chairman during the Burns-Miller years. By that, we mean that the monetary authority would have followed the policy rule dictated by the average $\gamma_{\Pi,t}$ during Greenspan's times while starting from the same states than Burns-Miller and suffering the same shocks (both structural and of volatility). We repeat this exercise with all possible combinations: Greenspan in Volcker's times, Burns-Miller in Greenspan's times, Burns-Miller in Volcker's times, Volcker in Greenspan's times, and finally Volcker in Burns-Millers times.

In table 10.1 we report the mean inflation, output growth, and federal funds rates in the observed data and in the counterfactual ones. Contrary to many previous assessments, Greenspan does not look as a particularly tough chairman. In Burns-Miller times, he would have deliver slightly higher average inflation, 6.83 versus the observed 6.23 and lower output growth, 1.89 versus observed 2.03. The difference is even bigger in Volcker's times, where average inflation would have been nearly 1.4 percent higher while output growth would have been virtually identical (1.34 versus 1.38). The key difference comes from the behavior of the federal funds rate, which would have raised only by 9 basis points on average if Greenspan would have driven the Fed instead of Volcker. Given the much higher inflation, this basically would had implied a much lower average level of real interest rates.

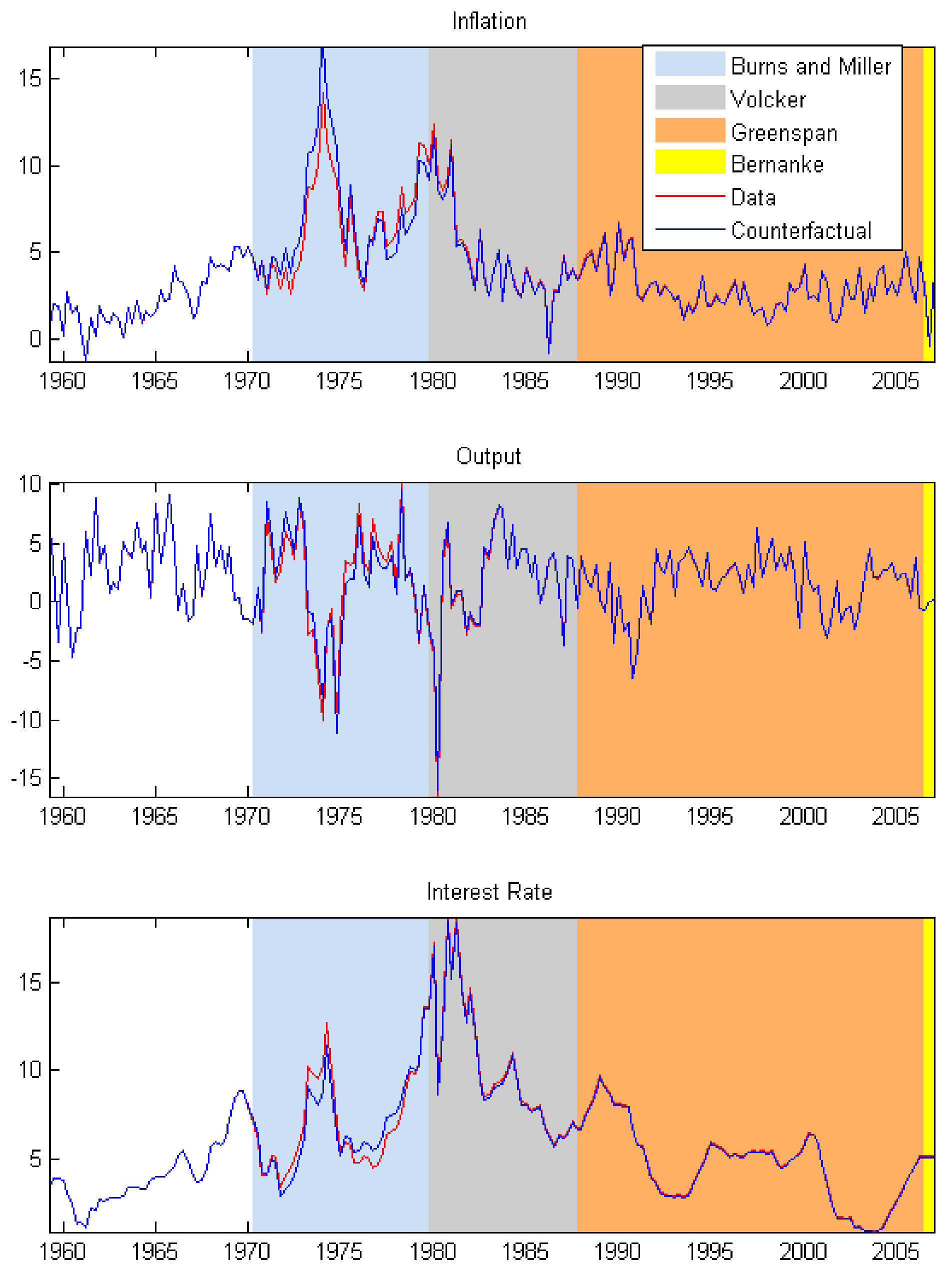
Table 10.1: Switching Chairmen, Data versus Counterfactural Histories

	Inflation		Output Growth		Federal Funds Rate	
	Data	Counter.	Data	Counter	Data	Counter
Greenspan to BM	6.2333	6.8269	2.0322	1.8881	6.5764	6.5046
Greenspan to Volcker	5.3584	6.7284	1.3846	1.3423	10.3338	10.4235
BM to Greenspan	2.9583	2.3355	1.5177	1.5277	4.7352	4.4529
BM to Volcker	5.3584	6.4132	1.3846	1.3560	10.3338	10.4126
Volcker to Greenspan	2.9583	-0.4947	1.5177	1.3751	4.7352	3.6560
Volcker to BM	6.2333	4.3604	2.0322	1.5010	6.5764	7.6479

The counterfactual of Burns-Miller in Greenspan and Volcker's times cast doubts on the usually malignant reputations of these two short-lived chairmen. Burns-Miller would have delivered even lower inflation than Greenspan and a bit higher output growth thanks to a slightly higher real federal funds rate. What is clear, also, is that Burns-Miller would have not matched the behavior of Volcker.

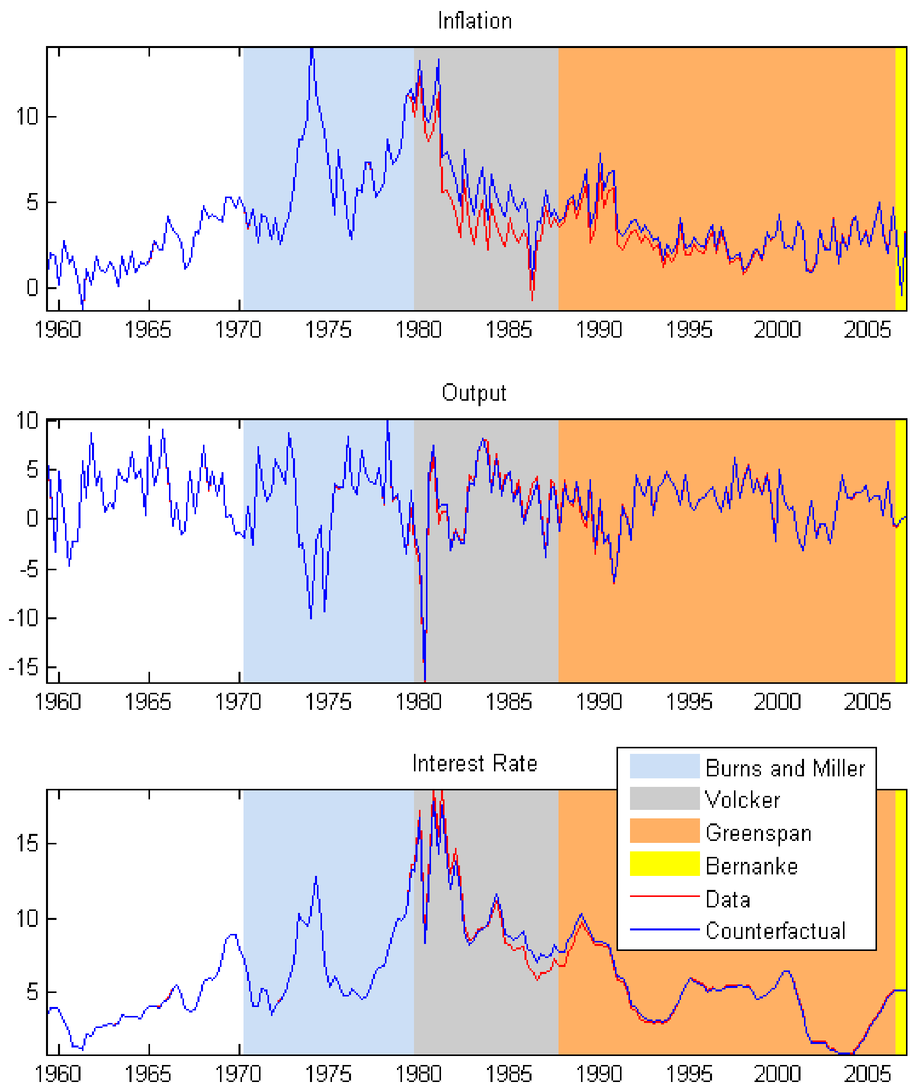
Our next exercise is to plot the whole counterfactual histories summarized in table 10.1. We find interesting to plot the whole history because changes in behavior of the economy in one period will propagate over time and we want to undersand, for example, how a Greenspan's legacy would have molded Volcker's tenure.

Our first plot is figure 10.2, where we put Greenspan in Burns-Miller times. As we can see clearly, from figure 10.2, Greenspan would have not been much of a difference in , somewhat lower inflation and a bit higher interest rates.



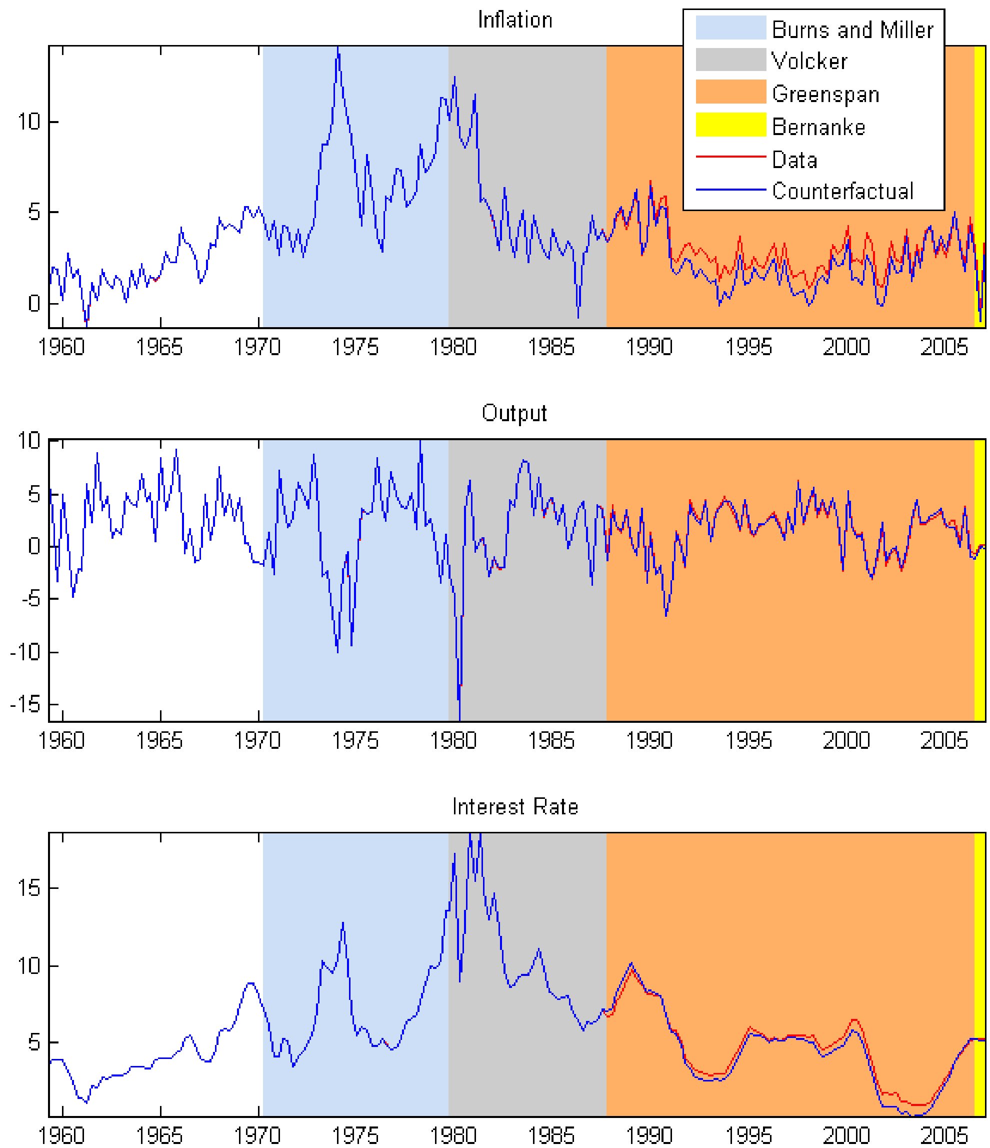
Chairman Greenspan during Burns-Miller years

In figure 9.4 we repeat the same exercise for Greenspan during Volcker's times. The main difference is a slower disinflation and lower interest rates with little impact on output.

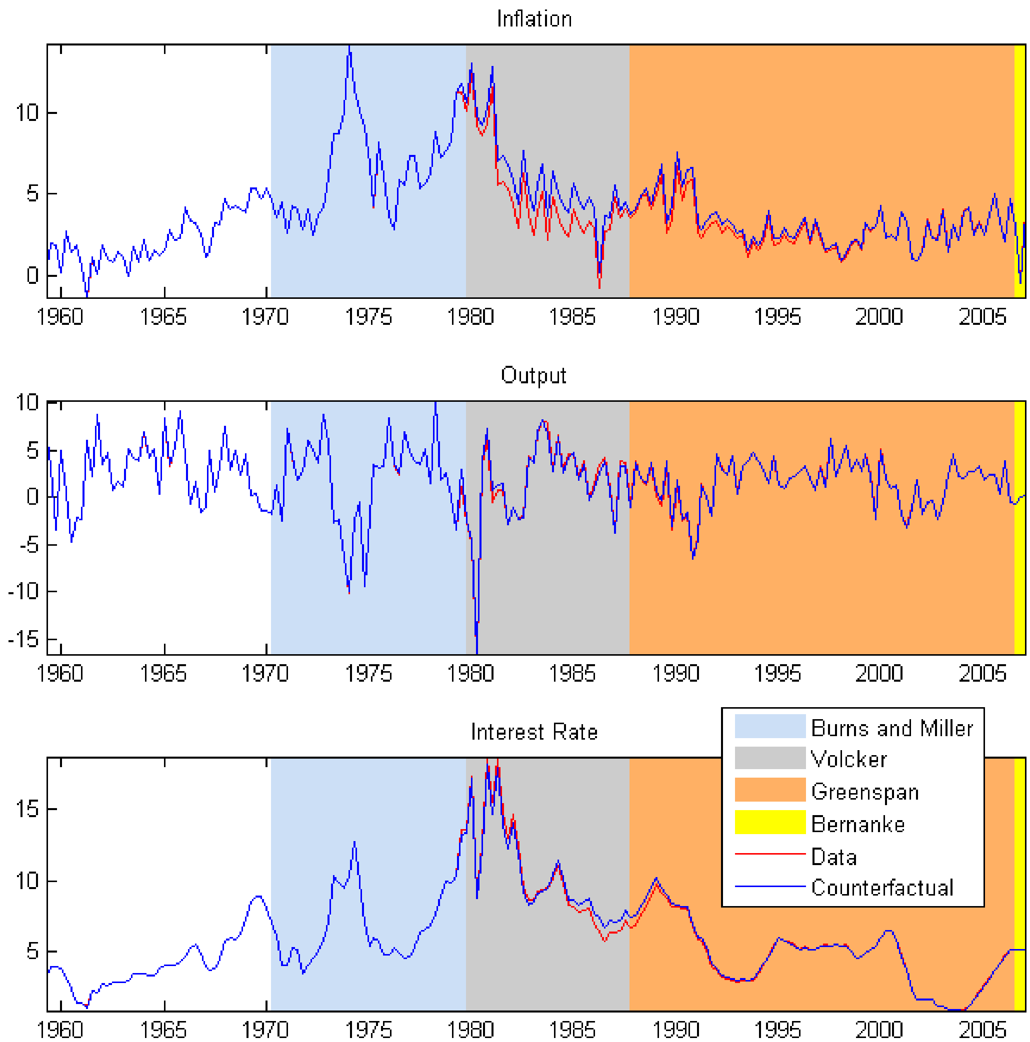


Chairman Greenspan during Volcker years

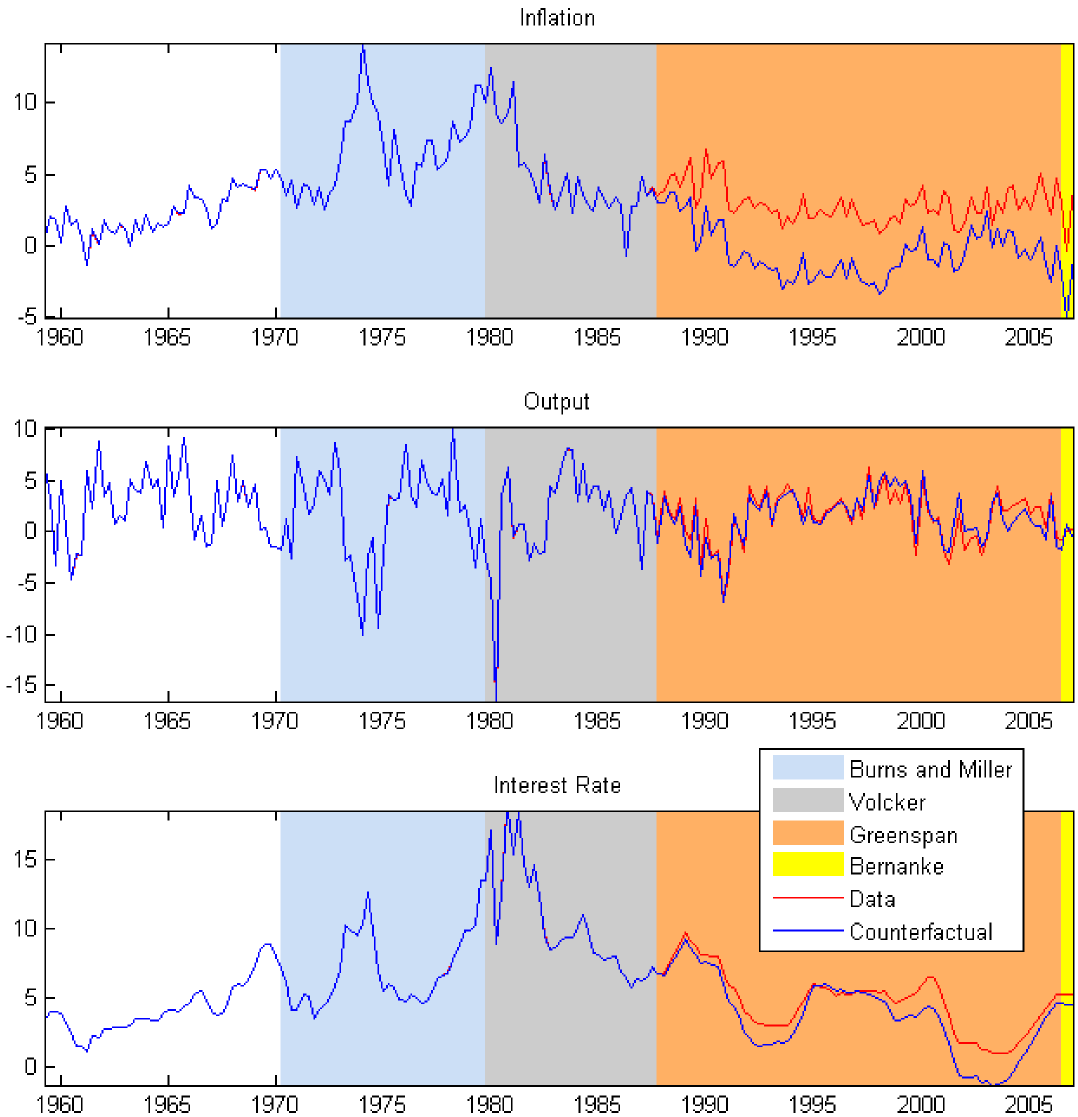
In figure 9.5, we move to Burns-Miller being resurrected at Greenspan times



Burns-Miller Chairmen during Greenspan years

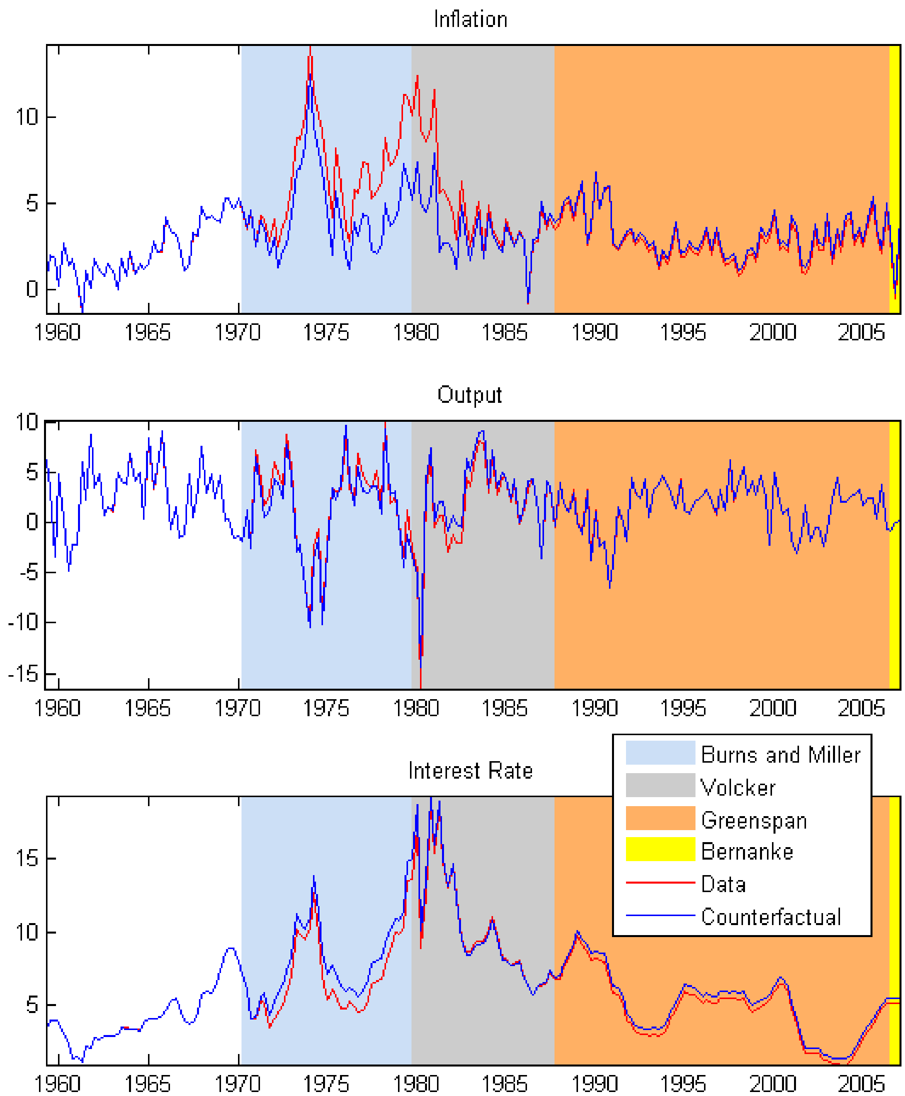


Burns-Miller Chairmen during Volcker years



Chairman Volcker during Greenspan years

A particularly interesting exercise if to check what would have happened if Reagan had decided to reappoint Volcker and not substitute him with Greenspan. The quick answer is much lower inflation and interest rates. This exercise has the problem that, during 2004-2005, the Fed should have been forced to have a negative federal funds rate. Forgetting for a moment about schemes to push the nominal interest rate below zero, this suggests that a nice extension of the model should take the existence of a zero-lower bound on interest rates seriously. This, however, would preclude us from using a perturbation method, and given the current computational frontier, make it impossible to estimate such model.



Chairman Volcker during Burns-Miller years

11. Conclusion

In this paper we have built and estimated a DSGE model with both stochastic volatility in the shocks and parameter drifting in policy rules. We have shown how such a rich model can be taken successfully to the data by taking advantage of our characterized of a higher order approximation of the decision rules of a large class of DSGE models. We can see many future applications for these tools.

With respect to our main empirical findings, a simple way to summarize them is to think about the recent monetary history of the U.S. as characterized by three eras:

- Little Fortune and little Virtue: Burns and Miller era, 1970-1979.
- Virtue but little Fortune: Volcker era, 1979-1987.
- Fortune but little Virtue: Greenspan era, 1987-2006.

Like all empirical work, our approach suffers from several problems. The most important, in our opinion, is the limitation of what we can learn from the data given our relatively short sample (Ploberger and Phillips, 2003, for a discussion of the problem of empirical limits for time series models in terms of information bounds). A source of information that can complement our quantitative investigation is a historical narrative based on the Federal Reserve statements and documents. We have in mind the type of work pioneered by Romer and Romer (2004) or Hetzel (2008). If the changes in policy uncovered by our estimates did, in fact, occurred, we should find telltale signs of them in all the record.

12. Appendix A: Theorem 1

Let us now prove theorem 1. In this theorem, we characterize the first and second order derivatives of the functions $h(\cdot)$ and $g(\cdot)$ evaluated at the non-stochastic steady state. We first show that the first partial derivatives of $h(\cdot)$ and $g(\cdot)$ with respect to any component of Σ_{t-1}, $\mathcal{U}_t$, or Λ evaluated at the non-stochastic steady is zero (or, in other words, that the the first order of the solution does not depend neither on volatility levels or shocks nor on the perturbation parameter. Second, it shows that. among many other results, the second partial derivative of $h(\cdot)$ and $g(\cdot)$ with respect to $u_{j,t}$ and any other state variable but $\varepsilon_{j,t}$ is also zero for any $j = \{1, \ldots, m\}$.

Before proceeding, note that we can write Z_t as a function of $Z_{t-1}, \Sigma_{t-1}, \mathcal{E}_t$, and $\mathcal{U}_t$:

$$\mathcal{Z}_t = \varsigma\left(\mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t\right)$$

and that Σ_t can be expressed as:

$$\Sigma_t = \vartheta\Sigma_{t-1} + \eta\mathcal{U}_t$$

where ϑ and η are both $m \times m$ diagonal matrices with diagonal elements equal to ϑ_i and η_i respectively. If we substitute the two functions (13) and (14) into (12) we get that:

$$F\left(\mathcal{S}_t, \mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t, \Lambda\right) \equiv$$

$$\mathbb{E}_t f\left(\begin{array}{c} g\left(h\left(\mathcal{S}_t, \mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t, \Lambda\right), \varsigma\left(\mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t\right), \vartheta\Sigma_{t-1} + \eta\mathcal{U}_t, \mathcal{E}_{t+1}, \mathcal{U}_{t+1}, \Lambda\right), \\ g\left(\mathcal{S}_t, \mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t, \Lambda\right), h\left(\mathcal{S}_t, \mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t, \Lambda\right), \mathcal{S}_t, \\ \varsigma\left(\varsigma\left(\mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t\right), \vartheta\Sigma_{t-1} + \eta\mathcal{U}_t, \mathcal{E}_{t+1}, \mathcal{U}_{t+1}\right), \varsigma\left(\mathcal{Z}_{t-1}, \Sigma_{t-1}, \mathcal{E}_t, \mathcal{U}_t\right) \end{array}\right) = 0.$$

To ease reading, we divide the proof in four parts, the first dealing with the first derivatives and the next three dealing with the second derivatives.

Proof, part 1. The first part of the proof deals with the first derivatives of (13) and (14) that are equal to zero. In particular, we want to show that:

$$\left[h_{\Sigma_{t-1}}\right]_j^{i_1} = \left[g_{\Sigma_{t-1}}\right]_j^{i_2} = \left[h_{\mathcal{U}_t}\right]_j^{i_1} = \left[g_{\mathcal{U}_t}\right]_j^{i_2} = \left[h_\Lambda\right]^{i_1} = \left[g_\Lambda\right]^{i_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, m\}$.

We show this result in three steps that basically repeat the same argument based on homogeneity of a system of linear equations:

1. We can write the derivative of $i-th$ element of F with respect to the $j-th$ element

of Σ_{t-1} as:

$$\left[F_{\Sigma_{t-1}}\right]_j^i = \left[f_{\mathcal{Y}_{t+1}}\right]_{i_2}^i \left(\left[g_{\mathcal{S}_t}\right]_{i_1}^{i_2} \left[h_{\Sigma_{t-1}}\right]_j^{i_1} + \left[g_{\Sigma_{t-1}}\right]_j^{i_2} \vartheta_j\right) + \left[f_{\mathcal{Y}_t}\right]_{i_2}^i \left[g_{\Sigma_{t-1}}\right]_j^{i_2} + \left[f_{\mathcal{S}_{t+1}}\right]_{i_1}^i \left[h_{\Sigma_{t-1}}\right]_j^{i_1} = 0$$

for $i \in \{1, \ldots, k+n+m\}$ and $j \in \{1, \ldots, m\}$. This is an homogenous system on $\left[h_{\Sigma_{t-1}}\right]_j^{i_1}$ and $\left[g_{\Sigma_{t-1}}\right]_j^{i_2}$ for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, m\}$. Thus:

$$\left[h_{\Sigma_{t-1}}\right]_j^{i_1} = \left[g_{\Sigma_{t-1}}\right]_j^{i_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, m\}$.

2. We can write the derivative of $i-th$ element of F with respect to the $j-th$ element of $\mathcal{U}_t$ as:

$$\left[F_{\mathcal{U}_t}\right]_j^i = \left[f_{\mathcal{Y}_{t+1}}\right]_{i_2}^i \left(\left[g_{\mathcal{S}_t}\right]_{i_1}^{i_2} \left[h_{\mathcal{U}_t}\right]_j^{i_1} + \left[g_{\Sigma_{t-1}}\right]_j^{i_2} \eta_j\right) + \left[f_{\mathcal{Y}_t}\right]_{i_2}^i \left[g_{\mathcal{U}_t}\right]_j^{i_2} + \left[f_{\mathcal{S}_{t+1}}\right]_{i_1}^i \left[h_{\mathcal{U}_t}\right]_j^{i_1} = 0$$

for $i \in \{1, \ldots, k+n+m\}$ and $j \in \{1, \ldots, m\}$. Since we have already shown that $\left[g_{\Sigma_{t-1}}\right]_j^{i_2} = 0$ for $i_2 \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, m\}$, this is an homogenous system on $\left[h_{\mathcal{U}_t}\right]_j^{i_1}$ and $\left[g_{\mathcal{U}_t}\right]_j^{i_2}$ for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, m\}$, Thus,:

$$\left[h_{\mathcal{U}_t}\right]_j^{i_1} = \left[g_{\mathcal{U}_t}\right]_j^{i_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, m\}$.

3. Finally, we can write the derivative of $i-th$ element of F with respect to Λ as:

$$\left[F_{\Lambda}\right]^i = \left[f_{\mathcal{Y}_{t+1}}\right]_{i_2}^i \left(\left[g_{\mathcal{S}_t}\right]_{i_1}^{i_2} \left[h_{\Lambda}\right]^{i_1} + \left[g_{\Lambda}\right]^{i_2}\right) + \left[f_{\mathcal{Y}_t}\right]_{i_2}^i \left[g_{\Lambda}\right]^{i_2} + \left[f_{\mathcal{S}_{t+1}}\right]_{i_1}^i \left[h_{\Lambda}\right]^{i_1} = 0$$

for $i \in \{1, \ldots, k+n+m\}$. Since this is an homogenous system on $\left[h_{\Lambda}\right]^{i_1}$ and $\left[g_{\Lambda}\right]^{i_2}$ for $i_1 \in \{1, \ldots, n\}$ and $i_2 \in \{1, \ldots, k\}$, we have that:

$$\left[h_{\Lambda}\right]^{i_1} = \left[g_{\Lambda}\right]^{i_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$ and $i_2 \in \{1, \ldots, k\}$.

■

Proof, part 2. The second part of the proof deals with the cross derivatives of (13) and (14) with respect to Λ and any of $\mathcal{S}_t$, $\mathcal{Z}_{t-1}$, Σ_{t-1}, $\mathcal{E}_t$, or $\mathcal{U}_t$ and it shows that all of them are equal to zero. In particular, we want to show that:

$$[h_{\Lambda,\mathcal{S}_t}]_j^{i_1} = [g_{\Lambda,\mathcal{S}_t}]_j^{i_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, n\}$ and:

$$\left[h_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_1} = \left[g_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_2} = \left[h_{\Lambda,\Sigma_{t-1}}\right]_j^{i_1} = \left[g_{\Lambda,\Sigma_{t-1}}\right]_j^{i_2} = [h_{\Lambda,\mathcal{E}_t}]_j^{i_1} = [g_{\Lambda,\mathcal{E}_t}]_j^{i_2} = [h_{\Lambda,\mathcal{U}_t}]_j^{i_1} = [g_{\Lambda,\mathcal{U}_t}]_j^{i_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, m\}$.

We show this result in five steps that, as in part 1 of the proof, exploit the homogeneity of a system of linear equations (and where we have already taken advantage of the terms that we know from the part 1 of the proof that they are equal to zero and eliminate them from our expressions):

1. We consider the cross derivative of $i-th$ element of F with respect to Λ and the $j-th$ element of $\mathcal{S}_t$:

   $$[F_{\Lambda,\mathcal{S}_t}]_j^i = \left[f_{\mathcal{Y}_{t+1}}\right]_{i_2}^i \left([g_{\mathcal{S}_t}]_{i_1}^{i_2} [h_{\Lambda,\mathcal{S}_t}]_j^{i_1} + [g_{\Lambda,\mathcal{S}_t}]_{i_1}^{i_2} [h_{\mathcal{S}_t}]_j^{i_1}\right) + [f_{\mathcal{Y}_t}]_{i_2}^i [g_{\Lambda,\mathcal{S}_t}]_j^{i_2} + \left[f_{\mathcal{S}_{t+1}}\right]_{i_1}^i [h_{\Lambda,\mathcal{S}_t}]_j^{i_1} = 0$$

   for $i \in \{1, \ldots, k+n+m\}$ and $j \in \{1, \ldots, n\}$. This is an homogenous system on $[h_{\Lambda,\mathcal{S}_t}]_j^{i_1}$ and $[g_{\Lambda,\mathcal{S}_t}]_j^{i_2}$ for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, n\}$. Thus:

   $$[h_{\Lambda,\mathcal{S}_t}]_j^{i_1} = [g_{\Lambda,\mathcal{S}_t}]_j^{i_2} = 0$$

   for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, n\}$.

2. We consider the cross derivative of $i-th$ element of F with respect to Λ and the $j-th$ element of $\mathcal{Z}_{t-1}$:

   $$\begin{aligned}\left[F_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^i = \left[f_{\mathcal{Y}_{t+1}}\right]_{i_2}^i &\left([g_{\mathcal{S}_t}]_{i_1}^{i_2} \left[h_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_1} + [g_{\Lambda,\mathcal{S}_t}]_{i_1}^{i_2} \left[h_{\mathcal{Z}_{t-1}}\right]_j^{i_1} + \rho_j \left[g_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_2}\right) \\ &+ [f_{\mathcal{Y}_t}]_{i_2}^i \left[g_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_2} + \left[f_{\mathcal{S}_{t+1}}\right]_{i_1}^i \left[h_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_1} = 0\end{aligned}$$

   for $i \in \{1, \ldots, k+n+m\}$ and $j \in \{1, \ldots, m\}$. Since $[g_{\Lambda,\mathcal{S}_t}]_j^{i_2} = 0$ for $i_2 \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, n\}$, this is an homogenous system on $\left[h_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_1}$ and $\left[g_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_2}$ for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, m\}$. Hence:

   $$\left[h_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_1} = \left[g_{\Lambda,\mathcal{Z}_{t-1}}\right]_j^{i_2} = 0$$

   for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j \in \{1, \ldots, m\}$.

3. We consider the cross derivative of $i-th$ element of F with respect to Λ and the $j-th$ element of Σ_{t-1}:

$$\begin{aligned}\left[F_{\Lambda,\Sigma_{t-1}}\right]^i_j = \left[f_{\mathcal{Y}_{t+1}}\right]^i_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\Lambda,\Sigma_{t-1}}\right]^{i_1}_j + \left[g_{\Lambda,\Sigma_{t-1}}\right]^{i_2}_j \vartheta_j\right)\\ + \left[f_{\mathcal{Y}_t}\right]^i_{i_2}\left[g_{\Lambda,\Sigma_{t-1}}\right]^{i_2}_j + \left[f_{\mathcal{S}_{t+1}}\right]^i_{i_1}\left[h_{\Lambda,\Sigma_{t-1}}\right]^{i_1}_j = 0\end{aligned}$$

for $i \in \{1,\ldots,k+n+m\}$ and $j \in \{1,\ldots,m\}$. This is an homogenous system on $\left[h_{\Lambda,\Sigma_{t-1}}\right]^{i_1}_j$ and $\left[g_{\Lambda,\Sigma_{t-1}}\right]^{i_2}_j$ for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j \in \{1,\ldots,m\}$. Hence:

$$\left[h_{\Lambda,\Sigma_{t-1}}\right]^{i_1}_j = \left[g_{\Lambda,\Sigma_{t-1}}\right]^{i_2}_j = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j \in \{1,\ldots,m\}$.

4. We consider the cross derivative of $i-th$ element of F with respect to Λ and the $j-th$ element of $\mathcal{E}_t$:

$$\begin{aligned}\left[F_{\Lambda,\mathcal{E}_t}\right]^i_j = \left[f_{\mathcal{Y}_{t+1}}\right]^i_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\Lambda,\mathcal{E}_t}\right]^{i_1}_j + \left[g_{\Lambda,\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\mathcal{E}_t}\right]^{i_1}_j + \exp^{\vartheta_j\sigma_{j,t-1}}\left[g_{\Lambda,\mathcal{Z}_{t-1}}\right]^{i_2}_j\right)\\ + \left[f_{\mathcal{Y}_t}\right]^i_{i_2}\left[g_{\Lambda,\mathcal{E}_t}\right]^{i_2}_j + \left[f_{\mathcal{S}_{t+1}}\right]^i_{i_1}\left[h_{\Lambda,\mathcal{E}_t}\right]^{i_1}_j = 0\end{aligned}$$

for $i \in \{1,\ldots,k+n+m\}$ and $j \in \{1,\ldots,m\}$. Since $\left[g_{\Lambda,\mathcal{Z}_{t-1}}\right]^{i_2}_j = 0$ for $i_2 \in \{1,\ldots,k\}$ and $j \in \{1,\ldots,m\}$ and $\left[g_{\Lambda,\mathcal{S}_t}\right]^{i_2}_j = 0$ for $i_2 \in \{1,\ldots,k\}$ and $j \in \{1,\ldots,n\}$, this is an homogenous system on $\left[h_{\Lambda,\mathcal{E}_t}\right]^{i_1}_j$ and $\left[g_{\Lambda,\mathcal{E}_t}\right]^{i_2}_j$ for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j \in \{1,\ldots,m\}$. Thus:

$$\left[h_{\Lambda,\mathcal{E}_t}\right]^{i_1}_j = \left[g_{\Lambda,\mathcal{E}_t}\right]^{i_2}_j = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j \in \{1,\ldots,m\}$.

5. We consider the cross derivative of $i-th$ element of F with respect to Λ and the $j-th$ element of $\mathcal{U}_t$:

$$\left[F_{\Lambda,\mathcal{U}_t}\right]^i_j = \left[f_{\mathcal{Y}_{t+1}}\right]^i_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\Lambda,\mathcal{U}_t}\right]^{i_1}_j + \eta_j\left[g_{\Lambda,\Sigma_{t-1}}\right]^{i_2}_j\right) + \left[f_{\mathcal{Y}_t}\right]^i_{i_2}\left[g_{\Lambda,\mathcal{U}_t}\right]^{i_2}_j + \left[f_{\mathcal{S}_{t+1}}\right]^i_{i_1}\left[h_{\Lambda,\mathcal{U}_t}\right]^{i_1}_j = 0$$

for $i \in \{1,\ldots,k+n+m\}$ and $j \in \{1,\ldots,m\}$. Since we have shown that $\left[g_{\Lambda,\Sigma_{t-1}}\right]^{i_2}_j = 0$ for $i_2 \in \{1,\ldots,k\}$ and $j \in \{1,\ldots,m\}$, we have that the above system is an homogenous on $\left[h_{\Lambda,\mathcal{U}_t}\right]^{i_1}_j$ and $\left[g_{\Lambda,\mathcal{U}_t}\right]^{i_2}_j$ for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j \in \{1,\ldots,m\}$. Then:

$$\left[h_{\Lambda,\mathcal{U}_t}\right]^{i_1}_j = \left[g_{\Lambda,\mathcal{U}_t}\right]^{i_2}_j = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j \in \{1,\ldots,m\}$.

■

Proof, part 3. The third part of the proof deals with the cross derivatives of (13) and (14) with respect to Σ_{t-1} and any of $\mathcal{S}_t$, $\mathcal{Z}_{t-1}$, Σ_{t-1}, or $\mathcal{E}_t$ and it shows that all of them are equal to zero with one exception. In particular, we want to show that:

$$\left[h_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, $j_1 \in \{1,\ldots,n\}$, and $j_2 \in \{1,\ldots,m\}$,

$$\left[h_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = \left[h_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = \left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j_1, j_2 \in \{1,\ldots,m\}$, and

$$\left[h_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, and $j_1, j_2 \in \{1,\ldots,m\}$ if $j_1 \neq j_2$.

We show this result in four steps (and where we have already taken advantage of the terms that we know from the part 1 of the proof that they are equal to zero and eliminate them from our expressions):

1. We consider the cross derivative of $i-th$ element of F with respect to j_1-th element of $\mathcal{S}_t$ and the j_2-th element of Σ_{t-1}:

$$\begin{aligned}\left[F_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i}_{j_1,j_2} = \left[f_{\mathcal{Y}_{t+1}}\right]^{i}_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} + \left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{i_1,j_2}\left[h_{\mathcal{S}_t}\right]^{i_1}_{j_1}\vartheta_{j_2}\right)\\ + \left[f_{\mathcal{Y}_t}\right]^{i}_{i_2}\left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} + \left[f_{\mathcal{S}_{t+1}}\right]^{i}_{i_1}\left[h_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = 0\end{aligned}$$

   for $i \in \{1,\ldots,k+n+m\}$, $j_1 \in \{1,\ldots,n\}$, and $j_2 \in \{1,\ldots,m\}$. This is an homogenous system on $\left[h_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}$ and $\left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}$ for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, $j_1 \in \{1,\ldots,n\}$, and $j_2 \in \{1,\ldots,m\}$. Therefore:

$$\left[h_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{i_2,j_2} = 0$$

   for $i_1 \in \{1,\ldots,n\}$, $i_2 \in \{1,\ldots,k\}$, $j_1 \in \{1,\ldots,n\}$, and $j_2 \in \{1,\ldots,m\}$.

2. We consider the cross derivative of $i-th$ element of F with respect to j_1-th element

of $\mathcal{Z}_{t-1}$ and the j_2-th element of Σ_{t-1}:

$$\left[F_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i}_{j_1,j_2}$$
$$= \left[f_{\mathcal{Y}_{t+1}}\right]^{i}_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}+\left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{i_1,j_2}\left[h_{\mathcal{Z}_{t-1}}\right]^{i_1}_{j_1}\vartheta_{j_2}+\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}\vartheta_{j_2}\rho_{j_1}\right)$$
$$+\left[f_{\mathcal{Y}_t}\right]^{i}_{i_2}\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}+\left[f_{\mathcal{S}_{t+1}}\right]^{i}_{i_1}\left[h_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}=0$$

for $i\in\{1,\ldots,k+n+m\}$, and $j_1,j_2\in\{1,\ldots,m\}$. Since we just found that $\left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}=0$ for $i_2\in\{1,\ldots,k\}$, $j_1\in\{1,\ldots,n\}$, and $j_2\in\{1,\ldots,m\}$, this is an homogenous system on $\left[h_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}$ and $\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{i_2,j_2}$ for $i_1\in\{1,\ldots,n\}$, $i_2\in\{1,\ldots,k\}$, and $j_1,j_2\in\{1,\ldots,m\}$. Therefore:

$$\left[h_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}=\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{i_2,j_2}=0$$

for $i_1\in\{1,\ldots,n\}$, $i_2\in\{1,\ldots,k\}$, and $j_1,j_2\in\{1,\ldots,m\}$.

3. We consider the cross derivative of $i-th$ element of F with respect to j_1-th element of Σ_{t-1} and the j_2-th element of Σ_{t-1}:

$$\left[F_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i}_{j_1,j_2}=$$
$$\left[f_{\mathcal{Y}_{t+1}}\right]^{i}_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}+\left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}\vartheta_{j_1}\vartheta_{j_2}\right)$$
$$+\left[f_{\mathcal{Y}_t}\right]^{i}_{i_2}\left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}+\left[f_{\mathcal{S}_{t+1}}\right]^{i}_{i_1}\left[h_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}=0$$

for $i\in\{1,\ldots,k+n+m\}$ and $j_1,j_2\in\{1,\ldots,m\}$. This is an homogenous system on $\left[h_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}$ and $\left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}$ for $i_1\in\{1,\ldots,n\}$, $i_2\in\{1,\ldots,k\}$, and $j_1,j_2\in\{1,\ldots,m\}$, therefore:

$$\left[h_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}=\left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}=0$$

for $i_1\in\{1,\ldots,n\}$, $i_2\in\{1,\ldots,k\}$, and $j_1,j_2\in\{1,\ldots,m\}$.

4. We consider the cross derivative of $i-th$ element of F with respect to j_1-th element of $\mathcal{E}_t$ and the j_2-th element of Σ_{t-1} if $j_1\neq j_2$:

$$\left[F_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i}_{j_1,j_2}=$$
$$\left[f_{\mathcal{Y}_{t+1}}\right]^{i}_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}+\left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{i_1,j_2}\left[h_{\mathcal{E}_t}\right]^{i_1}_{j_1}\vartheta_{j_2}+\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}\exp^{\vartheta_{j_1}\sigma_{j_1,t-1}}\vartheta_{j_2}\right)$$
$$+\left[f_{\mathcal{Y}_t}\right]^{i}_{i_2}\left[g_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}+\left[f_{\mathcal{S}_{t+1}}\right]^{i}_{i_1}\left[h_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}=0$$

for $i \in \{1, \ldots, k+n+m\}$ and $j_1, j_2 \in \{1, \ldots, m\}$. Since we know that $\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = \left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{j,j_2} = 0$ for $i_2 \in \{1, \ldots, k\}$, $j \in \{1, \ldots, n\}$, and $j_1, j_2 \in \{1, \ldots, m\}$, this is an homogenous system on $\left[h_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2}$ and $\left[g_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}$ for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j_1, j_2 \in \{1, \ldots, m\}$ if $j_1 \neq j_2$. Therefore:

$$\left[h_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = \left[g_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j_1, j_2 \in \{1, \ldots, m\}$ if $j_1 \neq j_2$.

Note that if $j_1 = j_2$, we have that:

$$\begin{aligned}
&\left[F_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i}_{j_1,j_1} = \left[f_{\mathcal{Y}_{t+1}}\right]^{i}_{i_2} * \\
&* \left([g_{\mathcal{S}_t}]^{i_2}_{i_1} \left[h_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_1} + \vartheta_{j_1} \left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{i_1,j_1} [h_{\mathcal{E}_t}]^{i_1}_{j_1} + \left(\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_1} + \left[g_{\mathcal{Z}_{t-1}}\right]^{i_2}_{j_1} \right) \exp^{\vartheta_{j_1}\sigma_{j_1,t-1}} \vartheta_{j_1} \right) \\
&+ [f_{\mathcal{Y}_t}]^{i}_{i_2} \left[g_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_1} + \left[f_{\mathcal{S}_{t+1}}\right]^{i}_{i_1} \left[h_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_1} \\
&+ \left([f_{\mathcal{Z}_t}]^{i}_{j_1} + \rho_{j_1} \left[f_{\mathcal{Z}_{t+1}}\right]^{i}_{j_1} \right) \exp^{\vartheta_{j_1}\sigma_{j_1,t-1}} \vartheta_{j_1} = 0
\end{aligned}$$

and since $[f_{\mathcal{Z}_t}]^{i}_{j_1}$ and $\left[f_{\mathcal{Z}_{t+1}}\right]^{i}_{j_1}$ are different from zero in general for $i \in \{1, \ldots, k+n+m\}$ and $j_1 \in \{1, \ldots, m\}$, we have that this system is not homogenous and

$$\left[h_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_1} = \left[g_{\mathcal{E}_t,\Sigma_{t-1}}\right]^{i_1}_{j_1,j_1} \neq 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j_1 \in \{1, \ldots, m\}$.

■

Proof, part 4. The fourth, and final, part of the proof deals with the cross derivatives of (13) and (14) with respect to $\mathcal{U}_t$ and any of $\mathcal{S}_t$, $\mathcal{Z}_{t-1}$, Σ_{t-1}, $\mathcal{E}_t$, or $\mathcal{U}_t$ and it shows that all of them are equal to zero with one exception. In particular, we want to show that:

$$[h_{\mathcal{S}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_2} = [g_{\mathcal{S}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, $j_1 \in \{1, \ldots, n\}$, and $j_2 \in \{1, \ldots, m\}$,

$$\left[h_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_2}_{j_1,j_2} = \left[h_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} = \left[g_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_2}_{j_1,j_2} = [h_{\mathcal{U}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_2} = [g_{\mathcal{U}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j_1, j_2 \in \{1, \ldots, m\}$, and

$$[h_{\mathcal{E}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_2} = [g_{\mathcal{E}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, $j_1, j_2 \in \{1, \ldots, m\}$, and $j_1 \neq j_2$.

Again, we follow the same steps for each part of the result than before and use our previous findings regarding which terms are zero.

1. We consider the cross derivative of $i-th$ element of F with respect to j_1-th element of S_t and the j_2-th element of $\mathcal{U}_t$:

$$\left[F_{\mathcal{S}_t,\mathcal{U}_t}\right]^i_{j_1,j_2} = \left[f_{\mathcal{Y}_{t+1}}\right]^i_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\mathcal{S}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} + \left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{i_1,j_2}\left[h_{\mathcal{S}_t}\right]^{i_1}_{j_1}\eta_{j_2}\right)$$
$$+\left[f_{\mathcal{Y}_t}\right]^i_{i_2}\left[g_{\mathcal{S}_t,\mathcal{U}_t}\right]^{i_2}_{j_1,j_2} + \left[f_{\mathcal{S}_{t+1}}\right]^i_{i_1}\left[h_{\mathcal{S}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} = 0$$

for $i \in \{1, \ldots, k+n+m\}$, $j_1 \in \{1, \ldots, n\}$, and $j_2 \in \{1, \ldots, m\}$. Since $\left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = 0$ for $i_2 \in \{1, \ldots, k\}$, $j_1 \in \{1, \ldots, n\}$, and $j_2 \in \{1, \ldots, m\}$, this is an homogenous system on $\left[h_{\mathcal{S}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}$ and $\left[g_{\mathcal{S}_t,\mathcal{U}_t}\right]^{i_2}_{j_1,j_2}$.Therefore:

$$\left[h_{\mathcal{S}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{S}_t,\mathcal{U}_t}\right]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, $j_1 \in \{1, \ldots, n\}$, and $j_2 \in \{1, \ldots, m\}$.

2. We consider the cross derivative of $i-th$ element of F with respect to j_1-th element of $\mathcal{Z}_{t-1}$ and the j_2-th element of $\mathcal{U}_t$:

$$\left[F_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^i_{j_1,j_2} =$$
$$\left[f_{\mathcal{Y}_{t+1}}\right]^i_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} + \eta_{j_2}\left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{i_1,j_2}\left[h_{\mathcal{Z}_t}\right]^{i_1}_{j_1} + \rho_{j_1}\eta_{j_2}\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}\right)$$
$$+\left[f_{\mathcal{Y}_t}\right]^i_{i_1}\left[g_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} + \left[f_{\mathcal{S}_{t+1}}\right]^i_{i_1}\left[h_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} = 0$$

for $i \in \{1, \ldots, k+n+m\}$, and $j_1, j_2 \in \{1, \ldots, m\}$. Since $\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = \left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{j,j_2} = 0$ for $i_2 \in \{1, \ldots, k\}$, $j \in \{1, \ldots, n\}$, and $j_1, j_2 \in \{1, \ldots, m\}$, this is an homogenous system on $\left[h_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}$ and $\left[g_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}$ for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j_1, j_2 \in \{1, \ldots, m\}$. Therefore:

$$\left[h_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2} = \left[g_{\mathcal{Z}_{t-1},\mathcal{U}_t}\right]^{i_2}_{j_1,j_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j_1, j_2 \in \{1, \ldots, m\}$.

3. We consider the cross derivative of $i-th$ element of F with respect to j_1-th element

of Σ_{t-1} and the j_2-th element of $\mathcal{U}_t$:

$$\begin{gathered}\left[F_{\Sigma_{t-1},\mathcal{U}_t}\right]^i_{j_1,j_2} = \\ \left[f_{\mathcal{Y}_{t+1}}\right]^i_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}+\left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}\vartheta_{j_1}\eta_{j_2}\right) \\ +\left[f_{\mathcal{Y}_t}\right]^i_{i_1}\left[g_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}+\left[f_{\mathcal{S}_{t+1}}\right]^i_{i_1}\left[h_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}=0\end{gathered}$$

for $i\in\{1,\ldots,k+n+m\}$, and $j_1,j_2\in\{1,\ldots,m\}$. Since $\left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}=0$ for $i_2\in\{1,\ldots,k\}$, and $j_1,j_2\in\{1,\ldots,m\}$, this is an homogenous system on $\left[h_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}$ and $\left[g_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_2}_{j_1,j_2}$ for $i_1\in\{1,\ldots,n\}$, $i_2\in\{1,\ldots,k\}$, and $j_1,j_2\in\{1,\ldots,m\}$. Therefore:

$$\left[h_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}=\left[g_{\Sigma_{t-1},\mathcal{U}_t}\right]^{i_2}_{j_1,j_2}=0$$

for $i_1\in\{1,\ldots,n\}$, $i_2\in\{1,\ldots,k\}$, $j_1,j_2\in\{1,\ldots,m\}$.

4. We consider the cross derivative of $i-th$ element of F with respect to j_1-th element of $\mathcal{U}_t$ and the j_2-th element of $\mathcal{U}_t$:

$$\begin{gathered}\left[F_{\mathcal{U}_t,\mathcal{U}_t}\right]^i_{j_1,j_2} = \\ \left[f_{\mathcal{Y}_{t+1}}\right]^i_{i_2}\left(\eta_{j_1}\eta_{j_2}\left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}+\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\mathcal{U}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}\right) \\ +\left[f_{\mathcal{Y}_t}\right]^i_{i_2}\left[g_{\mathcal{U}_t,\mathcal{U}_t}\right]^{i_2}_{j_1,j_2}+\left[f_{\mathcal{S}_{t+1}}\right]^i_{i_1}\left[h_{\mathcal{U}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}=0\end{gathered}$$

for $i\in\{1,\ldots,k+n+m\}$ and $j_1,j_2\in\{1,\ldots,m\}$. Since $\left[g_{\Sigma_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}=0$ for $i_1\in\{1,\ldots,k\}$ and $j_1,j_2\in\{1,\ldots,m\}$, this is an homogenous system on $\left[h_{\mathcal{U}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}$ and $\left[g_{\mathcal{U}_t,\mathcal{U}_t}\right]^{i_2}_{j_1,j_2}$ for $i_1\in\{1,\ldots,n\}$, $i_2\in\{1,\ldots,k\}$, and $j_1,j_2\in\{1,\ldots,m\}$. Therefore:

$$\left[h_{\mathcal{U}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}=\left[g_{\mathcal{U}_t,\mathcal{U}_t}\right]^{i_2}_{j_1,j_2}=0$$

for $i_1\in\{1,\ldots,n\}$, $i_2\in\{1,\ldots,k\}$, and $j_1,j_2\in\{1,\ldots,m\}$.

5. Finally, consider the cross derivative of $i-th$ element of F with respect to j_1-th element of $\mathcal{E}_t$ and the j_2-th element of $\mathcal{U}_t$ if $j_1\neq j_2$:

$$\begin{gathered}\left[F_{\mathcal{E}_t,\mathcal{U}_t}\right]^i_{j_1,j_2} = \\ \left[f_{\mathcal{Y}_{t+1}}\right]^i_{i_2}\left(\left[g_{\mathcal{S}_t}\right]^{i_2}_{i_1}\left[h_{\mathcal{E}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}+\eta_{j_2}\left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{i_1,j_2}\left[h_{\mathcal{E}_t}\right]^{i_1}_{j_1}+\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2}\exp^{\vartheta_{j_1}\sigma_{j_1,t-1}}\eta_{j_2}\right) \\ +\left[f_{\mathcal{Y}_t}\right]^i_{i_2}\left[g_{\mathcal{E}_t,\mathcal{U}_t}\right]^{i_2}_{j_1,j_2}+\left[f_{\mathcal{S}_{t+1}}\right]^i_{i_1}\left[h_{\mathcal{E}_t,\mathcal{U}_t}\right]^{i_1}_{j_1,j_2}=0\end{gathered}$$

for $i \in \{1, \ldots, k+n+m\}$ and $j_1, j_2 \in \{1, \ldots, m\}$. Since $\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_2} = \left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{j,j_2} = 0$ for $i_2 \in \{1, \ldots, k\}$, $j \in \{1, \ldots, n\}$, and $j_1, j_2 \in \{1, \ldots, m\}$, this is an homogenous system on $[h_{\mathcal{E}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_2}$ and $[g_{\mathcal{E}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_2}$ for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j_1, j_2 \in \{1, \ldots, m\}$ if $j_1 \neq j_2$. Therefore:

$$[h_{\mathcal{E}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_2} = [g_{\mathcal{E}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_2} = 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j_1, j_2 \in \{1, \ldots, m\}$ if $j_1 \neq j_2$.

Note that if $j_1 = j_2$, we have that:

$$\begin{gathered}[F_{\mathcal{E}_t,\mathcal{U}_t}]^{i}_{j_1,j_1} = \\ \left[f_{\mathcal{Y}_{t+1}}\right]^{i}_{i_2}\left(\exp^{\vartheta_{j_1}\sigma_{j_1,t-1}}\eta_{j_1}\left(\left[g_{\mathcal{Z}_{t-1},\Sigma_{t-1}}\right]^{i_2}_{j_1,j_1} + \left[g_{\mathcal{Z}_{t-1}}\right]^{i_2}_{j_1}\right) + \eta_{j_1}\left[g_{\mathcal{S}_t,\Sigma_{t-1}}\right]^{i_2}_{i_1,j_1}[h_{\mathcal{E}_t}]^{i_1}_{j_1} + [g_{\mathcal{S}_t}]^{i_2}_{i_1}[h_{\mathcal{E}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_1}\right) \\ + [f_{\mathcal{Y}_t}]^{i}_{i_2}[g_{\mathcal{E}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_1} + \left[f_{\mathcal{S}_{t+1}}\right]^{i}_{i_1}[h_{\mathcal{E}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_1} \\ + \exp^{\vartheta_{j_1}\sigma_{j_1,t-1}}\eta_{j_1}\left([f_{\mathcal{Z}_t}]^{i}_{j_1} + \rho_{j_1}\left[f_{\mathcal{Z}_{t+1}}\right]^{i}_{j_1}\right) = 0\end{gathered}$$

and since $[f_{\mathcal{Z}_t}]^{i}_{j_1}$ and $\left[f_{\mathcal{Z}_{t+1}}\right]^{i}_{j_1}$ are different from zero in general for $i \in \{1, \ldots, k+n+m\}$ and $j_1 \in \{1, \ldots, m\}$ we have that this system is not homogenous and hence:

$$[h_{\mathcal{E}_t,\mathcal{U}_t}]^{i_2}_{j_1,j_1} = [g_{\mathcal{E}_t,\mathcal{U}_t}]^{i_1}_{j_1,j_1} \neq 0$$

for $i_1 \in \{1, \ldots, n\}$, $i_2 \in \{1, \ldots, k\}$, and $j_1 \in \{1, \ldots, m\}$. ∎

13. Appendix B: Computation

In this appendix we provide some more details regarding the computation of the paper. We generate all the derivatives required by our higher-order perturbation with `Mathematica` 6.0. In that way, we do not need to recompute the derivatives, the most time-intensive step, for each set of parameter values in our estimation. Once we have all the relevant derivatives, we export them automatically into Fortran files. This whole process takes about 3 hours.

Then, we compile the resulting files with the `Intel Fortran Compiler` version 10.1.025 with `IMSL`. Previous versions failed to compile our project because of the length of some of the expression. Compilation takes about 18 hours. The project has 1798 files and occupies 2.33 Gbytes of memory.

The next step is, for given parameter values, to compute the first- and second-order approximation to the decision rules around the deterministic steady-state using the analytic derivatives we found before. For this task, Fortran takes around 5 seconds. Once we have the solution, we approximate the likelihood using the Particle Filter with $10,000$ particles. This number delivered a good compromise between accuracy and time to compute the likelihood. The evaluation of one likelihood requires 22 seconds in a Dell Server with 8 processors. Once you have the likelihood evaluation, we guess new parameter values and we restart again. This means that drawing 5,000 times from the posterior (even forgetting about the initial search over a grid of parameter values), takes around 38 hours.

It is important to emphasize that the `Mathematica` and `Fortran` code were highly optimized in order to 1) keep the size of the project within reasonable dimensions (otherwise, the compiler cannot sparse the files and, even when it can, it delivers code that it is too inefficient) and 2) provide a fast computation of the likelihood.

Perhaps the most important task in that optimization was the parallelization of the `Fortran` code using `OPENMP` as well as the compilation options: `OG` (global optimizations) and `Loop Unroll`. In addition, we tailored specialized code to perform the matrix multiplications required in the first- and second-order terms of our model solution.

Implementing corollary 1 requires the solution of a linear system of equations and the computation of a Jacobian. For our particular application, we found that the following sequence of `LAPACK` operations delivered the fastest solution:

1. `DGESV` (computes the solution to a real system of linear equations $\mathtt{A} * \mathtt{X} = \mathtt{B}$).
2. `DGETRI` (computes the inverse of a matrix using the LU factorization from the previous line).
3. `DGETRF` (helps to compute the determinant of the inverse from the previous line).

Without the parallelization and our optimized code, the solution of the model and evaluation of its likelihood take about 70 seconds.

With respect to the Random-Walk Metropolis-Hastings, we performed an intensive process of fine-tuning of the chain, both in terms of initial conditions as in terms of getting the right acceptance level. The only other important remark is to remember that as pointed out by McFadden (1989) and Pakes and Pollard (1989), we must keep the random numbers used for resampling in the Particle filter constant across draws of the Markov Chain. This is required to achieve stochastic equicontinuity, and even if this condition is not strictly necessary in a Bayesian framework, it reduces the numerical variance of the procedure, which was a serious concern for us given the complexity of our problem.

14. Appendix C: Construction of Data

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

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

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

References

[1] 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.

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

[3] Bernanke, B.S. and I. Mihov (1998). "Measuring Monetary Policy." *Quarterly Journal of Economics* 113, 869-902.

[4] Blanchard, O.J. and J. Simon (2001). "The Long and Large Decline in U.S. Output Volatility." *Brookings Papers on Economic Activity* 2001-1, 187-207.

[5] Canova, F. (2006). "You can Use VARs for Structural Analyses. A Comment to VARs and Great Moderation." Working paper CREI.

[6] Canova, F. and L. Gambetti (2009). "Structural Changes in the US Economy: Is there a Role for Monetary Policy?" *Journal of Economic Dynamics and Control* 33, 477-490

[7] Christoffel, K., G. Coenen, and A. Warne (2008) "The New Area-Wide Model of the Euro Area: A Micro-Founded Open-Economy Model for Forecasting and Policy Analysis." ECB Working Paper No. 944.

[8] Clarida, R., J. Galí, and M. Gertler (2000). "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." Quarterly Journal of Economics 115, 147-180.

[9] Cogley, T. and T.J. Sargent (2001). "Evolving Post-World War II U.S. Inflation Dynamics." *NBER Macroeconomics Annual*, vol. 16.

[10] Cogley, T. and T.J. Sargent (2005). "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S." *Review of Economic Dynamics* 8, 262-302.

[11] DeJong, D.N., H. Dharmarajan, R. Liesenfeld, and J.F. Richard (2007). "An Efficient Filtering Approach to Likelihood Approximation for State-Space Representations." *Mimeo*, University of Pittsburgh.

[12] Del Moral P. and J. Jacod (2002). "The Monte-Carlo Method for Filtering with Discrete Time Observations. Central Limit Theorems." in T. J. Lyons and T. S. Salisbury (eds.), *Numerical Methods and Stochastics*. The Fields Institute Communications, American Mathematical Society.

[13] Dynan, K.E., D.W. Elmendorf, and D.E. Sichel (2006). "Can financial innovation help to explain the reduced volatility of economic activity?," *Journal of Monetary Economics* 53, 123-150.

[14] Edge, R., M.T. Kiley, and J.-P. Laforte (2007). “Documentation of the Research and Statistics Division’s Estimated DSGE Model of the U.S. Economy: 2006 Version.” Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series 2007-53.

[15] Farmer, R.E., D.F. Waggoner, and T. Zha (2006a). “Minimal State Variable Solutions to Markov-Switching Rational Expectations Models.” Mimeo, Federal Reserve Bank of Atlanta.

[16] Farmer, R.E., D.F. Waggoner, and T. Zha (2006b). “Indeterminacy in a Forward Looking Regime Switching Model.” Mimeo, Federal Reserve Bank of Atlanta.

[17] Fermanian, J.D. and B. Salanié (2004). “A Nonparametric Simulated Maximum Likelihood Estimation Method.” *Econometric Theory* 20, 701-734.

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

[19] Fernandez-Villaverde, J. and J. Rubio-Ramirez (2006). “Solving DSGE Models with Perturbation Methods and a Change of Variables.” Journal of Economic Dynamics and Control 30, 2509-2531.

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

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

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

[23] Fernández-Villaverde, J., P. Guerrón-Quintana, and J. Rubio-Ramírez (2008). “The New Macroeconometrics: A Bayesian Approach.” in A. O’Hagan and M. West (eds.) Handbook of Applied Bayesian Analysis, Oxford University Press.

[24] Fiorentini, G., E. Sentanta, and N. Shephard (2004). “Likelihood Estimation of Latent Generalised ARCH Structures.” *Econometrica* 72, 1481-1517.

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

[26] Galí, J. and L. Gambetti (2009). “On the Sources of the Great Moderation.” *American Economic Journal: Macroeconomics* 1, 26-57.

[27] Geweke, J. (1998). “Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication”. *Staff Report* 249, Federal Reserve Bak of Minneapolis.

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

[29] 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.

[30] Guerrón-Quintana, P. (2009a). "Money Demand Heterogeneity and the Great Moderation External Link." *Journal of Monetary Economics* 56, 255-266.

[31] Guerrón-Quintana, P. (2009b). "What You Match Does Matter: The Effects of Data on DSGE Estimation." Forthcoming, *Journal of Applied Econometrics.*

[32] Hetzel, R.L. (2008). *The Monetary Policy of the Federal Reserve: a History.* Cambridge University Press.

[33] Hurwicz, L. (1962). "On the Structural Form of IInterdependent Systems." Logic, Methodology and Philosophy of Science, Proceedings of the 1960 International Congress, 232-39.

[34] Jeffreys H. (1961). *The Theory of Probability* (3rd ed). Oxford University Press.

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

[36] Kim, S., N. Shephard, and S. Chib (1998). "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models." *Review of Economic Studies* 65, 361-93.

[37] Klenow, P.J. and B.A. Malin (2009). "Microeconomic Evidence on Price-Setting." Chapter prepared for the *Handbook of Monetary Economic*s.

[38] Koop G., M. Hashem Pesaran, and S. Potter (1996). "Impulse Response Analysis in Nonlinear Multivariate Models." *Journal of Econometrics* 74, 199-147.

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

[40] McFadden, D.L. (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica* 57, 995-1026.

[41] Negro, M.D. and F. Schorfheide. "Forming Priors for DSGE Models (and How It Affects the Assessment of Nominal Rigidities)." *Journal of Monetary Economics* 55, 1191-1208.

[42] Pakes, A. and D. Pollard (1989). "Simulation and the Asymptotics of Optimization Estimators." *Econometrica* 57, 1027-1057.

[43] Phillips, P.C.B. and W. Ploberger (2003). "Empirical Limits for Time Series Econometric Models." *Econometrica* 71, 627-673.

[44] Primiceri, G.E. (2005). "Time Varying Structural Vector Autoregressions and Monetary Policy". *Review of Economic Studies* 72, 821-852.

[45] Ramey, V.A. and D.J. Vine (2006). "Declining Volatility in the U.S. Automobile Industry." *American Economic Review* 96, 1876-1889.

[46] Rogerson, R. and J. Wallenius (2009). "Micro and Macro Elasticities in a Life Cycle Model with Taxes." *Journal of Economic Theory*, forthcoming.

[47] Romer, C.D. and D. H. Romer (2004). "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review* 1055-1084.

[48] Schreft, S. (1990). "Credit Controls: 1980." *Federal Reserve Bank of Richmond Economic Review* 1990, 6, 25-55.

[49] Schwarz, G.E. (1978). "Estimating the Dimension of a Model." *Annals of Statistics* 6, 461–464.

[50] Shephard, N. (2008). "Stochastic Volatility." The New Palgrave Dictionary of Economics.

[51] Sims, C.A. and T. Zha (2006). "Were There Regime Switches in U.S. Monetary Policy?" American Economic Review vol 96, 54-81.

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

[53] Stock, J.H. and M.W. Watson (2002). "Has the Business Cycle Changed, and Why?" *NBER Macroeconomics Annual* 17, 159-218.

[54] Stock, J.H. and M.W. Watson (2007). "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking* 39, 3-33.

[55] Uhlig, H. (1997). "Bayesian Vector Autoregressions with Stochastic Volatility." *Econometrica* 65, 59-73.