HAS THE FEDERAL RESERVE'S INFLATION TARGET CHANGED?

ZHENG LIU, DANIEL F. WAGGONER, AND TAO ZHA

Abstract. We confront a variety of medium-scale regime-switching DSGE models against U.S. macroeconomic time series data. Our goal is to employ a unified Bayesian framework for these models to test empirical evidence of regime changes in the Federal Reserve's inflation target in the post-war period when heteroscedastic shock disturbances are properly taken into account.

I. Introduction

This paper aims to contribute to a recently active line of research that analyzes the evolution of monetary policy and its potential effects on the economy. Using a small DSGE model, Clarida, Galí, and Gertler (2000), Lubik and Schorfheide (2004), and Boivin and Giannoni (2006), among others, find that the U.S. monetary policy rule has switched significantly for the better from the pre-Volcker regime to the post-Volcker regime. On the other hand, Stock and Watson (2003) report that it is hard to detect direct evidence of changes in reduced-form coefficients in their loosely parameterized time-series models. In line with the finding by Stock and Watson (2003), Canova and Gambetti (2004), Cogley and Sargent (2005), and Primiceri (2005) also document small drifts in their vector autoregression (VAR) coefficients. By explicitly identifying different monetary policy regimes, moreover, Sims and Zha (2006) find that once heteroscedasticity in shock variances is properly accounted for, the data no longer favor changes in monetary policy or changes in other parts of the economy.[1]

Date: December 20, 2007.

Key words and phrases. Macroeconomic volatility, regime switches, structural approach, Bayesian estimation.

JEL classification: E32, E42, E52.

All empirical results are preliminary and tentative. We thank Tim Cogley, Wouter Den Haan, Jordi Gali, Giorgio Primiceri, Frank Schorfheide, Chris Sims, and participants at the conference "How Much Structure in Empirical Models" organized by Fabio Canova and held in Barcelona in November 2007. We also thank Eric Wang for his research assistance on cluster computing in the Linux operating system. The views expressed herein are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

[1]See Leeper and Zha (2003) for other evidence.

The Sims and Zha (2006) results lend support to a vast amount of monetary policy research using dynamic stochastic general equilibrium (DSGE) models based on the assumption that the response coefficients in the monetary policy rule remain constant over time. Their results suggest that if one allows for a richer, more loosely parameterized structure, the model is less likely to favor changes in U.S. monetary policy. Indeed, Smets and Wouters (2007) and Justiniano and Primiceri (2006) use DSGE models (that are considerably larger than a standard small DSGE model) to show that there is little evidence of changes in the response coefficients in U.S. monetary policy by splitting the sample into two sub-samples, the pre-Volcker and post-Volcker periods. Arguably, a splitting of the sample ignores the expectation effects that agents anticipate future changes in monetary policy. In a recent work, however, Liu, Waggoner, and Zha (2007) show that if one allows economic agents in a DSGE model to take changes in the policy coefficients into account in their expectation formation, it is even harder to obtain the effects of such policy changes on macroeconomic volatility.

While all these findings support the assumption that the response coefficients in monetary policy have not changed much, it does not mean that monetary policy has not changed in other forms. In an important paper, Schorfheide (2005) finds strong evidence of regime switches in the U.S. inflation target even if variances of policy shocks are allowed to change regimes. Furthermore, he shows that by allowing regime switches in the target, the probability for the estimated inflation coefficient in the Taylor rule to lie below one (corresponding to a dovish monetary policy regime) is less than 10%. Since the work of Schorfheide (2005) is based on a simple, small DSGE model, a crucial question is whether his findings continue to hold when one allows for a richer structure in DSGE modeling. If regime switches in the inflation target turn out to be empirically important in a larger DSGE model, one needs to take explicit account of such changes.

To this end, we study a large set of variations of our DSGE model by accomplishing a sequence of three tasks. First, we construct a medium-scale DSGE model along the line of Altig, Christiano, Eichenbaum, and Linde (2004) and Smets and Wouters (2007). We choose such a medium-scale model over the small-scale New-Keynesian model commonly used in the literature because it pushes the limits of usefulness of DSGE modeling and such a medium-scale model seems necessary to fit the data well (Guerron-Quintana, 2007). Second, we extend our constant-parameter DSGE model by allowing shock variances to follow a Markov-switching process. A recent work by Justiniano and Primiceri (2006) introduces time-varying shock volatilities into a DSGE model and finds an important role of volatility changes in explaining the reduction of macroeconomic volatility. Third, we allow the inflation target to switch regime over time in addition to heteroblastic shock variances. As implied by Schorfheide (2005), with the inflation target switching regime, the model is likely to rely less on shock volatilities to generate the 1970s observations of the dynamics of inflation, output, and other macroeconomic variables.[2] Unlike Sims and Zha (2006) where the number of policy coefficients is relatively large, our way of modeling policy changes gives a tightly parameterized model that has the best potential to find the importance of policy changes in the presence of heteroscedastic variances if it exists.

We believe that this line of research represents a necessary and important step to understand both the nature of disturbances responsible for changes in macroeconomic volatility and the source of changes in monetary policy. It also helps to resolve the uncertainty about the importance of regime changes in the inflation target in light of the work by Schorfheide (2005). Above all, it enables us to study all these models in a unified, coherent Bayesian framework in which each model is evaluated by its marginal likelihood. Our framework is flexible because regime switches can apply to shock variances and inflation targets simultaneously or we can let a regime-switching process for variances be modeled independently of a regime-switching process for targets, as shown by Sims, Waggoner, and Zha (2006). We can also make a regime switch permanent (by making one of the regimes an absorbing state) to test for evidence of a once-for-all change in the inflation target or in shock variances or both.

While our framework is flexible, the resulting models are complex and push the limits of what our computational and analytical capacity can handle. The posterior density function is extremely non-Gaussian, making it difficult to find the posterior peak. The posterior peak in turn is important as a starting point for initializing an efficient MCMC algorithm. The results we report in this draft of the paper are incomplete at this stage, because of the extended computing time required by our analysis.

Nonetheless, our preliminary results so far show that allowing for such heteroscedasticity in shock variances improves the model's fit considerably, as found in Sims and Zha (2006). We measure this improvement by likelihood (multiplied by the prior) values. We report our results and point out some identification issues related to DSGE models.

[2]Econometrically, Choi (2002) and Beyer and Farmer (2005) argue that allowing regime shifts can potentially aid identification of other parameters in the model.

II. Relation to Other Literature

There are two strands of literature that are relevant to our work. One strand is related to medium-scale DSGE modelling with no regime switching. Besides the works discussed in Introduction, works by Levin, Onatski, Williams, and Williams (2006) and Del Negro, Schorfheide, Smets, and Wouters (2007) offer a useful reference for the estimates of some key parameters in a DSGE monetary model.

The other strand emphasizes changes in the inflation target as a representation of important shifts in the conduct of U.S. monetary policy. Favero and Rovelli (2003) estimate a three-equation New Keynesian macroeconomic model and detect a one-time shift in the inflation target around 1979 when Volcker became the Chairman of the Federal Reserve System. Such a one-time shift represents a permanent regime switch in the inflation target process, corresponding to a special (degenerate) case of the regime-switching process studied in our paper. Erceg and Levin (2003) studies the implications of agents' inability to disentangle persistent and transitory shifts in the Federal Reserve's inflation target on the dynamics of inflation and output following the Volcker disinflation. Ireland (2005) estimates a small New Keynesian monetary model to draw inferences about the dynamic behaviors of the Federal Reserve's inflation target and detects evidence of substantial time variations in the unobserved target inflation rate. All these studies assume that shock variances remain constant over time and there is no econometric study of how important changes in the target are relative to shock volatilities in explaining the time series of macroeconomic variables. For reasons of feasibility and tractability, we follow these works and take regime changes in the inflation target as an exogenous process.[3] Unlike these works, however, we study a variety of medium-scale DSGE models to determine whether the Federal Reserve's inflation target in the post-war period has switched regimes in the presence of heteroscedastic disturbances.

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

[3] One exception is Sargent, Williams, and Zha (2006) who derive changes in the inflation target endogenously in the context of a learning and escape model where the central bank has a misspecified model.

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 $q_{ij} = Prob(s_{t+1} = i | s_t = j)$ 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

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

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

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

and that

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

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

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

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

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

where $K_t^f(j)$ and $L_t^f(j)$ are the inputs of capital and the composite skill, and λ_z is the growth rate of the labor-augmenting technological change. The variable Z_t denotes a neutral technology shock, which follows a stationary process

$$\ln Z_t = (1 - \rho_z) \ln Z + \rho_z \ln Z_{t-1} + \sigma_{zt} \varepsilon_{zt}, \tag{6}$$

where $\rho_z \in (-1, 1)$ measures the persistence, $\sigma_{zt} \equiv \sigma_z(s_t^*)$ denotes the regime-switching standard deviation, and ε_{zt} is an i.i.d. white noise process with a zero mean and a unit variance. The parameters α_1 and α_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 ξ_p. Following Woodford (2003), CEE (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

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

where $\pi_t = \bar{P}_t / \bar{P}_{t-1}$ is the inflation rate between $t-1$ and t, π is the steady-state inflation rate, and γ_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

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

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

$$\chi^p_{t,t+i} = \begin{cases} \Pi^i_{k=1}\pi^{\gamma_p}_{t+k-1}\pi^{1-\gamma_p} & \text{if } i \geq 1 \\ 1 & \text{if } i = 0. \end{cases} \tag{9}$$

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

$$\mathrm{E}_t \sum_{i=0}^{\infty} \xi^i_p D_{t,t+i} Y^d_{t+i}(j) \frac{1}{\mu_{p,t+i}-1} \left[\mu_{p,t+i}\Phi_{t+i}(j) - P_t(j)\chi^p_{t,t+i}\right] = 0, \tag{10}$$

where $\Phi_{t+i}(j) = \partial V_{t+i}(j)/\partial Y^d_{t+i}(j)$ denotes the marginal cost function. In the absence of markup shocks, μ_{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

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

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

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

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.[4] In what follows, we impose this assumption and omit the household index for consumption and investment.

[4]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.

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

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

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

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

where $\rho_a \in (-1,1)$ is the persistence parameter, $\sigma_{at} \equiv \sigma_a(s_t^*)$ is the regime-switching standard deviation, and ε_{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

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

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, Π_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,

$$Q_t = \lambda_q^t q_t, \tag{16}$$

where λ_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

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

where $\rho_q \in (-1,1)$ is the persistence parameter, $\sigma_{qt} \equiv \sigma_q(s_t^*)$ is the regime-switching standard deviation, and ε_{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 (Forthcoming).

The capital stock evolves according to the law of motion

$$K_t = (1-\delta)K_{t-1} + \left[1 - S(I_t/I_{t-1})\right] I_t, \tag{18}$$

where $\delta \in (0,1)$ denotes the depreciation rate of the capital stock. The function $S(\cdot)$ represents the adjustment cost in capital accumulation. We assume that $S(\cdot)$ is convex and satisfies $S(\lambda_q \lambda_*) = S'(\lambda_q \lambda_*) = 0$, where $\lambda_* = (\lambda_q^{\alpha_1} \lambda_z^{\alpha_2})^{\frac{1}{1-\alpha_1}}$.

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 ξ_w of households cannot re-optimize their wage decisions and, among those who cannot re-optimize, a fraction γ_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

$$W_t(h) = \pi_{t-1}^{\gamma_w} \pi^{1-\gamma_w} \lambda_* W_{t-1}(h). \tag{19}$$

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

$$\mathrm{E}_t \sum_{i=0}^{\infty} \xi_w^i D_{t,t+i} L_{t+i}^d(h) \frac{1}{\mu_{w,t+i} - 1} \left[\mu_{w,t+i} MRS_{t+i}(h) - W_t(h) \chi_{t,t+i}^w\right] = 0, \tag{20}$$

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

$$\chi_{t,t+i}^w \equiv \begin{cases} \Pi_{k=1}^i \pi_{t+k-1}^{\gamma_w} \pi^{1-\gamma_w} \lambda_*^i & \text{if } i \geq 1 \\ 1 & \text{if } i = 0. \end{cases} \tag{21}$$

In the absence of wage-markup shocks, μ_{wt} would be a constant and (20) 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 (20) implies that the nominal wage is a markup over the contemporaneous marginal rate of substitution.

The optimal choice of bond holdings leads to the equilibrium relation

$$D_{t,t+1} = \beta \frac{A_{t+1} U_{c,t+1}}{A_t U_{ct}} \frac{\bar{P}_t}{\bar{P}_{t+1}}, \tag{22}$$

where U_{ct} denotes the marginal utility of consumption. This equation states that the intertemporal marginal rate of substitution equals the price of the state contingent bond. The return to the risk-free nominal bond, that is, the nominal interest rate is then given by $R_t = [\mathrm{E}_t D_{t,t+1}]^{-1}$. It follows from (22) that

$$1 = \beta \mathrm{E}_t \left[\frac{A_{t+1} U_{c,t+1}}{A_t U_{ct}} \frac{R_t}{\pi_{t+1}} \right], \tag{23}$$

which is the intertemporal Euler equation for the risk-free nominal bond.

The optimal investment decision is described by

$$\frac{A_t U_{ct}}{Q_t} = q_{kt} \left[1 - S(\lambda_{It}) - \lambda_{It} S'(\lambda_{It}) \right] + \beta \mathrm{E}_t q_{k,t+1} S'(\lambda_{I,t+1}) \lambda_{I,t+1}^2, \tag{24}$$

where $\lambda_{It} = I_t / I_{t-1}$ and q_{kt} is the Lagrangian multiplier of the law of motion of capital stock (18). The left hand side of the equation gives the marginal value of consuming a unit of the final good (in terms of investment goods). The right hand side of the equation gives the marginal value of investing the unit, which consists of the value of the increased level of new capital net of adjustment cost and the expected present value of reduced adjustment cost in the next period for having more capital in place.

The optimal capital accumulation rule is described by

$$q_{kt} = \beta(1-\delta) \mathrm{E}_t q_{k,t+1} + \beta \mathrm{E}_t U_{c,t+1} \left[r_{k,t+1} u_{t+1} - \frac{a(u_{t+1})}{Q_{t+1}} \right]. \tag{25}$$

The cost of acquiring an extra unit of capital is q_{kt} today. The benefit of having this extra unit of capital consists of the discounted expected future re-sale value and the rental value net of utilization cost.

The optimal choice of capital utilization rate is described by

$$\frac{a'(u_t)}{Q_t} = r_{kt}, \tag{26}$$

which equates the marginal cost (the left hand side) and the marginal benefit (the right hand side) of increased utilization rate.

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. We assume that the detrended government spending $\tilde{G}_t \equiv \frac{G_t}{\lambda^{*t}}$ follows a stationary stochastic process given by

$$\ln \tilde{G}_t = (1-\rho_g)\ln \tilde{G} + \rho_g \ln \tilde{G}_{t-1} + \sigma_{gt}\varepsilon_{gt} + \rho_{gz}\sigma_{zt}\varepsilon_{zt}, \tag{27}$$

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

$$R_t = \kappa R_{t-1}^{\rho_r} \left[\left(\frac{\pi_t}{\pi^*(s_t)} \right)^{\phi_\pi} \tilde{Y}_t^{\phi_y} \right]^{1-\rho_r} e^{\sigma_{rt}\varepsilon_{rt}}, \tag{28}$$

where $R_t = [\mathrm{E}_t D_{t,t+1}]^{-1}$ denotes the nominal interest rate, $\pi^*(s_t)$ denotes the regime-dependent inflation target, and $\tilde{Y}_t = Y_t/\lambda_*^t$ denotes the detrended output. The constant terms κ, ρ_r, ϕ_π, and ϕ_y are policy parameters. The term ε_{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 seven shocks ε_{wt}, ε_{pt}, ε_{zt}, ε_{qt}, ε_{at}, ε_{rt}, and ε_{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

$$C_t + \frac{1}{Q_t}[I_t + a(u_t)K_{t-1}] + G_t = Y_t, \tag{29}$$

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

$$G_{pt} Y_t = Z_t (u_t K_{t-1})^{\alpha_1} (\lambda_z^t L_t)^{\alpha_2}, \tag{30}$$

with $G_{pt} \equiv \int_0^1 \left(\frac{P_t(j)}{P_t}\right)^{-\frac{\mu_{pt}}{\mu_{pt}-1}\frac{1}{\alpha_1+\alpha_2}} dj$ 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

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

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 λ_q. Thus, the transformed variable $\tilde{r}_{kt} = r_{kt} Q_t$, that is, the rental price in consumption unit, is stationary. Further, since consumption is growing at the rate λ_*, the marginal utility of consumption is declining at the same rate, 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

$$\pi^*(s_t) = [\pi^*(1),\ \pi^*(2)] e_{s_t}, \tag{31}$$

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

$$e_{s_t} = \begin{bmatrix} \mathbf{1}\{s_t = 1\} \\ \mathbf{1}\{s_t = 2\} \end{bmatrix}, \tag{32}$$

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:

$$e_{s_t} = Q e_{s_{t-1}} + v_t, \tag{33}$$

where Q is the transition matrix of the Markov switching process and the innovation vector has the property that $\mathrm{E}_{t-1} v_t = 0$. In the steady state, $v_t = 0$ so that (33) defines the ergodic probabilities for the Markov process and, from (31), the steady-state inflation π is the ergodic mean of the inflation target. Given π, the derivations for the rest of the steady-state equilibrium conditions are straightforward, as we show in Appendix C.

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

$$\hat{\pi}_t - \gamma_p \hat{\pi}_{t-1} = \frac{\kappa_p}{1+\bar{\alpha}\theta_p}(\hat{\mu}_{pt} + \hat{mc}_t) + \beta \mathrm{E}_t[\hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t], \tag{34}$$

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

$$\hat{mc}_t = \frac{1}{\alpha_1+\alpha_2}[\alpha_1(\hat{r}_{kt} - \hat{q}_t) + \alpha_2 \hat{w}_t - \hat{z}_t] + \bar{\alpha}\hat{y}_t. \tag{35}$$

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

$$\hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t - \gamma_w \hat{\pi}_{t-1} = \frac{\kappa_w}{1+\eta\theta_w}(\hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t) + \beta \mathrm{E}_t[\hat{w}_{t+1} - \hat{w}_t + \hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t], \tag{36}$$

where $\hat{w}_t$ denotes the log-deviations of the real wage, $\hat{mrs}_t = \eta \hat{l}_t - \hat{U}_{ct}$ denotes the marginal rate of substitution between leisure and consumption, and $\kappa_w \equiv \frac{(1-\beta\xi_w)(1-\xi_w)}{\xi_w}$ is a constant. To help understand the economics of this equation, we rewrite it as

$$\hat{\pi}^w_t - \gamma_w \hat{\pi}_{t-1} = \frac{\kappa_w}{1+\eta\theta_w}(\hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t) + \beta \mathrm{E}_t(\hat{\pi}^w_{t+1} - \gamma_w \hat{\pi}_t), \tag{37}$$

where $\hat{\pi}^w_t = \hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t$ 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:

$$\hat{q}_{kt} = S''(\lambda_I)\lambda_I^2 \left[\Delta\hat{i}_t + \Delta\hat{q}_t - \beta\mathrm{E}_t[\Delta\hat{i}_{t+1} + \Delta\hat{q}_{t+1}]\right], \tag{38}$$

$$\hat{q}_{kt} = \mathrm{E}_t\left\{\Delta\hat{a}_{t+1} + \Delta\hat{U}_{c,t+1} - \Delta\hat{q}_{t+1} + \frac{\beta}{\lambda_I}\left[(1-\delta)\hat{q}_{k,t+1} + \tilde{r}_k\hat{r}_{k,t+1}\right]\right\}, \tag{39}$$

$$\hat{r}_{kt} = \sigma_u \hat{u}_t, \tag{40}$$

$$0 = \mathrm{E}_t\left[\Delta\hat{a}_{t+1} + \Delta\hat{U}_{c,t+1} + \hat{R}_t - \hat{\pi}_{t+1}\right], \tag{41}$$

$$\hat{k}_t = \frac{1-\delta}{\lambda_I}[\hat{k}_{t-1} + \hat{q}_{t-1} - \hat{q}_t] + \left(1 - \frac{1-\delta}{\lambda_I}\right)\hat{i}_t, \tag{42}$$

$$\hat{y}_t = c_y\hat{c}_t + i_y\hat{i}_t + u_y\hat{u}_t + g_y\hat{g}_t, \tag{43}$$

$$\hat{y}_t = \hat{z}_t + \alpha_1[\hat{k}_{t-1} + \hat{u}_t + \hat{q}_{t-1}] + \alpha_2\hat{l}_t, \tag{44}$$

$$\hat{w}_t - \hat{r}_{kt} = \hat{k}_{t-1} + \hat{u}_t + \hat{q}_{t-1} - \hat{q}_t - \hat{l}_t, \tag{45}$$

where (38) is the linearized investment decision equation with $\hat{q}_{kt}$ denoting the shadow value of existing capital (i.e., Tobin's Q) and the Δ denoting the first-difference operator (so that $\Delta x_t = x_t - x_{t-1}$); (39) is the linearized capital Euler equation; (40) 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; (41) is the linearized bond Euler equation; (42) is the linearized law of motion for the capital stock; (43) is the linearized aggregate resource constraint, with the steady-state ratios given by $c_y = \frac{\tilde{C}}{\tilde{Y}}$, $i_y = \frac{\tilde{I}}{\tilde{Y}}$, $u_y = \frac{1}{\lambda_q\lambda_*}\frac{\tilde{r}_k\tilde{K}}{\tilde{Y}}$, and $g_y = \frac{\tilde{G}}{\tilde{Y}}$; (44) is the linearized aggregate production function; and (45) is the linearized factor demand relation.

Finally, the linearized interest rate rule is given by

$$\hat{R}_t = \rho_r\hat{R}_{t-1} + (1-\rho_r)\left[\phi_\pi(\hat{\pi}_t - \hat{\pi}^*(s^t)) + \phi_y\hat{y}_t\right] + \sigma_{rt}\varepsilon_{rt}, \tag{46}$$

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. To compute the equilibrium dynamics, we use the relation that

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

where the vector e_{s_t} is defined in (32) and follows a vector AR(1) process described in (33).

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 26 variables. Adding the four lagged variables $\hat{y}_{t-1}, \hat{c}_{t-1}, \hat{i}_{t-1}$, and $\hat{w}_{t-1}$ to the list gives a total of 30 variables. We denote all these state variables by the vector f_t where f_t is so arranged that the first 9 variables are $\hat{y}_t, \hat{c}_t, \hat{i}_t, \hat{w}_t, \hat{\pi}_t, \hat{\ell}_t, \hat{R}_t, \hat{k}_t$, and $\hat{q}_t$ and the last 4 variables are $\hat{y}_{t-1}, \hat{c}_{t-1}, \hat{i}_{t-1}$, and $\hat{w}_{t-1}$. The solution to our DSGE model leads to the following VAR(1) form of state equations

$$f_t = c(s_t, s_{t-1}) + Ff_{t-1} + C(s_t^*)\epsilon_t, \tag{47}$$

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

$C(s_t^*)$ is a matrix function of $\sigma_{rt}(s_t^*)$, $\sigma_{pt}(s_t^*)$, $\sigma_{wt}(s_t^*)$, $\sigma_{gt}(s_t^*)$, $\sigma_{zt}(s_t^*)$, $\sigma_{at}(s_t^*)$, and $\sigma_{qt}(s_t^*)$.

It follows from (47) 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 (see Schorfheide (2005)), then the composite regime collapses to (s_t, s_{t-1}). To simply our notation and make an analytical exposition 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 time-series observations on seven U.S. aggregate variables: real per capita GDP (Y_t^{Data}), real per capita consumption (C_t^{Data}), real per capita investment (I_t^{Data}), real wage (w_t^{Data}), the quarterly GDP-deflator inflation rate (π_t^{Data}), per capita hours (L_t^{Data}), and the (annualized) federal funds rate ($\text{FFR}_t^{\text{Data}}$). These data are related to the model through the following vector of observable variables:

$$y_t = \begin{bmatrix} \Delta \ln Y_t^{\text{Data}} \\ \Delta \ln C_t^{\text{Data}} \\ \Delta \ln I_t^{\text{Data}} \\ \Delta \ln w_t^{\text{Data}} \\ \ln \pi_t^{\text{Data}} \\ \ln L_t^{\text{Data}} \\ \frac{\text{FFR}_t^{\text{Data}}}{400} \end{bmatrix}.$$

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

$$y_t = a + Hf_t,$$

where

The vector a and matrix H are

$$a = \begin{bmatrix} \ln \lambda_* \\ \ln \lambda_* \\ \ln \lambda_* \\ \ln \lambda_* \\ \ln \pi \\ \ln L \\ \ln R \end{bmatrix}, \ H = \left[\begin{array}{c:c:c} \underset{7\times7}{I} & \underset{7\times19}{0} & \begin{bmatrix} -\underset{4\times4}{I} \\ \underset{3\times4}{0} \end{bmatrix} \end{array} \right]. \tag{48}$$

For the rest of this section, we describe our empirical strategy in general terms so that the method can be applied to other DSGE models.

V.1. **General setup for estimation.** Consider a regime-switching DSGE model with s_t following a Markov-switching process. Let θ 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, $n = 7$. The vector y_t is connected to the state vector f_t through (48). 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, θ, Q, S_t) be a collection of random variables where

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

and H is a finite set with h elements and is usually taken to be the set $\{1, \cdots, h\}$. 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 θ 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, θ, 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.

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

for $t > 0$.

V.2. **Propositions for Hamilton filter.** Given $p(y_t \mid Y_{t-1}, \theta, Q, s_t)$ for all t, the following propositions follow from Condition 1 (Hamilton, 1989; Chib, 1996; Sims, Waggoner, and Zha, 2006).

Proposition 1.

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

for $t > 0$.

Proposition 2.

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

for $t > 0$.

Proposition 3.

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

for $0 \leq t < T$.

V.3. **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

$$p(y_t \mid Y_{t-1}, \theta, Q, s_t)\,. \tag{49}$$

It follows from the rules of conditioning that

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

Using (49) and the above equation, one can show that the likelihood function of Y_T is

$$\begin{aligned} p(Y_T \mid \theta, Q) &= \prod_{t=1}^{T} p(y_t \mid Y_{t-1}, \theta, Q) \\ &= \prod_{t=1}^{T} \left[\sum_{s_t \in H} p(y_t \mid Y_{t-1}, \theta, Q, s_t)\, p(s_t \mid Y_{t-1}, \theta, Q)\right]. \end{aligned} \tag{50}$$

We assume that $p(s_0 \mid Y_0, \theta, Q) = \frac{1}{h}$ for every $s_0 \in H$.[5] Given this initial condition, the likelihood function (50) can be evaluated recursively, using Propositions 1 and 2.

[5]The conventional assumption for $p(s_0 \mid \theta, Q)$ is the ergodic distribution of Q, if it exists. This convention, however, precludes the possibility of allowing for an absorbing regime or state.

V.4. **Posterior distributions.** The prior for all the parameters is denoted by $p(\theta, Q)$, which will be discussed further in Section V.7. By the Bayes rule, it follows from (50) that the posterior distribution of (θ, Q) is

$$p(\theta, Q \mid Y_T) \propto p(\theta, Q) p(Y_T \mid \theta, Q). \tag{51}$$

The posterior density $p(\theta, Q \mid Y_T)$ 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 $p(\theta, Q, S_T \mid Y_T)$ by sampling alternately from the following conditional posterior distributions:

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

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

$$\begin{aligned} p\left(S_T \mid Y_T, \theta, Q\right) &= p\left(s_T \mid Y_T, \theta, Q\right) p\left(S_{T-1} \mid Y_T, \theta, Q, S_T^T\right) \\ &= p\left(s_T \mid Y_T, \theta, Q\right) \prod_{t=0}^{T-1} p\left(s_t \mid Y_T, \theta, Q, S_{t+1}^T\right) \end{aligned} \tag{52}$$

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

$$\begin{aligned} p\left(s_t \mid Y_T, \theta, Q, S_{t+1}^T\right) &= p\left(s_t \mid Y_t, \theta, Q, s_{t+1}\right) \\ &= \frac{p\left(s_t, s_{t+1} \mid Y_t, \theta, Q\right)}{p\left(s_{t+1} \mid Y_t, \theta, Q\right)} \\ &= \frac{p\left(s_{t+1} \mid Y_t, \theta, Q, s_t\right) p\left(s_t \mid Y_t, \theta, Q\right)}{p\left(s_{t+1} \mid Y_t, \theta, Q\right)} \\ &= \frac{q_{s_{t+1}, s_t} p\left(s_t \mid Y_t, \theta, Q\right)}{p\left(s_{t+1} \mid Y_t, \theta, Q\right)}. \end{aligned} \tag{53}$$

The conditional density $p\left(s_t \mid Y_T, Z_T, \theta, Q, S_{t+1}^T\right)$ 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 (53). It can be seen from (52) that draws of S_T this way come from $\Pr(S_T | Y_T, \theta)$.

V.5. **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, θ, Q by using the following fact

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

where $p\left(s_t \mid Y_t, \theta, Q, s_{t+1}\right)$ can be evaluated according to (53).

V.6. **The model's fit.** 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. The marginal data density is defined as

$$p(Y_T) = \int p(Y_T \mid \theta, Q) p(\theta)\, d\theta dQ, \tag{54}$$

where $p(Y_T \mid \theta, Q)$ can be evaluated according to (50). 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 (54) numerically is based on a theorem that states

$$p(Y_T)^{-1} = \int_{\Theta} \frac{h(\theta, Q)}{p(Y_T \mid \theta, Q) p(\theta, Q)} p(\theta, Q \mid Y_T) d\theta dQ, \tag{55}$$

where Θ 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 Θ. Denote

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

A numerical evaluation of the integral on the right hand side of (55) can be accomplished in principle through the Monte Carlo (MC) integration

$$\hat{p}(Y_T)^{-1} = \frac{1}{N} \sum_{i=1}^{N} m(\theta^{(i)}, Q^{(i)}), \tag{56}$$

where $(\theta^{(i)}, Q^{(i)})$ is the i^{th} draw of (θ, Q) from the posterior distribution $p(\theta, Q \mid Y_T)$. 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 large 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 (2006) 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.

Because the posterior density function is very non-Gaussian and complicated in shape, it is all the more important to find the posterior mode by maximizing the value of (51). The estimate at 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 (2006) 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.[6] 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.[7] 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

[6] The `csminwel` program can be found on `http://sims.princeton.edu/yftp/optimize/`.

[7] 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.

least a magnitude of hundreds of log values. For each DSGE model, the peak value of the posterior kernel and the estimates at the mode are reported.

V.7. **Prior distributions.** We fix 3 parameters before estimating the model. We fix the capital depreciation rate δ at 0.025 and the steady-state government spending to output ratio g_y at 0.18. As noted by Smets and Wouters (2007), these parameters are difficult to estimate unless capital stock and government spending are included in the measurement equations (see also Justiniano and Primiceri (2006)). We also normalize and fix the steady-state hours worked L at 1/3. We estimate all the remaining parameters, including the steady-state wage markup μ_w, which is fixed in Smets and Wouters (2007). Tables 1 and 2 summarize the prior distributions for the structural parameters and the shock parameters.

Our choice of the prior means for the structural parameters reported in Table 1 is primarily based on standard calibrations. We begin with the preference parameters b, η, and β. In light of the calibration in Boldrin, Christiano, and Fisher (2001) and Christiano, Eichenbaum, and Evans (2005), we set the prior mean for the habit persistence parameter to 0.7 with a standard deviation of 0.2. We set the prior mean of η to 2.0 with a standard deviation of 0.5. This value of η corresponds to a Frisch labor elasticity of 0.5, which is commonly used in the standard calibration. We set the prior mean of the subjective discount rate to 1.6% per annual, with a standard deviation of 0.2. Next, we discuss the prior distribution for the technology parameters α_1, α_2, λ_q, λ_*, σ_u, and S''. According to Chari, Kehoe, and McGrattan (2000), the value for the cost share of the firm-specific factor is typically set to 1/3, so that the share of the primary factors is $\alpha_1 + \alpha_2 = 2/3$; further, it is often assumed that roughly 1/3 of the primary factor income goes to capital input and 2/3 goes to labor input. Accordingly, we set the prior means for α_1 and α_2 to 0.2 and 0.467, respectively. The trend growth rate of the investment-specific technological change is set to 0.3 with a standard deviation of 0.1, corresponding to an annual growth rate of 1.2%. The trend growth rate of the neutral technological change is set to 0.5 with a standard deviation of 0.1, corresponding to an annual growth rate of 2.0%. We set the prior mean for the utilization parameter σ_u to 1.5 with a standard deviation of 0.5 and the prior mean for the adjustment cost parameter S'' to 4.0 with a standard deviation of 1.5. These values are based on Christiano, Eichenbaum, and Evans (2005) and are also used by Smets and Wouters (2007). 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 μ_p, the average wage markup μ_w, the Calvo probabilities of non-adjustment in pricing ξ_p and in wage-setting ξ_w, and the indexation parameters γ_p and γ_w. Under our prior distribution, the average price markup and wage markup are 20% each, with a standard deviation of 0.15; the price and nominal wage contracts both last on average for a year; and half of firms (households) index their prices (nominal wages) to past inflation. These parameter values are commonly used in the DSGE literature. Finally, we discuss the coefficients in the monetary policy rule, including ρ_r, ϕ_π, and ϕ_y. In light of the study by Clarida, Galí, and Gertler (2000), we set the prior mean for the interest-rate smoothing parameter ρ_r to 0.6 with a standard deviation of 0.2, the prior mean for the coefficient in front of inflation to 2.0 with a standard deviation of 1.0, and the prior mean for the coefficient in front of detrended output to 0.4 with a standard deviation of 0.25.

Following Smets and Wouters (2007), we assume that the AR(1) coefficient in each of the seven shock processes, as well as the MA(1) coefficients in the price and wage markup processes follow the Beta distribution with a mean of 0.5 and a standard deviation of 0.2. We assume the volatility parameter in each of the shock processes follows the Inverse Gamma distribution $InvGam(\alpha, \beta)$ where $\alpha = 0.65$ and $\beta = 0.15$. For these hyperparameter values, the first two moments of the prior distribution do not exist (thus, we display the asterisk symbol in the table). But these values imply a more diffuse prior than Smets and Wouters (2007) and Justiniano and Primiceri (2006), with the .90 probability interval being [0.066, 19.60]. Such a diffuse prior is needed to allow for possible large changes in shock variances across regimes, as found in Sims and Zha (2006).

VI. Estimation Results

The last three columns in Tables 1 and 2 summarize our posterior estimates of the structural and shock parameters for three alternative models. In the first model, we consider the case with constant parameters and no regime shifts. In the second model, we introduce regime shifts in the volatilities of the shocks, while the inflation target remains constant. In the third and final model, we allow regime shifts in both the shock volatilities and the inflation target.

VI.1. **Model I: no regime switching.** We first discuss the posterior estimates for the model with no regime shifts. This model is similar to that in Smets and Wouters (2007), with four 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 (e.g., equipment and software) has been declining for most of the postwar period, while in Smets and Wouters (2007), the investment-specific technology shock has no trend component. Third, our model features a preference shock, which enters all intertemporal decisions, including the 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 is independent of other intertemporal shocks. Finally, 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. These difference turns out to be important and renders our estimation results quite different from their studies.

Table 1 reveals that, overall, the data appear to be informative about many structural parameters. Among the three preference parameters, the posterior mode for habit persistence is 0.994, much higher than its prior. The posterior mode for η is 2.069, slightly higher than the prior. The estimate for β is about 0.9953, slightly lower than the prior of 0.9960. Among the technology parameters, the posterior mode for α_1 is similar to its prior, but the posterior for α_2 is much higher than its prior (0.795 v. 0.467). Without a priori restrictions on the importance of firm-specific factors (or equivalently, restrictions on $\alpha_1 + \alpha_2$), the data seem to prefer to have constant returns in the production function. With constant returns, the model loses one source of real rigidities and the propagation of shocks will have to rely more on greater degrees of nominal rigidities and more persistence in the shock processes, as we discuss below. The estimated trend growth rate for the investment-specific technological change is about 1.24% per annual, slightly higher than the prior; the estimated growth rate for the neutral technological change is about 1.71% per annual, somewhat lower than the prior. The curvature parameter in the utilization function is estimated at 2.05, much higher than the prior of 1.5. The investment adjustment cost parameter has a posterior mode of 3.28, somewhat lower than the prior mean of 4.0. Among the parameters that describe the price and wage setting behaviors, the posterior estimates are mostly far from the priors, suggesting that the data are indeed informative on these parameters. The price markup is estimated at 1.0001, which is much smaller than standard calibration; the wage markup is estimated at 1.425, which is much higher than the prior mean of 1.2 and which implies an elasticity of substitution between differentiated labor skills of about 3.35, which lies within the range calibrated by Huang and Liu (2002). The posterior estimates also suggest that price contracts last on average for about 12 quarters and nominal wage contracts last on average for slightly less than 9 quarters. These estimated price and wage rigidities are greater than many other studies and may reflect the lack of internal propagation mechanisms in the model. Interestingly, the estimated degrees of dynamic indexation are modest. The price indexation parameter is estimated at 0.179 and the wage indexation parameter is 0.535. Our posterior estimates of the policy parameters suggest that interest-rate smoothing is important, with a posterior mode for ρ_r of 0.939. The estimates also suggest that the interest rate rule responds aggressively to deviations of inflation from its target with a response coefficient of 1.444 and it responds to detrended output modestly, with a response coefficient of 0.591. Finally, the inflation target is estimated at 2.18% per annual, which is smaller than the prior mean of 4% per annual.

Our posterior estimates of the shock parameters reveal that all but the preference shock are very persistent, with the AR(1) coefficient greater than 0.9. The preference shock is less persistent, with an AR(1) coefficient of 0.576. The MA(1) coefficients in the price markup and wage markup processes are both sizable, at 0.666 and 0.651 respectively. In addition, the government spending shock also responds to the neutral technology shock, with a response coefficient of 0.585. Although we assume the same prior distribution for all the shock volatility parameters, we obtain very disperse posterior estimates for these volatility parameters. The estimates reveal that the wage markup shock, the price markup shock, the preference shock are the three largest shocks (in size), the monetary policy shock and the neutral technology shock are the two smallest shocks, and the investment-specific technology shock and the government spending shock lie somewhere in between.

VI.2. **Model II: regime shifts in shock variances.** We now discuss the posterior estimates for the model with regime shifts in shock volatilities, but with a constant inflation target. The estimates are reported in Table 1, the second to the last column.

Introducing regime shifts in shock volatilities influence the estimates for many of the structural parameters. The most important change compared to Model I with constant parameters is that, in the current model, the exogenous price and wage stickiness parameters are much smaller, so that the price and wage contract durations are much shorter. The estimates for ξ_p and ξ_w are now 0.828 and 0.794 instead of 0.919 and 0.886. These new estimates imply that the price and wage contracts last on average for 5.8 quarters and 4.8 quarters respectively. They are still longer than those estimated by many other studies, but are much more realistic than the estimates obtained in the model with constant volatilities. Compared to Model I, the estimated habit persistence here is smaller (0.911 v. 0.994); the capital share is larger (0.225 v. 0.203) and accordingly, the labor share is smaller (0.770 v. 0.795), although the data still prefer a constant returns production function; the inverse Frisch elasticity is slightly lower (1.956 v. 2.069); the average wage markup is slightly higher (1.440 v. 1.425); the price and wage indexation parameters are both higher (0.225 and 0.605 v. 0.179 and 0.535); and the estimated inflation target is slightly higher (0.550 v. 0.545). Other estimates of the structural parameters are similar to those in Model I.

Introducing regime shifts in shock volatilities also influence the estimates for the shock persistence parameters. Compared to the constant volatility case in Model I, the estimated AR(1) coefficient for the price markup shock becomes small (0.881 v. 0.962) and that for the wage markup shock becomes larger (0.971 v. 0.940); the estimated MA(1) coefficients in both the price markup and the wage markup shock processes are larger (0.707 and 0.879 v. 0.666 and 0.651); the preference shock becomes more persistent (0.645 v. 0.576) while the investment-specific technology shock becomes less persistent (0.832 v. 0.911). What is more interesting are the estimated realizations of the volatility parameters across the two regimes. The estimates reveal overall large reductions of the shock volatilities when moving from the first regime to the second: the volatility falls by at least one order of magnitude for all but the price and wage markup shocks. The volatilities of the price and wage markup shocks also fall substantially (from 0.936 and 0.905 to 0.426 and 0.537).

Finally, the estimates of the transition probabilities for the shock regimes are summarized by the matrix

$$\hat{Q} = \begin{bmatrix} 0.639 & 0.121 \\ 0.361 & 0.879 \end{bmatrix}. \tag{57}$$

Thus, the second regime (i.e., the regime with low shock volatilities) is more persistent.

VI.3. **Model III: regime shifts in both variances and the inflation target.** We consider a case in which there are two regimes for the inflation target and two regimes for shock variances. The Markov-switching process for the inflation target is independent of that for shock variances. While a cluster of computers are still running to find the peak of the posterior density for this case, I have so far been unable to get the posterior value at its peak higher than than the value for Model II. The highest peak value we have thus obtained is 4908.40, slightly lower than the posterior value for

Model II.[8] At this stage, it appears difficult to improve the fit of the model by allowing regime switches in the inflation target.

VII. Model Analysis

We now examine the propagation mechanisms in the DSGE model.

VII.1. **Impulse responses.** We first plot the impulse responses of the seven observable variables following each of the seven shocks under our estimated parameters. We focus on the model with regime shifts in shock volatilities (i.e., Model II).

Figures 1 and 2 plot the impulse responses of the seven observable variables following each of the seven shocks under the first shock regime, and Figures 3 and 4 plot the impulse responses under the second shock regime. The first shock regime features high volatilities of the shocks, as our posterior estimates suggest. Figure 1 shows that, following a tightening of monetary policy ('MP'), the nominal interest rate rises while the inflation rate, output, and other real variables all fall. Under the estimated parameters, the monetary policy shock generates a hump-shaped response of output and a weak hump for inflation. A price-markup shock ('PM') leads to a fall in output and other real variables and a rise in inflation. In this sense, the markup shock creates an unfavorable tradeoff between inflation and output. Following the rise in inflation, the nominal interest rate also rises since the monetary policy adjust interest rate aggressively against changes in inflation. A wage-markup shock ('WM') leads to a persistent decline in output, consumption, investment, and hours, but a rise in the real wage and inflation. A higher wage markup leads to higher real wages and thus higher real marginal cost for firms, and firms respond by raising prices when they can re-optimize pricing decisions. Following a government spending shock ('g'), output rises for several quarters, while consumption, investment, and the real wage all fall. The increase in government spending crowds out private consumption. To meet the higher demand for goods from the government, households have to work harder and average hours rise. The inflation rate initially rises modestly because of the increase in aggregate demand for goods and then declines after 5 quarters of the shock since the persistent decline in the real wage (and hence in the real marginal cost) begins to dominate the rise in aggregate demand. Following the rise in both output and inflation, the nominal interest rate rises. Since the nominal interest rate rises more than does the expected inflation rate, the real interest rate rises and investment falls. Figure 2 shows that,

[8]Note that the prior makes the posterior peak value of Model III slightly lower than that of Model II.

under the first regime, a positive neutral technology shock ('tech') leads to a persistent rise in output, consumption, investment, and the real wage. Inflation falls because the marginal cost of production declines following the improvement in productivity. Hours decline following the improvement in productivity, which is consistent with the predictions from the standard sticky-price models such as (Galí, 1999) and Basu, Fernald, and Kimball (2006). A positive preference shock ('pref') leads to a rise in output, consumption, the real wage, hours, and inflation, but a fall in investment. Households face higher marginal utility of consumption and thus invest less. Since both inflation and output rise, the nominal interest rate also rises. Finally, a positive investment-specific technology shock ('inv') leads to a rise in output, the real wage, and the inflation rate. Investment falls on impact and rises persistently thereafter, so do hours worked. The decline in investment is not surprising since the shock raises the trend growth rate of the investment-specific technology, and here investment measure deviations of actual investment from trend. The persistent rise in hours following the investment shock is consistent with the VAR evidence produced by Fisher (2006). As investment becomes more productive, saving increases and consumption stays flat initially; over time, the wealth effect raises consumption.

Figures 3 and 4 plot the impulse responses under the second shock regime, where the size of the shocks are much smaller than under the first regime. The overall qualitative patterns of the impulse responses resembles those under the first regime.

VII.2. **Variance decompositions.** Table 3 reports the forecast error variance decomposition for output and inflation under each of the two shock regimes at various forecasting horizons. Under the first shock regime, most of the output fluctuations in the short-run (between 4 quarters and 8 quarters) are driven mainly by the monetary policy shock. In the medium run (12 quarters to 16 quarters), the neutral and investment-specific technology shocks are both essential. Under this regime, the monetary policy shock plays an essential role in generating the inflation dynamics, although in the medium run, the neutral technology shock explains at least one half of the forecast variance of inflation.

Under the second shock regime, the dynamics of both output and inflation are primarily driven by two shocks: the wage markup shock and the neutral technology shock. Taken together, these two shocks account for $60-70\%$ of output variances in the short run and $85-90\%$ in the medium run. These two shocks also explain more than 80% of inflation fluctuations at the 4 quarter forecasting horizon and more than 90% thereafter.

VII.3. **Model Comparison.** Computing marginal likelihood for the models we study is a major computational task, which we have not undertaken at this point. We have instead computed maximum log posterior density (LPD) values, which can be used to compare the models by the Schwarz criterion. As reported in Table 4, The DSGE model with no regime shifts (Model I) does better than the VAR model with 4 lags, consistent with the findings reported by Smets and Wouters (2007). The VAR model is estimated with the log level data (not the log differences as in our DSGE models) using the Sims and Zha (1998) prior and the Gibbs sampler of Waggoner and Zha (2003). Furthermore, the model with regime shifts in shock variances (Model II) fits to the data better than the constant-parameter DSGE model. As for the DSGE model with 2-regime variances and 2-regime inflation targets (Model III), we have so far been unable to obtain the posterior peak value higher than that for Model II. We plan to compare more models with different numbers of regimes and different types of regimes in future research.

VIII. Conclusion

Despite a hard problem, we have demonstrated that it is possible or even feasible to estimate a medium-scale regime-switching DSGE model. While our results are regrettably tentative and more calculations are yet to be finished, our goal is to show that this line of research is promising.

Appendix A. Derivation of Optimizing Decisions

A.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:

$$\text{Max}_{\{C_t, I_t, K_t, u_t, B_{t+1}\}} \quad \mathrm{E}\sum_{t=0}^{\infty} \beta^t A_t \left\{ \log(C_t - bC_{t-1}) - \frac{\psi}{1+\eta} L_{t+i}^d(h)^{1+\eta} \right\} \tag{A1}$$

subject to

$$\bar{P}_t C_t + \frac{\bar{P}_t}{Q_t}(I_t + a(u_t)K_{t-1}) + \mathrm{E}_t D_{t,t+1} B_{t+1} \leq W_t(h) L_t^d(h) + \bar{P}_t r_{kt} u_t K_{t-1} + \Pi_t + B_t + T_t, \tag{A2}$$

$$K_t = (1-\delta)K_{t-1} + \left[1 - S\left(\frac{I_t}{I_{t-1}}\right)\right] I_t, \tag{A3}$$

Denote by μ_t the Lagrangian multiplier for the budget constraint (A2) and by μ_{kt} the Lagrangian multiplier for the capital accumulation equation (A3). The first order conditions for the utility-maximizing problem are given by

$$
\begin{aligned}
A_t U_{ct} &= \mu_t \bar{P}_t, && \text{(A4)}\\
D_{t,t+1} &= \beta \frac{\mu_{t+1}}{\mu_t}, && \text{(A5)}\\
\frac{\mu_t \bar{P}_t}{Q_t} &= \mu_{kt} \left\{1 - S(\lambda_{It}) - S'(\lambda_{It})\lambda_{It}\right\} + \beta \mathrm{E}_t \mu_{k,t+1} S'(\lambda_{I,t+1})(\lambda_{I,t+1})^2 && \text{(A6)}\\
\mu_{kt} &= \beta \mathrm{E}_t \left[\mu_{k,t+1}(1-\delta) + \mu_{t+1}\bar{P}_{t+1} r_{k,t+1} u_{t+1} - \frac{\mu_{t+1}\bar{P}_{t+1}}{Q_{t+1}} a(u_{t+1})\right], && \text{(A7)}\\
r_{kt} &= \frac{a'(u_t)}{Q_t}, && \text{(A8)}
\end{aligned}
$$

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, (A4) and (A6) imply that

$$
\frac{1}{Q_t} = \frac{q_{kt}}{Q_t}\left\{1 - S(\lambda_{It}) - S'(\lambda_{It})\lambda_{It}\right\} + \beta \mathrm{E}_t \frac{q_{k,t+1}}{Q_{t+1}} \frac{A_{t+1}U_{c,t+1}}{A_t U_{ct}} S'(\lambda_{I,t+1})(\lambda_{I,t+1})^2. \quad \text{(A9)}
$$

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 μ_t and μ_{kt}, the capital Euler equation (A7) can be rewritten as

$$
\frac{q_{kt}}{Q_t} = \beta \mathrm{E}_t \frac{A_{t+1}U_{c,t+1}}{A_t U_{ct}} \left[(1-\delta)\frac{q_{k,t+1}}{Q_{t+1}} + r_{k,t+1}u_{t+1} - \frac{a(u_{t+1})}{Q_{t+1}}\right]. \quad \text{(A10)}
$$

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 μ_t, the first-order condition with respect to bond holding can be written as

$$
D_{t,t+1} = \beta \frac{A_{t+1}U_{c,t+1}}{A_t U_{ct}} \frac{\bar{P}_t}{\bar{P}_{t+1}}. \quad \text{(A11)}
$$

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

$$
\frac{1}{R_t} = \beta \mathrm{E}_t \left[\frac{A_{t+1}U_{c,t+1}}{A_t U_{ct}} \frac{\bar{P}_t}{\bar{P}_{t+1}}\right]. \quad \text{(A12)}
$$

In each period t, a fraction ξ_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

$$\mathrm{E}_t \sum_{i=0}^{\infty} \beta^i \xi_w^i A_{t+i}[\log(C_{t+i} - bC_{t+i-1}) - \frac{\psi}{1+\eta} L_{t+i}^d(h)^{1+\eta}] + \tag{A13}$$

$$\mu_{t+i}[W_t(h)\chi_{t,t+i}^w L_{t+i}^d(h) + m_{t+i}], \tag{A14}$$

where the labor demand schedule is given by

$$L_{t+i}^d(h) = \left(\frac{W_t(h)\chi_{t,t+i}^w}{\bar{W}_{t+i}}\right)^{-\theta_{wt}} L_{t+i}, \quad \theta_{wt} = \frac{\mu_{wt}}{\mu_{wt}-1}, \tag{A15}$$

the term m_t is given by

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

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

$$\chi_{t,t+i}^w \equiv \begin{cases} \Pi_{k=1}^i \pi_{t+k-1}^{\gamma_w} \pi^{1-\gamma_w} \lambda_*^i & \text{if } i \geq 1 \\ 1 & \text{if } i = 0. \end{cases} \tag{A16}$$

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

$$\mathrm{E}_t \sum_{i=0}^{\infty} (\beta\xi_w)^i \left\{ -A_{t+i}\psi L_{t+i}^d(h)^\eta \frac{\partial L_{t+i}^d(h)}{\partial W_t(h)} + \mu_{t+i}(1-\theta_{w,t+i})\chi_{t,t+i}^w L_{t+i}^d(h) \right\} = 0, \tag{A17}$$

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

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

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 (A11), we can rewrite the first-order condition for wage setting as

$$\mathrm{E}_t \sum_{i=0}^{\infty} \xi_w^i D_{t,t+i} L_{t+i}^d(h) \frac{1}{\mu_{w,t+i}-1} \left\{ \mu_{w,t+i} MRS_{t+i}(h) - \chi_{t,t+i}^w W_t(h) \right\} = 0. \tag{A18}$$

A.2. **Firms' optimizing decisions.** Pricing decisions are staggered across firms. In each period, a fraction ξ_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

$$P_t(j) = \pi_{t-1}^{\gamma_p}\pi^{1-\gamma_p}P_{t-1}(j), \tag{A19}$$

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

$$\mathrm{Max}_{P_t(j)} \quad \mathrm{E}_t\sum_{i=0}^{\infty}\xi_p^i D_{t,t+i}[P_t(j)\chi_{t,t+i}^p Y_{t+i}^d(j) - V_{t+i}(j)], \tag{A20}$$

subject to

$$Y_{t+i}^d(j) = \left(\frac{P_t(j)\chi_{t,t+i}^p}{\bar{P}_{t+i}}\right)^{-\frac{\mu_{p,t+i}}{\mu_{p,t+i}-1}} Y_{t+i}, \tag{A21}$$

where $V_{t+i}(j)$ is the cost function and the term $\chi_{t,t+i}^p$ comes from the price-updating rule (A19) and is given by

$$\chi_{t,t+i}^p = \begin{cases} \Pi_{k=1}^i \pi_{t+k-1}^{\gamma_p}\pi^{1-\gamma_p} & \text{if } i \geq 1 \\ 1 & \text{if } i = 0. \end{cases} \tag{A22}$$

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

$$\mathrm{E}_t\sum_{i=0}^{\infty}\xi_p^i D_{t,t+i}Y_{t+i}^d(j)\frac{1}{\mu_{p,t+i}-1}\left[\mu_{p,t+i}\Phi_{t+i}(j) - P_t(j)\chi_{t,t+i}^p\right] = 0, \tag{A23}$$

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, μ_{pt} would be a constant and (A23) 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

$$\Phi_t(j) = \left[\frac{\tilde{\alpha}}{Z_t}(\bar{P}_t r_{kt})^{\alpha_1}\left(\frac{\bar{W}_t}{\lambda_z^t}\right)^{\alpha_2}\right]^{\frac{1}{\alpha_1+\alpha_2}} Y_t(j)^{\frac{1}{\alpha_1+\alpha_2}-1}, \tag{A24}$$

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

$$\bar{W}_t = \Phi_t(j)\alpha_2\frac{Y_t(j)}{L_t^f(j)}, \tag{A25}$$

$$\bar{P}_t r_{kt} = \Phi_t(j)\alpha_1\frac{Y_t(j)}{K_t^f(j)}. \tag{A26}$$

It follows that

$$\frac{\bar{W}_t}{\bar{P}_t r_{kt}} = \frac{\alpha_2}{\alpha_1}\frac{K_t^f(j)}{L_t^f(j)}, \quad \forall j \in [0,1]. \tag{A27}$$

A.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

$$C_t + \frac{1}{Q_t}[I_t + a(u_t)K_{t-1}] + G_t = Y_t, \tag{A28}$$

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

$$G_{pt} Y_t = Z_t (u_t K_{t-1})^{\alpha_1} (\lambda_z^t L_t)^{\alpha_2}, \tag{A29}$$

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.

A.4. **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:

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

A.4.1. *Stationary pricing decisions.* In terms of the stationary variables, we can rewrite the optimal pricing decision (A23) as

$$\mathrm{E}_t \sum_{i=0}^{\infty} (\beta \xi_p)^i A_{t+i} \tilde{U}_{c,t+i} \tilde{Y}_{t+i}^d(j) \frac{1}{\mu_{p,t+i} - 1} [\mu_{p,t+i} \phi_{t+i}(j) - p_t^* Z_{t,t+i}^p] = 0. \tag{A30}$$

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

$$Z_{t,t+i}^p = \frac{\chi_{t,t+i}^p}{\prod_{k=1}^i \pi_{t+k}} \tag{A31}$$

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

$$
\begin{aligned}
\phi_{t+i}(j) &= \left[\frac{\tilde{\alpha}}{Z_{t+i}}\left(\frac{\tilde{r}_{k,t+i}}{q_{t+i}\lambda_q^{t+i}}\right)^{\alpha_1}\left(\tilde{w}_{t+i}\frac{\lambda_*^{t+i}}{\lambda_z^{t+i}}\right)^{\alpha_2}\right]^{\frac{1}{\alpha_1+\alpha_2}} Y_{t+i}^d(j)^{\frac{1}{\alpha_1+\alpha_2}-1} \\
&= \left[\frac{\tilde{\alpha}}{Z_{t+i}}\left(\frac{\tilde{r}_{k,t+i}}{q_{t+i}}\right)^{\alpha_1}(\tilde{w}_{t+i})^{\alpha_2}\right]^{\frac{1}{\alpha_1+\alpha_2}} \tilde{Y}_{t+i}^d(j)^{\frac{1}{\alpha_1+\alpha_2}-1}. \qquad \text{(A32)}
\end{aligned}
$$

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

$$
\begin{aligned}
\tilde{Y}_{t+i}^d(j) &= \left[\frac{P_t(j)\chi_{t,t+i}^p}{\bar{P}_{t+i}}\right]^{-\theta_{p,t+i}} \tilde{Y}_{t+i} \\
&= \left[p_t^*\frac{\bar{P}_t}{\bar{P}_{t+i}}\chi_{t,t+i}^p\right]^{-\theta_{p,t+i}} \tilde{Y}_{t+i} \\
&= [p_t^* Z_{t,t+i}^p]^{-\theta_{p,t+i}}\tilde{Y}_{t+i}. \qquad \text{(A33)}
\end{aligned}
$$

Combining (A32) and (A33), we have

$$
\phi_{t+i}(j) = \tilde{\phi}_{t+i}[p_t^* Z_{t,t+i}^p]^{-\theta_{p,t+i}\bar{\alpha}}(\tilde{Y}_{t+i})^{\bar{\alpha}}, \qquad \text{(A34)}
$$

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

$$
\tilde{\phi}_{t+i} \equiv \left[\frac{\tilde{\alpha}}{Z_{t+i}}\left(\frac{\tilde{r}_{k,t+i}}{q_{t+i}}\right)^{\alpha_1}\tilde{w}_{t+i}^{\alpha_2}\right]^{\frac{1}{\alpha_1+\alpha_2}}. \qquad \text{(A35)}
$$

Given these relations, we can rewrite the optimal pricing rule (A30) in terms of stationary variables

$$
\mathrm{E}_t\sum_{i=0}^{\infty}(\beta\xi_p)^i\frac{A_{t+i}\tilde{U}_{c,t+i}\tilde{Y}_{t+i}^d(j)}{\mu_{p,t+i}-1}[\mu_{p,t+i}\tilde{\phi}_{t+i}[p_t^*Z_{t,t+i}^p]^{-\theta_{p,t+i}\bar{\alpha}}(\tilde{Y}_{t+i})^{\bar{\alpha}} - p_t^*Z_{t,t+i}^p] = 0, \qquad \text{(A36)}
$$

where $\tilde{\phi}$ is defined in (A35).

A.4.2. *Stationary wage setting decision.* Using (A4) and (A11), we can rewrite the optimal wage-setting decision (A18) as

$$
\mathrm{E}_t\sum_{i=0}^{\infty}(\beta\xi_w)^i\frac{A_{t+i}U_{c,t+i}}{A_tU_{ct}}\frac{\bar{P}_t}{\bar{P}_{t+i}}L_{t+i}^d(h)\frac{1}{\mu_{w,t+i}-1}[\mu_{w,t+i}\psi\frac{L_{t+i}^d(h)^\eta}{U_{c,t+i}}\bar{P}_{t+i} - W_t(h)\chi_{t,t+i}^w] = 0, \qquad \text{(A37)}
$$

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

$$L^d_{t+i}(h) = \left[\frac{W_t(h)\chi^w_{t,t+i}}{\bar{W}_{t+i}}\right]^{-\theta_{w,t+i}} L_{t+i} \tag{A38}$$

$$= \left[w^*_t \frac{\bar{W}_t}{\bar{W}_{t+i}}\chi^w_{t,t+i}\right]^{-\theta_{w,t+i}} L_{t+i} \tag{A39}$$

$$= \left[w^*_t \frac{\tilde{w}_t \bar{P}_t \lambda^t_*}{\tilde{w}_{t+i}\bar{P}_{t+i}\lambda^{t+i}_*}\chi^w_{t,t+i}\right]^{-\theta_{w,t+i}} L_{t+i} \tag{A40}$$

$$= \left[\frac{w^*_t \tilde{w}_t}{\tilde{w}_{t+i}} \frac{\chi^w_{t,t+i}}{\prod_{k=1}^i \pi_{t+k}\lambda^i_*}\right]^{-\theta_{w,t+i}} L_{t+i} \tag{A41}$$

$$\equiv \left[\frac{w^*_t \tilde{w}_t}{\tilde{w}_{t+i}} Z^w_{t,t+i}\right]^{-\theta_{w,t+i}} L_{t+i}, \tag{A42}$$

with $Z^w_{t,t+i}$ defined as

$$Z^w_{t,t+i} = \frac{\chi^w_{t,t+i}}{\prod_{k=1}^i \pi_{t+k}\lambda^i_*}. \tag{A43}$$

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 (A37) in terms of the stationary variables. With some cancelations, we obtain

$$\mathrm{E}_t \sum_{i=0}^{\infty}\prod_{k=1}^{i}(\beta\xi_w)^i \frac{A_{t+i}\tilde{U}_{c,t+i}L^d_{t+i}(h)}{\mu_{w,t+i}-1}\left\{\mu_{w,t+i}\psi\left[\frac{w^*_t\tilde{w}_t}{\tilde{w}_{t+i}}Z^w_{t,t+i}\right]^{-\eta\theta_{w,t+i}}\frac{L^\eta_{t+i}}{\tilde{U}_{c,t+i}} - w^*_t\tilde{w}_t Z^w_{t,t+i}\right\} = 0. \tag{A44}$$

A.4.3. *Other stationary equilibrium conditions.* We now rewrite the rest of the equilibrium conditions in terms of stationary variables.

First, the optimal investment decision equation (A9) can be written as

$$1 = q_{kt}\left\{1 - S(\lambda_{It}) - S'(\lambda_{It})\lambda_{It}\right\} + \frac{\beta}{\lambda_q\lambda_*}\mathrm{E}_t q_{k,t+1}\frac{q_t}{q_{t+1}}\frac{A_{t+1}\tilde{U}_{c,t+1}}{A_t\tilde{U}_{ct}}S'(\lambda_{I,t+1})(\lambda_{I,t+1})^2, \tag{A45}$$

where

$$\lambda_{It} = \frac{I_t}{I_{t-1}} = \frac{\tilde{I}_t Q_t \lambda^t_*}{\tilde{I}_{t-1}Q_{t-1}\lambda^{t-1}_*} = \frac{\tilde{I}_t q_t}{\tilde{I}_{t-1}q_{t-1}}\lambda_q\lambda_*. \tag{A46}$$

Second, the capital Euler equation (A10) can be written as

$$q_{kt} = \frac{\beta}{\lambda_q\lambda_*}\mathrm{E}_t\frac{A_{t+1}\tilde{U}_{c,t+1}}{A_t\tilde{U}_{ct}}\frac{q_t}{q_{t+1}}\left[(1-\delta)q_{k,t+1} + \tilde{r}_{k,t+1}u_{t+1} - a(u_{t+1})\right]. \tag{A47}$$

Third, the optimal capacity utilization decision (A8) is equivalent to

$$\tilde{r}_{kt} = a'(u_t). \tag{A48}$$

Fourth, the intertemporal bond Euler equation (A12) can be written as

$$\frac{1}{R_t} = \frac{\beta}{\lambda_*} \mathrm{E}_t \left[\frac{A_{t+1}\tilde{U}_{c,t+1}}{A_t \tilde{U}_{ct}} \frac{1}{\pi_{t+1}} \right]. \tag{A49}$$

Fifth, the law of motion for capital stock in (A3) can be written as

$$\tilde{K}_t = (1-\delta)\frac{q_{t-1}\tilde{K}_{t-1}}{q_t \lambda_q \lambda_*} + [1 - S(\lambda_{It})]\tilde{I}_t. \tag{A50}$$

Sixth, the aggregate resource constraint is given by

$$\tilde{C}_t + \tilde{I}_t + \frac{q_{t-1}}{q_t \lambda_q \lambda_*} a(u_t)\tilde{K}_{t-1} + \tilde{G}_t = \tilde{Y}_t. \tag{A51}$$

Seventh, the aggregate production function (A29) can be written as

$$G_{pt}\tilde{Y}_t = Z_t \left[\frac{u_t q_{t-1}\tilde{K}_{t-1}}{\lambda_q \lambda_*} \right]^{\alpha_1} L_t^{\alpha_2}. \tag{A52}$$

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

$$\frac{\tilde{w}_t}{\tilde{r}_{kt}} = \frac{1}{\lambda_q \lambda_*} \frac{\alpha_2}{\alpha_1} \frac{u_t \tilde{K}_{t-1}}{L_t} \frac{q_{t-1}}{q_t}. \tag{A53}$$

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

$$R_t = \kappa R_{t-1}^{\rho_r} \left[\left(\frac{\pi_t}{\pi^*(s_t)} \right)^{\phi_\pi} \tilde{Y}_t^{\phi_y} \right]^{1-\rho_r} e^{\sigma_{rt}\varepsilon_{rt}}. \tag{A54}$$

Appendix B. 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 (A36) reduces to

$$\frac{1}{\mu_p} = \left[\frac{\tilde{\alpha}}{Z} \left(\frac{\tilde{r}_k}{q} \right)^{\alpha_1} \tilde{w}^{\alpha_2} \right]^{\frac{1}{\alpha_1+\alpha_2}} \tilde{Y}^{\bar{\alpha}}. \tag{A55}$$

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 (A44) reduces to

$$\tilde{w} = \mu_w \frac{\psi L^\eta}{\tilde{U}_c}, \tag{A56}$$

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 (A26) for capital input together with the capital market clearing condition imply that

$$\tilde{r}_k = \frac{\alpha_1}{\mu_p} \frac{\tilde{Y} \lambda_q \lambda_*}{\tilde{K}}. \tag{A57}$$

The rest of the steady-state equilibrium conditions for the private sector come from (A45) -(A53) and are summarized below:

$$\begin{aligned}
1 &= q_k, && \text{(A58)}\\
\frac{\lambda_q \lambda_*}{\beta} &= 1 - \delta + \tilde{r}_k, && \text{(A59)}\\
\tilde{r}_k &= a'(1), && \text{(A60)}\\
R &= \frac{\lambda_*}{\beta}\pi, && \text{(A61)}\\
\frac{\tilde{I}}{\tilde{K}} &= 1 - \frac{1-\delta}{\lambda_q \lambda_*}, && \text{(A62)}\\
\tilde{Y} &= \tilde{C} + \tilde{I} + \tilde{G}, && \text{(A63)}\\
\tilde{Y} &= Z \left(\frac{q\tilde{K}}{\lambda_q \lambda_*} \right)^{\alpha_1} L^{\alpha_2}, && \text{(A64)}\\
\frac{\tilde{w}}{\tilde{r}_k} &= \frac{1}{\lambda_q \lambda_*} \frac{\alpha_2}{\alpha_1} \frac{\tilde{K}}{L}. && \text{(A65)}
\end{aligned}$$

Appendix C. Linearized equilibrium conditions

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

C.1. **Linearizing the price setting rule.** Log-linearizing the price setting rule (A36) around the steady state, we get

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

where

$$\hat{\hat{\phi}}_{t+i} = \frac{1}{\alpha_1 + \alpha_2}[\alpha_1(\hat{r}_{k,t+i} - \hat{q}_{t+i}) + \alpha_2\hat{w}_{t+i} - \hat{z}_{t+i}]. \tag{A66}$$

Collecting terms to get

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

Further simplifying

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

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

$$\begin{aligned}\frac{1+\bar{\alpha}\theta_p}{1-\beta\xi_p}\hat{p}^*_t &= \hat{\mu}_{pt} + \hat{mc}_t - (1+\theta_p\bar{\alpha})\hat{Z}^p_{t,t} \\ &+ \beta\xi_p\mathrm{E}_t[\hat{\mu}_{p,t+1} + \hat{mc}_{t+1} - (1+\theta_p\bar{\alpha})\hat{Z}^p_{t,t+1}] \\ &+ (\beta\xi_p)^2\mathrm{E}_t[\hat{\mu}_{p,t+2} + \hat{mc}_{t+2} - (1+\theta_p\bar{\alpha})\hat{Z}^p_{t,t+2}] + \ldots\end{aligned}$$

Forwarding this relation one period to get

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

Moving the $Z^p_{t,t+i}$ terms to the left, we have

$$\begin{aligned}
&\frac{1+\bar{\alpha}\theta_p}{1-\beta\xi_p}\hat{p}^*_t + (1+\bar{\alpha}\theta_p)\mathrm{E}_t[\hat{Z}^p_{t,t} + \beta\xi_p\hat{Z}^p_{t,t+1} + ...] = \hat{\mu}_{pt} + \hat{mc}_t \\
&\quad +\beta\xi_p\mathrm{E}_t[\hat{\mu}_{p,t+1} + \hat{mc}_{t+1}] \\
&\quad +(\beta\xi_p)^2\mathrm{E}_t[\hat{\mu}_{p,t+2} + \hat{mc}_{t+2}] + \ldots \\
&\quad = \hat{\mu}_{pt} + \hat{mc}_t \\
&\quad +\beta\xi_p\left[\frac{1+\bar{\alpha}\theta_p}{1-\beta\xi_p}\mathrm{E}_t\hat{p}^*_{t+1} + (1+\bar{\alpha}\theta_p)\mathrm{E}_t[\hat{Z}^p_{t+1,t+1} + \beta\xi_p\hat{Z}^p_{t+1,t+2} + ...]\right],
\end{aligned}$$

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

$$\begin{aligned}
\frac{1+\bar{\alpha}\theta_p}{1-\beta\xi_p}\hat{p}^*_t &= \hat{\mu}_{pt} + \hat{mc}_t + \beta\xi_p\frac{1+\bar{\alpha}\theta_p}{1-\beta\xi_p}\mathrm{E}_t[\hat{p}^*_{t+1} \\
&+ (1+\bar{\alpha}\theta_p)\beta\xi_p\mathrm{E}_t\sum_{i=0}^{\infty}(\beta\xi_p)^i[\hat{Z}^p_{t+1,t+i+1} - \hat{Z}^p_{t,t+i+1}]. \qquad \text{(A67)}
\end{aligned}$$

Using the definition of $Z^p_{t,t+i}$ in (A31), we obtain

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

Thus,

$$\hat{Z}^p_{t+1,t+i+1} - \hat{Z}^p_{t,t+i+1} = \hat{\pi}_{t+1} - \gamma_{pt}\hat{\pi}_t,$$

and the Z^p terms in (A67) can be reduced to

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

Substituting this result into (A67), we obtain

$$\hat{p}^*_t = \frac{1-\beta\xi_p}{1+\bar{\alpha}\theta_p}(\hat{\mu}_{pt} + \hat{mc}_t) + \beta\xi_p\mathrm{E}_t\hat{p}^*_{t+1} + \beta\xi_p\mathrm{E}_t[\hat{\pi}_{t+1} - \gamma_{pt}\hat{\pi}_t]. \qquad \text{(A68)}$$

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

$$1 = \xi_p\left[\frac{1}{\pi_t}\pi^{\gamma_p}_{t-1}\pi^{1-\gamma_p}\right]^{\frac{1}{1-\mu_{pt}}} + (1-\xi_p)(p^*_t)^{\frac{1}{1-\mu_{pt}}}, \qquad \text{(A69)}$$

the linearized version of which is given by

$$\hat{p}^*_t = \frac{\xi_p}{1-\xi_p}(\hat{\pi}_t - \gamma_p\hat{\pi}_{t-1}). \qquad \text{(A70)}$$

Using (A70) to substitute out the $\hat{p}_t^*$ in (A68), we obtain

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

or

$$\hat{\pi}_t - \gamma_p \hat{\pi}_{t-1} = \frac{\kappa_p}{1+\bar{\alpha}\theta_p}(\hat{\mu}_{pt} + \hat{mc}_t) + \beta \mathrm{E}_t[\hat{\pi}_{t+1} - \gamma_p \hat{\pi}_t], \tag{A71}$$

where the real marginal cost is given by

$$\hat{mc}_t = \frac{1}{\alpha_1+\alpha_2}[\alpha_1(\hat{r}_{k,t+i} - \hat{q}_{t+i}) + \alpha_2 \hat{w}_{t+i} - \hat{z}_{t+i}] + \bar{\alpha}\hat{y}_t. \tag{A72}$$

and the term κ_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.

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

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

Collecting terms to get

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

Further simplifying

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

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

$$\begin{aligned}\frac{1+\eta\theta_w}{1-\beta\xi_w}(\hat{w}_t^* + \hat{w}_t) &= \hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t + (1+\eta\theta_w)(\hat{w}_t - \hat{Z}_{t,t}^w) \\ &+ \beta\xi_w \mathrm{E}_t[\hat{\mu}_{w,t+1} + \hat{mrs}_{t+1} - \hat{w}_{t+1} + (1+\eta\theta_w)(\hat{w}_{t+1} - \hat{Z}_{t,t+1}^w)] \\ &+ (\beta\xi_w)^2 \mathrm{E}_t[\hat{\mu}_{w,t+2} + \hat{mrs}_{t+2} - \hat{w}_{t+2} + (1+\eta\theta_w)(\hat{w}_{t+2} - \hat{Z}_{t,t+2}^w)] + \ldots\end{aligned}$$

Forwarding this relation one period to get

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

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

$$\begin{aligned}&\frac{1+\eta\theta_w}{1-\beta\xi_w}(\hat{w}_t^* + \hat{w}_t) + (1+\eta\theta_w)\mathrm{E}_t[\hat{Z}_{t,t}^w + \beta\xi_w\hat{Z}_{t,t+1}^w + ...] = \hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t + (1+\eta\theta_w)\hat{w}_t \\ &\quad +\beta\xi_w\mathrm{E}_t[\hat{\mu}_{w,t+1} + \hat{mrs}_{t+1} - \hat{w}_{t+1} + (1+\eta\theta_w)\hat{w}_{t+1}] \\ &\quad +(\beta\xi_w)^2\mathrm{E}_t[\hat{\mu}_{w,t+2} + \hat{mrs}_{t+2} - \hat{w}_{t+2} + (1+\eta\theta_w)\hat{w}_{t+2}] + \ldots \\ &\quad = \hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t + (1+\eta\theta_w)\hat{w}_t \\ &\quad +\beta\xi_w\mathrm{E}_t\left[\frac{1+\eta\theta_w}{1-\beta\xi_w}(\hat{w}_{t+1}^* + \hat{w}_{t+1}) + (1+\eta\theta_w)[\hat{Z}_{t+1,t+1}^w + \beta\xi_w\hat{Z}_{t+1,t+2}^w + ...]\right],\end{aligned}$$

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

$$\begin{aligned}\frac{1+\eta\theta_w}{1-\beta\xi_w}(\hat{w}_t^* + \hat{w}_t) &= \hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t + (1+\eta\theta_w)\hat{w}_t + \beta\xi_w\frac{1+\eta\theta_w}{1-\beta\xi_w}\mathrm{E}_t(\hat{w}_{t+1}^* + \hat{w}_{t+1}) \\ &+ (1+\eta\theta_w)\beta\xi_w\mathrm{E}_t\sum_{i=0}^{\infty}(\beta\xi_w)^i[\hat{Z}_{t+1,t+i+1}^w - \hat{Z}_{t,t+i+1}^w]. \qquad \text{(A73)}\end{aligned}$$

Using the definition of $Z_{t,t+i}^w$ in (A43), we obtain

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

Thus,

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

and the Z^w terms in (A73) can be reduced to

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

Substituting this result into (A73), we obtain

$$\hat{w}_t^* + \hat{w}_t = \frac{1-\beta\xi_w}{1+\eta\theta_w}(\hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t) + (1-\beta\xi_w)\hat{w}_t + \beta\xi_w \mathrm{E}_t(\hat{w}_{t+1}^* + \hat{w}_{t+1}) + \beta\xi_w \mathrm{E}_t[\hat{\pi}_{t+1} - \gamma_{wt}\hat{\pi}_t]. \tag{A74}$$

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

$$1 = \xi_w \left[\frac{\tilde{w}_{t-1}}{\tilde{w}_t}\frac{1}{\pi_t}\pi_{t-1}^{\gamma_w}\pi^{1-\gamma_w}\right]^{\frac{1}{1-\mu_{wt}}} + (1-\xi_w)(w_t^*)^{\frac{1}{1-\mu_{wt}}}, \tag{A75}$$

the linearized version of which is given by

$$\hat{w}_t^* = \frac{\xi_w}{1-\xi_w}(\hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t - \gamma_w \hat{\pi}_{t-1})]. \tag{A76}$$

Using (A76) to substitute out the $\hat{w}_t^*$ in (A74), we obtain

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

or

$$\begin{aligned}
\hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t - \gamma_w \hat{\pi}_{t-1} = \frac{\kappa_w}{1+\eta\theta_w}(\hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t) + \\
\beta \mathrm{E}_t[\hat{w}_{t+1} - \hat{w}_t + \hat{\pi}_{t+1} - \gamma_w \hat{\pi}_t],
\end{aligned} \tag{A77}$$

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

$$\pi_t^w = \frac{\bar{W}_t}{\bar{W}_{t-1}} = \frac{\tilde{w}_t \bar{P}_t \lambda_*^t}{\tilde{w}_{t-1}\bar{P}_{t-1}\lambda_*^{t-1}} = \frac{\tilde{w}_t}{\tilde{w}_{t-1}}\pi_t \lambda_*. \tag{A78}$$

The log-linearized version is given by

$$\hat{\pi}_t^w = \hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t.$$

Thus, the optimal wage decision (A77) is equivalent to

$$\hat{\pi}_t^w - \gamma_w \hat{\pi}_{t-1} = \frac{\kappa_w}{1+\eta\theta_w}(\hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t) + \beta \mathrm{E}_t(\hat{\pi}_{t+1}^w - \gamma_w \hat{\pi}_t). \tag{A79}$$

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

C.3. **Linearizing other stationary equilibrium conditions.** Taking total differentiation in the investment decision equation (A45) and using the steady-state conditions that $S(\lambda_I) = S'(\lambda_I) = 0$ and that $\lambda_I = \lambda_q \lambda_*$, we obtain

$$\hat{q}_{kt} = S''(\lambda_I)\lambda_I^2 \left[\hat{\lambda}_{It} - \beta \mathrm{E}_t \hat{\lambda}_{I,t+1}\right], \tag{A80}$$

where, from (A46), we have

$$\hat{\lambda}_{It} = \hat{i}_t - \hat{i}_{t-1} + \hat{q}_t - \hat{q}_{t-1}. \tag{A81}$$

Taking total differentiation in the capital Euler equation (A47) 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

$$\hat{q}_{kt} = \mathrm{E}_t \left\{ \Delta \hat{a}_{t+1} + \Delta \hat{U}_{c,t+1} - \Delta \hat{q}_{t+1} + \frac{\beta}{\lambda_I}\left[(1-\delta)\hat{q}_{k,t+1} + \tilde{r}_k \hat{r}_{k,t+1}\right] \right\}, \tag{A82}$$

where $\Delta x_{t+1} \equiv x_{t+1} - x_t$ denotes the growth rate of the variable x.

The linearized capacity utilization decision equation (A48) is given by

$$\hat{r}_{kt} = \sigma_u \hat{u}_t, \tag{A83}$$

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 (A49) is given by

$$0 = \mathrm{E}_t \left[\Delta \hat{a}_{t+1} + \Delta \hat{U}_{c,t+1} + \hat{R}_t - \hat{\pi}_{t+1}\right]. \tag{A84}$$

Log-linearize the capital law of motion (A50) leads to

$$\hat{k}_t = \frac{1-\delta}{\lambda_q \lambda_*}[\hat{k}_{t-1} + \hat{q}_{t-1} - \hat{q}_t] + \frac{\tilde{I}}{\tilde{K}}\hat{i}_t. \tag{A85}$$

To obtain the linearized resource constraint, we take total differentiation of (A51) to obtain

$$\hat{y}_t = c_y \hat{c}_t + i_y \hat{i}_t + u_y \hat{u}_t + g_y \hat{g}_t, \tag{A86}$$

where $c_y = \frac{\tilde{C}}{\tilde{Y}}$, $i_y = \frac{\tilde{I}}{\tilde{Y}}$, $u_y = \frac{1}{\lambda_q \lambda_*}\frac{\tilde{r}_k \tilde{K}}{\tilde{Y}}$, and $g_y = \frac{\tilde{G}}{\tilde{Y}}$.

Log-linearizing the aggregate production, we get

$$\hat{y}_t = \hat{z}_t + \alpha_1[\hat{k}_{t-1} + \hat{u}_t + \hat{q}_{t-1}] + \alpha_2 \hat{l}_t. \tag{A87}$$

The linearized version of the factor demand relation (A53) is given by

$$\hat{w}_t - \hat{r}_{kt} = \hat{k}_{t-1} + \hat{u}_t + \hat{q}_{t-1} - \hat{q}_t - \hat{l}_t. \tag{A88}$$

Finally, linearizing the interest rate rule (A54) gives

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1-\rho_r)\left[\phi_\pi(\hat{\pi}_t - \hat{\pi}^*(s_t)) + \phi_y \hat{y}_t\right] + \sigma_{rt}\varepsilon_{rt}, \tag{A89}$$

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:

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

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

C.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.

$$\hat{\pi}_t - \gamma_p\hat{\pi}_{t-1} = \frac{\kappa_p}{1+\bar{\alpha}\theta_p}(\hat{\mu}_{pt} + \hat{mc}_t) + \beta\mathrm{E}_t[\hat{\pi}_{t+1} - \gamma_p\hat{\pi}_t], \tag{A90}$$

$$\begin{aligned} \hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t - \gamma_w\hat{\pi}_{t-1} = {} & \frac{\kappa_w}{1+\eta\theta_w}(\hat{\mu}_{wt} + \hat{mrs}_t - \hat{w}_t) + \\ & \beta\mathrm{E}_t[\hat{w}_{t+1} - \hat{w}_t + \hat{\pi}_{t+1} - \gamma_w\hat{\pi}_t]. \end{aligned} \tag{A91}$$

$$\hat{q}_{kt} = S''(\lambda_I)\lambda_I^2\left[\Delta\hat{i}_t + \Delta\hat{q}_t - \beta\mathrm{E}_t(\Delta\hat{i}_{t+1} + \Delta\hat{q}_{t+1})\right], \tag{A92}$$

$$\hat{q}_{kt} = \mathrm{E}_t\left\{\Delta\hat{a}_{t+1} + \Delta\hat{U}_{c,t+1} - \Delta\hat{q}_{t+1} + \frac{\beta}{\lambda_I}\left[(1-\delta)\hat{q}_{k,t+1} + \tilde{r}_k\hat{r}_{k,t+1}\right]\right\}, \tag{A93}$$

$$\hat{r}_{kt} = \sigma_u\hat{u}_t, \tag{A94}$$

$$0 = \mathrm{E}_t\left[\Delta\hat{a}_{t+1} + \Delta\hat{U}_{c,t+1} + \hat{R}_t - \hat{\pi}_{t+1}\right], \tag{A95}$$

$$\hat{k}_t = \frac{1-\delta}{\lambda_I}[\hat{k}_{t-1} + \hat{q}_{t-1} - \hat{q}_t] + \left(1 - \frac{1-\delta}{\lambda_I}\right)\hat{i}_t, \tag{A96}$$

$$\hat{y}_t = c_y\hat{c}_t + i_y\hat{i}_t + u_y\hat{u}_t + g_y\hat{g}_t, \tag{A97}$$

$$\hat{y}_t = \hat{z}_t + \alpha_1[\hat{k}_{t-1} + \hat{u}_t + \hat{q}_{t-1}] + \alpha_2\hat{l}_t, \tag{A98}$$

$$\hat{w}_t - \hat{r}_{kt} = \hat{k}_{t-1} + \hat{u}_t + \hat{q}_{t-1} - \hat{q}_t - \hat{l}_t, \tag{A99}$$

$$\hat{R}_t = \rho_r\hat{R}_{t-1} + (1-\rho_r)\left[\phi_\pi(\hat{\pi}_t - \hat{\pi}^*(s^t)) + \phi_y\hat{y}_t\right] + \sigma_{rt}\varepsilon_{rt}, \tag{A100}$$

where

$$\hat{mc}_t = \frac{1}{\alpha_1+\alpha_2}[\alpha_1(\hat{r}_{kt}-\hat{q}_t)+\alpha_2\hat{w}_t-\hat{z}_t]+\bar{\alpha}\hat{y}_t, \tag{A101}$$

$$\hat{mrs}_t = \eta\hat{l}_t-\hat{U}_{ct}, \tag{A102}$$

$$\hat{U}_{ct} = \frac{\beta b(1-\rho_a)}{\lambda_*-\beta b}\hat{a}_t-\frac{\lambda_*}{(\lambda_*-b)(\lambda_*-\beta b)}(\lambda_*\hat{c}_t-b\hat{c}_{t-1}) + \frac{\beta b}{(\lambda_*-b)(\lambda_*-\beta b)}(\lambda_*\mathrm{E}_t\hat{c}_{t+1}-b\hat{c}_t), \tag{A103}$$

$$\hat{\pi}^*(s_t) = [\hat{\pi}^*(1),\ \hat{\pi}^*(2)]e_{s_t}, \quad e_{s_t}=Qe_{s_{t-1}}+v_t, \tag{A104}$$

$$\tag{A105}$$

and the steady-state variables are given by

$$\tilde{r}_k = \frac{\lambda_I}{\beta}-(1-\delta), \tag{A106}$$

$$u_y \equiv \frac{\tilde{r}_k\tilde{K}}{\tilde{Y}\lambda_I}=\frac{\alpha_1}{\mu_p}, \tag{A107}$$

$$i_y = [\lambda_I-(1-\delta)]\frac{\alpha_1}{\mu_p\tilde{r}_k}, \tag{A108}$$

$$c_y = 1-i_y-g_y, \tag{A109}$$

with $\lambda_I \equiv \lambda_q\lambda_*$ and g_y calibrated to match the average ratio of government spending to real GDP.

To compute the equilibrium, we eliminate $\hat{u}_t$ by using (A97), leaving 10 equations (A90)-(A96) and (A98)-(A100) 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.

Table 1. Prior Distribution and Posterior Mode of Structural Parameters

	Prior			Posterior		
Parameter	Distribution	Mean	Std	Model I	Model II	Model III
b	Beta	0.7	0.2	0.994	0.939	
η	Gamma	2.0	0.5	2.069	1.954	
$100\left(\frac{1}{\beta}-1\right)$	Beta	0.4	0.2	0.476	0.149	
α_1	Beta	0.2	0.1	0.203	0.194	
α_2	Beta	0.467	0.1	0.795	0.806	
$100(\lambda_q-1)$	Gamma	0.3	0.1	0.310	0.234	
$100(\lambda_*-1)$	Gamma	0.5	0.1	0.428	0.386	
σ_u	Gamma	1.5	0.5	2.045	1.992	
S''	Gamma	4.0	1.5	3.282	3.341	
μ_p-1	Gamma	0.2	0.15	0.0001	0.019	
μ_w-1	Gamma	0.2	0.15	0.425	0.374	
ξ_p	Beta	0.75	0.1	0.919	0.858	
ξ_w	Beta	0.75	0.1	0.886	0.750	
γ_p	Beta	0.5	0.2	0.179	0.305	
γ_w	Beta	0.5	0.2	0.535	0.455	
ρ_r	Beta	0.6	0.2	0.939	0.799	
ϕ_π	Normal	2.0	1.0	1.444	1.656	
ϕ_y	Gamma	0.4	0.25	0.591	0.035	
$100(\pi^*(1)-1)$	Gamma	1.0	0.5	0.545	0.586	
$100(\pi^*(2)-1)$	Gamma	1.0	0.5	0.545	0.586	

Note: Model I: no regime shifts; Model II: regime shifts in shock volatilities; Model III: regime shifts in both shock volatilities and the inflation target.

Table 2. Prior Distribution and Posterior Mode of Shock Parameters

	Prior			Posterior		
Parameter	Distribution	Mean	Std	Model I	Model II	Model III
ρ_p	Beta	0.5	0.2	0.962	0.932	
ϕ_p	Beta	0.5	0.2	0.666	0.832	
ρ_w	Beta	0.5	0.2	0.940	0.974	
ϕ_w	Beta	0.5	0.2	0.651	0.892	
ρ_g	Beta	0.5	0.2	0.999	0.999	
ρ_{gz}	Beta	0.5	0.2	0.585	0.726	
ρ_a	Beta	0.5	0.2	0.576	0.227	
ρ_q	Beta	0.5	0.2	0.911	0.947	
ρ_z	Beta	0.5	0.2	0.999	0.999	
$\sigma_r(1)$	Inverse Gamma	$*$	$*$	0.0028	0.783	
$\sigma_p(1)$	Inverse Gamma	$*$	$*$	0.9001	0.860	
$\sigma_w(1)$	Inverse Gamma	$*$	$*$	1.0506	0.941	
$\sigma_g(1)$	Inverse Gamma	$*$	$*$	0.0274	0.813	
$\sigma_a(1)$	Inverse Gamma	$*$	$*$	0.4311	0.828	
$\sigma_q(1)$	Inverse Gamma	$*$	$*$	0.0511	0.793	
$\sigma_z(1)$	Inverse Gamma	$*$	$*$	0.0082	0.143	
$\sigma_r(2)$	Inverse Gamma	$*$	$*$	0.0028	0.003	
$\sigma_p(2)$	Inverse Gamma	$*$	$*$	0.9001	0.076	
$\sigma_w(2)$	Inverse Gamma	$*$	$*$	1.0506	0.510	
$\sigma_g(2)$	Inverse Gamma	$*$	$*$	0.0274	0.018	
$\sigma_a(2)$	Inverse Gamma	$*$	$*$	0.4311	0.060	
$\sigma_q(2)$	Inverse Gamma	$*$	$*$	0.0511	0.015	
$\sigma_z(2)$	Inverse Gamma	$*$	$*$	0.0082	0.007	

Table 3. Forecast Error Variance Decomposition

Regime I							
Output							
Horizon	MPolicy	PMarkup	WMarkup	GSpend	Productivity	Preference	Invest.
4Q	91.6462	0.2302	0.0168	1.6402	1.8317	0.0767	4.5582
8Q	75.0992	0.4912	0.0747	1.8678	7.0822	0.0188	15.3662
12Q	34.7132	0.7972	0.2567	4.2389	23.6818	0.0049	36.3073
16Q	2.6804	0.6372	0.4509	7.4381	44.0213	0.0012	44.7708
Inflation							
Horizon	MPolicy	PMarkup	WMarkup	GSpend	Productivity	Preference	Invest.
4Q	93.4191	0.2390	0.0480	0.0038	3.8626	0.0587	2.3688
8Q	86.3970	0.0789	0.1193	0.0416	6.5481	0.0687	6.7464
12Q	28.4258	0.3483	0.9855	1.3849	50.6854	0.3329	17.8371
16Q	40.0244	0.5365	0.9725	2.1447	56.0087	0.2844	0.0288

Regime II							
Output							
Horizon	MPolicy	PMarkup	WMarkup	GSpend	Productivity	Preference	Invest.
4Q	8.1943	12.4563	33.7673	5.5876	26.5742	2.7503	10.6700
8Q	2.0384	8.0674	45.6472	1.9317	31.1916	0.2042	10.9195
12Q	0.3085	4.2872	51.3534	1.4354	34.1504	0.0173	8.4477
16Q	0.0140	2.0147	53.0391	1.4809	37.3242	0.0024	6.1247
Inflation							
Horizon	MPolicy	PMarkup	WMarkup	GSpend	Productivity	Preference	Invest.
4Q	4.5989	7.1183	53.2097	0.0072	30.8540	1.1591	3.0529
8Q	2.1133	1.1681	65.6975	0.0387	25.9881	0.6741	4.3201
12Q	0.0908	0.6734	70.8722	0.1686	26.2763	0.4267	1.4920
16Q	0.1269	1.0292	69.4064	0.2591	28.8143	0.3617	0.0024

Table 4. Marginal Likelihood of VAR Model vs. DSGE Models

Model	Log-likelihood
VAR(4)	4495.00
DSGE Model I	4759.07
DSGE Model II	4911.38
DSGE Model III	4908.40

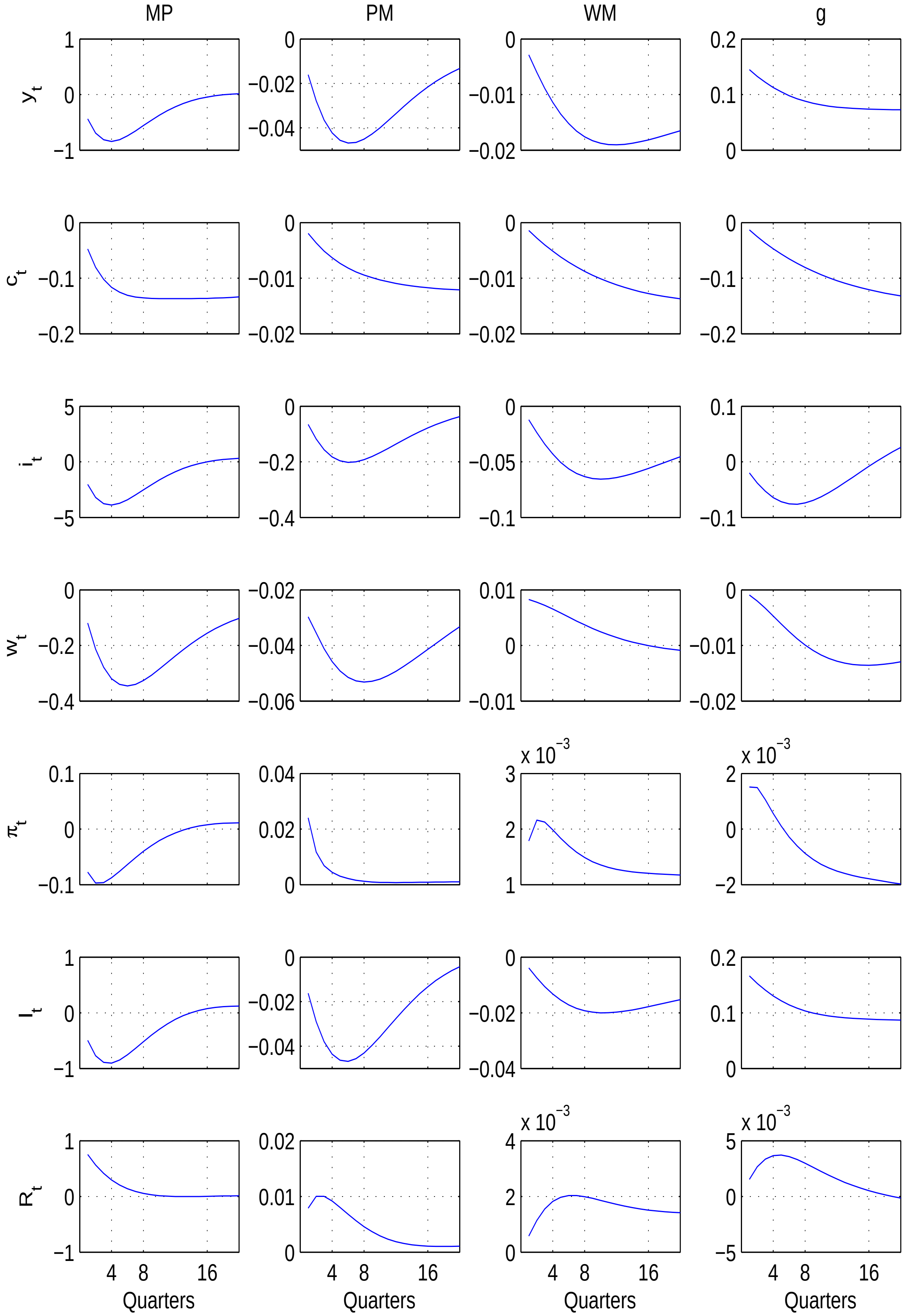

FIGURE 1. Impulse responses in Model II under the first regime: Part a.

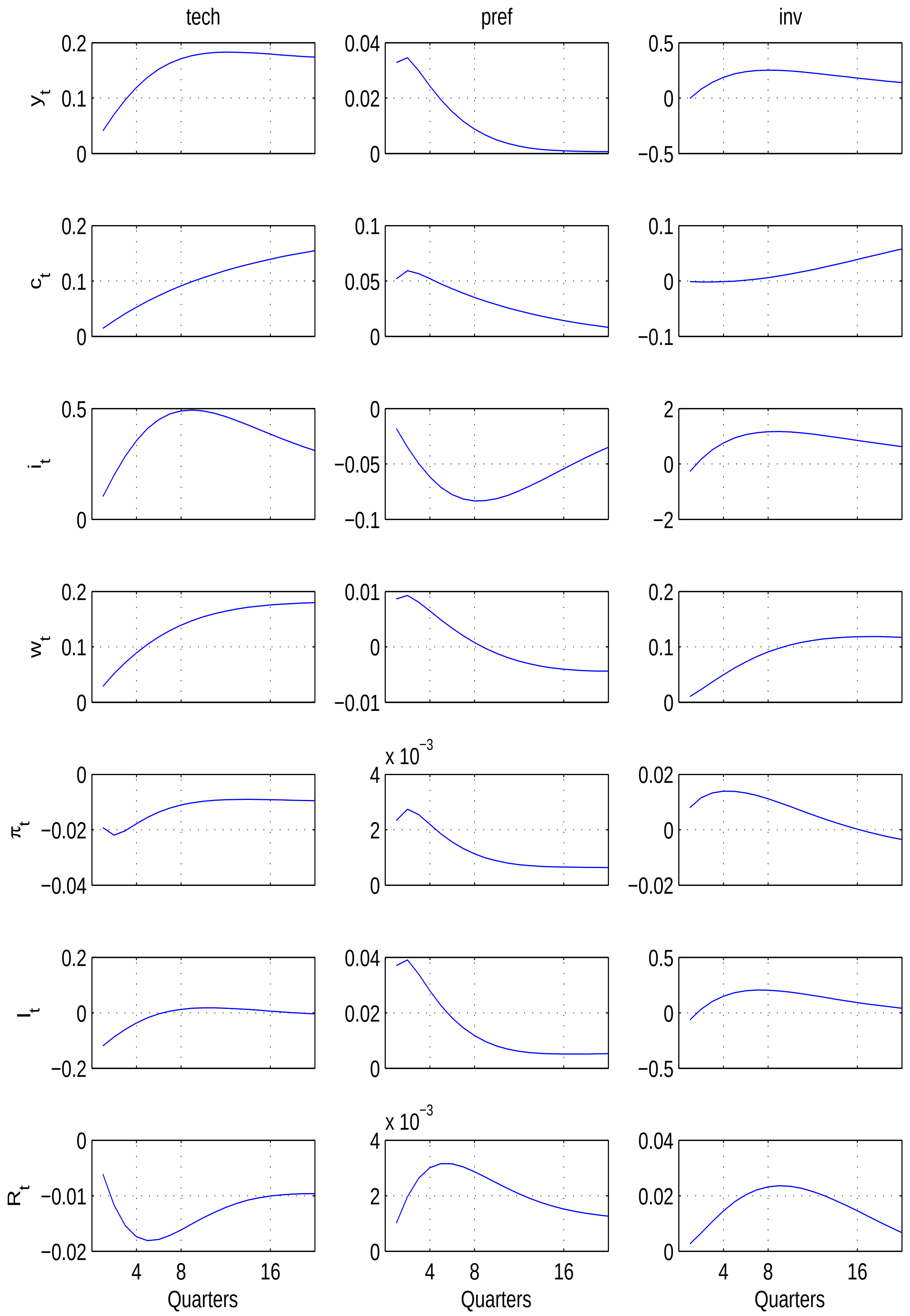

FIGURE 2. Impulse responses in Model II under the first regime: Part b.

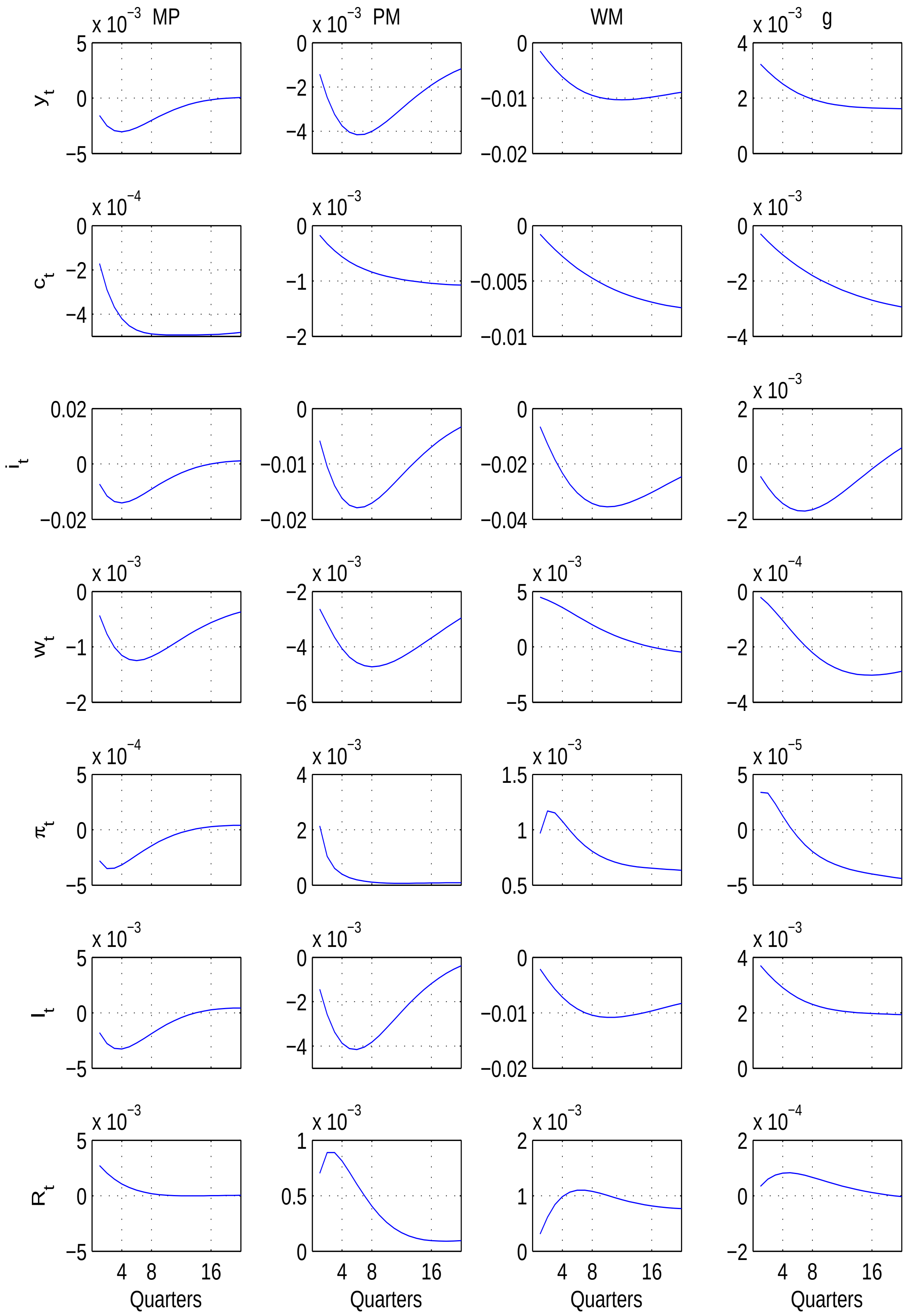

FIGURE 3. Impulse responses in Model II under the second regime: Part a.

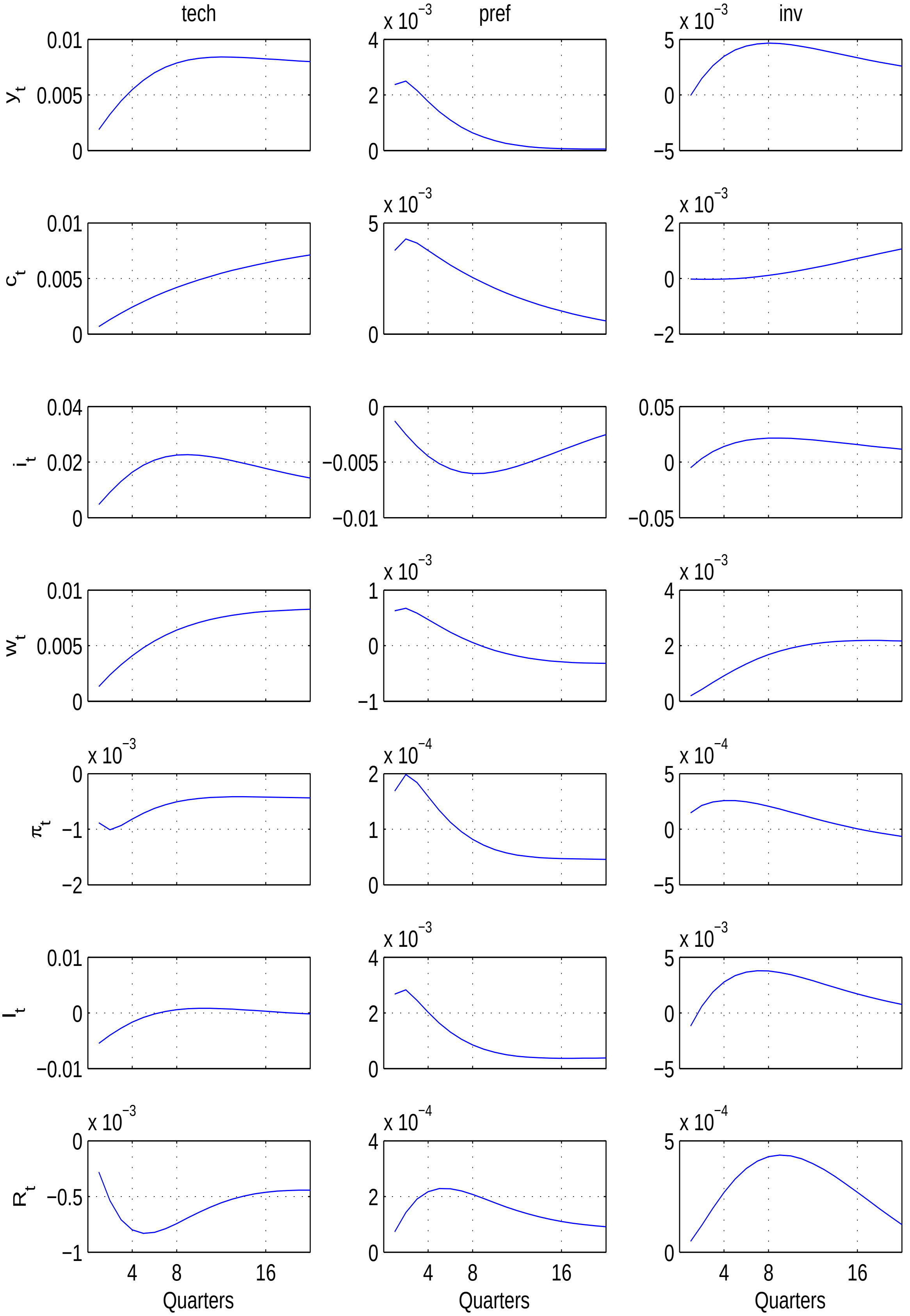

FIGURE 4. Impulse responses in Model II under the second regime: Part b.

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., J. G. Fernald, and M. S. Kimball (2006): "Are Technology Improvements Contractionary?," *American Economic Review*, 96(5), 1418–1448.

Beyer, A., and R. E. A. Farmer (2005): "Identifying the Monetary Transmission Mechanism using Structural Breaks," Manuscript, UCLA.

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(3), 383–398.

Canova, F., and L. Gambetti (2004): "Bad luck or bad policy? On the time variations of US monetary policy," Manuscript, IGIER and Universítá 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.

Chib, S. (1996): "Calculating Posterior Distributions and Model Estimates in Markov Mixture Models," *Journal of Econometrics*, 75, 79–97.

Choi, I. (2002): "Structural Change and Seemingly Unidentified Structural Equations," *Econometric Theory*, 18, 744–775.

Christiano, L., M. Eichenbaum, and C. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Clarida, R., J. Galí, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115(1), 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(2), 262–302.

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.

Erceg, C. J., and A. T. Levin (2003): "Imperfect Credibility andn Inflation Persistence," *Journal of Monetary Economics*, 50, 915–944.

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-Villaverde, J., and J. Rubio-Ramirez (Forthcoming): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*.

Fisher, J. D. M. (2006): "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks," *Journal of Political Economy*, 114(3), 413–451.

Galí, J. (1999): "Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?," *American Economic Review*, 89(1), 249–271.

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.

Geweke, J. (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication," *Econometric Reviews*, 18(1), 1–73.

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

Guerron-Quintana, P. A. (2007): "What You Match Does Matter: The Effects of Data on DSGE Estimation," Manuscript, North Carolina State University.

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 Analsis*. 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.

Kim, C.-J., and C. R. Nelson (1999): *State-Space Models with Regime Switching*. MIT Press, London, England and Cambridge, Massachusetts.

LEEPER, E. M., AND T. ZHA (2003): "Modest Policy Interventions," *Journal of Monetary Economics*, 50(8), 1673–1700.

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., D. F. WAGGONER, AND T. ZHA (2007): "Asymmetric Expectation Effects of Regime Shifts and the Great Moderation," Manuscript.

LUBIK, T. A., AND F. SCHORFHEIDE (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94(1), 190–219.

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.

SARGENT, T. J., N. WILLIAMS, AND T. ZHA (2006): "Shocks and Government Beliefs: The Rise and Fall of American Inflation," *American Economic Review*, 96(4), 1193–1224.

SCHORFHEIDE, F. (2005): "Learning and Monetary Policy Shifts," *Review of Economic Dynamics*, 8(2), 392–419.

SIMS, C. A., D. F. WAGGONER, AND T. ZHA (2006): "Methods for Inference in Large Multiple-Equation Markov-Switching Models," Federal Reserve Bank of Atlanta Working Paper 2006-22.

SIMS, C. A., AND T. ZHA (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, 39(4), 949–968.

——— (2006): "Were There Regime Switches in U.S. Monetary Policy?," *American Economic Review*, 96(1), 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.

WAGGONER, D. F., AND T. ZHA (2003): "A Gibbs Sampler for Structural Vector Autoregressions," *Journal of Economic Dynamics and Control*, 28(2), 349–366.

Woodford, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press, Princeton, NJ.

Emory University, Federal Reserve Bank of Atlanta, Federal Reserve Bank of Atlanta