

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32

**Supplement to “Behavioral Learning Equilibria in New Keynesian Models”**

CARS HOMMES

Bank of Canada and CeNDEF, Amsterdam School of Economics, University of Amsterdam,  
Netherlands

KOSTAS MAVROMATIS

De Nederlandsche Bank and Amsterdam School of Economics, University of Amsterdam,  
Netherlands

TOLGA ÖZDEN

Bank of Canada

MEI ZHU

Institute for Advanced Research and School of Economics, Shanghai University of Finance and  
Economics and the Key Laboratory of Mathematical Economics (SUFE), Ministry of Education,  
Shanghai 200433, China

---

Cars Hommes: [C.H.Hommes@uva.nl](mailto:C.H.Hommes@uva.nl)

Kostas Mavromatis: [K.Mavromatis@dnb.nl](mailto:K.Mavromatis@dnb.nl)

Tolga Özden: [tozden@bank-banque-canada.ca](mailto:tozden@bank-banque-canada.ca)

Mei Zhu: [Zhu.Mei@mail.shufe.edu.cn](mailto:Zhu.Mei@mail.shufe.edu.cn)

The views expressed in this paper do not represent the positions of the Bank of Canada, De Nederlandsche Bank or the Eurosystem.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32

## APPENDIX

## APPENDIX A: THEORETICAL RESULTS

A.1 *Mean of the rational expectations equilibrium*

Using (2.10–2.11) and (2.15–2.18), the mean of the REE satisfies

$$\begin{aligned}
 \bar{\mathbf{x}}^* &= (\mathbf{I} - \mathbf{c}_1)^{-1}(\mathbf{c}_0 + \mathbf{c}_2\bar{\mathbf{u}}) \\
 &= (\mathbf{I} - \mathbf{c}_1)^{-1}(\mathbf{I} - \mathbf{b}_1\mathbf{c}_1 - \mathbf{b}_1)^{-1}(\mathbf{b}_0 + \mathbf{b}_1\mathbf{c}_2\mathbf{a}) + (\mathbf{I} - \mathbf{c}_1)^{-1}\mathbf{c}_2(\mathbf{I} - \boldsymbol{\rho})^{-1}\mathbf{a} \\
 &= (\mathbf{I} - \mathbf{c}_1)^{-1}(\mathbf{I} - \mathbf{b}_1\mathbf{c}_1 - \mathbf{b}_1)^{-1}[\mathbf{b}_0 + (\mathbf{b}_1\mathbf{c}_2(\mathbf{I} - \boldsymbol{\rho}) + (\mathbf{I} - \mathbf{b}_1\mathbf{c}_1 - \mathbf{b}_1)\mathbf{c}_2)(\mathbf{I} - \boldsymbol{\rho})^{-1}\mathbf{a}] \\
 &= [(\mathbf{I} - \mathbf{b}_1\mathbf{c}_1 - \mathbf{b}_1)(\mathbf{I} - \mathbf{c}_1)]^{-1}[\mathbf{b}_0 + \mathbf{b}_3(\mathbf{I} - \boldsymbol{\rho})^{-1}\mathbf{a}] \\
 &= (\mathbf{I} - \mathbf{b}_1 - \mathbf{b}_2)^{-1}[\mathbf{b}_0 + \mathbf{b}_3(\mathbf{I} - \boldsymbol{\rho})^{-1}\mathbf{a}].
 \end{aligned}$$

A.2 *Autocorrelations in the multivariate linear model*

The purpose of this appendix is to compute the first-order autocorrelation coefficients of the linear stochastic stationary system (2.24) and to show that these are continuous functions with respect to  $(\beta_1, \beta_2, \dots, \beta_n)$  and the other parameters.

Define  $\mathbf{X}'_t = [\mathbf{x}'_t, \mathbf{u}'_t] - [\bar{\mathbf{x}}', \bar{\mathbf{u}}']$ . Rewrite model (2.24) as

$$\mathbf{X}_t = \widehat{\mathbf{B}}(\boldsymbol{\beta})\mathbf{X}_{t-1} + \widehat{\mathbf{C}}\boldsymbol{\eta}_t, \quad (\text{A.1})$$

where  $\boldsymbol{\eta}'_t = [\mathbf{v}'_t, \boldsymbol{\varepsilon}'_t]$ ,  $\widehat{\mathbf{B}}(\boldsymbol{\beta}) = \begin{pmatrix} \mathbf{b}_1\boldsymbol{\beta}^2 + \mathbf{b}_2\mathbf{b}_3\boldsymbol{\rho} & \\ \mathbf{0} & \boldsymbol{\rho} \end{pmatrix}$ ,  $\widehat{\mathbf{C}} = \begin{pmatrix} \mathbf{b}_4 & \mathbf{b}_3 \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$ . The variance-covariance matrix  $\widehat{\boldsymbol{\Gamma}}(0)$  and the autocovariance matrix  $\widehat{\boldsymbol{\Gamma}}(1)$  satisfy

$$\widehat{\boldsymbol{\Gamma}}(0) = E[\mathbf{X}_t\mathbf{X}'_t] = \widehat{\mathbf{B}}(\boldsymbol{\beta})\widehat{\boldsymbol{\Gamma}}(0)\widehat{\mathbf{B}}'(\boldsymbol{\beta}) + \widehat{\mathbf{C}}\boldsymbol{\Sigma}_\eta\widehat{\mathbf{C}}', \quad (\text{A.2})$$

$$\widehat{\boldsymbol{\Gamma}}(1) = E[\mathbf{X}_t\mathbf{X}'_{t+1}] = \widehat{\boldsymbol{\Gamma}}(0)\widehat{\mathbf{B}}'(\boldsymbol{\beta}), \quad (\text{A.3})$$

where  $\boldsymbol{\Sigma}_\eta = \begin{pmatrix} \boldsymbol{\Sigma}_v & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_\varepsilon \end{pmatrix}$ .

In order to obtain an expression for  $\widehat{\Gamma}(0)$ , we use column stacks of matrices. Suppose  $\text{vec}(\mathbf{K})$  is the vectorization of a matrix  $\mathbf{K}$  and  $\otimes$  is the Kronecker product.<sup>1</sup> Under the assumption that all eigenvalues of  $\mathbf{b}_1\beta^2 + \mathbf{b}_2$  and  $\rho$  are inside the unit circle, based on a property of Kronecker product,<sup>2</sup> it is easy to see that all eigenvalues of  $\widehat{\mathbf{B}}(\beta) \otimes \widehat{\mathbf{B}}(\beta)$  lie inside the unit circle and hence  $[\mathbf{I} - \widehat{\mathbf{B}}(\beta) \otimes \widehat{\mathbf{B}}(\beta)]^{-1}$  exist. Therefore,

$$\text{vec}(\widehat{\Gamma}(0)) = [\mathbf{I} - \widehat{\mathbf{B}}(\beta) \otimes \widehat{\mathbf{B}}(\beta)]^{-1}(\widehat{\mathbf{C}} \otimes \widehat{\mathbf{C}})\text{vec}(\boldsymbol{\Sigma}_\eta). \quad (\text{A.4})$$

From (A.4) we can obtain an expression for  $\widehat{\Gamma}(0)$ , and using (A.3), an expression for  $\widehat{\Gamma}(1)$  can be obtained.

Based on the properties of matrix operations, it is easy to see that the entries of matrices  $\widehat{\Gamma}(0)$  and  $\widehat{\Gamma}(1)$  are continuous functions with respect to  $(\beta_1, \beta_2, \dots, \beta_n)$  and the other parameters. Let  $\boldsymbol{\Omega} = \text{diag}[\gamma_{11}(0), \gamma_{22}(0), \dots, \gamma_{nn}(0)]$  (a diagonal matrix), where  $\gamma_{ii}(0)$  ( $i = 1, \dots, n$ ) are the first  $n$  diagonal entries of  $\widehat{\Gamma}(0)$ . Let  $\mathbf{L} = \text{diag}[\gamma_{11}(1), \gamma_{22}(1), \dots, \gamma_{nn}(1)]$  (a diagonal matrix), where  $\gamma_{ii}(1)$  ( $i = 1, \dots, n$ ) are the first  $n$  diagonal entries of  $\widehat{\Gamma}(1)$ . Thus the first-order autocorrelation coefficients of the linear stochastic stationary system (2.24)  $\mathbf{G} = \mathbf{L}\boldsymbol{\Omega}^{-1}$  are continuous functions with respect to  $(\beta_1, \beta_2, \dots, \beta_n)$  and the other parameters.

For example, in the case  $n = 1$ , following the procedures above with Mathematica or Matlab software, one obtains

$$G(\beta) = \left[ (b_3^2\sigma_\varepsilon^2 + b_4^2\sigma_v^2)(b_1\beta^2 + b_2) + (b_3^2\sigma_\varepsilon^2 - b_4^2\sigma_v^2)(b_1\beta^2 + b_2)^2\rho - b_4^2\sigma_v^2(b_1\beta^2 + b_2)\rho^2 + b_4^2\sigma_v^2(b_1\beta^2 + b_2)^2\rho^3 \right] / \left[ b_3^2\sigma_\varepsilon^2(1 + (b_1\beta^2 + b_2)\rho) + b_4^2\sigma_v^2(1 - (b_1\beta^2 + b_2)\rho)(1 - \rho^2) \right]$$

<sup>1</sup>One property of column stacks is that the column stack of a product of three matrices is  $\text{vec}(ABC) = (\mathbf{C}' \otimes \mathbf{A})\text{vec}(\mathbf{B})$ . For more details on this and related properties, see Magnus and Neudecker (2019, Chapter 2) and Evans and Honkapohja (2001, Section 5.7).

<sup>2</sup>The eigenvalues of  $\check{\mathbf{A}} \otimes \check{\mathbf{B}}$  are the  $mn$  numbers  $\lambda_r\mu_s$ ,  $r = 1, 2, \dots, m$ ,  $s = 1, 2, \dots, n$  where  $\lambda_1, \dots, \lambda_m$  are the eigenvalues of  $m \times m$  matrix  $\check{\mathbf{A}}$  and  $\mu_1, \dots, \mu_n$  are the eigenvalues of  $n \times n$  matrix  $\check{\mathbf{B}}$  (see Lancaster and Tismenetsky, 1985).

In the special case  $b_2 = 0$  and  $b_4 = 1$ , this expression is exactly the same as the first-order autocorrelation in Hommes and Zhu (2014), which was calculated using another approach.

### A.3 Proof of Proposition 2 (stability under SAC-learning)

This appendix derives the E-stability conditions for a BLE  $(\alpha^*, \beta^*)$ . Set  $\gamma_t = (1 + t)^{-1}$ . For the learning dynamics in (2.28) and (2.8),<sup>3</sup> since all functions are smooth, the SAC-learning rule satisfies the conditions (A.1–A.3) of Section 6.2.1 in Evans and Honkapohja (2001, p. 124). In order to check the conditions (B.1–B.2) of Section 6.2.1 in Evans and Honkapohja (2001, p. 125), we rewrite the system in matrix form as

$$\mathbf{X}_t = \tilde{\mathbf{A}}(\theta_{t-1})\mathbf{X}_{t-1} + \tilde{\mathbf{B}}(\theta_{t-1})\mathbf{W}_t,$$

where  $\theta'_t = (\alpha_t, \beta_t, \mathbf{R}_t)$ ,  $\mathbf{X}'_t = (1, \mathbf{x}'_t, \mathbf{x}'_{t-1}, \mathbf{u}'_t)$  and  $\mathbf{W}'_t = (1, \mathbf{v}'_t, \epsilon'_t)$ ,

$$\tilde{\mathbf{A}}(\theta) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{b}_0 + \mathbf{b}_1(\mathbf{I} - \beta^2)\alpha + \mathbf{b}_3\mathbf{a} & \mathbf{b}_1\beta^2 + \mathbf{b}_2 & \mathbf{0} & \mathbf{b}_3\rho \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{a} & \mathbf{0} & \mathbf{0} & \rho \end{pmatrix},$$

$$\tilde{\mathbf{B}}(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ \mathbf{0} & \mathbf{b}_4 & \mathbf{b}_3 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

Based on the properties of eigenvalues (see, e.g., Evans and Honkapohja, 2001, p. 117), all the eigenvalues of  $\tilde{\mathbf{A}}(\theta)$  include 0 (multiple  $n + 1$ ), the eigenvalues of  $\rho$  and  $\mathbf{b}_1\beta^2 + \mathbf{b}_2$ . Thus, based on the assumptions, all the eigenvalues of  $\tilde{\mathbf{A}}(\theta)$  lie inside the unit circle. Moreover, it is easy to see all the other conditions in Section 6.2.1 in Evans and Honkapohja (2001) are also satisfied.

<sup>3</sup>For convenience of theoretical analysis, one can set  $\mathbf{S}_{t-1} = \mathbf{R}_t$ .

Since  $\mathbf{x}_t$  is stationary, then the limits

$$\sigma_i^2 := \lim_{t \rightarrow \infty} E(x_{i,t} - \alpha_i)^2, \quad \sigma_{x_i x_{i,-1}}^2 := \lim_{t \rightarrow \infty} E(x_{i,t} - \alpha_i)(x_{i,t-1} - \alpha_i)$$

exist and are finite. Hence, according to Section 6.2.1 in Evans and Honkapohja (2001, p. 126), the associated ODE is

$$\begin{cases} \frac{d\boldsymbol{\alpha}}{d\tau} = \bar{\mathbf{x}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) - \boldsymbol{\alpha}, \\ \frac{d\boldsymbol{\beta}}{d\tau} = \mathbf{R}^{-1}[\mathbf{L} - \boldsymbol{\beta}\boldsymbol{\Omega}] = \mathbf{R}^{-1}\boldsymbol{\Omega}[\mathbf{L}\boldsymbol{\Omega}^{-1} - \boldsymbol{\beta}], \\ \frac{d\mathbf{R}}{d\tau} = \boldsymbol{\Omega} - \mathbf{R}, \end{cases} \quad (\text{A.5})$$

where  $\mathbf{R}$  is a diagonal matrix with the  $i$ -th diagonal entry  $R_i$  and  $\boldsymbol{\Omega}$ ,  $\mathbf{L}$  are also diagonal matrices, as defined in Section 2. As shown in Evans and Honkapohja (2001), a BLE corresponds to a fixed point of the following ODE (A.6):

$$\begin{cases} \frac{d\boldsymbol{\alpha}}{d\tau} = \bar{\mathbf{x}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) - \boldsymbol{\alpha}, \\ \frac{d\boldsymbol{\beta}}{d\tau} = \mathbf{G} - \boldsymbol{\beta}. \end{cases} \quad (\text{A.6})$$

Note that  $\boldsymbol{\beta}$  and  $\mathbf{G}$  are both diagonal matrices. The Jacobian matrix of A.6 is, in fact, equivalent to

$$\begin{pmatrix} (\mathbf{I} - \mathbf{b}_1\boldsymbol{\beta}^{*2} - \mathbf{b}_2)^{-1}(\mathbf{b}_1 + \mathbf{b}_2 - \mathbf{I}) & \boldsymbol{\rho} \\ \mathbf{0} & \mathbf{DG}_{\boldsymbol{\beta}}(\boldsymbol{\beta}^*) - \mathbf{I} \end{pmatrix},$$

where  $\mathbf{DG}_{\boldsymbol{\beta}}$  is a Jacobian matrix with the  $(i, j)$ -th entry  $\frac{\partial G_i}{\partial \beta_j}$ , and the form of matrix  $\boldsymbol{\rho}$  is omitted since it is not needed in the proof. Therefore, if all the eigenvalues of  $(\mathbf{I} - \mathbf{b}_1\boldsymbol{\beta}^{*2} - \mathbf{b}_2)^{-1}(\mathbf{b}_1 + \mathbf{b}_2 - \mathbf{I})$  have negative real parts and all the eigenvalues of  $\mathbf{DG}_{\boldsymbol{\beta}}(\boldsymbol{\beta}^*)$  have real parts less than 1, the SAC-learning  $(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t)$  converges to the BLE  $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$  as time  $t$  tends to  $\infty$ .

#### A.4 Eigenvalues of matrix $B\beta^2$ inside the unit circle

This appendix shows the sufficiency condition for the existence of a BLE of (2.40) in Corollary 1. The characteristic polynomial of  $B\beta^2$  is given by  $h(\nu) = \nu^2 + c_1\nu + c_2$ , where

$$c_1 = -\frac{\beta_y^2 + [\gamma\varphi + \lambda(1 + \varphi\phi_y)]\beta_\pi^2}{1 + \gamma\varphi\phi_\pi + \varphi\phi_y}, \quad c_2 = \frac{\lambda\beta_y^2\beta_\pi^2}{1 + \gamma\varphi\phi_\pi + \varphi\phi_y}.$$

Both of the eigenvalues of  $B\beta^2$  are inside the unit circle if and only if both of the following conditions hold (see [Elaydi, 2005](#)):

$$h(1) > 0, \quad h(-1) > 0, \quad |h(0)| < 1.$$

It is easy to see  $h(-1) > 0, |h(0)| < 1$  for any  $\beta_i \in [-1, 1]$ . Note that

$$\begin{aligned} h(1) &= \frac{(1 - \beta_y^2)(1 - \lambda\beta_\pi^2) + \gamma\varphi\phi_\pi + \varphi\phi_y - (\gamma\varphi + \lambda\varphi\phi_y)\beta_\pi^2}{1 + \gamma\varphi\phi_\pi + \varphi\phi_y}, \\ &\geq \frac{\varphi[\gamma(\phi_\pi - 1) + (1 - \lambda)\phi_y]}{1 + \gamma\varphi\phi_\pi + \varphi\phi_y}. \end{aligned}$$

Thus, if  $\gamma(\phi_\pi - 1) + (1 - \lambda)\phi_y > 0$ , then  $h(1) > 0$ . Therefore, both eigenvalues of  $B\beta^2$  lie inside the unit circle for all  $\beta_i \in [-1, 1]$ .

#### A.5 First-order autocorrelations in the baseline NK model

This appendix gives the expressions for the first-order autocorrelation coefficients for the output gap and inflation in the NK baseline model. Through complicated calculations, the following expressions in terms of the structural parameters are obtained:

$$G_1(\beta_y, \beta_\pi) = \frac{\tilde{f}_1}{\tilde{g}_1}, \tag{A.7}$$

$$G_2(\beta_y, \beta_\pi) = \frac{\tilde{f}_2}{\tilde{g}_2}, \tag{A.8}$$

1 where

$$\begin{aligned}
 \tilde{f}_1 = \sigma_y^2 & \left\{ (\rho + \lambda_1 + \lambda_2 - \lambda\beta_\pi^2)[1 - \lambda\beta_\pi^2(\rho + \lambda_1 + \lambda_2)] + [\lambda\beta_\pi^2(\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2) - \right. \\
 & \left. \rho\lambda_1\lambda_2][(\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2) - \lambda\beta_\pi^2\rho\lambda_1\lambda_2] \right\} + \sigma_\pi^2 \left\{ (\varphi\phi_\pi(\rho + \lambda_1 + \lambda_2) - \varphi\beta_\pi^2) \right. \\
 & \left. [\varphi\phi_\pi - \varphi\beta_\pi^2(\rho + \lambda_1 + \lambda_2)] + [\varphi\beta_\pi^2(\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2) - \varphi\phi_\pi\rho\lambda_1\lambda_2] \right. \\
 & \left. [\varphi\phi_\pi(\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2) - \varphi\beta_\pi^2\rho\lambda_1\lambda_2] \right\},
 \end{aligned}$$

$$\begin{aligned}
 \tilde{g}_1 = \sigma_y^2 & \left\{ [(1 + \lambda^2\beta_\pi^4) - 2\lambda\beta_\pi^2(\rho + \lambda_1 + \lambda_2) + (1 + \lambda^2\beta_\pi^4)(\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2)] \right. \\
 & \left. - \rho\lambda_1\lambda_2[(1 + \lambda^2\beta_\pi^4)(\rho + \lambda_1 + \lambda_2) - 2\lambda\beta_\pi^2(\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2) + (1 + \lambda^2\beta_\pi^4)\rho\lambda_1\lambda_2] \right\} \\
 & + \sigma_\pi^2 \left\{ [((\varphi\phi_\pi)^2 + \varphi^2\beta_\pi^4) - 2\varphi\phi_\pi\varphi\beta_\pi^2(\rho + \lambda_1 + \lambda_2) + ((\varphi\phi_\pi)^2 + \varphi^2\beta_\pi^4) \right. \\
 & \left. (\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2)] - \rho\lambda_1\lambda_2[ ((\varphi\phi_\pi)^2 + \varphi^2\beta_\pi^4)(\rho + \lambda_1 + \lambda_2) - 2\varphi\phi_\pi\varphi\beta_\pi^2 \right. \\
 & \left. (\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2) + ((\varphi\phi_\pi)^2 + \varphi^2\beta_\pi^4)\rho\lambda_1\lambda_2] \right\}, \tag{A.9}
 \end{aligned}$$

$$\begin{aligned}
 \tilde{f}_2 = \sigma_y^2 & \left\{ \gamma^2[(\rho + \lambda_1 + \lambda_2) - \rho\lambda_1\lambda_2(\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2)] \right\} + \sigma_\pi^2 \left\{ [(1 + \varphi\phi_y)(\rho + \lambda_1 + \lambda_2) \right. \\
 & \left. - \beta_y^2] \cdot [(1 + \varphi\phi_y) - \beta_y^2(\rho + \lambda_1 + \lambda_2)] + [\beta_y^2(\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2) - (1 + \varphi\phi_y)\rho\lambda_1\lambda_2] \cdot \right. \\
 & \left. [(1 + \varphi\phi_y)(\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2) - \beta_y^2\rho\lambda_1\lambda_2] \right\},
 \end{aligned}$$

$$\begin{aligned}
 \tilde{g}_2 = \sigma_y^2 & \left\{ \gamma^2[1 + \rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2 - \rho\lambda_1\lambda_2(\rho + \lambda_1 + \lambda_2) - (\rho\lambda_1\lambda_2)^2] \right\} \\
 & + \sigma_\pi^2 \left\{ [((1 + \varphi\phi_y)^2 + \beta_y^4) - 2(1 + \varphi\phi_y)\beta_y^2(\rho + \lambda_1 + \lambda_2) + ((1 + \varphi\phi_y)^2 + \beta_y^4) \right. \\
 & \left. (\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2)] - \rho\lambda_1\lambda_2[ ((1 + \varphi\phi_y)^2 + \beta_y^4)(\rho + \lambda_1 + \lambda_2) - 2(1 + \varphi\phi_y)\beta_y^2 \cdot \right. \\
 & \left. (\rho\lambda_1 + \rho\lambda_2 + \lambda_1\lambda_2) + ((1 + \varphi\phi_y)^2 + \beta_y^4)\rho\lambda_1\lambda_2] \right\}, \tag{A.10}
 \end{aligned}$$

$$\lambda_1 + \lambda_2 = \frac{\beta_y^2 + (\gamma\varphi + \lambda + \lambda\varphi\phi_y)\beta_\pi^2}{1 + \gamma\varphi\phi_\pi + \varphi\phi_y}, \tag{A.11}$$

$$\lambda_1\lambda_2 = \frac{\lambda\beta_y^2\beta_\pi^2}{1 + \gamma\varphi\phi_\pi + \varphi\phi_y}. \tag{A.12}$$

From these expressions, it is easy to see that  $G_1(\beta_y, \beta_\pi)$  and  $G_2(\beta_y, \beta_\pi)$  are analytic functions with respect to  $\beta_y$  and  $\beta_\pi$ , *independent of  $\alpha$* .

Finally, the covariance between output gap and inflation is given as

$$\begin{aligned}
E(y_t \pi_t) = & \left( -\sigma_y^2 \gamma (-(1 + \gamma \varphi \phi_\pi + \varphi \phi_y)(1 + \gamma \varphi \phi_\pi + \varphi \phi_y + \beta_y^2 \rho) + \beta_\pi^4 \lambda (1 + \gamma \varphi \phi_\pi \right. \\
& + \varphi \phi_y + \beta_y^2 \rho)(\lambda + \gamma \varphi + \lambda \varphi \phi_y) + \beta_\pi^2 \rho [\beta_y^4 \lambda + \beta_y^2 \lambda \rho (1 + \gamma \varphi \phi_\pi + \varphi \phi_y) + \gamma \varphi \\
& (-1 + \lambda \phi_\pi)(1 + \gamma \varphi \phi_\pi + \varphi \phi_y)] - \beta_y^2 \beta_\pi^6 \lambda^2 \rho (\beta_y^2 \lambda + \rho(\lambda + \gamma \varphi + \lambda \varphi \phi_y))) + \sigma_\pi^2 \varphi \\
& (-\phi_\pi (1 + \gamma \varphi \phi_\pi + \varphi \phi_y) [-\beta_y^4 - \beta_y^2 \gamma \rho \varphi \phi_\pi + (1 + \varphi \phi_y)(1 + \gamma \varphi \phi_\pi + \varphi \phi_y)] \\
& + \beta_y^2 \beta_\pi^6 \lambda \rho [-\gamma \varphi (-\beta_y^2 + \rho + \rho \varphi \phi_y) + \lambda \rho (\beta_y^4 - (1 + \varphi \phi_y)^2)] \\
& + \beta_\pi^4 (\gamma \varphi (1 - \beta_y^2 \rho + \varphi \phi_y)(1 + \gamma \varphi \phi_\pi + \varphi \phi_y) + \lambda (-1 + \beta_y^2 \rho - \varphi(\gamma \phi_\pi + \phi_y)) \\
& (\beta_y^4 - (1 + \varphi \phi_y)^2) + \beta_y^2 \lambda^2 \rho \phi_\pi (-\beta_y^4 + (1 + \varphi \phi_y)^2)) + \beta_\pi^2 \rho [-\beta_y^6 \lambda \rho \phi_\pi \\
& + \beta_y^2 \lambda \rho \phi_\pi (1 + \varphi \phi_y)(1 + \gamma \varphi \phi_\pi + \varphi \phi_y) - (-1 + \lambda \phi_\pi) \\
& (1 + \varphi \phi_y)^2 (1 + \gamma \varphi \phi_\pi + \varphi \phi_y) + \beta_y^4 (-1 - \varphi(\gamma \phi_\pi + \varphi_y) + \lambda(\phi_\pi + \varphi \phi_\pi \phi_y))] \Big) / \\
& \left( (-1 + \rho^2)(-1 + \beta_y^2 \beta_\pi^2 \lambda - \varphi(\gamma \phi_\pi + \phi_y))(1 + \beta_y^2 \rho (-1 + \beta_\pi^2 \lambda \rho) + \gamma \varphi \phi_\pi + \varphi \phi_y \right. \\
& - \beta_\pi^2 \rho (\lambda + \gamma \varphi + \lambda \varphi \phi_y)) (\beta_y^4 (-1 + \beta_\pi^4 \lambda^2) + 2 \beta_y^2 \beta_\pi^2 \gamma \varphi (-1 + \lambda \phi_\pi) \\
& \left. + (1 + \gamma \varphi \phi_\pi + \varphi \phi_y)^2 - \beta_\pi^4 (\lambda + \gamma \varphi + \lambda \varphi \phi_y)^2) \right). \tag{A.13}
\end{aligned}$$

### A.6 E-Stability of BLE for a baseline NK model

This appendix shows the E-stability condition in Corollary 2. Based on Proposition 2, we only need to show that both of the eigenvalues of  $(\mathbf{I} - \mathbf{B}\beta^2)^{-1}(\mathbf{B} - \mathbf{I})$  have negative real parts if  $\gamma(\phi_\pi - 1) + (1 - \lambda)\phi_y > 0$ .

The characteristic polynomial of  $(\mathbf{I} - \mathbf{B}\beta^2)^{-1}(\mathbf{B} - \mathbf{I})$  is given by  $h(\nu) = \nu^2 - c_1\nu + c_2$ , where  $c_1$  is the trace and  $c_2$  is the determinant of matrix  $(\mathbf{I} - \mathbf{B}\beta^2)^{-1}(\mathbf{B} - \mathbf{I})$ . Direct calculation shows that

$$c_1 = \frac{-(1 - \lambda)(1 - \beta_y^2) - 2\varphi(\gamma \phi_\pi + \phi_y) + \varphi(\gamma + \lambda \phi_y)(1 + \beta_\pi^2)}{\Delta (1 + \gamma \varphi \phi_\pi + \varphi \phi_y)}, \tag{A.14}$$



$$c_2 = \frac{\varphi[\gamma(\phi_\pi - 1) + (1 - \lambda)\phi_y]}{\Delta(1 + \gamma\varphi\phi_\pi + \varphi\phi_y)}, \quad (\text{A.15})$$

where  $\Delta = \frac{(1 - \beta_y^2)(1 - \lambda\beta_\pi^2) + \gamma\varphi\phi_\pi + \varphi\phi_y - (\gamma\varphi + \lambda\varphi\phi_y)\beta_\pi^2}{1 + \gamma\varphi\phi_\pi + \varphi\phi_y}$ .

Both of the eigenvalues of  $(\mathbf{I} - \mathbf{B}\beta^2)^{-1}(\mathbf{B} - \mathbf{I})$  have negative real parts if and only if  $c_1 < 0$  and  $c_2 > 0$  (these conditions are obtained by applying the *Routh-Hurwitz criterion theorem* (see [Brock and Malliaris, 1989](#))). If  $\gamma(\phi_\pi - 1) + (1 - \lambda)\phi_y > 0$ , from Appendix A.4, it is easy to see that  $\Delta > 0$ . Furthermore,

$$c_1 \leq \frac{-2\varphi[(\gamma(\phi_\pi - 1) + (1 - \lambda)\phi_y)]}{\Delta(1 + \gamma\varphi\phi_\pi + \varphi\phi_y)} < 0, \quad c_2 > 0.$$

## APPENDIX B: ITERATIVE E-STABILITY ALGORITHM FOR BLE

This section discusses the Iterative E-stability algorithm used in the approximation of BLE. The first-order autocorrelation coefficients  $\beta^*$  in a BLE are functions in terms of the structural parameters  $\mu$ , which satisfy the nonlinear equilibrium conditions  $G(\beta^*, \mu) = \beta^*$  in (2.41). In order to find a BLE for a given  $\mu$ , we use a simple fixed-point iteration, which is formalized below in Algorithm I.

---

### Algorithm I: Approximation of a BLE using Iterative E-stability

---

Denote the set of structural parameters by  $\mu$ , and the first-order autocorrelation function for a given  $\mu$  by  $G(\beta^{(k)}, \mu)$ .

- Step (0): Initialize the vector of learning parameters at  $\beta^{(0)}$ .
- Step (I): At each iteration  $k$ , using the first-order autocorrelation functions, update the vector of learning parameters as

$$\beta^{(k)} = G(\beta^{(k-1)}, \mu), \quad (\text{B.1})$$

where  $G(\beta^{(k-1)}, \mu)$  is known from iteration  $k - 1$ .

- Step (II): Terminate if  $\|\beta^{(k)} - \beta^{(k-1)}\|_p < \epsilon$ , for a small scalar  $\epsilon > 0^4$  and a suitable norm distance  $\|\cdot\|_p$ , otherwise repeat Step (I).
-

A BLE  $(\mathbf{0}, \beta^*)$  is locally stable under (B.1) if all eigenvalues of  $DG_\beta(\beta^*)$  lie inside the unit circle. Then the equilibrium is said to be *iteratively E-stable*. When Algorithm I terminates for some  $K$  at a small pre-specified  $\epsilon$ , we say that it has converged to  $\beta^{(K)}$ . Note that if Algorithm I converges, it converges to an approximate BLE since

$$\|\beta^{(K+1)} - \beta^{(K)}\| < \epsilon \Rightarrow \|G(\beta^{(K)}) - \beta^{(K)}\| < \epsilon \Rightarrow G(\beta^{(K)}) \approx \beta^{(K)}.$$

Given a vector of initial values for the first-order autocorrelation coefficients of the forward-looking variables, we use  $N = 200$  iterations under (2.41) for each parameter draw and use the resulting fixed-point as an approximate BLE. The algorithm typically takes less than 50 iterations to converge for a given parameter draw, hence  $N = 200$  is a conservative value.

## APPENDIX C: REDUCED-FORM MATRICES

Recall that we consider linear DSGE models in the following general form, as described in Section 2.1:

$$\mathbf{x}_t = \mathbf{F}(\mathbf{x}_{t+1}^e, \mathbf{u}_t, \mathbf{v}_t) = \mathbf{b}_0 + \mathbf{b}_1 \mathbf{x}_{t+1}^e + \mathbf{b}_2 \mathbf{x}_{t-1} + \mathbf{b}_3 \mathbf{u}_t + \mathbf{b}_4 \mathbf{v}_t, \quad (\text{C.1})$$

$$\mathbf{u}_t = \mathbf{a} + \rho \mathbf{u}_{t-1} + \boldsymbol{\varepsilon}_t. \quad (\text{C.2})$$

This appendix derives the reduced-form matrices  $\widehat{\mathbf{A}}$ ,  $\widehat{\mathbf{B}}$  and  $\widehat{\mathbf{C}}$  for the equilibrium models, and  $\widehat{\mathbf{A}}_t$ ,  $\widehat{\mathbf{B}}_t$  and  $\widehat{\mathbf{C}}_t$  for the learning models. The reduced-form matrices are used for the Kalman filter discussed in Appendix D.

### C.1 Equilibrium Models

*REE* Using the MSV solution, expectations under REE are given by

$$E_t \mathbf{x}_{t+1} = \mathbf{c}_0 + \mathbf{c}_2 \mathbf{a} + \mathbf{c}_1 \mathbf{x}_t + \mathbf{c}_2 \rho \mathbf{u}_t. \quad (\text{C.3})$$

<sup>4</sup>In the remainder of this paper, we use the common  $L^1$ -Norm as our norm distance, i.e.,  $\|\beta^{(k)} - \beta^{(k-1)}\|_p = \sum_{j=1}^N |\beta_j^{(k)} - \beta_j^{(k-1)}|$ .

1 Plugging back into (C.1) and rewriting yields

$$2 \quad \mathbf{X}_t = \widehat{\mathbf{A}} + \widehat{\mathbf{B}}\mathbf{X}_{t-1} + \widehat{\mathbf{C}}\boldsymbol{\eta}_t, \quad (C.4)$$

$$3 \quad \text{where } \widehat{\mathbf{A}} = \begin{bmatrix} I - \mathbf{b}_1\mathbf{c}_1 - (\mathbf{b}_1\mathbf{c}_2\rho + \mathbf{b}_3) & \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_0 + \mathbf{b}_1\mathbf{c}_0 + \mathbf{b}_1\mathbf{c}_2\mathbf{a} \\ \mathbf{a} \end{bmatrix},$$

$$4 \quad \widehat{\mathbf{B}} = \begin{bmatrix} I - \mathbf{b}_1\mathbf{c}_1 - (\mathbf{b}_1\mathbf{c}_2\rho + \mathbf{b}_3) & \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_2 \ 0 \\ 0 \ \rho \end{bmatrix},$$

$$5 \quad \widehat{\mathbf{C}} = \begin{bmatrix} I - \mathbf{b}_1\mathbf{c}_1 - (\mathbf{b}_1\mathbf{c}_2\rho + \mathbf{b}_3) & \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_4 \ 0 \\ 0 \ I \end{bmatrix}, \mathbf{X}'_t = [\mathbf{x}'_t, \mathbf{u}'_t], \boldsymbol{\eta}'_t = [\mathbf{v}'_t, \boldsymbol{\varepsilon}'_t].$$

6 *BLE* Expectations under BLE are given by

$$7 \quad E_t \mathbf{x}_{t+1} = \boldsymbol{\alpha}^* + \boldsymbol{\beta}^{*2}(\mathbf{x}_{t-1} - \boldsymbol{\alpha}^*). \quad (C.5)$$

8 Plugging back into (C.1) and rewriting yields

$$9 \quad \mathbf{X}_{t+1} = \widehat{\mathbf{A}} + \widehat{\mathbf{B}}\mathbf{X}_{t-1} + \widehat{\mathbf{C}}\boldsymbol{\eta}_t, \quad (C.6)$$

$$10 \quad \text{where } \widehat{\mathbf{A}} = \begin{bmatrix} I - \mathbf{b}_3 & \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_0 + \mathbf{b}_1\boldsymbol{\alpha}^* - \mathbf{b}_1\boldsymbol{\beta}^{*2}\boldsymbol{\alpha}^* \\ \mathbf{a} \end{bmatrix}, \widehat{\mathbf{B}} = \begin{bmatrix} I - \mathbf{b}_3 & \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_1\boldsymbol{\beta}^{*2} + \mathbf{b}_2 \ 0 \\ 0 \ \rho \end{bmatrix},$$

$$11 \quad \widehat{\mathbf{C}} = \begin{bmatrix} I - \mathbf{b}_3 & \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_4 \ 0 \\ 0 \ I \end{bmatrix}.$$

## 12 C.2 Learning Models

13 *SAC-learning* Expectations under SAC-learning are given by

$$14 \quad E_t \mathbf{x}_{t+1} = \boldsymbol{\alpha}_{t-1} + \boldsymbol{\beta}_{t-1}^2(\mathbf{x}_{t-1} - \boldsymbol{\alpha}_{t-1}). \quad (C.7)$$

15 Plugging back into (C.1) and rewriting yields

$$16 \quad \mathbf{X}_t = \widehat{\mathbf{A}}_{t-1} + \widehat{\mathbf{B}}_{t-1}\mathbf{X}_{t-1} + \widehat{\mathbf{C}}_{t-1}\boldsymbol{\eta}_t, \quad (C.8)$$

$$\text{where } \hat{\mathbf{A}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_0 + \mathbf{b}_1 \alpha_{t-1} - \mathbf{b}_1 \beta_{t-1}^2 \alpha_{t-1} \\ \mathbf{a} \end{bmatrix},$$

$$\hat{\mathbf{B}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_1 \beta_{t-1}^2 + \mathbf{b}_2 \ 0 \\ 0 \quad \boldsymbol{\rho} \end{bmatrix},$$

$$\hat{\mathbf{C}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_4 \ 0 \\ 0 \quad I \end{bmatrix}.$$

*AR(2)-learning* Expectations under AR(2)-learning are given by

$$E_t \mathbf{x}_{t+1} = \alpha_{t-1} + \beta_{1,t-1} \mathbf{x}_{t-1} + \beta_{2,t-1} \mathbf{x}_{t-2}. \quad (\text{C.9})$$

Plugging back into (C.1) and rewriting yields

$$\mathbf{X}_t = \tilde{\mathbf{A}}_{t-1} + \tilde{\mathbf{B}}_{t-1} \mathbf{X}_{t-1} + \tilde{\mathbf{C}}_{t-1} \mathbf{X}_{t-2} + \tilde{\mathbf{D}}_{t-1} \boldsymbol{\eta}_t, \quad (\text{C.10})$$

$$\text{where } \tilde{\mathbf{A}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_0 + \mathbf{b}_1 \alpha_{t-1} \\ \mathbf{a} \end{bmatrix}, \tilde{\mathbf{B}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_1 \beta_{1,t-1} + \mathbf{b}_2 \ 0 \\ 0 \quad \boldsymbol{\rho} \end{bmatrix},$$

$$\tilde{\mathbf{C}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_1 \beta_{2,t-1} \ 0 \\ 0 \quad 0 \end{bmatrix}, \tilde{\mathbf{D}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_4 \ 0 \\ 0 \quad I \end{bmatrix}.$$

This can be further rewritten as

$$\tilde{\mathbf{X}}_t = \hat{\mathbf{A}}_{t-1} + \hat{\mathbf{B}}_{t-1} \tilde{\mathbf{X}}_{t-1} + \hat{\mathbf{C}}_{t-1} \boldsymbol{\eta}_t, \quad (\text{C.11})$$

$$\text{where } \tilde{\mathbf{X}}_t = \begin{bmatrix} \mathbf{X}_t \\ \mathbf{X}_{t-1} \end{bmatrix}, \hat{\mathbf{A}}_{t-1} = \begin{bmatrix} I - \tilde{\mathbf{B}}_{t-1} \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{A}}_{t-1} \\ 0 \end{bmatrix}, \hat{\mathbf{B}}_{t-1} = \begin{bmatrix} I - \tilde{\mathbf{B}}_{t-1} \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} 0 \ \tilde{\mathbf{C}}_{t-1} \\ I \ 0 \end{bmatrix},$$

$$\hat{\mathbf{C}}_{t-1} = \begin{bmatrix} I - \tilde{\mathbf{B}}_{t-1} \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{D}}_{t-1} \\ 0 \end{bmatrix}.$$

*Pseudo-MSV and VAR(1)-learning* Expectations under pseudo-MSV and VAR(1)-learning are given by

$$E_t \mathbf{x}_{t+1} = \gamma_{0,t-1} + \gamma_{1,t-1} \mathbf{x}_{t-1} + \gamma_{2,t-1} \boldsymbol{\rho} u_{t-1}. \quad (\text{C.12})$$

1 Plugging back into (C.1) and rewriting yields

$$2 \quad \mathbf{X}_t = \widehat{\mathbf{A}}_{t-1} + \widehat{\mathbf{B}}_{t-1} \mathbf{X}_{t-1} + \widehat{\mathbf{C}}_{t-1} \boldsymbol{\eta}_t, \quad (C.13) \quad 3$$

$$4 \quad \text{where } \widehat{\mathbf{A}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_0 + \mathbf{b}_1 \gamma_{0,t-1} \\ \mathbf{a} \end{bmatrix}, \widehat{\mathbf{B}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_1 \gamma_{1,t-1} + \mathbf{b}_1 \mathbf{b}_2 \gamma_{2,t-1} \boldsymbol{\rho} \\ 0 \quad \boldsymbol{\rho} \end{bmatrix}, \quad 4$$

$$5 \quad \widehat{\mathbf{C}}_{t-1} = \begin{bmatrix} I - \mathbf{b}_3 \\ 0 \quad I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_4 \quad 0 \\ 0 \quad I \end{bmatrix}. \quad 6$$

7  
8  
9 VAR(1)-learning is the special case with  $\gamma_{2,t-1} = 0$ .

#### 10 APPENDIX D: KALMAN FILTER AND ESTIMATION

11  
12 This section describes the Kalman filter used in the estimation of equilibrium  
13 and adaptive learning models.

##### 14 D.1 Kalman Filter for Equilibrium Models

15  
16 For the REE model, expectations  $\mathbf{x}_{t+1}^e$  are pinned down by the equilibrium con-  
17 ditions (2.15)–(2.18). The fixed point of the system (2.15)–(2.18) is computed us-  
18 ing standard methods, as in Uhlig (1995).

19 For the BLE model, expectations  $\mathbf{x}_{t+1}^e$  are pinned down by the equilibrium  
20 conditions (consistency requirements) in (2.27). The fixed point associated with  
21 (2.27) is approximated by the iterative E-stability algorithm in (2.41), which con-  
22 verges to the corresponding equilibrium persistence values  $\boldsymbol{\beta}^*$  (see Appendix B).

23 Both REE and BLE models can be written in the following recursive form:  
24

$$25 \quad \mathbf{X}_t = \widehat{\mathbf{A}} + \widehat{\mathbf{B}} \mathbf{X}_{t-1} + \widehat{\mathbf{C}} \boldsymbol{\eta}_t, \quad (D.1) \quad 25$$

26  
27 with  $\mathbf{X}'_t = [\mathbf{x}'_t, \mathbf{u}'_t]$ ,  $\boldsymbol{\eta}'_t = [\mathbf{v}'_t, \boldsymbol{\varepsilon}'_t]$  and  $\widehat{\mathbf{A}}$ ,  $\widehat{\mathbf{B}}$  and  $\widehat{\mathbf{C}}$  matrices of structural parameters.  
28 REE and BLE models are characterized by different matrices  $\widehat{\mathbf{B}}$  and  $\widehat{\mathbf{C}}$ , as de-  
29 scribed in Appendix C.1. Given the linear structure of both models, the likelihood  
30 function can be evaluated using a standard Kalman filter.

31 We denote the initial state vector and state covariance matrix by  $\mathbf{X}_{0|0}$  and  $\mathbf{P}_{0|0}$ ,  
32 respectively, while L refers to the number of shocks. The measurement equations

are denoted by  $\mathbf{Y}_t = \bar{\phi} + \phi_1 \mathbf{X}_t$ , with  $\mathbf{Y}_t$  denoting the vector of observable variables.<sup>5</sup> For every period  $t$ , the Kalman filter recursions are given as follows:

$$\left\{ \begin{array}{l} \mathbf{X}_{t|t-1} = \hat{\mathbf{A}} + \hat{\mathbf{B}}\mathbf{X}_{t-1|t-1}, \\ \mathbf{P}_{t|t-1} = \hat{\mathbf{B}}\mathbf{P}_{t-1|t-1}\hat{\mathbf{B}}' + \hat{\mathbf{C}}\Sigma_\eta\hat{\mathbf{C}}', \\ v_t = \mathbf{Y}_t - \bar{\phi} - \phi_1\mathbf{X}_{t|t-1}, \\ \Sigma_t = \phi_1\mathbf{P}_{t|t-1}\phi_1', \\ \mathbf{X}_{t|t} = \mathbf{X}_{t|t-1} + \mathbf{P}_{t|t-1}\phi_1'\Sigma_t^{-1}v_t, \\ \mathbf{P}_{t|t} = \mathbf{P}_{t|t-1}\phi_1'\Sigma_t^{-1}\phi_1\mathbf{P}_{t|t-1}, \\ L(y_t) = -\frac{L}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma_t| - \frac{1}{2}(v_t'\Sigma_t^{-1}v_t). \end{array} \right. \quad (\text{D.2})$$

## D.2 Kalman Filter for Learning Models

Learning models can be represented as a recursive linear system after plugging in the expectations

$$\mathbf{X}_t = \hat{\mathbf{A}}_{t-1} + \hat{\mathbf{B}}_{t-1}\mathbf{X}_{t-1} + \hat{\mathbf{C}}_{t-1}\boldsymbol{\eta}_t, \quad (\text{D.3})$$

with time-varying matrices  $\hat{\mathbf{B}}_{t-1}$ ,  $\hat{\mathbf{C}}_{t-1}$  and perceived mean vector  $\hat{\mathbf{A}}_{t-1}$ , where the time-variation comes from agents' PLM coefficients. The coefficients are updated every period using the SAC-learning or constant gain recursive least squares algorithms in (3.2)–(3.6). Given the  $t - 1$  timing structure discussed in Section 3.1, the learning models admit a conditionally linear structure given the belief coefficients. Accordingly, the likelihood function can be evaluated using the standard Kalman filter recursions conditional on the belief coefficients.

We denote the initial perceived covariance matrix and initial belief coefficients that appear in the learning algorithms (3.2)–(3.6) by  $R_0$  and  $\theta_0$ . For every period

<sup>5</sup>The SW07 model considered in this paper consists of GDP, consumption, wage and investment growth, CPI inflation, hours worked and interest rates.

t, the conditionally linear Kalman filter recursions are given as follows:

$$\left\{ \begin{array}{l}
 \textbf{Kalman filter step:} \\
 \mathbf{X}_{t|t-1} = \widehat{\mathbf{A}}_{t-1} + \widehat{\mathbf{B}}_{t-1} \mathbf{X}_{t-1|t-1}, \\
 \mathbf{P}_{t|t-1} = \widehat{\mathbf{B}}_{t-1} \mathbf{P}_{t-1|t-1} \widehat{\mathbf{B}}_{t-1}' + \widehat{\mathbf{C}}_{t-1} \Sigma_{\eta} \widehat{\mathbf{C}}_{t-1}', \\
 v_t = \mathbf{Y}_t - \bar{\phi} - \phi_1 \mathbf{X}_{t|t-1}, \\
 \Sigma_t = \phi_1 \mathbf{P}_{t|t-1} \phi_1', \\
 \mathbf{X}_{t|t} = \mathbf{X}_{t|t-1} + \mathbf{P}_{t|t-1} \phi_1' \Sigma_t^{-1} v_t, \\
 \mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} \phi_1' \Sigma_t^{-1} \phi_1 \mathbf{P}_{t|t-1}, \\
 L(y_t | R_{t-1}, \theta_{t-1}) = -\frac{L}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma_t| - \frac{1}{2} (v_t' \Sigma_t^{-1} v_t),
 \end{array} \right. \quad \left\{ \begin{array}{l}
 \textbf{Learning step:} \\
 R_t = T_1(R_{t-1}, \mathbf{X}_{t|t}), \\
 \theta_t = T_2(\theta_{t-1}, \mathbf{X}_{t|t}, R_t),
 \end{array} \right. \quad (\text{D.4})$$

where  $T_1$  and  $T_2$  correspond to the updating equations for belief coefficients, as outlined in (3.2)–(3.6). In this context, we treat the learning models as a *temporary equilibrium* system, where the matrices  $\widehat{\mathbf{B}}_{t-1}$  and  $\widehat{\mathbf{C}}_{t-1}$  and vector  $\widehat{\mathbf{A}}_{t-1}$  are updated every period. The Kalman filter step treats the model as a linear system for a given  $\theta_t$  and returns the likelihood function and state variables  $\mathbf{X}_{t|t}$ . The state variables are used as an input to update the belief coefficients in the learning step.

### D.3 Estimation

The model parameters are estimated using standard Bayesian likelihood methods. This consists of a posterior mode search step and a Monte Carlo Markov Chain (MCMC) step, which are summarized below.

1. Prior distributions of the estimated parameters are specified. For the SW07 model used in this paper, the prior distributions are summarized in Table E.1. We denote the prior distribution function by  $p(\boldsymbol{\mu})$ .
2. The likelihood function  $p(\boldsymbol{\mu}) * L(y_t | \boldsymbol{\mu})$  is maximized using standard iterative gradient descent algorithms. We use the *csmmwel* algorithm (Sims, 1999) available in MATLAB software.

- 1 • The algorithm iteratively updates the parameters  $\mu$  until convergence. 1  
2 We denote each iteration of parameter draws by  $\mu^n$ . 2
  - 3 • Equilibrium models: For every parameter draw  $\mu^n$ , the equilibrium is 3  
4 calculated by finding the fixed point of (2.15)–(2.18) and (2.27) for REE 4  
5 and BLE models, respectively. For the REE model the fixed point is 5  
6 found by using Uhlig’s method (1995). For the BLE model, the iterative 6  
7 E-stability algorithm in (B.1) with 200 iterations is used. Given the 7  
8 reduced-form matrices  $\hat{A}(\mu^n)$ ,  $\hat{B}(\mu^n)$  and  $\hat{C}(\mu^n)$ , the likelihood function 8  
9  $p(\mu^n) * L(y_t|\mu^n)$  is computed using the Kalman filter recursions in (D.2). 9  
10  
11 • Learning models: For every parameter draw  $\mu^n$ , the reduced-form ma- 11  
12 trices  $\hat{A}_{t-1}(\mu^n)$ ,  $\hat{B}_{t-1}(\mu^n)$  and  $\hat{C}_{t-1}(\mu^n)$  are re-calculated at every period 12  
13  $t$  of the Kalman filter in (D.4). Given the reduced-form matrices, the like- 13  
14 lihood function  $p(\mu^n) * L(y_t|\mu^n)$  is computed. 14  
15  
16 • The parameter values  $\mu^n$  are iteratively updated until the likelihood 15  
17 function  $p(\mu^n) * L(y_t|\mu^n)$  converges. The optimized parameter values  $\mu^*$  16  
18 are referred to as the posterior mode. 17  
18
- 19 3. For each model, the MCMC algorithm is initialized at  $(\mu^*, c\Sigma_\mu^*)$ , where  $\Sigma_\mu^*$  19  
20 denotes the covariance matrix of  $\mu^*$ .  $c$  is a scaling coefficient tuned to obtain 20  
21 an average acceptance ratio between 30 and 45%. For each model, we use 21  
22 two parallel chains with 500000 draws and discard the first half as a burn- 22  
23 in sample. The second half of the chains is used to compute the posterior 23  
24 moments reported in Tables 1 and 3. 24

25 For the BLE model, the initial values of the first-order autocorrelation co- 25  
26 efficients of the AR(1) beliefs  $\beta^{(0)}$ , i.e., Step (0) of Algorithm I in Appendix 26  
27 B, are fixed prior to the estimation. For forward-looking variables that are 27  
28 observable, i.e., inflation  $\pi_t$  and hours worked  $l_t$ , the initial values are set 28  
29 to the corresponding first-order sample-autocorrelation over the estimation 29  
30 period. For the remaining latent forward-looking variables, we take the un- 30  
31 conditional first-order autocorrelations implied by the estimated REE. We 31  
32



keep the initial values  $\beta^{(0)}$  fixed at these values for all parameter draws  $\mu^n$ . Parameter draws where the fixed-point iteration fails to converge to a stationary equilibrium are discarded.

## APPENDIX E: THE SMETS-WOUTERS 2007 MODEL

### E.1 Model Descriptions

The model consists of 13 equations linearized around the steady-state growth path, supplemented with seven exogenous structural shocks. We deviate from the benchmark model by slightly restricting the parameter space of the model, where we assume all shocks follow an AR(1) process.<sup>6</sup> In this section we briefly outline the resulting linearized model economy that is used in our estimation. To start with the demand side of the economy, the aggregate resource constraint is given by

$$\begin{cases} \tilde{y}_t = c_y c_t + i_y i_t + z_y z_t + \epsilon_t^g, \\ \epsilon_t^g = \rho_g \epsilon_{t-1}^g + \eta_t^g, \end{cases} \quad (\text{E.1})$$

where  $\tilde{y}_t, c_t, i_t$  and  $z_t$  are the output, consumption, investment and capital utilization rate, respectively, while  $c_y, i_y$  and  $z_y$  are the steady-state shares in output of the respective variables. The second equation in (E.1) defines the exogenous spending shock  $\epsilon_t^g$ , where  $\eta_t^g$  is an i.i.d.-normal disturbance for spending. The consumption Euler equation is given by

$$\begin{cases} c_t = c_1 c_{t-1} + (1 - c_1) \mathbb{E}_t c_{t+1} + c_2 (l_t - \mathbb{E}_t l_{t+1}) - c_3 (r_t - \mathbb{E}_t \pi_{t+1}) + \epsilon_t^b, \\ \epsilon_t^b = \rho_b \epsilon_{t-1}^b + \eta_t^b, \end{cases} \quad (\text{E.2})$$

with  $c_1 = \frac{\lambda}{\gamma} / (1 + \frac{\lambda}{\gamma}), c_2 = (\sigma_c - 1)(w_{ss} l_{ss} / c_{ss}) / (\sigma_c (1 + \frac{\lambda}{\gamma})), c_3 = (1 - \frac{\lambda}{\gamma}) / ((1 + \frac{\lambda}{\gamma}) \sigma_c)$ , where  $\lambda, \gamma$  and  $\sigma_c$  denote the habit formation in consumption, steady state-growth rate and the elasticity of intertemporal substitution, respectively, while

<sup>6</sup>In particular, the benchmark model has more structure on the exogenous shocks, where the mark-up shocks each follow an ARMA(1,1) process and the technology and government spending shocks follow a VAR(1) process. We refer the reader to SW for more details about the microfoundations of the benchmark model.

$x_{ss}$  corresponds to the steady-state level of a given variable  $x$ . The equation implies that current consumption is a weighted average of the past and expected future consumption, expected growth in hours worked and the ex-ante real interest rate.  $\epsilon_t^b$  corresponds to the risk premium shock modeled as an AR(1) process, where  $\eta_t^b$  is an i.i.d-normal disturbance. Next, the investment Euler equation is defined as

$$\begin{cases} i_t = i_1 i_{t-1} + (1 - i_1) \mathbb{E}_t i_{t+1} + i_2 q_t + \epsilon_t^i, \\ \epsilon_t^i = \rho_i \epsilon_{t-1}^i + \eta_t^i, \end{cases} \quad (\text{E.3})$$

with  $i_1 = \frac{1}{1+\beta\gamma}$ ,  $i_2 = \frac{1}{(1+\beta\gamma)(\gamma^2\phi)}$ , where  $\bar{\beta} = \beta\gamma^{-\sigma_c}$ ,  $\phi$  is the steady-state elasticity of capital adjustment cost and  $\beta$  is the HH discount factor.  $q_t$  denotes the real value of existing capital stock. Similar to the consumption Euler, the equation implies that investment is a weighted average of past and expected future consumption, as well as the real value of existing capital stock.  $\epsilon_t^i$  represents the AR(1) investment shock, where  $\eta_t^i$  is an i.i.d-normal disturbance. The value of the capital-arbitrage equation is given by

$$q_t = q_1 \mathbb{E}_t q_{t+1} + (1 - q_1) \mathbb{E}_t r_{t+1}^k - (r_t - \mathbb{E}_t \pi_{t+1}) + \frac{1}{c_3} \epsilon_t^b, \quad (\text{E.4})$$

with  $q_1 = \bar{\beta}(1 - \delta)$ , implying the real value of capital stock is a weighted average of its expected future value and expected real rental rate on capital, net of ex-ante real interest rate and the risk premium shock. The production function is characterized as

$$\begin{cases} \tilde{y}_t = \phi_p (\alpha k_t^s + (1 - \alpha) l_t + \epsilon_t^a), \\ \epsilon_t^a = \rho_a \epsilon_{t-1}^a + \eta_t^a, \end{cases} \quad (\text{E.5})$$

where  $k_t^s$  denotes the capital services used in production,  $\alpha$  is the share of capital in production and  $\phi_p$  is (one plus) the share of fixed costs in production.  $\epsilon_t^a$  denotes the AR(1) total factor productivity shock. Capital is assumed to be the sum of the previous amount of capital services used and the degree of capital utilization. Hence,

$$k_t^s = k_{t-1} + z_t. \quad (\text{E.6})$$

Moreover, the degree of capital utilization is a positive function of the degree of rental rate,  $z_t = z_1 r_t^k$ , with  $z_1 = \frac{1-\psi}{\psi}$ ,  $\psi$  being the elasticity of the capital utilization adjustment cost. Next, the equation for installed capital is given by

$$k_t = k_1 k_{t-1} + (1 - k_1) i_t + k_2 \epsilon_t^i, \quad (\text{E.7})$$

with  $k_1 = \frac{1-\delta}{\gamma}$ ,  $k_2 = (1 - \frac{1-\delta}{\gamma})(1 + \bar{\beta}\gamma)\gamma^2\phi$ . The price mark-up equation is given by

$$\mu_t^p = \alpha(k_t^s - l_t) + \epsilon_t^a - w_t, \quad (\text{E.8})$$

which means the price mark-up  $\mu_t^p$  is the marginal product of the labor net of the current wage. The NKPC is characterized as

$$\begin{cases} \pi_t = \pi_1 \mathbb{E}_t \pi_{t+1} - \pi_2 \mu_t^p + \epsilon_t^p, \\ \epsilon_t^p = \rho_p \epsilon_{t-1}^p + \eta_t^p, \end{cases} \quad (\text{E.9})$$

with  $\pi_1 = \bar{\beta}\gamma$ ,  $\pi_2 = (1 - \beta\gamma\xi_p)(1 - \xi_p)/[\xi_p((\phi_p - 1)\epsilon_p + 1)]$ , where  $\xi_p$  corresponds to the degree of price stickiness, while  $\epsilon_p$  denotes the Kimball goods market aggregator. The equation implies that current inflation is determined by the expected future inflation, the price mark-up and the AR(1) price mark-up shock  $\epsilon_t^p$ , where  $\eta_t^p$  is an i.i.d-normal disturbance. The rental rate of capital is given by

$$r_t^k = -(k_t - l_t) + w_t, \quad (\text{E.10})$$

which implies the rental rate of capital is decreasing in the capital-labor ratio and increasing in the real wage. The wage mark-up is given as the real wages net of the marginal rate of substitution between working and consuming. Hence,

$$\mu_t^w = w_t - (\sigma_l l_t + \frac{1}{1 - \lambda/\gamma}(c_t - \frac{\lambda}{\gamma} c_{t-1})), \quad (\text{E.11})$$

where  $\sigma_l$  denotes the elasticity of labor supply. The real wage equation is given by

$$\begin{cases} w_t = w_1 w_{t-1} + (1 - w_1)(\mathbb{E}_t w_{t+1} + \mathbb{E}_t \pi_{t+1}) - w_2 \mu_t^w + \epsilon_t^w, \\ \epsilon_t^w = \rho_w \epsilon_{t-1}^w + \eta_t^w, \end{cases} \quad (\text{E.12})$$

with  $w_1 = 1/(1 + \bar{\beta}\gamma)$  and  $w_2 = ((1 - \bar{\beta}\gamma\xi_w)(1 - \xi_w)/(\xi_w(\phi_w - 1)\epsilon_w + 1))$ . Hence, the real wage is a weighted average of the past and expected wage, expected inflation, the wage mark-up and the wage mark-up shock  $\epsilon_t^w$ , where  $\eta_t^w$  is an i.i.d-normal disturbance. Finally, monetary policy is assumed to follow a standard generalized Taylor rule:

$$\begin{cases} r_t = \rho r_{t-1} + (1 - \rho)(\phi_\pi \pi_t + \phi_y y_t) + \phi_{\Delta y}(\Delta y_t) + \epsilon_t^r, \\ \epsilon_t^r = \rho_r \epsilon_{t-1}^r + \eta_t^r, \end{cases} \quad (\text{E.13})$$

where  $y_t$  denotes the output gap and  $\epsilon_t^r$  is the AR(1) monetary policy shock, with  $\eta_t^r$  the i.i.d-normal disturbance. Hence, the monetary policy responds with output gap growth on top of inflation and the output gap. In this paper, following the approach in Slobodyan and Wouters (2012a), the output gap is defined as the deviation of output from the underlying productivity process, i.e.,  $y_t = \tilde{y}_t - \Phi_p \epsilon_t^a$ . The prior distributions used for all estimated parameters are provided in Table E.1.

We use U.S. historical quarterly macroeconomic data for the period 1966:I–2007:IV. The observable variables used in the estimation are the (log-) difference of real GDP ( $y_t^{data}$ ), real consumption ( $c_t^{data}$ ), real investment ( $inv_t^{data}$ ), real wage ( $w_t^{data}$ ), log hours worked ( $l_t^{data}$ ), inflation ( $\pi_t^{data}$ ) and the federal funds rate ( $r_t^{data}$ ) for the U.S economy. The measurement equations are given as

$$\begin{cases} d(\log(y_t^{data})) = \bar{\gamma} + (y_t - y_{t-1}), \\ d(\log(c_t^{data})) = \bar{\gamma} + (c_t - c_{t-1}), \\ d(\log(inv_t^{data})) = \bar{\gamma} + (inv_t - inv_{t-1}), \\ d(\log(w_t^{data})) = \bar{\gamma} + (w_t - w_{t-1}), \\ \log(l_t^{data}) = \bar{l} + l_t, \\ (\log(\pi_t^{data})) = \bar{\pi} + \pi_t, \\ (\log(r_t^{data})) = \bar{r} + r_t. \end{cases} \quad (\text{E.14})$$

Fixed Parameters				
$\delta$	0.025			
$\phi_w$	1.5			
$g$	0.18			
$\epsilon_p$	10			
$\epsilon_w$	10			
Prior				
	Distribution	Mean	Var.	
Parameters related to nominal and real frictions				
$\phi$	Normal	4	1.5	
$\sigma_c$	Normal	1.5	0.375	
$\lambda$	Beta	0.7	0.1	
$\xi_w$	Beta	0.5	0.1	
$\sigma_l$	Normal	2	0.75	
$\xi_p$	Beta	0.5 (0.75)	0.1 (0.05)	
$\psi$	Beta	0.5	0.15	
$\phi_p$	Normal	1.25	0.125	
$\iota_p$	Normal	0.5	0.15	
$\iota_w$	Normal	0.5	0.15	
Policy related parameters				
$\phi_\pi$	Normal	1.5	0.25	
$\rho$	Beta	0.75	0.1	
$\phi_y$	Normal	0.125	0.05	
$\phi_{\Delta y}$	Normal	0.125	0.05	
Steady-state related parameters				
$\bar{\pi}$	Gamma	0.625	0.1	
$\bar{\beta}$	Gamma	0.25	0.1	
$\bar{l}$	Normal	0	2	
$\bar{\gamma}$	Normal	0.4	0.1	
$\alpha$	Normal	0.3	0.05	
Parameters related to shock persistence				
$\rho_a$	Beta	0.5	0.2	
$\rho_b$	Beta	0.5	0.2	
$\rho_g$	Beta	0.5	0.2	
$\rho_i$	Beta	0.5	0.2	
$\rho_r$	Beta	0.5	0.2	
$\rho_p$	Beta	0.5	0.2	
$\rho_w$	Beta	0.5	0.2	
$\rho_{ga}$	Beta	0.5	0.2	
Shock variance parameters				
$\eta_a$	Inv. Gamma	0.1	2	
$\eta_b$	Inv. Gamma	0.1	2	
$\eta_g$	Inv. Gamma	0.1	2	
$\eta_i$	Inv. Gamma	0.1	2	
$\eta_r$	Inv. Gamma	0.1	2	
$\eta_p$	Inv. Gamma	0.1	2	
$\eta_w$	Inv. Gamma	0.1	2	
$\gamma$	Gamma	0.035	0.015	

TABLE E.1. Fixed parameters and the prior distributions of the estimated parameters for the Smets-Wouters (2007) model.

1 The construction of the time series follow the same steps as in Smets and Wouters 1  
2 (2007). 2

### 4 E.2 Deviations from the Original Model 4

5 Our model follows the original Smets and Wouters (2007) structure with mi- 5  
6 nor deviations. First, we use CPI inflation as our inflation measure instead of the 6  
7 GDP deflator used in the original model. Second, we define the output gap in the 7  
8 model as the deviation of output from its natural level based on the productivity 8  
9 process.<sup>7</sup> Third, the observable variables used in the estimation are the (log-) dif- 9  
10 ference of real GDP, real consumption, real investment, real wages, (log-) hours 10  
11 worked, inflation and the federal funds rate for the U.S. economy. The model 11  
12 structure is the same as the original SW except for three minor deviations. The 12  
13 first is the definition of the output gap: in the original model, this is the deviation 13  
14 of output from its potential level, defined as output in the presence of flexible 14  
15 prices and wages. Instead, we follow Slobodyan and Wouters (2012) and define 15  
16 output gap as the deviation of output from its natural level based on the pro- 16  
17 ductivity process.<sup>8</sup> The second deviation involves the exogenous price and wage 17  
18 mark-up shocks, which follow ARMA(1,1) processes in the original model. How- 18  
19 ever, as shown in Slobodyan and Wouters (2012a), mark-up shocks are typically 19  
20 reduced to near white noise processes once learning dynamics are introduced. 20  
21 In these cases, the AR(1) and MA(1) parameters are typically locally unidenti- 21  
22 fied. Therefore we shut off the MA component of these shocks. The third differ- 22  
23 ence pertains to the prior distribution of price stickiness  $\xi_p$ , which we tighten 23  
24 from  $\xi_p \sim \text{Beta}(0.5, 0.2)$  to  $\xi_p \sim \text{Beta}(0.75, 0.05)$ . This follows from our observa- 24  
25 tions that the approximation algorithm may fail to find an equilibrium and thus 25  
26 break down for small values of  $\xi_p$ . Further, for the learning models, we have an 26  
27 additional estimated parameter  $\gamma$ , i.e., the constant gain value. This is assigned 27  
28 a prior of  $\gamma \sim \text{Gamma}(0.035, 0.015)$ , which closely follows the assumption in Slo- 28  
29 bodyan and Wouters (2012). The remainder of the model remains unchanged and 29  
30

31 <sup>7</sup>This is the approach used in Slobodyan and Wouters (2012). 31

32 <sup>8</sup>See Appendix E for further details. 32

1 consists of 13 equations with 7 forward-looking variables, 7 exogenous shocks, 1  
2 and 7 state variables that enter into the model equations with a lag. There are 2  
3 35 estimated parameters including the constant gain for the adaptive learning 3  
4 models. 4

## APPENDIX: REFERENCES

7 Brock, W. A. and A. G. Malliaris (1989). Differential equations, stability and chaos 7  
8 in dynamic economics. *Advanced Textbooks in Economics* (90A16 BROd). [9] 8

9 Elaydi, S. (2005). *An introduction to difference equations, 3rd edition*. New York: 9  
10 Springer. [6] 10

11 Lancaster, P. and M. Tismenetsky (1985). *The theory of matrices: with applica-* 11  
12 *tions*. Elsevier. [3] 12

13 Magnus, J. R. and H. Neudecker (2019). *Matrix differential calculus with applica-* 13  
14 *tions in statistics and econometrics*. John Wiley & Sons. [3] 14

15 Sims, C. (1999). Matlab optimization software. *QM&RBC Codes*. [15] 15  
16

17 Slobodyan, S. and R. Wouters (2012). Learning in an estimated medium-scale 17  
18 DSGE model. *Journal of Economic Dynamics and Control* 36(1), 26–46. [22] 18  
19

20 Smets, F. and R. Wouters (2007). Shocks and frictions in US business cycles: a 20  
21 Bayesian DSGE approach. *American Economic Review* 97(3), 586–606. [22] 21

22 Uhlig, H. F. (1995). A toolkit for analyzing nonlinear dynamic stochastic mod- 22  
23 els easily. *Institute for Empirical Macroeconomics, Federal Reserve Bank of Min-* 23  
24 *neapolis, 55480-0291*. [13, 16] 24  
25

26  
27  
28  
29  
30  
31  
32

26  
27  
28  
29  
30  
31  
32