

Liu, Zheng; Waggoner, Daniel F.; Zha, Tao

**Working Paper**

## Sources of macroeconomic fluctuations: A regime-switching DSGE approach

Working Paper, No. 2009-3a

**Provided in Cooperation with:**

Federal Reserve Bank of Atlanta

*Suggested Citation:* Liu, Zheng; Waggoner, Daniel F.; Zha, Tao (2010) : Sources of macroeconomic fluctuations: A regime-switching DSGE approach, Working Paper, No. 2009-3a, Federal Reserve Bank of Atlanta, Atlanta, GA

This Version is available at:

<https://hdl.handle.net/10419/70635>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

Sources of Macroeconomic Fluctuations:  
A Regime-Switching DSGE Approach

Zheng Liu, Daniel F. Waggoner, and Tao Zha

Working Paper 2009-3a  
October 2010

## Sources of Macroeconomic Fluctuations: A Regime-Switching DSGE Approach

Zheng Liu, Daniel F. Waggoner, and Tao Zha

Working Paper 2009-3a  
October 2010

**Abstract:** We examine the sources of macroeconomic economic fluctuations by estimating a variety of medium-scale dynamic stochastic general equilibrium (DSGE) models within a unified framework that incorporates regime switching both in shock variances and in the inflation target. Our general framework includes a number of different model features studied in the literature. We propose an efficient methodology for estimating regime-switching DSGE models. The model that best fits the U.S. time-series data is the one with synchronized shifts in shock variances across two regimes and the fit does not rely on strong nominal rigidities. We find little evidence of changes in the inflation target. We identify three types of shocks that account for most of macroeconomic fluctuations: shocks to total factor productivity, wage markup, and the capital depreciation rate.

JEL classification: C11, C51, E32, E42, E52

Key words: systematic analysis, regime switching, depreciation shock, efficient estimation methods

---

The authors thank Craig Burnside, Larry Christiano, Tim Cogley, Chris Erceg, Marco Del Negro, Wouter Den Haan, Martin Ellison, Jesus Fernandez-Villaverde, Jordi Gali, Marc Giannoni, Michael Golosov, Pat Higgins, Alejandro Justiniano, Soyoung Kim, Junior Maih, Christian Matthes, Ulrich Müller, Andy Levin, Lee Ohanian, Pietro F. Peretto, Giorgio Primiceri, Frank Schorfheide, Chris Sims, Harald Uhlig, and seminar participants at the Bank of Korea, the National Bureau of Economic Research (NBER) summer institute, the University of California, Berkeley, the Society for Economic Dynamics (SED), and Duke University for helpful discussions and comments. Eric Wang provided valuable assistance in grid computing. The views expressed here are the authors' and not necessarily those of the Federal Reserve Banks of Atlanta or San Francisco or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Zheng Liu, Research Department, Federal Reserve Bank of San Francisco, 101 Market Street, San Francisco, CA 94105, 415-974-3328, zheng.liu@sf.frb.org; Daniel F. Waggoner, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4740, 404-498-8278, daniel.f.waggoner@atl.frb.org; or Tao Zha, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4740, 404-498-8353, tzha@earthlink.net.

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's Web site at [frbatlanta.org/pubs/WP/](http://frbatlanta.org/pubs/WP/). Use the WebScriber Service at [frbatlanta.org](http://frbatlanta.org) to receive e-mail notifications about new papers.

## I. INTRODUCTION

We examine the sources of macroeconomic fluctuations by estimating a number of regime-switching models using modern Bayesian techniques in a unified dynamic stochastic general equilibrium (DSGE) framework. The standard approach to analyzing business-cycle fluctuations is the use of constant-parameter medium-scale DSGE models (Altig, Christiano, Eichenbaum, and Linde, 2004; Christiano, Eichenbaum, and Evans, 2005; Levin, Onatski, Williams, and Williams, 2006; Smets and Wouters, 2007; Del Negro, Schorfheide, Smets, and Wouters, 2007). In this paper we generalize the standard approach by allowing time variations in shock variances and in the central bank's inflation target according to Markov-switching processes. These time variations appear to be present in the U.S. macroeconomic time series. An important question is how significant the time variations are when we fit the data to relatively large DSGE models with rich dynamic structures and shock processes that are economically interpretable. If the answer is positive, the next question is in what dimension the time variations matter. To answer these questions, we estimate a number of alternative models nested in this general framework using the Bayesian method and we compare the fit of these models to the time series data in the postwar U.S. economy. The best-fit model is then used to identify shocks that are important in driving macroeconomic fluctuations.

Our approach yields several new results. We find strong empirical evidence in favor of the DSGE model with two regimes in shock variances, where regime shifts in the variances are synchronized. The models with constant parameters (i.e., no regime shifts), with independent regime shifts in shock variances, or with more than two regimes do not fit to the data as well. In our preferred model (i.e., the best-fit model) with two synchronized shock regimes, the high-volatility regime was frequently observed in the period from the early 1970s through the mid-1980s, while the low-volatility regime prevailed in most of the period from the mid-1980s through 2007. This finding is broadly consistent with the well-known fact that the U.S. economy experienced a general reduction in macroeconomic volatilities during the latter sample period (Stock and Watson, 2003).

The fit of our preferred regime-switching DSGE model does not rely on strong nominal rigidities. In particular, our estimates imply that the durations of the price and nominal wage contracts last no more than 2 quarters of a year—much shorter than those reported in the constant-parameter DSGE models in previous studies (Smets and Wouters, 2007). This finding highlights the sensitivity of the estimates of some key

structural parameters obtained in models with no regime switching to specifications of the shock processes. When we allow the shock variances to switch regimes, the model relies less on nominal rigidities to fit to the data.

Neither does the fit of our preferred model reply on regime shifts in the inflation target. Allowing the inflation target to shift between two regimes—either synchronized with or independent of the shock regime switching—does not improve the model’s marginal data density. This finding is robust to a variety of model specifications and it is consistent with the conclusion from other works about changes in monetary policy in general (Stock and Watson, 2003; Canova and Gambetti, 2004; Cogley and Sargent, 2005; Primiceri, 2005; Sims and Zha, 2006; Justiniano and Primiceri, 2008). We focus on studying changes in the inflation target instead of changes in monetary policy’s response to inflation for both conceptual and computational reasons. When agents take into account changes in monetary policy’s response to inflation in forming their expectations, a solution method to the model is nonstandard (Liu, Waggoner, and Zha, 2009). Indeed, it would be computationally infeasible for us to estimate a large set of DSGE models like what we do in the current paper since the solution would require an iterative algorithm that can be time-consuming in Monte Carlo simulations. Furthermore, indeterminacy is more prevalent in this kind of regime-switching model than in the standard DSGE model (Farmer, Waggoner, and Zha, 2009). For these reasons, we follow Schorfheide (2005) and Ireland (2005) and focus on examining changes in the inflation target to give the model the best chance to detect changes in monetary policy. Although we can apply the standard method to solving our regime-switching DSGE models (as shown in Section V), we have nonetheless pushed the limits of our computational and analytical capacity because of a large set of regime-switching models we have estimated.

In the best-fit model, we identify three types of shocks that are important for macroeconomic fluctuations. These are a shock to the growth rate of the total factor productivity (TFP), a shock to wage markups, and a shock to the capital depreciation rate. Taken together, these three shocks account for about 70–80% of the variances of aggregate output, investment, and inflation at business cycle frequencies. Other shocks such as monetary policy shocks, investment-specific technology shocks, and price markup shocks are not as important. The TFP shocks and the wage markup shocks should be familiar to a student of the DSGE literature, but the capital depreciation shock is new. We provide some economic interpretations of the depreciation shock in Section VII.3.

In what follows, we briefly discuss our contributions in relation to the literature in Section II. We then present, in Section III, the general regime-switching DSGE framework. In Section IV, we present the system of equilibrium conditions and discuss our solution methods. In Section V, we describe the data and our empirical approach. As a methodological contribution, we propose an efficient methodology for estimating regime-switching DSGE models; we summarize and discuss several modern methods for obtaining accurate estimates of marginal data densities for relatively large DSGE models. In Section VI, we compare the fit of a number of models nested by our general DSGE framework, identify the best-fit model, and report posterior estimates of the parameters in this model. In Section VII, we discuss the economic implications of our estimates in the best-fit model and identify the key sources of shocks that drive macroeconomic fluctuations. We conclude in Section VIII.

## II. RELATED LITERATURE

The debate in the literature on the sources of macroeconomic fluctuations gives emphasis to whether shifts in monetary policy are the main sources of macroeconomic volatilities (Clarida, Galí, and Gertler, 2000; Lubik and Schorfheide, 2004; Stock and Watson, 2003; Sims and Zha, 2006; Bianchi, 2008; Gambetti, Pappa, and Canova, 2008) or whether shocks in investment-specific technology are more important than other shocks in driving macroeconomic fluctuations (Fisher, 2006; Smets and Wouters, 2007; Justiniano and Primiceri, 2008). Much of the disagreement stems from the use of different frameworks and different empirical methods. Part of the literature focuses on reduced-form econometric models, part of it on small-scale DSGE models, and part of it on medium-scale DSGE models. Some models assume homogeneity in shock variances; others assume that shock variances are time-varying. Some models are estimated with different subsamples to reflect shifts in policy or in shock variances; other models are estimated with the entire sample. Given these differences in the model framework and in the empirical approach, it is difficult to draw a firm conclusion about the sources of macroeconomic fluctuations. The goal of the current paper is to provide a systematic examination of the sources of macroeconomic fluctuations in one unified DSGE framework that allows for regime shifts in shock variances and in monetary policy.

Our approach differs from that employed in the literature in several aspects. First, we aim at fully characterizing the uncertainty across different models by examining different versions of the DSGE model for robust analysis to substantiate our conclusion.

Although estimating a large set of models has not been performed in the literature, we think it is necessary to examine the robustness of a conclusion like ours about potential sources of macroeconomic fluctuations.

Second, our approach does not require splitting the sample to examine changes in monetary policy, although it nests sampling-splitting as a special case. Unlike Sims and Zha (2006) where the number of VAR parameters is relatively large and the inflation target is implicit, our way of modeling policy changes takes the inflation target explicitly and gives a tightly parameterized model that has the best chance to detect the importance of policy changes, if it exists, in generating business-cycle fluctuations.

Third and methodologically, for fairly large DSGE models, especially for regime-switching DSGE models, the posterior distribution tends to be very non-Gaussian, making it very challenging to search for the global peak. We improve on earlier works such as Cogley and Sargent (2005) and Justiniano and Primiceri (2006) by obtaining the estimate of parameters at the posterior mode for each model. We show that economic implications can be seriously distorted if the estimates are based on a lower posterior peak.

Fourth, there is a strand of literature that emphasizes changes in the inflation target as a representation of important shifts in the conduct of U.S. monetary policy (for example, Favero and Rovelli (2003); Erceg and Levin (2003); Schorfheide (2005); Ireland (2005)). Unlike the earlier works, we study a variety of fairly large DSGE models to avoid potential mis-specifications.

Finally, we use three new methods for computing marginal data densities in model comparison. Since these methods are based on different statistical foundations, it is essential that all these methods give a numerically similar result to ensure that the estimate of a marginal data density is unbiased and accurate (Sims, Waggoner, and Zha, 2008).

### III. THE MODEL

The model economy is populated by a continuum of households, each endowed with a unit of differentiated labor skill indexed by  $i \in [0, 1]$ ; and a continuum of firms, each producing a differentiated good indexed by  $j \in [0, 1]$ . The monetary authority follows a feedback interest rate rule, under which the nominal interest rate is set to respond to its own lag and deviations of inflation and output from their targets. The policy regime  $s_t$  represented by the time-varying inflation target switches between a finite number of regimes contained in the set  $\mathcal{S}$ , with the Markov transition probabilities summarized

by the matrix  $Q = [q_{ij}]$ , where  $q_{ij} = \text{Prob}(s_{t+1} = i | s_t = j)$  for  $i, j \in \mathcal{S}$ . The economy is buffeted by several sources of shocks. The variance of each shock switches between a finite number of regimes denoted by  $s_t^* \in \mathcal{S}^*$  with the transition matrix  $Q^* = [q_{ij}^*]$ .

**III.1. The aggregation sector.** The aggregation sector produces a composite labor skill denoted by  $L_t$  to be used in the production of each type of intermediate goods and a composite final good denoted by  $Y_t$  to be consumed by each household. The production of the composite skill requires a continuum of differentiated labor skills  $\{L_t(i)\}_{i \in [0,1]}$  as inputs, and the production of the composite final good requires a continuum of differentiated intermediate goods  $\{Y_t(j)\}_{j \in [0,1]}$  as inputs. The aggregation technologies are given by

$$L_t = \left[ \int_0^1 L_t(i)^{\frac{1}{\mu_{wt}}} di \right]^{\mu_{wt}}, \quad Y_t = \left[ \int_0^1 Y_t(j)^{\frac{1}{\mu_{pt}}} dj \right]^{\mu_{pt}}, \quad (1)$$

where  $\mu_{wt}$  and  $\mu_{pt}$  determine the elasticity of substitution between the skills and between the goods, respectively. Following Smets and Wouters (2007), we assume that

$$\ln \mu_{wt} = (1 - \rho_w) \ln \mu_w + \rho_w \ln \mu_{w,t-1} + \sigma_{wt} \varepsilon_{wt} - \phi_w \sigma_{w,t-1} \varepsilon_{w,t-1} \quad (2)$$

and that

$$\ln \mu_{pt} = (1 - \rho_p) \ln \mu_p + \rho_p \ln \mu_{p,t-1} + \sigma_{pt} \varepsilon_{pt} - \phi_p \sigma_{p,t-1} \varepsilon_{p,t-1}, \quad (3)$$

where, for  $j \in \{w, p\}$ ,  $\rho_j \in (-1, 1)$  is the AR(1) coefficient,  $\phi_j$  is the MA(1) coefficient,  $\sigma_{jt} \equiv \sigma_j(s_t^*)$  is the regime-switching standard deviation, and  $\varepsilon_{jt}$  is an i.i.d. white noise process with a zero mean and a unit variance. We interpret  $\mu_{wt}$  and  $\mu_{pt}$  as the wage markup and price markup shocks.

The representative firm in the aggregation sector faces perfectly competitive markets for the composite skill and the composite good. The demand functions for labor skill  $i$  and for good  $j$  resulting from the optimizing behavior in the aggregation sector are given by

$$L_t^d(i) = \left[ \frac{W_t(i)}{\bar{W}_t} \right]^{-\frac{\mu_{wt}}{\mu_{wt}-1}} L_t, \quad Y_t^d(j) = \left[ \frac{P_t(j)}{\bar{P}_t} \right]^{-\frac{\mu_{pt}}{\mu_{pt}-1}} Y_t, \quad (4)$$

where the wage rate  $\bar{W}_t$  of the composite skill is related to the wage rates  $\{W_t(i)\}_{i \in [0,1]}$  of the differentiated skills by  $\bar{W}_t = \left[ \int_0^1 W_t(i)^{1/(1-\mu_{wt})} di \right]^{1-\mu_{wt}}$  and the price  $\bar{P}_t$  of the composite good is related to the prices  $\{P_t(j)\}_{j \in [0,1]}$  of the differentiated goods by  $\bar{P}_t = \left[ \int_0^1 P_t(j)^{1/(1-\mu_{pt})} dj \right]^{1-\mu_{pt}}$ .



III.2. **The intermediate good sector.** The production of a type  $j$  good requires labor and capital inputs. The production function is given by

$$Y_t(j) = K_t^f(j)^{\alpha_1} [Z_t L_t^f(j)]^{\alpha_2}, \quad (5)$$

where  $K_t^f(j)$  and  $L_t^f(j)$  are the inputs of capital and the composite skill and the variable  $Z_t$  denotes a neutral technology shock, which follows the stochastic process

$$Z_t = \lambda_z^t z_t, \quad \ln z_t = (1 - \rho_z) \ln z + \rho_z \ln z_{t-1} + \sigma_{zt} \varepsilon_{zt}, \quad (6)$$

where  $\rho_z \in (-1, 1)$  measures the persistence,  $\sigma_{zt} \equiv \sigma_z(s_t^*)$  denotes the regime-switching standard deviation, and  $\varepsilon_{zt}$  is an i.i.d. white noise process with a zero mean and a unit variance. The parameters  $\alpha_1$  and  $\alpha_2$  measure the cost shares the capital and labor inputs. Following Chari, Kehoe, and McGrattan (2000), we introduce some real rigidity by assuming the existence of some firm-specific factors (such as land), so that  $\alpha_1 + \alpha_2 \leq 1$ .

Each firm in the intermediate-good sector is a price-taker in the input market and a monopolistic competitor in the product market where it sets a price for its product, taking the demand schedule in (4) as given. We follow Calvo (1983) and assume that pricing decisions are staggered across firms. The probability that a firm cannot adjust its price is given by  $\xi_p$ . Following Woodford (2003), Christiano, Eichenbaum, and Evans (2005), and Smets and Wouters (2007), we allow a fraction of firms that cannot re-optimize their pricing decisions to index their prices to the overall price inflation realized in the past period. Specifically, if the firm  $j$  cannot set a new price, its price is automatically updated according to

$$P_t(j) = \pi_{t-1}^{\gamma_p} \pi^{1-\gamma_p} P_{t-1}(j), \quad (7)$$

where  $\pi_t = \bar{P}_t / \bar{P}_{t-1}$  is the inflation rate between  $t-1$  and  $t$ ,  $\pi$  is the steady-state inflation rate, and  $\gamma_p$  measures the degree of indexation.

A firm that can renew its price contract chooses  $P_t(j)$  to maximize its expected discounted dividend flows given by

$$E_t \sum_{i=0}^{\infty} \xi_p^i D_{t,t+i} [P_t(j) \chi_{t,t+i}^p Y_{t+i}^d(j) - V_{t+i}(j)], \quad (8)$$

where  $D_{t,t+i}$  is the period- $t$  present value of a dollar in a future state in period  $t+i$ ,  $V_{t+i}(j)$  is the cost function, and the term  $\chi_{t,t+i}^p$  comes from the price-updating rule (7) and is given by

$$\chi_{t,t+i}^p = \begin{cases} \prod_{k=1}^i \pi_{t+k-1}^{\gamma_p} \pi^{1-\gamma_p} & \text{if } i \geq 1 \\ 1 & \text{if } i = 0. \end{cases} \quad (9)$$

In maximizing its profit, the firm takes as given the demand schedule  $Y_{t+i}^d(j) = \left(\frac{P_t(j)\chi_{t,t+i}^p}{P_{t+i}}\right)^{-\frac{\mu_{p,t+i}}{\mu_{p,t+i}-1}} Y_{t+i}$ . The first order condition for the profit-maximizing problem yields the optimal pricing rule

$$\mathbb{E}_t \sum_{i=0}^{\infty} \xi_p^i D_{t,t+i} Y_{t+i}^d(j) \frac{1}{\mu_{p,t+i} - 1} [\mu_{p,t+i} \Phi_{t+i}(j) - P_t(j) \chi_{t,t+i}^p] = 0, \quad (10)$$

where  $\Phi_{t+i}(j) = \partial V_{t+i}(j) / \partial Y_{t+i}^d(j)$  denotes the marginal cost function. In the absence of markup shocks,  $\mu_{pt}$  would be a constant and (10) implies that the optimal price is a markup over an average of the marginal costs for the periods in which the price will remain effective. Clearly, if  $\xi_p = 0$  for all  $t$ , that is, if prices are perfectly flexible, then the optimal price would be a markup over the contemporaneous marginal cost.

Cost-minimizing implies that the marginal cost function is given by

$$\Phi_t(j) = \left[ \tilde{\alpha} (\bar{P}_t r_{kt})^{\alpha_1} \left( \frac{\bar{W}_t}{Z_t} \right)^{\alpha_2} \right]^{\frac{1}{\alpha_1 + \alpha_2}} Y_t(j)^{\frac{1}{\alpha_1 + \alpha_2} - 1}, \quad (11)$$

where  $\tilde{\alpha} \equiv \alpha_1^{-\alpha_1} \alpha_2^{-\alpha_2}$  and  $r_{kt}$  denotes the real rental rate of capital input. The conditional factor demand functions imply that

$$\frac{\bar{W}_t}{\bar{P}_t r_{kt}} = \frac{\alpha_2}{\alpha_1} \frac{K_t^f(j)}{L_t^f(j)}, \quad \forall j \in [0, 1]. \quad (12)$$

**III.3. Households.** There is a continuum of households, each endowed with a differentiated labor skill indexed by  $h \in [0, 1]$ . Household  $h$  derives utility from consumption and leisure. We assume that there exists financial instruments that provide perfect insurance for the households in different wage-setting cohorts, so that the households make identical consumption and investment decisions despite that their wage incomes may differ due to staggered wage setting.<sup>1</sup> In what follows, we impose this assumption and omit the household index for consumption and investment.

The utility function for household  $h \in [0, 1]$  is given by

$$\mathbb{E} \sum_{t=0}^{\infty} \beta^t A_t \left\{ \ln(C_t - bC_{t-1}) - \frac{\Psi}{1 + \eta} L_t(h)^{1+\eta} \right\}, \quad (13)$$

---

<sup>1</sup>To obtain complete risk-sharing among households in different wage-setting cohorts does not rely on the existence of such (implicit) financial arrangements. As shown by Huang, Liu, and Phaneuf (2004), the same equilibrium dynamics can be obtained in a model with a representative household (and thus complete insurance) consisting of a large number of worker members. The workers supply their homogenous labor skill to a large number of employment agencies, who transform the homogenous skill into differentiated skills and set nominal wages in a staggered fashion.

where  $\beta \in (0, 1)$  is a subjective discount factor,  $C_t$  denotes consumption,  $L_t(h)$  denotes hours worked,  $\eta > 0$  is the inverse Frish elasticity of labor hours, and  $b$  measures the importance of habit formation. The variable  $A_t$  denotes a preference shock, which follows the stationary process

$$\ln A_t = (1 - \rho_a) \ln A + \rho_a \ln A_{t-1} + \sigma_{at} \varepsilon_{at}, \quad (14)$$

where  $\rho_a \in (-1, 1)$  is the persistence parameter,  $\sigma_{at} \equiv \sigma_a(s_t^*)$  is the regime-switching standard deviation, and  $\varepsilon_{at}$  is an i.i.d. white noise process with a zero mean and a unit variance.

In each period  $t$ , the household faces the budget constraint

$$\begin{aligned} \bar{P}_t C_t + \frac{\bar{P}_t}{Q_t} [I_t + a(u_t) K_{t-1}] + E_t D_{t,t+1} B_{t+1} \leq \\ W_t(h) L_t^d(h) + \bar{P}_t r_{kt} u_t K_{t-1} + \Pi_t + B_t + T_t. \end{aligned} \quad (15)$$

In the budget constraint,  $I_t$  denotes investment,  $B_{t+1}$  is a nominal state-contingent bond that represents a claim to one dollar in a particular event in period  $t + 1$ , and this claim costs  $D_{t,t+1}$  dollars in period  $t$ ;  $W_t(h)$  is the nominal wage for  $h$ 's labor skill,  $K_{t-1}$  is the beginning-of-period capital stock,  $u_t$  is the utilization rate of capital,  $\Pi_t$  is the profit share, and  $T_t$  is a lump-sum transfer from the government. The function  $a(u_t)$  captures the cost of variable capital utilization. Following Altig, Christiano, Eichenbaum, and Linde (2004) and Christiano, Eichenbaum, and Evans (2005), we assume that  $a(u)$  is increasing and convex. The term  $Q_t$  denotes the investment-specific technological change. Following Greenwood, Hercowitz, and Krusell (1997), we assume that  $Q_t$  contains a deterministic trend and a stochastic component. In particular,

$$Q_t = \lambda_q^t q_t, \quad (16)$$

where  $\lambda_q$  is the growth rate of the investment-specific technological change and  $q_t$  is an investment-specific technology shock, which follows a stationary process given by

$$\ln q_t = (1 - \rho_q) \ln q + \rho_q \ln q_{t-1} + \sigma_{qt} \varepsilon_{qt}, \quad (17)$$

where  $\rho_q \in (-1, 1)$  is the persistence parameter,  $\sigma_{qt} \equiv \sigma_q(s_t^*)$  is the regime-switching standard deviation, and  $\varepsilon_{qt}$  is an i.i.d. white noise process with a zero mean and a unit variance. The importance of investment-specific technological change is also documented in Fisher (2006) and Fernandez-Villaverde and Rubio-Ramirez (2007).

The capital stock evolves according to the law of motion

$$K_t = (1 - \delta_t) K_{t-1} + [1 - S(I_t/I_{t-1})] I_t, \quad (18)$$

where the function  $S(\cdot)$  represents the adjustment cost in capital accumulation. We assume that  $S(\cdot)$  is convex and satisfies  $S(\lambda_q \lambda_z^*) = S'(\lambda_q \lambda_z^*) = 0$ , where  $\lambda^* = (\lambda_q^{\alpha_1} \lambda_z^{\alpha_2})^{\frac{1}{1-\alpha_1}}$  is the steady-state growth rate of output and consumption. The term  $\delta_t$  denotes the depreciation rate of the capital stock and follows the stationary stochastic process

$$\ln \delta_t = (1 - \rho_d) \ln \delta + \rho_d \ln \delta_{t-1} + \sigma_{dt} \varepsilon_{dt}, \quad (19)$$

where  $\rho_e \in (-1, 1)$  is the persistence parameter,  $\sigma_{dt} \equiv \sigma_d(s_t^*)$  is the regime-switching standard deviation, and  $\varepsilon_{dt}$  is the white noise innovation with a zero mean and a unit variance. We introduce this time variation in the depreciation rate to capture the difference between economic depreciation (reflecting in part an unobserved quality improvement in equipment) and physical depreciation.

The household takes prices and all wages but its own as given and chooses  $C_t$ ,  $I_t$ ,  $K_t$ ,  $u_t$ ,  $B_{t+1}$ , and  $W_t(h)$  to maximize (13) subject to (15) - (18), the borrowing constraint  $B_{t+1} \geq -\underline{B}$  for some large positive number  $\underline{B}$ , and the labor demand schedule  $L_t^d(h)$  described in (4).

The wage-setting decisions are staggered across households. In each period, a fraction  $\xi_w$  of households cannot re-optimize their wage decisions and, among those who cannot re-optimize, a fraction  $\gamma_w$  of them index their nominal wages to the price inflation realized in the past period. In particular, if the household  $h$  cannot set a new nominal wage, its wage is automatically updated according to

$$W_t(h) = \pi_{t-1}^{\gamma_w} \pi^{1-\gamma_w} \lambda_{t-1,t}^* W_{t-1}(h), \quad (20)$$

where  $\lambda_{t-1,t}^* \equiv \frac{\lambda_t^*}{\lambda_{t-1}^*}$ , with  $\lambda_t^* \equiv (Q_t^{\alpha_1} Z_t^{\alpha_2})^{\frac{1}{1-\alpha_1}}$  denoting the trend growth rate of aggregate output (and the real wage). If a household  $h \in [0, 1]$  can re-optimize its nominal wage-setting decision, it chooses  $W(h)$  to maximize the utility subject to the budget constraint (15) and the labor demand schedule in (4). The optimal wage-setting decision implies that

$$\mathbb{E}_t \sum_{i=0}^{\infty} \xi_w^i D_{t,t+i} L_{t+i}^d(h) \frac{1}{\mu_{w,t+i} - 1} [\mu_{w,t+i} MRS_{t+i}(h) - W_t(h) \chi_{t,t+i}^w] = 0, \quad (21)$$

where  $MRS_t(h)$  denotes the marginal rate of substitution between leisure and income for household  $h$  and  $\chi_{t,t+i}^w$  is defined as

$$\chi_{t,t+i}^w \equiv \begin{cases} \prod_{k=1}^i \pi_{t+k-1}^{\gamma_w} \pi^{1-\gamma_w} \lambda_{t,t+i}^* & \text{if } i \geq 1 \\ 1 & \text{if } i = 0. \end{cases}, \quad (22)$$

where  $\lambda_{t,t+i}^* \equiv \frac{\lambda_{t+i}^*}{\lambda_t^*}$ . In the absence of wage-markup shocks,  $\mu_{wt}$  would be a constant and (21) implies that the optimal wage is a constant markup over a weighted average

of the marginal rate of substitution for the periods in which the nominal wage remains effective. If  $\xi_w = 0$ , then the nominal wage adjustments are flexible and (21) implies that the nominal wage is a markup over the contemporaneous marginal rate of substitution. We derive the rest of the household's optimizing conditions in a technical appendix available upon request.

**III.4. The government and monetary policy.** The government follows a Ricardian fiscal policy, with its spending financed by lump-sum taxes so that  $\bar{P}_t G_t = T_t$ , where  $G_t$  denotes the government spending in final consumption units. Denote by  $\tilde{G}_t \equiv \frac{G_t}{\lambda_t^*}$  the detrended government spending, where

$$\lambda_t^* \equiv (Z_t^{\alpha_2} Q_t^{\alpha_1})^{\frac{1}{1-\alpha_1}}. \quad (23)$$

We assume that  $\tilde{G}_t$  follows the stationary stochastic process

$$\ln \tilde{G}_t = (1 - \rho_g) \ln \tilde{G} + \rho_g \ln \tilde{G}_{t-1} + \sigma_{gt} \varepsilon_{gt} + \rho_{gz} \sigma_{zt} \varepsilon_{zt}, \quad (24)$$

where we follow Smets and Wouters (2007) and assume that the government spending shock responds to productivity shocks.

Monetary policy is described by a feedback interest rate rule that allows the possibility of regime switching in the inflation target. The interest rate rule is given by

$$R_t = \kappa R_{t-1}^{\rho_r} \left[ \left( \frac{\pi_t}{\pi^*(s_t)} \right)^{\phi_\pi} \left( \frac{Y_t}{\lambda_t^*} \right)^{\phi_y} \right]^{1-\rho_r} e^{\sigma_{rt} \varepsilon_{rt}}, \quad (25)$$

where  $R_t = [E_t D_{t,t+1}]^{-1}$  denotes the nominal interest rate and  $\pi^*(s_t)$  denotes the regime-dependent inflation target. The constant terms  $\kappa$ ,  $\rho_r$ ,  $\phi_\pi$ , and  $\phi_y$  are policy parameters. The term  $\varepsilon_{rt}$  denotes the monetary policy shock, which follows an i.i.d. normal process with a zero mean and a unit variance. The term  $\sigma_{rt} \equiv \sigma_r(s_t^*)$  is the regime-switching standard deviation of the monetary policy shock. We assume that the 8 shocks  $\varepsilon_{wt}$ ,  $\varepsilon_{pt}$ ,  $\varepsilon_{zt}$ ,  $\varepsilon_{qt}$ ,  $\varepsilon_{dt}$ ,  $\varepsilon_{at}$ ,  $\varepsilon_{rt}$ , and  $\varepsilon_{gt}$  are mutually independent.

**III.5. Market clearing and equilibrium.** In equilibrium, markets for bond, composite labor, capital stock, and composite goods all clear. Bond market clearing implies that  $B_t = 0$  for all  $t$ . Labor market clearing implies that  $\int_0^1 L_t^f(j) dj = L_t$ . Capital market clearing implies that  $\int_0^1 K_t^f(j) dj = u_t K_{t-1}$ . Composite goods market clearing implies that

$$C_t + \frac{1}{Q_t} [I_t + a(u_t) K_{t-1}] + G_t = Y_t, \quad (26)$$

where aggregate output is related to aggregate primary factors through the aggregate production function

$$G_{pt}Y_t = (u_t K_{t-1})^{\alpha_1} (Z_t L_t)^{\alpha_2}, \quad (27)$$

with  $G_{pt} \equiv \int_0^1 \left( \frac{P_t(j)}{\bar{P}_t} \right)^{-\frac{\mu_{pt}}{\mu_{pt}-1} \frac{1}{\alpha_1+\alpha_2}} dj$  measuring the price dispersion.

Given fiscal and monetary policy, an *equilibrium* in this economy consists of prices and allocations such that (i) taking prices and all nominal wages but its own as given, each household's allocation and nominal wage solve its utility maximization problem; (ii) taking wages and all prices but its own as given, each firm's allocation and price solve its profit maximization problem; (iii) markets clear for bond, composite labor, capital stock, and final goods.

#### IV. EQUILIBRIUM DYNAMICS

**IV.1. Stationary equilibrium and the deterministic steady state.** We focus on a stationary equilibrium with balanced growth. On a balanced growth path, output, consumption, investment, capital stock, and the real wage all grow at constant rates, while hours remain constant. Further, in the presence of investment-specific technological change, investment and capital grow at a faster rate. To induce stationarity, we transform variables so that

$$\tilde{Y}_t = \frac{Y_t}{\lambda_t^*}, \quad \tilde{C}_t = \frac{C_t}{\lambda_t^*}, \quad \tilde{w}_t = \frac{W_t}{\bar{P}_t \lambda_t^*}, \quad \tilde{I}_t = \frac{I_t}{Q_t \lambda_t^*}, \quad \tilde{K}_t = \frac{K_t}{Q_t \lambda_t^*},$$

where  $\lambda_t^*$  is the underlying trend for output, consumption, and the real wage given by (23).

Along the balanced growth path, as noted by Greenwood, Hercowitz, and Krusell (1997), the real rental price of capital keeps falling since the capital-output ratio keeps rising. The rate at which the rental price is falling is given by  $\lambda_q$ . Thus, the transformed variable  $\tilde{r}_{kt} = r_{kt} Q_t$ , that is, the rental price in consumption unit, is stationary. Further, the marginal utility of consumption is declining, so we define  $\tilde{U}_{ct} = U_{ct} \lambda_t^*$  to induce stationarity.

The steady state in the model is the stationary equilibrium in which all shocks are shut off, including the “regime shocks” to the inflation target. To derive the steady state, we represent the finite Markov switching process with a vector AR(1) process (Hamilton, 1994). Specifically, the inflation target can be written as

$$\pi^*(s_t) = [\pi^*(1), \pi^*(2)] e_{s_t}, \quad (28)$$

where  $\pi^*(j)$  is the inflation target in regime  $j \in \{1, 2\}$  and

$$e_{s_t} = \begin{bmatrix} \mathbf{1}\{s_t = 1\} \\ \mathbf{1}\{s_t = 2\} \end{bmatrix}, \quad (29)$$

with  $\mathbf{1}\{s_t = j\} = 1$  if  $s_t = j$  and 0 otherwise. As shown in Hamilton (1994), the random vector  $e_{s_t}$  follows an AR(1) process:

$$e_{s_t} = Qe_{s_{t-1}} + v_t, \quad (30)$$

where  $Q$  is the transition matrix of the Markov switching process and the innovation vector has the property that  $E_{t-1}v_t = 0$ . In the steady state,  $v_t = 0$  so that (30) defines the ergodic probabilities for the Markov process and, from (28), the steady-state inflation  $\pi$  is the ergodic mean of the inflation target. Given  $\pi$ , the derivations for the rest of the steady-state equilibrium conditions are straightforward.

**IV.2. Linearized equilibrium dynamics.** To solve for the equilibrium dynamics, we log-linearize the equilibrium conditions around the deterministic steady state. We use a hatted variable  $\hat{x}_t$  to denote the log-deviations of the stationary variable  $X_t$  from its steady-state value (i.e.,  $\hat{x}_t = \ln(X_t/X)$ ).

Linearizing the optimal pricing decision rule implies that<sup>2</sup>

$$\hat{\pi}_t - \gamma_p \hat{\pi}_{t-1} = \frac{\kappa_p}{1 + \bar{\alpha}\theta_p} (\hat{\mu}_{pt} + \hat{m}c_t) + \beta E_t[\hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t], \quad (31)$$

where  $\theta_p \equiv \frac{\mu_p}{\mu_p - 1}$ ,  $\kappa_p \equiv \frac{(1 - \beta\xi_p)(1 - \xi_p)}{\xi_p}$ ,  $\bar{\alpha} \equiv \frac{1 - \alpha_1 - \alpha_2}{\alpha_1 + \alpha_2}$ , and

$$\hat{m}c_t = \frac{1}{\alpha_1 + \alpha_2} [\alpha_1 \hat{r}_{kt} + \alpha_2 \hat{w}_t] + \bar{\alpha} \hat{y}_t. \quad (32)$$

This is the standard price Phillips-curve relation generalized to allow for partial dynamic indexation. In the special case without indexation (i.e.,  $\gamma_p = 0$ ), this relation reduces to the standard forward-looking Phillips curve relation, under which the price inflation depends on the current-period real marginal cost and the expected future inflation. In the presence of dynamic indexation, the price inflation also depends on its own lag.

Linearizing the optimal wage-setting decision rule implies that

$$\hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t - \gamma_w \hat{\pi}_{t-1} = \frac{\kappa_w}{1 + \eta\theta_w} (\hat{\mu}_{wt} + \hat{m}r_s t - \hat{w}_t) + \beta E_t[\hat{w}_{t+1} - \hat{w}_t + \hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t], \quad (33)$$

where  $\hat{w}_t$  denotes the log-deviations of the real wage,  $\hat{m}r_s t = \eta \hat{l}_t - \hat{U}_{ct}$  denotes the marginal rate of substitution between leisure and consumption,  $\theta_w \equiv \frac{\mu_w}{\mu_w - 1}$ , and  $\kappa_w \equiv$

---

<sup>2</sup>Derivations of the linearized equilibrium conditions are available upon request.

$\frac{(1-\beta\xi_w)(1-\xi_w)}{\xi_w}$  is a constant. To help understand the economics of this equation, we rewrite this relation in terms of the nominal wage inflation:

$$\begin{aligned}\hat{\pi}_t^w - \gamma_w \hat{\pi}_{t-1} &= \frac{\kappa_w}{1 + \eta\theta_w} (\hat{\mu}_{wt} + m\hat{r}s_t - \hat{w}_t) + \beta E_t (\hat{\pi}_{t+1}^w - \gamma_w \hat{\pi}_t) \\ &+ \frac{1}{1 - \alpha_1} [\alpha_1 (\Delta \hat{z}_t - \beta E_t \Delta \hat{z}_{t+1}) + \alpha_2 (\Delta \hat{q}_t - \beta E_t \Delta \hat{q}_{t+1})].\end{aligned}\quad (34)$$

where  $\hat{\pi}_t^w = \hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t + \Delta \hat{\lambda}_t^*$  denotes the nominal wage inflation. This nominal-wage Phillips curve relation parallels that of the price-Phillips curve and has similar interpretations.

The rest of the linearized equilibrium conditions are summarized below:

$$\begin{aligned}\hat{q}_{kt} &= S''(\lambda_I) \lambda_I^2 \left\{ \Delta \hat{i}_t - \beta E_t \Delta \hat{i}_{t+1} \right. \\ &\quad \left. + \frac{1}{1 - \alpha_1} [\alpha_2 (\Delta \hat{z}_t - \beta E_t \Delta \hat{z}_{t+1}) + \Delta \hat{q}_t - \beta E_t \Delta \hat{q}_{t+1}] \right\},\end{aligned}\quad (35)$$

$$\begin{aligned}\hat{q}_{kt} &= E_t \left\{ \Delta \hat{a}_{t+1} + \Delta \hat{U}_{c,t+1} - \frac{1}{1 - \alpha_1} [\alpha_2 \Delta \hat{z}_{t+1} + \Delta \hat{q}_{t+1}] \right. \\ &\quad \left. + \frac{\beta}{\lambda_I} [(1 - \delta) \hat{q}_{k,t+1} - \delta \hat{\delta}_{t+1} + \tilde{r}_k \hat{r}_{k,t+1}] \right\},\end{aligned}\quad (36)$$

$$\hat{r}_{kt} = \sigma_u \hat{u}_t, \quad (37)$$

$$0 = E_t \left[ \Delta \hat{a}_{t+1} + \Delta \hat{U}_{c,t+1} - \frac{1}{1 - \alpha_1} [\alpha_2 \Delta \hat{z}_{t+1} + \alpha_1 \Delta \hat{q}_{t+1}] + \hat{R}_t - \hat{\pi}_{t+1} \right], \quad (38)$$

$$\hat{k}_t = \frac{1 - \delta}{\lambda_I} \left[ \hat{k}_{t-1} - \frac{1}{1 - \alpha_1} (\alpha_2 \Delta \hat{z}_t + \Delta \hat{q}_t) \right] - \frac{\delta}{\lambda_I} \hat{\delta}_t + \left( 1 - \frac{1 - \delta}{\lambda_I} \right) \hat{i}_t, \quad (39)$$

$$\hat{y}_t = c_y \hat{c}_t + i_y \hat{i}_t + u_y \hat{u}_t + g_y \hat{g}_t, \quad (40)$$

$$\hat{y}_t = \alpha_1 \left[ \hat{k}_{t-1} + \hat{u}_t - \frac{1}{1 - \alpha_1} (\alpha_2 \Delta \hat{z}_t + \Delta \hat{q}_t) \right] + \alpha_2 \hat{l}_t, \quad (41)$$

$$\hat{w}_t = \hat{r}_{kt} + \hat{k}_{t-1} + \hat{u}_t - \frac{1}{1 - \alpha_1} (\alpha_2 \Delta \hat{z}_t + \Delta \hat{q}_t) - \hat{l}_t, \quad (42)$$

where (35) is the linearized investment decision equation with  $\hat{q}_{kt}$  denoting the shadow value of existing capital (i.e., Tobin's Q) and the  $\Delta$  denoting the first-difference operator (so that  $\Delta x_t = x_t - x_{t-1}$ ); (36) is the linearized capital Euler equation; (37) is the linearized capacity utilization decision equation with  $\sigma_u \equiv \frac{a''(1)}{a'(1)}$  denoting the curvature the function  $a(u)$  evaluated at the steady state; (38) is the linearized bond Euler equation; (39) is the linearized law of motion for the capital stock; (40) is the linearized aggregate resource constraint, with the steady-state ratios given by  $c_y = \frac{\tilde{C}}{\tilde{Y}}$ ,  $i_y = \frac{\tilde{I}}{\tilde{Y}}$ ,  $u_y = \frac{\tilde{r}_k \tilde{K}}{\tilde{Y} \lambda_I}$ , and  $g_y = \frac{\tilde{G}}{\tilde{Y}}$ ; (41) is the linearized aggregate production function; and (42) is the linearized factor demand relation.



Finally, the linearized interest rate rule is given by

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1 - \rho_r) [\phi_\pi (\hat{\pi}_t - \hat{\pi}^*(s_t)) + \phi_y \hat{y}_t] + \sigma_{rt} \varepsilon_{rt}, \quad (43)$$

where the term  $\hat{\pi}^*(s_t) \equiv \log \pi^*(s_t) - \log \pi$  denotes the deviations of the inflation target from its ergodic mean.

## V. ESTIMATION APPROACH

We estimate the parameters in our model using the Bayesian method. We describe a general empirical strategy so that the method can be applied to other regimes-switching DSGE models. As shown in the appendices, our model contains twenty seven variables. Adding the five lagged variables  $\hat{y}_{t-1}, \hat{c}_{t-1}, \hat{i}_{t-1}, \hat{w}_{t-1}$ , and  $\hat{q}_{t-1}$  to the list gives a total of thirty three variables. We denote all these state variables by the vector  $f_t$  where  $f_t$  is so arranged that the first eight variables are  $\hat{y}_t, \hat{c}_t, \hat{i}_t, \hat{w}_t, \hat{q}_t, \hat{\pi}_t, \hat{\ell}_t$ , and  $\hat{R}_t$  and the last five variables are  $\hat{y}_{t-1}, \hat{c}_{t-1}, \hat{i}_{t-1}, \hat{w}_{t-1}$ , and  $\hat{q}_{t-1}$ .

We apply the relation (28) to the policy rule (43), where the vector  $e_{s_t}$  defined in (29) follows a vector AR(1) process described in (30). Expanding the log-linearized system with the additional variables represented by  $e_{s_t}$  maintains the log-linear form in which all coefficients are constant (i.e., independent of regime changes). A standard solution technique, such as the method proposed by Sims (2002), can be directly utilized to solve our DSGE model. The solution leads to the following VAR(1) form of state equations

$$f_t = c(s_t, s_{t-1}) + F f_{t-1} + C(s_t^*) \epsilon_t, \quad (44)$$

where  $\epsilon_t = [\epsilon_{rt}, \epsilon_{pt}, \epsilon_{wt}, \epsilon_{gt}, \epsilon_{zt}, \epsilon_{at}, \epsilon_{dt}, \epsilon_{qt}]'$ , and  $c(s_t)$  is a vector function of the inflation targets  $\pi^*(s_t)$  and  $\pi^*(s_{t-1})$  and the elements in the transition matrix  $Q$ , and  $C(s_t^*)$  is a matrix function of  $\sigma_{rt}(s_t^*), \sigma_{pt}(s_t^*), \sigma_{wt}(s_t^*), \sigma_{gt}(s_t^*), \sigma_{zt}(s_t^*), \sigma_{at}(s_t^*), \sigma_{dt}(s_t^*)$ , and  $\sigma_{qt}(s_t^*)$ .

It follows from (44) that the solution to our DSGE model depends on the composite regime  $(s_t, s_{t-1}, s_t^*)$ . If  $s_t^*$  is assumed to be the same as  $(s_t, s_{t-1})$  (see Schorfheide (2005)), then the composite regime collapses to  $s_t$ . To simplify our notation and keep analytical expressions tractable, we use  $s_t$  to represent a composite regime that includes  $(s_t, s_{t-1}, s_t^*)$  as a special case for the rest of this section.

Our estimation is based on the 1959:I-2007:IV quarterly time-series observations on 8 U.S. aggregate variables:<sup>3</sup> real per capita GDP ( $Y_t^{\text{Data}}$ ), real per capita consumption

---

<sup>3</sup>We did not include the sample after 2007 because it is beyond the scope of this paper to address the current financial crisis and the effect of monetary policy at the lower zero bound.

( $C_t^{\text{Data}}$ ), real per capita investment ( $I_t^{\text{Data}}$ ), real wage ( $w_t^{\text{Data}}$ ), the investment-specific technology (i.e., the biased technology  $Q_t$ ), the quarterly GDP-deflator inflation rate ( $\pi_t^{\text{Data}}$ ), per capita hours ( $L_t^{\text{Data}}$ ), and the (annualized) federal funds rate ( $\text{FFR}_t^{\text{Data}}$ ). Note that  $I_t^{\text{Data}}$  corresponds to  $\frac{I_t}{Q_t}$  in the model (i.e., investment measured in units of consumption goods); a detailed description of the data is in Appendix A. These data are represented by the following vector of observable variables:

$$y_t = \left[ \Delta \ln Y_t^{\text{Data}}, \Delta \ln C_t^{\text{Data}}, \Delta \ln I_t^{\text{Data}}, \Delta \ln w_t^{\text{Data}}, \ln \pi_t^{\text{Data}}, \Delta \ln Q_t^{\text{Data}}, \ln L_t^{\text{Data}}, \frac{\text{FFR}_t^{\text{Data}}}{400} \right]'$$

The observable vector is connected to the model (state) variables through the measurement equations

$$y_t = a + Hf_t,$$

where

$$a = \left[ \ln \lambda_*, \ln \lambda_*, \ln \lambda_*, \ln \lambda_*, \ln \pi, \ln \lambda_q, \ln L, \ln R \right]'. \quad (45)$$

Given the aforementioned regime-switching state space form, one can estimate the model following the general estimation methodology of Sims, Waggoner, and Zha (2008).<sup>4</sup>

**V.1. Three methods for computing marginal data densities.** To evaluate the model's fit to the data and compare it to the fit of other models, one wishes to compute the marginal data density implied by the model. To keep the notation simple, let  $\theta$  represent a vector of all model parameters except the transition matrix and  $Q$  be a collection of all free parameters in the transition matrix. The marginal data density is defined as

$$p(Y_T) = \int p(Y_T | \theta, Q) p(\theta) d\theta dQ, \quad (46)$$

where the likelihood function  $p(Y_T | \theta, Q)$  can be evaluated recursively. For many empirical models, the modified harmonic mean (MHM) method of Gelfand and Dey (1994) is a widely used method to compute the marginal data density. The MHM method used to approximate (46) numerically is based on a theorem that states

$$p(Y_T)^{-1} = \int_{\Theta} \frac{h(\theta, Q)}{p(Y_T | \theta, Q)p(\theta, Q)} p(\theta, Q | Y_T) d\theta dQ, \quad (47)$$

where  $\Theta$  is the support of the posterior probability density and  $h(\theta, Q)$ , often called a *weighting* function, is any probability density whose support is contained in  $\Theta$ . Denote

$$m(\theta, Q) = \frac{h(\theta, Q)}{p(Y_T | \theta, Q)p(\theta, Q)}.$$

---

<sup>4</sup>The method details are also provided in an independent technical appendix to this article, which is available on <http://home.earthlink.net/tzha01/workingPapers/wp.html>.

A numerical evaluation of the integral on the right hand side of (47) can be accomplished in principle through the Monte Carlo (MC) integration

$$\hat{p}(Y_T)^{-1} = \frac{1}{N} \sum_{i=1}^N m(\theta^{(i)}, Q^{(i)}), \quad (48)$$

where  $(\theta^{(i)}, Q^{(i)})$  is the  $i^{\text{th}}$  draw of  $(\theta, Q)$  from the posterior distribution  $p(\theta, Q | Y_T)$ . If  $m(\theta, Q)$  is bounded above, the rate of convergence from this MC approximation is likely to be practical.

Geweke (1999) proposes a Gaussian function for  $h(\cdot)$  constructed from the posterior simulator. The likelihood and posterior density functions for our medium-scale DSGE model turn out to be quite non-Gaussian and there exist zeros of the posterior pdf in the interior points of the parameter space. In this case, the standard MHM procedure tends to be unreliable as the MCMC draws are likely to be dominated by a few draws as the number of draws increase. Sims, Waggoner, and Zha (2008) proposes a truncated non-Gaussian weighting function for  $h(\cdot)$  to remedy the problem. This weighting function seems to work well for the non-Gaussian posterior density.

In addition to the method of Sims, Waggoner, and Zha (2008), we use the unpublished method developed by Ulrich Müller at Princeton University. To summarize Müller's method for computing the marginal data density, we introduce the following notation. Let  $\theta$  be an  $n \times 1$  vector of random variables,  $p(\theta)$  be the target pdf, whose probability density is of unknown form, and  $p^*(\theta)$  be the target kernel where  $p(\theta) = c^* p^*(\theta)$ . Thus, our objective is to obtain an accurate estimate of the positive constant  $c^*$ . Let  $h(\theta)$  be an approximate or weighting pdf and  $c$  be a positive real number. Define the function  $f(c)$  as follows:

$$f(c) = E_h \left[ \mathbf{1} \left\{ \frac{cp^*(\theta)}{h(\theta)} < 1 \right\} \left( 1 - \frac{cp^*(\theta)}{h(\theta)} \right) \right] - E_g \left[ \mathbf{1} \left\{ \frac{h(\theta)}{cp^*(\theta)} < 1 \right\} \left( 1 - \frac{h(\theta)}{cp^*(\theta)} \right) \right].$$

One can show that this function has the following properties:

- $f(c)$  is monotonically decreasing in  $c$ ;
- $f(0) = 1$  and  $f(\infty) = -1$ .

Given these properties, one can use a bisection method to find an estimate of  $c^*$  where  $f(c^*) = 0$ .

A third method we use is bridge sampling of Meng and Wong (1996). The bridge-sampling method has been often regarded as one of the most reliable methods for

computing the Bayes factor. Since these three methods are developed from different mathematical relationships, we recommend using all these methods to ensure that the estimated value of the marginal data density is numerically similar across methods.

Because the posterior density function is very non-Gaussian and complicated in shape, it is all the more important to find the posterior mode via an optimization routine. The estimate of the mode not only represents the most likely value (and thus the posterior estimate) but also serves as a crucial starting point for initializing different chains of MCMC draws.

For various DSGE models studied in this paper, finding the mode has proven to be a computationally challenging task. The optimization method we use combines the block-wise BFGS algorithm developed by Sims, Waggoner, and Zha (2008) and various constrained optimization routines contained in the commercial IMSL package. The block-wise BFGS algorithm, following the idea of Gibbs sampling and EM algorithm, breaks the set of model parameters into subsets and uses Christopher A. Sims's `csminwel` program to maximize the likelihood of one set of the model's parameters conditional on the other sets.<sup>5</sup> Maximization is iterated at each subset until it converges. Then the optimization iterates between the block-wise BFGS algorithm and the IMSL routines until it converges. The convergence criterion is the square root of machine epsilon.

Thus far we have described the optimization process for only one starting point.<sup>6</sup> Our experience is that without such a thorough search, one can be easily misled to a much lower posterior value (e.g., a few hundreds lower in log value than the posterior peak). We thus use a set of cluster computing tools described in Ramachandran, Urazov, Waggoner, and Zha (2007) to search for the posterior mode. We begin with a grid of 100 starting points; after convergence, we perturb each maximum point in both small and large steps to generate additional 20 new starting points and restart the optimization process again; the posterior estimates attain the highest posterior density value. The other converged points typically have much lower likelihood values by at least a magnitude of hundreds of log values. For each DSGE model, the peak value of the posterior kernel and the mode estimates are reported.

---

<sup>5</sup>The `csminwel` program can be found on <http://sims.princeton.edu/yftp/optimize/>.

<sup>6</sup>For the no-switching (constant-parameter) DSGE model, it takes a couple of hours to find the posterior peak. While the model with two-regime shock variances takes about 20 hours to converge, the model with two-regime inflation targets and two-regime two-regime shock variances takes four times longer.

V.2. **Priors.** We set three parameters a priori. We set the steady-state government spending to output ratio at  $g_y = 0.18$ . We follow Justiniano and Primiceri (2006) and fix the persistence of the government spending shock process at  $\rho_g = 0.99$ . As noted by Smets and Wouters (2007), all these government parameter are difficult to estimate unless government spending is included in the set of measurement equations. Finally, we normalize and fix the steady-state hours worked at  $L = 0.2$ . We estimate all the remaining parameters. Tables 3 and 4 summarize the prior distributions for the structural parameters and the shock parameters.

Our priors are chosen to be more flexible and less tight than those in the previous literature. Specifically, instead of specifying the mean and the standard deviation, we use the 90% probability interval to back out the hyperparameter values of the prior distribution.<sup>7</sup> The intervals are generally set wide enough to allow the possibility of multiple posterior peaks (Del Negro and Schorfheide, 2008). Our approach is also necessary to deal with skewed distributions and allow for some reasonable hyperparameter values in certain distributions (such as the Inverse-Gamma) where the first two moments may not exist. The probability intervals reported in Table 3 cover the calibrated value of each parameter.

We begin with the preference parameters  $b$ ,  $\eta$ , and  $\beta$ . Our prior for the habit-persistence parameter  $b$  follows the Beta distribution. We choose the 2 hyper-parameters of the Beta distribution such that the lower bound for  $b$  (0.05) has a cumulative probability of 5% and the upper bound (0.948) has a cumulative probability of 95%. This 90% probability interval for  $b$  covers the values used by most economists (for example, Boldrin, Christiano, and Fisher (2001) and Christiano, Eichenbaum, and Evans (2005)). Our prior for the inverse Frisch elasticity  $\eta$  follows the Gamma distribution. We choose the 2 hyper-parameters of the Gamma distribution such that the lower bound (0.2) and the upper bound (10.0) of  $\eta$  correspond to the 90% probability interval. This prior range for  $\eta$  implies that the Frisch elasticity lies between 0.1 and 5, a range broad enough to cover the values based on both microeconomic evidence (Pencavel, 1986) and macroeconomic studies (Rupert, Rogerson, and Wright, 2000). Our prior for the transformed subjective discount factor  $\chi_\beta \equiv 100(\frac{1}{\beta} - 1)$  follows the Gamma distribution, with the hyper-parameters appropriately chosen such that the bounds for the 90% probability interval of  $\chi_\beta$  are 0.2 and 4.0. The implied value of  $\beta$  lies in the range between 0.9615 and 0.998, which nests the values obtained by Smets

---

<sup>7</sup>The program for backing out the hyperparameter values of a given prior can be found in <http://home.earthlink.net/~tzha02/ProgramCode/programCode.html>.

and Wouters (2007) ( $\beta = 0.9975$ ) and Altig, Christiano, Eichenbaum, and Linde (2004) ( $\beta = 0.9926$ ).

Next, we discuss the prior distributions for the technology parameters  $\alpha_1$ ,  $\alpha_2$ ,  $\lambda_q$ ,  $\lambda_*$ ,  $\sigma_u$ ,  $S''$ , and  $\delta$ . Our priors for the labor share and capital share both follow the Beta distribution with the restriction  $\alpha_1 + \alpha_2 \leq 1$  so that the production technology requires firm-specific factors (Chari, Kehoe, and McGrattan, 2000). Specifically, the bounds for the  $\alpha_1$  values in the 90% probability interval are 0.15 and 0.35 and those for  $\alpha_2$  are 0.35 and 0.75. With the restriction  $\alpha_1 + \alpha_2 \leq 1$ , however, the joint 90% probability region would be somewhat different. We assume that the priors for the (transformed) trend growth rates of the investment-specific technology and the neutral technology both follow the Gamma distribution, with the 5% and 95% bounds given by 0.1 and 1.5 respectively. These values imply that, with 90% probability, the prior values for the trend growth rates  $\lambda_q$  and  $\lambda_*$  lie in the range between 1.001 and 1.015 (corresponding to annual rates of 0.4% and 6%, respectively). We assume that the priors for the capacity utilization parameter  $\sigma_u$  and the investment adjustment cost parameter  $S''$  both follow the Gamma distribution, with the lower bounds given by 0.5 and 0.1 and the upper bounds given by 3.0 and 5.0, respectively. These 90% probability ranges cover the values obtained, for example, by Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007). We assume that the prior for the average annualized depreciation rate follows the Beta distribution with the 90% probability range lying between 0.05 and 0.20.

Third, we discuss the prior distributions for the parameters that characterize price and nominal wage setting in the model. These include the average price markup  $\mu_p$ , the average wage markup  $\mu_w$ , the Calvo probabilities of non-adjustment in pricing  $\xi_p$  and in wage-setting  $\xi_w$ , and the indexation parameters  $\gamma_p$  and  $\gamma_w$ . The priors for the net markups  $\mu_p - 1$  and  $\mu_w - 1$  both follow the Gamma distribution with the 90% probability range covering the values between 0.01 and 0.5. This range covers most of the calibrated values of the markup parameters used in the literature (e.g., Basu and Fernald (2002), Rotemberg and Woodford (1997), Huang and Liu (2002)). The priors for the price and wage duration parameters  $\xi_p$  and  $\xi_w$  both follow the Beta distribution with the 90% probability range between 0.1 and 0.75. Under this prior distribution, the nominal contract durations vary, with 90% probability, between 1.1 quarters and 4 quarters. This range covers the values of the frequencies of price and wage adjustments used in the literature (e.g., Bils and Klenow (2004), Taylor (1999)). The priors for the indexation parameters  $\gamma_p$  and  $\gamma_w$  both follow the uniform distribution with the 90%

probability range lying between 0.05 and 0.95. In this sense, we have loose priors on these indexation parameters, the range of which covers those used in most studies (e.g., Christiano, Eichenbaum, and Evans (2005), Smets and Wouters (2007), and Woodford (2003)).

Finally, we discuss the coefficients in the monetary policy rule, including  $\rho_r$ ,  $\phi_\pi$ , and  $\phi_y$ . The prior for the interest-rate smoothing parameter  $\rho_r$  follows the Beta distribution with the 90% probability range between 0.05 and 0.948. The prior for the inflation coefficient  $\phi_\pi$  follows the Gamma distribution with the 90% probability range between 0.5 and 5.0. The prior for the output coefficient  $\phi_y$  follows the Gamma distribution with the 90% probability range between 0.05 and 3.0. This range includes the values obtained by Clarida, Galí, and Gertler (2000) and others. These prior values allow for an indeterminacy region. When the equilibrium is indeterminate, we follow Boivin and Giannoni (2006) and use the MSV solution. In our estimation, however, there is practically little probability for the parameters to be in the indeterminate region.

Our priors for the AR(1) coefficients for the neutral and biased technology shocks  $\rho_q$  and  $\rho_z$  are uniformly distributed in the  $[0, 1]$  interval. The AR(1) coefficients for all other shocks and the MA(1) coefficients for the price and wage markup shocks follow the Beta distribution with the 5%-95% probability range given by  $[0.05, 0.948]$ . The prior for the parameter  $\rho_{gz}$  follows the Gamma distribution with the 90% probability range given by  $[0.2, 3.0]$ . The standard deviations of each of the 8 shocks follow the Inverse Gamma distribution with the 90% probability range given by  $[0.0005, 1.0]$ . This probability range implies a more agnostic prior than Smets and Wouters (2007) and Justiniano and Primiceri (2006). Such an agnostic prior is needed to allow for possible large changes in shock variances across regimes, as found in Sims and Zha (2006).

We have experimented with different priors. In one alternative prior, we follow the literature and make a prior on the persistence parameters in shock processes much tighter towards zero, such as the Beta(1, 2) probability density. Our conclusions hold true for these priors as well.

## VI. EMPIRICAL RESULTS

In this section, we report our main empirical findings. We compare in Section VI.1 the empirical fit of a variety of models nested by our general regime-switching DSGE framework. We then report in Section VI.2 the estimation results in our best-fit model and highlight the difference of these estimates from some alternative models.

VI.1. **Model Fit.** The first set of results to discuss is measures of model fit, with the comparison based on maximum log posterior densities adjusted by the Schwarz criterion.<sup>8</sup> Table 1 reports Schwarz criteria for different versions of our DSGE model (the column “Baseline”) and for models with the restriction that all the persistence parameters in both price markup and wage markup processes are set to zero (the column “Restricted”).

Table 1 shows that the model with regime shifts in shock variances only (DSGE-2v) is the best-fit model, much better than the constant-parameter DSGE model (DSGE-con). The Schwarz criterion for the baseline DSGE-2v model is 5963.03, compared to 5859.71 for the DSGE-con model. When we allow the inflation target to switch regimes while holding the shock variances constant (DSGE-2c), the model’s fit does not improve upon the constant-parameter DSGE model. When we allow both the inflation target and shock variances to switch regimes with the same Markov process (i.e., regime switching is synchronized), the model (DSGE-2cv) does better than the one with regime switching in the inflation target alone, but it does not improve upon the baseline DSGE-2v model with regime shifts in the shock variances only. When we relax the assumption that switches in the shock regime and those in the inflation target regime are synchronized and compute the Schwarz criterion for the model with the target regime and the shock regime independent of each other (DSGE-2c2v), we find that the model’s fit does not improve relative to either the DSGE-2cv model with synchronized regime shifts in the inflation target and the shock variances or the baseline DSGE-2v model with synchronized regime shifts in shock variances only. We have also examined the possibility of 3 shock regimes instead of 2. We find that the 3-regime model (DSGE-3v) does not improve upon the baseline 2-regime model (DSGE-2v).

We have also estimated models with shock variances following independent Markov switching processes. This scenario approximates stochastic volatility models, where each shock variance has its own independent stochastic process (Tauchen, 1986; Sims, Waggoner, and Zha, 2008). In addition, we have grouped a subset of shock variances having the same Markov processes. None of these models fits to the data better than our baseline DSGE-2v model. For example, when we allow regimes associated with the variances of the two technology shocks to be independent of the regime switching processes of the other shock variances (DSGE-2v2v), we obtain a Schwarz criterion of 5958.18, which is lower than that of the baseline DSGE-2v model (5963.03). In short,

---

<sup>8</sup>The Schwarz criterion is similar to the Laplace approximation used by Smets and Wouters (2007).



the data favor the parsimoniously-parameterized model with shock variances switching regimes simultaneously.

The last column in Table 1 shows that the model with regime changes in shock variances only continues to dominate all the other models, when the persistence parameters in both price and wage markup shock processes are restricted to zero. In particular, the model with the target switching regimes (DSGE-2c) does not improve upon the constant-parameter model. Of course, all these restricted models fit to the data much worse than the corresponding baseline models, implying that persistent shock processes are important in fitting the data.

Finally, we have estimated a number of models with persistence parameters in other shock processes set to zero and with habit and indexation parameters set to zero. The model with synchronized regimes in shock variances continue to outperform other models in fitting the data.

The relative performance of the alternative DSGE models in fitting the data does not change when we look at the marginal data density (MDD). Table 2 reports the MDD for each of the alternative models. The table shows that the model with simultaneous regime shifts in shock variances (DSGE-2v) is the best-fit model not only in terms of the Schwarz criterion, but also in terms of the marginal data density. In particular, the DSGE-2v model's MDD is 5832.38, much higher than that of the DSGE-con model (whose MDD is 5741.24). The model with regime switching in the inflation target alone (DSGE-2c) slightly outperforms the constant parameter model, but substantially under-performs the DSGE-2v model. With regime shifts in shock variances, introducing regime shifts in the inflation target synchronized with regime shifts in shock variances (DSGE-2cv) or allowing the inflation target to follow a Markov switching process independent of shock regimes (DSGE-2c2v) does not improve the marginal data density relative to the DSGE-2v model.<sup>9</sup>

**VI.2. Estimates of Structural Parameters.** We first discuss our best-fit model "DSGE-2v." The model is similar to that in Smets and Wouters (2007) with six notable exceptions. First, we introduce a source of real rigidity in the form of firm-specific factors, which replaces the kinked demand curves considered by Smets and Wouters (2007). Second, we introduce trend growth in the investment-specific technological change to better capture the data, in which the relative price of investment goods

---

<sup>9</sup>The good fit represented by DSGE-2cv comes entirely from significant shifts in shock variances. The estimated inflation targets are 2.18% for one regime and 1.70% for the other regime and the difference between these two targets are statistically insignificant.

(e.g., equipment and software) has been declining for most of the postwar period, while in Smets and Wouters (2007) the investment-specific technological changes have no trend component. We use the observed time series of biased technological changes in our estimation, while Smets and Wouters (2007) treat these changes as a latent variable. Third, we introduce the depreciation shock that acts as a wedge in the capital-accumulation Euler equation. Fourth, the preference shock in our model enters all intertemporal decisions, including choices of the nominal bond, the capital stock, and investment, while Smets and Wouters (2007) introduce a “risk-premium shock” that enters the bond Euler equation only and does not affect other intertemporal decisions. Fifth, in the interest rate rule, we assume that the nominal interest rate responds to deviations of inflation from its target and detrended output, while in Smets and Wouters (2007) the interest rate rule targets inflation, output gap, and the growth rate of output gap. Finally, we allow for heteroscedasticity of structural shocks to obtain the accurate estimate of the role of a particular shock in explaining macroeconomic fluctuations. All these distinctions may explain some of the differences between our estimated results and theirs.

Tables 3 and 4 report the estimates of the model parameters. The data are informative about many structural parameters. Among the three preference parameters, the estimate for habit persistence ( $b$ ) is 0.91 with the tight error bands. The estimate for  $\eta$  is 2.89, implying a Frisch elasticity of 0.35 and consistent with most microeconomic studies. The probability interval indicates that  $\eta$  can be as high as 8.38. The estimate for the subjective discount factor  $\beta$  is 0.998 (the same as the value obtained by Smets and Wouters (2007)) with the tight probability interval [0.996, 0.999].

Among the technology parameters, the estimate for  $\alpha_1$  (0.153) with the upper error band (0.216) close to the estimate obtained by Smets and Wouters (2007) (0.19). Because of the constraint  $\alpha_1 + \alpha_2 \leq 1$ , the estimate for  $\alpha_2$  is (0.835). These posterior estimates suggest that the data prefer a model specification with (near) constant-returns production technology. The estimated trend growth rate for the investment-specific technological change ( $\lambda_q$ ) is 4% per annum, slightly higher than the calibrated value obtained by Greenwood, Hercowitz, and Krusell (1997) because we include the data in the late 1990s until 2007 when the investment-specific technological improvement was the fastest in the sample. The estimate for the trend growth rate of the neutral technological change ( $\lambda_*$ ) is 0.95% per annum. There is a large amount of uncertainty about these trend estimates as shown in the last two columns of Table 3. The curvature parameter in the utilization function ( $\sigma_u$ ) is estimated at 2.26, substantially

lower than the value obtained by Justiniano and Primiceri (2006) (7.13), but higher than the values estimated by Altig, Christiano, Eichenbaum, and Linde (2004) (2.02) and by Smets and Wouters (2007) (1.174). The error bands show a large amount of uncertainty around the estimate of this parameter. The investment adjustment cost parameter ( $S''$ ) is estimated to be 2.0, lower than those obtained in the literature. Unlike most studies in the literature that fix the value of the capital depreciation rate a priori, we allow the depreciation rate  $\delta$  to follow a stationary stochastic process and estimate the parameter in the process. The estimated average annual depreciation rate is 13.4%, which is remarkably close to the standard calibration value in the real business cycle literature, but the error bands are very wide, implying the great uncertainty about this estimate.

Among the pricing and wage setting parameters, the estimated average price markup ( $\mu_p$ ) is about 1.0, which is consistent with the studies by Hall (1988), Basu and Fernald (1997), and Rotemberg and Woodford (1999), who argue that the pure economic profit is close to zero. It is also similar to the estimate obtained by Altig, Christiano, Eichenbaum, and Linde (2004), but much smaller than the value estimated by Justiniano and Primiceri (2006). Our estimate for the average wage markup ( $\mu_w$ ) is 1.06, which is lower than the calibrated value (Huang and Liu, 2002) and the estimated value (Justiniano and Primiceri, 2006), but is similar to the value used by Christiano, Eichenbaum, and Evans (2005). The uncertainty about the wage markup parameter, judged by the .90 probability bands, is much larger than that about the price markup parameter. The estimated price and wage stickiness parameters ( $\xi_p = 0.412$  and  $\xi_w = 0.213$ ) imply that, on average, price contracts last for less than 2 quarters and nominal wage contracts have an even shorter duration, which is slightly more than 1 quarter. Our estimated nominal contract duration is consistent with the microeconomic studies such as Bils and Klenow (2004). The estimated dynamic indexation is unimportant for price setting ( $\gamma_p = 0.178$ ) but very important for nominal wage setting ( $\gamma_w = 1.0$ ). The .90 probability intervals indicate that while the price indexation is tightly estimated, the uncertainty about the nominal wage indexation is extremely large.

As shown in Tables 3, the estimated wage stickiness parameter lies below the lower bound of the .90 probability interval. This phenomenon occurs because the posterior distribution around the mode for this parameter is on the thin ridge and because there are many local peaks that give a significant probability to regions containing the values above the estimated wage stickiness parameter. While it is impossible to graph this phenomenon in a high dimensional parameter space like ours, we display in Figure 1

the joint distribution of the wage stickiness parameter and the average wage markup parameter after integrating out all other parameters. As one can see, the multiple local peaks give much of the probability to the values of the wage stickiness parameters greater than the estimate at the posterior mode. Because the two-dimensional distribution displayed in Figure 1 integrates out all other parameters, the distribution is already skewed toward the values of the wage stickiness parameters greater than 0.2. Nonetheless, the picture demonstrates clearly the nature of thin ridges and multiple local peaks inherent in the posterior distribution.

The estimates of policy parameters suggest that interest-rate smoothing is important; the estimate of  $\rho_r$  is 0.82 with a narrow probability interval. The policy response to deviations of inflation from its target in the interest rule ( $\phi_\pi$ ) is 1.655 with the lower probability bound still significantly above 1.0. Policy does not respond much to detrended output and the parameter ( $\phi_y$ ) is tightly estimated. The inflation target ( $\pi^*$ ) is estimated at 2.28% per annum.

The estimated results for shock processes are reported in Table 4. The AR(1) coefficients for all shocks except the preference shock ( $\rho_a$ ) are above 0.9, although the lower probability bounds for some coefficients are substantially below (0.9). The preference shock is almost i.i.d.. The MA(1) coefficients in the price markup and wage markup processes ( $\phi_p$  and  $\phi_w$ ) are both sizable. The estimates are 0.698 and 0.749 and the corresponding .90 probability intervals support these high values. The government spending shock responds to the neutral technology shock; the response coefficient ( $\rho_{gz}$ ) is 0.894 with a wide probability interval. Although the prior distributions for all the shock variances are the same, the posterior estimates are very disperse. The depreciation shock ( $\sigma_d$ ) and the wage markup shock ( $\sigma_w$ ) have the largest variances; the monetary policy shock ( $\sigma_r$ ) and the two types of technology shocks ( $\sigma_z$  and  $\sigma_q$ ) have the smallest variances. The .90 probability intervals indicate that the marginal posterior distribution of a shock variance is skewed to the right. This shape is expected as the variance is bounded below by zero below and has no upward bound.

As shown in Table 4, the estimated shock variances in the second regime are substantially smaller than those in the first regime. The estimated transition probabilities are summarized by the matrix

$$\hat{Q} = \begin{bmatrix} 0.8072 & 0.0598 \\ 0.1928 & 0.9402 \end{bmatrix}, \quad (49)$$

where the elements in each column sum to one. The second regime (i.e., the regime with low shock variances) is more persistent and, as shown in Figure 2, covers most

of the period since Greenspan became Chairman of the Federal Reserve Board. This result is even stronger when one take into account the error bands, where the lower bound of  $q_{22}$  is higher than the upper bound of  $q_{11}$ .

Figure 3 plots the marginal posterior distribution of some key parameters. The local peaks shown in the marginal distribution of the inflation target are the direct outcome of the integrated effect of the non-Gaussian joint posterior distribution of all parameters that has thin ridges and multiple peaks. Most of the probability, however, concentrates between 2% and 4%. The marginal distribution of the response coefficient to inflation in the Taylor indicates that there is practically no probability for indeterminate equilibria for our model.

The marginal distribution of the price-stickiness parameter implies that the price rigidity is much smaller than what is obtained in the previous literature. The posterior mode is near the lower tail of the marginal distribution. The joint distribution, as illustrated in Figure 1, has a thin ridge and many local peaks. After integrating out all other parameters, the marginal distribution of the wage-stickiness parameter shows a local peak around 0.7. The majority of the probability, however, lies below the value 0.6.

There are two reasons why we obtain estimates that imply shorter durations of price and wage contracts than those obtained in the literature such as Altig, Christiano, Eichenbaum, and Linde (2004) and Smets and Wouters (2007). First, our estimates suggest that the price markup is very small, implying that the demand curve for differentiated goods is very flat. Thus, a small increase in the relative price can lead to large declines in relative output demand. Even if firms can re-optimize their pricing decisions very frequently, they choose not to adjust their relative prices too much. In this sense, the small average markup and thus the large demand elasticity become a source of strategic complementarity in firms' pricing decisions. Second, unlike Altig, Christiano, Eichenbaum, and Linde (2004) who use a minimum-distance estimator that matches the model's impulse responses to those in the data, we use full-information maximum likelihood estimation. This difference is important because Altig, Christiano, Eichenbaum, and Linde (2004) find that, while a shock to neutral technology leads to rapid adjustments in prices, a shock to monetary policy leads to small and gradual price adjustments. Under their estimation approach, matching the impulse responses following the monetary policy shock is important so that price adjustments have to be small and gradual. Our estimation approach differs from theirs and we find that the most important shocks are those to neutral technology, capital depreciation,

and wage markup, all of which lead to rapid adjustments in prices. Consequently, our estimated durations of nominal contracts are shorter than those in the literature.

The last row of Figure 3 displays the marginal posterior distributions of the investment technology trend and the wage indexation. The distribution of the investment technology trend puts a significant amount of probability around 4%, consistent with the data on the relative price of investment. The distribution of the wage indexation parameter is most interesting. While the estimate is at 1.0, there is considerable uncertainty around the wage indexation parameter so that the estimate of 1.0 is very imprecise. This result implies that our estimation does not necessarily support a strong wage indexation.

## VII. ECONOMIC IMPLICATIONS

We now discuss the economic implications of our best-fit model. We first examine, in Section VII.1, the role of the various shocks in driving macroeconomic fluctuations through variance decompositions. We then present, in Section VII.2, impulse responses of several key aggregate variables to each of the shocks that we identify as important for macroeconomic fluctuations. Finally, we provide some economic interpretations of the key sources of shocks and in particular, the capital depreciation shock.

**VII.1. Variance decompositions.** Tables 5 and 6 report variance decompositions in forecast errors of output, investment, hours, the real wage, and inflation under the two shock regimes at different forecasting horizons for our best-fit model. As we have discussed in Section VI.2, the wage markup shock and the depreciation shock have the largest variances among all eight structural shocks. The neutral technology shock is of considerable interest because of the debate in the recent literature on its dynamic effects on the labor market variables (e.g., Galí (1999), Christiano, Eichenbaum, and Vigfusson (2003), Uhlig (2004), and Liu and Phaneuf (2007)).

As we can see, capital depreciation shocks, neutral technology shocks, and wage markup shocks play an important role in driving business cycle fluctuations under both regimes. Taken together, these three types of shocks account for 70 – 80% of the fluctuations in output, investment, hours, and inflation under each regime for the forecast horizons beyond eight quarters. Monetary policy shock accounts for a sizable fraction of inflation fluctuations under the first regime but otherwise it is unimportant. The price markup shock contributes to about 15 – 30% of the real wage fluctuations under both regimes. It is also somewhat important for inflation fluctuations under the second regime. The remaining three shocks, including the government spending shock,

the preference shock, and the biased technology shock are unimportant in explaining macroeconomic fluctuations.

**VII.2. Impulse responses.** To gain intuition about the model's transmission mechanisms, we analyze impulse responses of selected variables following some of the structural shocks. In particular, we focus on the dynamic effects of a wage markup shock, a neutral technology shock, and a depreciation shock on output, investment, the real wage, the inflation rate, hours, and the nominal interest rate. These shocks, as we discuss in the previous section, are the most important driving sources of macroeconomic fluctuations. Since the impulse responses display the same patterns for both shock regimes except the scaling effect, we report the impulse responses only for the second regime.

Figure 4 displays the impulse responses following a one-standard-deviation shock to the capital depreciation rate. The increase in the depreciation rate reduces the value of capital accumulation and raises utilization and the rental price of capital; thus investment falls. Since the expected stock of capital wealth declines, the negative wealth effect leads to a fall in consumption as well. Consequently, aggregate output falls. The decline in output leads to a decline in hours. The decline in hours and in consumption lowers the marginal rate of substitution between labor and consumption, so that the households' desired wage falls. Thus, the equilibrium real wage declines as well. The fall in the real wage reduces the firms' marginal cost so that inflation declines. Through the Taylor rule, the nominal interest rate declines as well. As the .90 probability error bands show, all the responses are statistically significant.

Figure 5 reports the impulse responses following a one-standard-deviation shock to the investment-specific technology. The biased shock raises the efficiency of investment, investment goods today become cheaper, and current consumption becomes more expensive. This type of shock, unlike the depreciation shock or the neutral technology shock, shifts resources from consumption to investment. Consequently, investment rises and consumption declines. Hours declines initially due to the costly adjustment in investment as well as the habit formation. After the second quarter, the increase in demand for investment gradually leads to a rise in hours and the real wage. The rise in labor hours helps produce more output. Utilization and the rental price of capital rise as well. All the responses are well estimated, judged by the .90 probability error bands. In contrast to the responses to the depreciation shock, the biased technology shock generates opposite movements in output and consumption in the short run and

consequently its impact on the macroeconomy is much smaller (by comparing the scales in Figure 4 and those in Figure 5).

Both the capital depreciation shock and the investment-specific technology shock enter the intertemporal capital accumulation decision. But we find that this biased technology shock is much less important for macroeconomic fluctuations than the depreciation shock. This finding is different from that in Justiniano, Primiceri, and Tambalotti (2008), mainly because we use direct observations on the biased technology shock in our estimation while they do not.

Figure 6 displays the impulse responses following a one-standard-deviation shock to the neutral technology (i.e., the total factor productivity, or TFP). The positive neutral technology shock raises output, consumption, investment, utilization of capital, and the real wage. All these responses are statistically significant for the most part. The shock should lower inflation and, through the Taylor rule, the nominal interest rate. But the error bands are wide so that the estimates are insignificant.

The neutral technology shock leads to a statistically significant decline in hours worked. The decline in hours here, however, is not a direct consequence of price stickiness. Even with much more frequent price adjustments, we find that the positive neutral technology shock leads to a decline in hours (not reported). Instead, the investment adjustment cost (as well as the habit formation to a less extent) plays an important role in generating the decline in hours. If the investment adjustment cost parameter is small, we find that the model generates an increase in hours following the neutral technology shock (not reported), regardless of whether prices are sticky or not. Thus, our finding does not support the view that the contractionary effect of a neutral technology shock arises from the price stickiness. It is consistent with Francis and Ramey (2005), who argue that a real business cycle model with habit persistence and investment adjustment cost can generate a decline in hours following a positive neutral technology shock.

Figure 7 reports the impulse responses following a one-standard-deviation shock to the wage markup. An increase in the wage markup raises the households' desired real wage. The households who can adjust their nominal wage raise their nominal wage. The increase in the nominal wage raises the firms' marginal cost so that inflation rises and real aggregate demand falls. It follows that aggregate output, investment, and hours decline. Consequently, the rental price of capital and utilization rise. Through the interest-rate rule, the rise in inflation leads to an increase in the nominal interest rate. All these responses are statistically significant.



**VII.3. What is a shock to capital depreciation?** The variance decompositions indicate that the TFP shock, the wage markup shock, and the depreciation shock are the most important sources of macroeconomic fluctuations. The TFP shock and the wage markup shock should be familiar to many researchers, but the capital depreciation shock is new. Given its importance in accounting for the macroeconomic fluctuations in our model, it is useful to provide economic interpretations of the depreciation shock.

Like the TFP shock or any other shocks in this class of models, the depreciation shock is of reduced form that captures some “deeper” sources of disturbances and possibly microeconomic frictions that distort intertemporal capital accumulation decisions. Greenwood, Hercowitz, and Krusell (1997) draw a mapping between investment-specific technological changes and economic depreciation (as opposed to physical depreciation) of capital. They note that the economic depreciation rate rises when the equipment price relative to the consumption price is expected to decline in the future. As the equipment price is expected to fall, existing capital is worth less and investors have incentive to postpone investment to future periods, leading to a contraction in current economic activity, as does our depreciation shock.

Our depreciation shock also closely resembles the capital quality shock in Justiniano, Primiceri, and Tambalotti (2008) and Gertler and Kiyotaki (2010), who interpret their capital quality shock as representing some exogenous changes in the value of capital. One possible microeconomic interpretation is that a large number of goods are produced using good-specific capital. In each period, as a fraction of goods becomes obsolete randomly, the capital used for producing those obsolete goods becomes worthless. In aggregate, the law of motion for capital would feature a depreciation shock or similarly a capital quality shock to reflect the economic obsolescence of capital.

Thus, we view the depreciation shock as a stand in for economic obsolescence of capital. Unlike other intertemporal wedges such as the investment-specific technology shock (or biased technology shock), the depreciation shock in our model generates positive comovement between consumption, investment, hours, and the real wage. Our empirical results in general suggest that the depreciation shock, along with the standard TFP shock and wage markup shock, is an important driving source of business cycle fluctuations in the U.S. economy.

## VIII. CONCLUSION

We have studied a variety of fairly large DSGE models within a unified framework to reexamine the sources of observed macroeconomic fluctuations in the post-WWII U.S.

economy. Our econometric estimation suggests that heteroscedasticity in shock disturbances are important and that changes in shock variances take place simultaneously rather than independently. Three types of shocks stand out as the most important sources of macroeconomic volatilities: a depreciation shock that functions as an intertemporal wedge in capital accumulation, a total factor productivity shock that acts as an efficiency wedge, and a wage markup shock that serves as an intratemporal labor supply wedge. We do not find evidence of changes in the inflation target, nor do we find support for strong nominal rigidities in prices and nominal wages. These findings are robust across a large set of regime-switching models.

#### APPENDIX A. DETAILED DATA DESCRIPTION

All data are either taken directly from the Haver Analytics Database or constructed by Patrick Higgins at the Federal Reserve Bank of Atlanta. The construction methods developed or used by Patrick Higgins, available on request, will be briefly described below.

The model estimation is based on quarterly time-series observations on 8 U.S. aggregate variables during the sample period 1959:Q1–2007:Q4. The 8 variables are real per capita GDP ( $Y_t^{\text{Data}}$ ), real per capita consumption ( $C_t^{\text{Data}}$ ), real per capita investment ( $I_t^{\text{Data}}$ ) in capital goods, real wage ( $w_t^{\text{Data}}$ ), the quarterly GDP-deflator inflation rate ( $\pi_t^{\text{Data}}$ ), per capita hours ( $L_t^{\text{Data}}$ ), the federal funds rate ( $\text{FFR}_t^{\text{Data}}$ ), and the inverse of the relative price of investment ( $Q_t^{\text{Data}}$ ).

These series are derived from the original data in the Haver Analytics Database (with the relevant data codes provided) or from the constructed data.

- $Y_t^{\text{Data}} = \frac{\text{GDPH}}{\text{POP25-64}}$ .
- $C_t^{\text{Data}} = \frac{(\text{CN@USECON} + \text{CS@USECON}) * 100 / \text{JGDP}}{\text{POP25-64}}$ .
- $I_t^{\text{Data}} = \frac{(\text{CD@USECON} + \text{F@USECON}) * 100 / \text{JGDP}}{\text{POP25-64}}$ .
- $w_t^{\text{Data}} = \frac{\text{LXNFC@USECON} / 100}{\text{JGDP}}$ .
- $\pi_t^{\text{Data}} = \frac{\text{JGDP}_t}{\text{JGDP}_{t-1}}$ .
- $L_t^{\text{Data}} = \frac{\text{LXNFMH@USECON}}{\text{POP25-64}}$ .
- $\text{FFR}_t^{\text{Data}} = \frac{\text{FFED@USECON}}{400}$ .
- $Q_t^{\text{Data}} = \frac{\text{JGDP}}{\text{TornPriceInv4707CV}}$ .

The original data, the constructed data, and their sources are described as follows.

**POP25-64:** civilian noninstitutional population with ages 25-64 by eliminating breaks in population from 10-year censuses and post 2000 American Community Surveys using “error of closure” method. This fairly simple method was used by

the Census Bureau to get a smooth population monthly population series. This smooth series reduces the unusual influence of drastic demographic changes.

**GDPH:** real gross domestic product (2000 dollars). Source: BEA.

**CN@USECON:** nominal personal consumption expenditures: nondurable goods. Source: BEA

**CS@USECON:** nominal consumption expenditures: services. Source: BEA.

**CD@USECON:** nominal personal consumption expenditures: durable goods. Source: BEA.

**F@USECON:** nominal private fixed investment. Source: BEA.

**JGDP:** gross domestic product: chain price index (2000=100). Source: BEA.

**LXNFC@USECON:** nonfarm business sector: compensation per hour (1992=100). Source: BLS.

**LXNFH@USECON:** nonfarm business sector: hours of all persons (1992=100). Source: BLS.

**FFED@USECON:** annualized federal funds effective rate. Source: FRB.

**TornPriceInv4707CV:** investment deflator. The Tornquist procedure is used to construct this deflator as a weighted aggregate index from the four quality-adjusted price indexes: private nonresidential structures investment, private residential investment, private nonresidential equipment & software investment, and personal consumption expenditures on durable goods. Each price index is a weighted one from a number of individual price series within this categories. For each individual price series from 1947 to 1983, we use Gordon (1990)'s quality-adjusted price index. Following Cummins and Violante (2002), we estimate an econometric model of Gordon's price series as a function of a time trend and a few NIPA indicators (including the current and lagged values of the corresponding NIPA price series); the estimated coefficients are then used to extrapolate the quality-adjusted price index for each individual price series for the sample from 1984 to 2007. These constructed price series are annual. Denton (1971)'s method is used to interpolate these annual series on a quarterly frequency. The Tornquist procedure is then used to construct each quality-adjusted price index from the appropriate interpolated quarterly price series.

TABLE 1. Schwarz Criterion for the Set of DSGE Models

Model	Baseline	Restricted
DSGE-con	5859.71	5811.14
DSGE-2v	5963.03	5920.01
DSGE-2c	5853.13	5805.27
DSGE-2cv	5960.71	5907.00
DSGE-2c2v	5958.78	5913.85
DSGE-2v2v	5958.18	5912.22
DSGE-3v	5950.73	5926.91

*Note:* Column 1 lists the models studied: the DSGE model with all parameters that are constant across time (DSGE-con), the DSGE model with two regimes in shock variances (DSGE-2v), the DSGE model with two regimes in the inflation target only (DSGE-2c), the DSGE model with two common regimes for both shock variances and the inflation target (DSGE-2cv), and the DSGE model with two independent Markov processes, one controlling two regimes in shock variances and the other controlling two regimes in the inflation target (DSGE-2c2v), the DSGE model with two independent Markov processes, one controlling two regimes in variances of two technology shocks and the other controlling two regimes in variances of all the other shocks (DSGE-2v2v), and the DSGE model with three regimes in shock variances (DSGE-3v). Column 2 reports the posterior densities at the posterior mode, adjusted by Schwarz criterion. Column 3 displays the posterior densities evaluated at the posterior modes for models with the persistence parameters in both the price and wage markup processes set to zero.

TABLE 2. Comprehensive Measures of Model Fits

Model	Marginal Data Density
DSGE-con	5741.24
DSGE-2v	5832.38
DSGE-2c	5739.32
DSGE-2cv	5832.60
DSGE-2c2v	5830.84
DSGE-2v2v	5826.95
DSGE-3v	5813.91

TABLE 3. Prior and Posterior Distributions of Structural Parameters for the model “DSGE-2v.”

Parameter	Prior			Posterior		
	Distribution	5%	95%	Mode	5%	95%
$b$	Beta	0.05	0.948	0.907	0.8898	0.9533
$\alpha_1$	Beta	0.15	0.35	0.163	0.1701	0.2162
$\alpha_2$	Beta	0.35	0.75	0.835	0.7348	0.8101
$\eta$	Gamma	0.2	10.0	2.888	2.7151	8.3848
$100(\lambda_q - 1)$	Gamma	0.1	1.5	1.000	0.7663	1.5001
$100(\lambda_* - 1)$	Gamma	0.1	1.5	0.237	0.0876	0.4414
$100(\beta^{-1} - 1)$	Beta	0.2	4.0	0.175	0.0518	0.3915
$\sigma_u$	Gamma	0.5	3.0	2.263	1.3343	5.5196
$S''$	Gamma	0.5	5.0	2.000	1.2086	3.6145
$\mu_p - 1$	Gamma	0.01	0.50	0.000	0.0001	0.0032
$\mu_w - 1$	Gamma	0.01	0.50	0.060	0.0166	0.4151
$4\delta$	Beta	0.05	0.2	0.134	0.0956	0.5928
$\xi_p$	Beta	0.1	0.75	0.412	0.2312	0.6202
$\gamma_p$	Beta	0.05	0.95	0.178	0.0261	0.3390
$\xi_w$	Beta	0.1	0.75	0.213	0.2482	0.6779
$\gamma_w$	Beta	0.05	0.95	1.000	0.2275	0.9724
$\rho_r$	Beta	0.05	0.948	0.816	0.7923	0.8726
$\phi_\pi$	Gamma	0.5	5.0	1.655	1.4119	2.2899
$\phi_y$	Gamma	0.05	3.0	0.043	0.0241	0.1168
$400 \log \pi^*$	Gamma	1.0	8.0	2.283	1.2228	6.1118

*Note:* “5%” and “95%” demarcate the bounds of the 90% probability interval. “DSGE-2v” denotes the model with two regimes in shock variances.

TABLE 4. Prior and Posterior Distributions of Shock Parameters for the model “DSGE-2v.”

Parameter	Prior			Posterior		
	Distribution	5%	95%	Mode	5%	95%
$\rho_p$	Beta	0.05	0.948	0.949	0.8524	0.9642
$\phi_p$	Beta	0.05	0.948	0.698	0.5009	0.8439
$\rho_w$	Beta	0.05	0.948	0.999	0.8047	0.9973
$\phi_w$	Beta	0.05	0.948	0.749	0.6597	0.9307
$\rho_{gz}$	Gamma	0.2	3.0	0.894	0.4555	1.3748
$\rho_a$	Beta	0.05	0.948	0.107	0.0471	0.4010
$\rho_q$	Beta	0.05	0.95	0.994	0.9875	0.9973
$\rho_z$	Beta	0.05	0.95	0.992	0.9832	0.9984
$\rho_d$	Beta	0.05	0.948	0.934	0.9172	0.9820
$\sigma_r(1)$	Inverse Gamma	0.0005	1.0	0.004	0.0038	0.0066
$\sigma_r(2)$	Inverse Gamma	0.0005	1.0	0.001	0.0012	0.0016
$\sigma_p(1)$	Inverse Gamma	0.0005	1.0	0.039	0.0312	0.1426
$\sigma_p(2)$	Inverse Gamma	0.0005	1.0	0.028	0.0211	0.0819
$\sigma_w(1)$	Inverse Gamma	0.0005	1.0	0.255	0.2708	2.9007
$\sigma_w(2)$	Inverse Gamma	0.0005	1.0	0.144	0.1452	1.6253
$\sigma_g(1)$	Inverse Gamma	0.0005	1.0	0.041	0.0332	0.0566
$\sigma_g(2)$	Inverse Gamma	0.0005	1.0	0.021	0.0192	0.0247
$\sigma_z(1)$	Inverse Gamma	0.0005	1.0	0.010	0.0099	0.0167
$\sigma_z(2)$	Inverse Gamma	0.0005	1.0	0.006	0.0064	0.0083
$\sigma_a(1)$	Inverse Gamma	0.0005	1.0	0.043	0.0375	0.1371
$\sigma_a(2)$	Inverse Gamma	0.0005	1.0	0.037	0.0337	0.0819
$\sigma_q(1)$	Inverse Gamma	0.0005	1.0	0.007	0.0064	0.0109
$\sigma_q(2)$	Inverse Gamma	0.0005	1.0	0.002	0.0026	0.0034
$\sigma_d(1)$	Inverse Gamma	0.0005	1.0	0.193	0.1442	1.0925
$\sigma_d(2)$	Inverse Gamma	0.0005	1.0	0.099	0.0702	0.4989
$q_{11}$	Dirichlet	0.589	0.991	0.807	0.6045	0.8749
$q_{22}$	Dirichlet	0.589	0.991	0.940	0.9063	0.9769

*Note:* “5%” and “95%” demarcate the bounds of the 90% probability interval. “DSGE-2v” denotes the model with two regimes in shock variances.

TABLE 5. Forecast Error Variance Decomposition: Regime I

Horizon	MP	PM	WM	GS	Ntech	Pref	Btech	Dep
Output								
4Q	5.1443	4.1486	21.7621	12.8076	21.7050	1.4694	0.2106	32.7525
8Q	2.6618	4.1472	36.6626	6.3157	23.4311	0.5821	0.3289	25.8707
16Q	1.2227	2.9278	49.3857	3.4905	25.7249	0.2658	0.3942	16.5885
20Q	0.9756	2.5166	52.4044	2.9881	26.1211	0.2121	0.4119	14.3702
Investment								
4Q	8.2082	5.7979	14.4682	0.7438	4.7909	0.5182	1.5667	63.9061
8Q	4.6292	6.2905	22.8817	1.1375	7.1533	0.7034	2.2140	54.9904
16Q	3.2021	5.9545	30.5410	1.2310	10.2411	0.6839	3.6330	44.5134
20Q	3.0318	5.7789	32.0371	1.1764	11.1034	0.6478	4.2787	41.9460
Hours								
4Q	6.2409	4.6560	33.5268	18.7046	3.9659	2.0046	0.0815	30.8198
8Q	3.4031	4.8188	58.8519	10.9127	1.8040	0.9368	0.1388	19.1339
16Q	1.7212	3.0812	76.6957	6.8522	0.9926	0.4698	0.1114	10.0761
20Q	1.3929	2.5430	79.8568	6.0511	0.8162	0.3841	0.0938	8.8619
Real wage								
4Q	6.2726	14.9794	24.5772	0.2213	34.1127	1.1717	0.2982	18.3668
8Q	5.6839	19.8200	12.8581	0.1161	31.7275	0.5643	0.2978	28.9321
16Q	3.3235	20.6873	6.9996	0.1603	36.8344	0.3290	0.3967	31.2691
20Q	2.8272	19.7771	5.9603	0.1726	39.2705	0.2857	0.4580	31.2486
Inflation								
4Q	17.1586	11.2160	34.3253	1.1835	0.1755	1.1923	0.5994	34.1493
8Q	17.1334	9.3782	36.9222	1.1124	0.1509	1.0689	0.7772	33.4568
16Q	14.3407	7.9953	40.9412	0.9402	0.1589	0.9094	0.7557	33.9585
20Q	12.8011	7.2902	42.2714	0.8484	0.2053	0.8251	0.6802	35.0783

*Note:* Columns 2 – 9 correspond to the shocks: the monetary policy shock (MP), the price markup shock (PM), the wage markup shock (WM), the government spending shock (GS), the neutral technology shock (Ntech), the preference shock (Pref), the biased technology shock (Btech), and the depreciation shock (Dep).

TABLE 6. Forecast Error Variance Decomposition: Regime II

Horizon	MP	PM	WM	GS	Ntech	Pref	Btech	Dep
Output								
4Q	1.4623	6.8572	22.2957	10.9572	27.1732	3.5101	0.0921	27.6522
8Q	0.7325	6.6368	36.3663	5.2312	28.4007	1.3462	0.1392	21.1470
16Q	0.3285	4.5746	47.8282	2.8228	30.4437	0.6002	0.1629	13.2390
20Q	0.2611	3.9167	50.5532	2.4070	30.7917	0.4770	0.1696	11.4238
Investment								
4Q	2.6142	10.7376	16.6082	0.7130	6.7202	1.3870	0.7675	60.4524
8Q	1.3975	11.0425	24.8966	1.0335	9.5109	1.7845	1.0280	49.3065
16Q	0.9411	10.1760	32.3509	1.0888	13.2559	1.6891	1.6423	38.8559
20Q	0.8887	9.8499	33.8462	1.0379	14.3342	1.5957	1.9291	36.5185
Hours								
4Q	1.8550	8.0476	35.9182	16.7333	5.1919	5.0074	0.0373	27.2093
8Q	0.9744	8.0232	60.7354	9.4043	2.2750	2.2543	0.0611	16.2724
16Q	0.4847	5.0454	77.8436	5.8075	1.2310	1.1118	0.0482	8.4277
20Q	0.3918	4.1596	80.9619	5.1229	1.0112	0.9081	0.0406	7.4039
Real wage								
4Q	1.5771	21.9006	22.2722	0.1675	37.7754	2.4758	0.1153	13.7161
8Q	1.4263	28.9214	11.6296	0.0877	35.0659	1.1900	0.1150	21.5641
16Q	0.8150	29.4979	6.1863	0.1183	39.7808	0.6781	0.1496	22.7740
20Q	0.6917	28.1379	5.2561	0.1271	42.3184	0.5875	0.1724	22.7089
Inflation								
4Q	5.3155	20.2044	38.3260	1.1035	0.2395	3.1041	0.2856	31.4214
8Q	5.3827	17.1324	41.8078	1.0518	0.2089	2.8220	0.3755	31.2189
16Q	4.4593	14.4569	45.8850	0.8799	0.2176	2.3764	0.3614	31.3634
20Q	3.9610	13.1172	47.1437	0.7901	0.2798	2.1456	0.3237	32.2388



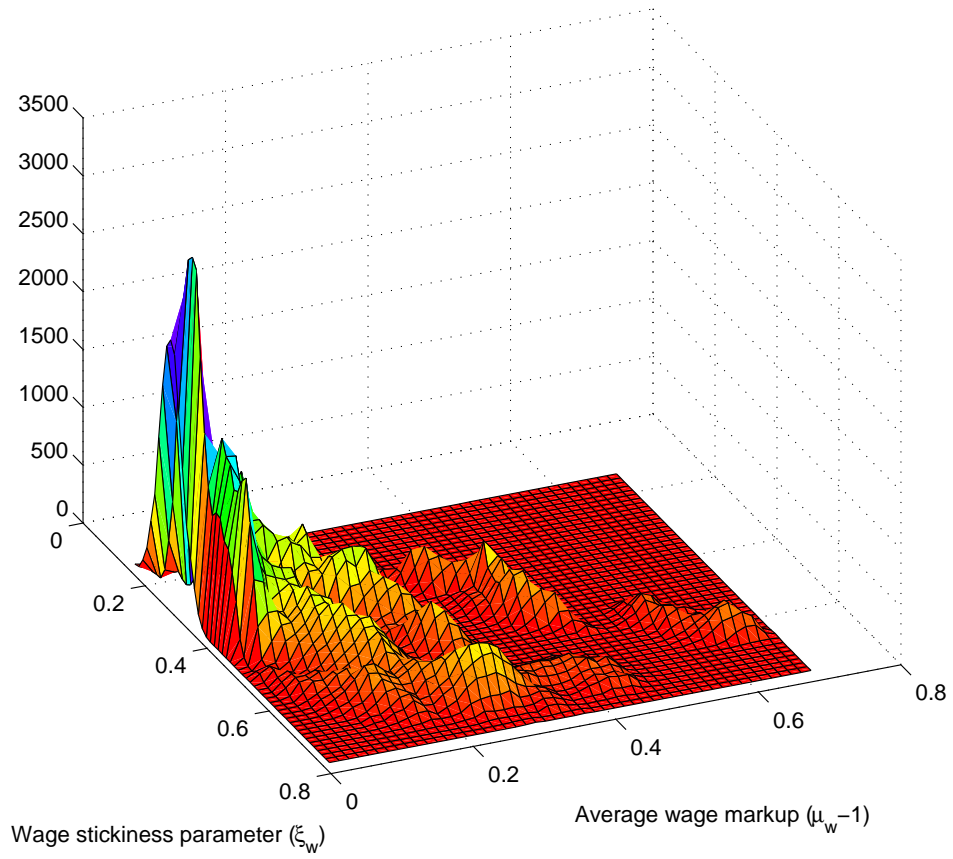


FIGURE 1. The histogram plot for the joint posterior distribution of the wage stickiness parameter and the average wage markup parameter for the model DSGE-2v.

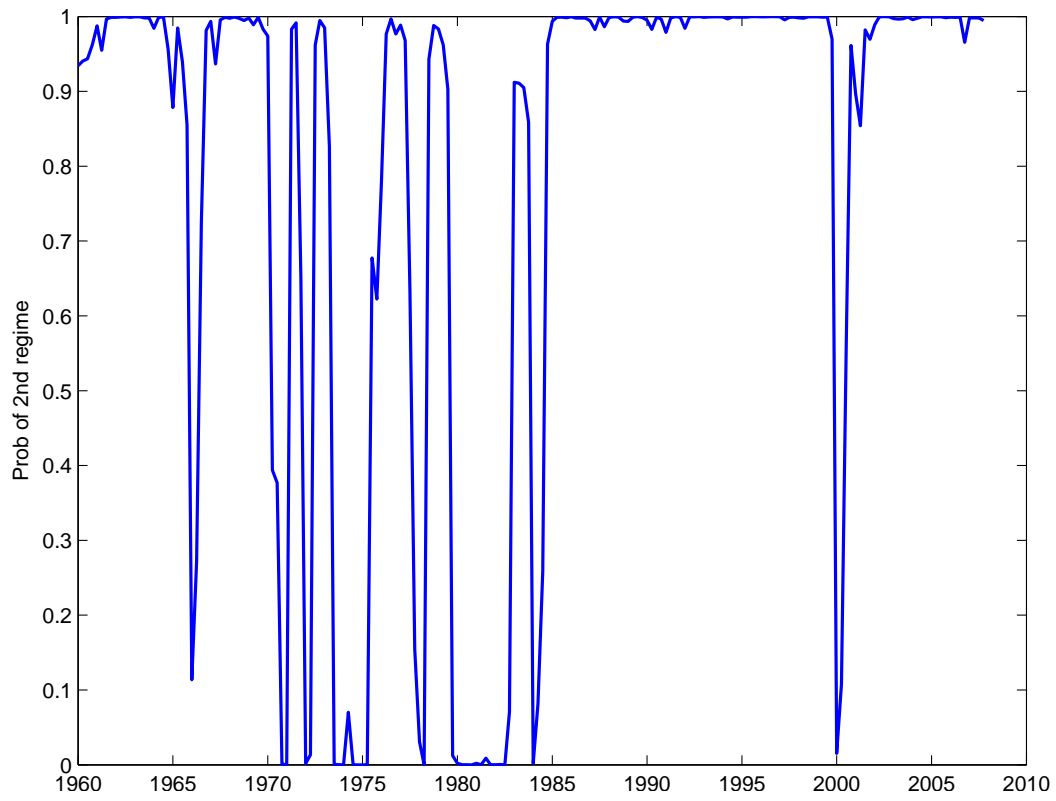


FIGURE 2. Posterior probabilities of the second regime for the model DSGE-2v.

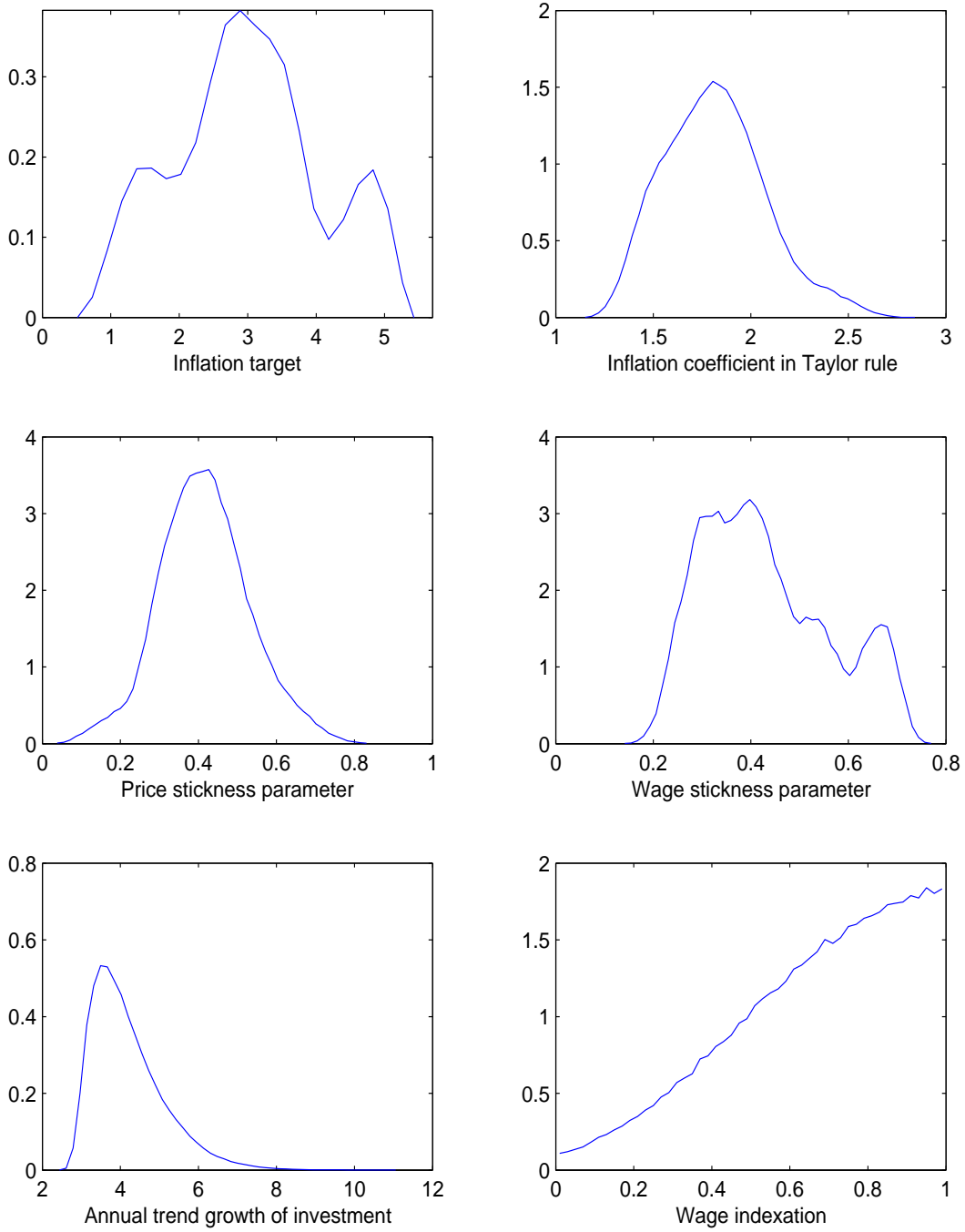


FIGURE 3. Marginal posterior distributions of some key parameters for the DSGE-2v model.

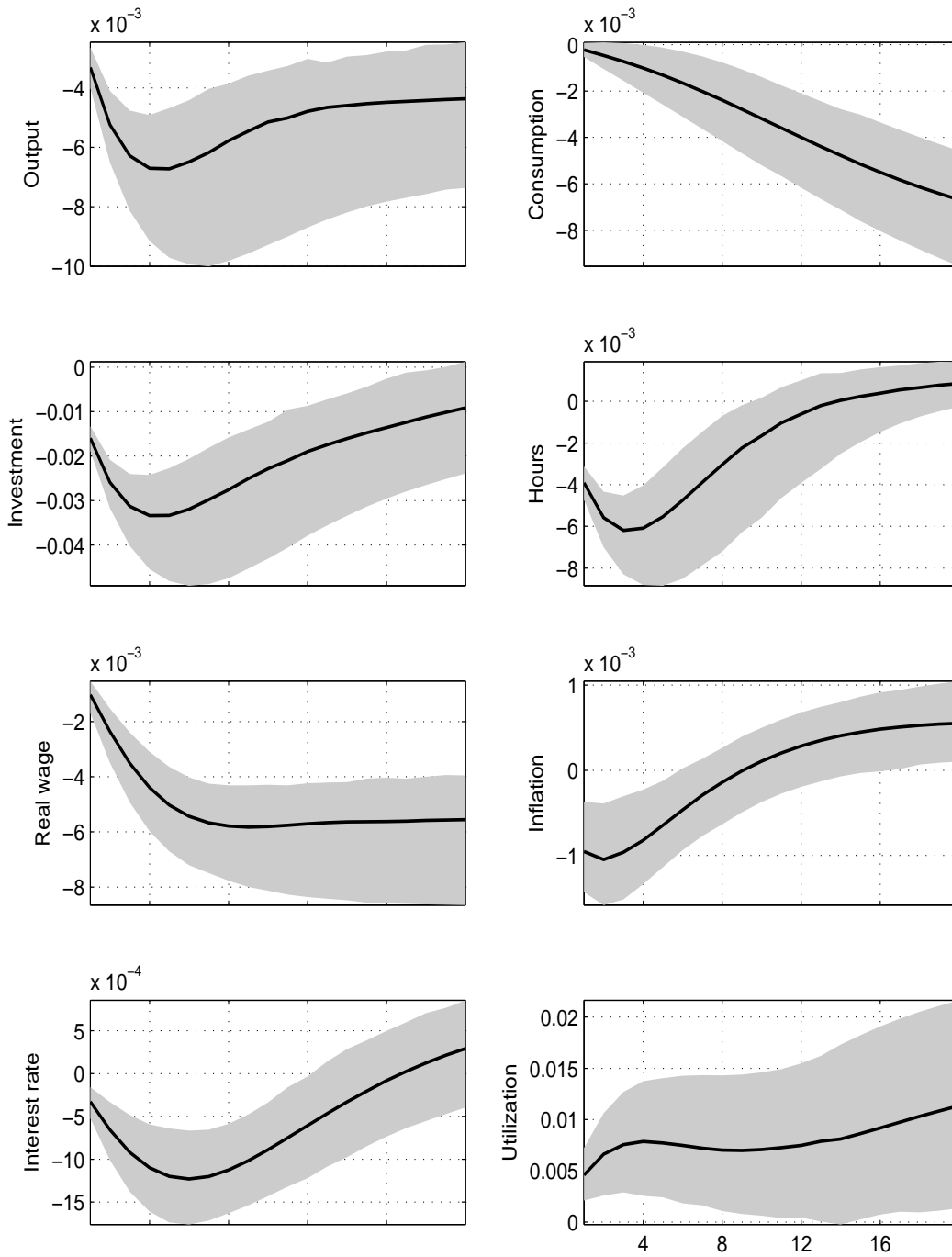


FIGURE 4. Impulse responses to a depreciation shock in the second regime. The shaded area represents .90 probability point-wise error bands and the thick line represents the median estimate.

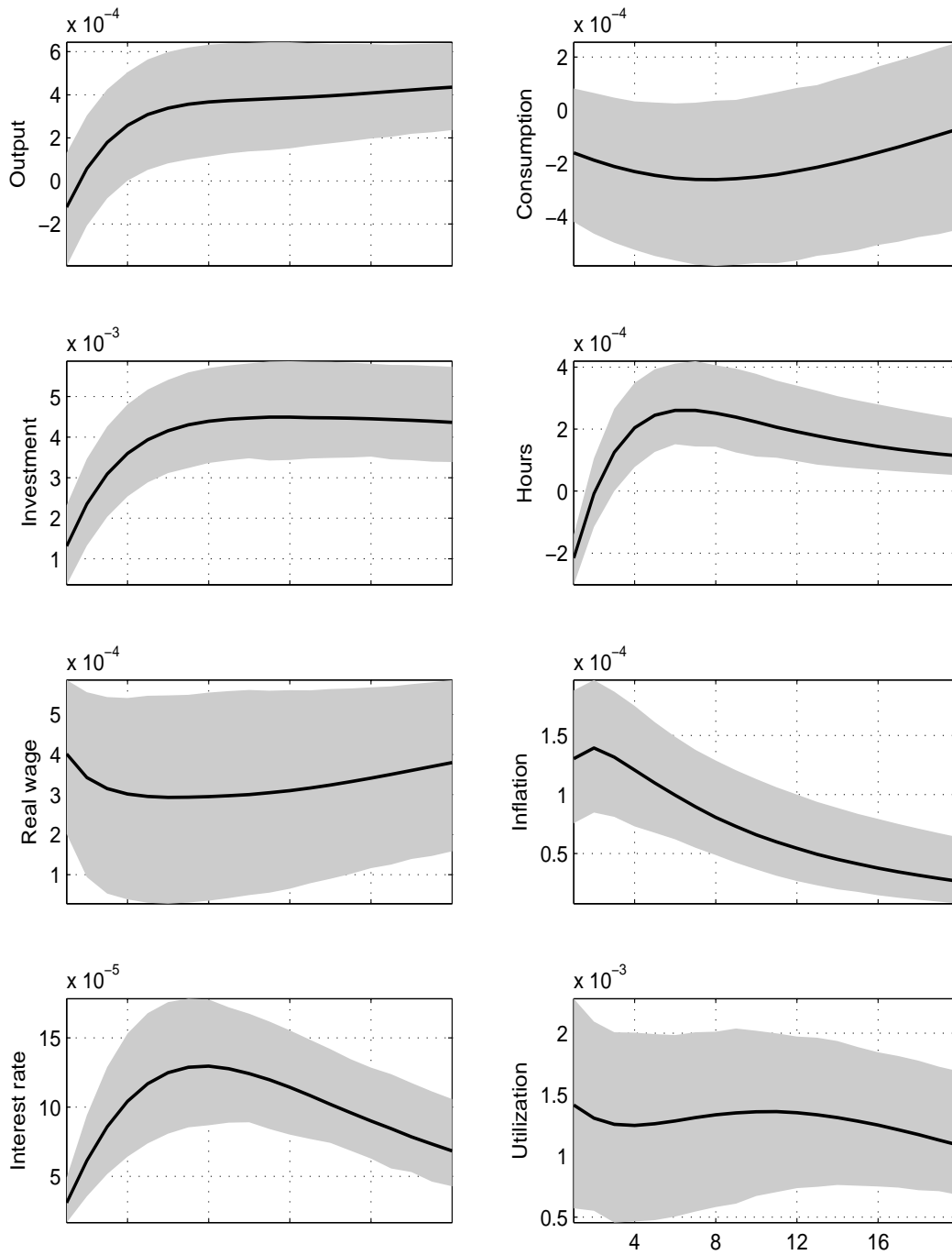


FIGURE 5. Impulse responses to an investment-specific technology shock in the second regime. The shaded area represents .90 probability point-wise error bands and the thick line represents the median estimate.

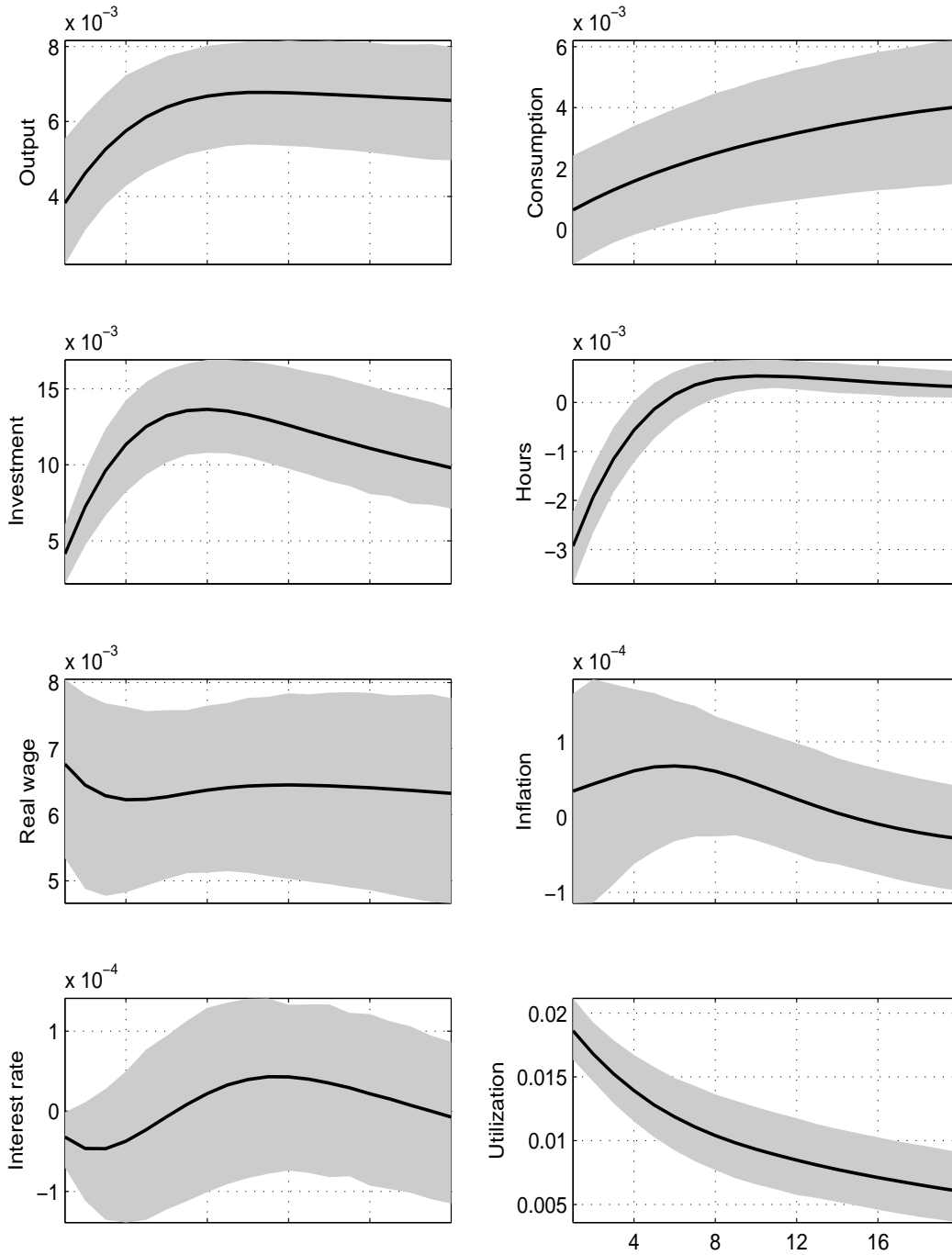


FIGURE 6. Impulse responses to a neutral technology shock in the second regime. The shaded area represents .90 probability point-wise error bands and the thick line represents the median estimate.

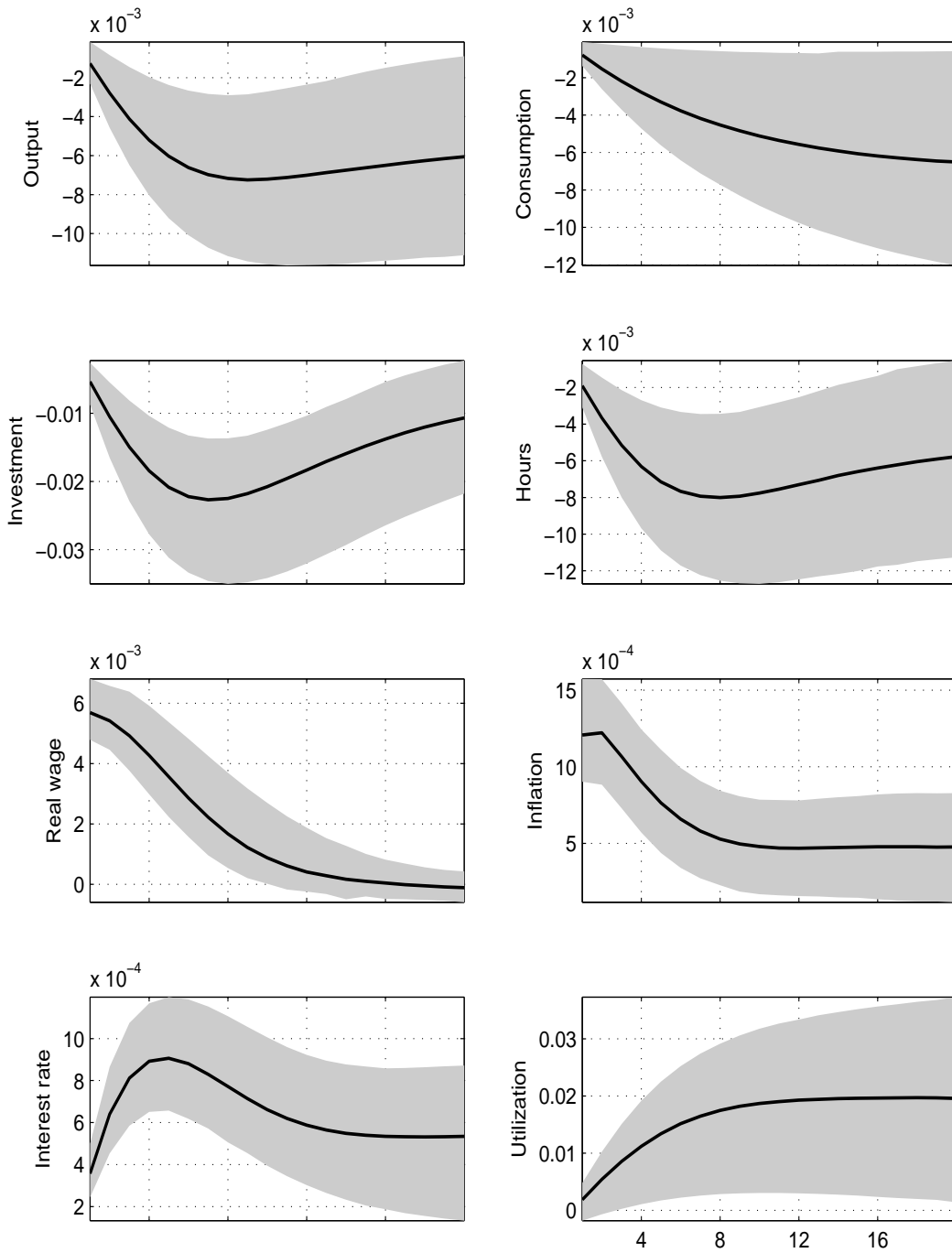


FIGURE 7. Impulse responses to a wage markup shock in the second regime. The shaded area represents .90 probability point-wise error bands and the thick line represents the median estimate.

## REFERENCES

- ALTIG, D., L. J. CHRISTIANO, M. EICHENBAUM, AND J. LINDE (2004): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," Federal Reserve Bank of Cleveland Working Paper 04-16.
- BASU, S., AND J. G. FERNALD (1997): "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 105(2), 249–283.
- (2002): "Aggregate Productivity and Aggregate Technology," *European Economic Review*, 46(6), 963–991.
- BIANCHI, F. (2008): "Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics," Manuscript, Princeton.
- BILS, M., AND P. J. KLENOW (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112(5), 947–985.
- BOIVIN, J., AND M. GIANNONI (2006): "Has Monetary Policy Become More Effective?," *Review of Economics and Statistics*, 88(3), 445–462.
- BOLDRIN, M., L. J. CHRISTIANO, AND J. D. FISHER (2001): "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review*, 91(1), 149–166.
- CALVO, G. (1983): "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12, 383–398.
- CANOVA, F., AND L. GAMBETTI (2004): "Bad luck or bad policy? On the time variations of US monetary policy," Manuscript, IGER and Università Bocconi.
- CHARI, V., P. J. KEHOE, AND E. R. MCGRATTAN (2000): "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?," *Econometrica*, 68(5), 1151–1180.
- CHRISTIANO, L., M. EICHENBAUM, AND C. EVANS (2005): "Nominal Rigidities and the Dynamics Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, 1–45.
- CHRISTIANO, L. J., M. EICHENBAUM, AND R. VIGFUSSON (2003): "What Happens After A Technology Shock?," NBER Working Paper 9819.
- CLARIDA, R., J. GALÍ, AND M. GERTLER (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, CXV, 147–180.
- COGLEY, T., AND T. J. SARGENT (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," *Review of Economic Dynamics*, 8, 262–302.



- CUMMINS, J. G., AND G. L. VIOLANTE (2002): "Investment-Specific Technical Change in the United States (1947-2000): Measurement and Macroeconomic Consequences," *Review of Economic Dynamics*, 5, 243–284.
- DEL NEGRO, M., AND F. SCHORFHEIDE (2008): "Forming Priors for DSGE Models (and How It Affects the Assessment of Nominal Rigidities)," Manuscript, Federal Reserve Bank of New York and University of Pennsylvania.
- DEL NEGRO, M., F. SCHORFHEIDE, F. SMETS, AND R. WOUTERS (2007): "On the Fit and Forecasting Performance of New Keynesian Models," *Journal of Business and Economic Statistics*, 25(2), 123–143.
- DENTON, F. T. (1971): "Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization," *Journal of the American Statistical Association*, 66, 99–102.
- ERCEG, C. J., AND A. T. LEVIN (2003): "Imperfect Credibility and Inflation Persistence," *Journal of Monetary Economics*, 50, 915–944.
- FARMER, R. E., D. F. WAGGONER, AND T. ZHA (2009): "Understanding Markov-Switching Rational Expectations Models," *Journal of Economic Theory*, 144, 1849–1867.
- FAVERO, C. A., AND R. ROVELLI (2003): "Macroeconomic Stability and the Preferences of the Fed: A Formal Analysis, 1961–98," *Journal of Money, Credit, and Banking*, 35, 545–556.
- FERNANDEZ-VILLAYERDE, J., AND J. RUBIO-RAMIREZ (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74, 1059–1087.
- FISHER, J. D. M. (2006): "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks," *Journal of Political Economy*, 114(3), 413–451.
- FRANCIS, N., AND V. A. RAMEY (2005): "Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited," *Journal of Monetary Economics*, 52, 1379–1399.
- GALÍ, J. (1999): "Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?," *American Economic Review*, 89(1), 249–271.
- GAMBETTI, L., E. PAPPÀ, AND F. CANOVA (2008): "The Structural Dynamics of US Output and Inflation: What Explains the Changes?," *Journal of Money, Credit, and Banking*, 40, 369–388.

- GELFAND, A. E., AND D. K. DEY (1994): "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society (Series B)*, 56, 501–514.
- GERTLER, M., AND N. KIYOTAKI (2010): "Financial Intermediation and Credit Policy in Business Cycle Analysis," Manuscript, New York University and Princeton University.
- GEWEKE, J. (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication," *Econometric Reviews*, 18(1), 1–73.
- GORDON, R. J. (1990): *The Measurement of Durable Goods Prices*. University of Chicago Press, Chicago, Illinois.
- GREENWOOD, J., Z. HERCOWITZ, AND P. KRUSELL (1997): "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, 87, 342–362.
- HALL, R. E. (1988): "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96, 921–947.
- HAMILTON, J. D. (1994): *Times Series Analysis*. Princeton University Press, Princeton, NJ.
- HUANG, K. X., AND Z. LIU (2002): "Staggered Price-Setting, Staggered Wage-Setting, and Business Cycle Persistence," *Journal of Monetary Economics*, 49(2), 405–433.
- HUANG, K. X., Z. LIU, AND L. PHANEUF (2004): "Why Does the Cyclical Behavior of Real Wages Change Over Time?," *American Economic Review*, 94(4), 836–856.
- IRELAND, P. N. (2005): "Changes in the Federal Reserve's Inflation Target: Causes and Consequences," Manuscript, Boston College.
- JUSTINIANO, A., AND G. E. PRIMICERI (2006): "The Time Varying Volatility of Macroeconomic Fluctuations," NBER Working Paper No. 12022.
- (2008): "The Time Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 98(3), 604–641.
- JUSTINIANO, A., G. E. PRIMICERI, AND A. TAMBALOTTI (2008): "Investment Shocks and Business Cycles," Manuscript, Northwestern University.
- LEVIN, A. T., A. ONATSKI, J. C. WILLIAMS, AND N. WILLIAMS (2006): "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models," in *NBER Macroeconomics Annual 2005.*, ed. by M. Gertler, and K. Rogoff, pp. 229–287. MIT Press, Cambridge, MA.
- LIU, Z., AND L. PHANEUF (2007): "Technology Shocks and Labor Market Dynamics: Some Evidence and Theory," *Journal of Monetary Economics*, 54(8), 2534–2553.

- LIU, Z., D. F. WAGGONER, AND T. ZHA (2009): "Asymmetric Expectation Effects of Regime Shifts in Monetary Policy," *Review of Economic Dynamics*, 12(2), 284–303.
- LUBIK, T. A., AND F. SCHORFHEIDE (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94(1), 190–219.
- MENG, X.-L., AND W. H. WONG (1996): "Simulating Ratios of Normalizing Constants via a Simple Identity: A Theoretical Exploration," *Statistica Sinica*, 6, 831–860.
- PENCAVEL, J. H. (1986): "The Labor Supply of Men: A Survey," in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter, and R. Layard, pp. 3–102. Elsevier, Amsterdam, North Holland.
- PRIMICERI, G. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.
- RAMACHANDRAN, K., V. URAZOV, D. F. WAGGONER, AND T. ZHA (2007): "EcoSystem: A Set of Cluster Computing Tools for a Class of Economic Applications," Manuscript, Computing College at Georgia Institute of Technology.
- ROTEMBERG, J. J., AND M. WOODFORD (1997): "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke, and J. J. Rotemberg, pp. 297–346. The MIT Press, Cambridge (MA).
- (1999): "The Cyclical Behavior of Prices and Costs," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1B. Elsevier, Amsterdam.
- RUPERT, P., R. ROGERSON, AND R. WRIGHT (2000): "Homework in labor economics: Household production and intertemporal substitution," *Journal of Monetary Economics*, 46(3), 557–579.
- SCHORFHEIDE, F. (2005): "Learning and Monetary Policy Shifts," *Review of Economic Dynamics*, 8(2), 392–419.
- SIMS, C. A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1), 1–20.
- SIMS, C. A., D. F. WAGGONER, AND T. ZHA (2008): "Methods for Inference in Large Multiple-Equation Markov-Switching Models," *Journal of Econometrics*, 146(2), 255–274.
- SIMS, C. A., AND T. ZHA (2006): "Were There Regime Switches in US Monetary Policy?," *American Economic Review*, 96, 54–81.
- SMETS, F., AND R. WOUTERS (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606.

- STOCK, J. H., AND M. W. WATSON (2003): "Has the Business Cycles Changed? Evidence and Explanations," in *Monetary Policy and Uncertainty: Adapting to a Changing Economy*, pp. 9–56. Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming.
- TAUCHEN, G. (1986): "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20, 177–181.
- TAYLOR, J. B. (1999): "Staggered Price and Wage Setting in Macroeconomics," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, pp. 1009–1050. Elsevier, Amsterdam, North Holland.
- UHLIG, H. (2004): "Do Technology Shocks Lead to a Fall in Total Hours Worked?," *Journal of the European Economic Association*, 2(2-3), 361–71.
- WOODFORD, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton, New Jersey.

FEDERAL RESERVE BANK OF SAN FRANCISCO, FEDERAL RESERVE BANK OF ATLANTA, FEDERAL RESERVE BANK OF ATLANTA AND EMORY UNIVERSITY

**TECHNICAL APPENDIX**  
**SOURCES OF MACROECONOMIC FLUCTUATIONS: A**  
**REGIME-SWITCHING DSGE APPROACH**  
*(NOT INTENDED FOR PUBLICATION)*

ZHENG LIU, DANIEL F. WAGGONER, AND TAO ZHA

In this appendix, we derive the optimizing decisions, describe the stationary equilibrium, and derive the log-linearized equilibrium conditions in the paper entitled “Sources of Macroeconomic Fluctuations: A Regime-Switching DSGE Approach” by Liu, Waggoner, and Zha.

I. THE OPTIMIZING DECISIONS

**I.1. Households’ optimizing decisions.** Each household chooses consumption, investment, new capital stock, capacity utilization, and next-period bond to solve the following utility maximizing problem:

$$\text{Max}_{\{C_t, I_t, K_t, u_t, B_{t+1}\}} \mathbb{E} \sum_{t=0}^{\infty} \beta^t A_t \left\{ \log(C_t - bC_{t-1}) - \frac{\psi}{1+\eta} L_{t+i}^d(h)^{1+\eta} \right\} \quad (1)$$

subject to

$$\bar{P}_t C_t + \frac{\bar{P}_t}{Q_t} (I_t + a(u_t) K_{t-1}) + \mathbb{E}_t D_{t,t+1} B_{t+1} \leq W_t(h) L_t^d(h) + \bar{P}_t r_{kt} u_t K_{t-1} + \Pi_t + B_t + T_t, \quad (2)$$

$$K_t = (1 - \delta_t) K_{t-1} + \left[ 1 - S \left( \frac{I_t}{I_{t-1}} \right) \right] I_t, \quad (3)$$

---

*Date:* February 25, 2010.

This appendix is *not* intended for publication.

Denote by  $\mu_t$  the Lagrangian multiplier for the budget constraint (2) and by  $\mu_{kt}$  the Lagrangian multiplier for the capital accumulation equation (3). The first order conditions for the utility-maximizing problem are given by

$$A_t U_{ct} = \mu_t \bar{P}_t, \quad (4)$$

$$D_{t,t+1} = \beta \frac{\mu_{t+1}}{\mu_t}, \quad (5)$$

$$\frac{\mu_t \bar{P}_t}{Q_t} = \mu_{kt} \{1 - S(\lambda_{It}) - S'(\lambda_{It})\lambda_{It}\} + \beta \mathbf{E}_t \mu_{k,t+1} S'(\lambda_{I,t+1})(\lambda_{I,t+1})^2 \quad (6)$$

$$\mu_{kt} = \beta \mathbf{E}_t \left[ \mu_{k,t+1} (1 - \delta_{t+1}) + \mu_{t+1} \bar{P}_{t+1} r_{k,t+1} u_{t+1} - \frac{\mu_{t+1} \bar{P}_{t+1}}{Q_{t+1}} a(u_{t+1}) \right], \quad (7)$$

$$r_{kt} = \frac{a'(u_t)}{Q_t}, \quad (8)$$

where  $\lambda_{It} \equiv I_t/I_{t-1}$ .

Let  $q_{kt} \equiv Q_t \frac{\mu_{kt}}{\mu_t \bar{P}_t}$  denote the shadow price of capital stock (in units of investment goods). Then, (4) and (6) imply that

$$\frac{1}{Q_t} = \frac{q_{kt}}{Q_t} \{1 - S(\lambda_{It}) - S'(\lambda_{It})\lambda_{It}\} + \beta \mathbf{E}_t \frac{q_{k,t+1}}{Q_{t+1}} \frac{A_{t+1} U_{c,t+1}}{A_t U_{ct}} S'(\lambda_{I,t+1})(\lambda_{I,t+1})^2. \quad (9)$$

Thus, in the absence of adjustment cost or in the steady-state equilibrium where  $S(\lambda_I) = S'(\lambda_I) = 0$ , we have  $q_{kt} = 1$ . One can interpret  $q_{kt}$  as Tobin's Q.

By eliminating the Lagrangian multipliers  $\mu_t$  and  $\mu_{kt}$ , the capital Euler equation (7) can be rewritten as

$$\frac{q_{kt}}{Q_t} = \beta \mathbf{E}_t \frac{A_{t+1} U_{c,t+1}}{A_t U_{ct}} \left[ (1 - \delta_{t+1}) \frac{q_{k,t+1}}{Q_{t+1}} + r_{k,t+1} u_{t+1} - \frac{a(u_{t+1})}{Q_{t+1}} \right]. \quad (10)$$

The cost of acquiring a marginal unit of capital is  $q_{kt}/Q_t$  today (in consumption unit). The benefit of having this extra unit of capital consists of the expected discounted future resale value and the rental value net of utilization cost.

By eliminating the Lagrangian multiplier  $\mu_t$ , the first-order condition with respect to bond holding can be written as

$$D_{t,t+1} = \beta \frac{A_{t+1} U_{c,t+1}}{A_t U_{ct}} \frac{\bar{P}_t}{\bar{P}_{t+1}}. \quad (11)$$

Denote by  $R_t = [\mathbf{E}_t D_{t,t+1}]^{-1}$  the interest rate for a one-period risk-free nominal bond. Then we have

$$\frac{1}{R_t} = \beta \mathbf{E}_t \left[ \frac{A_{t+1} U_{c,t+1}}{A_t U_{ct}} \frac{\bar{P}_t}{\bar{P}_{t+1}} \right]. \quad (12)$$

In each period  $t$ , a fraction  $\xi_w$  of households re-optimize their nominal wage setting decisions. Those households who can re-optimize wage setting chooses the nominal

wage  $W_t(h)$  to maximize

$$\mathbb{E}_t \sum_{i=0}^{\infty} \beta^i \xi_w^i A_{t+i} [\log(C_{t+i} - bC_{t+i-1}) - \frac{\psi}{1+\eta} L_{t+i}^d(h)^{1+\eta}] + \quad (13)$$

$$\mu_{t+i} [W_t(h) \chi_{t,t+i}^w L_{t+i}^d(h) + m_{t+i}], \quad (14)$$

where the labor demand schedule is given by

$$L_{t+i}^d(h) = \left( \frac{W_t(h) \chi_{t,t+i}^w}{\bar{W}_{t+i}} \right)^{-\theta_{wt}} L_{t+i}, \quad \theta_{wt} = \frac{\mu_{wt}}{\mu_{wt} - 1}, \quad (15)$$

the term  $m_t$  is given by

$$m_t = \bar{P}_t r_{kt} u_t K_{t-1} + \Pi_t + B_t + T_t - \bar{P}_t C_t - \frac{\bar{P}_t}{Q_t} (I_t + a(u_t) K_{t-1}) - \mathbb{E}_t D_{t,t+1} B_{t+1},$$

and the term  $\chi_{t,t+i}^w$  is given by

$$\chi_{t,t+i}^w \equiv \begin{cases} \prod_{k=1}^i \pi_{t+k-1}^{\gamma_w} \pi^{1-\gamma_w} \lambda_{t,t+i}^* & \text{if } i \geq 1 \\ 1 & \text{if } i = 0, \end{cases} \quad (16)$$

where  $\lambda_{t,t+i}^* \equiv \frac{\lambda_{t+i}^*}{\lambda_t^*}$ .

The first-order condition for the wage-setting problem is given by

$$\mathbb{E}_t \sum_{i=0}^{\infty} (\beta \xi_w)^i \left\{ -A_{t+i} \psi L_{t+i}^d(h)^\eta \frac{\partial L_{t+i}^d(h)}{\partial W_t(h)} + \mu_{t+i} (1 - \theta_{w,t+i}) \chi_{t,t+i}^w L_{t+i}^d(h) \right\} = 0, \quad (17)$$

where

$$\frac{\partial L_{t+i}^d(h)}{\partial W_t(h)} = -\theta_{w,t+i} \frac{L_{t+i}^d(h)}{W_t(h)} = -\frac{\mu_{w,t+i}}{\mu_{w,t+i} - 1} \frac{L_{t+i}^d(h)}{W_t(h)}.$$

Factoring out the common terms and rearranging, we obtain

$$\mathbb{E}_t \sum_{i=0}^{\infty} (\beta \xi_w)^i \frac{\mu_{t+i}}{\mu_t} L_{t+i}^d(h) \frac{1}{\mu_{w,t+i} - 1} \left\{ \mu_{w,t+i} \frac{\psi A_{t+i} L_{t+i}^d(h)^\eta}{\mu_{t+i}} - \chi_{t,t+i}^w W_t(h) \right\} = 0.$$

Let  $MRS_t(h) \equiv \frac{\psi A_t L_t^d(h)^\eta}{\mu_t}$  denote the marginal rate of substitution between leisure and income. Then, using (11), we can rewrite the first-order condition for wage setting as

$$\mathbb{E}_t \sum_{i=0}^{\infty} \xi_w^i D_{t,t+i} L_{t+i}^d(h) \frac{1}{\mu_{w,t+i} - 1} \left\{ \mu_{w,t+i} MRS_{t+i}(h) - \chi_{t,t+i}^w W_t(h) \right\} = 0. \quad (18)$$

**I.2. Firms' optimizing decisions.** Pricing decisions are staggered across firms. In each period, a fraction  $\xi_p$  of firms can re-optimize their pricing decisions and the other fraction  $1 - \xi_p$  of firms mechanically update their prices according to the rule

$$P_t(j) = \pi_{t-1}^{\gamma_p} \pi^{1-\gamma_p} P_{t-1}(j), \quad (19)$$

If a firm can re-optimize, it chooses  $P_t(j)$  to solve

$$\text{Max}_{P_t(j)} \quad \mathbb{E}_t \sum_{i=0}^{\infty} \xi_p^i D_{t,t+i} [P_t(j) \chi_{t,t+i}^p Y_{t+i}^d(j) - V_{t+i}(j)], \quad (20)$$

subject to

$$Y_{t+i}^d(j) = \left( \frac{P_t(j) \chi_{t,t+i}^p}{\bar{P}_{t+i}} \right)^{-\frac{\mu_{p,t+i}}{\mu_{p,t+i}-1}} Y_{t+i}, \quad (21)$$

where  $V_{t+i}(j)$  is the cost function and the term  $\chi_{t,t+i}^p$  comes from the price-updating rule (19) and is given by

$$\chi_{t,t+i}^p = \begin{cases} \prod_{k=1}^i \pi_{t+k-1}^{\gamma_p} \pi^{1-\gamma_p} & \text{if } i \geq 1 \\ 1 & \text{if } i = 0. \end{cases} \quad (22)$$

The first order condition for the profit-maximizing problem yields the optimal pricing rule

$$\mathbb{E}_t \sum_{i=0}^{\infty} \xi_p^i D_{t,t+i} Y_{t+i}^d(j) \frac{1}{\mu_{p,t+i} - 1} [\mu_{p,t+i} \Phi_{t+i}(j) - P_t(j) \chi_{t,t+i}^p] = 0, \quad (23)$$

where  $\Phi_{t+i}(j) = \partial V_{t+i}(j) / \partial Y_{t+i}^d(j)$  denotes the marginal cost function. In the absence of markup shocks,  $\mu_{pt}$  would be a constant and (23) implies that the optimal price is a markup over an average of the marginal costs for the periods in which the price will remain effective. Clearly, if  $\xi_p = 0$  for all  $t$ , that is, if prices are perfectly flexible, then the optimal price would be a markup over the contemporaneous marginal cost.

Cost-minimizing implies that the marginal cost function is given by

$$\Phi_t(j) = \left[ \tilde{\alpha} (\bar{P}_t r_{kt})^{\alpha_1} \left( \frac{\bar{W}_t}{Z_t} \right)^{\alpha_2} \right]^{\frac{1}{\alpha_1 + \alpha_2}} Y_t(j)^{\frac{1}{\alpha_1 + \alpha_2} - 1}, \quad (24)$$

where  $\tilde{\alpha} \equiv \alpha_1^{-\alpha_1} \alpha_2^{-\alpha_2}$  and  $r_{kt}$  denotes the real rental rate of capital input. The conditional factor demand functions are given by

$$\bar{W}_t = \Phi_t(j) \alpha_2 \frac{Y_t(j)}{L_t^f(j)}, \quad (25)$$

$$\bar{P}_t r_{kt} = \Phi_t(j) \alpha_1 \frac{Y_t(j)}{K_t^f(j)}. \quad (26)$$



It follows that

$$\frac{\bar{W}_t}{\bar{P}_t r_{kt}} = \frac{\alpha_2}{\alpha_1} \frac{K_t^f(j)}{L_t^f(j)}, \quad \forall j \in [0, 1]. \quad (27)$$

**I.3. Market clearing.** In equilibrium, markets for bond, composite labor, capital stock, and composite goods all clear. Bond market clearing implies that  $B_t = 0$  for all  $t$ . Labor market clearing implies that  $\int_0^1 L_t^f(j) dj = L_t$ . Capital market clearing implies that  $\int_0^1 K_t^f(j) dj = u_t K_{t-1}$ . Composite goods market clearing implies that

$$C_t + \frac{1}{Q_t} [I_t + a(u_t) K_{t-1}] + G_t = Y_t, \quad (28)$$

where aggregate output is related to aggregate primary factors through the aggregate production function

$$G_{pt} Y_t = (u_t K_{t-1})^{\alpha_1} (Z_t L_t)^{\alpha_2}, \quad (29)$$

with  $G_{pt} \equiv \int_0^1 \left( \frac{P_t(j)}{\bar{P}_t} \right)^{-\frac{\mu_{pt}}{\mu_{pt}-1} \frac{1}{\alpha_1 + \alpha_2}} dj$  measuring the price dispersion.

## II. STATIONARY EQUILIBRIUM CONDITIONS

Since both the neutral technology and the investment-specific technology are growing over time, we transform the appropriate variables to induce stationarity. In particular, we denote by  $\tilde{X}_t$  the stationary counterpart of the variable  $X_t$  and we make the following transformations:

$$\begin{aligned} \tilde{Y}_t &= \frac{Y_t}{\lambda_t^*}, & \tilde{C}_t &= \frac{C_t}{\lambda_t^*}, & \tilde{I}_t &= \frac{I_t}{Q_t \lambda_t^*}, & \tilde{G}_t &= \frac{G_t}{\lambda_t^*}, & \tilde{K}_t &= \frac{K_t}{Q_t \lambda_t^*}, \\ \tilde{w}_t &= \frac{\bar{W}_t}{\bar{P}_t \lambda_t^*}, & \tilde{r}_{kt} &= r_{kt} Q_t, & \tilde{U}_{ct} &= U_{ct} \lambda_t^*, \end{aligned}$$

where the underlying trend for output is given by

$$\lambda_t^* \equiv (Z_t^{\alpha_2} Q_t^{\alpha_1})^{\frac{1}{1-\alpha_1}}.$$

**II.1. Stationary pricing decisions.** In terms of the stationary variables, we can rewrite the optimal pricing decision (23) as

$$E_t \sum_{i=0}^{\infty} (\beta \xi_p)^i A_{t+i} \tilde{U}_{c,t+i} \tilde{Y}_{t+i}^d(j) \frac{1}{\mu_{p,t+i} - 1} [\mu_{p,t+i} \phi_{t+i}(j) - p_t^* Z_{t,t+i}^p] = 0. \quad (30)$$

In this equation,  $\tilde{Y}_{t+i}^d(j) = \frac{Y_{t+i}^d(j)}{\lambda_{t+i}^*}$  denotes the detrended output demand;  $p_t^* \equiv \frac{P_t(j)}{\bar{P}_t}$  denotes the relative price for optimizing firms, which does not have a  $j$  index since all

optimizing firms make identical pricing decisions in a symmetric equilibrium; the term  $Z_{t,t+i}^p$  is defined as

$$Z_{t,t+i}^p = \frac{\chi_{t,t+i}^p}{\prod_{k=1}^i \pi_{t+k}} \quad (31)$$

and finally, the term  $\phi_{t+i}(j) \equiv \frac{\Phi_{t+i}(j)}{P_{t+i}}$  denotes the real unit cost function, which is given by

$$\begin{aligned} \phi_{t+i}(j) &= \left[ \tilde{\alpha} \left( \frac{\tilde{r}_{k,t+i}}{Q_{t+i}} \right)^{\alpha_1} \left( \tilde{w}_{t+i} \frac{\lambda_{t+i}^*}{Z_{t+i}} \right)^{\alpha_2} \right]^{\frac{1}{\alpha_1 + \alpha_2}} Y_{t+i}^d(j)^{\frac{1}{\alpha_1 + \alpha_2} - 1} \\ &= [\tilde{\alpha} (\tilde{r}_{k,t+i})^{\alpha_1} (\tilde{w}_{t+i})^{\alpha_2}]^{\frac{1}{\alpha_1 + \alpha_2}} \tilde{Y}_{t+i}^d(j)^{\frac{1}{\alpha_1 + \alpha_2} - 1}. \end{aligned} \quad (32)$$

The demand schedule  $\tilde{Y}_{t+i}^d(j)$  for the optimizing firm  $j$  is related to the relative price and aggregate output through

$$\begin{aligned} \tilde{Y}_{t+i}^d(j) &= \left[ \frac{P_t(j) \chi_{t,t+i}^p}{\bar{P}_{t+i}} \right]^{-\theta_{p,t+i}} \tilde{Y}_{t+i} \\ &= \left[ p_t^* \frac{\bar{P}_t}{\bar{P}_{t+i}} \chi_{t,t+i}^p \right]^{-\theta_{p,t+i}} \tilde{Y}_{t+i} \\ &= [p_t^* Z_{t,t+i}^p]^{-\theta_{p,t+i}} \tilde{Y}_{t+i}. \end{aligned} \quad (33)$$

Combining (32) and (33), we have

$$\phi_{t+i}(j) = \tilde{\phi}_{t+i} [p_t^* Z_{t,t+i}^p]^{-\theta_{p,t+i} \bar{\alpha}} (\tilde{Y}_{t+i})^{\bar{\alpha}}, \quad (34)$$

where  $\bar{\alpha} \equiv \frac{1 - \alpha_1 - \alpha_2}{\alpha_1 + \alpha_2}$  and

$$\tilde{\phi}_{t+i} \equiv [\tilde{\alpha} (\tilde{r}_{k,t+i})^{\alpha_1} (\tilde{w}_{t+i})^{\alpha_2}]^{\frac{1}{\alpha_1 + \alpha_2}}. \quad (35)$$

Given these relations, we can rewrite the optimal pricing rule (30) in terms of stationary variables

$$E_t \sum_{i=0}^{\infty} (\beta \xi_p)^i \frac{A_{t+i} \tilde{U}_{c,t+i} \tilde{Y}_{t+i}^d(j)}{\mu_{p,t+i} - 1} [\mu_{p,t+i} \tilde{\phi}_{t+i} [p_t^* Z_{t,t+i}^p]^{-\theta_{p,t+i} \bar{\alpha}} (\tilde{Y}_{t+i})^{\bar{\alpha}} - p_t^* Z_{t,t+i}^p] = 0, \quad (36)$$

where  $\tilde{\phi}$  is defined in (35).

**II.2. Stationary wage setting decision.** Using (4) and (11), we can rewrite the optimal wage-setting decision (18) as

$$E_t \sum_{i=0}^{\infty} (\beta \xi_w)^i \frac{A_{t+i} U_{c,t+i}}{A_t U_{ct}} \frac{\bar{P}_t}{\bar{P}_{t+i}} L_{t+i}^d(h) \frac{1}{\mu_{w,t+i} - 1} [\mu_{w,t+i} \psi \frac{L_{t+i}^d(h)^\eta}{U_{c,t+i}} \bar{P}_{t+i} - W_t(h) \chi_{t,t+i}^w] = 0, \quad (37)$$

where the labor demand schedule  $L_{t+i}^d(h)$  is related to aggregate variables through

$$L_{t+i}^d(h) = \left[ \frac{W_t(h)\chi_{t,t+i}^w}{\bar{W}_{t+i}} \right]^{-\theta_{w,t+i}} L_{t+i} \quad (38)$$

$$= \left[ w_t^* \frac{\bar{W}_t}{\bar{W}_{t+i}} \chi_{t,t+i}^w \right]^{-\theta_{w,t+i}} L_{t+i} \quad (39)$$

$$= \left[ w_t^* \frac{\tilde{w}_t \bar{P}_t \lambda_t^*}{\tilde{w}_{t+i} \bar{P}_{t+i} \lambda_{t+i}^*} \chi_{t,t+i}^w \right]^{-\theta_{w,t+i}} L_{t+i} \quad (40)$$

$$= \left[ \frac{w_t^* \tilde{w}_t}{\tilde{w}_{t+i}} \frac{\chi_{t,t+i}^w}{\prod_{k=1}^i \pi_{t+k} \lambda_{t,t+i}^*} \right]^{-\theta_{w,t+i}} L_{t+i} \quad (41)$$

$$\equiv \left[ \frac{w_t^* \tilde{w}_t}{\tilde{w}_{t+i}} Z_{t,t+i}^w \right]^{-\theta_{w,t+i}} L_{t+i}, \quad (42)$$

with  $Z_{t,t+i}^w$  defined as

$$Z_{t,t+i}^w = \frac{\chi_{t,t+i}^w}{\prod_{k=1}^i \pi_{t+k} \lambda_{t,t+i}^*}. \quad (43)$$

Further, we can rewrite the individual optimal nominal wage  $W_t(h)$  as

$$W_t(h) = w_t^* \bar{W}_t = w_t^* \tilde{w}_t \bar{P}_t \lambda_t^*.$$

Given these relations, we can rewrite the wage setting rule (37) in terms of the stationary variables. With some cancelations, we obtain

$$\mathbb{E}_t \sum_{i=0}^{\infty} \prod_{k=1}^i (\beta \xi_w)^i \frac{A_{t+i} \tilde{U}_{c,t+i} L_{t+i}^d(h)}{\mu_{w,t+i} - 1} \left\{ \mu_{w,t+i} \psi \left[ \frac{w_t^* \tilde{w}_t}{\tilde{w}_{t+i}} Z_{t,t+i}^w \right]^{-\eta \theta_{w,t+i}} \frac{L_{t+i}^\eta}{\tilde{U}_{c,t+i}} - w_t^* \tilde{w}_t Z_{t,t+i}^w \right\} = 0. \quad (44)$$

**II.3. Other stationary equilibrium conditions.** We now rewrite the rest of the equilibrium conditions in terms of stationary variables.

First, the optimal investment decision equation (9) can be written as

$$1 = q_{kt} \{1 - S(\lambda_{It}) - S'(\lambda_{It}) \lambda_{It}\} + \beta \mathbb{E}_t q_{k,t+1} \frac{\lambda_t^* Q_t}{\lambda_{t+1}^* Q_{t+1}} \frac{A_{t+1} \tilde{U}_{c,t+1}}{A_t \tilde{U}_{ct}} S'(\lambda_{I,t+1}) (\lambda_{I,t+1})^2, \quad (45)$$

where

$$\lambda_{It} = \frac{I_t}{I_{t-1}} = \frac{\tilde{I}_t Q_t \lambda_t^*}{\tilde{I}_{t-1} Q_{t-1} \lambda_{t-1}^*}. \quad (46)$$

Second, the capital Euler equation (10) can be written as

$$q_{kt} = \beta \mathbb{E}_t \frac{A_{t+1} \tilde{U}_{c,t+1}}{A_t \tilde{U}_{ct}} \frac{\lambda_t^* Q_t}{\lambda_{t+1}^* Q_{t+1}} [(1 - \delta_{t+1}) q_{k,t+1} + \tilde{r}_{k,t+1} u_{t+1} - a(u_{t+1})]. \quad (47)$$

Third, the optimal capacity utilization decision (8) is equivalent to

$$\tilde{r}_{kt} = a'(u_t). \quad (48)$$

Fourth, the intertemporal bond Euler equation (12) can be written as

$$\frac{1}{R_t} = \beta E_t \left[ \frac{\lambda_t^*}{\lambda_{t+1}^*} \frac{A_{t+1} \tilde{U}_{c,t+1}}{A_t \tilde{U}_{ct}} \frac{1}{\pi_{t+1}} \right]. \quad (49)$$

Fifth, the law of motion for capital stock in (3) can be written as

$$\tilde{K}_t = (1 - \delta_t) \frac{\lambda_{t-1}^* Q_{t-1}}{\lambda_t^* Q_t} \tilde{K}_{t-1} + [1 - S(\lambda_{It})] \tilde{I}_t. \quad (50)$$

Sixth, the aggregate resource constraint is given by

$$\tilde{C}_t + \tilde{I}_t + \frac{\lambda_{t-1}^* Q_{t-1}}{\lambda_t^* Q_t} a(u_t) \tilde{K}_{t-1} + \tilde{G}_t = \tilde{Y}_t. \quad (51)$$

Seventh, the aggregate production function (29) can be written as

$$G_{pt} \tilde{Y}_t = \left[ \frac{\lambda_{t-1}^* Q_{t-1}}{\lambda_t^* Q_t} u_t \tilde{K}_{t-1} \right]^{\alpha_1} L_t^{\alpha_2}. \quad (52)$$

Eighth, firms' cost-minimizing implies that, in the stationary equilibrium, we have

$$\frac{\tilde{w}_t}{\tilde{r}_{kt}} = \frac{\alpha_2}{\alpha_1} \frac{\lambda_{t-1}^* Q_{t-1}}{\lambda_t^* Q_t} \frac{u_t \tilde{K}_{t-1}}{L_t}. \quad (53)$$

Finally, we rewrite the interest rate rule here for convenience of referencing:

$$R_t = \kappa R_{t-1}^{\rho_r} \left[ \left( \frac{\pi_t}{\pi^*(s_t)} \right)^{\phi_\pi} \tilde{Y}_t^{\phi_y} \right]^{1-\rho_r} e^{\sigma_{rt} \varepsilon_{rt}}. \quad (54)$$

### III. STEADY STATE

A deterministic steady state is an equilibrium in which all stochastic shocks are shut off. Our model contains a non-standard ‘‘shock’’: the Markov regime switching in monetary policy regime and the shock regime. In computing the steady-state equilibrium, we shut off all shocks, including the regime shocks. Since there is a mapping between any finite-state Markov switching process and a vector AR(1) process (Hamilton, 1994), shutting off the regime shocks in the steady state is equivalent to setting the innovations in the AR(1) process to its unconditional mean (which is zero). In such a steady state, all stationary variables are constant.

In the steady state,  $p^* = 1$  and  $Z^p = 1$ , so that the price setting rule (36) reduces to

$$\frac{1}{\mu_p} = [\tilde{\alpha} \tilde{r}_k^{\alpha_1} \tilde{w}^{\alpha_2}]^{\frac{1}{\alpha_1 + \alpha_2}} \tilde{Y}^{\tilde{\alpha}}. \quad (55)$$

That is, the real marginal cost is constant and equals the inverse markup.

Similarly, in the steady state,  $w^* = 1$  and  $Z^w = 1$ , so that the wage setting rule (44) reduces to

$$\tilde{w} = \mu_w \frac{\psi L^\eta}{\tilde{U}_c}, \quad (56)$$

which says that the real wage is a constant markup over the marginal rate of substitution between leisure and consumption.

Given that the steady-state markup, and thus the steady-state real marginal cost, is a constant, the conditional factor demand function (26) for capital input together with the capital market clearing condition imply that

$$\tilde{r}_k = \frac{\alpha_1 \tilde{Y} \lambda_q \lambda^*}{\mu_p \tilde{K}}. \quad (57)$$

The rest of the steady-state equilibrium conditions for the private sector come from (45) - (53) and are summarized below:

$$1 = q_k, \quad (58)$$

$$\frac{\lambda_q \lambda^*}{\beta} = 1 - \delta + \tilde{r}_k, \quad (59)$$

$$\tilde{r}_k = a'(1), \quad (60)$$

$$R = \frac{\lambda^*}{\beta} \pi, \quad (61)$$

$$\frac{\tilde{I}}{\tilde{K}} = 1 - \frac{1 - \delta}{\lambda_q \lambda^*}, \quad (62)$$

$$\tilde{Y} = \tilde{C} + \tilde{I} + \tilde{G}, \quad (63)$$

$$\tilde{Y} = \left( \frac{\tilde{K}}{\lambda_q \lambda^*} \right)^{\alpha_1} L^{\alpha_2}, \quad (64)$$

$$\frac{\tilde{w}}{\tilde{r}_k} = \frac{1}{\lambda_q \lambda^*} \frac{\alpha_2 \tilde{K}}{\alpha_1 L}. \quad (65)$$

#### IV. LINEARIZED EQUILIBRIUM CONDITIONS

We now describe our procedure to linearize the stationary equilibrium conditions around the deterministic steady state.

IV.1. **Linearizing the price setting rule.** Log-linearizing the price setting rule (36) around the steady state, we get

$$\begin{aligned} & \mathbb{E}_t \ln \sum_{i=0}^{\infty} (\beta \xi_p)^i \exp \left\{ \hat{a}_{t+i} + \hat{u}_{c,t+i} + \hat{y}_{t+i}^d(h) - \frac{\mu_p}{\mu_p - 1} \hat{\mu}_{p,t+i} + \hat{\mu}_{p,t+i} + \right. \\ & \qquad \qquad \qquad \left. \hat{\phi}_{t+i} - \theta_p \bar{\alpha} [\hat{p}_t^* + \hat{Z}_{t,t+i}^p] + \bar{\alpha} \hat{y}_{t+i} \right\} \\ & \approx \mathbb{E}_t \ln \sum_{i=0}^{\infty} (\beta \xi_p)^i \exp \left\{ \hat{a}_{t+i} + \hat{u}_{c,t+i} + \hat{y}_{t+i}^d(h) - \frac{\mu_p}{\mu_p - 1} \hat{\mu}_{p,t+i} + \hat{p}_t^* + \hat{Z}_{t,t+i}^p \right\}, \end{aligned}$$

where

$$\hat{\phi}_{t+i} = \frac{1}{\alpha_1 + \alpha_2} [\alpha_1 \hat{r}_{k,t+i} + \alpha_2 \hat{w}_{t+i}]. \quad (66)$$

Collecting terms to get

$$\begin{aligned} & \mathbb{E}_t \sum_{i=0}^{\infty} (\beta \xi_p)^i \left\{ \hat{\mu}_{p,t+i} + \hat{\phi}_{t+i} - \theta_p \bar{\alpha} [\hat{p}_t^* + \hat{Z}_{t,t+i}^p] + \bar{\alpha} \hat{y}_{t+i} \right\} \\ & \approx \mathbb{E}_t \sum_{i=0}^{\infty} (\beta \xi_p)^i \left\{ \hat{p}_t^* + \hat{Z}_{t,t+i}^p \right\}. \end{aligned}$$

Further simplifying

$$\frac{1 + \theta_p \bar{\alpha}}{1 - \beta \xi_p} \hat{p}_t^* = \mathbb{E}_t \sum_{i=0}^{\infty} (\beta \xi_p)^i \left\{ \hat{\mu}_{p,t+i} + \hat{\phi}_{t+i} + \bar{\alpha} \hat{y}_{t+i} - (1 + \theta_p \bar{\alpha}) \hat{Z}_{t,t+i}^p \right\}.$$

Denote  $\hat{m}c_{t+i} \equiv \hat{\phi}_{t+i} + \bar{\alpha} \hat{y}_{t+i}$ . Expanding the infinite sum in the above equation, we get

$$\begin{aligned} \frac{1 + \bar{\alpha} \theta_p}{1 - \beta \xi_p} \hat{p}_t^* &= \hat{\mu}_{pt} + \hat{m}c_t - (1 + \theta_p \bar{\alpha}) \hat{Z}_{t,t}^p \\ &+ \beta \xi_p \mathbb{E}_t [\hat{\mu}_{p,t+1} + \hat{m}c_{t+1} - (1 + \theta_p \bar{\alpha}) \hat{Z}_{t,t+1}^p] \\ &+ (\beta \xi_p)^2 \mathbb{E}_t [\hat{\mu}_{p,t+2} + \hat{m}c_{t+2} - (1 + \theta_p \bar{\alpha}) \hat{Z}_{t,t+2}^p] + \dots \end{aligned}$$

Forwarding this relation one period to get

$$\begin{aligned} \frac{1 + \bar{\alpha} \theta_p}{1 - \beta \xi_p} \hat{p}_{t+1}^* &= \hat{\mu}_{p,t+1} + \hat{m}c_{t+1} - (1 + \theta_p \bar{\alpha}) \hat{Z}_{t+1,t+1}^p \\ &+ \beta \xi_p \mathbb{E}_{t+1} [\hat{\mu}_{p,t+2} + \hat{m}c_{t+2} - (1 + \theta_p \bar{\alpha}) \hat{Z}_{t+1,t+2}^p] \\ &+ (\beta \xi_p)^2 \mathbb{E}_{t+1} [\hat{\mu}_{p,t+3} + \hat{m}c_{t+3} - (1 + \theta_p \bar{\alpha}) \hat{Z}_{t+1,t+3}^p] + \dots \end{aligned}$$

Moving the  $Z_{t,t+i}^p$  terms to the left, we have

$$\begin{aligned}
 & \frac{1 + \bar{\alpha}\theta_p}{1 - \beta\xi_p} \hat{p}_t^* + (1 + \bar{\alpha}\theta_p) \mathbb{E}_t[\hat{Z}_{t,t}^p + \beta\xi_p \hat{Z}_{t,t+1}^p + \dots] = \hat{\mu}_{pt} + \hat{m}c_t \\
 & \quad + \beta\xi_p \mathbb{E}_t[\hat{\mu}_{p,t+1} + \hat{m}c_{t+1}] \\
 & \quad + (\beta\xi_p)^2 \mathbb{E}_t[\hat{\mu}_{p,t+2} + \hat{m}c_{t+2}] + \dots \\
 & = \hat{\mu}_{pt} + \hat{m}c_t \\
 & \quad + \beta\xi_p \left[ \frac{1 + \bar{\alpha}\theta_p}{1 - \beta\xi_p} \mathbb{E}_t \hat{p}_{t+1}^* + (1 + \bar{\alpha}\theta_p) \mathbb{E}_t[\hat{Z}_{t+1,t+1}^p + \beta\xi_p \hat{Z}_{t+1,t+2}^p + \dots] \right],
 \end{aligned}$$

Since  $\hat{Z}_{t,t}^p = 0$ , we have

$$\begin{aligned}
 \frac{1 + \bar{\alpha}\theta_p}{1 - \beta\xi_p} \hat{p}_t^* & = \hat{\mu}_{pt} + \hat{m}c_t + \beta\xi_p \frac{1 + \bar{\alpha}\theta_p}{1 - \beta\xi_p} \mathbb{E}_t \hat{p}_{t+1}^* \\
 & \quad + (1 + \bar{\alpha}\theta_p) \beta\xi_p \mathbb{E}_t \sum_{i=0}^{\infty} (\beta\xi_p)^i [\hat{Z}_{t+1,t+i+1}^p - \hat{Z}_{t,t+i+1}^p]. \tag{67}
 \end{aligned}$$

Using the definition of  $Z_{t,t+i}^p$  in (31), we obtain

$$\begin{aligned}
 \hat{Z}_{t,t+i+1}^p & = -[\hat{\pi}_{t+i+1} - \gamma_p \hat{\pi}_{t+i} + \dots + \hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t] \\
 \hat{Z}_{t+1,t+i+1}^p & = -[\hat{\pi}_{t+i+1} - \gamma_p \hat{\pi}_{t+i} + \dots + \hat{\pi}_{t+2} - \gamma_p \hat{\pi}_{t+1}].
 \end{aligned}$$

Thus,

$$\hat{Z}_{t+1,t+i+1}^p - \hat{Z}_{t,t+i+1}^p = \hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t,$$

and the  $Z^p$  terms in (67) can be reduced to

$$\sum_{i=0}^{\infty} (\beta\xi_p)^i [\hat{Z}_{t+1,t+i+1}^p - \hat{Z}_{t,t+i+1}^p] = \frac{1}{1 - \beta\xi_p} [\hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t].$$

Substituting this result into (67), we obtain

$$\hat{p}_t^* = \frac{1 - \beta\xi_p}{1 + \bar{\alpha}\theta_p} (\hat{\mu}_{pt} + \hat{m}c_t) + \beta\xi_p \mathbb{E}_t \hat{p}_{t+1}^* + \beta\xi_p \mathbb{E}_t [\hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t]. \tag{68}$$

This completes log-linearizing the optimal price setting equation. We now log-linearize the price index relation. In an symmetric equilibrium, the price index relation is given by

$$1 = \xi_p \left[ \frac{1}{\pi_t} \pi_{t-1}^{\gamma_p} \pi^{1-\gamma_p} \right]^{\frac{1}{1-\mu_{pt}}} + (1 - \xi_p) (p_t^*)^{\frac{1}{1-\mu_{pt}}}, \tag{69}$$

the linearized version of which is given by

$$\hat{p}_t^* = \frac{\xi_p}{1 - \xi_p} (\hat{\pi}_t - \gamma_p \hat{\pi}_{t-1}). \tag{70}$$

Using (70) to substitute out the  $\hat{p}_t^*$  in (68), we obtain

$$\begin{aligned} & \frac{\xi_p}{1 - \xi_p} [\hat{\pi}_t - \gamma_p \hat{\pi}_{t-1}] \\ &= \frac{1 - \beta \xi_p}{1 + \bar{\alpha} \theta_p} (\hat{\mu}_{pt} + \hat{m}c_t) \\ &+ \beta \xi_p \frac{\xi_p}{1 - \xi_p} \text{E}_t [\hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t] + \beta \xi_p \text{E}_t [\hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t], \end{aligned}$$

or

$$\hat{\pi}_t - \gamma_p \hat{\pi}_{t-1} = \frac{\kappa_p}{1 + \bar{\alpha} \theta_p} (\hat{\mu}_{pt} + \hat{m}c_t) + \beta \text{E}_t [\hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t], \quad (71)$$

where the real marginal cost is given by

$$\hat{m}c_t = \frac{1}{\alpha_1 + \alpha_2} [\alpha_1 \hat{r}_{k,t+i} + \alpha_2 \hat{w}_{t+i}] + \bar{\alpha} \hat{y}_t. \quad (72)$$

and the term  $\kappa_p$  is given by

$$\kappa_p \equiv \frac{(1 - \beta \xi_p)(1 - \xi_p)}{\xi_p}$$

This completes the derivation of the price Phillips curve.

**IV.2. Linearizing the optimal wage setting rule.** Log-linearizing this wage decision rule, we get

$$\begin{aligned} & \text{E}_t \ln \sum_{i=0}^{\infty} (\beta \xi_w)^i \exp \left\{ \hat{a}_{t+i} + \hat{u}_{c,t+i} + \hat{l}_{t+i}^d(h) - \frac{\mu_w}{\mu_w - 1} \hat{\mu}_{w,t+i} + \hat{\mu}_{w,t+i} - \right. \\ & \quad \left. \eta \theta_w [\hat{w}_t^* + \hat{w}_t - \hat{w}_{t+i} + \hat{Z}_{t,t+i}^w] + \eta \hat{l}_{t+i} - \hat{u}_{c,t+i} \right\} \\ & \approx \text{E}_t \ln \sum_{i=0}^{\infty} (\beta \xi_w)^i \exp \left\{ \hat{a}_{t+i} + \hat{u}_{c,t+i} + \hat{l}_{t+i}^d(h) - \frac{\mu_w}{\mu_w - 1} \hat{\mu}_{w,t+i} + \hat{w}_t^* + \hat{w}_t + \hat{Z}_{t,t+i}^w \right\}. \end{aligned}$$

Collecting terms to get

$$\begin{aligned} & \text{E}_t \sum_{i=0}^{\infty} (\beta \xi_w)^i \left\{ \hat{\mu}_{w,t+i} - \eta \theta_w [\hat{w}_t^* + \hat{w}_t - \hat{w}_{t+i} + \hat{Z}_{t,t+i}^w] + \eta \hat{l}_{t+i} - \hat{u}_{c,t+i} \right\} \\ & \approx \text{E}_t \sum_{i=0}^{\infty} (\beta \xi_w)^i \left\{ \hat{w}_t^* + \hat{w}_t + \hat{Z}_{t,t+i}^w \right\}. \end{aligned}$$

Further simplifying

$$\frac{1 + \eta \theta_w}{1 - \beta \xi_w} (\hat{w}_t^* + \hat{w}_t) = \text{E}_t \sum_{i=0}^{\infty} (\beta \xi_w)^i \left\{ \hat{\mu}_{w,t+i} + \eta \hat{l}_{t+i} - \hat{u}_{c,t+i} + \eta \theta_w \hat{w}_{t+i} - (1 + \eta \theta_w) \hat{Z}_{t,t+i}^w \right\}.$$



Denote  $m\hat{r}s_{t+i} \equiv \eta\hat{l}_{t+i} - \hat{u}_{c,t+i}$ . Expanding the infinite sum in the above equation, we get

$$\begin{aligned} \frac{1 + \eta\theta_w}{1 - \beta\xi_w}(\hat{w}_t^* + \hat{w}_t) &= \hat{\mu}_{wt} + m\hat{r}s_t - \hat{w}_t + (1 + \eta\theta_w)(\hat{w}_t - \hat{Z}_{t,t}^w) \\ &+ \beta\xi_w \mathbf{E}_t[\hat{\mu}_{w,t+1} + m\hat{r}s_{t+1} - \hat{w}_{t+1} + (1 + \eta\theta_w)(\hat{w}_{t+1} - \hat{Z}_{t,t+1}^w)] \\ &+ (\beta\xi_w)^2 \mathbf{E}_t[\hat{\mu}_{w,t+2} + m\hat{r}s_{t+2} - \hat{w}_{t+2} + (1 + \eta\theta_w)(\hat{w}_{t+2} - \hat{Z}_{t,t+2}^w)] + \dots \end{aligned}$$

Forwarding this relation one period to get

$$\begin{aligned} \frac{1 + \eta\theta_w}{1 - \beta\xi_w}(\hat{w}_{t+1}^* + \hat{w}_{t+1}) &= \hat{\mu}_{w,t+1} + m\hat{r}s_{t+1} - \hat{w}_{t+1} + (1 + \eta\theta_w)(\hat{w}_{t+1} - \hat{Z}_{t+1,t+1}^w) \\ &+ \beta\xi_w \mathbf{E}_{t+1}[\hat{\mu}_{w,t+2} + m\hat{r}s_{t+2} - \hat{w}_{t+2} + (1 + \eta\theta_w)(\hat{w}_{t+2} - \hat{Z}_{t+1,t+2}^w)] \\ &+ (\beta\xi_w)^2 \mathbf{E}_{t+1}[\hat{\mu}_{w,t+3} + m\hat{r}s_{t+3} - \hat{w}_{t+3} + (1 + \eta\theta_w)(\hat{w}_{t+3} - \hat{Z}_{t+1,t+3}^w)] + \dots \end{aligned}$$

Moving the  $Z_{t,t+i}^w$  terms to the left, we have

$$\begin{aligned} \frac{1 + \eta\theta_w}{1 - \beta\xi_w}(\hat{w}_t^* + \hat{w}_t) + (1 + \eta\theta_w)\mathbf{E}_t[\hat{Z}_{t,t}^w + \beta\xi_w\hat{Z}_{t,t+1}^w + \dots] &= \hat{\mu}_{wt} + m\hat{r}s_t - \hat{w}_t + (1 + \eta\theta_w)\hat{w}_t \\ &+ \beta\xi_w \mathbf{E}_t[\hat{\mu}_{w,t+1} + m\hat{r}s_{t+1} - \hat{w}_{t+1} + (1 + \eta\theta_w)\hat{w}_{t+1}] \\ &+ (\beta\xi_w)^2 \mathbf{E}_t[\hat{\mu}_{w,t+2} + m\hat{r}s_{t+2} - \hat{w}_{t+2} + (1 + \eta\theta_w)\hat{w}_{t+2}] + \dots \\ &= \hat{\mu}_{wt} + m\hat{r}s_t - \hat{w}_t + (1 + \eta\theta_w)\hat{w}_t \\ &+ \beta\xi_w \mathbf{E}_t \left[ \frac{1 + \eta\theta_w}{1 - \beta\xi_w}(\hat{w}_{t+1}^* + \hat{w}_{t+1}) + (1 + \eta\theta_w)[\hat{Z}_{t+1,t+1}^w + \beta\xi_w\hat{Z}_{t+1,t+2}^w + \dots] \right], \end{aligned}$$

Since  $\hat{Z}_{t,t}^w = 0$ , we have

$$\begin{aligned} \frac{1 + \eta\theta_w}{1 - \beta\xi_w}(\hat{w}_t^* + \hat{w}_t) &= \hat{\mu}_{wt} + m\hat{r}s_t - \hat{w}_t + (1 + \eta\theta_w)\hat{w}_t + \beta\xi_w \frac{1 + \eta\theta_w}{1 - \beta\xi_w} \mathbf{E}_t(\hat{w}_{t+1}^* + \hat{w}_{t+1}) \\ &+ (1 + \eta\theta_w)\beta\xi_w \mathbf{E}_t \sum_{i=0}^{\infty} (\beta\xi_w)^i [\hat{Z}_{t+1,t+i+1}^w - \hat{Z}_{t,t+i+1}^w]. \end{aligned} \quad (73)$$

Using the definition of  $Z_{t,t+i}^w$  in (43), we obtain

$$\begin{aligned} \hat{Z}_{t,t+i+1}^w &= -[\hat{\pi}_{t+i+1} - \gamma_w \hat{\pi}_{t+i} + \dots + \hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t] \\ \hat{Z}_{t+1,t+i+1}^w &= -[\hat{\pi}_{t+i+1} - \gamma_w \hat{\pi}_{t+i} + \dots + \hat{\pi}_{t+2} - \gamma_w \hat{\pi}_{t+1}]. \end{aligned}$$

Thus,

$$\hat{Z}_{t+1,t+i+1}^w - \hat{Z}_{t,t+i+1}^w = \hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t,$$

and the  $Z^w$  terms in (73) can be reduced to

$$\sum_{i=0}^{\infty} (\beta\xi_w)^i [\hat{Z}_{t+1,t+i+1}^w - \hat{Z}_{t,t+i+1}^w] = \frac{1}{1 - \beta\xi_w} [\hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t].$$

Substituting this result into (73), we obtain

$$\hat{w}_t^* + \hat{w}_t = \frac{1 - \beta\xi_w}{1 + \eta\theta_w}(\hat{\mu}_{wt} + m\hat{r}s_t - \hat{w}_t) + (1 - \beta\xi_w)\hat{w}_t + \beta\xi_w E_t(\hat{w}_{t+1}^* + \hat{w}_{t+1}) + \beta\xi_w E_t[\hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t]. \quad (74)$$

This completes log-linearizing the wage decision equation. We now log-linearize the wage index relation. In an symmetric equilibrium, the wage index relation is given by

$$1 = \xi_w \left[ \frac{\tilde{w}_{t-1}}{\tilde{w}_t} \frac{1}{\pi_t} \pi_t^{\gamma_w} \pi_t^{1-\gamma_w} \right]^{\frac{1}{1-\mu_{wt}}} + (1 - \xi_w)(w_t^*)^{\frac{1}{1-\mu_{wt}}}, \quad (75)$$

the linearized version of which is given by

$$\hat{w}_t^* = \frac{\xi_w}{1 - \xi_w}(\hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t - \gamma_w \hat{\pi}_{t-1}). \quad (76)$$

Using (76) to substitute out the  $\hat{w}_t^*$  in (74), we obtain

$$\begin{aligned} \hat{w}_t + \frac{\xi_w}{1 - \xi_w}[\hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t - \gamma_w \hat{\pi}_{t-1}] \\ = \frac{1 - \beta\xi_w}{1 + \eta\theta_w}(\hat{\mu}_{wt} + m\hat{r}s_t - \hat{w}_t) + (1 - \beta\xi_w)\hat{w}_t \\ + \beta\xi_w E_t \left\{ \hat{w}_{t+1} + \frac{\xi_w}{1 - \xi_w}[\hat{w}_{t+1} - \hat{w}_t + \hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t] \right\} + \beta\xi_w E_t[\hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t], \end{aligned}$$

or

$$\begin{aligned} \hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t - \gamma_w \hat{\pi}_{t-1} = \frac{\kappa_w}{1 + \eta\theta_w}(\hat{\mu}_{wt} + m\hat{r}s_t - \hat{w}_t) + \\ \beta E_t[\hat{w}_{t+1} - \hat{w}_t + \hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t], \end{aligned} \quad (77)$$

where  $\kappa_w \equiv \frac{(1 - \beta\xi_w)(1 - \xi_w)}{\xi_w}$ .

To help understand the economics behind this equation, we define the nominal wage inflation as

$$\pi_t^w = \frac{\bar{W}_t}{\bar{W}_{t-1}} = \frac{\tilde{w}_t \bar{P}_t \lambda_t^*}{\tilde{w}_{t-1} \bar{P}_{t-1} \lambda_{t-1}^*} = \frac{\tilde{w}_t}{\tilde{w}_{t-1}} \pi_t \lambda_{t-1,t}^*. \quad (78)$$

The log-linearized version is given by

$$\hat{\pi}_t^w = \hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t + \Delta \hat{\lambda}_t^*,$$

where  $\Delta x_t = x_t - x_{t-1}$  is the first-difference operator and  $\hat{\lambda}_t^* = \frac{1}{1 - \alpha_1}(\alpha_1 \hat{q}_t + \alpha_2 \hat{z}_t)$ . Thus, the optimal wage decision (77) is equivalent to

$$\begin{aligned} \hat{\pi}_t^w - \gamma_w \hat{\pi}_{t-1} = \frac{\kappa_w}{1 + \eta\theta_w}(\hat{\mu}_{wt} + m\hat{r}s_t - \hat{w}_t) + \beta E_t(\hat{\pi}_{t+1}^w - \gamma_w \hat{\pi}_t) \\ + \frac{1}{1 - \alpha_1}[\alpha_1(\Delta \hat{z}_t - \beta E_t \Delta \hat{z}_{t+1}) + \alpha_2(\Delta \hat{q}_t - \beta E_t \Delta \hat{q}_{t+1})]. \end{aligned} \quad (79)$$

This nominal-wage Phillips curve relation parallels that of the price-Phillips curve and has similar interpretations.

**IV.3. Linearizing other stationary equilibrium conditions.** Taking total differentiation in the investment decision equation (45) and using the steady-state conditions that  $S(\lambda_I) = S'(\lambda_I) = 0$ , we obtain

$$\hat{q}_{kt} = S''(\lambda_I)\lambda_I^2 \left[ \hat{\lambda}_{It} - \beta \mathbb{E}_t \hat{\lambda}_{I,t+1} \right], \quad (80)$$

which, combined with the definition of the investment growth rate

$$\hat{\lambda}_{It} = \Delta \hat{i}_t + \frac{1}{1 - \alpha_1} [\Delta \hat{q}_t + \alpha_2 \Delta \hat{z}_t], \quad (81)$$

implies the linearized investment decision equation in the text.

Taking total differentiation in the capital Euler equation (47) and using the steady-state conditions that  $\tilde{q}_k = 1$ ,  $u = 1$ ,  $a(1) = 0$ ,  $\tilde{r}_k = a'(1)$ , and  $\frac{\beta}{\lambda_I}(1 - \delta + \tilde{r}_k) = 1$ , we obtain

$$\hat{q}_{kt} = \mathbb{E}_t \left\{ \Delta \hat{a}_{t+1} + \Delta \hat{U}_{c,t+1} - \Delta \hat{\lambda}_{t+1}^* - \Delta \hat{q}_{t+1} + \frac{\beta}{\lambda_I} \left[ (1 - \delta) \hat{q}_{k,t+1} - \delta \hat{\delta}_{t+1} + \tilde{r}_k \hat{r}_{k,t+1} \right] \right\}, \quad (82)$$

which, upon substituting the expressions for the  $\Delta \hat{\lambda}_t^*$  and  $\Delta \hat{q}_t$ , implies the linearized capital Euler equation in the text.

The linearized capacity utilization decision equation (48) is given by

$$\hat{r}_{kt} = \sigma_u \hat{u}_t, \quad (83)$$

where  $\sigma_u \equiv \frac{a''(1)}{a'(1)}$  is the curvature parameter for the capacity utility function  $a(u)$  evaluated at the steady state.

The linearized intertemporal bond Euler equation (49) is given by

$$0 = \mathbb{E}_t \left[ \Delta \hat{a}_{t+1} + \Delta \hat{U}_{c,t+1} - \Delta \hat{\lambda}_{t+1}^* + \hat{R}_t - \hat{\pi}_{t+1} \right], \quad (84)$$

which, along with the definition of the exogenous term  $\Delta \hat{\lambda}_{t+1}^*$ , implies the linearized bond Euler equation in the text.

Log-linearize the capital law of motion (50) leads to

$$\hat{k}_t = \frac{1 - \delta}{\lambda_I} [\hat{k}_{t-1} - \Delta \hat{\lambda}_t^* - \Delta \hat{q}_t] - \frac{\delta}{\lambda_I} \hat{\delta}_t + \frac{\tilde{I}}{\tilde{K}} \hat{i}_t, \quad (85)$$

which implies the linearized capital law of motion in the text.

To obtain the linearized resource constraint, we take total differentiation of (51) to obtain

$$\hat{y}_t = c_y \hat{c}_t + i_y \hat{i}_t + u_y \hat{u}_t + g_y \hat{g}_t, \quad (86)$$

where  $c_y = \frac{\tilde{C}}{\tilde{Y}}$ ,  $i_y = \frac{\tilde{I}}{\tilde{Y}}$ ,  $u_y = \frac{\tilde{r}_k \tilde{K}}{\tilde{Y} \lambda_I}$ , and  $g_y = \frac{\tilde{G}}{\tilde{Y}}$ .

Log-linearizing the aggregate production function (52), we get

$$\begin{aligned} \hat{y}_t &= \alpha_1 [\hat{k}_{t-1} + \hat{u}_t - \Delta \hat{\lambda}_t^* - \Delta \hat{q}_t] + \alpha_2 \hat{l}_t \\ &= \alpha_1 \left[ \hat{k}_{t-1} + \hat{u}_t - \frac{1}{1 - \alpha_1} (\alpha_2 \Delta \hat{z}_t + \Delta \hat{q}_t) \right] + \alpha_2 \hat{l}_t. \end{aligned} \quad (87)$$

The linearized version of the factor demand relation (53) is given by

$$\begin{aligned} \hat{w}_t &= \hat{r}_{kt} + \hat{k}_{t-1} + \hat{u}_t - \Delta \hat{\lambda}_t^* - \Delta \hat{q}_t - \hat{l}_t \\ &= \hat{r}_{kt} + \hat{k}_{t-1} + \hat{u}_t - \frac{1}{1 - \alpha_1} (\alpha_2 \Delta \hat{z}_t + \Delta \hat{q}_t) - \hat{l}_t \end{aligned} \quad (88)$$

Finally, linearizing the interest rate rule (54) gives

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1 - \rho_r) [\phi_\pi (\hat{\pi}_t - \hat{\pi}^*(s_t)) + \phi_y \hat{y}_t] + \sigma_{rt} \varepsilon_{rt}, \quad (89)$$

where

$$\hat{\pi}^*(s_t) \equiv \log \pi^*(s_t) - \log \pi.$$

Note that, with regime-switching inflation target, we have

$$\hat{\pi}^*(s_t) = \mathbf{1}\{s_t = 1\} \hat{\pi}^*(1) + \mathbf{1}\{s_t = 2\} \hat{\pi}^*(2) = [\hat{\pi}^*(1), \hat{\pi}^*(2)] e_{s_t},$$

where

$$e_{s_t} = \begin{bmatrix} \mathbf{1}\{s_t = 1\} \\ \mathbf{1}\{s_t = 2\} \end{bmatrix}.$$

It is useful to use the result that the random vector  $e_{s_t}$  follows an AR(1) process:

$$e_{s_t} = Q e_{s_{t-1}} + v_t,$$

where  $Q$  is the Markov transition matrix of the regime and  $E_{t-1} v_t = 0$ .

**IV.4. Summary of linearized equilibrium conditions.** We now summarize the linearized equilibrium conditions to be used for solving and estimating the model. These conditions are listed below.

$$\hat{\pi}_t - \gamma_p \hat{\pi}_{t-1} = \frac{\kappa_p}{1 + \bar{\alpha}\theta_p} (\hat{\mu}_{pt} + \hat{m}c_t) + \beta \text{E}_t [\hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t], \quad (90)$$

$$\begin{aligned} \hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t - \gamma_w \hat{\pi}_{t-1} &= \frac{\kappa_w}{1 + \eta\theta_w} (\hat{\mu}_{wt} + \hat{m}r_s t - \hat{w}_t) + \\ &\beta \text{E}_t [\hat{w}_{t+1} - \hat{w}_t + \hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t]. \end{aligned} \quad (91)$$

$$\begin{aligned} \hat{q}_{kt} &= S''(\lambda_I) \lambda_I^2 \left\{ \Delta \hat{i}_t + \frac{1}{1 - \alpha_1} (\Delta \hat{q}_t + \alpha_2 \Delta \hat{z}_t) \right. \\ &\quad \left. - \beta \text{E}_t \left[ \Delta \hat{i}_{t+1} + \frac{1}{1 - \alpha_1} (\Delta \hat{q}_{t+1} + \alpha_2 \Delta \hat{z}_{t+1}) \right] \right\} \end{aligned} \quad (92)$$

$$\begin{aligned} \hat{q}_{kt} &= \text{E}_t \left\{ \Delta \hat{a}_{t+1} + \Delta \hat{U}_{c,t+1} - \frac{1}{1 - \alpha_1} [\alpha_2 \Delta \hat{z}_{t+1} + \Delta \hat{q}_{t+1}] \right. \\ &\quad \left. + \frac{\beta}{\lambda_I} \left[ (1 - \delta) \hat{q}_{k,t+1} - \delta \hat{\delta}_{t+1} + \tilde{r}_k \hat{r}_{k,t+1} \right] \right\}, \end{aligned} \quad (93)$$

$$\hat{r}_{kt} = \sigma_u \hat{u}_t, \quad (94)$$

$$0 = \text{E}_t \left[ \Delta \hat{a}_{t+1} + \Delta \hat{U}_{c,t+1} - \frac{1}{1 - \alpha_1} [\alpha_2 \Delta \hat{z}_{t+1} + \alpha_1 \Delta \hat{q}_{t+1}] + \hat{R}_t - \hat{\pi}_{t+1} \right], \quad (95)$$

$$\hat{k}_t = \frac{1 - \delta}{\lambda_I} \left[ \hat{k}_{t-1} - \frac{1}{1 - \alpha_1} (\alpha_2 \Delta \hat{z}_t + \Delta \hat{q}_t) \right] - \frac{\delta}{\lambda_I} \hat{\delta}_t + \left( 1 - \frac{1 - \delta}{\lambda_I} \right) \hat{i}_t, \quad (96)$$

$$\hat{y}_t = c_y \hat{c}_t + i_y \hat{i}_t + u_y \hat{u}_t + g_y \hat{g}_t, \quad (97)$$

$$\hat{y}_t = \alpha_1 \left[ \hat{k}_{t-1} + \hat{u}_t - \frac{1}{1 - \alpha_1} (\alpha_2 \Delta \hat{z}_t + \Delta \hat{q}_t) \right] + \alpha_2 \hat{l}_t, \quad (98)$$

$$\hat{w}_t = \hat{r}_{kt} + \hat{k}_{t-1} + \hat{u}_t - \frac{1}{1 - \alpha_1} (\alpha_2 \Delta \hat{z}_t + \Delta \hat{q}_t) - \hat{l}_t, \quad (99)$$

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1 - \rho_r) [\phi_\pi (\hat{\pi}_t - \hat{\pi}^*(s^t)) + \phi_y \hat{y}_t] + \sigma_{rt} \varepsilon_{rt}, \quad (100)$$

where

$$\hat{m}c_t = \frac{1}{\alpha_1 + \alpha_2} [\alpha_1 \hat{r}_{kt} + \alpha_2 \hat{w}_t] + \bar{\alpha} \hat{y}_t, \quad (101)$$

$$\hat{m}r_s t = \eta \hat{l}_t - \hat{U}_{ct}, \quad (102)$$

$$\begin{aligned} \hat{U}_{ct} &= \frac{\beta b (1 - \rho_a)}{\lambda_* - \beta b} \hat{a}_t - \frac{\lambda_*}{(\lambda_* - b)(\lambda_* - \beta b)} [\lambda_* \hat{c}_t - b(\hat{c}_{t-1} - \Delta \hat{\lambda}_t^*)] \\ &\quad + \frac{\beta b}{(\lambda_* - b)(\lambda_* - \beta b)} [\lambda_* \text{E}_t (\hat{c}_{t+1} + \Delta \hat{\lambda}_{t+1}^*) - b \hat{c}_t], \end{aligned} \quad (103)$$

$$\hat{\pi}^*(s_t) = [\hat{\pi}^*(1), \hat{\pi}^*(2)] e_{s_t}, \quad e_{s_t} = Q e_{s_{t-1}} + v_t, \quad (104)$$

$$(105)$$

and the steady-state variables are given by

$$\tilde{r}_k = \frac{\lambda_I}{\beta} - (1 - \delta), \quad (106)$$

$$u_y \equiv \frac{\tilde{r}_k \tilde{K}}{\tilde{Y} \lambda_I} = \frac{\alpha_1}{\mu_p}, \quad (107)$$

$$i_y = [\lambda_I - (1 - \delta)] \frac{\alpha_1}{\mu_p \tilde{r}_k}, \quad (108)$$

$$c_y = 1 - i_y - g_y, \quad (109)$$

with  $\lambda_I \equiv (\lambda_q \lambda_z^{\alpha_2})^{\frac{1}{1-\alpha_1}}$ ,  $\lambda_* \equiv (\lambda_z^{\alpha_2} \lambda_q^{\alpha_1})^{\frac{1}{1-\alpha_1}}$ ,  $\Delta \hat{\lambda}_t^* \equiv \frac{1}{1-\alpha_1}(\alpha_1 \Delta \hat{q}_t + \alpha_2 \Delta \hat{z}_t)$ , and  $g_y$  calibrated to match the average ratio of government spending to real GDP.

Recall that  $\theta_p \equiv \frac{\mu_p}{\mu_p-1}$ ,  $\Delta x_t = x_t - x_{t-1}$ ,  $\kappa_p \equiv \frac{(1-\beta\xi_p)(1-\xi_p)}{\xi_p}$ ,  $\bar{\alpha} \equiv \frac{1-\alpha_1-\alpha_2}{\alpha_1+\alpha_2}$ ,  $\theta_w \equiv \frac{\mu_w}{\mu_w-1}$ ,  $\kappa_w \equiv \frac{(1-\beta\xi_w)(1-\xi_w)}{\xi_w}$ , and  $\hat{\pi}_t^w = \hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t + \Delta \hat{\lambda}_t^*$ ,

To compute the equilibrium, we eliminate  $\hat{u}_t$  by using (97), leaving 10 equations (90)-(96) and (98)-(100) with 10 variables  $\hat{\pi}_t$ ,  $\hat{w}_t$ ,  $\hat{i}_t$ ,  $\hat{q}_{kt}$ ,  $\hat{r}_{kt}$ ,  $\hat{c}_t$ ,  $\hat{k}_t$ ,  $\hat{y}_t$ ,  $\hat{l}_t$ , and  $\hat{R}_t$ . Out of these 10 variables, we have 7 observable variables, that is, all but  $\hat{q}_{kt}$ ,  $\hat{r}_{kt}$ , and  $\hat{k}_t$ , for our estimation. We also include the biased technology shock  $\hat{q}_t$  in the set of observable variables.

## V. GENERAL SETUP FOR ESTIMATION

In this section, we describe our empirical strategy in general terms so that the method can be applied to any state-space-form model.

Consider a regime-switching DSGE model with  $s_t$  following a Markov-switching process. Let  $\theta$  be a vector of all the model parameters except the transition matrix for  $s_t$ . Let  $y_t$  be an  $n \times 1$  vector of observable variables. In our case,  $n = 8$ . The vector  $y_t$  is connected to the state vector  $f_t$ . For our regime-switching DSGE model, this state-space representation implies a non-standard Kalman-filter problem as discussed in Kim and Nelson (1999).

Let  $(Y_t, \theta, Q, S_t)$  be a collection of random variables where

$$\begin{aligned} Y_t &= (y_1, \dots, y_t) \in (\mathbb{R}^n)^t, \\ \theta &= (\theta_i)_{i \in H} \in (\mathbb{R}^r)^h, \\ Q &= (q_{i,j})_{(i,j) \in H \times H} \in \mathbb{R}^{h^2}, \\ S_t &= (s_0, \dots, s_t) \in H^{t+1}, \\ S_{t+1}^T &= (s_{t+1}, \dots, s_T) \in H^{T-t}, \end{aligned}$$

and  $H$  is a finite set with  $h$  elements and is usually taken to be the set  $\{1, \dots, h\}$ . Because  $s_t$  represents a composite regime,  $h$  can be greater than the actual number of regimes at time  $t$ . The matrix  $Q$  is the Markov transition matrix and  $q_{i,j}$  is the probability that  $s_t$  is equal to  $i$  given that  $s_{t-1}$  is equal to  $j$ . The matrix  $Q$  is restricted to satisfy

$$q_{i,j} \geq 0 \text{ and } \sum_{i \in H} q_{i,j} = 1.$$

The object  $\theta$  is a vector of all the model parameters except the elements in  $Q$ . The object  $S_t$  represents a sequence of unobserved regimes or states. We assume that

$(Y_t, \theta, Q, S_t)$  has a joint density function  $p(Y_t, \theta, Q, S_t)$ , where we use the Lebesgue measure on  $(\mathbb{R}^n)^t \times (\mathbb{R}^r)^h \times \mathbb{R}^{h^2}$  and the counting measure on  $H^{t+1}$ . This density satisfies the following key condition.

*Condition 1.*

$$p(s_t | Y_{t-1}, \theta, Q, S_{t-1}) = q_{s_t, s_{t-1}}$$

for  $t > 0$ .

**V.1. Propositions for Hamilton filter.** Given  $p(y_t | Y_{t-1}, \theta, Q, s_t)$  for all  $t$ , the following propositions follow from Condition 1 (Hamilton, 1989; Chib, 1996; Sims, Waggoner, and Zha, 2008).

*Proposition 1.*

$$p(s_t | Y_{t-1}, \theta, Q) = \sum_{s_{t-1} \in H} q_{s_t, s_{t-1}} p(s_{t-1} | Y_{t-1}, \theta, Q)$$

for  $t > 0$ .

*Proposition 2.*

$$p(s_t | Y_t, \theta, Q) = \frac{p(y_t | Y_{t-1}, \theta, Q, s_t) p(s_t | Y_{t-1}, \theta, Q)}{\sum_{s_{t-1} \in H} p(y_t | Y_{t-1}, \theta, Q, s_t) p(s_t | Y_{t-1}, \theta, Q)}$$

for  $t > 0$ .

*Proposition 3.*

$$p(s_t | Y_t, \theta, Q, s_{t+1}) = p(s_t | Y_T, \theta, Q, S_{t+1}^T)$$

for  $0 \leq t < T$ .

**V.2. Likelihood.** We follow the standard assumption in the literature that the initial data  $Y_0$  is taken as given. Using Kim and Nelson (1999)'s Kalman-filter updating procedure, we obtain the conditional likelihood function at time  $t$

$$p(y_t | Y_{t-1}, \theta, Q, s_t). \tag{110}$$

It follows from the rules of conditioning that

$$\begin{aligned} p(y_t, | Y_{t-1}, \theta, Q) &= \sum_{s_t \in H} p(y_t, s_t | Y_{t-1}, \theta, Q) \\ &= \sum_{s_t \in H} p(y_t | Y_{t-1}, \theta, Q, s_t) p(s_t | Y_{t-1}, \theta, Q). \end{aligned}$$

Using (110) and the above equation, one can show that the likelihood function of  $Y_T$  is

$$\begin{aligned} p(Y_T | \theta, Q) &= \prod_{t=1}^T p(y_t | Y_{t-1}, \theta, Q) \\ &= \prod_{t=1}^T \left[ \sum_{s_t \in H} p(y_t | Y_{t-1}, \theta, Q, s_t) p(s_t | Y_{t-1}, \theta, Q) \right]. \end{aligned} \quad (111)$$

We assume that  $p(s_0 | Y_0, \theta, Q) = \frac{1}{h}$  for every  $s_0 \in H$ .<sup>1</sup> Given this initial condition, the likelihood function (111) can be evaluated recursively, using Propositions 1 and 2.

**V.3. Posterior distributions.** The prior for all the parameters is denoted by  $p(\theta, Q)$ , which will be discussed further in the main text of the article. By the Bayes rule, it follows from (111) that the posterior distribution of  $(\theta, Q)$  is

$$p(\theta, Q | Y_T) \propto p(\theta, Q) p(Y_T | \theta, Q). \quad (112)$$

The posterior density  $p(\theta, Q | Y_T)$  is unknown and complicated; the Monte Carlo Markov Chain (MCMC) simulation directly from this distribution can be inefficient and problematic. One can, however, use the idea of Gibbs sampling to obtain the empirical joint posterior density  $p(\theta, Q, S_T | Y_T)$  by sampling alternately from the following conditional posterior distributions:

$$\begin{aligned} p(S_T | Y_T, \theta, Q), \\ p(Q | Y_T, S_T, \theta), \\ p(\theta | Y_T, Q, S_T). \end{aligned}$$

One can use the Metropolis-Hastings sampler to sample from the conditional posterior distributions  $p(\theta | Y_T, Q, S_T)$  and  $p(Q | Y_T, S_T, \theta)$ . To simulate from the distribution  $p(S_T | Y_T, \theta, Q)$ , we can see from the rules of conditioning that

$$\begin{aligned} p(S_T | Y_T, \theta, Q) &= p(s_T | Y_T, \theta, Q) p(S_{T-1} | Y_T, \theta, Q, S_T^T) \\ &= p(s_T | Y_T, \theta, Q) \prod_{t=0}^{T-1} p(s_t | Y_T, \theta, Q, S_{t+1}^T) \end{aligned} \quad (113)$$

---

<sup>1</sup>The conventional assumption for  $p(s_0 | \theta, Q)$  is the ergodic distribution of  $Q$ , if it exists. This convention, however, precludes the possibility of allowing for an absorbing regime or state.



where  $S_{t+1}^T = \{s_{t+1}, \dots, s_T\}$ . From Proposition 3,

$$\begin{aligned}
 p(s_t | Y_T, \theta, Q, S_{t+1}^T) &= p(s_t | Y_t, \theta, Q, s_{t+1}) \\
 &= \frac{p(s_t, s_{t+1} | Y_t, \theta, Q)}{p(s_{t+1} | Y_t, \theta, Q)} \\
 &= \frac{p(s_{t+1} | Y_t, \theta, Q, s_t) p(s_t | Y_t, \theta, Q)}{p(s_{t+1} | Y_t, \theta, Q)} \\
 &= \frac{q_{s_{t+1}, s_t} p(s_t | Y_t, \theta, Q)}{p(s_{t+1} | Y_t, \theta, Q)}.
 \end{aligned} \tag{114}$$

The conditional density  $p(s_t | Y_T, Z_T, \theta, Q, S_{t+1}^T)$  is straightforward to evaluate according to Propositions 1 and 2.

To draw  $S_T$ , we use the backward recursion by drawing the last state  $s_T$  from the terminal density  $p(s_T | Y_T, \theta, Q)$  and then drawing  $s_t$  recursively given the path  $S_{t+1}^T$  according to (114). It can be seen from (113) that draws of  $S_T$  this way come from  $\Pr(S_T | Y_T, \theta)$ .

**V.4. Marginal posterior density of  $s_t$ .** The smoothed probability of  $s_t$  given the values of the parameters and the data can be evaluated through backward recursions. Starting with  $s_T$  and working backward, we can calculate the probability of  $s_t$  conditional on  $Y_T, \theta, Q$  by using the following fact

$$\begin{aligned}
 p(s_t | Y_T, \theta, Q) &= \sum_{s_{t+1} \in H} p(s_t, s_{t+1} | Y_T, \theta, Q) \\
 &= \sum_{s_{t+1} \in H} p(s_t | Y_T, \theta, Q, s_{t+1}) p(s_{t+1} | Y_T, \theta, Q)
 \end{aligned}$$

where  $p(s_t | Y_t, \theta, Q, s_{t+1})$  can be evaluated according to (114).

## VI. COMPARING THE DSGE MODEL WITH THE BVAR MODEL

We compare our models with a four-lag BVAR model estimated with log level data. The maximum log posterior density with (Sims and Zha, 1998)'s prior is 5116.80; thus, by the Schwarz criterion, all the DSGE models fit to the data better. The marginal data density for this BVAR, however, is 5866.18, higher than the marginal data densities of DSGE models.

Figure 1 plots the time series of conditional likelihoods  $p(y_t | Y_{t-1}, \hat{\theta}, \hat{q})$ . The likelihoods are smaller for the DSGE model because it has fewer parameters. But the conditional likelihoods for both DSGE and BVAR models tend to move in tandem,

implying that the DSGE model can capture the similar dynamics to those generated by the BVAR model.

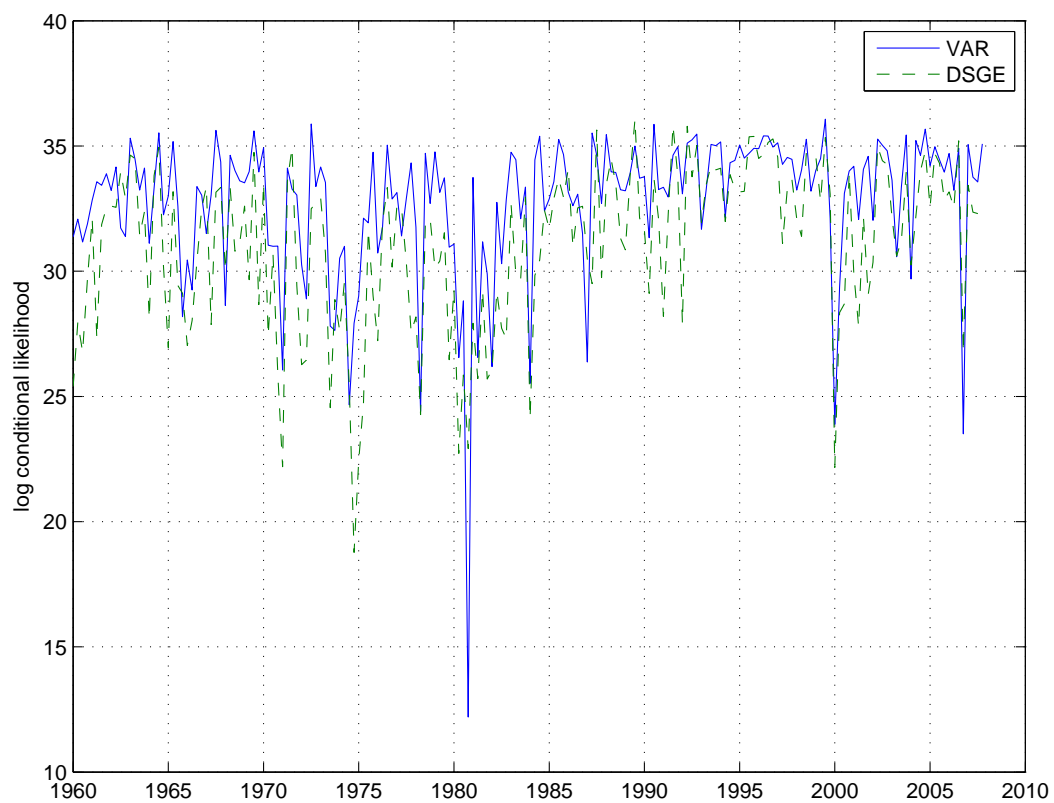


FIGURE 1. Conditional likelihoods for the DSGE-2v model and the BVAR model.

## REFERENCES

- CHIB, S. (1996): “Calculating Posterior Distributions and Model Estimates in Markov Mixture Models,” *Journal of Econometrics*, 75, 79–97.
- HAMILTON, J. D. (1989): “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle,” *Econometrica*, 57(2), 357–384.
- (1994): *Times Series Analysis*. Princeton University Press, Princeton, NJ.
- KIM, C.-J., AND C. R. NELSON (1999): *State-Space Models with Regime Switching*. MIT Press, London, England and Cambridge, Massachusetts.
- SIMS, C. A., D. F. WAGGONER, AND T. ZHA (2008): “Methods for Inference in Large Multiple-Equation Markov-Switching Models,” *Journal of Econometrics*, 146(2), 255–274.
- SIMS, C. A., AND T. ZHA (1998): “Bayesian Methods for Dynamic Multivariate Models,” *International Economic Review*, 39(4), 949–968.

FEDERAL RESERVE BANK OF SAN FRANCISCO, FEDERAL RESERVE BANK OF ATLANTA, FEDERAL RESERVE BANK OF ATLANTA AND EMORY UNIVERSITY