The Time-Varying Volatility of Macroeconomic Fluctuations

BY ALEJANDRO JUSTINIANO AND GIORGIO E. PRIMICERI*

We investigate the sources of the important shifts in the volatility of US macroeconomic variables in the postwar period. To this end, we propose the estimation of DSGE models allowing for time variation in the volatility of the structural innovations. We apply our estimation strategy to a large-scale model of the business cycle and find that shocks specific to the equilibrium condition of investment account for most of the sharp decline in volatility of the last two decades. (JEL C51, E32)

It has been well documented that the volatility of output, inflation, and several other macroeconomic variables of the US economy has exhibited a very high degree of time variation over the last 50 years (see, for instance, James H. Stock and Mark W. Watson 2003, or Christopher A. Sims and Tao Zha 2006). Perhaps the most notorious episode of a substantial volatility shift in recent US economic history is the “Great Moderation,” which corresponds to the sharp decline in the standard deviation of GDP as well as other macroeconomic and financial variables since the mid-1980s. While significant efforts have been devoted to determine the timing of the Great Moderation (see, among others, Chang-Jin Kim and Charles R. Nelson 1999; Margaret M. McConnell and Gabriel Perez-Quiros 2000; Marcelle Chauvet and Simon Potter 2001; Stock and Watson 2002; Ana Maria Herrera and Elena Pesavento 2005), there is still substantial disagreement about the origin of this common decline in volatility (see Stock and Watson 2003 for an overview).

In this paper, we investigate this issue by estimating a DSGE model in which the variance of the structural innovations is allowed to change over time. First, we describe an algorithm that allows for simultaneous inference on both the model’s parameters and the stochastic volatilities. Then, we apply our estimation strategy to a large-scale business cycle model of the US economy, along the lines of Frank Smets and Rafael Wouters (2003) and Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans (2005). The model exhibits a number of real and nominal frictions, and various shocks with a structural interpretation. The novelty of our setup is that all of these shocks have variances that can fluctuate over time.

We believe that this is an interesting innovation because it enables us to identify the sources of the changes in the volatility of the main macro variables during the postwar period. Thereafter,

* Justiniano: Economic Research, Federal Reserve Bank of Chicago, 230 S. LaSalle St., Chicago, IL 60604 (e-mail: ajustiniano@frbchi.org); Primiceri: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, NBER, and CEPR (e-mail: g-primiceri@northwestern.edu). We would like to thank Jinill Kim, David Lopez-Salido, Tommaso Monacelli, Ernst Schaumburg, Jim Stock, four anonymous referees, our discussants Jean Boivin, Hans Genberg, Thomas Lubik, Simon Potter, Juan Rubio-Ramirez, Kevin Salyer, Andrew Scott, and Noah Williams, and seminar participants at several universities and research institutions for comments. We are also grateful to Riccardo Di Cecio for providing some of the investment deflators data. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of Chicago or any other person associated with the Federal Reserve System.
we are able to shed light on the nature of the underlying disturbances responsible for changes in the variability of the US business cycle and, in particular, the Great Moderation.

The main conclusions we reach in this study are as follows. First, the exogenous structural disturbances hitting the US economy display substantial stochastic volatility. Nonetheless, the degree of time variation in variances differs considerably across shocks, being more pronounced for technology disturbances and, particularly, monetary policy shocks. Consequently, while stochastic volatility is present in all of the model's observed endogenous variables, different series exhibit contrasting patterns of fluctuation in their variances. Hence, it is not surprising that our approach delivers a substantially better fit of the data, compared not only to a homoskedastic model, but also to a specification that allows for a single jump in the volatilities.

Second, the decline in the volatility of output, investment, hours, and consumption in the early 1980s is largely driven by a change in the variance of the shock specific to the equilibrium condition of investment. This result is robust to various modifications of the baseline model, including those in which we allow for a jump in all model parameters and, in particular, a switch from passive to active monetary policy.

Broadly speaking, these shocks to the equilibrium condition of investment capture innovations specific to the return on capital or to the marginal efficiency of the investment technology. We suggest two particular interpretations of these disturbances, which we believe are useful to shed light on the Great Moderation. First, in our model these disturbances correspond either to investment-specific technological shocks or, equivalently, to shocks to the relative price of investment in terms of consumption goods. Our model is not rich enough, however, to exclude some alternative interpretations. Therefore, motivated by Ben S. Bernanke and Mark Gertler (1989) and Bernanke, Mark Gertler, and Simon Gilchrist (1999), we suggest a second, broader view of these disturbances as proxying for unmodeled investment financial frictions.

We rely on evidence outside our DSGE model to argue for the plausibility of both interpretations. In particular, in line with the first view, we document a decline in the standard deviation of the relative price of investment, as well as of investment-specific technology shocks when the latter are identified, as in Jonas D. Fisher (2006). Regarding the second view, and consistent with a recent line of research, we note that financial frictions became less binding at the beginning of the 1980s, following market deregulation and financial innovations that allowed firms and households increased access to credit markets (Gertler and Cara Lown 1999; Karen E. Dynan, Douglas W. Elmendorf, and Daniel E. Sichel 2006; Jeffrey R. Campbell and Zvi Hercowitz 2006).

More generally, our results suggest that efforts to understand the Great Moderation should focus on the dramatic changes in the investment equilibrium condition that occurred in the early 1980s.

From a methodological standpoint, this paper is related to the statistics literature on stochastic volatility models (for an overview, see Sangjoon Kim, Neil Shephard, and Siddhartha Chib 1998) and on partial non-Gaussian state-space models (Shephard 1994). Drawing from this literature, we develop an efficient algorithm, based on Bayesian Markov chain Monte Carlo (MCMC) methods, for the numerical evaluation of the posterior of the parameters of interest. Methodologically, the paper closest to ours is the recent contribution of Jesús Fernández-Villaverde and Juan Rubio-Ramírez (2007a) in nonlinear DSGE estimation. As discussed in Section I, their approach and ours are complementary as we analyze different models, using different solution methods and estimation algorithms.1

Regarding the application of these techniques, this paper is related to the large literature using estimated micro-founded models to understand the main sources of US business cycle

1 A related analysis is Jean-Philippe Laforte (2005), who models variances in a small-scale macro model as a Markov switching process.
fluctuations (see, for instance, Julio J. Rotemberg and Michael Woodford 1997; Peter N. Ireland 2004; David Altig et al. 2005; Christiano et al. 2005; Smets and Wouters 2007). As mentioned, however, we depart from previous work in this area by allowing for time variation in the volatility of the structural disturbances. In this respect, the paper closest to ours is perhaps Jean Boivin and Marc Giannoni (2006), although they allow only a one-time shift in parameters and variances, and their estimation is tailored to the analysis of changes in the effectiveness of monetary policy. In addition, the model of Boivin and Giannoni (2006) abstracts from investment dynamics, which instead turn out to be crucial in our study of the Great Moderation.

Our approach is also linked to the fairly large literature dealing with the estimation of vector autoregressions with heteroskedastic shocks (see, for example, Bernanke and Ilian Mihov 1998; Timothy Cogley and Thomas J. Sargent 2005; Primiceri 2005; Sims and Zha 2006; or Fabio Canova, Luca Gambetti, and Evi Pappa 2007). In contrast to this strand of work, one advantage of our analysis is that a fully fledged model provides an easier interpretation for the structural disturbances hitting the economy.

The paper is organized as follows. Section I presents the class of models that we deal with. Sections II and III illustrate our application to the model of the US business cycle and sketch the estimation technique. Sections IV and V discuss the estimation results and address the causes of the Great Moderation. Section VI provides two interpretations of our results. Section VII conducts a number of robustness checks and compares the fit of our baseline stochastic volatility model relative to alternative specifications, including some that allow for indeterminacy in the model solution. Section VIII concludes with some final remarks.

I. Stochastic Volatility in DSGE Models

Consider the general class of models summarized by the following system of equations:

\[
E_t[f(y_{t+1}, y_t, y_{t-1}, \eta_t, \theta)] = 0,
\]

where \(y_t\) is a \(k\times1\) vector of endogenous variables, \(\eta_t\) is an \(n\times1\) vector of exogenous disturbances, \(\theta\) is a \(p\times1\) vector of structural parameters, and \(E_t\) denotes the mathematical expectation operator, conditional on the information available at time \(t\). For example, (1) can be thought of as a collection of constraints and first-order conditions derived from a micro-founded model of consumer and/or firm behavior. The novelty of this paper is that the standard deviations of the elements of \(\eta_t\) are allowed to change over time. In particular, we make the assumption that

\[
\log \eta_t = \hat{\eta}_t = \Sigma_t \epsilon_t,
\]

\[
\epsilon_t \sim N(0, I_n),
\]

where \(N\) indicates the normal distribution, \(I_n\) denotes an \(n\times n\) identity matrix, and \(\Sigma_t\) is a diagonal matrix with the \(n\times1\) vector \(\sigma_t\) of time-varying standard deviations on the main diagonal. Following the stochastic volatility literature (see, for instance, Kim et al. 1998), we assume that each element of \(\sigma_t\) evolves (independently) according to the following stochastic processes:

\[
\log \sigma_{i,t} = (1 - \rho_{\sigma_i}) \log \sigma_i + \rho_{\sigma_i} \log \sigma_{i,t-1} + \nu_{i,t},
\]

\[
\nu_{i,t} \sim N(0, \omega_i^2) \quad i = 1, \ldots, n.
\]
Observe that modeling the logarithm of $\sigma_t$, as opposed to $\sigma_t$ itself, ensures that the standard deviation of the shocks remains positive at every point in time.

Our objective is to characterize the posterior distribution of the model structural parameters ($\theta$) and the time-varying volatility of the shocks ($[\sigma_t]_{t=1}^T$). Note that the model described by (1) is in general nonlinear and its solution must be approximated, as it does not have a closed-form expression. Our solution method is based on a log-linear approximation of (1) around the deterministic steady state. By deriving the log-linear approximation as a function of the heteroskedastic shock $\eta_t$ (as opposed to $\epsilon_t$), we are able to retain the minimal set of higher-order terms necessary for heteroskedasticity to play a role in the solution, and yet we can still use standard packages to solve for the resulting linear system of rational expectations equations.

In a recent related paper, Fernández-Villaverde and Rubio-Ramírez (2007a) deal with this class of models, adopting a second-order approximation of the solution and the particle filter to evaluate the likelihood function. The two approaches are complementary. Our approach is only first-order accurate and neglects the role of nonlinearities in the model. In those cases where nonlinearities are important, this approach may overstate the case for heteroskedasticity in the shocks. On the other hand, our method has the advantage of considerably simplifying the model solution and inference, allowing us to estimate a rich model of the business cycle with nominal and real rigidities which is substantially harder to handle with the methodology of (Fernández-Villaverde and Rubio-Ramírez 2007a). In this way, we can also evaluate the importance of some prominent explanations put forth for the Great Moderation, in particular the role of monetary policy.

II. The Model

We estimate a relatively large-scale model of the US business cycle, which has been shown to fit the data fairly well (Marco Del Negro et al. 2007). The model is based on work by Smets and Wouters (2003) and Christiano et al. (2005), to which the reader is referred for additional details. Our brief illustration of the model follows closely Del Negro et al. (2007).

A. Final Goods Producers

At every point in time $t$, perfectly competitive firms produce the final consumption good $Y_t$, using the intermediate goods $Y_i(i), i \in [0, 1]$ and the production technology

$$Y_t = \left[ \int_0^1 Y_i(i)^{1/\lambda_{p,t}} \, di \right]^{1+\lambda_{p,t}}.$$ 

$\lambda_{p,t}$ follows the exogenous stochastic process

$$\log \lambda_{p,t} = (1 - \rho_p) \log \lambda_{p,t} + \rho_p \log \lambda_{p,t-1} + \sigma_{p,t} \epsilon_{p,t},$$

where $\epsilon_{p,t}$ is $i.i.d. N(0, 1)$ and $\sigma_{p,t}$ evolves as in (2). Unless otherwise noticed, this property of a time-varying variance applies to all shocks in the model. Profit maximization and zero profit condition for the final goods producers imply the following relation between the price of the final good ($P_t$) and the prices of the intermediate goods ($P_i(i), i \in [0, 1]$):

$$P_t = \left[ \int_0^1 P_i(i)^{-1/\lambda_{p,t}} \, di \right]^{-\lambda_{p,t}}.$$
and the following demand function for the intermediate good $i$:

$$Y_t(i) = \left( \frac{P_t(i)}{P_t} \right)^{-(1+\lambda_{p,t})/\lambda_{p,t}} Y_t.$$

As a consequence, $\lambda_{p,t}$ will also correspond to the price markup over marginal costs for the firms producing intermediate goods.

**B. Intermediate Goods Producers**

A monopolistic firm produces the intermediate good $i$ using the following production function:

$$Y_t(i) = \max\{A_t^{1-\alpha}K_t(i)^{\alpha}L_t(i)^{1-\alpha} - A_tF; 0\},$$

where, as usual, $K_t(i)$ and $L_t(i)$ denote, respectively, the capital and labor input for the production of good $i$, $F$ represents a fixed cost of production, and $A_t$ is an exogenous stochastic process capturing the effects of technology. In particular, we model $A_t$ as a unit root process, with a growth rate $(z_t \equiv \log A_t/A_{t-1})$ that follows the exogenous process

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

As in Guillermo Calvo (1983), a fraction $\xi_p$ of firms cannot reoptimize their prices and, as we allow for indexation, set their prices following the rule

$$P_t(i) = P_{t-1}(i) \pi_{t-1}^{1-\nu} \pi^{1-\nu},$$

where $\pi_t$ is defined as $P_t/P_{t-1}$ and $\pi$ denotes the steady-state value of $\pi_t$. Subject to the usual cost minimization condition, reoptimizing firms choose their price ($\bar{P}_t(i)$) by maximizing the present value of future profits,

$$E_t \sum_{s=0}^{\infty} \xi^{s}_{p} \beta^{s} \lambda_{t+s} \left\{ \left[ \bar{P}_t(i) \Pi_{j=0}^{s} \pi_{t+s+j}^{1-\nu} \right] Y_{t+s}(i) - \left[ W_{t+s}L_{t+s}(i) + R_{t+s}^kK_{t+s}(i) \right] \right\},$$

where $\lambda_{t+s}$ is the marginal utility of consumption, and $W_t$ and $R_t^k$ denote, respectively, the wage and the rental cost of capital.

**C. Households**

Firms are owned by a continuum of households, indexed by $j \in [0, 1]$. As in Christopher J. Erceg, Dale W. Henderson, and Andrew T. Levin (2000), while each household is a monopolistic supplier of specialized labor $(L_t(j))$, a number of “employment agencies” combine households’ specialized labor into labor services available to the intermediate firms:

$$L_t = \left[ \int_0^1 L_t(j)^{1/(1+\lambda_u)} \, dj \right]^{1+\lambda_u}.$$
Profit maximization and a zero profit condition for the perfectly competitive employment agencies imply the following relation between the wage paid by the intermediate firms and the wage received by the supplier of specialized labor $L_s(j)$:

$$W_t = \left[ \int_0^1 W_t(j)^{-1/\lambda_c} \, dj \right]^{-\lambda_c},$$

and the following labor demand function for labor type $j$:

$$L_t(j) = \left( \frac{W_t(j)}{W_t} \right)^{-\lambda_c} L_t,$$

Each household maximizes the utility function\(^2\)

$$E_0 \sum_{t=0}^{\infty} b^t \beta^t \left[ \log(C_{t+x}(j)) - hC_{t+x-1}(j) \right] - \varphi_t \left( \frac{L_{t+x}(j)^{1+v}}{1+v} \right),$$

where $C_t(j)$ is consumption, $h$ is the “degree” of internal habit formation, $\varphi_t$ is a preference shock that affects the marginal disutility of labor, and $b_t$ is a “discount factor” shock affecting both the marginal utility of consumption and the marginal disutility of labor. These two shocks follow the stochastic processes

$$\log b_t = \rho_b \log b_{t-1} + \sigma_{b_t} \varepsilon_{b_t},$$
$$\log \varphi_t = (1 - \rho_\varphi) \log \varphi_{t-1} + \sigma_{\varphi_t} \varepsilon_{\varphi_t}.$$

The household budget constraint is given by

$$P_{t+x} C_{t+x}(j) + P_{t+x} I_{t+x}(j) + B_{t+x}(j) \leq R_{t+x-1} B_{t+x-1}(j) + Q_{t+x-1}(j) + \Pi_{t+x}$$
$$+ W_{t+x}(j)L_{t+x}(j) + R_{t+x}^k(j) u_{t+x}(j) \tilde{K}_{t+x-1}(j) - P_{t+x} a(u_{t+x}(j)) \tilde{K}_{t+x-1}(j),$$

where $I_t(j)$ is investment, $B_t(j)$ denotes holding of government bonds, $R_t$ is the gross nominal interest rate, $Q_t(j)$ is the net cash flow from participating in state contingent securities, and $\Pi_t$ is the per capita profit that households get from owning the firms. Households own capital and choose the capital utilization rate that transforms physical capital ($\tilde{K}_t(j)$) into effective capital

$$K_t(j) = u_t(j) \tilde{K}_{t-1}(j),$$

which is rented to firms at the rate $R^k_t(j)$. The cost of capital utilization is $a(u_{t+x}(j))$ per unit of physical capital. Following Altig et al. (2005), we assume that $u_t = 1$ and $a(u_t) = 0$ in steady state. In our partially nonlinear approximation of the model solution, only the curvature of the function $a$ in steady state needs to be specified, $\chi = a''(1)/a'(1)$. The usual physical capital accumulation equation is described by

$$\tilde{K}_t(j) = (1-\delta)\tilde{K}_{t-1}(j) + \mu_t \left( 1 - S \left( \frac{I_t(j)}{I_{t-1}(j)} \right) \right) I_t(j),$$

\(^2\) We assume a cashless limit economy as described in Woodford (2003).
where \( \delta \) denotes the depreciation rate and, as in Christian et al. (2005) and Altig et al. (2005), the function \( S \) captures the presence of adjustment costs in investment, with \( S' = 0 \) and \( S'' > 0 \). David Lucca (2005) shows that this formulation of the adjustment cost function is equivalent (up to a first-order approximation of the model) to a generalization of a time to build assumption. Following Jeremy Greenwood, Hercowitz, and Per Krusell (1997) and Fisher (2006), \( \mu_t \) can be interpreted as an investment-specific technology shock (or a shock to the production technology of capital goods), as well as a shock to the relative price of investment in terms of consumption goods. More generally, \( \mu_t \) can be thought of as a disturbance to the equilibrium condition of investment, given that it affects the return on capital. For space considerations (and somewhat abusing notation), we label this disturbance the “investment shock.” While we will return to the interpretation of this shock in Section VI, here we just assume that it evolves following the exogenous process

\[
\log \mu_t = \rho_{\mu} \log \mu_{t-1} + \sigma_{\mu,t} \epsilon_{\mu,t}.
\]

Following Erceg et al. (2000), in every period, a fraction \( \xi_w \) of households cannot reoptimize their wages and, therefore, set their wages following the indexation rule

\[
W_t(j) = W_{t-1}(j)(\pi_{t-1}e^\epsilon\omega)^{1-\omega}(\pi e^\epsilon)^{1-\omega}.
\]

The remaining fraction of reoptimizing households set their wages by maximizing

\[
E_t \sum_{s=0}^{\infty} \xi_w^s \beta^s b_{t+s} \left\{-\varphi_{t+s}(j)^{1+\gamma} \frac{L_{t+s}(j)^{1+\gamma}}{1 + \nu} \right\},
\]

subject to the labor demand function.

D. Monetary and Government Policies

Monetary policy sets short-term nominal interest rates following a Taylor type rule. In particular, the rule allows for interest rate smoothing and interest rate responses to deviations of inflation from the steady state and deviations of output from trend level:

\[
R_t = \left( \frac{R_{t-1}}{R} \right)^{\rho_R} \left( \frac{\pi_t}{\pi} \right)^{\phi_R} \left( \frac{Y_t/A_t}{Y/A} \right)^{\phi_y} \left[ 1 - \rho_R \right]^{1-\rho_R} e^{\epsilon_{R,t} + \epsilon_{R,t}},
\]

where \( R \) is the steady state for the gross nominal interest rate and \( \epsilon_{R,t} \) is a monetary policy shock. We also consider, and later discuss, an alternative specification of the policy rule, in which the monetary authority responds to the output gap, defined as the ratio between output and the level of output that would prevail in a flexible price and wage economy (see, for instance, Woodford 2003 or Levin et al. 2005).

Fiscal policy is assumed to be fully Ricardian, and public spending is given by

\[
G_t = \left( 1 - \frac{1}{g_t} \right) Y_t,
\]

where \( g_t \) is an exogenous disturbance following the stochastic process

\[
\log g_t = (1 - \rho_g) \log g + \rho_g \log g_{t-1} + \sigma_{g,t} \epsilon_{g,t}.
\]
E. Market Clearing

The resource constraint is given by

$$C_t + I_t + G_t + a(u_t)K_{t-1} = Y_t.$$ 

F. Steady State and Model Solution

Since the technology process $A_t$ is assumed to have a unit root, consumption, investment, capital, real wages, and output evolve along a stochastic growth path. Once the model is rewritten in terms of detrended variables, we can compute the nonstochastic steady state to approximate the solution around it. This procedure delivers a partial nonlinear state space model of the kind described in Shephard (1994).

We conclude the discussion of the model by specifying the vector of observables, completing the state space representation of our model:

$$[\Delta \log Y_t, \Delta \log C_t, \Delta \log I_t, \log L_t, \Delta \log \frac{W_t}{P_t}, \pi_t, R_t],$$

where $\Delta \log X_t$ denotes $\log X_t - \log X_{t-1}$.

III. Inference

A. The Data

We estimate the model using seven series of US quarterly data, as in Levin et al. (2005) and Del Negro et al. (2007). These series correspond to the vector of observable variables of our model, reported in Section III F. The sample for our dataset spans from 1954:III up to 2004:IV. All data are extracted from the Haver Analytics database (series mnemonics in parentheses). Following Del Negro et al. (2007), we construct real GDP by dividing the nominal series (GDP) by population (LF and LH) and the GDP Deflator (JGDP). Real series for consumption and investment are obtained in the same manner, although consumption corresponds only to personal consumption expenditures of nondurables (CN) and services (CS), while investment is the sum of personal consumption expenditures of durables (CD) and gross private domestic investment (I). Real wages correspond to nominal compensation per hour in the nonfarm business sector (LXNFC) divided by the GDP deflator. Our measure of labor is given by the log of hours of all persons in the nonfarm business sector (HNFBN) divided by population. Inflation is measured as the quarterly log difference in the GDP deflator, while for nominal interest rates we use the effective federal funds rate. Unlike Smets and Wouters (2003), Levin et al. (2005), or Boivin and Giannoni (2006), we do not demean or detrend any series.

B. Bayesian Inference

Estimation of these models by pure maximum likelihood is extremely challenging. Following a growing recent literature, we adopt a Bayesian approach to inference, integrating the sample information with weakly informative priors, which summarize additional information about the parameters (see, for instance, Smets and Wouters 2003, Levin et al. 2005, or Del Negro et al. 2007). One advantage of this approach is that it also ameliorates common numerical problems
related to both the flatness of the likelihood function in some regions of the parameter space and the existence of multiple local maxima.  

MCMC methods are used to characterize the posterior distribution of the model’s structural parameters ($\theta$), the time-varying volatility of the shocks ($\{\sigma_i\}_{i=1}^{T}$), and the coefficients of the volatility processes ($\{\rho_s, \omega^2\}$). Bayesian methods deal efficiently with the high dimension of the parameter space and the nonlinearities of the model, by splitting the original estimation problem into smaller and simpler blocks. In particular, the MCMC algorithm for this paper is carried out in three steps. First, a Metropolis step is used to draw from the posterior of the structural coefficients $\theta$. Drawing the sequence of time-varying volatilities $\sigma^T$ (conditional on $\theta, \alpha, \rho_s$, and $\omega^2$) is instead more involved and relies mostly on the method presented in Kim et al. (1998). It consists of transforming a nonlinear and non-Gaussian state space form into a linear and approximately Gaussian one, which allows the use of simulation smoothers such as those of Christopher K. Carter and Robert J. Kohn (1994) or James Durbin and Siem J. Koopman (2002). Simulating the conditional posterior of $[\sigma, \rho_s, \omega^2]$ is standard, since it is the product of independent normal-inverse-Gamma distributions. Further details of the estimation are relegated to Appendix A, while Appendix B discusses checks for the convergence of the algorithm.

C. Priors

As it is customary when taking DSGE models to the data, we fix a small number of the model parameters to values that are very common in the existing literature. In particular, we set the steady-state share of capital income ($\alpha$) to 0.3, the quarterly depreciation rate of capital ($\delta$) to 0.025, and the steady-state government spending to GDP ratio to 0.22, which corresponds to the average share of government spending in total GDP ($G_t/Y_t$) in our sample. Moreover, we set the autocorrelation of the mark-up shock ($\rho_p$) to zero. Two reasons motivate this choice: first, this parameter is weakly identified from the price indexation coefficient; second, shutting down this persistence mechanism helps the identification of indeterminacy in Section VII. Finally, we set all the autoregressive coefficients of the log-volatilities, $\rho_s$’s, to 1. The assumption that the volatilities follow geometric random walk processes serves two main purposes: on the one hand, it helps to reduce the number of free parameters of the model; on the other hand, it allows us to focus on lower frequency changes in the volatilities of the endogenous variables of our macroeconomic model.

The first three columns of Table 1 report our priors for the remaining parameters of the model. While most of these priors are relatively disperse and reflect previous results in the literature, a few of them deserve some further discussion. First, our baseline prior distribution assigns zero probability to the indeterminacy region of the parameter space, although we will relax this assumption in Section VII. Second, for all but one persistence parameters we use a Beta prior, with mean 0.6 and standard deviation 0.2. The only exception is neutral technology which includes a unit root already, and for this reason the prior for the autocorrelation of the growth rate of neutral technology, $\rho_z$, is centered at 0.4 instead.

Finally, following Del Negro et al. (2007), the priors for the standard deviations of the shocks are fairly disperse and chosen in order to generate realistic volatilities for the endogenous variables. These priors only enter the specification of the model without stochastic volatility that we estimate simply for comparison.  

---

3 See the survey article by Sungbae An and Frank Schorfheide (2007) for a detailed discussion of these issues.

4 In an earlier version, Justiniano and Primiceri (2005), we adopted, instead, a prior favoring high autocorrelation in the mark-up shock. Our results are robust to this alternative specification.

5 To be precise, the mean and the variance of these priors are also used to initialize the filter for the stochastic volatilities. Alternative values for the initialization leave our results unchanged.
The priors on the variance \( s^2 \) of the innovations to the log-volatility processes deserve some comment as well, as these coefficients are new in the DSGE literature. We chose an inverse-Gamma prior with mean equal to 0.01 for several reasons. First, assuming that the log-volatilities behave as random walks, this parameterization implies an average variation of about 25 percent over our sample of 40 years. We regard this as a conservative degree of time variation. Second, in the context of time-varying vector autoregressions, Primiceri (2005) has tested several prior specifications and concluded that this value attains the highest marginal likelihood. Nonetheless, we have assessed the sensitivity of our estimates to alternative specifications of the prior (especially for the variance of the innovation to the log-volatilities) and found that these modifications had no important influence on the results.

IV. Estimation Results

A. Parameter Estimates

The last three columns of Table I summarize the posterior distribution of the model coefficients, reporting posterior medians, standard deviations, and fifth and ninety-fifth percentiles computed with the draws of our posterior simulator. All coefficients estimates are fairly tight and seem for the most part in line with those reported in Levin et al. (2005) and Del Negro et al. (2007).

One important exception is the wage stickiness parameter \( \xi_w \), which is lower than previous estimates reported in the literature dealing with inference in DSGE models. In view of the welfare implications of wage rigidity (see, for instance, Levin et al. 2005), this variation in estimates may be important, although we do not explore this issue in the current paper. The inferred median of the Calvo price stickiness parameter \( \xi_p \) is approximately equal to 0.9, which is in line with the value found in Smets and Wouters (2003). This number is higher than recent estimates in micro studies (see, for instance, Mark Bils and Peter J. Klenow 2004), although the presence of indexation mechanisms (which assures that prices are actually changed in every period) makes the results potentially more consistent with the micro evidence on the high frequency of price changes.

For comparison, Table 1 also reports posterior medians, standard deviations, and fifth and ninety-fifth percentiles of a model estimated with time-invariant volatilities. Notice that most of the coefficient estimates are similar to the stochastic volatility model, with the exception perhaps of the parameter for the adjustment costs in investment \( S^0 \) which is slightly lower relative to the specification with time-varying volatility.

Finally, Table 2 shows that the coefficient estimates of the stochastic volatility model are quite robust to an alternative specification of the monetary policy rule in which the monetary authority responds to the output gap. From now on, for space considerations, we report only estimates for our baseline specification in which the policy authority responds to deviations of output from the neutral technology trend. Our choice is motivated by the better fit of this specification according to the marginal data density (which is roughly 20 log points higher compared to the gap rule). While a detailed discussion of model fit is presented in Section VIIA, it is important to stress that none of our results below depends on this choice.

---

Note that in a previous version of the paper (Justiniano and Primiceri 2005) we obtained a lower estimate of \( \xi_p \), as we allowed for autocorrelation in the mark-up shock. As already mentioned, our results are unaffected by this modification.
Table 1—Prior Densities and Posterior Estimates for the Time-Invariant Model and Baseline Model with Stochastic Volatility

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Description</th>
<th>Prior Density</th>
<th>Mean Std</th>
<th>Posterior time invariant</th>
<th>Posterior with stochastic volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Density</td>
<td>Mean Std</td>
<td>[5  , 95 ]</td>
<td>[5  , 95 ]</td>
</tr>
<tr>
<td>$\lambda_p$</td>
<td>Mark-up price elasticity</td>
<td>B</td>
<td>0.50 0.15</td>
<td>0.84 0.04 [0.77 0.91 ]</td>
<td>0.83 0.05 [0.75 0.91 ]</td>
</tr>
<tr>
<td>$\lambda_p$</td>
<td>Mark-up price elasticity</td>
<td>B</td>
<td>0.50 0.15</td>
<td>0.09 0.03 [0.05 0.14 ]</td>
<td>0.08 0.03 [0.04 0.13 ]</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>SS technology growth rate</td>
<td>N</td>
<td>0.50 0.03</td>
<td>0.43 0.02 [0.39 0.47 ]</td>
<td>0.43 0.02 [0.39 0.47 ]</td>
</tr>
<tr>
<td>$h$</td>
<td>Consumption habit</td>
<td>B</td>
<td>0.50 0.10</td>
<td>0.81 0.03 [0.76 0.86 ]</td>
<td>0.84 0.03 [0.79 0.88 ]</td>
</tr>
<tr>
<td>$\lambda_g$</td>
<td>SS mark-up goods prices</td>
<td>N</td>
<td>0.15 0.05</td>
<td>0.22 0.04 [0.16 0.28 ]</td>
<td>0.23 0.04 [0.17 0.29 ]</td>
</tr>
<tr>
<td>$\lambda_s$</td>
<td>SS mark-up wages</td>
<td>N</td>
<td>0.15 0.05</td>
<td>0.17 0.04 [0.10 0.25 ]</td>
<td>0.16 0.04 [0.09 0.24 ]</td>
</tr>
<tr>
<td>$\lambda_s$</td>
<td>SS mark-up wages</td>
<td>N</td>
<td>0.10 0.50</td>
<td>0.56 0.10 [0.40 0.71 ]</td>
<td>0.55 0.10 [0.39 0.71 ]</td>
</tr>
<tr>
<td>$\pi$</td>
<td>SS quarterly inflation</td>
<td>N</td>
<td>0.50 0.10</td>
<td>1.03 0.07 [0.91 1.15 ]</td>
<td>1.03 0.07 [0.90 1.15 ]</td>
</tr>
<tr>
<td>$r$</td>
<td>Real interest rate</td>
<td>N</td>
<td>0.50 0.10</td>
<td>1.03 0.07 [0.91 1.15 ]</td>
<td>1.03 0.07 [0.90 1.15 ]</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Inverse Frisch labor</td>
<td>G</td>
<td>2.00 0.75</td>
<td>1.59 0.35 [0.98 2.12 ]</td>
<td>1.59 0.48 [0.94 2.47 ]</td>
</tr>
<tr>
<td>$\xi_c$</td>
<td>Calvo prices</td>
<td>B</td>
<td>0.75 0.10</td>
<td>0.90 0.01 [0.88 0.92 ]</td>
<td>0.91 0.01 [0.89 0.93 ]</td>
</tr>
<tr>
<td>$\xi_c$</td>
<td>Calvo wages</td>
<td>B</td>
<td>0.75 0.10</td>
<td>0.61 0.05 [0.52 0.69 ]</td>
<td>0.66 0.05 [0.57 0.74 ]</td>
</tr>
<tr>
<td>$\chi$</td>
<td>Elasticity capital utilization costs</td>
<td>G</td>
<td>5.00 1.00</td>
<td>6.90 1.10 [5.25 8.91 ]</td>
<td>7.13 1.09 [5.45 9.02 ]</td>
</tr>
<tr>
<td>$\sigma^*$</td>
<td>Investment adjustment costs</td>
<td>G</td>
<td>3.00 0.75</td>
<td>2.72 0.48 [1.99 3.61 ]</td>
<td>3.30 0.57 [2.42 4.29 ]</td>
</tr>
<tr>
<td>$\phi_r$</td>
<td>Taylor rule inflation</td>
<td>N</td>
<td>1.70 0.30</td>
<td>1.92 0.13 [1.71 2.15 ]</td>
<td>1.90 0.14 [1.67 2.14 ]</td>
</tr>
<tr>
<td>$\phi_r$</td>
<td>Taylor rule output</td>
<td>G</td>
<td>0.13 0.10</td>
<td>0.10 0.02 [0.07 0.13 ]</td>
<td>0.08 0.02 [0.06 0.11 ]</td>
</tr>
<tr>
<td>$\rho_p$</td>
<td>Taylor rule smoothing</td>
<td>B</td>
<td>0.60 0.20</td>
<td>0.81 0.02 [0.78 0.84 ]</td>
<td>0.84 0.02 [0.80 0.87 ]</td>
</tr>
<tr>
<td>$\rho_p$</td>
<td>Taylor rule smoothing</td>
<td>B</td>
<td>0.60 0.20</td>
<td>0.28 0.06 [0.18 0.39 ]</td>
<td>0.32 0.06 [0.21 0.43 ]</td>
</tr>
<tr>
<td>$\rho_p$</td>
<td>Technology growth</td>
<td>B</td>
<td>0.60 0.20</td>
<td>0.98 0.00 [0.98 0.98 ]</td>
<td>0.98 0.00 [0.98 0.98 ]</td>
</tr>
<tr>
<td>$\rho_p$</td>
<td>Technology growth</td>
<td>B</td>
<td>0.60 0.20</td>
<td>0.87 0.03 [0.82 0.92 ]</td>
<td>0.92 0.02 [0.88 0.96 ]</td>
</tr>
<tr>
<td>$\rho_p$</td>
<td>Technology growth</td>
<td>B</td>
<td>0.60 0.20</td>
<td>0.90 0.03 [0.86 0.95 ]</td>
<td>0.89 0.04 [0.81 0.95 ]</td>
</tr>
<tr>
<td>$\rho_p$</td>
<td>Technology growth</td>
<td>B</td>
<td>0.60 0.20</td>
<td>0.84 0.05 [0.74 0.89 ]</td>
<td>0.83 0.05 [0.73 0.89 ]</td>
</tr>
<tr>
<td>$\sigma_i$</td>
<td>Monetary policy</td>
<td>I</td>
<td>0.15 0.15</td>
<td>0.25 0.01 [0.23 0.28 ]</td>
<td></td>
</tr>
<tr>
<td>$\sigma_i$</td>
<td>Technology growth</td>
<td>I</td>
<td>2.00 2.00</td>
<td>1.10 0.06 [1.01 1.20 ]</td>
<td></td>
</tr>
<tr>
<td>$\sigma_i$</td>
<td>Government spending</td>
<td>I</td>
<td>1.00 2.00</td>
<td>0.55 0.03 [0.51 0.61 ]</td>
<td></td>
</tr>
<tr>
<td>$\sigma_i$</td>
<td>Investment-specific</td>
<td>I</td>
<td>2.00 2.00</td>
<td>5.46 0.70 [4.43 6.76 ]</td>
<td></td>
</tr>
<tr>
<td>$\sigma_i$</td>
<td>Mark-up</td>
<td>I</td>
<td>0.15 0.15</td>
<td>0.17 0.01 [0.15 0.18 ]</td>
<td></td>
</tr>
<tr>
<td>$\sigma_i$</td>
<td>Labor disutility</td>
<td>I</td>
<td>4.00 2.00</td>
<td>10.41 1.11 [8.96 12.77]</td>
<td></td>
</tr>
<tr>
<td>$\sigma_i$</td>
<td>Intertemporal preference</td>
<td>I</td>
<td>2.00 2.00</td>
<td>3.13 0.38 [2.67 3.98 ]</td>
<td></td>
</tr>
</tbody>
</table>

(log) Likelihood at median: −1,891.7, −1,675.97

Notes: Calibrated coefficients: capital share ($s_c$) at 0.3, depreciation rate ($\delta$) at 0.025, $g$ at 1/0.77 (which implies a SS government share of 0.22), and persistence of mark-up shocks ($\rho_p$) set at zero. Relative to the text, $\gamma$ corresponds to a quarterly growth rate in the estimation and is therefore multiplied by 100. Meanwhile, $\pi$ and $R$ are expressed as net rates and multiplied by 100 as well. Finally, the standard deviations of the innovations are also scaled by 100 for the estimation. All these changes are reflected in the specification of the priors.

$^a$N stands for Normal, B Beta, G Gamma, and I Inverted-Gamma1 distribution.

$^b$Median, standard deviations and posterior percentiles from four chains of 140,000 draws each from the Random Walk Metropolis algorithm initialized from dispersed starting values around the mode. We discard the initial 40,000 draws.

$^c$Median, standard deviations and posterior percentiles from the Random Walk Metropolis within Gibbs algorithm for the model with stochastic volatility. Results based on 3 chains of 150,000 draws each, where we discarded the initial 50,000 and retain 1 in every 5 simulations from the remaining 100,000 draws.

B. Volatility Estimates

Figure 1 presents the plots of the time-varying standard deviations for the seven shocks of our model. Notice that the degree of stochastic volatility varies substantially across disturbances.

The standard deviation of the price mark-up shock ($\lambda_p$, Figure 1E) is relatively stable, while for the two taste shocks ($\varphi$, and $b_t$, Figures 1F and 1G, respectively) their volatilities exhibit some
moderate fluctuations over the sample. In contrast, the remaining four shocks exhibit a very important degree of time variation.

The exogenous disturbance showing the largest degree of stochastic volatility is the monetary policy shock \( e_{MP}^t \), Figure 1A, for which the difference between the lowest and the highest levels of the standard deviation is roughly 500 percent. Observe that the “Volcker episode”7 is perfectly captured in our estimates, as is the reduction in the volatility of monetary policy shocks during the Greenspan period. In addition, notice that the volatility of monetary policy shocks is relatively high during the 1970s. This might be explained, at least in part, by the fact that our baseline estimation does not allow for breaks in the coefficients of the Taylor rule. This issue is quite controversial in the existing literature: among others, Richard Clarida, Jordi Galí, and Mark Gertler2000 have argued in favor of these changes, while Sims and Zha2006 have concluded against them. We will return to these ideas in Section VII, where we will allow for shifts in the policy rule coefficients and analyze the robustness of our results to this alternative empirical specification.

Monetary policy shocks are not the only ones exhibiting a clear pattern of fluctuation in their standard deviations. The standard deviation of technology shocks \( z_t \), Figure 1B) seems to decrease by approximately one-third in the second part of the sample. This is potentially consistent with the observed reduction in the volatility of GDP in the last two decades, an issue addressed in more detail in the next section. A similar pattern is observed for the volatilities of the government spending shock \( g_t \), Figure 1C) and, particularly, the investment shock \( \mu_t \), Figure 1D).

Table 2—Posterior Estimates for Stochastic Volatility Model with Taylor Rule Responding to the Output Gap

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Description</th>
<th>Posterior</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i_p )</td>
<td>Price indexation</td>
<td>0.86 0.04 [ 0.79 , 0.93 ]</td>
</tr>
<tr>
<td>( i_w )</td>
<td>Wage indexation</td>
<td>0.11 0.03 [ 0.06 , 0.17 ]</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>SS technology growth rate</td>
<td>0.44 0.03 [ 0.39 , 0.48 ]</td>
</tr>
<tr>
<td>( h )</td>
<td>Consumption habit</td>
<td>0.80 0.03 [ 0.76 , 0.84 ]</td>
</tr>
<tr>
<td>( \lambda_p )</td>
<td>SS mark-up goods prices</td>
<td>0.27 0.03 [ 0.22 , 0.33 ]</td>
</tr>
<tr>
<td>( \lambda_w )</td>
<td>SS mark-up wages</td>
<td>0.19 0.04 [ 0.13 , 0.26 ]</td>
</tr>
<tr>
<td>( L_{SS} (\log) )</td>
<td>SS labor</td>
<td>396.64 0.51 [395.83 , 397.50 ]</td>
</tr>
<tr>
<td>( \pi )</td>
<td>SS quarterly inflation</td>
<td>0.62 0.09 [ 0.46 , 0.77 ]</td>
</tr>
<tr>
<td>( r )</td>
<td>SS real interest rate</td>
<td>1.04 0.08 [ 0.93 , 1.18 ]</td>
</tr>
<tr>
<td>( \nu )</td>
<td>Inverse Frisch labor</td>
<td>1.61 0.37 [ 1.13 , 2.35 ]</td>
</tr>
<tr>
<td>( \xi_p )</td>
<td>Calvo prices</td>
<td>0.94 0.03 [ 0.89 , 0.97 ]</td>
</tr>
<tr>
<td>( \xi_w )</td>
<td>Calvo wages</td>
<td>0.40 0.07 [ 0.31 , 0.52 ]</td>
</tr>
<tr>
<td>( \chi )</td>
<td>Elasticity capital utilization costs</td>
<td>6.88 1.00 [ 5.41 , 8.70 ]</td>
</tr>
<tr>
<td>( S^* )</td>
<td>Investment adjustment costs</td>
<td>3.37 0.49 [ 2.57 , 4.19 ]</td>
</tr>
<tr>
<td>( \phi_r )</td>
<td>Taylor rule inflation</td>
<td>1.61 0.27 [ 1.27 , 2.13 ]</td>
</tr>
<tr>
<td>( \phi_y )</td>
<td>Taylor rule output gap</td>
<td>0.23 0.09 [ 0.06 , 0.36 ]</td>
</tr>
<tr>
<td>( \rho_k )</td>
<td>Taylor rule smoothing</td>
<td>0.85 0.02 [ 0.82 , 0.88 ]</td>
</tr>
<tr>
<td>( \rho_k )</td>
<td>Technology growth</td>
<td>0.26 0.06 [ 0.16 , 0.36 ]</td>
</tr>
<tr>
<td>( \rho_s )</td>
<td>Government spending</td>
<td>0.98 0.00 [ 0.98 , 0.98 ]</td>
</tr>
<tr>
<td>( \rho_{s} )</td>
<td>Investment-specific</td>
<td>0.90 0.02 [ 0.86 , 0.94 ]</td>
</tr>
<tr>
<td>( \rho_{v} )</td>
<td>Labor disutility</td>
<td>0.94 0.03 [ 0.89 , 0.97 ]</td>
</tr>
<tr>
<td>( \rho_{v} )</td>
<td>Intertemporal preference</td>
<td>0.85 0.04 [ 0.77 , 0.91 ]</td>
</tr>
</tbody>
</table>

\( (\log) \) Likelihood at median \(-1,698.12\)

Note: Priors and calibration identical to those reported for the stochastic volatility baseline model in Table 1.

7 The “Volcker episode” refers to the high volatility of interest rates in the 1979–1983 period, due to the monetary targeting regime initiated by Federal Reserve Chairman Paul Volcker in response to the dramatic rise in US inflation in the 1970s.
Figure 1. Stochastic Volatility of Each Shock in DSGE Model

Note: Median and 5–95 percentiles for the time-varying volatility of each disturbance computed with the draws generated in the estimation of the baseline stochastic volatility model.
One contribution of our analysis is the ability to quantify how the importance of various shocks has changed over time in generating economic fluctuations. To this end, we analyze the variance decomposition of each series, which will allow us later to address the causes of the Great Moderation. We perform the variance decomposition exercise in the following way: for every draw of the parameters and the volatilities of the exogenous disturbances, we construct the implied variances of the (endogenous) observable variables, using the state space representation of the model solution. Then, we recompute the variances of the observables, by sequentially setting to zero the volatility of all disturbances but one, for all time periods. In this way, we are able to investigate the contribution of each shock to the variability of the observable variables. Notice that, since the variances are changing over time, our variance decomposition is a time-varying “object” as well. Due to space considerations, we do not present the graphs for all variance decompositions. Instead, we provide a complete characterization of the variance decomposition for GDP, while for the remaining series we report only the time-varying share of the variance explained by selected shocks.

Figure 2 presents the evolution of the variance shares of GDP growth attributed to each exogenous disturbance. Consistent with Greenwood et al. (2000) and Fisher (2005), the most important shock in explaining the variability of GDP growth seems to be the investment shock (Figure 2D). Indeed, at least in the first part of the sample, this disturbance explains roughly half the variance of GDP growth. Note, however, that the importance of this shock for output fluctuations declines over time. On average, neutral technology shocks explain 20 percent of the variance of GDP growth (Figure 2B). Labor preference shocks play a lesser role in output fluctuations earlier in the sample, although their importance has increased in the last two decades (Figure 2F). Other shocks are less central for output.

For the remaining series, Figure 3 plots the time-varying variance shares explained by selected shocks. A major portion of the variance of consumption is explained by the intertemporal shock to the discount factor (Figure 3A). Although not crucial for output, monetary policy and markup shocks are each quite important for the volatility of interest rates (Figure 3B) and inflation (Figure 3C). Notice, however, that a major portion of the volatility of inflation is also explained by the labor disutility and, particularly, the investment shock (Figure 3D and 3E). Moreover, as one would expect, the labor disutility and the investment shocks explain most of the variability of hours (Figure 3G) and investment (Figure 3F), respectively, while the neutral technology shock accounts for about 20 percent of the variance of real wages (Figure 3H). 8

Figure 4 plots the model’s implied spectral variance decomposition for the level of output in deviations from the model’s common stochastic trend. We consider periodicities between 8 and 32 quarters. Figure 4 corroborates the evidence on the importance of the investment shock, suggesting that this disturbance is crucial in explaining output fluctuations at business cycle frequencies (Figure 4D). Note, however, that the labor disutility shock also plays a very important role in this case (Figure 4F).

V. The Great Moderation

In two very influential papers, Kim and Nelson (1999) and McConnell and Perez-Quiros (2000) drew attention to the dramatic reduction in the volatility of US GDP, which has characterized the last two decades relative to the pre-1980s period. This change seems to be more abrupt than gradual (Kim and Nelson 1999; Stock and Watson 2002) and the break date is estimated to correspond approximately to 1984. In our sample, the standard deviation of GDP growth

---

8The complete set of variance decomposition graphs is available upon request.
Figure 2. Variance Decomposition for Output Growth

Note: Median share and associated 5–95 percentiles for each disturbance computed with the draws generated in the estimation of the baseline stochastic volatility model.
Figure 3. Selected Variance Decomposition for Other Series (Series, Shock)

Note: Median share and associated 5–95 percentiles for each disturbance computed with the draws generated in the estimation of the baseline stochastic volatility model.
Figure 4. Spectral Decomposition for Output

Note: Spectral decomposition for periodicities between 8 and 32 quarters for the level of output in deviation from the model implied stochastic trend. Median share and associated 5–95 percentiles for each disturbance computed with the draws generated in the estimation of the baseline stochastic volatility model.
over the 1984–2004 period is almost one-half of the standard deviation computed over the 1955–1983 sample. As mentioned, the literature has labeled this phenomenon the Great Moderation.

A number of hypotheses have been put forward to account for this decline in volatility, and exhaustive reviews can be found in Olivier J. Blanchard and John Simon (2001) and Stock and Watson (2002, 2003). The main explanations of this phenomenon can be broadly collected as corresponding to simple good luck, improvements in the conduct of monetary policy under the Volcker and Greenspan chairmanships or, alternatively, improvements in inventory management.

With regards to the last explanation, however, several authors have raised doubts about the improved inventory management hypothesis initially put forward by McConnell and Perez-Quiros (2000) and James A. Kahn and McConnell (2002). Moreover, our model is already quite involved, and the computational demand of our estimation algorithm is substantial. For these reasons, our analysis of the Great Moderation completely abstracts from this channel. We refer, instead, to the literature for an exhaustive review of the reasons why the inventories hypothesis has been questioned, both from a theoretical (Louis J. Maccini and Adrian Pagan 2003; Kahn and Julia K. Thomas 2007) and empirical (Stock and Watson 2002; Spencer Krane 2002; Shaghil Ahmed, Levin, and Beth A. Wilson 2004; Herrera and Pesavento 2005; or Valerie A. Ramey and Daniel J. Vine 2006) perspective.

There is more disagreement, instead, on the role of improved monetary policy in the Great Moderation, with some authors advocating its significant role (Clarida et al. 2000; Bernanke 2004; Luca Benati 2004; or Boivin and Giannoni 2006), and others presenting evidence against this explanation (see, for instance, Stock and Watson 2002, 2003; Ahmed et al. 2004; or Sims and Zha 2006). We defer a careful examination of how changes in the systematic conduct of monetary policy may have affected volatility until Section VII.

Instead, the starting point for the analysis of the Great Moderation undertaken in this section is the robust finding of Stock and Watson (2002, 2003), who conclude that “this reduction in volatility is associated with an increase in the precision of forecasts of output growth” (Stock and Watson 2002, 42). Notice that our framework is a natural candidate to understand the structural causes of this reduction in forecast errors. In fact, given that our methodology allows for time-varying volatilities and is based on a fully fledged model, it provides an interpretation for the structural disturbances hitting the economy.

Figure 5 plots the volatility of GDP growth implied by our model. There are at least two things to notice from the comparison between Figure 5 and Figure 8 (p. 626), which reports a simple ten-year moving window estimate of the standard deviation of GDP growth. First, the DSGE model somewhat overpredicts the level of the volatility of GDP growth during the entire postwar period. The overprediction would be far less severe if Figure 8 had been produced using a shorter window, but would never completely disappear. This problem is common to the time invariant version of the model and is therefore indicative of difficulties in simultaneously matching the levels of persistence, comovements, and volatilities observed in the data, even with state-of-the-art DSGE models (Del Negro et al. 2007). Second, nonetheless, the model captures remarkably well
the timing and the size of the Great Moderation, despite the abrupt nature of this fall in volatility. Observe that the volatility of GDP growth starts declining around 1981, which is slightly earlier than some estimates provided by the literature using models with discrete structural breaks. This is due to the specification of our time-varying volatility model, which tends to smooth out abrupt changes (see, for example, Boivin 2001). Therefore, we will later compare the fit of our baseline model with one that allows for a jump in the variance of shocks.

To assess the role played by each shock in accounting for the Great Moderation, we rely on counterfactual simulation exercises. Our approach consists of using our stochastic volatility model to simulate the variability of GDP growth under alternative paths for the standard deviation of each structural disturbance. These counterfactual simulations can be interpreted as the hypothetical pattern of the volatility of GDP growth in the period 1981–2004, had the standard deviation of that particular structural shock remained unchanged with respect to its 1980 level.9

Figure 6 presents the results of our counterfactual exercises. Our approach gives a very strong conclusion about the causes of the Great Moderation. As evident from Figure 6D, the main explanation for the Great Moderation seems to be the sharp reduction in the volatility of the investment shocks. That is, had the volatility of these disturbances remained at its 1980 level, then the standard deviation of GDP growth would have been substantially higher than the one observed in the 1981–2004 period. Finally, it is worth noting that changes in the volatility of the monetary policy shock had a significant, although much smaller, effect on the decline in the variance of output growth (Figure 6A).

Even if our main focus is on the volatility of GDP, we also perform a similar counterfactual experiment for the volatility of inflation, which is also characterized by a substantial decline around the early 1980s. Similarly to output, the main contributor to the lower variability of inflation is the investment shock (Figure 7D). We will analyze the ability of alternative specifications to capture this simultaneous decline in inflation and output volatility in Section VII.

VI. Interpretation of the Main Result

Our results suggest that the key to understanding the Great Moderation is to analyze the reasons for the decline in volatility of the investment shocks. Broadly speaking, these are innovations specific to the return on capital or the marginal efficiency of the investment technology. In this section, we suggest two interpretations for these shocks and rely on evidence outside our DSGE model to argue for the plausibility of both views.

A. Investment-Specific Technology Shocks and the Relative Price of Investment

A strict, model-based interpretation indicates that the investment shocks correspond either to investment-specific technological disturbances or, equivalently, to shocks to the relative price of investment in terms of consumption goods. In recent years, data on the latter have been used by other authors to proxy for investment-specific technology shocks (see, for instance, Greenwood et al. 1997, 2000; and Fisher 2006). However, this relative price is not used in our estimation, as we also wish to consider other possible interpretations of the investment shock. Nonetheless, we first verify that the decline in the volatility of our investment shocks is at least broadly in line with a decline in the variance of the relative price of investment in terms of consumption goods.

9 More precisely, we fix the volatility of each shock to the average of the time-varying standard deviation for all four quarters in 1980. A longer window does not affect our results.
Figure 6. Actual and Counterfactual Standard Deviation (Std) for Output Growth

Notes: Counterfactual std obtained by fixing, for the remainder of the sample, the std of each shock, one at a time, at the four-quarter-average level of that shock’s time-varying standard deviation in 1980. These are computed with the draws generated in the estimation of the baseline stochastic volatility model.
Figure 7. Actual and Counterfactual Standard Deviation (Std) for Inflation

Notes: Counterfactual std obtained by fixing for the remainder of the sample the std of each shock, one at a time, at the four-quarter-average level of that shock’s time-varying standard deviation in 1980. These are computed with the draws generated in the estimation of the baseline stochastic volatility model.
In particular, we construct this relative price using the chain-weighted deflators for our components of consumption (nondurables and services) and investment (durables and total private investment) and estimate the standard deviation of its growth rate using a simple ten-year moving window. Figure 8 plots the estimate of this time-varying standard deviation and makes clear that the volatility of this relative price has sharply decreased in the second part of the postwar sample. Not only does the timing of this decline coincide remarkably well with the onset of the Great Moderation, but the patterns in the volatilities of the relative price of investment and GDP growth (also reported in Figure 8) are also very similar throughout the entire postwar period. It is important to stress, however, that the relative price of investment to consumption has a strong downward trend in the data, which cannot be accounted for by our model, as we assume that the investment shock is stationary. Nevertheless, we regard the fact that our model provides a very similar insight (without using any data on the relative price of investment) as a remarkable result.

A possible criticism to simply looking at the standard deviation of the relative price of investment is that its decline could have come either from a reduction in the volatility of factors affecting the demand for investment goods (such as monetary policy) or from a lower volatility of factors affecting the supply of investment goods (as investment-specific technological change). Supporting evidence for the latter is provided by Fisher (2006), who uses a structural vector autoregression (SVAR) to identify the investment-specific technology shock as the only disturbance affecting the relative price of investment in the long run. Fisher (2006) estimates the SVAR in two separate subsamples (from 1955:I to 1979:II and from 1982:III to 2000:IV) and finds that the standard deviation of the identified investment-specific technology shock in the second subsample is approximately 70 percent lower. This suggests that interpreting the lower variability of investment shocks as reflecting a decline in the volatility of disturbances to investment-specific technology would be consistent with the data.

B. Financial Frictions

The measure of the relative price of investment analyzed above does not include financial costs related to the purchase of durable and capital goods. If external financing is an important determinant of purchases of this type of goods, the financial cost of borrowing contributes to the effective cost of investment in terms of consumption goods, which could be reflected in our $\mu_t$ shock. If we subscribe to this broader interpretation of this disturbance—which is admittedly outside our current DSGE model—a natural explanation of the Great Moderation would also be based on a reduction in financial frictions.

Interestingly, this purely “theoretical” hypothesis squares remarkably well with the empirical and anecdotal evidence about the expanded access to credit and borrowing for firms and

---

10 This assumption allows us to directly compare our results to a large literature on Bayesian estimation of DSGE models (see, for instance, Levin et al. 2005; Del Negro et al. 2007; or Smets and Wouters 2007). Assuming, instead, a nonstationary investment shock and using data on the relative price of investment in the estimation would require at least two important modifications: first, we would need to deflate all real variables by the consumption deflator (see Greenwood et al. 1997 or Fisher 2006), which would then become our measure of price inflation, providing perhaps a less accurate characterization of monetary policy in the Taylor rule; second, we would need some additional structural shocks or measurement errors for the observable variables (to avoid stochastic singularity). We leave this extension for future work.

11 The intuitive link between our shock to the relative price of investment ($\mu_t$) and financial frictions is even more evident in models that take into consideration agency costs for the financing of investment, such as that of Charles T. Carlstrom and Timothy S. Fuerst (1997). In particular, Bernanke et al. (1999) and, more recently, Christiano and Joshua M. Davis (2006) and V. V. Chari, Patrick J. Kehoe, and Ellen R. McGrattan (2007) have argued that unmodeled financial frictions of that kind might be captured in reduced form by disturbances to the real return on capital similar to the one we emphasize in this paper.
households since the beginning of the 1980s. Important elements of this transformation were the passing of the Depository Institutions Deregulation and Monetary Control Act (DIDMCA) in 1980, particularly the demise of Regulation Q, and the Garn-St Germain Depository Institutions Act of 1982 (see, for instance, Patric H. Hendershott 1990; Dynan et al. 2006; or Campbell and Hercowitz 2006). These changes allowed households unprecedented access to external financing (Campbell and Hercowitz 2006), which was further facilitated by the emergence of secondary mortgage markets (Jonathan McCarty and Richard W. Peach 2002; Joe Peek and James A. Wilcox 2006). Moreover, firms’ access to external financing was enhanced by the development of a market for bonds with below-investment grade ratings (Gertler and Lown 1999), as well as a decline in the cost of new equity issuances (Urban Jermann and Vincenzo Quadrini 2006).

Notice that this broader interpretation seems capable of addressing some additional salient features of the Great Moderation. In particular, as noted by Stock and Watson (2002), Dynan et al. (2006), and Justiniano and Primiceri (2005), the most drastic reduction in volatility has characterized the time series of durable goods, investment, and, especially, residential investment. In addition, it is noteworthy that these events coincided with a decline in credit spreads for mortgages and BAA-grade firms, which is suggestive of a decline in financial frictions.12

We want to stress that our model accounts for the link between financial frictions and decreased macroeconomic volatility at best in reduced form and therefore lacks the structure to understand the underlying transmission mechanisms. In addition, incorporating financial frictions in our model is likely to alter not only the investment margin that we emphasize in this paper, but also other dimensions of the model. This said, our results suggest that further structural analysis of the Great Moderation would probably benefit from incorporating an explicit role for these frictions.

12 For the mortgage market, for instance, the large and volatile spreads of the early 1980s correspond to the end of a market dominated by heavily regulated thrift institutions, in which credit availability was subject to large swings due to fluctuations in deposits (Michael G. Bradley, Stuart A. Gabriel, and Mark E. Wohar 1995). The transition to smaller spreads is commonly associated with the beginning of more efficient and integrated financial markets (Hendershott 1990; McCarty and Peach 2002; Dynan et al. 2006; or Calvin Schnure 2005).
VII. Model Fit and Robustness Issues: 
The Role of Changes in Monetary Policy and Private Sector Behavior

This section has two objectives. First, we evaluate the fit of our stochastic volatility model relative to a number of alternative specifications. Second, we analyze changes in the conduct of systematic monetary policy and shifts in the private sector behavior as potential alternative explanations of the Great Moderation. On this latter point, our counterfactual exercises indicate that, in the context of this model, the lower variability of US output is difficult to explain when considering only these changes in parameters and absent shifts in the variance of the disturbances. As for the assessment of fit, we find that our baseline stochastic volatility model outperforms all alternative specifications, even when we allow for a single shift in the variance of the shocks.

A. Model Fit

We assess the fit of our model using the marginal likelihood (or marginal data density), which corresponds to the posterior density integrated over the model’s parameters. From a Bayesian perspective, the marginal likelihood is the most comprehensive and accurate measure of fit, as it can be used to construct posterior odds on competing models.13 The first two rows of Table 3 report the log-marginal data density for our model with stochastic volatility, and its time invariant counterpart. As is evident, the values of the log-marginal likelihood are strongly in favor of our stochastic volatility model.

Although the data seem to beg for time variation in the variance of the shocks, a natural question is whether accounting for a single break in these variances is enough to capture this salient feature of the data. The answer to this question is no. The third row of Table 3 reports the value of the marginal likelihood for a model with a single break in the volatilities (set in the first quarter of 1984), which delivers a substantially lower fit than our stochastic volatility baseline.

Rows 4 and 5 of Table 3 refer to two alternative specifications, in which all coefficients are allowed to change in 1984. The next subsection discusses in detail these model variants, while here we simply stress that one of them (row 5) incorporates the possibility that monetary policy may have shifted from passive to active over the sample.14 Both models deliver a rather poor fit of the data, compared not only with our baseline model, but also with the model in which only variances shift (row 3). It is also noteworthy that the specification where policy always remains active outperforms the one in which monetary policy shifts from passive to active.

B. The Role of Changes in Monetary Policy and Private Sector Behavior

Two recent studies by Luca Benati and Paolo Surico (2006) and Thomas Lubik and Surico (2006) have questioned the usefulness of reduced form methods to analyze the role of monetary policy in the Great Moderation. These authors recommend, instead, to investigate this question in the context of structural models. Therefore, despite the poor fit of both models with time variation in the systematic part of monetary policy and private sector behavior, we are interested in using our DSGE model to address the possibility that these changes may account for the Great Moderation.

13 Technical details about the computation of the marginal likelihood are presented in Appendix A.
14 As will become clear in the next subsection, monetary policy is denoted as passive when the reaction to inflation is weak and it is not consistent with a determinate equilibrium.
We first discuss the estimation of the model with a one-time change in all coefficients. We then turn to counterfactual experiments that compare the decline in the variance of GDP growth and inflation observed in the data with those predicted by the model under alternative scenarios for the changes in coefficients across subsamples.

**Split Sample Estimation.**—We focus on the model in which all coefficients are allowed to shift across two subsamples: 1953:I–1983:IV and 1984:I–2004:IV.\(^{15}\) We estimate two variants of this model, one in which values of the policy reaction to inflation \(\phi_p\) are always consistent with a unique equilibrium, and an alternative model where \(\phi_p\) is, instead, consistent with multiple equilibria in the first subsample. Several authors have suggested a shift from passive to active monetary policy as a prominent explanation of the Great Moderation. Therefore, and since at least in principle the indeterminacy channel can explain the higher variance of GDP growth in the pre-Volcker period, we focus mainly on the model with indeterminacy.

In order to construct the likelihood function, we apply the methodology of Lubik and Schorfheide (2003), who show how to parameterize the continuum of possible solutions under indeterminacy.\(^{16}\) In this case, the model solution differs in two important ways from the standard solution under determinacy. First, the transmission mechanism of the fundamental shocks is not pinned down uniquely by the structural coefficients, and additional free parameters are needed to characterize the solution; second, sunspot shocks can contribute to the volatility of the model’s endogenous variables. The reader is referred to Appendix C for additional details about the prior and the model solution in the indeterminate case. Here we point out that, according to our prior, the direct effect of the sunspot shocks, combined with these changes in the transmission mechanism of the fundamental disturbances, results in substantially higher a priori model-implied standard deviations for all series relative to our benchmark determinate model. Moreover, we consider only

<table>
<thead>
<tr>
<th>Specification</th>
<th>Log marginal(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline stochastic volatility model(^b)</td>
<td>−1,824.6</td>
</tr>
<tr>
<td>Time-invariant model(^c)</td>
<td>−1,984.7</td>
</tr>
<tr>
<td>Determinate model with a single jump in volatilities</td>
<td>−1,925.8</td>
</tr>
<tr>
<td>Split model with a jump in all coefficients and active policy in first subsample</td>
<td>−1,947.9</td>
</tr>
<tr>
<td>Split model with a jump in all coefficients and passive policy in first subsample(^d)</td>
<td>−1,959.0</td>
</tr>
<tr>
<td>Stochastic volatility with Taylor rule responding to output gap(^e)</td>
<td>−1,843.6</td>
</tr>
<tr>
<td>Stochastic volatility model with two stochastic trends</td>
<td>−1,838.0</td>
</tr>
</tbody>
</table>

\(^a\)Log Marginal data density computed using the output of the MCMC simulators as described in Appendix A. Model favored by the data attains the highest marginal data density.

\(^b\)Parameter estimates shown in column 3 of Table 1.

\(^c\)Parameter estimates shown in column 2 of Table 1.

\(^d\)Parameter estimates shown in Table 4.

\(^e\)Parameter estimates shown in Table 2.

Full set of parameter estimates for the remaining models is available from the authors upon request.

---

\(^{15}\) We split the sample in 1984 because it corresponds to the onset of the Great Moderation. Justiniano and Primiceri (2005) perform a similar exercise excluding the 1979–1983 period and reach very similar conclusions.

\(^{16}\) When estimating the indeterminate model, we modify our prior for \(\phi_p\) for the first subsample following Lubik and Schorfheide (2004), and specify a Gamma distribution with mean equal to 1.1 and standard deviation equal to 0.5. This prior assigns roughly equal probability to the determinacy and indeterminacy regions of the parameter space as determined by \(\phi_p\).
regions of the parameter space in which the degree of indeterminacy is one, and indeterminacy is generated by low values of $\phi_\pi$, i.e., low interest rate reactions to inflation. This means that we effectively truncate our prior at the boundary of a multidimensional indeterminacy region.\footnote{In this model, it is not possible to characterize the region of indeterminacy analytically. While the strength of the policy reaction to inflation is the crucial parameter, extensive simulation analysis suggests that determinacy of equilibrium also hinges on the remaining coefficients of the Taylor rule. Moreover, indeterminacy also seems to occur for very large values of the degree of wage rigidity, which are substantially farther out from the right tail of our posterior estimates. As mentioned above, in our exercise, we divide the determinacy and indeterminacy regions of the parameter space based only on the response to inflation in the Taylor rule.}

Table 4 presents medians, fifth, and ninety-fifth percentiles of the posterior distribution of the coefficients estimated over the two subsamples under indeterminacy before 1984. Consistent with the results of the stochastic volatility model, there are important changes in the standard deviation of the shocks across sample periods. However, some of the remaining coefficient estimates exhibit differences across subsamples as well. By construction, this is particularly evident for the policy reaction to inflation, $\phi_\pi$, which displays a considerable increase, switching from a value implying multiple equilibria to a value consistent with a unique equilibrium (Clarida et al. 2000). Because of space considerations, we do not report estimates for the model in which policy is always active, but note simply that coefficient changes are less pronounced in this case.

\textit{Counterfactual Experiments}.—Our aim is to understand whether changes in monetary policy and private sector coefficients across subsamples can induce a reduction in the volatility of the endogenous variables similar to the one observed in the data. Once again, we focus on the results of the indeterminate model and briefly return to the determinate case toward the end of the section.

As a starting point, the first column of Table 5A shows that the unconditional standard deviation of GDP growth in the second sample period relative to the first is 0.76, while this ratio is 0.26 for inflation.\footnote{Notice that, compared to the data, the model of this section underpredicts the decline in volatility of GDP growth and overpredicts the decline in volatility of inflation. This is also related to the poor fit of this model without stochastic volatility, documented in Section VIIA.} In the context of our model, this decline can potentially be explained by three different sets of parameters: the standard deviations of the shocks, monetary policy coefficients, and the remaining structural coefficients. Our goal here is to isolate the contribution of the last two.

We begin by assessing the role of a policy shift from passive to active on the variability of output growth and inflation. To this end, we compare the model-implied standard deviation of these two variables when the coefficients of the Taylor rule estimated in the second subsample replace those of the first subsample, leaving all other coefficients unchanged. These counterfactual standard deviations are given in the second column of Table 5A (relative to the volatility estimates in the first subsample). Similarly, the third column of Table 5A corresponds to the relative standard deviations when we redo this counterfactual exercise but replace, in addition, the estimated coefficients of the private sector from the second sample period. Finally, the last column presents the ratio of standard deviations once we also replace the autocorrelation of the shocks from the second subsample.

Table 5A makes some important points. First, if monetary policy in the 1960s and 1970s had been as aggressive against inflation and as weak in response to real activity as in the later part of the sample, we would have probably experienced substantially lower inflation volatility. However, the variability of GDP growth would have been, if anything, even higher.\footnote{More precisely, the volatility of the output gap is substantial in our model, especially in the first subsample. The stronger response to inflation and weaker response to real activity would have somewhat lowered this volatility. However, the counterfactual policy would have also altered the correlation between the output gap and potential output in such a way that, as a result, we would have observed more volatile output.} This result is in line with Efrem Castelnuovo (2006) and Boivin and Giannoni (2006), who also note
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Std [ 5 , 95 ]</td>
<td>Median</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>Price indexation</td>
<td>0.55 0.07 [ 0.45 , 0.67 ]</td>
<td>0.66 0.08 [ 0.51 , 0.79 ]</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>Wage indexation</td>
<td>0.05 0.02 [ 0.02 , 0.09 ]</td>
<td>0.31 0.07 [ 0.20 , 0.43 ]</td>
</tr>
<tr>
<td>( \mu )</td>
<td>SS technology growth rate</td>
<td>0.48 0.02 [ 0.44 , 0.52 ]</td>
<td>0.47 0.02 [ 0.43 , 0.51 ]</td>
</tr>
<tr>
<td>( \nu )</td>
<td>Consumption habit</td>
<td>0.67 0.03 [ 0.61 , 0.72 ]</td>
<td>0.77 0.04 [ 0.71 , 0.83 ]</td>
</tr>
<tr>
<td>( \lambda_s )</td>
<td>SS mark-up goods prices</td>
<td>0.23 0.04 [ 0.16 , 0.29 ]</td>
<td>0.18 0.04 [ 0.11 , 0.24 ]</td>
</tr>
<tr>
<td>( \lambda_w )</td>
<td>SS mark-up wages</td>
<td>0.11 0.04 [ 0.07 , 0.20 ]</td>
<td>0.19 0.04 [ 0.12 , 0.26 ]</td>
</tr>
<tr>
<td>( \xi )</td>
<td>SS quarterly inflation</td>
<td>0.52 0.08 [ 0.40 , 0.66 ]</td>
<td>0.82 0.07 [ 0.71 , 0.95 ]</td>
</tr>
<tr>
<td>( \phi )</td>
<td>SS real interest rate</td>
<td>0.87 0.06 [ 0.77 , 0.98 ]</td>
<td>0.75 0.07 [ 0.64 , 0.86 ]</td>
</tr>
<tr>
<td>( \psi )</td>
<td>Inverse Frisch labor</td>
<td>0.58 0.12 [ 0.38 , 0.80 ]</td>
<td>1.92 0.61 [ 1.15 , 3.11 ]</td>
</tr>
<tr>
<td>( \xi_p )</td>
<td>Calvo prices</td>
<td>0.91 0.02 [ 0.88 , 0.94 ]</td>
<td>0.90 0.02 [ 0.88 , 0.93 ]</td>
</tr>
<tr>
<td>( \xi_c )</td>
<td>Calvo wages</td>
<td>0.69 0.04 [ 0.62 , 0.76 ]</td>
<td>0.45 0.08 [ 0.34 , 0.60 ]</td>
</tr>
<tr>
<td>( \chi )</td>
<td>Elasticity capital utilization costs</td>
<td>6.92 0.76 [ 5.73 , 8.15 ]</td>
<td>4.94 0.90 [ 3.70 , 6.63 ]</td>
</tr>
<tr>
<td>( S' )</td>
<td>Investment adjustment costs</td>
<td>1.48 0.19 [ 1.12 , 1.77 ]</td>
<td>2.83 0.53 [ 2.09 , 3.83 ]</td>
</tr>
<tr>
<td>( \Phi_r )</td>
<td>Taylor rule inflation</td>
<td>0.52 0.09 [ 0.37 , 0.68 ]</td>
<td>2.37 0.19 [ 2.08 , 2.70 ]</td>
</tr>
<tr>
<td>( \Phi_y )</td>
<td>Taylor rule output</td>
<td>0.19 0.03 [ 0.14 , 0.24 ]</td>
<td>0.02 0.01 [ 0.00 , 0.04 ]</td>
</tr>
<tr>
<td>( \rho_g )</td>
<td>Taylor rule smoothing</td>
<td>0.69 0.04 [ 0.62 , 0.74 ]</td>
<td>0.84 0.02 [ 0.80 , 0.87 ]</td>
</tr>
<tr>
<td>( \rho_i )</td>
<td>Technology growth</td>
<td>0.18 0.05 [ 0.10 , 0.26 ]</td>
<td>0.31 0.08 [ 0.16 , 0.44 ]</td>
</tr>
<tr>
<td>( \rho_b )</td>
<td>Government spending</td>
<td>0.97 0.01 [ 0.94 , 0.98 ]</td>
<td>0.98 0.00 [ 0.98 , 0.98 ]</td>
</tr>
<tr>
<td>( \rho_p )</td>
<td>Investment-specific technology</td>
<td>0.66 0.05 [ 0.58 , 0.74 ]</td>
<td>0.92 0.02 [ 0.89 , 0.95 ]</td>
</tr>
<tr>
<td>( \rho_e )</td>
<td>Mark-up</td>
<td>0.07 0.02 [ 0.03 , 0.11 ]</td>
<td>0.93 0.05 [ 0.83 , 0.97 ]</td>
</tr>
<tr>
<td>( \rho_t )</td>
<td>Labor disutility</td>
<td>0.93 0.02 [ 0.89 , 0.95 ]</td>
<td>0.83 0.07 [ 0.69 , 0.92 ]</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>Monetary policy</td>
<td>0.27 0.02 [ 0.25 , 0.31 ]</td>
<td>0.15 0.01 [ 0.13 , 0.17 ]</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>Technology growth</td>
<td>1.27 0.08 [ 1.14 , 1.40 ]</td>
<td>0.82 0.07 [ 0.71 , 0.92 ]</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>Government spending</td>
<td>0.71 0.05 [ 0.64 , 0.80 ]</td>
<td>0.48 0.04 [ 0.42 , 0.52 ]</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>Investment-specific</td>
<td>5.08 0.43 [ 4.26 , 5.68 ]</td>
<td>2.63 0.37 [ 2.10 , 3.32 ]</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>Market-up</td>
<td>0.19 0.01 [ 0.17 , 0.22 ]</td>
<td>0.13 0.01 [ 0.12 , 0.16 ]</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>Labor disutility</td>
<td>30.58 2.26 [ 26.59 , 34.25 ]</td>
<td>7.15 4.90 [ 4.29 , 16.24 ]</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>Intertemporal preference</td>
<td>2.82 0.12 [ 2.60 , 2.99 ]</td>
<td>2.17 0.40 [ 1.69 , 2.98 ]</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>Sunspot</td>
<td>0.08 0.01 [ 0.06 , 0.09 ]</td>
<td>0.71 0.04 [ 0.66 , 0.75 ]</td>
</tr>
<tr>
<td>( M_i )</td>
<td>-0.20 0.19 [ -0.51 , 0.12 ]</td>
<td>-0.40 0.19 [ -0.69 , 0.10 ]</td>
<td></td>
</tr>
<tr>
<td>( M_i )</td>
<td>-0.10 0.10 [ -0.26 , 0.08 ]</td>
<td>-0.15 0.16 [ -0.39 , 0.15 ]</td>
<td></td>
</tr>
<tr>
<td>( M_i )</td>
<td>-0.15 0.16 [ -0.39 , 0.15 ]</td>
<td>-0.04 0.05 [ -0.12 , 0.02 ]</td>
<td></td>
</tr>
<tr>
<td>( M_i )</td>
<td>-0.04 0.05 [ -0.12 , 0.02 ]</td>
<td>-0.76 0.19 [ -1.02 , -0.43 ]</td>
<td></td>
</tr>
<tr>
<td>( M_i )</td>
<td>-0.76 0.19 [ -1.02 , -0.43 ]</td>
<td>-0.03 0.01 [ -0.04 , -0.02 ]</td>
<td></td>
</tr>
<tr>
<td>( M_i )</td>
<td>-0.03 0.01 [ -0.04 , -0.02 ]</td>
<td>-0.23 0.07 [ -0.34 , -0.11 ]</td>
<td></td>
</tr>
</tbody>
</table>

(log) Likelihood at median \(-1.776.0\)

Notes: Calibrated parameters and priors for both subsamples are identical to those reported for the time-invariant model in Table 1, with the exception of the coefficient on inflation in the Taylor rule. For that coefficient, given that we allow for indeterminacy, the prior in the first subsample is a Gamma with mean 1.1 and std 0.5, following Lubik and Schorfheide (2004). For the second subsample we return to the original normal prior centered at 1.7 with dispersion 0.3 which is used throughout the paper. Indeterminacy results in eight new coefficients: the seven \( M \) parameters that modify the transmission of the fundamental shocks, and the standard deviation of the sunspot shock (\( \sigma_y \)). As in Lubik and Schorfheide (2004), we center our prior in the first subsample on the continuity solution and therefore specify a prior for each \( M \) as Normal(0,0.5). See Appendix C for details. For the sunspot, we choose an Inverse Gamma 1 prior with mean and dispersion equal to 0.15.

Median, standard deviations and posterior percentiles of 150,000 draws from the Random Walk Metropolis algorithm for the joint estimation of the model with indeterminacy in the first subsample. We discard the initial 50,000 draws.

that more aggressive monetary policy would have increased output variability, absent a change in the shocks’ process. One explanation for this result is that the model attributes a large share of output variability to “supply” shocks—i.e., the investment shocks and neutral technology disturbances—which in the first subsample imply a strong negative correlation between inflation.
and output. Hence, counterfactually imposing a policy response that is more accommodating toward output and substantially more aggressive toward inflation leads, all else equal, to an increase in the variability of output.

Our second counterfactual indicates that jointly considering changes in monetary policy and private sector behavior could have reduced the variability of inflation but, once again, not output growth. Our last counterfactual suggests that changes in the autocorrelation of the disturbances could have influenced the implied standard deviations as well. These changes, however, would have induced a higher volatility of output and all other variables absent a decline in the variance of the shocks, which is at odds with the data. On this last point, we note that this is mainly related to a substantial increase in the autocorrelation of the labor disutility shock in the post-1984 sample, which is almost fully compensated by an even more sizable reduction in its volatility.\footnote{Indeed, simulations not reported in Table 5 indicate that the two effects counteract each other almost entirely. These large changes in the persistence and volatility of the labor disutility shock are probably due to weak identification in the short subsamples. In fact, these changes are smaller under determinacy since it is well known that indeterminacy generates higher persistence.}

Table 5B performs similar counterfactual simulations for the split-sample model in which policy is active throughout the entire sample. In this case, $\phi_\pi$ is estimated at 1.63 for the first

| Table 5—Counterfactual Standard Deviations (std) of Output and Inflation |
|--------------------------|-----------------------------|-----------------|-----------------|
|                          | (1)                         | (2)             | (3)             | (4)             |
|                          | Estimated                   | Counterfactual: | Counterfactual: | Counterfactual: |
|                          |                             | Monetary policy | Monetary policy | All coefficients |
|                          |                             | and private sector | and private sector | except volatilities |
| Output                   | 0.76                        | 1.81            | 1.04            | 2.09            |
| Median and [5,95] posterior bands | [0.66 , 0.88] | [1.58 , 2.17] | [0.87 , 1.28] | [1.35 , 3.14] |
| Inflation                | 0.25                        | 0.25            | 0.28            | 0.58            |
| Median and [5,95] posterior bands | [0.17 , 0.34] | [0.17 , 0.32] | [0.19 , 0.37] | [0.36 , 1.00] |

\(\text{Panel A. Passive policy in first subsample}\)

|                          | (1)                         | (2)             | (3)             | (4)             |
|                          | Estimated                   | Counterfactual: | Counterfactual: | Counterfactual: |
|                          |                             | Monetary policy | Monetary policy | All coefficients |
|                          |                             | and private sector | and private sector | except volatilities |
| Output                   | 0.57                        | 1.07            | 1.13            | 1.31            |
| Median and [5,95] posterior bands | [0.49 , 0.66] | [0.98 , 1.19] | [0.90 , 1.44] | [0.97 , 1.78] |
| Inflation                | 0.69                        | 0.73            | 0.70            | 1.40            |
| Median and [5,95] posterior bands | [0.40 , 1.16] | [0.54 , 0.88] | [0.50 , 0.99] | [0.71 , 3.29] |

\(\text{Panel B. Active policy in first subsample}\)

Notes: Ratios of model implied stds (either estimated in the second subsample or counterfactual) to the estimated std in the first subsample. First subsample ends in 1983:IV, second subsample starts in 1984:I. Medians and [5,95] posterior probabilities computed using MCMC draws for the models in which all coefficients are allowed to change in 1984:I. Panel A: monetary policy is passive and equilibrium indeterminate in first subsample. Panel B: monetary policy is active in first subsample. Column (1): ratio of stds implied by the estimated models in the second (numerator) to the first (denominator) subsamples. Column (2): ratio of counterfactual std when in the first subsample we replace the Taylor rule coefficients estimated in the second subsample. Column (3): same as column (2), but in addition replace private sector coefficients with those estimated in the second subsample. Column (4): same as column (3), but in addition replace the autocorrelation coefficients of the shocks with those estimated in the second subsample.
subsample; hence the policy shift in 1984 is less dramatic than in the indeterminate case. Accordingly, counterfactual simulations exhibit milder variations in response to the policy change. Variations in autocorrelation coefficients are more subdued in this case as well, which is reflected in a less drastic change in variances in the last column. However, the general picture from this counterfactual exercise is broadly in line with the results of Table 5A.

Summarizing, changes in the systematic conduct of monetary policy may have contributed to the reduction in inflation variability. But Table 5 makes evident that neither changes in monetary policy nor the remaining model coefficients—other than a fall in the volatilities—can account for the observed simultaneous decline in the variability of output and inflation. This is true even when we allow for indeterminacy and multiple equilibria in the pre-1984 period, although the data would seem to favor the determinate case.

On balance, however, we mention some possible caveats to the results of this section. In particular, allowing for all coefficients to change may be prone to identification issues in a large-scale model like ours estimated on relatively short subsamples. This is perhaps more so for the model with passive policy, if the identification of indeterminacy is rather weak (Andreas Beyer and Roger Farmer 2007). While the risk of overfitting is significant for both models, this seems particularly the case for the indeterminate model, as reflected in the marginal data density reported in Table 3. Finally, one possible explanation for the poor fit of all models with policy shifts is the fact that we have forced the change in policy to occur at the time of the Great Moderation (i.e., in 1984). In this respect, a promising direction for future work would be to combine the time-varying volatility model of this paper with the time-varying coefficients specification of Fernández-Villaverde and Rubio-Ramírez (2007b).

The fact that we do not find a big role for improved monetary policy in the decline of output volatility is consistent with evidence provided by Stock and Watson (2003), Ahmed et al. (2004), Primiceri (2005), and Sims and Zha (2006). Unlike these authors, however, our analysis is conditional on a DSGE model, with all the pros and cons that this entails. One advantage of a structural approach is that counterfactual exercises involving changes in policy are easier to interpret. Of course, this comes at the cost of imposing a set of stronger identifying restrictions on the analysis. Therefore, we cannot exclude the possibility that there are other less standard models in which an explanation for the Great Moderation based on changes in monetary policy may fare better. For example, we suspect that deviating from full information rational expectations might be promising (see, for instance, Athanasios Orphanides and John C. Williams 2005; Erceg and Levin 2003; Carl E. Walsh 2005), although we do not entertain this possibility in this paper.

VIII. Concluding Remarks

In this paper we have estimated a large scale DSGE model of the US business cycle allowing for the volatility of the structural disturbances to change over time. Our results indicate that the volatility of several structural shocks has changed dramatically in the postwar period. However, the sharp reduction in the standard deviation of GDP growth that has characterized the last 20 years can be explained mostly by the decline in the variability of a single shock: the shock specific to the equilibrium condition of investment.

---

21 To account for problems of overfitting, rows 6 and 7 of Table 3 report the log-marginal likelihood for two additional versions of the model, one in which only monetary policy and volatilities undergo a one-time change in 1984, and another one in which only monetary policy can switch (in both cases from passive to active) at the same date. The fit of these two models is −1941.1 and −1983.4, respectively, which, in both cases, is lower than the fit of the model with a single shift in variances.
We have suggested two interpretations of this finding. The first view is strictly model-based and suggests that these disturbances correspond either to investment-specific technological shocks or, equivalently, to innovations in the relative price of investment in terms of consumption goods. In addition, since a broad interpretation of this relative price should include the financial costs of borrowing for the purchase of investment goods, we have also suggested an interpretation of the Great Moderation based on a decline in financial frictions.

More generally, our results point to dramatic changes in the investment equilibrium condition that played a prominent role in the Great Moderation.

**Appendix A: The Estimation Algorithm**

**A. The Standard Case: Homoskedastic Disturbances**

For the model without stochastic volatility, the estimation algorithm is a random walk Metropolis MCMC procedure, as suggested originally by Schorfheide (2000). We first use a maximization algorithm (Chris Sims’s csmvminwel) to find the posterior mode and to obtain an inverse Hessian, which then becomes the dispersion measure for our proposal distribution. This is done for multiple initial values (at least 30) drawn at random from our prior to ensure convergence of this initial search to a unique mode. We then scale this variance-covariance matrix to attain an acceptance rate close to 0.25, as it is usually suggested. Appendix B discusses convergence diagnostics which we also apply to this posterior simulator.

**B. Stochastic Volatility**

When the structural shocks exhibit stochastic volatility, this algorithm must be modified to account for inference on the unobserved stochastic volatilities. A Metropolis within Gibbs MCMC algorithm allows us to iteratively draw from the posterior densities of the DSGE model’s parameters, stochastic volatilities, and associated innovation variances. Some steps of the algorithm require data augmentation (Martin A. Tanner and Wing Hung Wong 1987) using the Gibbs sampler in order to draw other latent auxiliary variables (Chib 2001). For instance, as discussed below, generating a draw for the stochastic volatilities entails using a normal mixture approximation and sampling a set of latent indicators for the components of this mixture.

To illustrate the steps involved in sampling from the different blocks, let the vector $\theta$ collect all parameters of the DSGE model (other than the standard deviations of the structural disturbances of the time-invariant model) and notice that the solution of the log-linearized DSGE model leads to a state-space representation of the form

\begin{align}
\mathbf{x}_t &= D\mathbf{\xi}_t + \text{const}, \\
\mathbf{\xi}_t &= \mathbf{A}(\theta)\mathbf{\xi}_{t-1} + \mathbf{B}(\theta)\mathbf{\eta}_t,
\end{align}

where $\mathbf{x}_t$ represents a vector of observable variables and $\mathbf{\xi}_t$ denotes the vector of endogenous/state variables in log-deviation from the deterministic steady state. As discussed in Section I, the novelty of our framework is that the vector of structural innovations $\mathbf{\eta}_t$ (dimension $n \times 1$) is allowed to have a time-varying variance covariance matrix. Indexing each structural shock by $i$, the stochastic volatilities for each shock are modelled as

\begin{align}
\mathbf{\eta}_{i,t} &= \mathbf{\sigma}_{i,t}\mathbf{e}_{i,t},
\end{align}
\[ \log \sigma_{i,t} = (1 - \rho_\sigma) \log \sigma_i + \rho_\sigma \log \sigma_{i,t-1} + \nu_{i,t}, \]

(A4)

\( e_{i,t} \sim N(0, 1), \)

(A5)

\( \nu_{i,t} \sim N(0, \omega_i^2) \quad i = 1, \ldots, n. \)

Let the vector \( \mathbf{h}_i \), with entry \( i \) given by \( h_{i,t} = \log \sigma_{i,t} \), collect the log volatilities for all shocks at time \( t \), and stack the whole sample of stochastic volatilities into the matrix \( \mathbf{H}^T = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_n] \).

Finally, we denote the sample of structural shocks as \( \hat{\eta}^T = [\hat{\eta}_1, \hat{\eta}_2, \ldots, \hat{\eta}_n, \ldots, \hat{\eta}_T]' \) and the vector including all the fixed coefficients of the volatility processes by

\[ \Phi = [\sigma_1, \rho_{\sigma_1}, \omega_1^2, \ldots, \sigma_n, \rho_{\sigma_n}, \omega_n^2]. \]

Suppose that the MCMC algorithm has completed iteration \( g \) (> 0), producing samples \( \theta^{(g)} \), \( \mathbf{H}^{T(g)} \), and \( \Phi^{(g)} \) of the parameters of interest (individual elements of a vector are indexed by \( i \) while \( (g) \) indicates the current state of the chain). In iteration \( g + 1 \), the following five steps are used to generate a set of new draws.

**Step 1: Draw the structural shocks** \( \hat{\eta}^{T,(g+1)} \).

In order to generate a new sample of the stochastic volatilities we must first obtain a new draw of the structural shocks. This can be done easily using the efficient simulation smoother for disturbances developed by Durbin and Koopman (2002). The simulation smoother is applied to the state space representation given by (A1) and (A2).

**Step 2: Draw the stochastic volatilities** \( \mathbf{H}^{T,(g+1)} \).

With a draw of \( \hat{\eta}^T \) in hand, the system of nonlinear measurement equations in (A3) for each structural shock can be easily converted into a linear one by squaring and taking logarithms of every element. Due to the fact that the squared shocks \( \hat{\eta}_i^2 \) can be very small, an offset constant is used to make the estimation procedure more robust. Dropping the iteration indicators momentarily for ease of notation, this leads to the following approximating state space form:

(A6) \[ \tilde{\eta}_{i,t} = 2h_{i,t} + e_{i,t}, \]

(A7) \[ h_{i,t} = h_{i,t-1} + \nu_{i,t}, \]

where \( \tilde{\eta}_{i,t} = \log [(\hat{\eta}_{i,t})^2 + \bar{c}] ; \) \( \bar{c} \) is the offset constant (set to 0.001); and \( e_{i,t} = \log (e_{i,t}^2) \).

Observe that the \( e \)'s and the \( \nu \)'s are not correlated. The resulting system has a linear, but non-Gaussian, state space form, because the innovations in the measurement equations are distributed as a log \( \chi^2(1) \). In order to further transform the system in a Gaussian one, a mixture of normals approximation of the log \( \chi^2(1) \) distribution is used, as described in Kim et al. (1998). Under the assumption of orthogonality across the \( e \)'s (recall the variance covariance matrix of the \( e \)'s is the identity matrix), this implies that the variance covariance matrix of the \( \nu \)'s is also diagonal, which justifies using the same (independent) mixture of normals approximation for each innovation:

\[ f(e_{i,t}) = \sum_{k=1}^{K} q_k f_N(e_{i,t} | s_{i,t} = k) \quad i = 1, \ldots, n, \]

where \( s_{i,t} \) is the indicator variable selecting which member of the mixture of normals has to be used at time \( t \) for the innovation \( i \), \( q_k = \Pr(s_{i,t} = k) \), and \( f_N(\cdot) \) denotes the pdf of a normal
distribution. Kim et al. (1998) select a mixture of seven normal densities ($K = 7$) with component probabilities $q_k$, means $m_k = 1.2704$, and variances $r_j^2$, $j = 1, \ldots, 7$, chosen to match a number of moments of the log $\chi^2(1)$ distribution. For completeness the constants \{ $q_j, m_j, r_j^2$ \} are reported below:\textsuperscript{22}

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$q_j$</th>
<th>$m_j$</th>
<th>$r_j^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00730</td>
<td>-10.12999</td>
<td>5.79596</td>
</tr>
<tr>
<td>2</td>
<td>0.10556</td>
<td>-3.97281</td>
<td>2.61369</td>
</tr>
<tr>
<td>3</td>
<td>0.00002</td>
<td>-8.56686</td>
<td>0.34023</td>
</tr>
<tr>
<td>4</td>
<td>0.04395</td>
<td>2.77786</td>
<td>0.16735</td>
</tr>
<tr>
<td>5</td>
<td>0.34001</td>
<td>0.61942</td>
<td>0.64009</td>
</tr>
<tr>
<td>6</td>
<td>0.24566</td>
<td>1.79518</td>
<td>0.34023</td>
</tr>
<tr>
<td>7</td>
<td>0.25750</td>
<td>-1.08819</td>
<td>1.26261</td>
</tr>
</tbody>
</table>


Conditional on $s^{T,(g)}$, the system has an approximate linear and Gaussian state space form. Therefore, a new draw for the complete history of the volatility $H^{T,(g+1)}$ can be obtained recursively with the standard Gibbs sampler for state space forms using, for instance, the forward-backward recursion of Carter and Kohn (1994).

**Step 3: Draw the indicators of the mixture approximation $s^{T,(g+1)}$.**

A new sample of the indicators, $s^{(g+1),t}$, for the mixture is obtained conditional on $\hat{h}^{T,(g+1)}$ and $H^{T,(g+1)}$ by independently sampling each from the discrete density defined by

$$
Pr(s_{i,t}^{(g+1)} = j | \hat{h}^{T,(g+1)}, H_{i,t}^{(g+1)}) \propto q_j f (\hat{h}_{i,t}^{(g+1)} | 2h_{i,t}^{(g+1)} + m_j - 1.2704, r_j^2), j = 1, \ldots, 7.
$$

Consistent with the notation above, collect the indicators for which component of the mixture of the normal approximation to use for each structural shock and time period into a stacked matrix $s^{T,(g+1)} = [s_1^{(g+1)}, s_2^{(g+1)}, \ldots, s_{i,t}^{(g+1)}, \ldots, s_T^{(g+1)}]'$.

**Step 4: Draw the coefficients of the stochastic volatility processes.**

Having generated a sample $H^{T,(g+1)}$, the elements of the vector $\Phi^{(g+1)}$ can be generated easily from usual Normal inverse-Gamma distributions.

**Step 5: Draw the DSGE parameters $\theta^{(g+1)}$.**

As in the time invariant algorithm, a new candidate parameter $\theta^*$ is drawn from a proposal density. In this case, however, the computation of the likelihood used to construct the probability of acceptance depends on $H^{T,(g+1)}$. More formally, the candidate draw is accepted with probability

$$
a = \min \left\{ 1, \frac{\mathcal{L}(X|\theta^*, H^{T,(g+1)}) \pi(\theta^*)}{\mathcal{L}(X|\theta^{(g)}, H^{T,(g+1)}) \pi(\theta^{(g)})} \right\},
$$

where $X$ is the matrix of data and $\mathcal{L}(\cdot)$ and $\pi(\cdot)$ denote the likelihood and the prior distribution, respectively.

These five steps are repeated $N$ times, across multiple chains.

\textsuperscript{22} We abstract from the reweighting procedure used in Kim et al. (1998) to correct the minor approximation error.
We conclude by observing that the computational demands of this algorithm are substantial. As a benchmark, generating and storing 100 draws takes roughly 11 minutes in a computer using an AMD Athlon 64 processor 3000+ at 1.81 Ghz with 1 MB of RAM.

C. Computation of the Marginal Likelihood

As it is standard in the literature, for the computation of the marginal likelihood of the time-invariant model, we use the modified harmonic mean method of Alan E. Gelfand and Depak K. Dey (1994) and John F. Geweke (1998).

For the computation of the marginal likelihood of the model with stochastic volatility, we also use a version of the modified harmonic mean method. In fact, it is easy to show that

$$m(X) = \left[ \int \frac{f(\theta, H^T)}{\mathcal{L}(X|\theta, H^T) \cdot \pi(\theta, H^T)} \cdot p(\theta, H^T | X) \cdot d(\theta, H^T) \right]^{-1},$$

where $X$ is the matrix of data, $m(X)$ denotes the marginal data density, and $\mathcal{L}(\cdot)$, $\pi(\cdot)$, and $p(\cdot)$ denote the sampling, the prior, and the normalized posterior densities, respectively; $f(\cdot)$ can be any pdf with support contained in the support of the posterior density. For computational convenience, we choose $f(\theta, H^T) = f(\theta) \cdot f(H^T)$. Moreover, following Geweke (1998), $f(\theta)$ is chosen to be a truncated multivariate normal with mean and variance equal to the mean and variance of the posterior draws of $\theta$. Since the dimension of $H^T$ is very large, we have decided to set $f(H^T) = \pi(H^T)$. It follows that the marginal likelihood can be approximated by

$$\bar{m}_N(X) = \left[ \frac{1}{N} \sum_{j=1}^{N} \frac{f(\theta_j)}{p(X|\theta_j, H^T_j) \cdot \pi(\theta_j)} \right]^{-1},$$

where $\theta_j$ and $H^T_j$ are draws from the posterior distribution and we have used the fact that our priors for $\theta$ and $H^T$ are independent.

It is very easy to compute $\bar{m}(X)$, which converges (almost surely) to $m(X)$, such that it represents a consistent estimate of marginal likelihood (Michael A. Newton and Adrian E. Raftery 1994). It might not be stable, however, as it may not satisfy a Gaussian central limit theorem because the random variable $f(\theta_j)/[p(X|\theta_j, H^T_j) \cdot \pi(\theta_j)]$ can violate the assumption of a finite variance. Nevertheless, in the particular application carried out in this paper, we have noted that this method works quite well, delivering estimates of the marginal likelihood, which are almost identical across multiple simulation chains and models with slightly different priors.

We have, nonetheless, also tried to compute the marginal likelihood using a combination of the modified harmonic mean and the method of Chib (1995). Our experience indicates that this alternative method is very unreliable and extremely slow to converge. Indeed, different chains produced wide variation in estimates of the marginal likelihood, although posterior coefficient and volatility estimates were almost identical across runs.

APPENDIX B: CONVERGENCE

We assess the convergence of our posterior simulators using a battery of diagnostics. For the stochastic volatility model, we launch multiple chains of our Metropolis within Gibbs simulator from different starting values (drawn randomly from the prior). To check that these multiple chains agree in their characterization of the posterior distribution, we look at various sample
moments within and across chains. Several chains of different length initialized in this manner delivered roughly identical results when looking at means, medians, and posterior percentiles, as well as trace and kernel plots.

More formally, for the multiple chains used to generate the results in the paper, Table 6 reports potential scale reduction factors proposed by Stephen P. Brooks and Andrew Gelman (1998) both for variances and 90 percent intervals (first two columns). These numbers are very close to one and therefore well below the 1.2 benchmark widely used in practice as an upper bound for convergence. In addition, the last two columns test for the equality of means using Geweke’s (1992) estimator on the initial 20 percent and last 50 percent of the sample, using two alternative estimators for the serially correlated variance of the draws. At the 95 percent confidence level, we cannot reject the null hypothesis of equal means for all but one coefficient. For this single case, trace plots suggest that this is due to a few outliers in one of the chains, while means and medians across chains differ by less than 0.01.

**Appendix C: The Solution Method under Indeterminacy**

In order to solve the model, we log-linearize (1) and obtain

\[ \hat{G}_0(\theta)\hat{y}_t = \hat{G}_1(\theta)\hat{y}_{t-1} + \hat{G}_2(\theta)E_t\hat{y}_{t+1} + \hat{G}_3(\theta)\hat{h}_t, \]

where the “hat” denotes log deviations from the nonstochastic steady state, and \( \hat{G}_0, \hat{G}_1, \hat{G}_2 \) and \( \hat{G}_3 \) are matrices conformable to \( \hat{y} \) and \( \hat{h} \). We define a vector of endogenous forecast errors \( (\omega) \) and extend the vector of endogenous variables to include the expectational variables. Denoting this extended vector by \( \xi_t \), we can put the system in Sims’s (2002) canonical form:

\[ \Gamma_0(\theta)\xi_t = \Gamma_1(\theta)\xi_{t-1} + \Psi(\theta)\hat{h}_t + \Pi(\theta)\omega_t. \]

We can now solve for the endogenous forecast error with the methodology of Lubik and Schorfheide (2003, 2004) and write the solution of the model as

\[ \xi_t = \underbrace{A(\theta)\xi_{t-1} + B(\theta)\hat{h}_t}_{\text{part I}} + \underbrace{C(\theta)(\hat{M}\hat{h}_t + ss_t)}_{\text{part II}}. \]

Part I of this expression corresponds to the solution of the model under determinacy, which is completely pinned down by the structural coefficients of the model \( (\theta) \). Part II of the expression is the additional term that appears and parameterizes the continuum of equilibria under indeterminacy. Note that it includes a modification of the transmission mechanism of the fundamental shocks (parameterized by \( \hat{M} \)), as well as the effect of a sunspot shock \( (ss_t \sim N(0, \sigma^2_s)) \).

Details regarding the derivation of the solution and the interpretation of the matrices can be found in Lubik and Schorfheide (2004).

In the indeterminacy region, we consider only draws for which the degree of indeterminacy is one, and we discard all other draws. In this case, the dimensions of \( \hat{M} \) and \( \sigma_s \) are \( n \times 1 \) and \( 1 \times 1 \), respectively.

An important issue is, of course, the specification of a prior density for these additional free parameters. For \( \sigma_s \), we choose an inverse-Gamma prior with mean and standard deviation equal to 0.15. For \( \hat{M} \), we follow Lubik and Schorfheide (2004) and center our prior on \( \hat{M}^*(\theta) \), where \( \hat{M}^*(\theta) \) denotes the continuity solution, i.e., the value of \( \hat{M} \) for which the impact of fundamental shocks on endogenous variables is continuous at the boundary of the indeterminacy region. For a given parameter value \( \theta \) leading to indeterminacy, \( \hat{M}^*(\theta) \) is found by minimizing (using a least
squares criterion) the distance between the impact matrix under indeterminacy and the one at
the boundary of the determinacy region. Because of the dependence of \( M^* \) on \( u \), it is convenient
to center our prior on the continuity solution by, instead, specifying a zero mean prior
with standard deviation equal to 0.5 for \( M^* \). Hence, for each indeterminate draw, we
must find the corresponding parameter value at the boundary of the determinacy region. We
do this by gradually increasing \( \phi_u \), while keeping all other coefficients fixed until we obtain a
determinate solution. The shape of the determinacy and indeterminacy regions of our parameter
space is slightly more complex than in simpler three-equation New-Keynesian models. Therefore,
in very few cases, this procedure does not generate a determinate solution and we discard the
draw. For the model in which all coefficients are allowed to change

| Potential scale reduction factors (PSRF) for multiple chains below 1.2 or 1.1 are regarded as indicative of convergence. The first column reports PSRF using within and between variances. The second column reports PSRF based on empirical 90 percent interval lengths (length of total-sequence interval to mean length of within-sequence intervals), as proposed by Brooks and Gelman (1998). |
| \( p \)-values for the null hypothesis of equal means based on the initial 20 percent and the latter half of the draws when pooling all individual chains. See Geweke (1992) for details. We compute the corresponding Wald-statistics and associated \( p \)-values using batched means (Chib 2001) and autoregressive spectral estimates for the variance of the (serially correlated) draws. These alternative estimates of the long-run variance yield an overall identical picture regarding the equality of means for the pooled chain.

### REFERENCES


