Technical Appendix to "Monetary Policy Regime Shifts and Inflation Persistence"

Troy Davig and Taeyoung Doh

September 19, 2011

This appendix consists of three parts providing the technical details related to the model description and estimation. In the first part, we discuss determinacy restrictions of the baseline model in the paper. In the second part, we describe the derivation of equilibrium restrictions in an alternative specification with habit formation, a modified price adjustment cost function, and policy inertia in the Taylor rule. We also provide estimation results of regime-switching models with this alternative specification. Finally, we discuss convergence diagnostics of our MCMC output and the computational details of log-marginal data density in our estimation of the baseline model in the paper.

1 Determinacy Restrictions

The MSNK model is inherently nonlinear, which complicates the conditions for determinacy. Following Davig and Leeper(2007), we obtain the linear representation of the MSNK model and restrict our attention to the unique and bounded solution in the linear representation. For the illustrative purpose, first assume all shocks are i.i.d. and there are only two regimes. Then we can rewrite expectations as follows:

$$E_t\pi_{t+1} = E[\pi_{t+1}\,|s_t = i, \Omega_t^{-s}] = p_{i1}E[\pi_{1t+1}\,|\Omega_t^{-s}] + p_{i2}E[\pi_{2t+1}\,|\Omega_t^{-s}], \tag{1}$$

$$E_t y_{t+1} = E[y_{t+1}\,|s_t = i, \Omega_t^{-s}] = p_{i1}E[y_{1t+1}\,|\Omega_t^{-s}] + p_{i2}E[y_{2t+1}\,|\Omega_t^{-s}], \tag{2}$$

where $\pi_{it} = \pi_t(s_t = i, \varepsilon_t)$, $y_{it} = y_t(s_t = i, \varepsilon_t)$, and $\varepsilon_t = \begin{bmatrix} \widehat{a}_t & \widehat{u}_t & \widehat{e}_t \end{bmatrix}'$ for $i = 1, 2$.[1] The information set, $\Omega_t^{-s} = \; s_{t-1}, \ldots$, excludes the current regime, so $\Omega_t = \Omega_t^{-s} \cup \{s_t\}$. Distributing probability mass across the different conditional expectations for inflation, as in $(1) - (2)$, is the same approach as in Gordon and St-Amour (2000) and Bansal and Zhou(2002).

Next, define the forecast errors

$$\eta^{\pi}_{jt+1} = \pi_{jt+1} - E[\pi_{jt+1}\,|\Omega_t^{-s}], \tag{3}$$

$$\eta^{x}_{jt+1} = x_{jt+1} - E[x_{jt+1}\,|\Omega_t^{-s}], \tag{4}$$

for $j = 1, 2$. Substituting expectations, $(1) - (2)$, and the policy rule, shock processes into the equilibrium conditions of the model yields the stacked system

$$AY_t = BY_{t-1} + A\eta_t + Cu_t, \tag{5}$$

[1] Whether shocks are i.i.d. or serially correlated does not matter for determinacy, so is made in this section for convenience.

where

$$Y_t = \begin{bmatrix} \pi_{1t} \\ \pi_{2t} \\ x_{1t} \\ x_{2t} \end{bmatrix}, \qquad \eta_t = \begin{bmatrix} \eta_{1t}^{\pi} \\ \eta_{2t}^{\pi} \\ \eta_{1t}^{x} \\ \eta_{2t}^{x} \end{bmatrix}, \qquad u_t = \begin{bmatrix} \varepsilon_{at} \\ \varepsilon_{ut} \\ \varepsilon_{et} \end{bmatrix}, \tag{6}$$

and A, B and C are conforming matrices consisting of private sector parameters, policy parameters and the transition matrix. The stacked system has constant coefficient matrices, yet captures the impact potential regime changes in monetary policy have on expectation formation. Further, standard methods for solving linear rational expectations systems are applicable to (5), such as Blanchard and Kahn(1980).[2]

Necessary and sufficient conditions for determinacy, which is the existence of a unique and bounded solution to (5), is that all the generalized eigenvalues of (B, A) lie inside the unit circle. The determinacy conditions are intuitive. First, the passive monetary regime cannot be too passive, meaning the response to inflation can be less than one, but still has to be above some minimum threshold. And second, the passive regime cannot be too persistent, meaning that the expected duration of the regime must be below a given threshold. The determinacy conditions are joint restrictions over both monetary regimes, so the parameters governing the active regime affect the determinacy restrictions over the passive regime. Therefore, the more persistent or active the active regime is, the more persistent or passive the passive regime can be.

One caveat in using the linear representation to check the determinacy conditions is that there might be non-MSV solutions of the MSNK model even if determinacy restrictions are satisfied in the linear representation.[3] However, there are no known (necessary and sufficient) determinacy conditions for the class of bounded solutions of MSNK models. Farmer et al. (2009) and Cho (2009) suggest considering mean-square stable solutions rather than bounded solutions and provide determinacy conditions within the class of mean-square stable solutions of a quasi-linear version of MSNK models.[4] While these advances are certainly interesting, restricting our attention to the class of bounded solutions is of interest not only because of the uniqueness of equilibrium, but also because we can justify the log-linear approximation of the original fully nonlinear model only in the local neighborhood of the steady state by the implicit function theorem.[5] For this reason, we believe that the linear representation is useful for the analysis of MSNK models given our current knowledge.

2 Alternative Specification

2.1 Equilibrium Conditions

In the alternative model, the representative household chooses consumption (C_t) and labor supply (N_t) to maximize the following lifetime utility

[2] Mccallum (2004) proves the equivalence between MSV solutions and determinate (i.e. unique and non-explosive) solutions from solving a system of linear expectations difference equations for purely forward looking models. Davig and Leeper (2007) show the equivalence between the MSV solution and determinate solution of the stacked system in regime-switching rational expectation models.

[3] See Farmer et al. (2010a).

[4] The terminology is from Davig and Leeper (2010). The quasi-linear version is based on the log-linear approximation of the original MSNK model but keep nonlinearities due to regime-switching parts unlike the linear representation.

[5] See Woodford (2003) on this point.

$$E_t \sum_{t=0}^{\infty} \beta^t \left(\frac{\left(\frac{C_t - be^{\lambda} \overline{C_{t-1}}}{A_t} \right)^{1-\tau}}{1-\tau} - H_t \right),$$

where $\overline{C_{t-1}}$ denotes lagged aggregate consumption and b is a parameter determining the degree of habit formation, and λ is the average growth rate of log productivity (lnA_t).

The marginal utility of consumption at t is given by $\mu_t = (\frac{C_t - be^{\lambda}\overline{C_{t-1}}}{A_t})^{-\tau}\frac{1}{A_t}$. The household's maximization implies that the short-term nominal interest rate should satisfy the consumption Euler equation,

$$E_t(\beta \frac{\mu_{t+1}}{\mu_t} \frac{R_t}{\Pi_{t+1}}) = 1. \tag{7}$$

Log-linearizing this first-order condition, we obtain the following equation

$$\widehat{y}_t = \frac{E_t(\widehat{y}_{t+1})}{1+b} + \frac{b}{1+b}\widehat{y}_{t-1} + \frac{(1-b+\tau b)\rho_a - \tau b}{\tau(1+b)}\widehat{a}_t - \frac{(1-b)}{\tau(1+b)}(\widehat{R}_t - E_t(\widehat{\pi}_{t+1})). \tag{8}$$

We modify the specification for price adjustment costs of intermediate goods-producing firms to allow the internal dynamics to generate inflation persistence. Now, the cost of adjusting prices depends on not the steady state inflation rate but a weighted average of the steady state inflation rate and the previous period inflation rate

$$ac_{jt} = \frac{\varphi}{2}(\frac{p_{jt}}{p_{jt-1}} - (\Pi_{t-1})^{\gamma(s_{t-1})}(\Pi)^{1-\gamma(s_{t-1})})^2 Y_t. \tag{9}$$

While this specification looks somewhat arbitrary, it induces a similar log-linearized Phillips curve from a Calvo (1983) type model with dynamic indexation. Since the indexation parameter may not be invariant to different monetary policy regimes, we allow the parameter to depend on regimes.

Each intermediate-goods production firm maximizes the expected present value of profits

$$E_t(\sum_{s=0}^{\infty} \beta^s \frac{\mu_{t+s}}{\mu_t}[(\frac{p_{jt+s}}{P_{t+s}})^{1-\theta_{t+s}} - \psi_{t+s}(\frac{p_{jt+s}}{P_{t+s}})^{-\theta_{t+s}} - \frac{\varphi}{2}(\frac{p_{jt+s}}{p_{jt+s-1}} - (\Pi_{t+s-1})^{\gamma(s_{t+s-1})}(\Pi)^{1-\gamma(s_{t+s-1})})^2]Y_{t+s}).$$

where ψ_{t+s} denotes the real marginal cost $(W_t/P_t)/A_t$.

We can define the price markup in the absence of price adjustment costs by $f_t = \frac{\theta_t}{\theta_t - 1}$. The steady state markup is $f = \frac{\theta}{\theta-1}$.

The log-linearization of the first-order condition of each intermediate goods-production firm around the steady state results in the following Phillips curve

$$\widehat{\pi}_t - \gamma_t(s_{t-1})\widehat{\pi}_{t-1} = \beta(E_t(\widehat{\pi}_{t+1}) - \gamma_{t+1}(s_t)\widehat{\pi}_t) + \kappa(\frac{\widehat{y}_t - b\widehat{y}_{t-1} + b\widehat{a}_t}{1-b} + \frac{\widehat{f}_t}{\tau}). \tag{10}$$

where $\kappa = \frac{\tau}{\varphi(\Pi)^2(f-1)}$. In the estimation, we use a rescaled markup shock $\widehat{u}_t = \frac{\widehat{f}_t}{\tau}$.

We incorporate policy rule inertia into the Taylor rule. We assume that the monetary authority sets the short-term nominal rate using the following rule

$$R_t = (\overline{r}\Pi(\frac{\Pi_t}{\Pi})^{\alpha_p(s_t)}(\frac{Y_t}{A_t y^\star})^{\alpha_y(s_t)})^{1-\rho_R(s_t)}(R_{t-1})^{\rho_R(s_t)}(e_t). \tag{11}$$

where r is the steady state real interest rate given by $\frac{e^\lambda}{\beta}$. Conditioning on a give regime, the monetary policy rule in terms of the log-deviation from the steady state is linear and given by

$$\widehat{R}_t = \rho_R(s_t)\widehat{R}_{t-1} + (1-\rho_R(s_t))(\alpha_p(s_t)\widehat{\pi}_t + \alpha_y(s_t)\widehat{y}_t) + \widehat{e}_t. \tag{12}$$

Once we shut down habit formation ($b = 0$) and policy rule inertia ($\rho_R(s_t) = 0$) with $\gamma(s_{t-1}) = 0$ while allowing the persistence of a monetary policy shock, we are back in the baseline model.

The extended model contains lagged endogenous variables as state variables. Since finding MSV solutions by the expanded linear system in Davig and Leeper (2007) does not work in this case, we use a numerical method suggested by Farmer, Waggoner, and Zha (2010b) to find a MSV solution of the model.

2.2 Estimation Results of Regime Switching Models under the Alternative Specification

We applied MCMC methods to generate posterior output for regime-switching models under the alternative specification. However, after several draws, MCMC chains tend to reject every proposed draw. We implemented tailored randomized block MCMC methods in Chib and Ramamurthy (2010) in addition to the procedure that we used in the estimation of the baseline model in the paper. We still faced difficulties in generating accepted draws from a tailored proposal density. However, finding out posterior modes by using the following blockwise optimization methods worked fine for all the models.[6]

- **Step 1:** Divide a vector of parameters θ into several blocks using random permutation of indexes of parameters $\theta_1, \cdots, \theta_B$.
- **Step 2:** Maximize the posterior kernel with respect to θ_i , $(i = 1, \cdots, B)$ while parameters of other blocks fixed by simulated annealing. Denote the solution of this maximization problem by $\theta_i^\star$.
- **Step 3:** Perturb θ_i around $\theta_i^\star$ and use the simplex search method to see if further improvement of the posterior kernel is possible.
- **Step 4:** Repeat previous steps while parameter estimates converge.

After intensive search, we found reasonable estimates of posterior modes, which are robust to many trials of perturbed search. Table 1 provides information on parameter estimates. Interestingly, under the alternative specification, the posterior mode of the model with only volatility shifts has the highest log-likelihood value as shown in Table 2. However, this value is still lower than the log-likelihood at the posterior mode of the baseline four regime model. Hence, at least in

[6]The methods apply tailored randomized blocking in Chib and Ramamurthy (2010) for finding out point estimates.

terms of empirical fit, richer structures in the alternative specification do not necessarily provide improvements over the baseline specification.

3 Convergence Diagnostics and Marginal Likelihood Computation

We start a MCMC chain initialized at the mode of the posterior kernel found by a numerical optimization routine. We compute the mean (m) and the covariance (c) matrix based on 1 million draws from the MCMC chain. Then we run three different MCMC chains initialized at widely dispersed points around the mean (m) by using the covariance matrix (c) as the scaling matrix in the proposal density of the MCMC Chain. We burn-in the first 100,000 draws and use the remaining 1,400,000 draws for the posterior inference in the four-regime model.[7] To check the convergence of our MCMC output, we compute potential scale reduction factors (PSRFs) suggested by Brooks and Gelman (1998). Table 3 shows PSRFs which compare between and within variances of multiple chains. If these numbers are below 1.1 or 1.2, between chain variances are not much different from within chain variances, indicating that MCMC output converges to the stationary distribution. Except for the persistence of markup shock in the model with shifts in volatility only, PSRF statistics are below 1.2 for all the other parameters.

Marginal likelihood reported in the paper can be computed by the Monte Carlo integration of a probability density function of ϑ as follows.

$$p(Z^T)^{-1} = \int \frac{h(\vartheta)}{p(Z^T|\vartheta)p(\vartheta)} p(\vartheta|Z^T)d\vartheta \longrightarrow \widehat{p}(Z^T)^{-1} = \frac{1}{N}\sum_{i=1} \frac{h(\vartheta_i)}{p(Z^T|\vartheta_i)p(\vartheta_i)} \tag{13}$$

where ϑ_i is a posterior draw. Geweke (1999) proposes an implementation with $h(\cdot)$ a Gaussian density around the posterior mean. Sims et al. (2008) show that the Gaussian approximation for the proposal density to compute the marginal likelihood can be misleading and numerically unstable due to the non-Gaussian posterior distribution of parameters in models with regime-switching. They suggest the following elliptical distribution as an alternative.

$$g(\vartheta) = \frac{\Gamma(k/2)}{(2\pi)^{k/2}|det(\overline{S})|}\frac{f(r)}{r^{k-1}} \ , \ r = \sqrt{(\vartheta-\overline{\vartheta})'\overline{\Omega}^{-1}(\vartheta-\overline{\vartheta})} \ , \ \overline{S} = c\sqrt{\overline{\Omega}}$$

where $f(r)$ is any one-dimensional density defined on the positive reals which we can estimated based on posterior draws. $\overline{\Omega}$ is the sample covariance matrix computed by centering out posterior draws from the posterior mode. And c is a scaling parameter. Then the weighting function can be constructed as a truncated $g(\cdot)$ where the area with a very low posterior probability is truncated.

$$h(\vartheta) = \frac{\Xi_{\Theta_L}}{q_L} g(\vartheta) \ , \ \Theta_L = \{\vartheta : p(Z^T|\vartheta)p(\vartheta) > L\}$$

where Ξ_{Θ_L} is an indicator function which is 1 if ϑ belong to Θ_L and 0 otherwise. While Sims et al. (2008) provide an example where this methods is more robust than Geweke (1999), it does not work well for our purpose. Particularly, $g(\cdot)$ turns out to be very sensitive to the choice of a scaling parameter of $\overline{\Omega}$.[8] In contrast, Geweke (1999)'s method provides more reliable estimates of

[7] We burn-in the first 100,000 draws and use the remaining 1,900,000 draws in the model with only volatility shifts while we burn-in 500,000 draws use the remaining 500,000 draws in the model with only policy shifts.

[8] Sims et al. (2008) acknowledge that the choice of the scaling parameter is importatnt in the implementation of their method.

marginal likelihood in our cases. And the marginal likelihood computed by Sims et al. (2008)'s method also indicates the the best fitting model is the four regime model with a constant inflation target as implied by the marginal likelihood computed by Geweke's (1999) method.

Additional References

Bansal, R., and H. Zhou (2002): "Term Structure of Interest Rates with Regime Shifts," *Journal of Finance*, 57, pp.1997-2043.

Blanchard, O. J., and C. Kahn (1980): "The Solution of Linear Difference Models Under Rational Expectations," *Econometrica*, 48, pp.1305-1313.

Brooks, S., and A. Gelman (1998): "General Methods for Monitoring Convergence of Iterative Simulations," *Journal of Computational and Graphical Statistics*, 7(4), pp.434-455.

Calvo, G. A. (1983): "Staggered Prices in a Utility-maximizing Framework," *Journal of Moentary Economics*, 12(3), pp.383-398.

Chib, S., and S. Ramamurthy (2010): "Tailored Randomized Block MCMC Methods with Application to DSGE Models," *Journal of Econometrics*, 155(1), pp.19-38.

Cho, S. (2009): "Characterizing Markov-Switching Rational Expectations Models," Working Paper, Yonsei University.

Davig, T. and E. M. Leeper (2010): "Generalizing the Taylor Principle: Reply," *American Economic Review*, 100(1), pp.618-624.

Farmer, R. E., D. F. Waggoner, and T. Zha (2009): "Understanding Markov-Switching Rational Expectations Models," *Journal of Economic Theory*, 144, pp.1849-1867.

Farmer, R. E., D. F. Waggoner, and T. Zha (2010a): "Generalizing the Taylor Principle: Comment," *American Economic Review*, 100(1), pp.608-617.

Farmer, R. E. A., D. F. Waggoner, and T. Zha (2010b): "Minimal State Variable Solutions to Markov-Switching Rational Expectations Models", *Manuscript*, Federal Reserve Bank of Atlanta.

McCallum, B. (2004): "On the Relationship Between Determinate and MSV Solutions in Linear RE Models," *Economic Letters*, 84, pp.55-60.

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

Table 1: Posterior Modes under the Alternative Specification

Parameters	Prior 90% Interval	Posterior Mode		
		P1	P2	P3
τ	[0.84,2.13]	3.76	4.63	3.999
b	[0.07,0.97]	0.08	0.11	0.105
κ	[0.18,0.81]	0.21	0.11	0.163
β	[0.997,0.999]	0.999	0.999	0.999
γ	[0.180, 0.828]		0.029	
γ_1	[0.001,0.589]	0.007		0.007
γ_2	[0.414,0.999]	0.50		0.50
α_p	[1.075, 1.884]		1.378	
$\alpha_{p,1}$	[1.570,2.404]	2.093		1.992
$\alpha_{p,2}$	[0.917,1.080]	0.973		0.980
α_y	[0.027, 0.176]		0.083	
$\alpha_{y,1}$	[0.022,0.172]	0.121		0.088
$\alpha_{y,2}$	[0.023,0.173]	0.083		0.088
ρ_a	[0.001,0.593]	0.50	0.373	0.428
ρ_u	[0.413,0.999]	0.947	0.933	0.910
ρ_R	[0.179,0.839]		0.849	
$\rho_{R,1}$	[0.001,0.588]	0.806		0.861
$\rho_{R,2}$	[0.408,0.998]	0.804		0.773
σ_a	[0.0002,0.0008]	0.0075		
$\sigma_{a,1}$	[0.0033,0.0119]		0.0139	0.0134
$\sigma_{a,2}$	[0.0011,0.0040]		0.0069	0.0067
σ_u	[0.0016,0.0059]	0.0056		
$\sigma_{u,1}$	[0.0021,0.0080]		0.0079	0.0081
$\sigma_{u,2}$	[0.0010,0.0039]		0.0041	0.005
σ_e	[0.0016,0.0060]	0.0023		
$\sigma_{e,1}$	[0.0022,0.0080]		0.0034	0.0039
$\sigma_{e,2}$	[0.0010,0.0040]		0.0011	0.0012
$\ln A_0$	[9.387,9.711]	9.525	9.527	9.529
$y^\star$	[-0.0839,-0.0510]	-0.0672	-0.0673	-0.0695
$\pi^\star$	[0.007, 0.0102]	0.0075	0.0083	0.0072
λ	[0.0034, 0.0067]	0.0043	0.0045	0.0046
p_{11}	[0.82,0.98]	0.993		0.99
p_{22}	[0.81,0.98]	0.954		0.92
q_{11}	[0.82,0.98]		0.91	0.87
q_{22}	[0.82,0.98]		0.96	0.96

Notes: P1 allows switching only in monetary policy coefficients while P2 allows switching coefficients only in variance parameters of shocks. P3 allows switching for both policy coefficients and variances.

Table 2: Log Likelihood at the Posterior Mode of Each Model

	P1	P2	P3
Baseline Specification	2,743.3	2,770.8	2,785.4
Alternative Specification	2,702.3	2,784.6	2,779.8

Table 3: POTENTIAL SCALE REDUCTION FACTORS FOR MCMC OUTPUT

Parameter	P1	P 2	P3
α		1.026	
α_1	1.002		1.017
α_2	1.004		1.009
γ		1.008	
γ_1	1.001		1.001
γ_2	1.004		1.003
κ	1.000	1.002	1.012
β	1.00	1.010	1.000
τ	1.002	1.010	1.048
λ	1.000	1.041	1.005
Π	1.003	1.092	1.002
ρ_a	1.000	1.001	1.002
ρ_u	1.001	1.270	1.006
ρ_i	1.000	1.004	1.025
σ_a	1.001		
$\sigma_{a,1}$		1.002	1.004
$\sigma_{a,2}$		1.004	1.001
σ_u	1.005		
$\sigma_{u,1}$		1.017	1.006
$\sigma_{u,2}$		1.023	1.002
σ_e	1.001		
$\sigma_{e,1}$		1.002	1.001
$\sigma_{e,2}$		1.004	1.011
$\ln A_0$	1.005	1.045	1.000
y^*	1.008	1.099	1.000
p_{11}	1.005		1.001
p_{22}	1.001		1.001
q_{11}		1.001	1.001
q_{22}		1.002	1.000

Notes: P1 allows switching only in monetary policy coefficients while P2 allows switching coefficients only in variance parameters of shocks. P3 allows switching for both policy coefficients and variances.