

ESSAYS ON MACROECONOMETRICS

Kotbee Shin

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

Frank Schorfheide

Professor of Economics

Graduate Group Chairperson

George J. Mailath

Professor of Economics, Walter H. Annenberg Professor in the Social Sciences

Dissertation Committee

Cecilia Fieler, Professor of Economics

Urban Jermann, Professor of Finance

ESSAYS ON MACROECONOMETRICS

COPYRIGHT

2015

Kotbee Shin

To my beloved family

Acknowledgements

I am immensely indebted to my advisor, Frank Schorfheide. His insightful advice and support have inspired my research and helped me strive through every stage of my graduate studies. I am especially grateful that he has been patient with me even when I made painfully-slow progress. I would also like to thank my dissertation committee, Cecilia Fieler and Urban Jermann. Their valuable comments have helped shape my research greatly. I have also benefited from seminar participants at the University of Pennsylvania. I owe thanks to my undergraduate advisor, Chang-Jin Kim. His enthusiasm in economics motivated me to become a passionate researcher.

I was so fortunate to be surrounded by smart colleagues and friends at McNeil. I am also thankful to Kyungmin Kim, Soojin Kim, Eunice Yang, Eun-young Shim, Cezar Santos, Dave Weiss, Chen Han, Sophie Shin, and Minchul Shin for their friendship. Special thanks go to Dongho Song for substantive conversations about research.

Last and most importantly, I would like to thank my family. I cannot thank enough to my parents and parents-in-law for their unconditional love and sacrifice. My deepest gratitude goes to my husband, Kihwan. Without his love and encouragement, I could not have concluded this endeavor. Also I thank my daughter, Elyse to fill my life with great joy and love.

ABSTRACT

ESSAYS ON MACROECONOMETRICS

Kotbee Shin

Frank Schorfheide

This dissertation presents two essays on macroeconometrics. In the second chapter, I empirically compare alternative specifications of time-varying volatility in the context of linearized dynamic stochastic general equilibrium models. I consider time variation in the volatility of structural innovations in two ways: one in which the logarithm of the volatility is assumed to follow a simple autoregressive process (stochastic volatility) and the other in which the volatility follows a Markov-switching process. A comprehensive simulation study is presented to assess the fit and performance of two specifications. I show that modeling heteroscedasticity in a highly synchronized fashion across shocks may lead to distorted estimation of the volatility. In the empirical application to the United States data, stochastic volatility model delivers the best-fit and accounts for the heteroscedasticity present in the data well.

In the third chapter, I conduct a quantitative evaluation of the potential role of adaptive expectations in a two-country dynamic stochastic general equilibrium model. Under the learning mechanism economic agents are assumed to form their expectations of forward-looking variables using a simple vector autoregressive forecasting model. The agents estimate their vector autoregression based on past model variables and update the estimates every period via a constant gain learning algorithm. I show in a simulation study that the learning mechanism increases the volatility and persistence of the endogenous variables and that as the constant gain parameter grows larger, so do these increases. The two-country DSGE model is then estimated with data from the United States and Euro area. A comparison based on log marginal

data densities favors the learning over the rational expectations specification. The learning mechanism generates more persistent responses of variables to the monetary shocks. The improvement in terms of fitting the observed Dollar-Euro exchange rate dynamics is limited.

Contents

Contents	vii
List of Tables	viii
List of Figures	ix
I Introduction	1
II Regime Switching and Stochastic Volatility in DSGE Models	4
1 Introduction	4
2 A Benchmark Model	9
2.1 Firms	9
2.2 Households	11
2.3 Policy	14
2.4 Market Clearing	14
2.5 Exogenous Stochastic Process	14
2.6 Steady State and Model Solution	15
3 Simulation Study	16
3.1 Simulation Design	16
3.2 Model Comparison	17
4 Application to U.S. data	22

4.1	Estimation Approach	22
4.2	Estimation Results	23
5	Conclusion	27
6	Tables and Figures	28
 III Bayesian Estimation of a New Open Economy Model with Adaptive Expectations		 50
7	Introduction	50
8	The Model	53
8.1	Domestic Households	54
8.2	Domestic Producers	55
8.2.1	Domestic Final Producers	55
8.2.2	Domestic Intermediate Producers	58
8.2.3	Domestic Importers	60
8.3	Policy	61
8.4	Foreign Economy	62
8.5	Market clearing	64
8.6	Exogenous Stochastic Process	64
9	Learning Model	65
9.1	Learning Algorithm	65
9.2	Exchange Rate Determination under Learning	66
10	Simulation Study	68

11 Empirical Application	70
11.1 Posterior Distribution	72
11.2 Marginal Data Density Comparison	73
11.3 Impulse Response Function	74
11.4 Variance Decomposition	75
11.5 Posterior Predictive Checks	75
12 Conclusion	76
13 Tables and Figures	77
A Appendices	90
A.1 Estimation Algorithm	90
A.1.1 Stochastic Volatility in DSGE Models	90
A.1.2 A Four Regime-Switching in DSGE Models	94
Bibliography	97

List of Tables

1	Summary of Simulation Study	28
2	Log Median Likelihood	28
3	Summary of Prior Densities and Posterior Estimates	29
4	Summary of Parameters for Simulation	78
5	Standard Deviation of Simulated Models	79
6	Correlation of Depreciation to Inflation and Interest Rate Gap	79
7	Prior and Posterior Distribution	82
8	Marginal Data Density	83
9	Variance Decomposition from the Rational Expectations Model	83

List of Figures

1	True and Estimated Volatilities for the model SM1	30
2	True and Estimated Volatilities for the model SM2	31
3	True and Estimated Volatilities for the model SM3	32
4	Variance Decomposition for the model SM1	33
5	Variance Decomposition for the model SM2	34
6	Variance Decomposition for the model SM3	35
7	True and Estimated Volatilities for the model FM1	36
8	Posterior Distributions of non-Volatility Parameters: FM1	37
9	True and Estimated Volatilities for the model FM2	38
10	Posterior Distributions of non-Volatility Parameters: FM2.	39
11	Posterior Probability of the High Volatility Regime: FM2	40
12	Rolling Standard Deviations for U.S. Data	41
13	Estimated Standard Deviations: SV-DSGE	42
14	Estimated Standard Deviations: RS(2)-DSGE	43
15	Estimated Standard Deviations: RS(4)-DSGE	44
16	Variance Decomposition for U.S. Data	45
17	Posterior Probability of the High Volatility Regime: RS(2)-DSGE	46
18	Posterior Probability of the High Volatility Regime: RS(4)-DSGE	47
19	Posterior Probability: RS(4)-DSGE	48
20	Posterior Density of High- and Low- Volatility Regime Duration	49
21	Autocorrelation Function of the Real Exchange Rate from Simulation	80
22	Depreciation and Inflation, Interest Rate Gap	81
23	Impulse Response to U.S. Monetary Policy Shock for the Learning Model	84
24	Impulse Response to Euro Monetary Policy Shock for the Learning	

	Model	85
25	Impulse Response to U.S. Monetary Policy Shock	86
26	Impulse Response to Euro Monetary Policy Shock	87
27	Variance Decomposition for Depreciation in the Learning Model	88
28	Posterior Predictive Checks	89

Chapter I

Introduction

The estimation of dynamic stochastic general equilibrium (DSGE) models has been an important subject in macroeconomics. DSGE models are useful to understand the propagation mechanism of structural shocks to business cycle fluctuations and provide a tool for the quantitative analysis of policy experiments. Over the past few decades, there has been remarkable advance in theoretical and empirical DSGE models. DSGE models with various frictions and different types of shocks have been developed and they seem to reproduce the key features of data well in many dimensions. Despite the significant progress in DSGE model literature, there remains open issues on which specifications DSGE model should take to characterize macroeconomic observations. This paper aims to contribute to this literature. In this dissertation, I investigate and empirically compare the specifications of DSGE models to answer two macroeconomic questions: (1) what is the better specification between regime switching model and stochastic volatility model to represent the overall volatility reduction, quoted as “Great Moderation”?; (2) can adaptive expectations instead of rational expectations explain the exchange rate dynamics in the general equilibrium framework better?

In the second chapter, I compare the regime switching models and stochastic volatility models. Regime switching models have been a standard approach to identify the key source of a large decline in aggregate volatilities. The proponents of regime switching approach find it appealing since it is a parsimonious way of modeling the discrete jumps and can be potentially related to economic regimes with meaningful interpretations. An alternative time-varying volatility model is the stochastic volatility model that allows the volatility processes randomly fluctuate through time. Stochas-

tic volatility model is distinct from the regime switching model in two ways. First, while the volatility discontinuously shifts from one level to another in the regime switching model, it continuously changes with persistence in the stochastic volatility model. Second, the timing of volatility shifts across innovations is restricted in regime switching models, but stochastic volatility models can allow the independent movement of volatility processes. It is of importance to take a deeper look at the empirical performance of these two modeling approaches since a different specification of variance could lead to a different conclusion on the source of macroeconomic fluctuations. To do so, I simulate the large-scale DSGE models with regime switching volatility and with stochastic volatility by assuming several possible episodes of underlying data generation processes for volatility dynamics. The simulation study shows that the stochastic volatility specification provides results comparable to or better than regime switching models regardless of the underlying volatility patterns.

In the third chapter, different approaches to modeling the expectation formation is explored in the open economy context. Standard open economy DSGE models with rational expectations have had challenges in the exchange rate determination because the exchange rates are too volatile and persistent to be justified by economic fundamentals. The empirical shortcoming of rational expectations models arise from the tight link between the exchange rates and economic fundamentals.

Researchers have attempted to modify the open economy models with different ingredients to relax the link of the exchange rates from the rest of the economy. In this chapter, I focus on the expectation formation mechanism. Rational expectations hypothesis assume that agents have complete knowledge of economic environment and have model consistent expectations. By relaxing this strong informational assumption, adaptive expectations has been of growing interest to study many macroeconomic observed behaviors that has been hard to reconcile in the closed economy lit-

erature.* Under the adaptive expectations mechanism, economic agents are assumed to form subjective expectations with limited information and learn the structure of the economy over time. Since the agents' learning process relax the tight restrictions on the relationships of model variables by forecast biases, the short-run dynamics of the model variables under adaptive expectations could be substantially different from those under rational expectations.

To quantify the role of adaptive expectations in the open economy model, I conduct a simulation study and estimate the model using Bayesian techniques. A simulation shows that adaptive expectations mechanism increases the volatility and persistence of endogenous variables and allows the exchange rates to drift away from the uncovered interest parity equation. The estimation results provide evidence that the adaptive expectations is a potential channel to explain the persistence of variables. I also find that adaptive expectations also improve the fit of data.

*For example, see Bullard and Mitra (2001), Evans and Honkapojha (2001).

Chapter II

Regime Switching and Stochastic Volatility in DSGE Models

†

1 Introduction

In recent years, economists have produced a collection of methods to account for heteroscedasticity present in the U.S. aggregate data. The most notable example is the “Great Moderation” episode when the U.S. economy experienced a general reduction in macroeconomic volatilities. A branch of macro literature has presented empirical evidences in favor of heteroscedasticity in the shock variances. (Kim and Nelson, 1999; McConnell and Perez-Quiros, 2000; Stock and Watson, 2002; Sensier and van Dijk, 2004; Koopman, Lee, and Wong, 2006).

One class of time-varying models are Markov switching models. These models allow the time series to be in any of a finite number of distinct regimes, see Hamilton (1989) and Kim and Nelson (1999). The choice is attractive because of the parsimonious flexibility it provides in the specification of the distributions of the underlying structural shocks and requires fewer parameters to estimate. Markov switching models are especially appealing for characterizing the Great Moderation if there has been a discrete and comprehensive volatility reduction in macroeconomic variables (Stock and Watson, 2002; Chauvet and Potter, 2001; Sensier and van Dijk, 2004). Initiated

†This chapter is based on the joint work with Dongho Song.

by Hamilton (1989), Markov switching models are used extensively in business cycle analysis to characterize discrete changes in the volatilities (Sims and Zha, 2006; Davig and Doh, 2009; Liu, Waggoner, and Zha, 2010).

A more flexible specification is considered in stochastic volatility models, which allow the continuous change of volatility processes having the potential for moving one or two steps closer to complex reality. An advantage of stochastic volatility specification is that it can characterize continuous shifts in variance and does not restrict the system to switch between the same configuration. Methodologically, it is related to the statistics literature on stochastic volatility models (Jacquier, Polson, and Rossi, 1995; Kim, Shephard, Chib, 1998; Chib, Nardari, and Shephard 2006), but the recent contribution of Fernandez-Villaverde and Rubio-Ramirez (2007) and Justiniano and Primiceri (2008) indicates that stochastic volatility in general equilibrium models has been exploited in the business cycle literature. If the degree of time variability differs across volatility processes and the structural disturbances hitting the economy display substantial stochastic volatility, stochastic volatility specification is an effective way to accommodate changes in the volatility of the U.S. economy (Cogley and Sargent, 2005; Primiceri, 2005; Fernandez-Villaverde and Rubio-Ramirez, 2007; Justiniano and Primiceri, 2008; Creal, Koopman, and Zivot, 2010).

While numerous studies have found significant time variation in shock variances, the magnitude, the number of structural breaks, as well as the underlying causes of the Great Moderation still remain as one of the main open questions in macroeconomics. A lively debate unfolded between proponents of sudden change in volatility (Stock and Watson, 2002) versus gradual reduction in volatility (Blanchard and Simon, 2001). Some people are also concerned whether one or more structural breaks exist (see the discussion in Sensier and van Dijk, 2004). However, much of the disagreement also comes from the differences in the model framework and in the empirical

approach. In the absence of actual knowledge of the underlying structure of volatility processes, it is often difficult to decide which estimation algorithm is the preferred route to pursue. What is left unsaid in the literature is how model-dependent the conclusions are when identifying the sources of macroeconomic fluctuations. In more general terms, good modeling practice requires investigation of the robustness of a conclusion when the study includes some form of economic modeling. To my best knowledge, not much research has been done to minimize the sensitivity to model-dependent analyses of the sources of macroeconomic fluctuations by estimating a variety of structural models that assumes time-varying shock variances.

This provides a clear motivation to investigation. This paper incorporates time varying volatility structure in large-scale linearized DSGE economies and compares the Markov Regime-Switching (henceforth RS-DSGE) and Stochastic Volatility DSGE (henceforth SV-DSGE) models using both simulation study and empirical application with a strong emphasis on the specification of volatility dynamics. The goal of this paper is to provide a systematic examination of the performance of two competing models in explaining the driving sources of macroeconomic fluctuations. Investigation is important since macroeconomic implications can be seriously distorted if the competing models produce different volatility estimates. First, I investigate whether the outcomes of two models are similar and second, which model is more reliable. I believe that only after taking account of the model sensitivity, it is possible to draw a firm conclusion about the sources of macroeconomic fluctuations. This paper also tries to address the danger of relying exclusively on a model selection criterion that favors models that fit the data well. A common practice in the empirical DSGE literature is comparing models using marginal likelihood. From a Bayesian perspective, the marginal data density is the most comprehensive and accurate measure of fit and is needed for the comparison of non-nested Bayesian models. I illustrate with the

aid of simulation examples and empirical application that a situation where volatility dynamics are spuriously estimated but it survives the Bayesian model selection criterion for offering a parsimonious approximation and delivering a better time-series fit is possible

I propose a simulation study. The architecture of simulation study is designed to replicate salient features of U.S. business cycles and is implemented by using artificial dataset of 200 observations generated using a large scale DSGE model of Justiniano and Primiceri (2008) (henceforth, JP). The details are discussed in Section 3. Application to the simulated dataset will provide guidance on how well each competing model performs when the true model is in hand. In a simulation study, model performances are measured in three different ways. I compare the estimated volatility components to the true Data Generating Process (henceforth DGP), perform variance decomposition to examine the consequence of volatility misspecification, and compute the log marginal data density to measure the data fit. Next, I repeat the same steps with the aggregate U.S. data and use the simulation results as a benchmark to understand and interpret the empirical performance of each model.

The main empirical findings in the experiments are as follows. A two regime-switching DSGE (henceforth, RS(2)-DSGE) model seems inappropriate for drawing inferences about the volatility processes of business cycles. A common disadvantage of RS(2)-DSGE models is that they assume complete synchronization of Markov states across volatilities. Estimated volatilities can be very crude if the true DGP exhibits contrasting patterns of fluctuations. The natural step is to amend RS(2)-DSGE to allow additional degree of flexibility in the movement of volatilities. A four regime-switching DSGE (henceforth RS(4)-DSGE) model nests RS(2)-DSGE in this regard. Hence, I argue that RS(4)-DSGE may perform better than RS(2)-DSGE model for drawing inference about the volatility processes. Due to technical limitations, provid-

ing extra degree of flexibility in the RS-DSGE model is a challenging task. SV-DSGE model, on the contrary, promises great flexibility in modeling volatility dynamics and its performance is certainly not inferior to RS-DSGE's. However, the price for this flexibility is an increase in dimension. Aside from the over-parametrization problem, SV-DSGE estimates may exaggerate or discount time variation and stochastic movement in volatility. Its performance may deteriorate as the oscillation between volatility regimes increases and the difference between the regimes decreases.

In the application to U.S. aggregate data, I estimate the large scale DSGE model in JP using RS(2)-DSGE, RS(4)-DSGE, and SV-DSGE models. I have grouped a subset of shock variances having the same Markov processes. I allow regime associated with the variances of the monetary policy shock to be independent of the regime switching processes of the other shock variances. The empirical motivation is from Sensier and van Dijk (2004) and Cecchetti, Hooper, Kasman, Schoenholtz, and Watson (2007) that the volatility of inflation has undergone multiple structural breaks. I find that SV-DSGE model delivers a best-fit and accounts for the heteroscedasticity present in the data well. RS(4)-DSGE performs better than RS(2)-DSGE model in volatility specification, but does not improve upon the data fit. By synchronizing shifts in variances across two regimes, RS(2)-DSGE model may detect the timing of the volatility regime wrong and provide imprecise estimates for volatility processes while fitting to the data better than RS(4)-DSGE model. In order to provide robustness of empirical findings, I perform the variance decomposition and compute the marginal data density using the modified harmonic mean method of Geweke (1999).

Liu, Waggoner, and Zha (2010) is the closest paper to ours. They examine the sources of macroeconomic fluctuations by estimating a number of alternative regime-switching models using Bayesian methods in a unified DSGE framework and compare the fit to the time series data in the post war U.S. economy. Based on marginal data

density and Schwarz criterion, they find strong evidence in favor of the RS(2)-DSGE model where regime shifts in the variances are synchronized. They then use RS(2)-DSGE model to identify shocks that are important in driving macroeconomic fluctuations. My approach differs from Liu et al. (2010) since I extend the investigation by including stochastic volatility specification and examine the effectiveness of Bayesian model-selection criterion.

This chapter is organized as follows. Section 2 briefly explains the DSGE model framework, Section 3 constitutes a simulation study. Section 4 illustrates an application to U.S. aggregate data and Section 5 concludes. Technical details are summarized in Appendix.

2 A Benchmark Model

The model is based on JP and exhibits a number of real and nominal rigidities which has been shown to fit the data fairly well. For additional details, see JP. The basic elements of the model include a continuum of households, perfectly competitive final goods producers, and a continuum of monopolistic intermediate goods producers. Monetary policy follows a Taylor type rule and fiscal policy is assumed to be fully Ricardian. Here is the illustration of the JP model.

2.1 Firms

A monopolistic intermediate goods producing firm $i \in [0, 1]$ produces output according to:

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

where A_t is an exogenous measure of productivity that is the same across firms and F represents a fixed cost of production. As usual, $K_t(i)$ and $L_t(i)$ denote, respectively, the capital and labor input for the production of good i . A_t follows a unit root process, with a growth rate ($z_t \equiv \log \frac{A_t}{A_{t-1}}$) that follows:

$$z_t = (1 - \rho_z)\gamma + \rho_z z_{t-1} + \sigma_{zt} \varepsilon_{zt} \quad (1)$$

Firms follow a Calvo pricing mechanism when they set their prices. At the start of each period, a randomly selected fraction ξ_p of firms cannot reoptimize and set their prices according to:

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

where π_t is defined as $\frac{P_t}{P_{t-1}}$ and π is the steady-state value. Remaining fraction $1 - \xi_p$ of firms choose their prices by maximizing the present value of future profits:

$$E_t \sum_{s=0}^{\infty} \xi_p^s \beta^s \lambda_{t+s} \left\{ [\tilde{P}_t(i) (\prod_{j=0}^s \pi_{t-1+j}^{\iota_p} \pi^{1-\iota_p})] Y_{t+s}(i) - [W_{t+s} L_{t+s}(i) + R_{t+s}^k K_{t+s}(i)] \right\}$$

where λ_{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.

There is a representative final goods producing firm that produce the consumption goods using the intermediate goods and the following constant-returns-to-scale technology:

$$Y_t = \left[\int_0^1 Y_t(i)^{\frac{1}{1+\lambda_{pt}}} di \right]^{1+\lambda_{pt}}$$

where λ_{pt} follows the exogenous stochastic process

$$\log \lambda_{pt} = (1 - \rho_p) \log \lambda_p + \rho_p \log \lambda_{pt-1} + \sigma_{pt} \varepsilon_{pt} \quad (2)$$

Profit maximization problem for the final goods producer yields a demand for each intermediate good given by

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\frac{(1+\lambda_{pt})}{\lambda_{pt}}} Y_t$$

and the zero profit condition imply

$$P_t = \left[\int_0^1 P_t(i)^{-\frac{1}{\lambda_{pt}}} di \right]^{-\lambda_{pt}}.$$

2.2 Households

Firms are owned by a continuum of households, indexed by $j \in [0, 1]$. As in JP, while each household is a monopolistic supplier of specialized labor, a number of "employment agencies" combine households' specialized labor into labor services available to the intermediate firms:

$$L_t = \left[\int_0^1 L_t(j)^{\frac{1}{1+\lambda_w}} dj \right]^{1+\lambda_w}$$

Profit maximization problem for the employment agencies yields a demand for each labor given by

$$L_t(j) = \left(\frac{W_t(j)}{W_t} \right)^{-\frac{(1+\lambda_w)}{\lambda_w}} L_t$$

and the zero profit condition imply

$$W_t = \left[\int_0^1 W_t(j)^{-\frac{1}{\lambda_w}} dj \right]^{-\lambda_w}.$$

Household j 's preferences are representable by a lifetime utility functions:

$$E_t \sum_{s=0}^{\infty} \beta^s b_{t+s} \left\{ \log(C_{t+s}(j) - hC_{t+s-1}(j)) - \varphi_{t+s} \frac{L_{t+s}(j)^{1+\nu}}{1+\nu} \right\}$$

which is separable in consumption, $C_t(j)$, and labor $L_t(j)$. h is the degree of habit formation, φ_t is a preference shock that affects the marginal disutility of labor, and b_t is a discount factor shock affecting both marginal utility of consumption and the marginal disutility of labor. Both shocks follow exogenous stochastic processes

$$\log b_t = \rho_b \log b_{t-1} + \sigma_{bt} \varepsilon_{bt} \quad (3)$$

$$\log \varphi_t = (1 - \rho_\varphi) \log \varphi + \rho_\varphi \log \varphi_{t-1} + \sigma_{\varphi t} \varepsilon_{\varphi t}. \quad (4)$$

The j^{th} household's budget constraint is given by:

$$\begin{aligned} P_{t+s} C_{t+s}(j) + P_{t+s} I_{t+s}(j) + B_{t+s}(j) &\leq R_{t+s-1} B_{t+s-1}(j) + Q_{t+s-1}(j) + \Pi_{t+s} \\ &+ W_{t+s}(j) L_{t+s}(j) + R_{t+s}^k(j) u_{t+s}(j) \bar{K}_{t+s-1}(j) - P_{t+s} a(u_{t+s}(j)) \bar{K}_{t+s-1}(j) \end{aligned}$$

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

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

which is rented to firms at the rate $R_t^k(j)$. The cost of capital utilization is $a(u_{t+s}(j))$

per unit of physical capital. Following JP, I assume $u_t = 1$ and $a(u_t) = 0$ in steady state. In this partially nonlinear approximation of the model solution, only the curvature of the function in steady state needs to be specified, $\chi = \frac{a''(1)}{a'(1)}$. The usual physical capital accumulation equation is described by

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

where δ denotes the depreciation rate and S captures the investment adjustment cost, with $S' > 0$ and $S'' > 0$. μ_t can be interpreted as an investment-specific technology shock following Greenwood, Hercowitz, and Krusell (1997). Assume that this investment shock follows the exogenous stochastic process

$$\log \mu_t = \rho_\mu \log \mu_{t-1} + \sigma_{\mu t} \varepsilon_{\mu t}. \quad (5)$$

Households follow a Calvo pricing mechanism when they set wages. At the start of every period, a randomly selected ξ_w of households cannot reoptimize wages and set their wages according to the indexation rule:

$$W_t(j) = W_{t-1}(j)(\pi_{t-1}e^{z_{t-1}})^{\iota_w}(\pi e^\gamma)^{1-\iota_w}.$$

The remaining $1 - \xi_w$ of households set their wages by maximizing

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

subject to

$$L_t(j) = \left(\frac{W_t(j)}{W_t} \right)^{-\frac{(1+\lambda w)}{\lambda_w}} L_t.$$

2.3 Policy

The monetary authority sets the short-term nominal rate using the following rule

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi}\right)^{\phi_\pi} \left(\frac{Y_t/A_t}{Y/A}\right)^{\phi_Y} \right]^{1-\rho_R} e^{\sigma_{Rt}\varepsilon_{Rt}}, \quad (6)$$

where R is the steady state for the gross nominal interest rate and ε_{Rt} is a monetary policy shock.

Fiscal policy is, by assumption, 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 exogenous stochastic process

$$\log g_t = (1 - \rho_g) \log g + \rho_g \log g_{t-1} + \sigma_{gt}\varepsilon_{gt}. \quad (7)$$

2.4 Market Clearing

The resource constraint is given by

$$C_t + I_t + G_t + a(u_t)\bar{K}_{t-1} = Y_t$$

2.5 Exogenous Stochastic Process

With regard to exogenous stochastic process in (1)-(7), $\varepsilon_{jt} \sim iid N(0, 1)$ where $j \in \{z, p, b, \varphi, \mu, R, g\}$ and t denotes time. As for the standard deviations, $\sigma_{j\tau}$, I assume that τ is t in stochastic volatility model and $\tau \in \{\#regimes\}$ in regime switching

models. For instance, $\tau \in \{High\ volatility\ regime, Low\ volatility\ regime\}$ in two Markov regime-switching model. Following Kim, Shepard, and Chip (1998), I assume that each element of σ_{jt} evolves independently according to the following stochastic processes:

$$\begin{aligned} \log \sigma_{jt} &= (1 - \rho_{\sigma_j}) \log \sigma_j + \rho_{\sigma_j} \log \sigma_{jt-1} + \nu_{jt} \\ \nu_{jt} &\sim^{iid} N(0, \omega_j^2) \end{aligned}$$

2.6 Steady State and Model Solution

Since technology process A_t is a unit root process, after detrending consumption, investment, capital, real wages, and output, I am able to compute the nonstochastic steady state. Define the vector of relevant model endogenous variables α_t . Solution of the linear rational expectations system, $E_t[f(\alpha_{t+1}, \alpha_t, \alpha_{t-1}, \eta_t, \theta)] = 0$, where η_t is a vector of exogenous disturbances and θ is a vector of structural parameters, is obtained by running Chris Sims's code gensys.m. Then, the observable $y_t = [\Delta \log Y_t, \Delta \log C_t, \Delta \log I_t, \log L_t, \Delta \log \frac{W_t}{P_t}, \pi_t, R_t]$ can be expressed as a linear function of the endogenous model variables x_t

$$y_t = D + Z\alpha_t \tag{ME}$$

$$\alpha_t = T(\theta)\alpha_{t-1} + R(\theta)\eta_t \tag{TE}$$

Further details regarding the model solution is discussed in the appendix.

3 Simulation Study

3.1 Simulation Design

In this section I propose a simulation study for the performance evaluation of RS(2)-DSGE, RS(4)-DSGE, and SV-DSGE model.

I consider three simulation scenarios in which the key qualitative features of the Great Moderation are replicated. In the first set of simulation, labeled as SM1, whereas monetary shock displays a double hump-shaped pattern. A double hump-shaped pattern of monetary shock has its ground in the growing literature on Markov switching DSGE models (see, for more information, Davig and Doh, 2009; Bianchi, 2010; Liu, Waggoner, and Zha, 2010). All other structural shocks exhibit a one time permanent reduction in the volatility. Many empirical studies document a discrete drop in conditional variance of macroeconomic time series. For example, Stock and Watson (2002) argue that changes in the volatility around 1984 are comprehensive and best characterized as discrete break. The second simulation scenario (henceforth, SM2) complexifies volatility process one step further by allowing a subset of volatilities to have independent regimes. By doing so, I are able to examine the ability of each estimation method to recover the true volatility specification. The last simulation scenario (henceforth, SM3) is based on the argument that the high volatility regime is often associated with recession (French and Sichel, 1993; Hamilton and Susmel, 1994). Since there are relatively fewer number of recessions after 1990s, this may appear as volatility reduction. I let high volatility regimes manifest themselves at the NBER recession dates.

In all scenarios, I assume the true DGP to follow Markov processes since the volatility dynamics can be easily configured to replicate salient features of Great Moderation. Due to substantial computational demand and time constraint, during

a simulation study I shut down Metropolis-Hastings algorithm and give the true parameter values for non-volatility parameters. A full-blown estimation requires roughly three (two) days to complete 200,000 iteration for RS-DSGE (SV-DSGE) models which is the minimum of iteration number to achieve convergence. Since I are interested in the estimation of volatility dynamics, without loss of generality, I assume true values for non-volatility parameters are recovered through Metropolis-Hastings algorithm. All simulations are based on artificial dataset generated from a large scale DSGE model of JP and the posterior median estimates for non-volatility parameters in JP are used.

Next, I conduct some full-blown estimation to check the validity of the assumption that true values for non-volatility parameters are recovered through Metropolis-Hastings algorithm. I extend SM1 to a full-blown estimation scheme and denote as FM1. I also consider a case (henceforth, FM2) in which the true DGP follows stochastic volatility process. Here, I use the posterior volatility estimates of SV-DSGE in JP as DGP and try to estimate with RS(2)-DSGE model. This exercise centers on the following idea: if the true DGP of macroeconomic variables follow stochastic volatility process, how well RS(2)-DSGE models estimate the volatility dynamics? Table 1 summarizes the simulation study.

3.2 Model Comparison

Model performance is measured in three different ways; I compare the estimated volatility components to the true DGP, perform variance decomposition to examine the consequence of volatility misspecification, and compute the log marginal data density to measure the data fit. All figures are obtained from the remaining 10,000 posterior draws after discarding the initial 40,000 draws. Figure 1 through figure 3

plot the time-varying standard deviations for the structural shocks of each model. Top panels of figure 1, figure 2 and figure 3 juxtapose the volatility estimates from RS(4)-DSGE, SV-DSGE model, and true DGP. Bottom panels are constructed similarly but carry the estimates from RS(2)-DSGE model. Figure 4 to figure 6 present the evolution of the variance shares of GDP growth attributed to each structural shock.

Figure 1 is based on SM1. Note that the true DGP and the estimates from RS(4)-DSGE model are almost identical in the top panel. However, in the bottom panel, the estimated monetary shock demonstrates a one-time reduction in volatility as the other shocks do. This is inevitable because RS(2)-DSGE assumes all shock variance to switch regimes simultaneously. Due to limited space, I do not report in this paper but depending on the parameterization, I have cases where all estimated shock variances follow a double hump-shaped. This double hump-shaped pattern has been consistently documented in previous literatures (see Davig and Doh, 2009; Bianchi, 2010; Liu, Waggoner, and Zha, 2010). From this example, I cast doubt on the estimated volatility components from the RS(2)-DSGE model. I believe that when the model assumes synchronized regime shifts in the variances, it is more likely to produce spurious estimates. SV-DSGE model captures the time variation in volatility well.

Some interesting findings are shown in figure 2. I slightly modify SM1 by allowing independent regimes for government spending and price mark-up shocks. It is now called SM2. The estimates from RS(2)-DSGE, presented in the bottom panel, are misleading in that they spuriously detect the number of structural breaks in volatility. They can be very crude as the true DGP exhibits contrasting patterns of fluctuations. Considering that the number of structural breaks in macroeconomic time series is still controversial, this result deserves attention. RS(4)-DSGE allows additional degree of flexibility in the movement of volatilities, and hence performs better in capturing

volatility movements than RS(2)-DSGE. One remark is that when grouping a subset of shock variances to have the same Markov processes, I have to rely on empirical evidences to find the right combination. I will discuss this in more detail in the application to real data.

A common disadvantage of RS-DSGE models is that they assume perfect synchronization of Markov states across volatilities. SM1 and SM2 show how easily RS-DSGE model can be misleading when the evolutions of shock volatilities are separated. On the other hand, SV-DSGE performs quite well in both scenarios. Since I employ random-walk specification for the stochastic volatility processes, SV-DSGE can account for both a gradual decline and a sudden change in volatility.

In SM3, I try to identify cases when SV-DSGE model performs poorly as shown in figure 3. SV-DSGE estimates may exaggerate or discount time variation in volatility. Because SV-DSGE has a tendency to smooth out the patterns, when there are frequent oscillations the estimates can be misleading. I do not report in this paper, but when the differences between the two volatility regimes are small, the posterior credible interval of SV-DSGE widens and the time-invariant volatility hypothesis cannot be rejected.

What can go wrong if the estimated volatility processes are misspecified? I would like to address this issue by performing variance decomposition. This exercise is important since conclusions drawn from the estimated volatility components are dependent on the estimation methods and consequently variance decomposition results will differ. Variance decomposition is obtained by solving the following discrete Lyapunov equations:

$$\begin{aligned} Var(\alpha_t|\theta, Q_t) &= T(\theta)Var(\alpha_t|\theta, Q_t)T(\theta)' + R(\theta)Q_tR(\theta) \\ Var(y_t|\theta, Q_t) &= Z(\theta)Var(\alpha_t|\theta, Q_t)Z(\theta)' \end{aligned}$$

where Q_t is a regime-dependent variance-covariance matrix in RS-DSGE and is per se a stochastic volatility variance-covariance matrix in SV-DSGE. The contribution of shock i is obtained by setting to zero the volatility of all disturbances but one, σ_{it}^2 . Figure 4 through Figure 6 report the evolution of the variance shares of GDP growth attributed to each structural shocks. Variance decomposition results are mostly similar across models. Especially, SV-DSGE and RS(4)-DSGE models produce qualitatively similar results in all scenarios. However, SV-DSGE model slightly exaggerates the role of government spending shock in explaining the variability of GDP growth. The variance decomposition from RS(2)-DSGE relies on the poorly estimated volatilities and tend to exaggerate the role of monetary policy shock.

Table 2 reports the log likelihood for all combination of experiments. Since my simulation study shut down Metropolis-Hastings algorithm for non-volatility parameters, I use log median likelihood in model comparison. Note that I am integrating out the unobserved volatilities nor penalizing the likelihood with the number of parameters, instead I compute $\mathcal{L}(Y|\text{volatilities}; \text{true non-volatility parameters})$. Although this approach will favor models with many parameters, it may be a primitive way to understand the effectiveness of each model. I use marginal data density approach in the empirical application. SV-DSGE model outperforms others with the exception of SM3. RS(2)-DSGE model performs poorly in general, but if there is a synchronized regime switching in shock volatility, RS(2)-DSGE model is the best-fit model. Note that the performance of RS(4)-DSGE model is certainly better than RS(2)-DSGE model in all scenarios. SV-DSGE model delivers best-fit in two out of three scenarios that I considered.

In sum, I argue that when there is not enough knowledge about the volatility process, RS(2)-DSGE model may not be a good choice since it cannot minimize the impact of misspecification for the volatility dynamics. Simulation examples show that

allowing for additional degree of flexibility can cure this problem. SV-DSGE model can be a good candidate since it promises great flexibility in modeling volatility dynamics and delivers data-fit.

A Full-blown Estimation I report the findings from a full-blown estimation. Two sets of full-blown estimations are conducted by generating 200,000 draws. All findings are reported after discarding the initial 150,000 posterior draws. First, I modify the estimation algorithm to make inferences about the non-volatility parameters in SM1 setting. Notice that volatility estimates from SV-DSGE in figure 7 are very similar to those in figure 1. Estimation of non-volatility parameters does not change the inference of volatility processes. Figure 8 shows the kernel density estimation of Bayesian posterior distributions of non-volatility parameters. Except few parameters, the Bayesian credible set includes the true values.

Second, I use the posterior median estimates of volatility processes from JP as the true DGP and estimate with RS(2)-DSGE model. Figure 9 plots the estimated time-varying volatility components. As suggested in the simulation study, RS(2)-DSGE model detects spurious structural break in some volatility components. This is because regime shifts in the variances are synchronized. Figure 10 shows the kernel density estimation of Bayesian posterior distributions of non-volatility parameters. Compared to figure 8, more true parameter values is not contained in the Bayesian credible set. It is not clear at this point what roles volatility specifications play in consistent estimation of non-volatility parameter values. It might be that volatility misspecification affects Kalman gain and in turn compromises the validity of the Kalman filter. Figure 11 displays the posterior expected values of the high-volatility regime probability. This will be explained in more detail in Section 4.

4 Application to U.S. data

4.1 Estimation Approach

I estimate SV-DSGE, RS(2)-DSGE, and RS(4)-DSGE using the same prior distributions and dataset in JP 2008. The data comprises of seven series of U.S. quarterly aggregate variables; the growth rate of output, consumption, investment, real wage, the log of hours worked, annualized inflation, and nominal interest rates (for more details on data description, see JP 2008). I use the same priors for non-volatility parameters across three specifications in order to treat them equal a priori. I refer to Liu, Waggoner, and Zha (2010) for the choice of the prior distributions for the volatility parameters.

As discussed previously, a careful investigation is required to determine how to group a subset of shocks. I allow regime associated with the variance of monetary policy shock to be independent of the regime switching processes of the other shock variances based on previous literatures and on the rolling estimation of standard deviations presented in figure 12. Sensier and van Dijk (2004) find that 83% of the U.S. macroeconomic time series variables have experienced a break in the (un)conditional volatility, and in particular nominal variables such as inflation and interest rates experienced multiple volatility breaks. Cecchetti et al. (2007) report that the level and volatility of inflation display coincident hump-shaped patterns that allow us to date the start of the Great Inflation in the late-1960s and a synchronized Inflation Stabilization in the mid-1980s. The rolling estimation of standard deviations in figure 12 depicts the overall volatility movements. Consistent with Cecchetti et al. (2007), I also witness multiple fluctuations in two nominal variables, inflation and interest rates. This motivates us to assume that the regime associated with nominal shock variances is independent of the regime switching processes of the other shock variances. From

here on, I denote RS(4)-DSGE as the model with the regime associated with monetary shock being independent of other regimes.

Table 3 reports posterior medians and fifth and ninety-fifth percentiles of a model estimated with SV-DSGE, RS(2)-DSGE, and RS(4)-DSGE model. All posterior estimates are obtained by running a single block random-walk MH algorithm (RW-MH) for 400,000 iterations following a burn-in of 350,000 iterations. Calibrated parameters are capital share (α) at 0.3, depreciation rate (δ) at 0.025, SS government spending share (g) at 0.22, and persistent of mark-up shock (ρ_λ) at zero. Chib and Ramamurthy (2010) argue that the results from the RW-MH algorithm are not satisfactory due to slow convergence and often the algorithm does not work in many circumstances. According to Sims, Waggoner, and Zha (2008), due to the complexity inherent in high-dimensional Markov-switching models, the RW-MH algorithm can be very costly and sometimes take a couple of weeks to obtain an estimate that is close to the peak of the likelihood. Indeed, RW-MH algorithm needed around five days to complete 400,000 iterations for RS-DSGE models. I assess the convergence of RS-DSGE model using some different starting values and found that they delivered roughly similar results when looking at medians. However, due to substantial computational burden, the (informal) convergence test was limited. For SV-DSGE model, I verified the robustness of the algorithm by obtaining almost identical posterior median values in JP 2008.

4.2 Estimation Results

Parameter estimates are not entirely identical. While most of the parameter estimates from SV-DSGE are similar to ones reported in JP 2008, labor-related parameters of RS-DSGE models are inconsistently estimated. For instance, labor-related

parameters of RS-DSGE models are somewhat different. For example, labor disutility coefficient is higher in both RS(2)-DSGE and RS(4)-DSGE models. This may somehow generate lower labor disutility shock estimates than one from SV-DSGE model. Although this variation in estimates may be important, I do not explore it any more. Instead, I would like to focus on volatility estimates. Figure 13 through figure 15 plot the time-varying standard-deviations for the seven structural shocks of SV-DSGE, RS(2)-DSGE, and RS(4)-DSGE models, respectively. Note that figure 13 is roughly identical to the results in JP 2008. In figure 14, observe that the estimates of monetary policy, investment specific, and government spending shocks are roughly similar to those in SV-DSGE model, but the estimates of technology and intertemporal preference shocks behave very differently. While these two estimates from SV-DSGE model tend to show gradual decline over the time periods, corresponding estimates from RS(2)-DSGE model are characterized by multiple Markov-shifts. Since the true volatility process is unknown, I do not know which is closer to the truth. However, it is very unlikely that all volatility processes change magnitude and shape simultaneously. The estimates from RS(4)-DSGE model look quite similar to those from RS(2)-DSGE model. But two things stand out in figure 15. Since I allow the regime associated with the exogenous disturbance showing the largest degree of time variation (monetary policy shock) to be independent of the regime switching processes of other shock variances, the double-hump shaped pattern is most notable in monetary policy shock estimates. Also, relatively fewer high volatility regimes are realized since mid 1980s. This enables us to replicate the great reduction in volatility of the remaining disturbances around mid 1980s. The fact that estimates from RS(2)-DSGE and RS(4)-DSGE models are somewhat different indicates that independent movement across volatility processes are evident. This shows why one should be careful about using RS(2)-DSGE model to identify the sources of the changes in the volatility of

U.S. macroeconomic variables.

Figure 16 presents the variance decomposition for output growth. With some exception, each model delivers roughly similar results. Note that SV-DSGE model assumes variance shares attributed to each shock are more time-varying. Figure 17 and figure 18 show posterior expected values of the high volatility regime in RS(2)-DSGE and RS(4)-DSGE model. Notice that figure 17 looks as if the two posterior expected values in Figure 18 are combined in one figure. According to RS(2)-DSGE model, the high volatility regimes for monetary shock are observed in the mid-1960s and in the beginning of 2000s. (See also the figure 2 in Liu, Waggoner, and Zha (2010), page 40. They have the same figure like I do.) Figure 18 tells us that there was no high volatility regime for monetary shock at that time. Observe that the starting period of high volatility regime for monetary shock and that for the others do not coincide in early 2000s. By allowing additional degree of flexibility in RS-DSGE model, I am able to detect the timing of each volatility regime shift better. Figure 19 separately plots the regime probabilities in RS(4)-DSGE model. The second and third rows in Figure 19 imply that taking account of these two possible regimes can be important since both of them are significantly greater than zeros in probabilities.

I try to address the drawback of RS(2)-DSGE model in different direction. Using the posterior median values of SV-DSGE model, reported in Table 3 and figure 13, I generate an artificial dataset. The thought experiment centers on the following idea: provided that SV-DSGE model ideally captures the volatility dynamics, how is the performance of RS(2)-DSGE model. Figure 13 is now used as the true DGP and the estimates of monetary shock disturbance in the figure tell us that low volatility regime was present in the mid 1960s. However, the estimates from RS(2)-DSGE model detects the presence of high volatility regime for monetary shock at that period. This is shown in figure 9. Posterior expected values of the high volatility regime are

displayed in figure 11. This figure conveys wrong impression that monetary shock was in high volatility state around 1960s as well as other shocks. These evidences show that RS(2)-DSGE model can detect the timing of the high volatility regime wrong and provide imprecise estimates for volatility processes.

As it is standard in the literature, I assess the fit by computing the marginal data density as suggested in Geweke (1999). From a Bayesian point of view, the marginal data density comparison gives a comprehensive measure of fit on non-nested competing models. Details of the computation of the marginal data density is relegated to the technical appendix in JP 2008. However, I would like to point out that JP do not integrate out all latent variables numerically. In fact, JP choose $f(\theta, H^T) = f(\theta)f(H^T) = f(\theta)\pi(H^T)$ and assume $\pi(\theta, H^T) = \pi(\theta)\pi(H^T)$ for computational convenience:

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

Then, the marginal data density can be approximated by:

$$\bar{m}_N(Y) = \left[\frac{1}{N} \sum_{j=1}^N \frac{f(\theta_j)}{p(Y|\theta_j, H_j^T)\pi(\theta_j)} \right]^{-1}$$

where θ_j and H_j^T are from posterior distribution. I acknowledge the possible approximation error from following JP's method. But, at the same time I understand that their assumption allows us to compute $\bar{m}_N(Y)$ very easily.

The last two rows of Table 3 reports the log-marginal data density and the median likelihood values for SV-DSGE, RS(2)-DSGE, and RS(4)-DSGE model, respectively. Consistent with JP, I find that the values of the log-marginal likelihood are in favor of SV-DSGE model. I also get results in line with Liu, Waggoner, and Zha (2010)

that between the two specifications of RS-DSGE models, the data favor the parsimoniously parameterized model with shock variances switching regimes simultaneously. I suspect that volatility dynamics are spuriously estimated in RS(2)-DSGE model but find out that it performs pretty well in terms of the Bayesian model selection criterion for offering a parsimonious approximation and delivering a better data fit. Consistent with the simulation study, I find that the median likelihood value for SV-DSGE model is the highest while RS(4)-DSGE model is the lowest.

5 Conclusion

I have estimated a variety of large-scale DSGE models in which the variance of the structural shocks is time-varying. In particular, I evaluated the effectiveness of Markov-switching and stochastic volatility specification of volatility dynamics. Results from simulation indicate that allowing for too few regimes in the RS-DSGE model leads to poor volatility estimates. SV-DSGE model promises great flexibility in modeling volatility dynamics in a sense that it does not restrict the number of changes nor synchronization across shocks. In empirical application, SV-DSGE model delivers best-fit and accounts for the heteroscedasticity present in the data well. Among RS-DSGE models, the one with synchronized shifts in shock variances fits best, but may provide imprecise estimates for volatility processes. I have shown that model comparison based on the marginal likelihood approach can be misleading since parsimoniously parameterized model will always be favored regardless of its ability to capture volatility dynamics. The findings imply that DSGE models extended with stochastic volatility are good alternatives for understanding the evolving volatility dynamics of U.S. aggregate data since it can minimize the impact of misspecification for the volatility dynamics.

6 Tables and Figures

Table 1: Summary of Simulation Study

	Description	True DGP	Estimation
SM1	Monetary vs the Rest	RS(4)	Restricted ⁺
SM2	Monetary vs Gov't and Price Mark-up vs the Rest	RS(8)	Restricted ⁺
SM3	High volatility in recession	RS(2)	Restricted ⁺
FM1	Monetary vs the Rest	RS(4)	Full-blown
FM2	Stochastic Volatility Process	SV	Full-blown

Note. Restricted⁺: I only estimate volatility parameters.

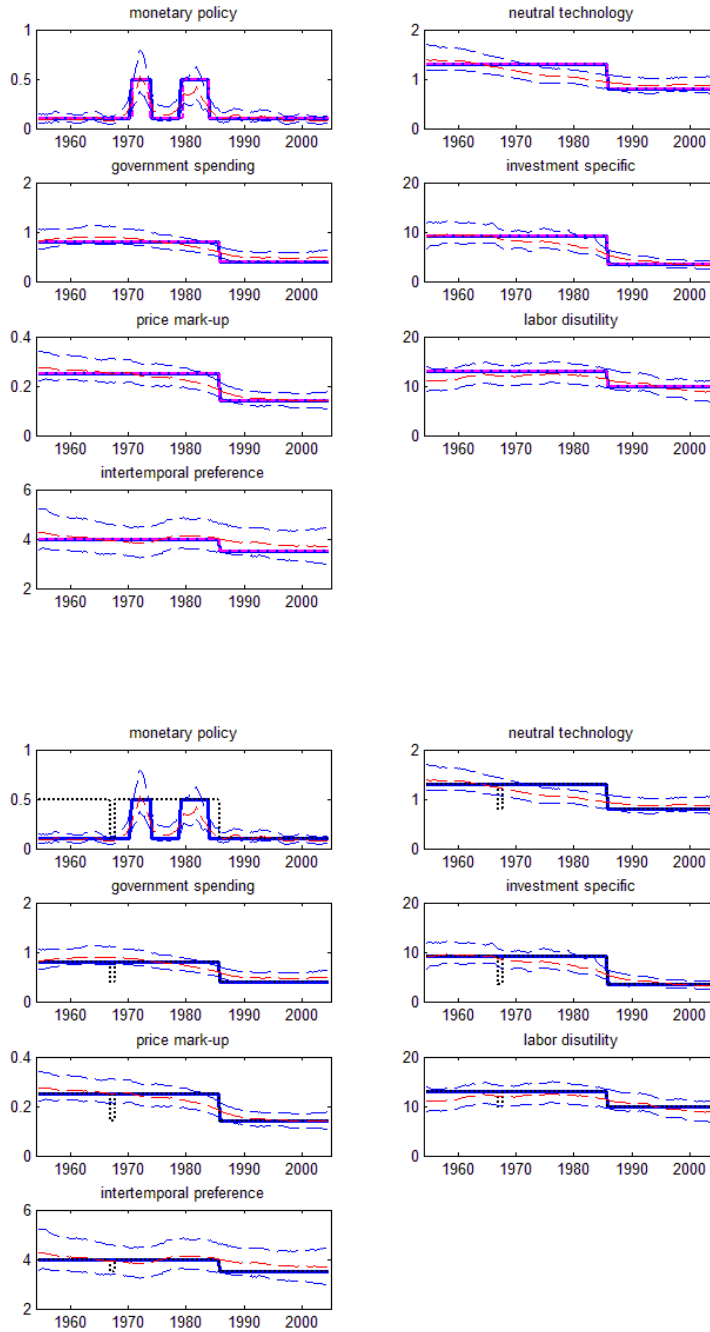
Table 2: Log Median Likelihood

DGP\Estimation	RS(2)-DSGE	RS(4)-DSGE	SV-DSGE
SM1	-1892.25	-1793.39	-1783.69
SM2	-1823.29	-1729.91	-1710.28
SM3	-1429.50	-1429.50	-1469.68

Table 3: Summary of Prior Densities and Posterior Estimates

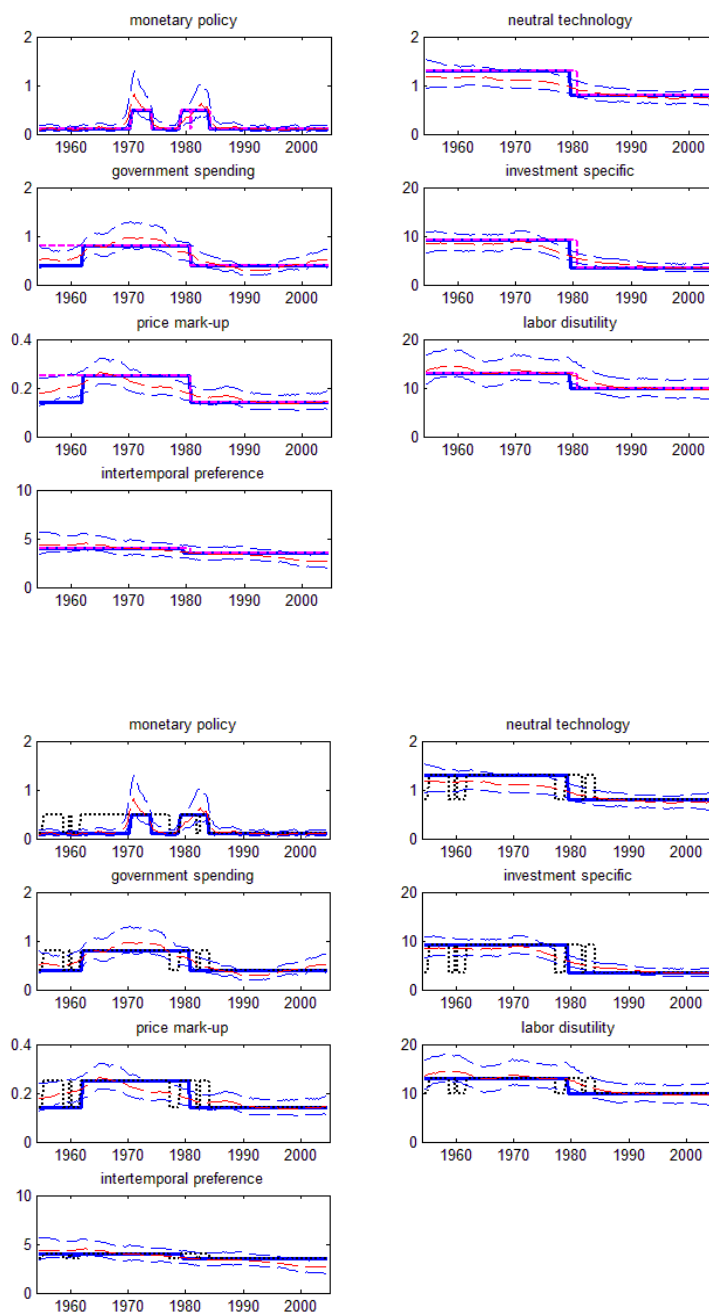
coefficient	Density	Prior			SV-DSGE			RS(2)-DSGE			RS(4)-DSGE		
		Mean	Std	Median	90% CI	Median	90% CI	90% CI	Median	90% CI	Median	90% CI	
α	-	0.30	0.00	-	-	-	-	-	-	-	-	-	
δ	-	0.025	0.00	-	-	-	-	-	-	-	-	-	
ι_p	B	0.50	0.15	0.83	[0.75, 0.91]	0.91	[0.83, 0.97]	0.82	[0.75, 0.89]				
ι_w	B	0.50	0.15	0.07	[0.04, 0.13]	0.01	[0.00, 0.03]	0.01	[0.00, 0.06]				
γ	N	0.50	0.03	0.43	[0.39, 0.47]	0.40	[0.36, 0.44]	0.40	[0.36, 0.44]				
h	B	0.50	0.10	0.84	[0.80, 0.88]	0.81	[0.78, 0.84]	0.80	[0.76, 0.82]				
λ_p	N	0.15	0.05	0.22	[0.16, 0.28]	0.25	[0.19, 0.31]	0.26	[0.20, 0.31]				
λ_w	N	0.15	0.05	0.16	[0.09, 0.23]	0.17	[0.11, 0.26]	0.18	[0.10, 0.24]				
L_{ss}	N	396.83	0.50	397.03	[396.31, 397.79]	397.09	[396.81, 397.45]	396.47	[396.31, 396.89]				
π	N	0.50	0.10	0.55	[0.37, 0.72]	0.60	[0.44, 0.72]	0.63	[0.44, 0.94]				
r	N	0.50	0.10	1.02	[0.92, 1.13]	1.23	[1.11, 1.36]	1.12	[1.00, 1.21]				
g_{ss}	-	0.22	0.00	-	-	-	-	-	-				
ν	G	2.00	0.75	1.49	[0.92, 2.37]	1.28	[0.72, 1.80]	0.66	[0.50, 0.91]				
ξ_p	B	0.75	0.10	0.91	[0.89, 0.93]	0.90	[0.88, 0.93]	0.91	[0.89, 0.94]				
ξ_w	B	0.75	0.10	0.66	[0.57, 0.73]	0.54	[0.40, 0.65]	0.54	[0.46, 0.59]				
χ	G	5.00	1.00	7.04	[5.42, 8.78]	7.68	[7.27, 8.16]	5.08	[4.95, 5.39]				
S''	G	3.00	0.75	3.19	[2.38, 4.22]	3.01	[2.44, 3.68]	1.98	[1.87, 2.15]				
Φ_p	N	1.70	0.30	1.87	[1.65, 2.14]	2.01	[1.87, 2.21]	1.80	[1.63, 1.95]				
Φ_y	G	0.13	0.10	0.08	[0.06, 0.11]	0.07	[0.05, 0.09]	0.07	[0.05, 0.09]				
ρ_R	B	0.60	0.20	0.84	[0.80, 0.87]	0.84	[0.81, 0.87]	0.81	[0.77, 0.84]				
ρ_z	B	0.40	0.20	0.33	[0.23, 0.43]	0.33	[0.25, 0.41]	0.27	[0.17, 0.34]				
ρ_g	B	0.60	0.20	0.99	[0.99, 0.99]	0.98	[0.98, 0.98]	0.98	[0.98, 0.98]				
ρ_μ	B	0.60	0.20	0.93	[0.88, 0.96]	0.95	[0.92, 0.97]	0.93	[0.90, 0.95]				
ρ_p	-	0.00	0.00	-	-	-	-	-	-				
ρ_φ	B	0.60	0.20	0.90	[0.82, 0.95]	0.96	[0.93, 0.98]	0.98	[0.97, 0.98]				
ρ_b	B	0.60	0.20	0.81	[0.70, 0.88]	0.90	[0.87, 0.93]	0.89	[0.83, 0.92]				
median likelihood				-1718.8		-1737.7		-1730.6					
MDD				-1846.9		-1867.0		-1904.7					

Figure 1: True and Estimated Volatilities for the model SM1



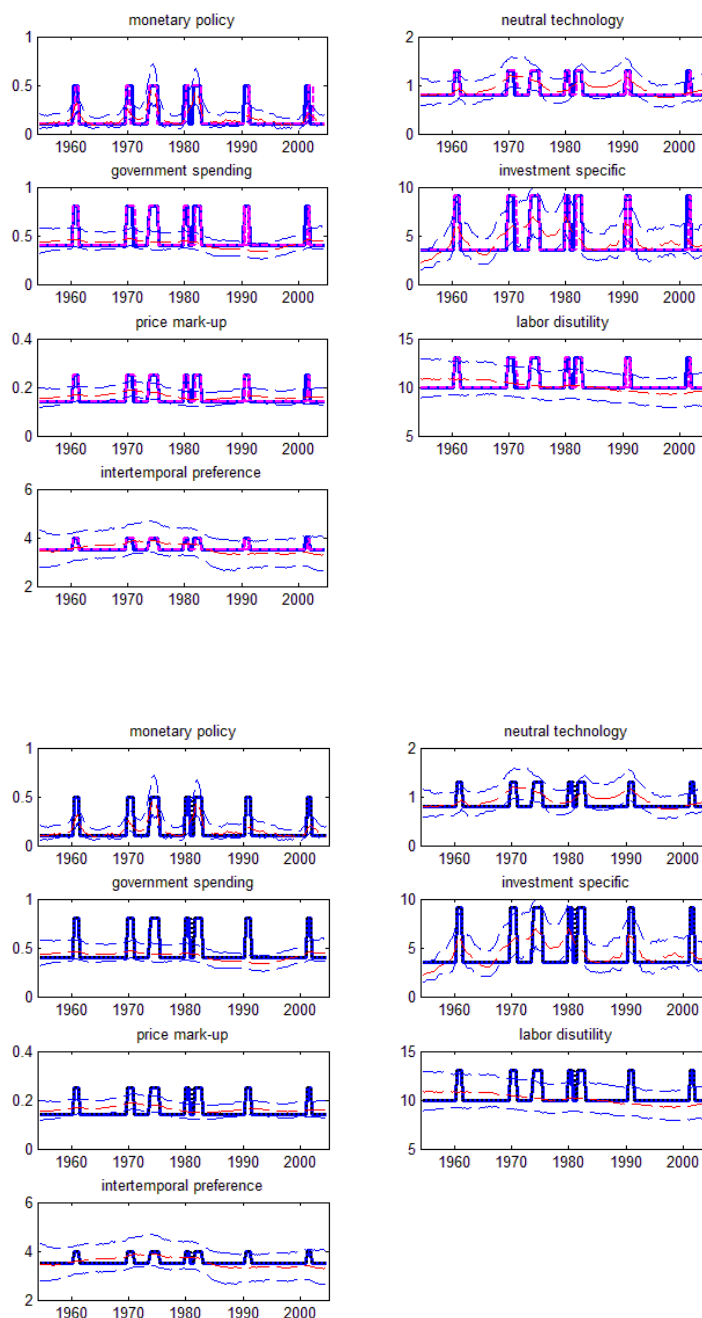
Note: In this figure I plot the true volatility processes (bold), median standard deviation and 90% credible intervals generated in the estimation of SV-DSGE (dashed), RS(2)-DSGE (black-dotted), and RS(4)-DSGE (pink-dashed).

Figure 2: True and Estimated Volatilities for the model SM2



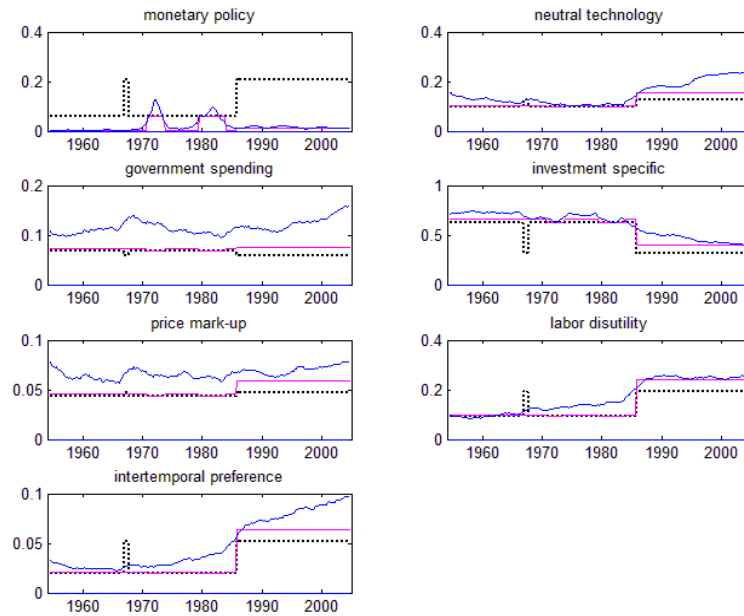
Note: In this figure I plot the true volatility processes (bold), median standard deviation and 90% credible intervals generated in the estimation of SV-DSGE (dashed), RS(2)-DSGE (black-dotted), and RS(4)-DSGE (pink-dashed).

Figure 3: True and Estimated Volatilities for the model SM3



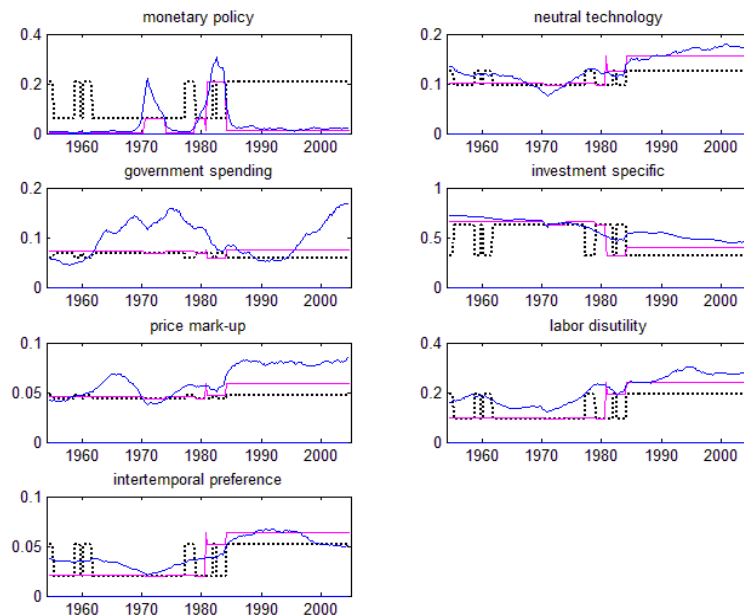
Note: In this figure I plot the true volatility processes (bold), median standard deviation and 90% credible intervals generated in the estimation of SV-DSGE (dashed), RS(2)-DSGE (black-dotted), and RS(4)-DSGE (pink-dashed).

Figure 4: Variance Decomposition for the model SM1



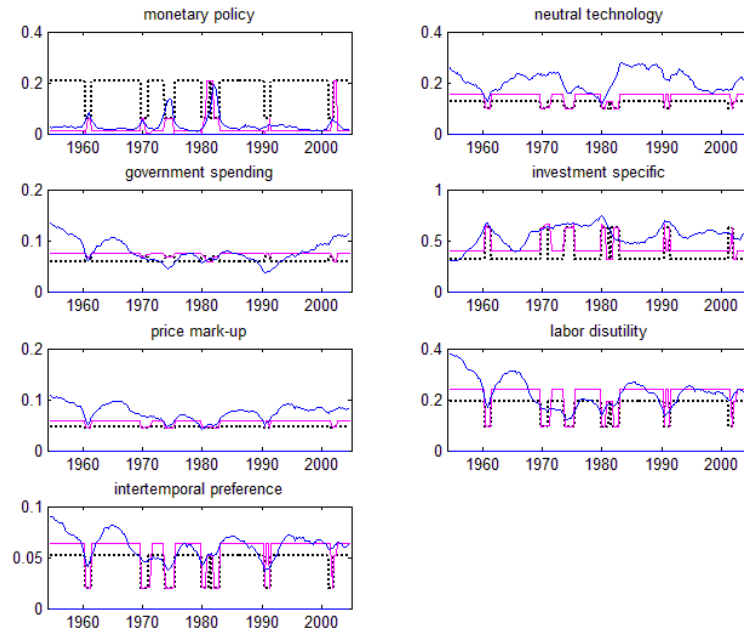
Note: This figure presents the contribution of each shock to the variability of GDP growth. Variance decomposition under RS(4)-DSGE (pink), SV-DSGE (blue), and RS(2)-DSGE model (black-dotted) are included.

Figure 5: Variance Decomposition for the model SM2



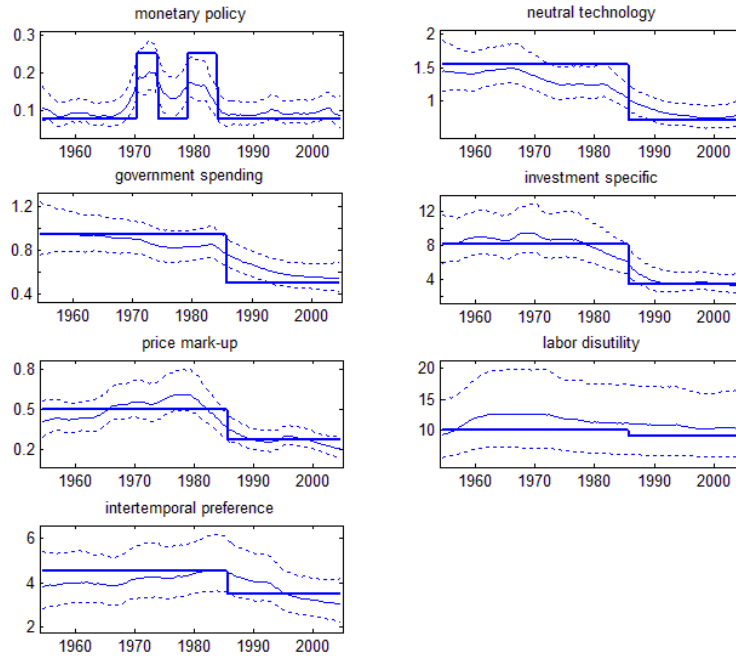
Note: This figure presents the contribution of each shock to the variability of GDP growth. Variance decomposition under RS(4)-DSGE (pink), SV-DSGE (blue), and RS(2)-DSGE model (black-dotted) are included.

Figure 6: Variance Decomposition for the model SM3



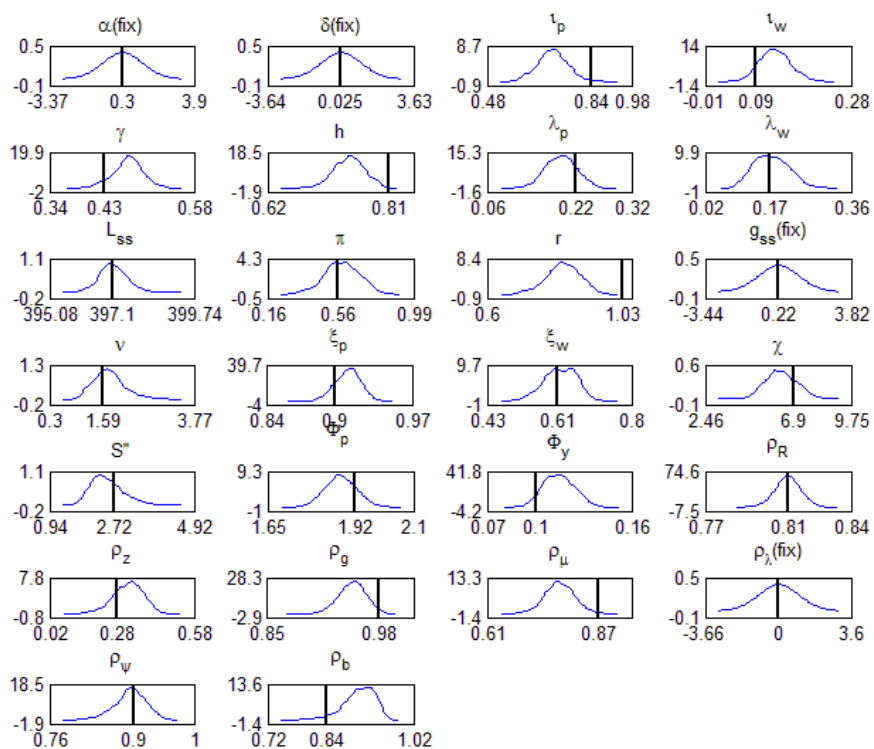
Note: This figure presents the contribution of each shock to the variability of GDP growth. Variance decomposition under RS(4)-DSGE (pink), SV-DSGE (blue), and RS(2)-DSGE model (black-dotted) are displayed.

Figure 7: True and Estimated Volatilities for the model FM1



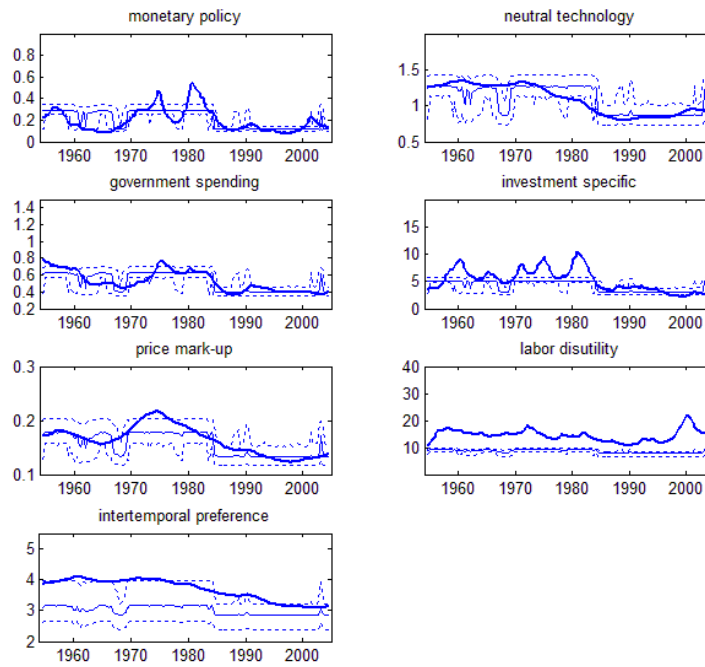
Notes: Bold lines represent true volatility processes. Median standard deviations(solid) and 90% credible intervals (dotted) are generated in the estimation of SV-DSGE model.

Figure 8: Posterior Distributions of non-Volatility Parameters: FM1



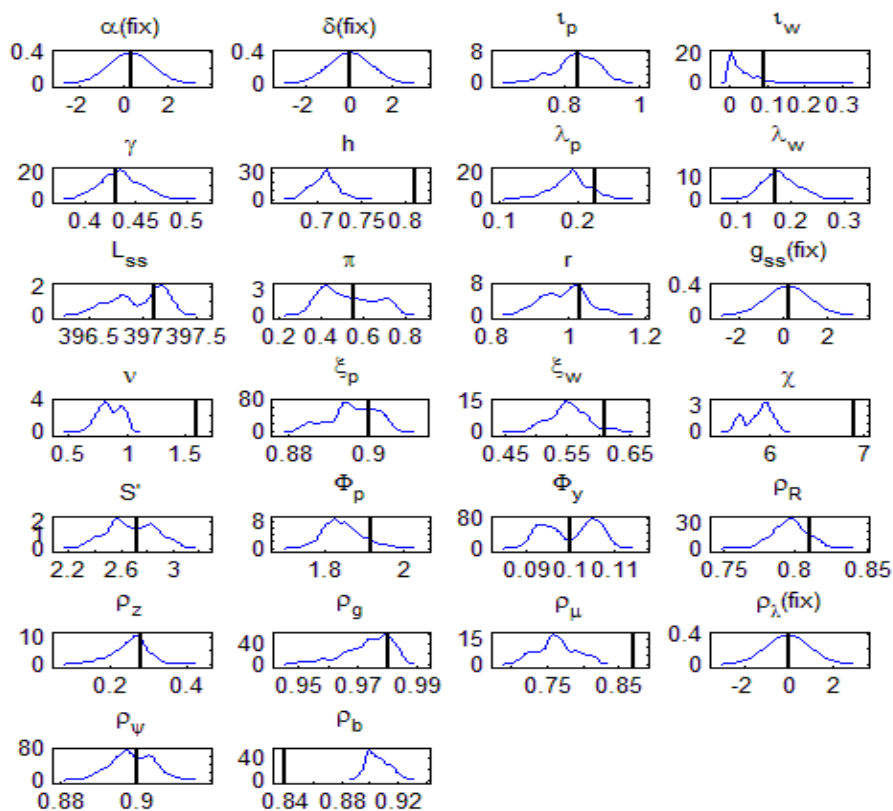
Note: This figure plots kernel density estimation of posterior distributions. True parameter values are indicated by vertical lines. DGP is RS(4)-DSGE and I estimate with SV-DSGE model.

Figure 9: True and Estimated Volatilities: FM2



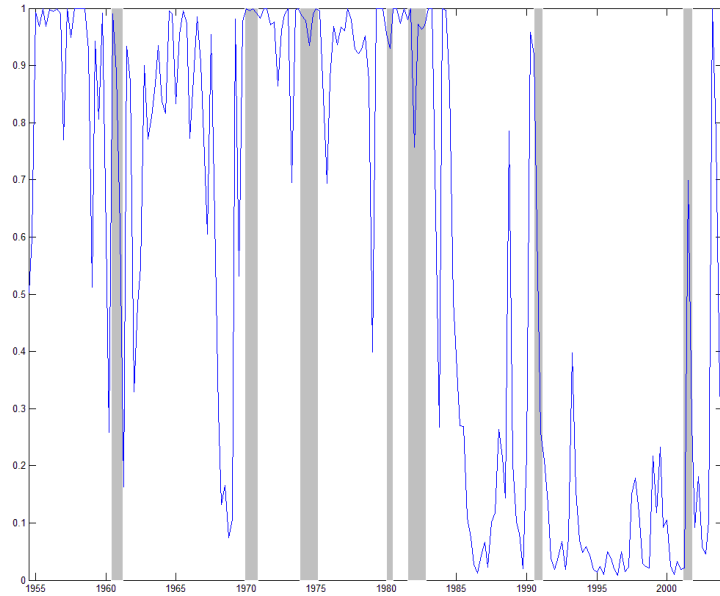
Notes: Bold lines represent true volatility processes, SV-DSGE. Median standard deviations (solid) and 90% credible intervals (dotted) are generated in the estimation of RS(2)-DSGE model.

Figure 10: Posterior Distributions of non-Volatility Parameters: FM2



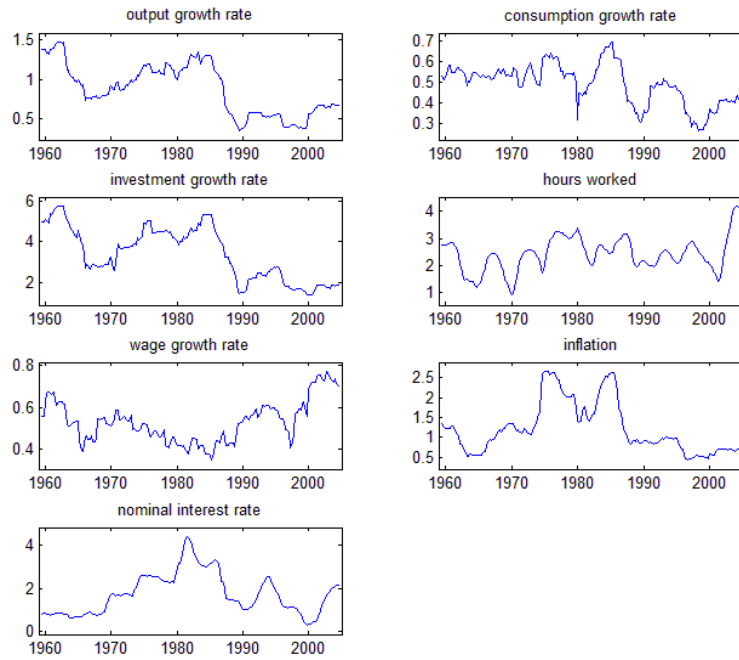
Note: This figure plots kernel density estimation of posterior distributions. True parameter values are indicated by vertical lines. DGP is SV-DSGE and I estimate with RS(2)-DSGE model.

Figure 11: Posterior Probability of the High Volatility Regime: FM2



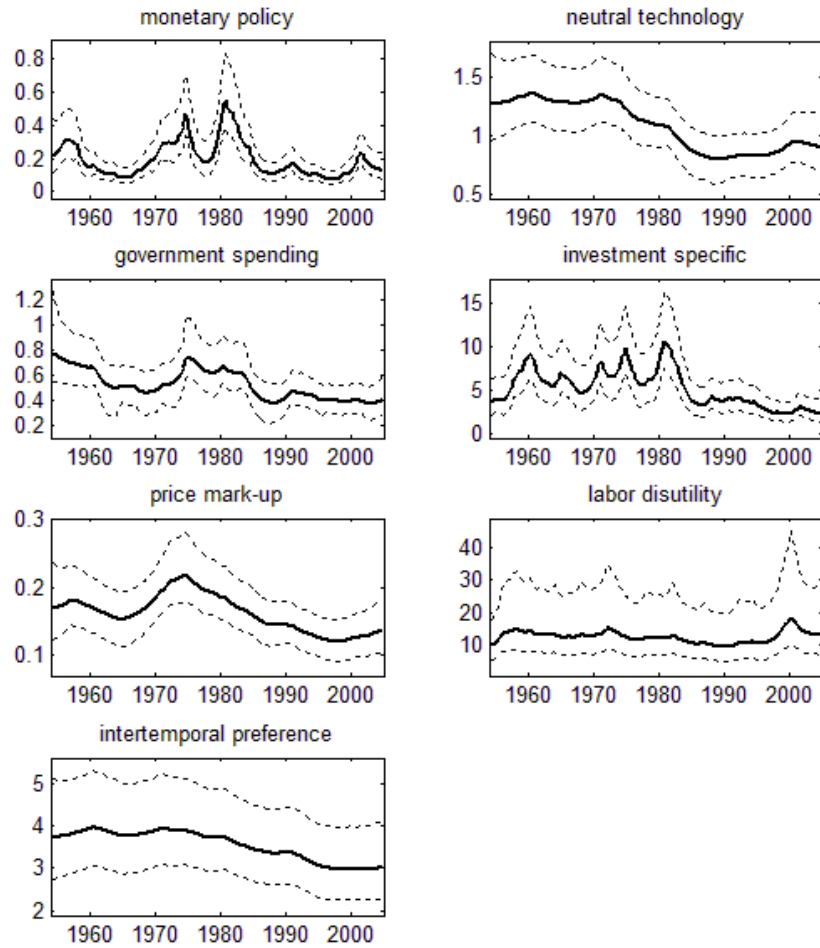
Notes: Shaded bars indicate NBER recessions and solid line represents posterior expected value of the high volatility regime in RS(2)-DSGE model. DGP is SV-DSGE and I estimate with RS(2)-DSGE model. The results are based on 200,000 posterior draws.

Figure 12: Rolling Standard Deviations for U.S. Data



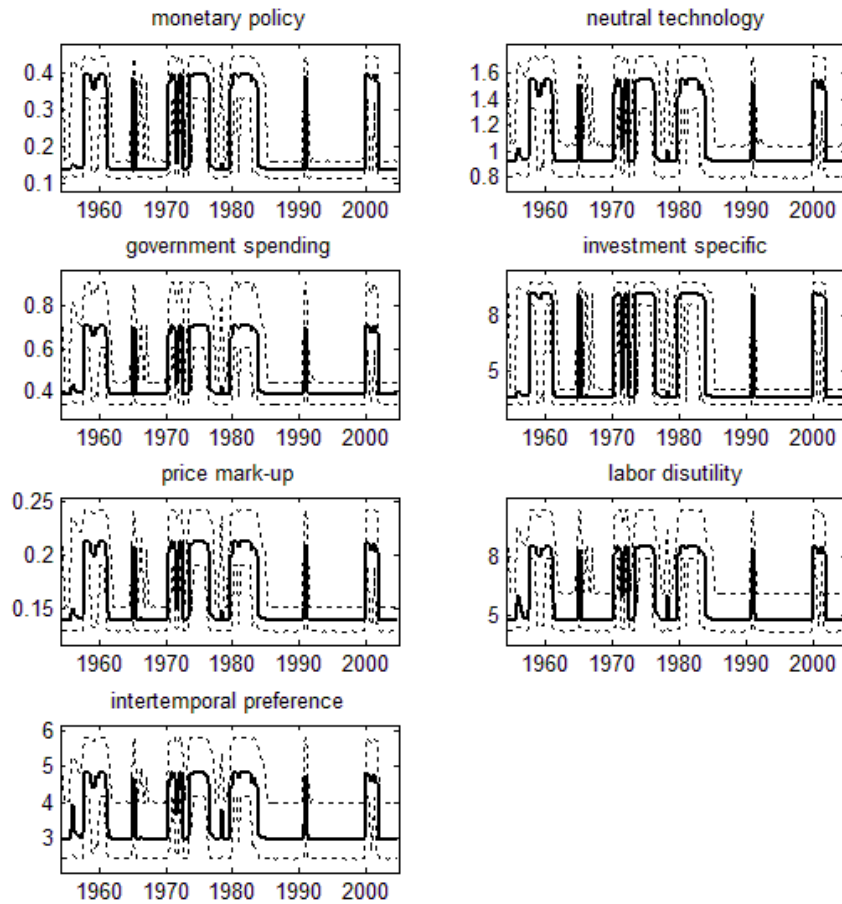
Notes: Rolling estimation window is twenty quarters for this figure.

Figure 13: Estimated Standard Deviations: SV-DSGE



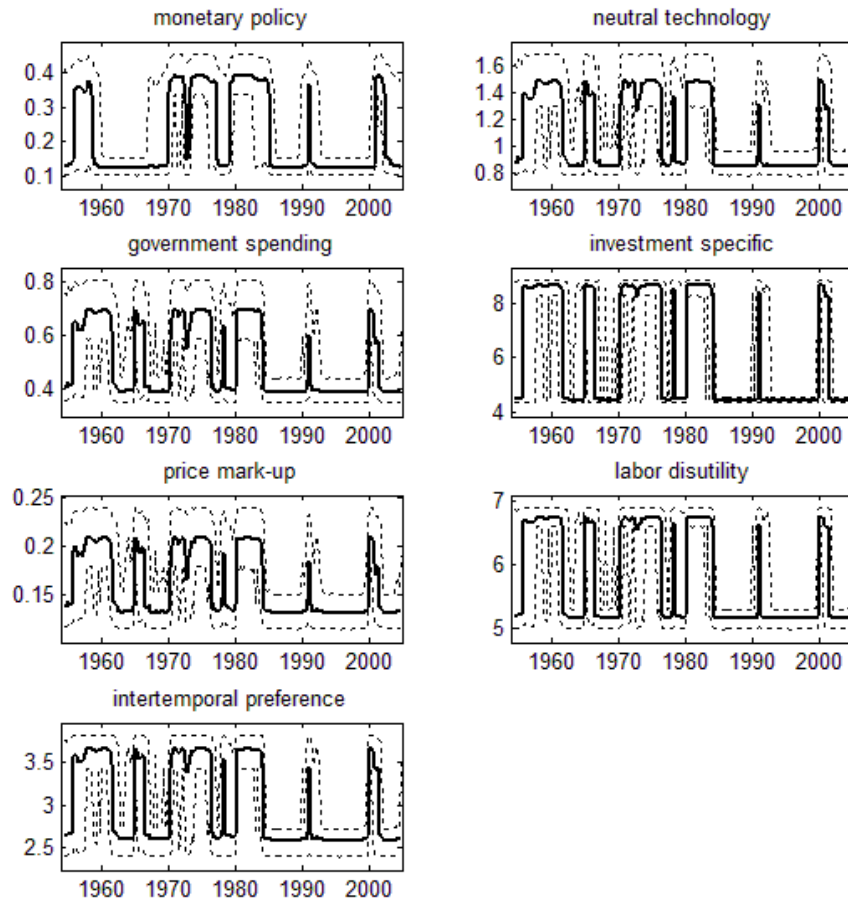
Notes: Median (bold) and 90% credible intervals (dotted) for the time-varying volatility of each disturbance computed with the draws generated in the estimation of SV-DSGE model.

Figure 14: Estimated Standard Deviations: RS(2)-DSGE



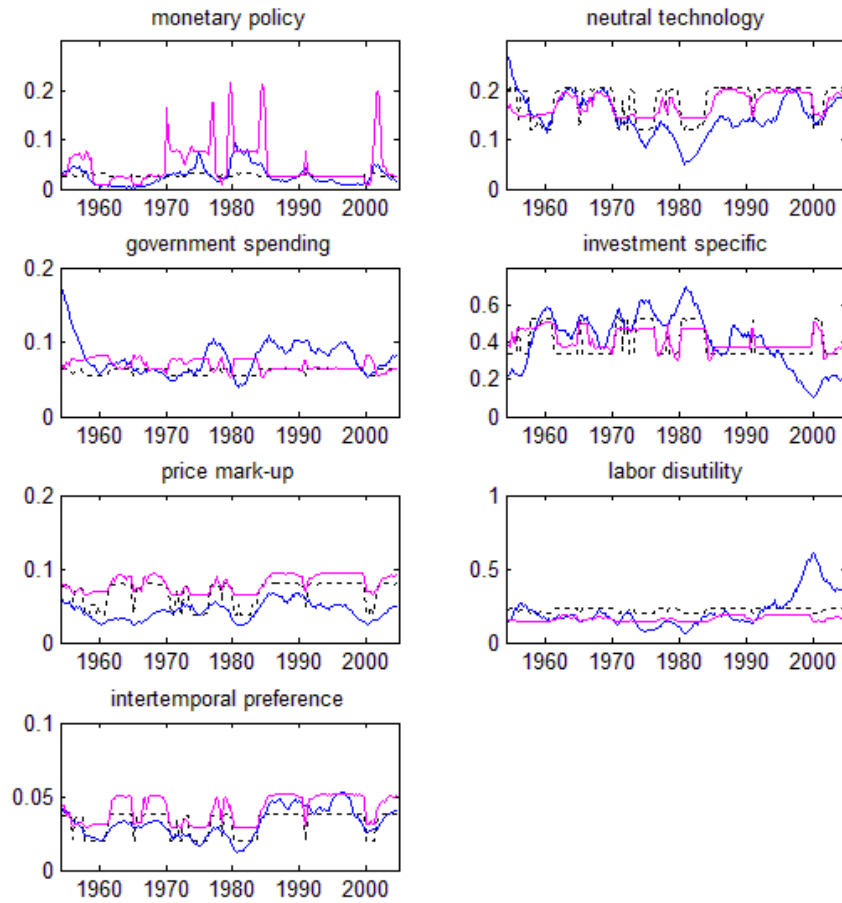
Notes: Median (bold) and 90% credible intervals (dotted) for the time-varying volatility of each disturbance computed with the draws generated in the estimation of RS(2)-DSGE model.

Figure 15: Estimated Standard Deviations: RS(4)-DSGE



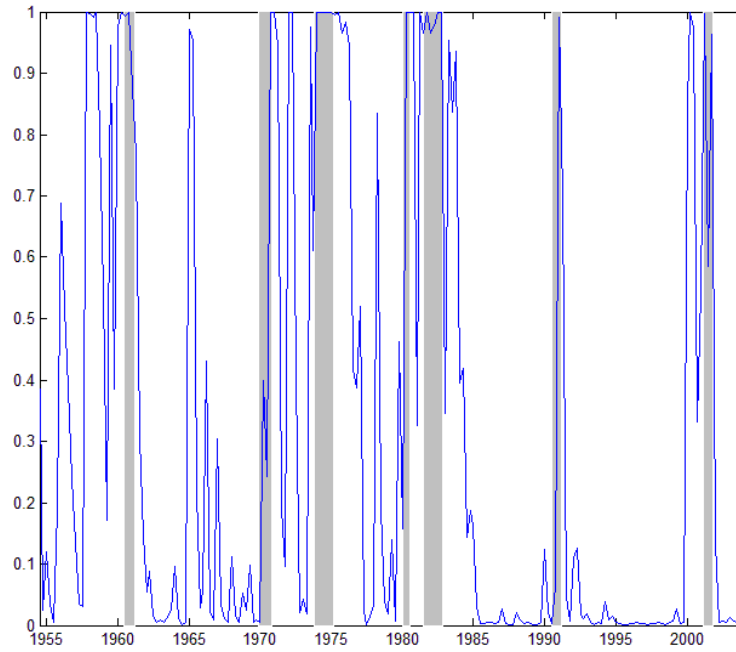
Notes: Median (bold) and 90% credible intervals (dotted) for the time-varying volatility of each disturbance computed with the draws generated in the estimation of RS(4)-DSGE model.

Figure 16: Variance Decomposition for U.S. Data



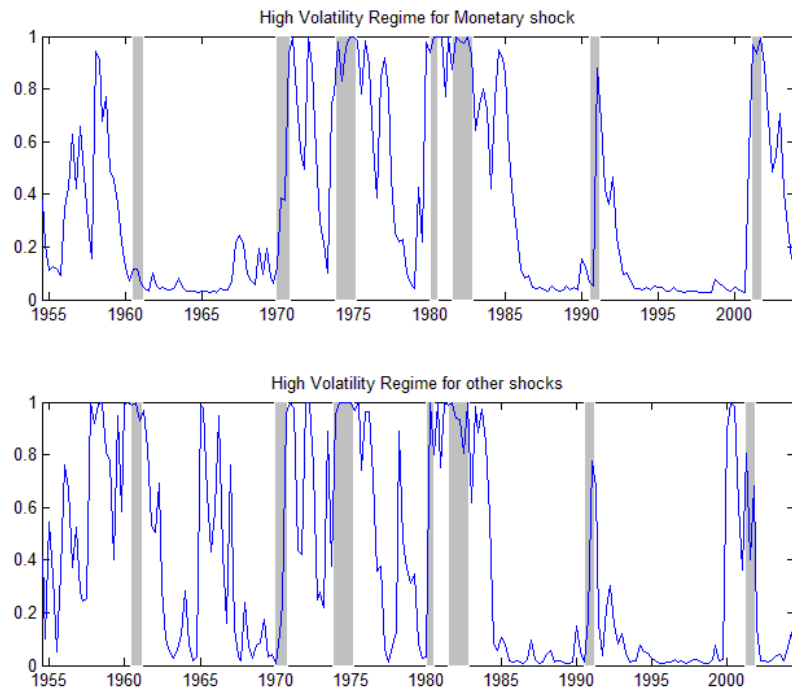
Notes: This figure presents the contribution of each shock to the variability of GDP growth. Variance decomposition under RS(4)-DSGE (pink), SV-DSGE (blue), and RS(2)-DSGE model (black-dotted) are displayed.

Figure 17: Posterior Probability of the High Volatility Regime: RS(2)-DSGE



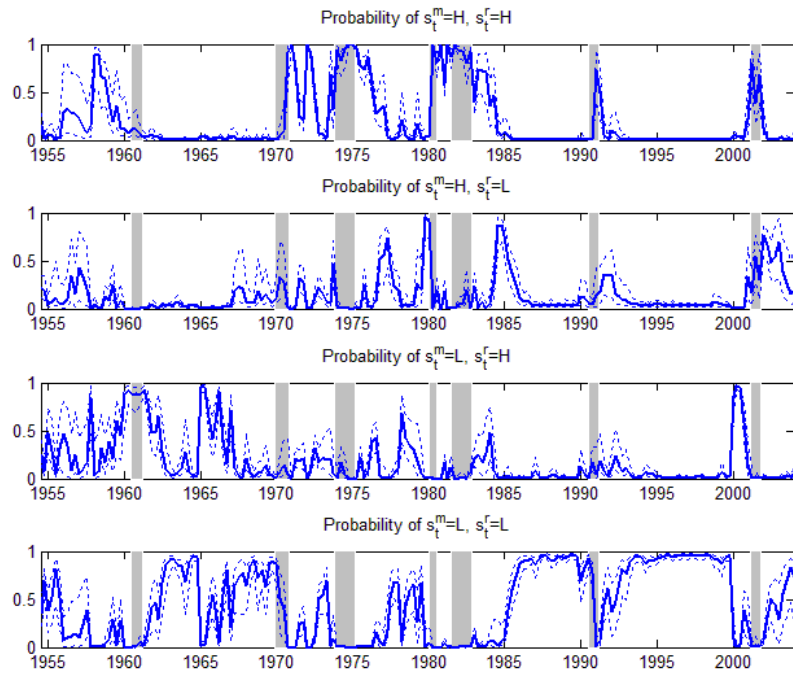
Notes: Shaded bars indicate NBER recessions and solid line represents posterior expected value of the high volatility regime in RS(2)-DSGE model. The results are based on 300,000 posterior draws.

Figure 18: Posterior Probability of the High Volatility Regime: RS(4)-DSGE



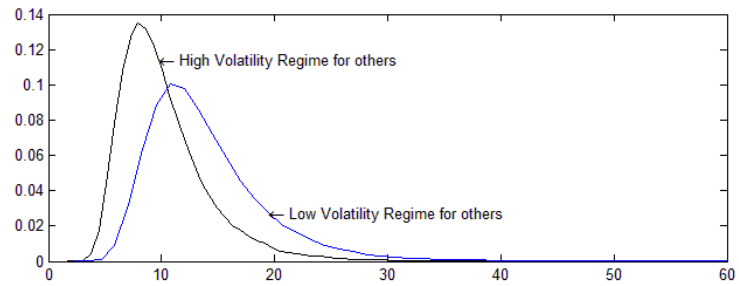
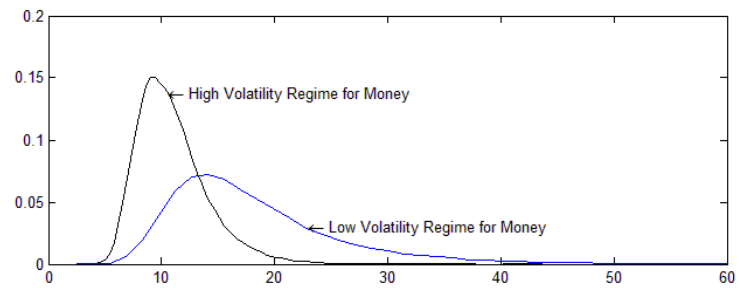
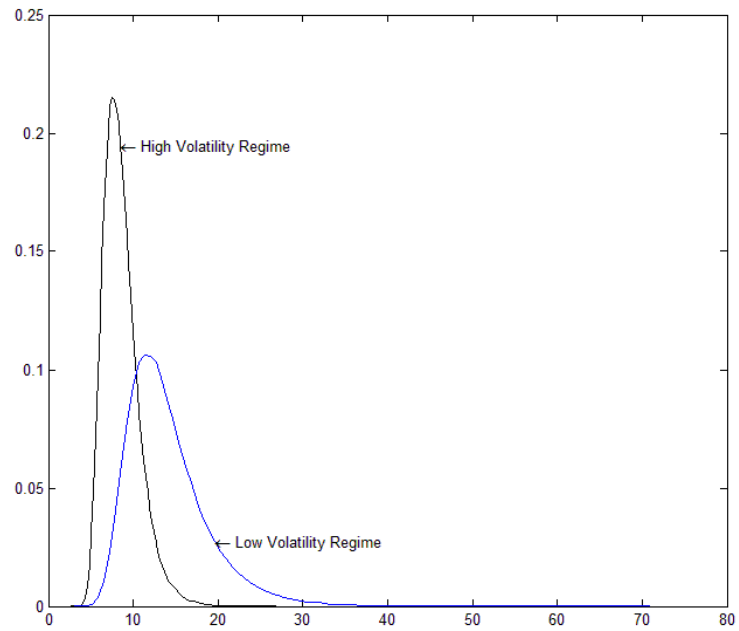
Notes. Solid line represents posterior probability of the high volatility regime for monetary shock and the other shocks respectively in RS(4)-DSGE model.

Figure 19: Posterior Probability: RS(4)-DSGE



Notes: Shaded bars indicate NBER recessions and bold lines are posterior probability of each regime in RS(4)-DSGE model and dashed lines are 90% credible intervals.

Figure 20: Posterior Density of High- and Low- Volatility Regime Duration



Notes: Top figure is posterior density of each regime for RS(2)-DSGE and two bottom figures are for RS(4)-DSGE.

Chapter III

Bayesian Estimation of a New Open Economy Model with Adaptive Expectations

7 Introduction

Explaining real exchange rate dynamics has been an long-lasting challenge in international economics. Exchange rates are more volatile and persistent than standard open-economy models can account for, and they appear to be disconnected from fundamentals in the short run. In one strand of the literature, there are a large number of theoretical studies which incorporate the endogenous sources that can make consumption unresponsive to the exchange rate. Some examples include nominal rigidities, pricing-to-market, introduction of durable goods and investment, and alternative asset market structure, e.g., the work of Betts and Devereux (2000), Chari, Kehoe and MaGrattan (2002), Monacelli (2005), Benigno (2009), and Engel and Wang (2011). In the other strand of the literature, relatively few empirical assessments of structural models have been done, as seen in Smets and Wouters (2002), Adolfson et al. (2001, 2007), Bergin (2003, 2006), Lubik and Schorfheide (2005). These papers estimate structural models with different frictions and specifications and document the importance of incomplete pass-through.

A variety of open economy models are built on is the rational expectations hypothesis that expectations have to be model-consistent. Motivated by the criticism

on rational expectations that assumes too much knowledge by economic agents, recent research has formulated ways of deviating from rational expectations. The most common approach is adaptive learning, which assumes that economic agents make their forecasts based on past observations and update the forecast every period, like an econometrician following Sargent (1993) and Evans and Honkapohja (2001). In the closed-economy general equilibrium framework, a variety of scholars shows that the learning mechanism amplifies the effects of stochastic shocks and gives a plausible explanation for inflation persistence. (Milani, 2005, 2007; Huang, Liu and Zha, 2009; Eusepi and Preston, 2011; Slobodyan and Wouters, 2012) Learning breaks the tight link between fundamental variables imposed by general equilibrium models and therefore can potentially reconcile business cycle patterns that were difficult to explain under rational expectations. In the open economy literature, the learning mechanism succeeds in replicating some aspects of exchange rate, e.g., Mark (2009), Lewis and Markiewicz (2009), Dieppe et al. (2013). Mark (2009) finds that the learning paths match the volatility and actual movement of the real deutschemark-dollar exchange rate from 1973 to 2005 better in a partial equilibrium model. Lewis and Markiewicz (2009) adapt a simple monetary model with dual learning and reproduce the excess volatility of the exchange rate return. Dieppe et al. (2013) use a multi-country euro area model with a limited information learning approach, and document different responses to an expansionary fiscal policy under learning and rational expectations.

The contribution of this paper is that I introduce adaptive expectations rather than rational expectations to relax the tight link between the exchange rate and fundamentals imposed by a two-country open economy model with nominal price rigidities, i.e., New Open Economy Macroeconomics. The model is the open-economy version of the canonical New Keynesian dynamic stochastic general equilibrium (DSGE) model extended with the existence of monopolistically competitive importers. The

importers' price-setting behavior introduces endogenous deviation from purchasing power parity, allowing for incomplete exchange rate pass-through. Under the learning mechanism, economic agents are assumed to form their expectations of forward-looking variables using a simple vector autoregressive forecasting model. The agents estimate their vector autoregression based on past model variables and update the estimates every period via a constant-gain learning algorithm. Constant-gain learning is widely used in learning literature due to its appealing features, namely, that a single parameter regulates the departure from rational expectation, and that the learning model nests the rational expectation model.

I conduct simulation exercises to compare the equilibrium paths implied by rational expectations and the learning mechanism. The simulation results show that the learning mechanism increases the volatility and persistence of the endogenous variables. By increasing the gain parameter in a constant gain algorithm, these increases become more pronounced. Since agents' subjective views of the economy change over time, the belief-updating process increases the overall volatilities. When agents observe a higher realization of variables than expected, the perceived persistence will be revised upward, leading to the additional persistence in the data generating process. The gradual and ongoing adjustment of beliefs is an endogenous source of persistence in learning models. The learning mechanism also improves the model in terms of uncovered interest rate parity by allowing the interest rate gap to deviate from the exchange rate depreciation by the expectational difference between rational expectations and the expectations formed from the learning process. For some values of the gain parameter, the learning mechanism reduces the correlation between the exchange rate depreciation and the interest rate gap as far as is found in the data.

The two-country open economy model is estimated with U.S. and Euro area data from 1983:Q3 to 2012:Q4 under the two different specifications of expectations: ra-

tional expectations and learning. The Bayesian estimation provides three key findings. First, the posterior distributions under the learning mechanism show that price stickiness parameters for the U.S. are much smaller than they are under rational expectations and that the learning mechanism reduces the importance of habit formation. This implies that the learning mechanism replaces the endogenous source of persistence in line with findings of Milani (2005, 2007) for closed economy DSGE models. Second, I compute posterior probabilities for the learning and the rational expectations specifications, and find that the data favors the model with constant gain learning. More specifically, the learning model better fits the output growth comovement between the U.S. and the Euro area and the correlation between output growth and inflation of the U.S., although the improvement in terms of fitting observed Dollar-Euro exchange rate dynamics is limited. Third, the model with the learning mechanism produces lagged and persistent responses of inflations and exchange rate depreciation to the monetary policy shocks, whereas the shocks dissipate quickly in the rational expectations model.

This chapter is organized as follows. Section 8 describes a small-scale two-country general equilibrium model and Section 9 illustrates the learning model and how the learning mechanism changes the exchange rate determination. Section 10 presents the simulation study, and the empirical application to U.S. and Euro-area data will be followed in Section 11. I conclude in Section 12.

8 The Model

In this section, I specify a small-scale two-country general equilibrium model, the extension of Monacelli (2005). Home country and foreign country are assumed to be symmetric in terms of preference and technology and will be denoted by H for

the home country and by F for the foreign country. Each economy is populated by a continuum of households, final good producers, intermediate good producers, importers and a monetary authority. As a matter of notation, subscripts H and F refer to the country where the good is produced and an asterisk '*' indicates the foreign variables.

When it comes to log-linearization, non-stationary worldwide technology shock induces the non-stationary trend on some variables, so I express the model with respect to detrended variable beforehand. I let the tilded variables denote the log-deviation from the steady state, i.e. $\tilde{x}_t = \ln\left(\frac{x_t}{x}\right)$.

8.1 Domestic Households

The domestic representative household consumes Dixit-Stiglitz aggregate C_t of domestic goods $C_{H,t}$ and imported goods $C_{F,t}$. Aggregate consumption is defined as

$$C_t = \left((1 - \alpha)^{\frac{1}{\mu}} C_{H,t}^{\frac{\mu-1}{\mu}} + \alpha^{\frac{1}{\mu}} C_{F,t}^{\frac{\mu-1}{\mu}} \right)^{\frac{\mu}{\mu-1}} \quad (8)$$

The household is a monopolistic labor supplier to the firms and derives the disutility from hours worked and maximizes

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \left(\frac{(C_t - h\gamma C_{t-1}) / A_{W,t}}{1 - \tau} \right)^{1-\tau} - L_t \right\} \quad (9)$$

subject to

$$P_{H,t} C_{H,t} + P_{F,t} C_{F,t} + E_t [Q_{t,t+1} D_{t+1}] \leq W_t L_t + D_t - T_t \quad (10)$$

where β is a discount factor, τ is a relative risk-aversion parameter, h is the habit formation parameter and W_t is the nominal wage. The household is assumed to derive utility from effective consumption relative to the level of non-stationary world-wide

technology, $A_{W,t}$, so that the economy evolves along a balance growth path even if the utility is additively separable in consumption and leisure. I also assume the existence of complete asset markets so that households have access to a complete set of the state-contingent claims denominated in the home country currency, which will be evaluated by the stochastic discount factor $Q_{t,t+1}$ at time t . T_t is the lump-sum tax imposed by the government to finance its purchase.

The household maximization problem implies the following optimality conditions:

$$P_t \lambda_t = \frac{1}{A_{W,t}} \left(\frac{C_t - h\gamma C_{t-1}}{A_{W,t}} \right)^{-\tau} - \beta E_t \left[\frac{h\gamma}{A_{W,t+1}} \left(\frac{C_{t+1} - h\gamma C_t}{A_{W,t+1}} \right)^{-\tau} \right] \quad (11)$$

$$1 = \lambda_t \frac{W_t}{P_t} \quad (12)$$

$$\frac{\lambda_t}{P_t} Q_{t,t+1} = \beta \frac{\lambda_{t+1}}{P_{t+1}} \quad (13)$$

Under the complete asset market assumption, the price of risk-free government bond is constructed by averaging the price of state-contingent claims as:

$$\frac{1}{R_t} = E [Q_{t,t+1}] = \beta E \left[\frac{\lambda_{t+1}}{\lambda_t} \frac{P_t}{P_{t+1}} \right] \quad (14)$$

where R_t is the nominal interest rate, which is the instrument of monetary authority.

8.2 Domestic Producers

8.2.1 Domestic Final Producers

The final goods producers combine home produced goods and imported goods according to a constant elasticity of substitution (CES) aggregation technology and sell

in a perfectly competitive market. Final good producers maximize the profit

$$\max P_t Y_t - P_{H,t} Y_{H,t} - P_{F,t} Y_{F,t} \quad (15)$$

subject to

$$Y_t = \left[(1 - \alpha)^{\frac{1}{\mu}} Y_{H,t}^{\frac{\mu-1}{\mu}} + \alpha^{\frac{1}{\mu}} Y_{F,t}^{\frac{\mu-1}{\mu}} \right]^{\frac{\mu}{\mu-1}} \quad (16)$$

where $0 < \alpha < 1$ is a import share parameter and μ determines the degree of substitutability of home goods and foreign goods. $Y_{H,t}$ is the aggregate of home goods sold in home country and $Y_{F,t}$ is the aggregate of imported goods sold in home country. The demand for home good and the demand for foreign good are given by

$$Y_{H,t} = (1 - \alpha) \left[\frac{P_{H,t}}{P_t} \right]^{-\mu} Y_t \quad (17)$$

$$Y_{F,t} = \alpha \left[\frac{P_{F,t}}{P_t} \right]^{-\mu} Y_t \quad (18)$$

The zero profit condition implies the aggregate price index as

$$P_t = \left[(1 - \alpha) P_{H,t}^{1-\mu} + \alpha P_{F,t}^{1-\mu} \right]^{\frac{1}{1-\mu}} \quad (19)$$

The log-linearized version of (19) yields

$$\begin{aligned} \tilde{\pi}_t &= (1 - \alpha) \tilde{\pi}_{H,t} + \alpha \tilde{\pi}_{F,t} \\ &= \tilde{\pi}_{H,t} - \alpha (\tilde{\pi}_{H,t} - \tilde{\pi}_{F,t}) \\ &= \tilde{\pi}_{H,t} - \alpha \Delta \tilde{q}_t \end{aligned} \quad (20)$$

where the log terms of trade q_t is defined by the price of exports relative to the price

of imports in terms of the domestic currency ($P_{H,t}/P_{F,t}$). When the exchange rate pass-through is complete, the exchange rate fluctuation is transmitted to the domestic inflation, scaled by the import share.

Each home good aggregate and foreign good aggregate are composed of a continuum of differentiated intermediate goods indexed by $j \in [0, 1]$

$$Y_{H,t} = \left(\int_0^1 Y_{H,t}(j)^{\frac{1}{1+\nu}} dj \right)^{1+\nu} \quad (21)$$

$$Y_{F,t} = \left(\int_0^1 Y_{F,t}(j)^{\frac{1}{1+\nu}} dj \right)^{1+\nu} \quad (22)$$

The profit maximization problems of home and foreign final good producers yield the demands for intermediate goods

$$Y_{H,t}(j) = \left(\frac{P_{H,t}(j)}{P_{H,t}} \right)^{-\frac{1+\nu}{\nu}} Y_{H,t} \quad (23)$$

$$Y_{F,t}(j) = \left(\frac{P_{F,t}(j)}{P_{F,t}} \right)^{-\frac{1+\nu}{\nu}} Y_{F,t} \quad (24)$$

and the zero profit conditions give the relationships between the aggregate good price and the individual intermediate home goods prices

$$P_{H,t} = \left(\int_0^1 P_{H,t}(j)^{-\frac{1}{\nu}} dj \right)^{-\nu} \quad (25)$$

$$P_{F,t} = \left(\int_0^1 P_{F,t}(j)^{-\frac{1}{\nu}} dj \right)^{-\nu} \quad (26)$$

8.2.2 Domestic Intermediate Producers

The intermediate good producers are monopolistically competitive. They produce differentiated goods according to a linear production function that uses only labor

$$Y_{H,t}(j) = A_{W,t}A_{H,t}L_t(j) \quad (27)$$

where $A_{W,t}$ is a non-stationary world-wide technology shock and $A_{H,t}$ is a stationary and country-specific technology shock.

Firms face the nominal rigidities following the Calvo-type pricing mechanism. At the beginning of the period, a fraction ξ_H of firms cannot re-optimize their price and adjust the prices to steady state inflation, π .

$$P_{H,t}(j) = P_{H,t-1}(j)\pi \quad (28)$$

The remaining fraction $(1 - \xi_H)$ of firms choose their prices by maximizing the present value of future profits.

$$\max E_t \sum_{s=0}^{\infty} \xi_H^s Q_{t,t+s} \{P_{H,t}(j)\pi^s Y_{H,t+s}(j) - W_{t+s}L_{t+s}(j)\} \quad (29)$$

subject to

$$\begin{aligned} Y_{H,t+s}(j) &= \left(\frac{\pi^s P_{H,t}(j)}{P_{H,t+s}} \right)^{-\frac{1+\nu}{\nu}} Y_{H,t+s} \\ Y_{H,t}(j) &= A_{W,t}A_{H,t}L_t(j) \end{aligned} \quad (30)$$

where $Q_{t,t+s}$ is the households' stochastic discount factor and W_{t+s} stands for the nominal wage. The first part of bracket is the revenue when they cannot adjust the price optimally. Since intermediate good firms are monopolistically competitive, they

face the downward-sloping demand function given by the profit maximization of home final good producers.

The optimal price for the firms who can reset the price is given by the first order condition:

$$P_{H,t}^{new}(j) = (1 + \nu) \frac{E_t \sum_{s=0}^{\infty} \xi_H^s Q_{t,t+s} W_{t+s} Y_{H,t+s}(j)}{E_t \sum_{s=0}^{\infty} \pi^s \xi_H^s Q_{t,t+s} Y_{H,t+s}(j)} \quad (31)$$

The aggregate price index for home good is defined as:

$$P_{H,t}^{-\frac{1}{\nu}} = (1 - \xi_H) P_{H,t}^{new}(j)^{-\frac{1}{\nu}} + \xi_H (\pi \hat{P}_{H,t-1}(j))^{-\frac{1}{\nu}} \quad (32)$$

By substituting the log-linear equation of (31) into log-linearized (32), one can derive the Phillips-curve relationship between domestic inflation and marginal cost

$$\tilde{\pi}_{H,t} = \beta E_t \tilde{\pi}_{H,t+1} + \kappa_H \tilde{m}c_t \quad (33)$$

where $\kappa_H \equiv \frac{(1-\xi_H)(1-\beta\xi_H)}{\xi_H}$ and $\tilde{m}c_t = -\tilde{\lambda}_t - \alpha(\tilde{p}_{H,t} - \tilde{p}_{F,t}) - \tilde{A}_{H,t}$. Using the household optimality condition (12) and the definition of overall price equation (19), the marginal cost can be expressed as follows:

$$\begin{aligned} \tilde{m}c_t &= \tilde{w}_t - \tilde{p}_{H,t} - \tilde{A}_{H,t} \\ &= -\tilde{\lambda}_t + \tilde{p}_t - \tilde{p}_{H,t} - \tilde{A}_{H,t} \\ &= -\tilde{\lambda}_t + (1 - \alpha)\tilde{p}_{H,t} + \alpha\tilde{p}_{F,t} - \tilde{p}_{H,t} - \tilde{A}_{H,t} \\ &= -\tilde{\lambda}_t - \alpha(\tilde{p}_{H,t} - \tilde{p}_{F,t}) - \tilde{A}_{H,t} \end{aligned} \quad (34)$$

8.2.3 Domestic Importers

The importers are assumed to purchase a homogenous good produced abroad at price sold in the foreign market and convert it into a differentiated good for the home market. Analogously to the intermediate good producers, importers are monopolistically competitive and face the nominal rigidities subject to Calvo-pricing mechanism, inducing the incomplete pass-through. A fraction ξ_F of importers adjust the prices according to the steady state inflation and the remaining fraction $(1 - \xi_F)$ of importers set their prices optimally to maximize the present value of expected profits.

$$\max E_t \sum_{s=0}^{\infty} \xi_F^s Q_{t,t+s} \{ P_{F,t}(j) \pi^s C_{F,t+s}(j) - e_{t+s} P_{F,t+s}^*(j) C_{F,t+s}(j) \} \quad (35)$$

subject to

$$C_{F,t}(j) = \left(\frac{P_{F,t}(j)}{P_{F,t}} \right)^{-\frac{1+\nu}{\nu}} C_{F,t} \quad (36)$$

Importers pay $e_{t+s} P_{F,t+s}^*(j)$ in their home currency at the world market, so that the law of one price holds at the border. Under monopolistic competition, importers charge a mark-up, which generates the deviation from the law of one price in the short-run.

The first-order condition yields:

$$P_{F,t}^{new}(j) = (1 + \nu) \frac{E_t \sum_{s=0}^{\infty} \xi_F^s Q_{t,t+s} e_{t+s} P_{F,t+s}^*(j) C_{F,t+s}(j)}{E_t \sum_{s=0}^{\infty} \pi^s \xi_F^s Q_{t,t+s} C_{F,t+s}(j)} \quad (37)$$

The aggregate domestic import price is defined as:

$$P_{F,t}^{-\frac{1}{\nu}} = (1 - \xi_F) P_{F,t}^{new}(j)^{-\frac{1}{\nu}} + \xi_F (\pi \hat{P}_{F,t-1}(j))^{-\frac{1}{\nu}} \quad (38)$$

Following Monacelli (2005), I define the measure of the law-of-one-price gap (l.o.p

gap) as

$$\Psi_{F,t} = \frac{e_t P_{F,t}^*}{P_{F,t}} \quad (39)$$

With incomplete pass-through, l.o.p gap captures the endogenous deviation from purchasing power parity. It will play a role in determining the dynamics of real exchange rate.

Combining the log-linear equations of (37) and (38) yields the Phillips-curve relationship between import price inflation and the l.o.p gap:

$$\tilde{\pi}_{F,t} = \beta E_t \tilde{\pi}_{F,t+1} + \kappa_F \tilde{\psi}_{F,t} \quad (40)$$

where $\kappa_F \equiv \frac{(1-\xi_F)(1-\beta\xi_F)}{\xi_F}$. The l.o.p gap acts as a marginal cost for the importers. When the nominal exchange rate rises, it creates the persistent increases of the l.o.p gap due to the nominal rigidities, which results in the persistent increase in import good inflation.

8.3 Policy

Monetary authority sets the short-term nominal interest rate by an interest-rate feedback rule with smoothing.

$$R_t = R_{t-1}^{\rho_R} \hat{R}_t^{1-\rho_R} \exp(\varepsilon_{Rt}) \quad (41)$$

where \hat{R}_t is the target rate and ε_{Rt} is the monetary policy shock. The central bank is assumed to respond to inflation, output growth, exchange rate depreciation :

$$\hat{R}_t = R^{ss} \left(\frac{\pi_t}{\pi^{ss}} \right)^{\psi_1} \left(\frac{Y_t}{\gamma Y_{t-1}} \right)^{\psi_2} \left(\frac{e_t}{e_{t-1}} \right)^{\psi_3} \quad (42)$$

The government spending is exogenously given as

$$\log G_t = \rho_g \log G_{t-1} + \varepsilon_{gt} \quad (43)$$

8.4 Foreign Economy

Since domestic and foreign economy are symmetric, the budget constraint of a consumer in the foreign country is given by

$$P_{H,t}^* C_{H,t}^* + P_{F,t}^* C_{F,t}^* + E_t \left[Q_{t,t+1} \frac{D_{t+1}^*}{e_t} \right] \leq W_t^* L_t^* + \frac{D_t^*}{e_t} - T_t^* \quad (44)$$

where D_t^* denotes the foreign household's holdings of the portfolio denominated in home country currency. The first-order condition with respect to portfolio choice is

$$\frac{\lambda_t^*}{P_t^*} \frac{Q_{t,t+1}}{e_t} = \beta \frac{\lambda_{t+1}^*}{P_{t+1}^*} \frac{1}{e_{t+1}} \quad (45)$$

The existence of complete international state-contingent claims implies the perfect risk sharing

$$\begin{aligned} Q_{t,t+1} &= \beta \frac{\lambda_{t+1}}{\lambda_t} \frac{P_t}{P_{t+1}} \\ &= \beta \frac{\lambda_{t+1}^*}{\lambda_t^*} \frac{P_t^*}{P_{t+1}^*} \frac{e_t}{e_{t+1}} \end{aligned} \quad (46)$$

The real exchange rate is defined as:

$$S_t = \frac{e_t P_t^*}{P_t} \quad (47)$$

and the log-linearization gives the purchasing power parity (PPP) relationship :

$$\Delta \tilde{s}_t = \Delta \tilde{e}_t + \pi_t^* - \pi_t \quad (48)$$

Using the definition of the real exchange rate, equation (46) implies that the ratio of marginal utilities of home and foreign countries adjusts the real exchange rate:

$$\tilde{s}_t = \tilde{\lambda}_t^* - \tilde{\lambda}_t \quad (49)$$

Also, by substituting definitions of terms of trade, the overall inflations and the l.o.p gap into (47) , one can derive the evolution of the real exchange rate :

$$\begin{aligned} \tilde{s}_t &= \tilde{e}_t + \tilde{p}_t^* - \tilde{p}_t \\ &= \tilde{e}_t + (1 - \alpha)\tilde{p}_{F,t}^* + \alpha\tilde{p}_{H,t}^* - (1 - \alpha)\tilde{p}_{H,t} - \alpha\tilde{p}_{F,t} \\ &= \tilde{e}_t + \tilde{p}_{F,t}^* - \tilde{p}_{F,t} - \alpha(\tilde{p}_{F,t}^* - \tilde{p}_{H,t}^*) - (1 - \alpha)(\tilde{p}_{H,t} - \tilde{p}_{F,t}) \\ &= \tilde{\psi}_{F,t} - \alpha\tilde{q}_t^* - (1 - \alpha)\tilde{q}_t \end{aligned} \quad (50)$$

Under complete international asset market, households can purchase the domestic and foreign government bond denominated in each country's currency and by no arbitrage condition, the expected returns on domestic government bond and foreign government bond have to be equalized. Since the return on the foreign government bond will be the nominal interest rate and the expected exchange rate depreciation, this implies a log-linear version of an uncovered interest rate parity condition (UIP) :

$$\tilde{R}_t - \tilde{R}_t^* = E_t \Delta \tilde{e}_{t+1} \quad (51)$$

8.5 Market clearing

The resource constraints for domestic country and foreign country are given by

$$Y_{H,t} = C_{H,t} + C_{H,t}^* + G_t \quad (52)$$

$$Y_{F,t} = C_{F,t} + C_{F,t}^* + G_t^* \quad (53)$$

That is, the goods produced in home country are consumed by domestic households, $C_{H,t}$, or exported to foreign country, $C_{H,t}^*$, or consumed by government.

8.6 Exogenous Stochastic Process

Non-stationary worldwide technology follows a random-walk with drift.

$$\tilde{A}_{W,t} = \ln \gamma + \tilde{A}_{W,t-1} + \tilde{z}_t$$

$$\tilde{z}_t = \rho_z \tilde{z}_{t-1} + \varepsilon_{zt}$$

The monetary policy shocks for each country, ε_{Rt} and ε_{R^*t} , are assumed to be i.i.d. Country-specific technology shocks and government spending shocks evolve according to the following stochastic processes :

$$\tilde{A}_{H,t} = \rho_{A_H} \tilde{A}_{H,t-1} + \varepsilon_{A_H t}$$

$$\tilde{A}_{F,t} = \rho_{A_F} \tilde{A}_{F,t-1} + \varepsilon_{A_F t}$$

$$\tilde{G}_{H,t} = \rho_{G_H} \tilde{G}_{H,t-1} + \varepsilon_{G_H t}$$

$$\tilde{G}_{F,t} = \rho_{G_F} \tilde{G}_{F,t-1} + \varepsilon_{G_F t}$$

9 Learning Model

9.1 Learning Algorithm

The log-linearized model described in the previous section can be stacked in a form of

$$\Omega_0(\theta)X_{t-1} + \Omega_1(\theta)X_t + \Pi_\epsilon\epsilon_t + \Pi_\eta\eta_t = 0, \quad (54)$$

where X_t is a vector of all endogenous variables and exogenous processes, ϵ_t is a vector of stochastic innovations, and η_t consists of the rational expectations forecast errors.

By imposing the rational expectation condition, the state equation (54) has a solution written as:

$$X_t = T(\theta)X_{t-1} + H(\theta)\epsilon_t, \quad (55)$$

where T and H are non-linear functions of structural parameters θ . X_t can be decomposed into state variables, X_t^s , and forward variables, X_t^f , which appear with a lead in the model. More specifically, in the model, agents have to forecast domestic inflation and import inflation for each country, domestic consumption for their home country, the real exchange rate, and the world-wide technology shock.

In this paper, I relax the rational expectation assumption that agents have the equilibrium-consistent forecasts. Instead, I assume that agents forecast the value of forward variables, only using the limited information set $\{X_{j-1}^f\}_{j=1}^{t-1}$. Agents form their forecast following a small forecasting rule in a vector autoregressive regression form at each period with data available. The forecasting rule is assumed as:

$$X_t^f = a_{t-1} + b_{t-1}X_{t-1}^f + e_t, \quad (56)$$

where the coefficients of the forecasting rule are called the beliefs. The equation

(56) represents the “Perceived Law of Motion” (PLM) of the agents. As the new data get available, agents update their beliefs according to the constant gain learning algorithm:

$$\Sigma_t = \Sigma_{t-1} + g(Z_{t-1}Z'_{t-1} - \Sigma_{t-1}), \text{ and} \quad (57)$$

$$\phi_t = \phi_{t-1} + g\Sigma_t^{-1}(X_t^f - Z_{t-1}\phi_{t-1}), \quad (58)$$

where $Z_{t-1} \equiv [1 \ X_{t-1}^f]$, $\phi_t \equiv [a_t, b_t]'$ stacks the beliefs, Σ_t denotes the second moments, and g is the gain parameter which determines the rate at which past observations are discounted. This algorithm places more weight on recent observations and geometrically discounts the weight to $g(1 - g)^t$ after t periods. Orphanides and Williams (2005) refer to this as “perpetual learning” since agents are alert to potential structural change and forget the past. As $g \rightarrow 0$, the learning model converges to the rational expectation model.

By substituting PLM into the state equation (54), I derive the “Actual Law of Motion” (ALM):

$$X_t = C_t(\theta, \phi_{t-1}) + T_t(\theta, \phi_{t-1})X_{t-1} + H_t(\theta, \phi_{t-1})\epsilon_t.$$

Now, T_t and H_t are functions of agents’ beliefs ϕ_{t-1} , as well as structural parameters θ .

9.2 Exchange Rate Determination under Learning

Under the learning mechanism, model optimality conditions have the subjective expectations denoted by \hat{E}_t in place of the rational expectations denoted by E_t . I will discuss the equations that determine the exchange rate dynamics imposed by the

model and how the learning mechanism affects the equilibrium path.

The real exchange rate is determined by the relative marginal utility:

$$s_t = \tilde{\lambda}_t^* - \tilde{\lambda}_t,$$

where $\tilde{\lambda}_t = -\frac{\tau}{1-h\beta}(\tilde{c}_t - h\tilde{c}_{t-1}) + \frac{h\beta}{1-h\beta}\hat{E}_t[\tau(\tilde{c}_{t+1} - h\tilde{c}_t) + \tilde{z}_{t+1}]$. Due to the habit formation, the marginal utility will be affected by the deviation from the rational expectations. Since the adaptive expectations are not firmly tied to the equilibrium, the link between relative consumption and the real exchange rate will be loosened.

The model also imposes the PPP equation that the price differential determines the exchange rate, shown as:

$$\tilde{e}_t = s_t - \tilde{p}_t^* + \tilde{p}_t,$$

Since the expectation term does not directly appear in PPP equation, but only through the real exchange rate s_t , the learning mechanism will not play a separate role.

The linearized UIP condition under the learning mechanism adds an additional term to the UIP equation under rational expectations. This is shown as:

$$\begin{aligned} \tilde{R}_t - \tilde{R}_t^* &= \hat{E}_t \Delta \tilde{e}_{t+1} \\ &= E_t \Delta \tilde{e}_{t+1} + (\hat{E}_t \Delta \tilde{e}_{t+1} - E_t \Delta \tilde{e}_{t+1}) \end{aligned}$$

It is well known that the linearized UIP equation has been rejected by the data. To relax this equation, McCallum and Nelson (1999, 2000) derive the extra term in the UIP equation as a time-varying risk premium omitted by linearization, and Jeanne and Rose (2002) consider noise traders. Under the learning mechanism, the forecast

errors will disconnect the exchange rate dynamics from the interest rate differential.

10 Simulation Study

Throughout the simulation study and the empirical application, I consider the observables consisting of seven variables: output growth, inflation, interest rate for home country, output growth, inflation, interest rate for foreign country and the exchange rate depreciation. The state-space representations under rational expectations and under the learning mechanism are summarized in the following table.

Rational Expectations Model	Learning Model
$Y_t = D(\theta) + Z(\theta)X_t$	$Y_t = D(\theta) + Z(\theta)X_t$
$X_t = T(\theta)X_{t-1} + H(\theta)\epsilon_t$	$X_t = T_t(\theta, \phi_{t-1})X_{t-1} + H_t(\theta, \phi_{t-1})\epsilon_t$
	$\phi_t = \phi_{t-1} + g\Sigma_t^{-1}(Z_t - Z_{t-1}\phi_{t-1})$
	$\Sigma_t = \Sigma_{t-1} + g(Z_{t-1}Z'_{t-1} - \Sigma_{t-1})$

All simulations are based on the same structural parameters as the medians of Lubik and Schorfheide (2005) in Table 4. I generate 1000-period-long data sets using the rational expectations model solution and the learning model solution. In order to provide discipline on constant gain parameters, I consider different values of constant gain parameters, (0.001, 0.002, and 0.01) that impose the weight on the data by half after 172, 86, and 17 years. I focus on small values to ensure the stationary path around the rational equilibrium, and these gain parameters are in line with Eusepi and Preston (2011), who used the interval of 0.0013 to 0.003. To initialize the belief process, I generate presample data from the rational expectations model solution, following Huang, Liu, and Zha (2008).

Table 5 shows that the learning mechanism affects the volatilities in economics variables. For instance of “g =0.001”, the learning model moderately reduces the

volatilities of output growth, interest rate for home country and the volatility of inflation for foreign country, but increases the variability of other variables including the exchange rate depreciation. As I adapt higher values of constant gain parameters, there is a tendency for the standard deviations of variables to grow because the belief updating process fluctuates more. This result is in line with the findings of Slobodyan and Wouters (2012) that the standard deviations of variables do not increase much for small value of gain parameters, and tend to increase variability of the level variable by the so-called “exits” or “large deviations” for higher value of gain parameters.

In terms of persistence, Figure 21 provides the autocorrelation of the real exchange rate from the rational expectation and learning models. The learning mechanism generates the additional persistence in the real exchange rate which is related to the increased persistence of foreign inflation. The reason for the additional persistence is that agents’ perceived inflation and the exchange rates are more persistent than under rational expectations. When I run regressions with forecasts of the exchange rates from the learning model and rational expectations model, the autoregressive coefficients on the past forecasts is 0.93 in the learning model whereas it is 0.79 in the rational expectations model.

Table 6 gives an idea of how the learning mechanism disconnects the exchange rate from fundamentals. The UIP and PPP equation columns report the correlation of the interest rate differentials and the exchange rate depreciation and the correlation of the inflation differentials and the exchange rate depreciation, respectively. In the rational expectation model, the interest rate differential shows a significantly positive correlation, 0.26 to the exchange rate depreciation while the data implies the correlation is close to zero. The learning mechanism decreases the correlation as the gain parameter get increased. When $g = 0.002$, the learning model can generate the closest correlation to the data. The correlations of inflation differentials to depreciation are

similarly larger than found in the data. As I discussed in previous section, learning does not play a role in the PPP relationship. The Backus-Smith column shows the correlation between relative consumption and the real exchange rate. As documented in the literature, I find that this correlation is 0.03 in the data, and that the standard rational expectations model predicts a correlation of 0.97. Learning models also perform poorly in this dimension, but notice that the learning mechanism shows a tendency of decreases toward the data moment.

11 Empirical Application

The model is fitted to U.S. data for the “home country” and Euro area data for the “foreign country”. The data are quarter-to-quarter output growths, annualized inflations, annualized interest rates for both countries, and bilateral nominal exchange rate depreciation in percentage terms from 1981:Q3 to 2012:Q4. U.S. data series are obtained from the database of the Federal Reserve Bank of St. Louis (FRED) and Euro area data are from the Area-Wide Model database. This latter database became a standard reference for Euro area data after Fagan, Henry, and Metre (2005). Euro area data are the weighted aggregates of individual countries based on constant gross domestic product (GDP) at market prices for 1995[‡]. For the exchange rates data, I use the U.S./Euro exchange rate from FRED after the introduction of the euro in 1999:Q1 and construct the synthetic exchange rate from the bilateral exchange rates of countries in the Euro area for the historical data. The bilateral exchange rates of individual countries are obtained from FRED, and the historical exchange rate depreciation is given as the weighted sum of exchange rate depreciations of countries

[‡]These weights are 0.036 on Belgium, 0.283 on Germany, 0.111 on Spain, 0.201 on France, 0.015 on Ireland, 0.195 on Italy, 0.003 on Luxembourg, 0.060 on Netherlands, 0.030 on Austria, 0.024 on Portugal, 0.017 on Finland and 0.025 on Greece.

with weights underlying in the AWM database:

$$\Delta \ln E_t = \sum_{i=1}^n f_i \Delta \ln E_{i,t}.$$

Figure 22 plots the nominal depreciation and the inflation and interest rate differentials. The shaded section in the top panel indicates the recent financial crisis, while the bottom panel magnifies the inflation differential and the interest rate differential. Notice the different patterns before and after the crisis. The nominal depreciation movement seems to be more correlated with fundamentals after the crisis. The correlation between the depreciation and inflation differential goes from 0.18 up to 0.34, and the correlation between the depreciation and interest rate differential changes even more, going from -0.11 to 0.17.

Initialization of Estimation I use the prior density from Lubik and Schorfheide (2005) for both the rational expectations model and the learning model. The import share parameter, α , is fixed at 0.13 since the data do not have information that can fit the trade. The gain parameter is set to 0.002. To get regression-based initial beliefs as in the simulation study, I use the smoothed states from the Kalman filter using pre-sample data from 1973:Q1 to 1981:Q2

$$X_{0|0}^f = a_0 + b_0 X_{-1|0}^f,$$

where $X_{0|0}^f = \left\{ X_{t|0}^f \right\}_{t=-k}^0$ and $X_{-1|0}^f = \left\{ X_{t-1|0}^f \right\}_{t=-k}^0$. In order to get the smoothed states, I run the rational expectation model with pre-sample data and save the mean and variance of states from the Kalman filter. In terms of Hamilton (1994), $X_{t|t} = E(X_t|I_t)$, $P_{t|t} = Var(X_t|I_t)$, $X_{t|t-1} = E(X_t|I_{t-1})$, and $P_{t|t-1} = Var(X_t|I_{t-1})$. Smoothed states are obtained by iterating the following algorithm from the end of

sample:

$$\begin{aligned}
 X_{t|T} &= X_{t|t} + P_{t|t}T'P_{t+1|t}^{-1}(X_{t+1|t} - TX_{t|t}) \\
 P_{t|T} &= P_{t|t} + P_{t|t}T'P_{t+1|t}^{-1}(P_{t+1|T} - P_{t+1|t})P_{t+1|t}^{-1}TP_{t|t}.
 \end{aligned}$$

11.1 Posterior Distribution

Posterior densities from a Bayesian estimation for the rational expectations model and the learning model are summarized in Table 7 along with the prior densities. Based on the random-walk Metropolis-Hastings algorithm, posterior distributions are explored by generating 100,000 draws and burning the initial 90% of draws.

There are some remarkable changes in estimated parameters. First, I find that the Calvo parameters for U.S. drop significantly in the learning model. The estimated nominal rigidity parameter for domestic good in the U.S. under rational expectations is 0.41, suggesting 1.7 quarters of price stickiness. Under constant-gain learning mechanism, that parameter is 0.16, suggesting 1.2 quarters of price stickiness. The nominal rigidity parameter for the import sector in the U.S. is significantly reduced to 0.09 (1-quarter price stickiness) with the learning mechanism, compared to 0.57 (2.3-quarter price stickiness) under rational expectations. The price stickiness parameters in the Euro area do not change much either in domestic production or in the import sector. These findings are in line with Vilagi (2007) that learning does not replace the structural sources of persistence in the Euro area. Second, the degree of habit formation significantly drops from 0.23 to 0.02. Third, the autoregressive coefficients of structural shocks decrease with the exception of Euro-area specific technology shocks. As a flip-side of the simulation study, learning replaces the structural sources of persistence and generates the endogenous persistence. Last, I obtain a reduction in the standard

deviations of four out of seven structural shocks. This implies that learning can add volatilities to variables as found in simulation exercises. The non-structural PPP shock which is designed to capture the model misspecification decreases slightly, but not significantly, suggesting that the improvement of the learning model is limited in explaining the excess volatility of the exchange rate. Also, two model specifications similarly match these moments. This can give support to the learning mechanism since the learning model predictions are generated with smaller structural persistence parameters and smaller exogenous shocks.

11.2 Marginal Data Density Comparison

As is standard in the Bayesian framework, I evaluate which model fits the data better by comparing the marginal data density. Due to the computational difficulty of the marginal data density, I adapt the modified Harmonic mean estimator suggested by Geweke (1999). The Harmonic mean estimator is based on the following identity

$$p(Y) = \left[\int \frac{f(\theta)}{L(Y|\theta)p(\theta)} p(\theta|Y) d\theta \right]^{-1},$$

where $p(\theta)$ is the prior density of θ and $f(\theta)$ is a function satisfying $\int f(\theta) d\theta = 1$. Then, the sample correspondence is given by:

$$\hat{p}_J(Y) = \left[\frac{1}{J} \sum_{j=1}^J \frac{f(\theta_j)}{L(Y|\theta_j)p(\theta_j)} \right]^{-1},$$

where θ_j are draws from the posterior distribution, $p(\theta|Y)$.

Table 8 reports the log marginal data densities (MDD) from the rational expectations and learning models with different gain parameters (0.001, 0.002, and 0.01).

The learning model with a constant gain of 0.002 delivers the log MDD of -1372.3, while the log MDD of the rational expectations model equals -1416. For the other values of constant gain parameters, I find the dominance of learning models. This implies that the data favor the constant gain learning specification over the rational expectations specification. Since the benchmark model with constant gain of 0.002 provides the best fit, it relieves the limitations of fixing the gain parameter instead of estimating it.

11.3 Impulse Response Function

The impulse response functions to the U.S. monetary policy shock from rational expectation model and the learning model are shown in from Figure 23 to Figure 26. Note that all the variables react with significant lags in learning model. It takes more than 20 quarters for the exchange rate depreciation to be adjusted in the learning model, visible as the hump-shaped response, while in the rational expectations model it instantly drops and dissipates. In the learning model, when the unanticipated monetary shock increases the interest rate, agents do not immediately recognize the source of the fluctuations. It takes time for agents to realize the nature of the shock based on the equilibrium relations of variables in their forecasting rules. Therefore, the responses are small at the beginning and propagate through the expectation formation. In the learning model, monetary policy shocks are absorbed in the expectation, and do not produce real effects on consumption and output. Instead, expectation affects the price movement, and thus the exchange rate. The responses of output to the monetary policy are contrary to that which closed-economy models predict. I also find that the effects of the Euro area monetary policy shock on the U.S. economy and the exchange rate are not significant.

11.4 Variance Decomposition

I report the variance decomposition results from the rational expectations model in Table 8 and from the learning model in Figure 27. The unconditional variance decomposition results are calculated with the medians of posterior density. Variance decomposition is obtained by solving the following discrete Lyapunov equations:

$$\begin{aligned} \text{Var}(X_t|\theta) &= T(\theta)\text{Var}(X_t|\theta)T(\theta)' + H(\theta)Q(\theta)H(\theta), \text{ and} \\ \text{Var}(Y_t|\theta) &= Z(\theta)\text{Var}(X_t|\theta)Z(\theta)', \end{aligned}$$

where $Q(\theta)$ is a variance-covariance matrix of structural shocks. The contribution of shock i is obtained by setting the volatility of all disturbances except the corresponding shock i to zero.

Under both the rational expectations and learning specifications, the non-structural PPP shock explains most of the exchange rate fluctuation. The second biggest contribution in the rational expectation model is made by the Euro and U.S. monetary policy shocks, followed by world-wide technology shock. In contrast, learning model attributes the movement of depreciation to the real shocks, U.S. technology shock, and world-wide technology shock by approximately 6%.

11.5 Posterior Predictive Checks

To assess the absolute fit of models, I conduct posterior predictive checks. Given the posterior densities from the rational expectations and the learning models, I simulate 10,000 samples of observations, each of which is 126 periods in length for each set of draws from posterior densities. These are known as predictive densities. I plot the kernel distributions to determine how well models with different expectations

specifications fit data statistics in Figure 28. In the first row, I show the correlation of output growths across countries, and the correlations of output growth and inflation in both the U.S. and Euro areas. Solid lines indicate predictive densities from the learning model, the dashed lines represent results from the rational expectations model, and the red dotted line denotes the actual data moments. The solid lines are clearly closer to the dotted "actual" moments than the dashed lines, indicating that the learning model explains better the business cycle fluctuation. The second row shows the exchange rate determination equations, the UIP relation, PPP relation, and Backus-Smith relation. Since I do not use consumption data, I calculate the correlation of real exchange rate depreciation and output growth differentials for the Backus-Smith relation. The results from the rational expectations and learning models are similar, though rational expectations model sometimes performs slightly better. This is not inconsistent with the simulation study because simulations are conducted under different parameterization for predictive checks. Thus, the performance of the learning model results from different structural parameters as well as the different expectation mechanism. The last row displays the standard deviation and the persistence of the real exchange rate depreciation.

12 Conclusion

I specify and estimate two-country New Open Economy Models under rational expectations and learning mechanism. A major feature of the learning model is that agents are assumed to use econometric models with historical data to make their forecasts as well as update their forecasts over time via a constant-gain learning algorithm. I find from a simulation study that the learning mechanism increases the volatilities of variables as the gain parameter rises. The self-fulfilling property of the belief-updating

process engenders a more persistent exchange rate and more persistent inflation series, as shown in Orphanides and Williams (2005). A Bayesian estimation with U.S. and Euro-area data implies that the learning mechanism substitutes the structural persistence sources of the model, e.g., the nominal price stickiness for the U.S. and habit formation, since learning generates additional persistence. Model comparison based on the marginal data density favors the learning specification over rational expectations, although the improvement in terms of fitting the observed Dollar-Euro exchange rate dynamics is limited. The performance of the learning mechanism surpasses the rational expectations model in explaining some business cycle statistics, such as output growth comovement between the U.S. and the Euro area and the correlation between output growth and inflation of the U.S. Impulse response functions from the learning model are lagged and persistent, since agents realize the source of fluctuation with a lag and adjust their forecasts in a recursive manner.

Yet the results from the learning model rely on the specification of agents' forecasting rules and the choice of initial belief. It is worth examining with different types of forecasting rules and different initial belief schemes. It would also be a plausible extension of the learning model to consider the endogenous switching of forecasting models based on the forecasting model performance, as in Lewis and Markiewicz (2009).

13 Tables and Figures

Table 4: Summary of Parameters for Simulation

Coefficient	Description	Value
ξ_H	Calvo parameter of domestic goods for home country	0.66
ξ_F	Calvo parameter of imported goods for home country	0.56
ξ_H^*	Calvo parameter of imported goods for foreign country	0.86
ξ_F^*	Calvo parameter of domestic goods for foreign country	0.76
α	import share	0.13
h	consumption habit	0.41
τ	relative risk aversion coefficient	3.76
μ	intratemporal substitution elasticity between home and foreign good	0.43
ψ_1	response to inflation gap for home country	1.41
ψ_2	response to output growth for home country	0.66
ψ_3	response to exchange rate depreciation for home country	0.03
ψ_1^*	response to inflation gap for foreign country	1.37
ψ_2^*	response to output growth for foreign country	1.27
ψ_3^*	response to exchange rate depreciation for foreign country	0.03
ρ_{A_H}	AR coefficient for country-specific technology shock of H	0.83
ρ_R	AR coefficient for monetary policy of H	0.76
ρ_g	AR coefficient for government spending shock of H	0.90
$\rho_{A_F}^*$	AR coefficient for country-specific technology shock of F	0.85
ρ_R^*	AR coefficient for monetary policy of F	0.84
ρ_g^*	AR coefficient for government spending shock of F	0.94
ρ_z	AR coefficient for world-wide technology shock	0.60
r^A	steady state interest rate	0.86
γ	world-wide technology growth rate	0.39
π^A	steady state inflation	3.16
σ_A	standard deviation of country-specific shock for H	1.66
σ_R	standard deviation of monetary policy shock for H	0.18
σ_g	standard deviation of government spending shock for H	0.50
σ_{A^*}	standard deviation of country-specific shock for F	2.61
σ_{R^*}	standard deviation of monetary policy shock for F	0.18
σ_{g^*}	standard deviation of government spending shock for F	0.62
σ_z	standard deviation of world-wide technology shock	0.35

Table 5: Standard Deviation of Simulated Models

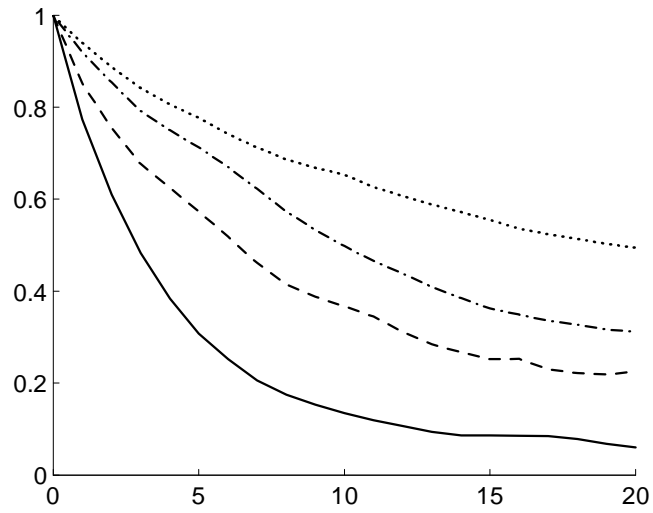
	R.E. model	Learning model		
		$g = 0.001$	$g = 0.002$	$g = 0.01$
Output growth for H	0.57	0.57	0.57	0.62
Inflation for H	1.22	1.64	1.79	2.28
Interest rate for H	1.53	2.40	2.71	2.57
Output growth for F	0.62	0.63	0.63	0.64
Inflation for F	0.32	0.33	0.41	0.61
Interest rate for F	0.68	0.69	0.64	0.92
Depreciation	1.64	1.71	1.80	1.89
Real exchange rate	1.17	1.65	1.69	1.70

Table 6: Correlation of Depreciation to Inflation and Interest Rate Gap

	UIP equation	PPP equation	Backus-Smith
Data	0.03	0.19	0.03
Rational Expectations	0.26	0.54	0.97
Learning Model $g = 0.001$	0.23	0.52	0.95
$g = 0.002$	0.05	0.58	0.93
$g = 0.01$	-0.03	0.72	0.92

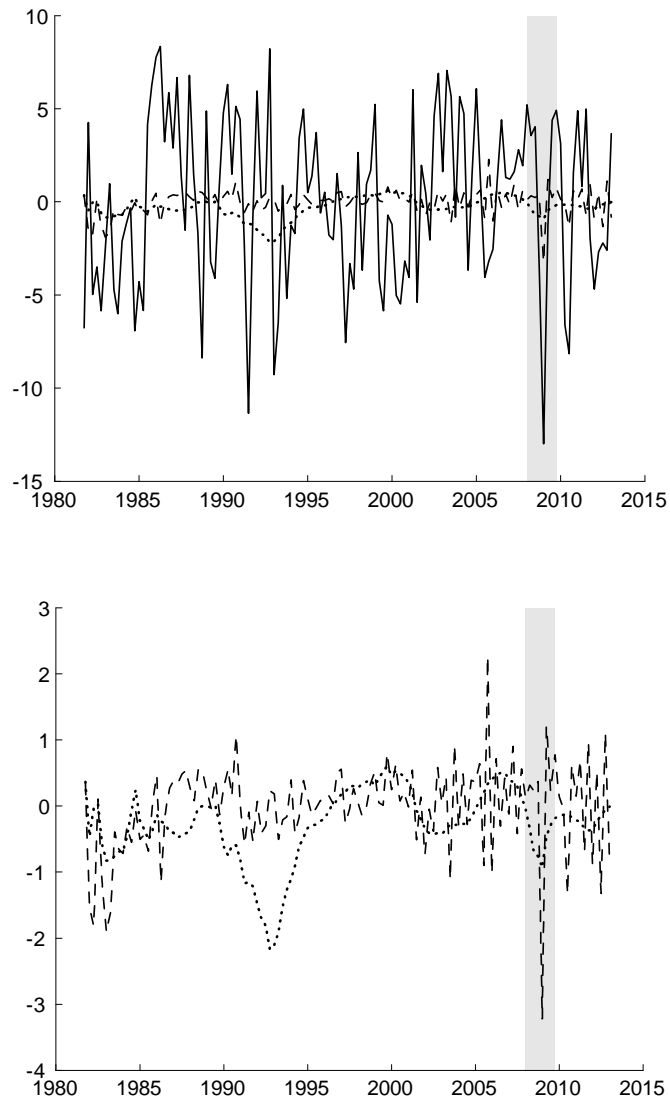
Note. UIP column reports the correlation between depreciation and the interest rate gap (i.e., $Corr(\Delta e_{t+1}, R_t - R_t^*)$). The PPP column reports the correlation between depreciation and the inflation gap (i.e., $Corr(\Delta e_t, \pi_t - \pi_t^*)$). The Backus-Smith column shows the correlation between relative consumption and the real exchange rate (i.e., $Corr(s_t, c_t - c_t^*)$).

Figure 21: Autocorrelation Function of the Real Exchange Rate from Simulation



Note. The solid line shows results from the rational expectations model. The dashed line denotes results from the learning model where $g = 0.001$; the dash-dot line for $g = 0.002$; and the dotted line for $g = 0.01$.

Figure 22: Depreciation and Inflation, Interest Rate Gap



Note. The top panel shows three series: the solid line shows nominal depreciation; the dotted line shows the inflation gap; and the dashed line shows the interest rate gap. The bottom panel shows a magnification of the inflation gap and the interest rate gap.

Table 7: Prior and Posterior Distribution

Coefficient	Prior			R.E. Model		Learning Model	
	Density	Mean	Std	Median	90% Interval	Median	90% Interval
ξ_H	Beta	0.50	0.15	0.41	[0.34 0.47]	0.16	[0.07 0.26]
ξ_F	Beta	0.50	0.15	0.57	[0.45 0.69]	0.09	[0.03 0.15]
ξ_H^*	Beta	0.75	0.15	0.84	[0.75 0.93]	0.84	[0.71 0.95]
ξ_F^*	Beta	0.75	0.15	0.62	[0.57 0.65]	0.60	[0.52 0.69]
h	Beta	0.30	0.10	0.23	[0.09 0.38]	0.02	[0.01 0.03]
τ	Gamma	2.00	0.50	3.39	[3.31 3.49]	4.64	[3.97 5.34]
μ	Gamma	1.00	0.50	0.52	[0.39 0.60]	0.11	[0.04 0.21]
ψ_1	Gamma	1.50	0.25	1.01	[1.00 1.04]	1.53	[1.22 1.82]
ψ_2	Gamma	0.50	0.25	0.53	[0.45 0.66]	0.77	[0.33 1.28]
ψ_3	Gamma	0.10	0.05	0.02	[0.00 0.03]	0.04	[0.02 0.06]
ψ_1^*	Gamma	1.50	0.25	1.02	[1.01 1.04]	1.43	[1.14 1.70]
ψ_2^*	Gamma	0.50	0.25	1.16	[1.08 1.32]	1.01	[0.55 1.56]
ψ_3^*	Gamma	0.10	0.05	0.06	[0.03 0.08]	0.03	[0.02 0.06]
ρ_{AH}	Beta	0.80	0.10	0.98	[0.96 0.99]	0.88	[0.79 0.95]
ρ_R	Beta	0.50	0.20	0.77	[0.74 0.80]	0.93	[0.91 0.95]
ρ_g	Beta	0.80	0.10	0.97	[0.95 0.99]	0.95	[0.91 0.98]
ρ_{AF}^*	Beta	0.60	0.20	0.87	[0.75 0.97]	0.92	[0.85 0.97]
ρ_R^*	Beta	0.50	0.20	0.82	[0.79 0.84]	0.94	[0.92 0.96]
ρ_g^*	Beta	0.80	0.10	0.97	[0.96 0.99]	0.92	[0.85 0.97]
ρ_z	Beta	0.66	0.15	0.83	[0.78 0.88]	0.82	[0.72 0.92]
r^A	Gamma	0.50	0.50	0.49	[0.35 0.57]	0.52	[0.01 1.20]
γ	Normal	0.40	0.20	0.37	[0.31 0.43]	0.44	[0.32 0.58]
π^A	Gamma	7.00	2.00	3.24	[3.09 3.35]	5.09	[3.58 6.44]
σ_{AH}	InvGamma	1.00	4.00	1.13	[0.96 1.27]	1.31	[1.01 1.66]
σ_R	InvGamma	1.00	4.00	0.23	[0.20 0.26]	0.18	[0.16 0.20]
σ_g	InvGamma	0.40	4.00	0.75	[0.67 0.84]	0.60	[0.52 0.68]
σ_{AF}	InvGamma	0.40	4.00	1.99	[1.78 2.32]	2.48	[1.50 4.56]
σ_{R^*}	InvGamma	1.00	4.00	0.19	[0.16 0.22]	0.12	[0.11 0.13]
σ_{g^*}	InvGamma	0.20	4.00	0.57	[0.51 0.64]	0.64	[0.55 0.73]
σ_z	InvGamma	0.50	4.00	0.25	[0.21 0.30]	0.21	[0.18 0.24]
σ_e	InvGamma	3.50	4.00	4.49	[4.34 4.70]	4.45	[4.00 4.98]

Table 8: Marginal Data Density

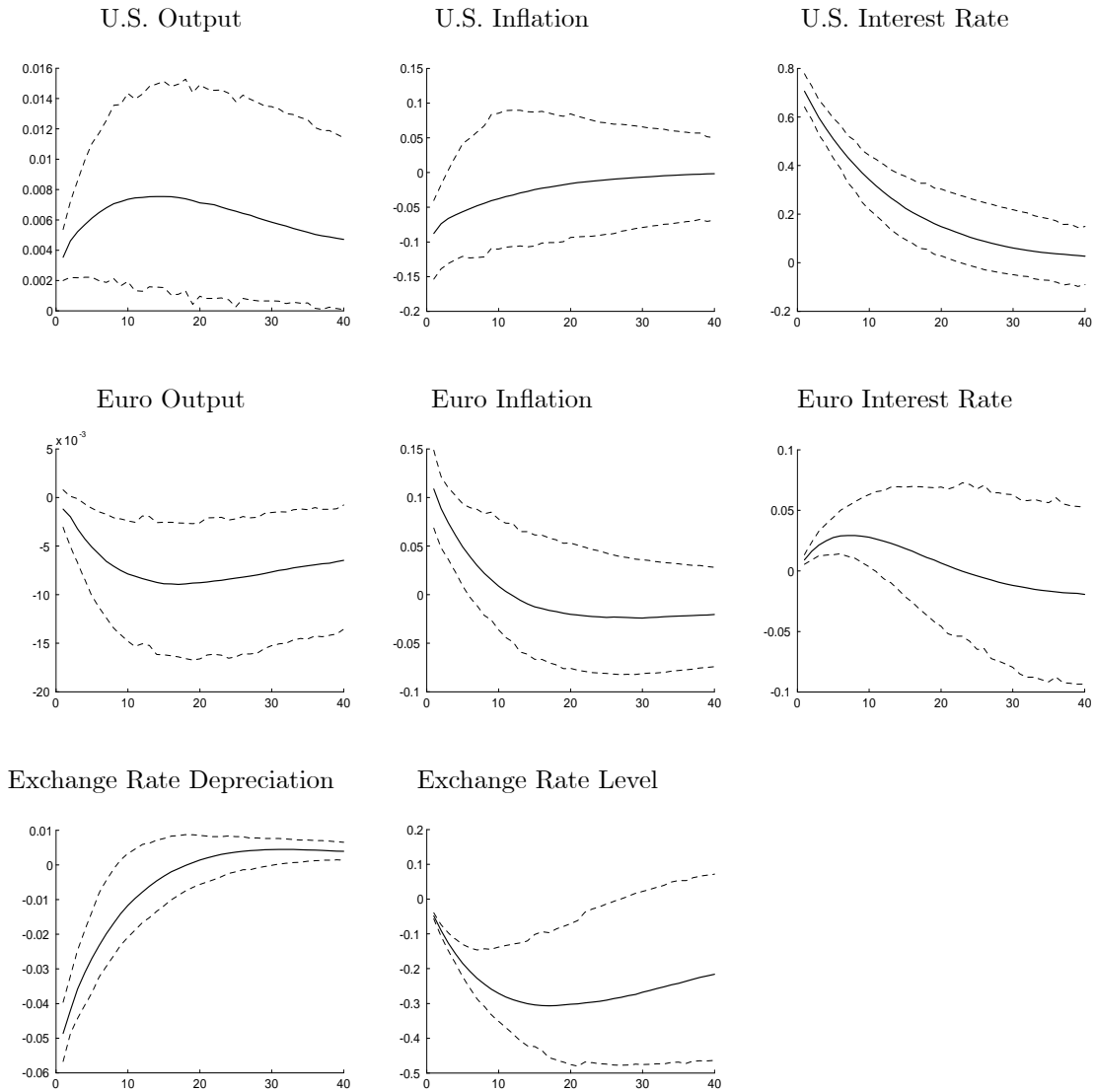
	R.E. Model	Learning Model		
		$g = 0.001$	$g = 0.002$	$g = 0.01$
MDD	-1416.0	-1377.3	-1372.3	-1397.4

Table 9: Variance Decomposition from the Rational Expectations Model

	output	inflation	int rate	output*	inflation*	int rate*	Depreciation
monetary policy	0.007	0.320	0.122	0.019	0.0444	0.115	0.042
technology	0.067	0.614	0.752	0.069	0.2029	0.031	0.005
govn't spending	0.918	0.083	0.191	0.002	0.2242	0.557	0.004
monetary policy*	0.002	0.006	0.013	0.151	0.155	0.077	0.044
technology*	0.001	0.036	0.092	0.046	0.015	0.027	0.000
govn't spending*	0.001	0.009	0.016	0.386	0.130	0.407	0.008
world-wide tech	0.002	0.053	0.076	0.307	0.322	0.038	0.036
PPP shock	0.000	0.002	0.000	0.011	0.014	0.004	0.861

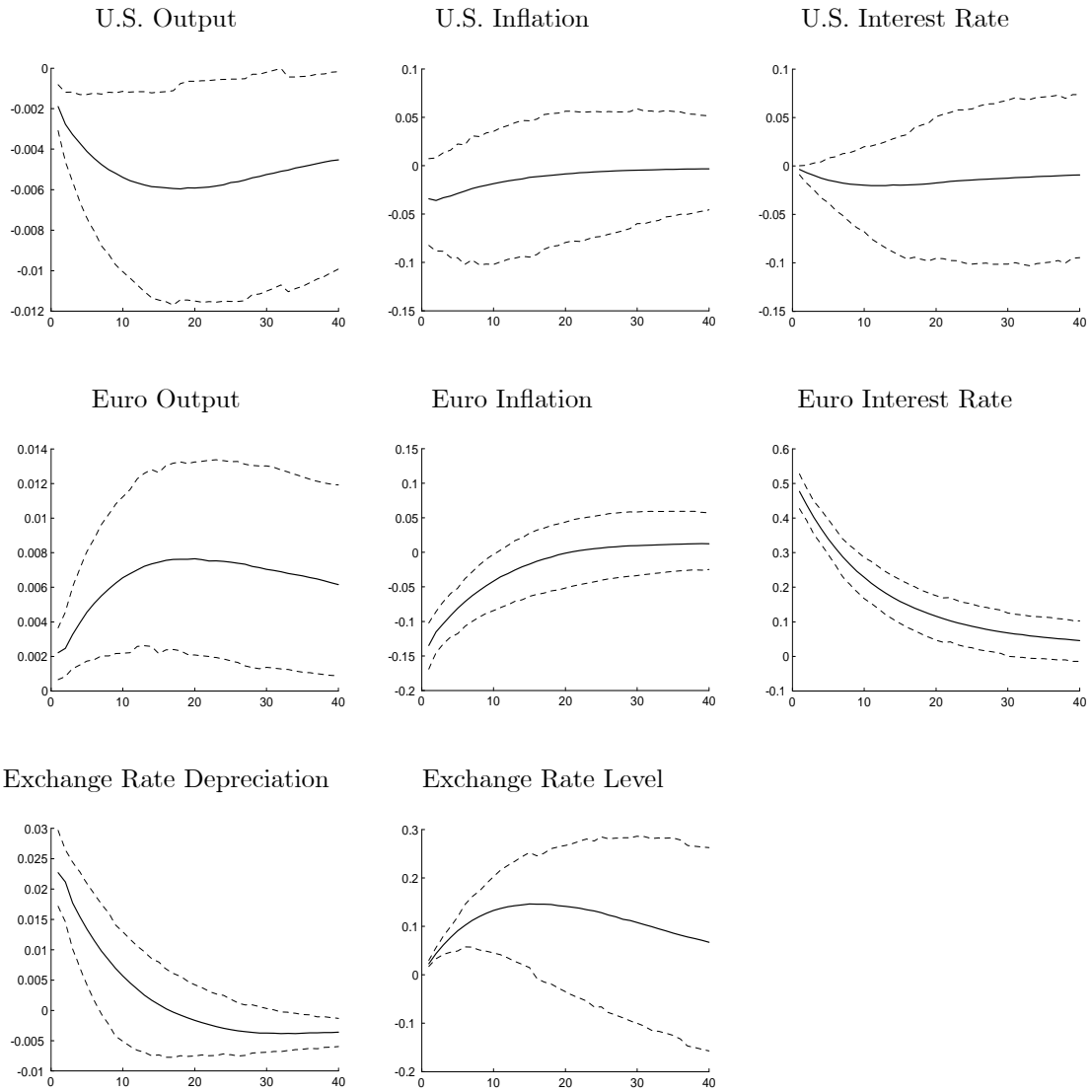
Note: The asterisks denote the foreign(Euro-area) variables.

Figure 23: Impulse Response to U.S. Monetary Policy Shock for the Learning Model



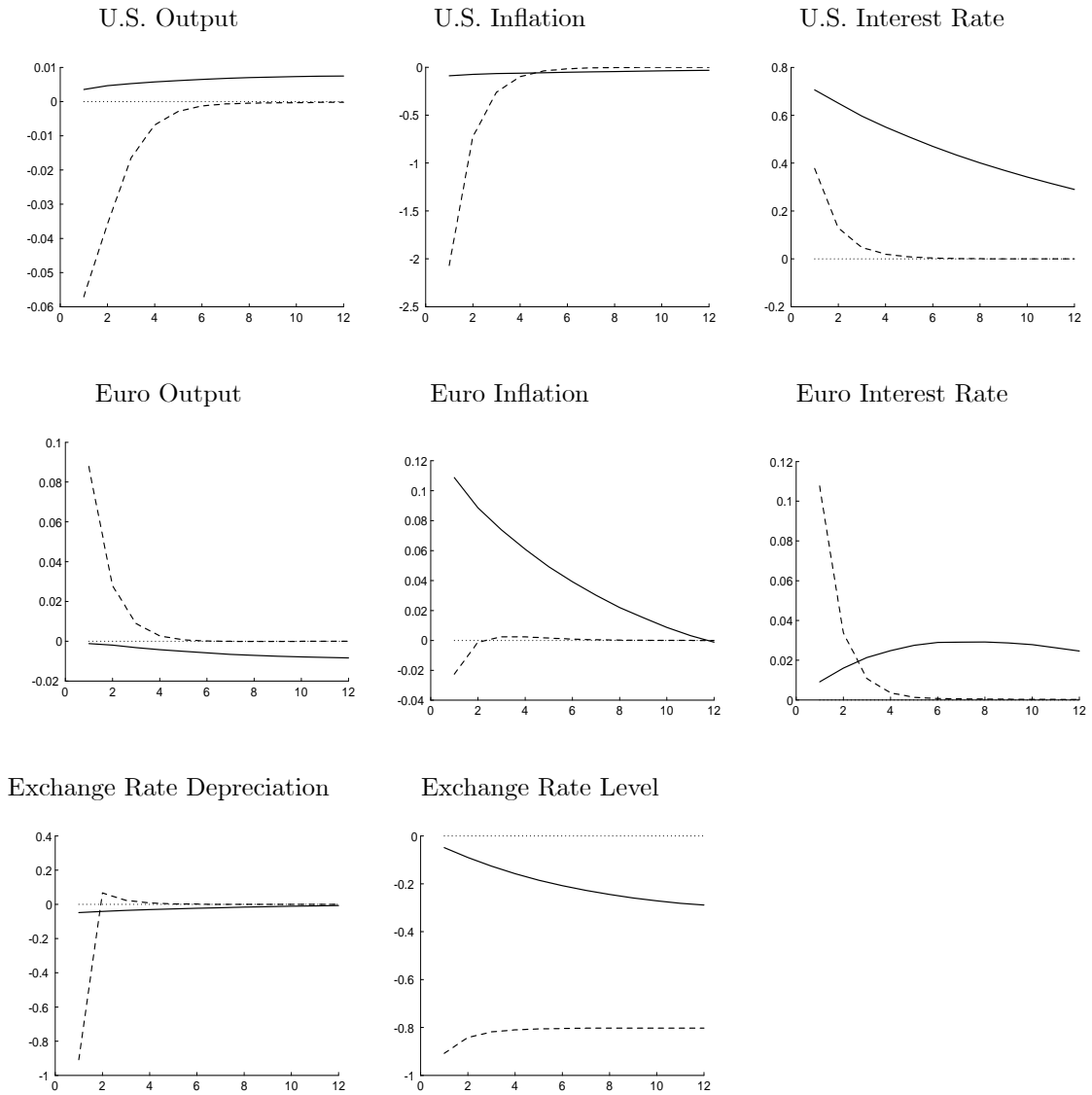
Note. Solid lines represent the median and dotted lines represent the 90% confidence interval of impulse response functions. All responses are in percentage and those for inflation and interest rate are annualized.

Figure 24: Impulse Response to Euro Monetary Policy Shock for the Learning Model



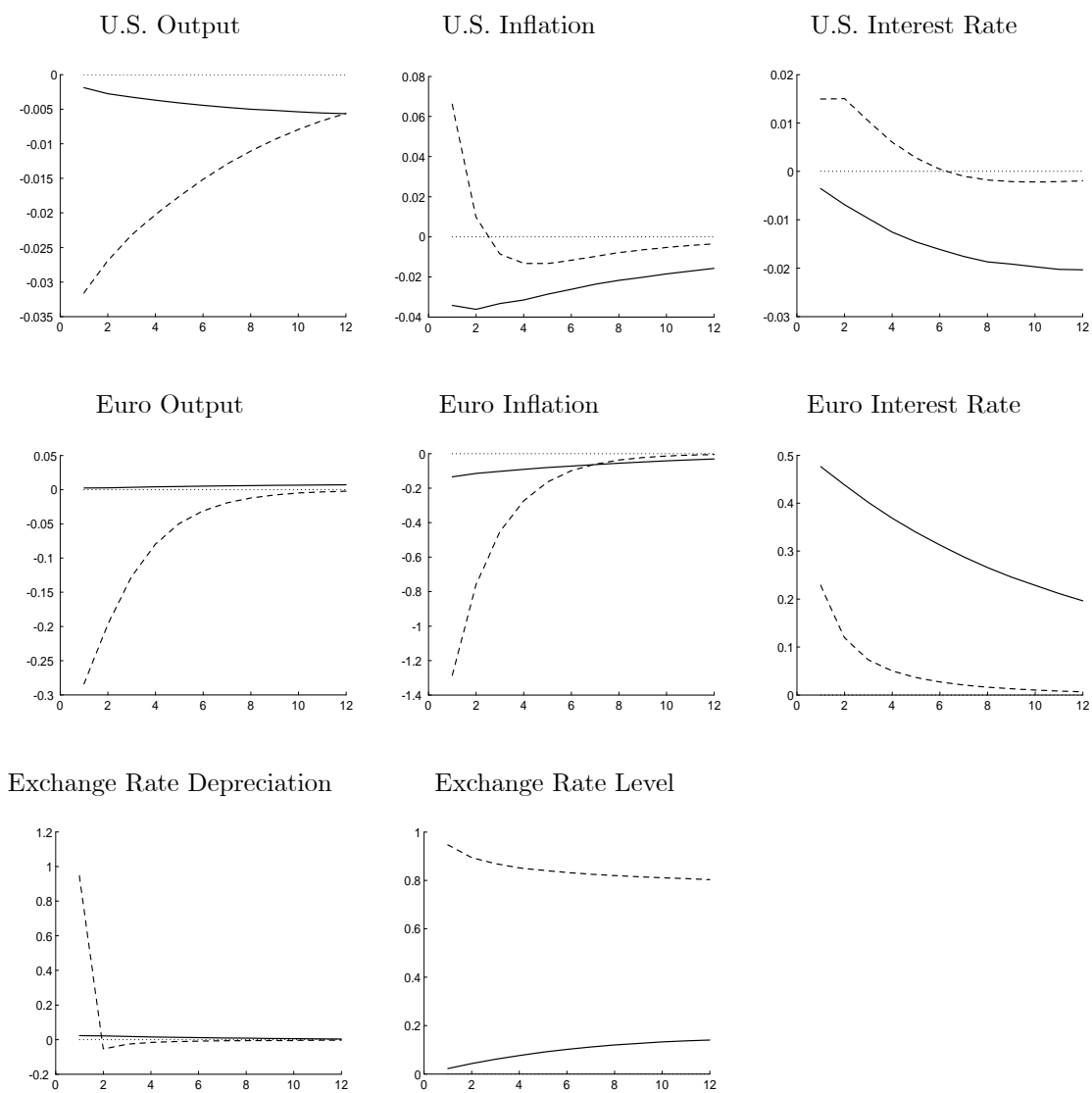
Note: Solid lines represent the median and dotted lines represent the 90% confidence interval of impulse response functions. All responses are in percentage and those for inflation and interest rate are annualized.

Figure 25: Impulse Response to U.S. Monetary Policy Shock



Note. Solid lines correspond to results from the learning model, and dashed lines correspond to results from the rational expectations model. All responses are in percentage and those for inflation and interest rate are annualized.

Figure 26: Impulse Response to Euro Monetary Policy Shock



Note: Solid lines correspond to results from the learning model, and dashed lines correspond to results from the rational expectations model. All responses are in percentage and those for inflation and interest rate are annualized.

Figure 27: Variance Decomposition for Depreciation in the Learning Model

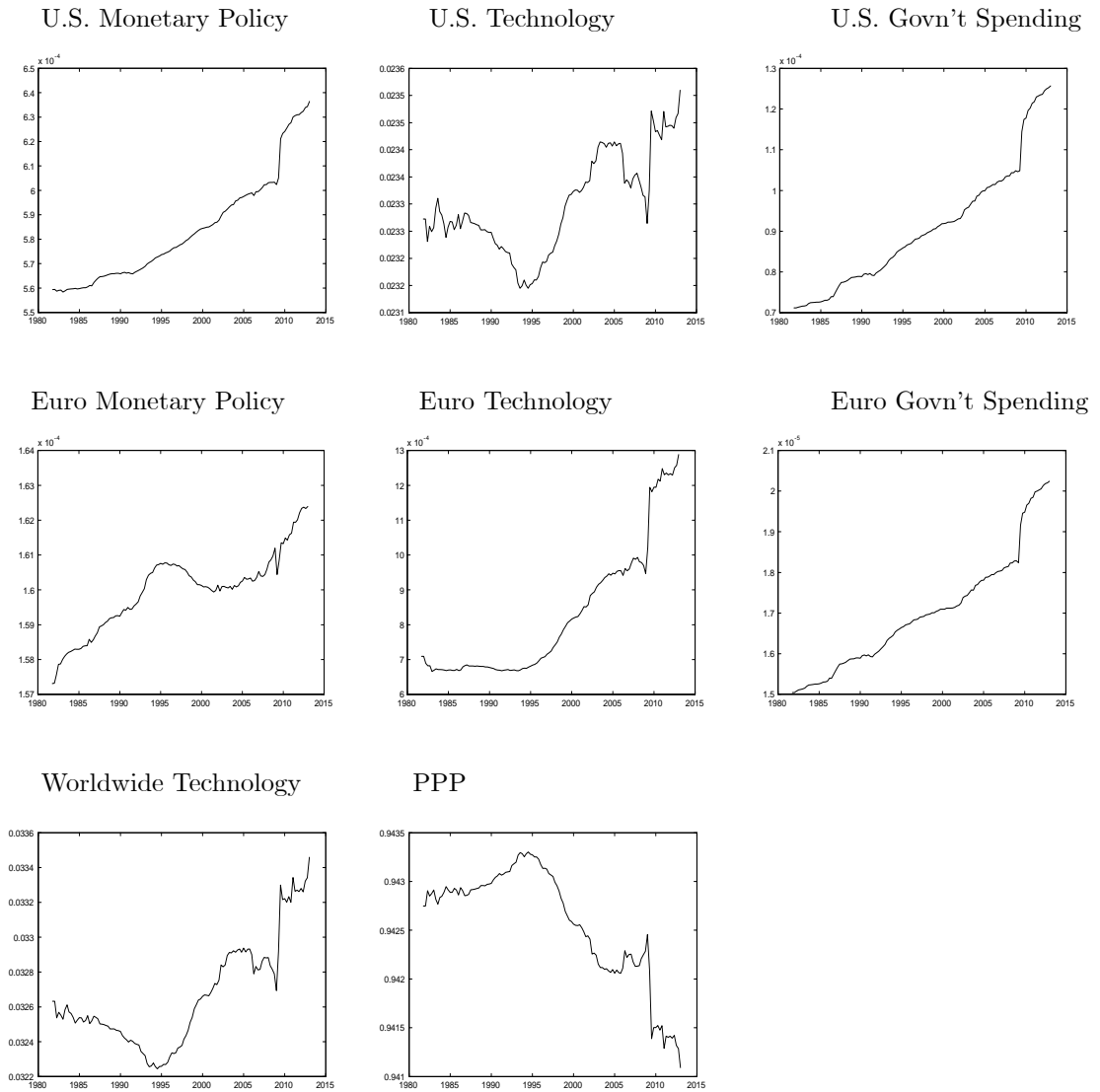
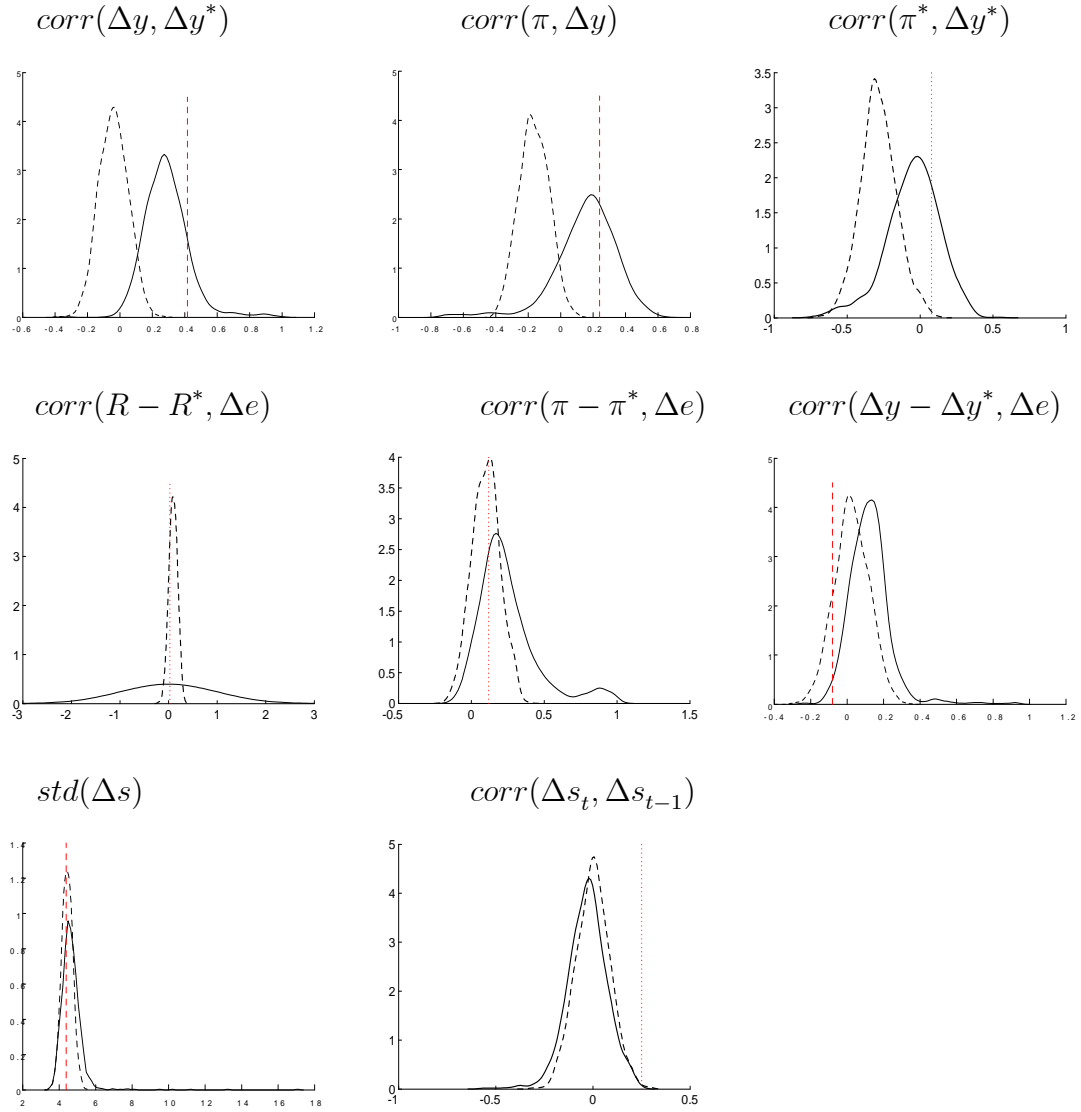


Figure 28: Posterior Predictive Checks



Note: Solid lines represent kernel densities from the learning model; the dashed lines represent kernel densities from the rational expectations model; and the red dotted vertical lines are actual data moments.

A Appendices

A.1 Estimation Algorithm

A.1.1 Stochastic Volatility in DSGE Models

A state-space representation of log-linearized DSGE model is given by:

$$y_t = D + Z\alpha_t \quad (\text{ME})$$

$$\alpha_t = T(\theta)\alpha_{t-1} + R(\theta)\eta_t \quad (\text{TE})$$

The stochastic volatilities for each structural shock (j) are, where $j \in \{z, p, b, \varphi, \mu, R, g\}$,

$$\begin{aligned} \eta_{jt} &= \sigma_{jt}\varepsilon_{jt} \\ \log \sigma_{jt} &= (1 - \rho_{\sigma_j}) \log \sigma_j + \rho_{\sigma_j} \log \sigma_{jt-1} + \nu_{jt}, \quad \nu_{jt} \sim N(0, \omega_j^2) \end{aligned}$$

In order to reduce the number of free parameters, I assume that the stochastic volatilities are random walk processes by setting the autoregressive coefficients $\rho_{\sigma_j} = 1$. Let the element of vector h_t be $h_{jt} = \log(\sigma_{jt})$. Then, $H^T = [h_1, \dots, h_T]'$ is a T-by-7 vector matrix. Similarly, denote the sample of structural shocks as $\eta^T = [\eta_1, \dots, \eta_T]'$. Collect the remaining coefficients of the stochastic volatility processes by V , where j^{th} element is ω_j^2 .

Initialization of the Algorithm. Generate $\{H^{T0}, V^{T0}, \theta^0, \eta^{T0}\}$. H^{T0} is constructed recursively using the posterior median value of the time-invariant standard deviations for the structural shocks as initial values. V^{T0} is squared values of time-invariant standard deviations. $\theta^0 \sim N(\hat{\theta}, c\Sigma^{-1})$ where $\hat{\theta}$ and Σ^{-1} are posterior mode and variance respectively, obtained by using Chris Sims's maximization algorithm. c is a scaling

parameter (set to 2). Finally, η^{T_0} is obtained by using simulation smoother developed by Durbin and Koopman (2002), which will be explained later.

Simulation Smoother by Durbin and Koopman (2002). Consider the linear Gaussian form where $\varepsilon_t \sim N(0, H_t)$ and $\eta_t \sim N(0, Q_t)$. Note that I assume $y_t = y_t - D$ and $H_t = 0$.

$$\begin{aligned} y_t &= Z\alpha_t + \varepsilon_t & (*) \\ \alpha_t &= T(\theta)\alpha_{t-1} + R(\theta)\eta_t \end{aligned}$$

The distribution of w is $w \sim N(0, \Omega)$, $\Omega = \text{diag}(H_1, Q_1, \dots, H_T, Q_T)$.

Step 1 Draw a random vector w^+ from density $p(w)$ and use it to generate y^+ by means of recursion (*) with w replaced by w^+ , where the recursion is initialized by the draw $\alpha_1^+ \sim N(a_1, P_1)$.

Step 2 Compute $\hat{w} = E(w|y)$ and $\hat{w}^+ = E(w^+|y^+)$ by means of standard Kalman Filtering and disturbance smoothing:

Kalman Filter

$$\begin{aligned} v_t = y_t - Za_t \quad F_t = ZP_tZ' + H_t \quad K_t = TP_tZ'F_t^{-1} \\ L_t = T - K_tZ \quad a_{t+1} = Ta_t + K_tv_t \quad P_{t+1} = TP_tL_t' + RQ_tR' \end{aligned}$$

Disturbance Smoothing

$$\begin{aligned} \hat{w}_t &= \begin{bmatrix} H_tF_t^{-1} & -H_tK_t' \\ 0 & Q_tR' \end{bmatrix} \begin{pmatrix} v_t \\ 0 \end{pmatrix} \\ r_{t-1} &= Z'F_t^{-1}v_t + L_t'r_t \end{aligned}$$

Step 3 Take $\tilde{w} = \hat{w} - \hat{w}^+ + w^+$.

Random Walk Metropolis Hastings Algorithm. At the beginning of each iteration g , given $\{H^{Tg-1}, \theta^{g-1}\}$

Step 1 Draw a candidate, θ^* from proposal distribution, $N(\theta^{g-1}, c_1 \Sigma^{-1})$

Step 2 Evaluate the likelihood at the candidate θ^* , conditional on H^{Tg-1} :

$$r = \min \left\{ \frac{L(Y|\theta^*, H^{Tg-1})}{L(Y|\theta^{g-1}, H^{Tg-1})}, 1 \right\}$$

accept the proposal with probability r or keep θ^{g-1} with $1 - r$ probability

Step 3 Sample the structural innovations η^{Tg} using simulation smoother conditional on θ^g

Step 4 Draw the stochastic volatilities using η^{Tg} .

Denote $\tilde{\eta}_{jt} = \log[\hat{\eta}_{jt}^2 + 0.001]$ and $e_{jt} = \log(\varepsilon_{jt}^2)$. This leads to the following approximating state-space form:

$$\begin{aligned} \tilde{\eta}_{jt} &= 2h_{jt} + e_{jt} \\ h_{jt} &= h_{jt-1} + \nu_{jt} \end{aligned}$$

Since the innovations in the measurement equation follow $\log\chi^2(1)$, I transform the system in a Gaussian one using a mixture of normal approximations as described in Kim, Shephard, and Chib (1998):

$$f(e_{jt}) = \sum_{k=1}^7 q_k f_N(e_{jt} | s_{jt} = k)$$

where s_{jt} is the indicator variable selecting which member of the mixture of normals has to be used at time t for the innovation j , $q_k = \Pr(s_{jt} = k)$, and f_N is the pdf of

a normal distribution.

w	$q_j = \Pr(w = j)$	m_j	r_j^2
1	0.00730	-10.12999	5.79596
2	0.10556	-3.97281	2.61369
3	0.00002	-8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	-1.08819	1.26261

Source : Kim, Shepard, and Chib (1998)

Conditional on s^{Tg} , the system has an approximate linear Gaussian state-space form. Hence, standard algorithm can be applied. A new draw for the complete history of H^{Tg} can be obtained recursively with the standard Gibbs sampler for state-space forms using the forward-backward recursion of Carter and Kohn (1994). Having generated H^{Tg} , the elements of the vector V^{Tg} can be generated from Normal inverse-Gamma distributions. A new sample of indicators s_{jt}^g is obtained conditional on η^{Tg} and H^{Tg} :

$$\Pr(s_{jt}^g = i | \eta_{jt}^g, h_{jt}^g) \propto q_j f_N(\tilde{\eta}_{jt}^g | 2h_{jt}^g + m_i - 1.2704, r_i^2), \quad i = 1, \dots, 7$$

Step 5 Set $g - 1 = g$ and go to Step 1 and repeat until convergence is achieved

A.1.2 A Four Regime-Switching in DSGE Models

A state-space representation of log-linearized DSGE model is given by:

$$y_t = D + Z\alpha_t \quad (\text{ME})$$

$$\alpha_t = T(\theta)\alpha_{t-1} + R(\theta)\eta_t \quad (\text{TE})$$

The volatility for each structural shock, j , follow Markov process:

$$\begin{aligned} \eta_{jt} &= \sigma_{jst}\varepsilon_{jt} & \varepsilon_{jt} &\sim N(0, 1) \\ \sigma_{jst} &\in \{\sigma_{jH}, \sigma_{jL}\} \end{aligned}$$

While the majority of the existing Markov regime-switching DSGE models assumes regime-switching behavior to be synchronized across shocks, I create flexibility by allowing some shock processes to be independent of the rest. For instance, in RS(4)-DSGE model, I separate monetary policy shock from the remaining shocks. Suppose the regime governing the volatility of the shock process, s_t^i , follows a two-state Markov chain:

$$P^i = \begin{bmatrix} p_{HH}^i & 1 - p_{HH}^i \\ 1 - p_{LL}^i & p_{LL}^i \end{bmatrix} \quad i \in \{Money, Rest\}$$

where $p_{HH}^i = \Pr[s_t^i = High | s_{t-1}^i = High]$. Let P^{Money} and P^{Rest} denote two transition matrices respectively and assume that each regime switching system consists of two regimes. The number of all the possible regimes from two independent structures is four and the resulting transition matrix is

$$P = P^M \otimes P^R$$

where \otimes denotes the Kronecker product and

$$\begin{aligned}
s_t &= HH \quad \text{if } s_t^M = High, s_t^R = High \\
s_t &= HL \quad \text{if } s_t^M = High, s_t^R = Low \\
s_t &= LH \quad \text{if } s_t^M = Low, s_t^R = High \\
s_t &= LL \quad \text{if } s_t^M = Low, s_t^R = Low
\end{aligned}$$

Inference on Volatility Regimes. The following procedure by Kim and Nelson (1999) provides the probability of s_t conditional on the information up to t and appropriate approximation of likelihood as a by-product

Step 1 Run the Kalman filter to get $\alpha_{t|t-1}^{i,j} (\equiv E(a_t | s_t = j, s_{t-1} = i, I_{t-1}))$ and $P_{t|t-1}^{i,j} (\equiv V(a_t | s_t = j, s_{t-1} = i, I_{t-1}))$

Step 2 Update the probability of the current regime after observing data at time t :

$$\begin{aligned}
\Pr[s_t, s_{t-1} | I_t] &= \frac{f(y_t, s_t, s_{t-1} | I_{t-1})}{f(y_t | I_{t-1})} \\
&= \frac{f(y_t | s_t, s_{t-1}, I_{t-1}) \Pr[s_t, s_{t-1} | I_{t-1}]}{\sum_{s_t} \sum_{s_{t-1}} f(y_t | s_t, s_{t-1}, I_{t-1}) \Pr[s_t, s_{t-1} | I_{t-1}]} \\
\Pr[s_t | I_t] &= \sum_{s_{t-1}} \Pr[s_t, s_{t-1} | I_t]
\end{aligned}$$

Step 3 Using these probability terms, collapse 4×4 posteriors into 4×1 :

$$\alpha_{t|t}^j = \frac{\sum_{i=1}^4 \Pr[s_t^i = j, s_{t-1} = i | I_t] \alpha_{t|t-1}^{i,j}}{\Pr[s_t = j | I_t]}$$

$$p_{t|t}^j = \frac{\sum_{i=1}^4 \Pr[s_t^i = j, s_{t-1} = i | I_t] \left(P_{t|t}^{i,j} + (\alpha_{t|t}^j - \alpha_{t|t}^{i,j})(\alpha_{t|t}^j - \alpha_{t|t}^{i,j})' \right)}{\Pr[s_t = j | I_t]}$$

I sample $s^T \equiv [s_1, \dots, s_T]'$ by backward recursion, using $\Pr[s_t | I_t]$.

Gibbs-Sampler for Transition Probability. I use the conjugate priors for transition probability:

parameter	Distribution	90% Interval
p_{HH}^M	<i>Beta</i>	[0.8459 0.9440]
p_{LL}^M	<i>Beta</i>	[0.8084 0.9679]
p_{HH}^R	<i>Beta</i>	[0.8459 0.9440]
p_{LL}^R	<i>Beta</i>	[0.8084 0.9679]

Let parameters for priors a_0 and b_0 . Then, posterior distribution of p_{kk}^i is given by:

$$\mathcal{B}(a_0 + \sum_{t=1}^T I(s_t^i = k | s_{t-1}^i = k), b_0 + \sum_{t=1}^T I(s_t^i = l | s_{t-1}^i = k)).$$

Bibliography

- [1] Adam, Klaus, Albert Marcet, and Juan Pablo Nicolini (2008): “Stock market volatility and learning,” *European Central Bank Working Paper Series*, No 0862.
- [2] Adolfson, Malin (2001): “Monetary Policy with Incomplete Exchange Rate Pass-Through,” No. 127. *Sveriges Riksbank Working Paper Series*.
- [3] Adolfson, M., Laseen, S., Linde, J., and Villani, M. (2007): “Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through,” *Journal of International Economics* 72: 481-511.
- [4] An, Sungbae and Schorfheide, Frank (2007): “Bayesian Analysis of DSGE models,” *Econometric Reviews* 26(2-4): 113–172.
- [5] Benigno, Pierpaolo (2009): “Price Stability with Imperfect Financial Integration,” *Journal of Money, Credit and Banking* 41.s1: 121-149.
- [6] Bergin, Paul (2003): “Putting the ‘New Open Economy Macroeconomics’ to a test,” *Journal of International Economics* 60: 3-34.
- [7] Bergin, Paul (2006): “How Well Can the New Open Economy Macroeconomics Explain the Exchange Rate and Current Account?,” *Journal of International Money and Finance* 25: 675-701.
- [8] Betts, Caroline, and Devereux, Michael B. (2000): “Exchange Rate Dynamics in a Model of Pricing-to-Market,” *Journal of International Economics* 50.1: 215-244.

- [9] Bianchi, Francesco (2009): “Regime Switches, Agents’ Beliefs, and Post-World War II U.S. Macroeconomic Dynamics,” Economics Research Initiatives at Duke (ERID) Working Paper No. 40.
- [10] Blanchard, Olivier and Simon, John (2001): “The Long and Large Decline in U.S. Output Volatility,” Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, 32(2001-1): 135-174.
- [11] Branch, William A., and George W. Evans (2006): “A Simple Recursive Forecasting Model,” *Economics Letters* 91.2: 158-166.
- [12] Bullard, James, and Kaushik Mitra (2002): “Learning About Monetary Policy Rules,” *Journal of Monetary Economics* 49.6: 1105-1129.
- [13] Campa, Jose Manuel and Goldberg, Linda S. (2005): “Exchange Rate Pass-Through into Import Prices,” *Review of Economics and Statistics* 87.4: 679-690.
- [14] Canova, Fabio (1998): “Detrending and Business Cycle Facts,” *Journal of Monetary Economics*, 41(3): 475-512.
- [15] Canova, Fabio, Luca Gambetti, and Evi Pappa (2008): “The Structural Dynamics of U.S. Output and Inflation: What Explains the Changes,” *Journal of Money, Credit, and Banking*, 40: 369-388.
- [16] Cecchetti, Stephen G., Hooper P., Kasman Bruce, Schoenholtz K.L. and Watson, Mark W. (2007): “understanding the evolving inflation process,” At the U.S. Monetary Policy Forum, Washington, D.C.
- [17] Chari, Varadarajan V., Patrick J. Kehoe, and Ellen R. McGrattan (2002): “Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?,” *The Review of Economic Studies* 69.3: 533-563.

- [18] Chauvet, Marcelle and Potter, Simon (2001): “Recent Changes in the US Business Cycle,” Federal Reserve Bank of New York. Manuscript.
- [19] Chib, Siddhartha, Nardari, Federico, and Shephard, Neil (2006): “Analysis of High Dimensional Multivariate Stochastic Volatility Models,” *Journal of Econometrics*, 134(2): 341-371.
- [20] Chib, Siddhartha and Ramamurthy, Srikanth (2010): “Tailored Randomized-block MCMC Methods with Application to DSGE Models,” *Journal of Econometrics*, 155: 19-38.
- [21] Cogley, Timothy and Sargent, Thomas J. (2005): “Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S,” *Review of Economic Dynamics*, 8(2): 262-302.1
- [22] Corsetti, Giancarlo (2007): “New Open Economy Macroeconomics,” *mimeo*.
- [23] Corsetti, Giancarlo, Luca Dedola, and Sylvain Leduc (2008): “High Exchange-Rate Volatility and Low Pass-Through,” *Journal of Monetary Economics* 55.6: 1113-1128.
- [24] Creal, Drew, Koopman, Siem Jan and Zivot, Eric (2010): “Extracting a Robust US Business Cycle Using a Time-varying Multivariate Model-based Bandpass Filter,” *Journal of Applied Econometrics*, 25: 695-719.
- [25] Davig, Troy and Doh, Taeyoung (2009): “Monetary Policy Regime Shifts and Inflation Persistence,” Federal Reserve Bank of Kansas City Working Paper.
- [26] Devereux, Michael B., and Charles Engel (2002): “Exchange Rate Pass-Through, Exchange Rate Volatility, and Exchange Rate Disconnect,” *Journal of Monetary economics* 49.5: 913-940.

- [27] Devereux, Michael B., and Charles Engel (2006): “Expectations and Exchange Rate Policy.” *National Bureau of Economic Research*, No. w12213.
- [28] Dieppe, Alistair, Alberto Gonzalez Pandiella, Stephen Hall, and Alpo Willman (2013): “Limited Information Minimal State Variable Learning in a Medium-Scale Multi-Country Model,” *Economic Modelling* 33: 808-825.
- [29] Durbin, James and Koopman, Siem Jan (2001): *Time series analysis by state space methods*, Oxford University Press, Oxford.
- [30] Engel, Charles, and Jian Wang (2011): “International Trade in Durable Goods: Understanding Volatility, Cyclicalities, and Elasticities,” *Journal of International Economics* 83.1: 37-52.
- [31] Eusepi, Stefano and Preston, Bruce (2011): “Expectations, Learning, and Business Cycle Fluctuations,” *American Economic Review* 101: 2844-2872.
- [32] Evans, George W. and Honkapohja, Seppo (2008): “Expectations, Learning, and Monetary Policy: An Overview of Recent Research,” *Central Bank of Chile Working Papers*.
- [33] Fagan, Gabriel, Jerome Henry, and Ricardo Mestre (2005): “An Area-Wide Model for the Euro Area,” *Economic Modelling* 22.1: 39-59.
- [34] Fernandez-Villaverde, Jesus and Rubio-Ramirez, Juan F. (2007a): “Estimating Macroeconomic Models: A Likelihood Approach,” *Review of Economic Studies*, 74(4): 1059-1087.
- [35] Fernandez-Villaverde, Jesus and Rubio-Ramirez, Juan F. (2007b): “How Structural Are Structural Parameters?,” NBER Working Paper 13166.

- [36] Fernandez-Villaverde, Jesus, Guerron-Quintana, Pablo, and Rubio-Ramirez, Juan F. (2010): “Fortune or Virtue: Time-Variandt Volatilities Versus Parameter Drifting in U.S. Data,” Federal Reserve Bank of Philadelphia Working Paper.
- [37] Gaus, E. and Ramamurthy, S. (2012): “Estimation of Constant Gain Learning Models,” *mimeo*.
- [38] Geweke, John (1999): “Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication,” *Econometric Reviews*, 18(1): 1-73.
- [39] Gourinchas, Pierre-Olivier, and Aaron Tornell (2004): “Exchange Rate Puzzles and Distorted Beliefs,” *Journal of International Economics* 64.2: 303-333.
- [40] Hamilton, James D (1989): “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle,” *Econometrica*, 57(2): 357-84.
- [41] Huang, Kevin XD, Zheng Liu, and Tao Zha (2009): “Learning, Adaptive Expectations and Technology Shocks,” *The Economic Journal* 119.536: 377-405.
- [42] Jacquier, Eric, Polson, Nicholas G. and Rossi, Peter E. (1995): “Models and Priors for Multivariate Stochastic Volatility,” CIRANO Working Papers, 95s-18
- [43] Justiniano, Alejandro and Primiceri, Giorgio E. (2008): “The Time-Varying Volatility of Macroeconomic Fluctuations,” *American Economic Review*, 98(3): 604-41.
- [44] Kim, Chang-Jin and Nelson, Charles R. (1999): *State-Space Models with Regime Switching*, MIT Press, Cambridge, Massachusetts.

- [45] Kim, Chang-Jin and Nelson, Charles R. (1999): “Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle,” *The Review of Economics and Statistics*, 81(4): 608-616.
- [46] Kim, Sangjoon, Shephard, Neil, and Chib, Siddhartha (1998): “Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models,” *Review of Economic Studies*, 65(3): 361-93.
- [47] Kim, Young Se (2009): “Exchange Rates and Fundamentals Under Adaptive Learning,” *Journal of Economic Dynamics and Control* 33: 843-863.
- [48] Lane, Philip R. (2001): “The New Open Macroeconomics: A Survey,” *Journal of International Economics* 54: 235-266.
- [49] Lewis, Vivien and Markiewicz, Agnieszka (2009): “Model Misspecification, Learning and the Exchange Rate Disconnect Puzzle,” *The B.E. Journal of Macroeconomics* 9. no 1.
- [50] Liu, Zheng, Waggoner, Daniel F. and Zha, Tao (2010): “Sources of the Great Moderation: Shocks, Frictions, or Monetary Policy?,” Federal Reserve Bank of Atlanta Working Paper.
- [51] Lubik, Thomas A. and Schorfheide, Frank (2004): “Testing for Indeterminacy: An Application to U.S. Monetary Policy,” *American Economic Review*, 94(1): 190–217.
- [52] Lubik, Thomas and Schorfheide, Frank (2005): “A Bayesian Look at New Open Economy Macroeconomics,” *NBER Macroeconomics Annual* Vol 20.

- [53] Mark, Nelson C. (2009): “Changing Monetary Policy Rules, Learning, and Real Exchange Rate Dynamics,” *Journal of Money, Credit and Banking* 41: 1047-1070.
- [54] McConnell, Margaret M. and Perez-Quiros, Gabriel (2000): “Output Fluctuations in the United States: What Has Changed since the Early 1980’s?,” *American Economic Review*, 90(5); 1464-1476.
- [55] Milani, Fabio (2005): “Adaptive Learning and Inflation Persistence,” *Manuscript*, University of California, Irvine.
- [56] Milani, Fabio (2007): “Expectations, Learning and Macroeconomic Persistence,” *Journal of Monetary Economics* 54: 2065-2082.
- [57] Milani, Fabio (2009): “Adaptive Learning and Macroeconomic Inertia in the Euro Area,” *Journal of Common Market Studies* 47.3: 579-599.
- [58] Monacelli, Tommaso (2005): “Monetary Policy in a Low Pass-Through Environment,” *Journal of Money, Credit and Banking*: 1047-1066.
- [59] Obstfeld, Maurice, and Kenneth Rogoff (2000): “The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?,” *NBER Macroeconomics Annual*, Volume 15. MIT press, 2001. 339-412.
- [60] Opazo, Luis (2006): “The Backus-Smith Puzzle: the Role of Expectations,” *Central Bank of Chile Working Paper* 395.
- [61] Orphanides, Athanasios and John Williams, C. (2005): “The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations,” *Journal of Economic dynamics and control* 29.11: 1927-1950.

- [62] Primiceri, Giorgio E. (2005): “Time Varying Structural Vector Autoregressions and Monetary Policy,” *The Review of Economic Studies*, 72: 821-852.
- [63] Rabanal, Pau, and Vicente Tuesta (2006): “Euro-Dollar Real Exchange Rate Dynamics in an Estimated Two-Country Model: What is Important and What is Not,” *CEPR Discussion Paper* No. 5957.
- [64] Sensier, Marianne and van Dijk, Dick (2004): “Testing for Volatility Changes in U.S. Macroeconomic Time Series,” *The Review of Economics and Statistics*, 86(3): 833-839.
- [65] Sims, Christopher A., Waggoner, Daniel F. and Zha, Tao (2008): “Methods for Inference in Large Multiple-equation Markov-switching Models,” *Journal of Econometrics*, 146(2): 255-274. October
- [66] Sims, Christopher A. and Zha, Tao (2006): “Were There Regime Switches in U.S. Monetary Policy?,” *American Economic Review*, 96(1): 54-81.
- [67] Slobodyan, Sergey and Wouters, Raf (2012): “Learning in a medium-scale DSGE model with expectations based on small forecasting models,” *American Economic Journal: Macroeconomics* 4.2: 65-101.
- [68] Slobodyan, Sergey and Wouters, Raf (2012): “Learning in an Estimated Medium-Scale DSGE Model,” *Journal of Economic Dynamics and Control* 36: 26-46.
- [69] Smets and Wouters (2002): “Openness, Imperfect Exchange Rate Pass-Through and Monetary Policy,” *Journal of Monetary Economics* 49: 947-981.
- [70] Stock, James H (2003): “The Econometric Analysis of Business Cycles,” *Medium Econometrisch Toepassingen* (special issue honoring J. Tinbergen), forthcoming

- [71] Stock, James H. and Watson, Mark W. (2002): “Has the Business Cycle Changed and Why?,” NBER Macroeconomics Annual 2002, 17: 159-230.
- [72] Vilagi, Balazs (2007): “Adaptive Learning and Macroeconomic Persistence: Comparing DSGE models of the Euro Area,” *Magyar Nemzeti Bank mimeo*.