Online Appendix to "Simple Models and Biased Forecasts"

Pooya Molavi[*]

February 21, 2025

B Omitted Details for the Business-Cycle Economy

This appendix details the problems of different agents in the business-cycle application of Section 5.

Final-good producers

The final good Y_t is produced by competitive firms by combining a continuum of intermediate goods, indexed by i, according to the CES production function

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

where λ_p denotes the elasticity of substitution. Profit maximization and the zero-profit condition imply that the price of the final good is given by the price index

$$P_t = \left[\int_0^1 P_t(i)^{\frac{1}{\lambda_p}} di\right]^{\lambda_p},$$

where $P_t(i)$ denotes the price of intermediate good i. The demand for good i is given by the isoelastic demand schedule

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

[*]Northwestern University, pmolavi@kellogg.northwestern.edu.

Intermediate-goods producers

A monopolist produces intermediate good i according to the production function

$$Y_t(i) = \max\left\{a_t K_t(i)^\alpha \left(\gamma^t L_t(i)\right)^{1-\alpha} - \gamma^t F, 0\right\},$$

where $K_t(i)$ and $L_t(i)$ denote the capital and labor inputs of the monopolist, respectively, F is a fixed cost of production, chosen so that profits are zero along the balanced-growth path, γ denotes the exogenous rate of labor-augmenting technological progress, and a_t is a stationary TFP shock, which follows the AR(1) process $\log a_t = \rho_a \log a_{t-1} + \varepsilon_{at}$ with ε_{at} i.i.d. $\mathcal{N}(0, \sigma_a^2)$.

Intermediate-goods producers are subject to nominal frictions à la Calvo. Each period the price of a randomly-selected fraction ξ_p of intermediate goods grows at rate π, where π denotes the value of gross inflation rate along the balanced-growth path. The remaining intermediate-goods producers choose their prices $P_t(i)$ optimally by maximizing the present-discounted value of future profits,

$$E_{pt}\left[\sum_{s=0}^{\infty} \xi_p^s \beta^s \Lambda_{t+s}\Big(\pi^s P_t(i) Y_{t+s}(i) - W_{t+s} L_{t+s}(i) - r_{t+s} K_{t+s}(i)\Big)\right],$$

subject to the demand curve

$$Y_{t+s}(i) = \left(\frac{\pi^s P_t(i)}{P_{t+s}}\right)^{-\frac{1+\lambda_p}{\lambda_p}} Y_{t+s},$$

where Λ_t is the marginal utility of nominal income, W_t is the nominal wage, r_t is the rental rate of capital, and E_{pt} denotes the time-t forecasts of intermediate-goods producers about the path $\{\Lambda_{t+s}, W_{t+s}, r_{t+s}, P_{t+s}, Y_{t+s}, a_{t+s}\}_{s\geq 1}$ of variables they take as given.

Investment firms

The capital stock of the economy is owned by competitive investment firms. They take the rental rate of capital and the price of the final good as given and maximize the present-discounted value of profits

$$E_{it}\left[\sum_{s=0}^{\infty} \beta^s \Lambda_{t+s}\left(r_{t+s} K_{t+s} - P_{t+s} I_{t+s}\right)\right],$$

subject to the capital accumulation equation

$$K_{t+1} = (1-\delta)K_t + \mu_t \left(I_t - S_k \left(\frac{I_t}{K_t} \right) K_t \right),$$

where I_t is investment, K_t denotes the physical capital, E_{it} denotes the time-t forecasts of investment firms, $S_k(\cdot)$ represents the adjustment cost function, and μ_t is the investment shock, which follows the AR(1) process $\log \mu_t = \rho_\mu \log \mu_{t-1} + \varepsilon_{\mu t}$ with $\varepsilon_{\mu t}$ is i.i.d. $\mathcal{N}(0, \sigma_\mu^2)$. I assume that the adjustment cost satisfies $S_k = S_k' = 0$ and $S_k'' = \varsigma_k > 0$ along the balanced-growth path.[1] I also assume there is no spot market for installed capital.[2]

Employment agencies

There is a continuum of households, indexed by j, each of which is a monopolistic supplier of a specialized type of labor. A competitive employment agency combines specialized labor into a homogeneous labor input using the CES function

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

where λ_w denotes the elasticity of substitution among differentiated types of labor. Profit maximization by employment agencies and the zero-profit condition imply that the price of the homogeneous labor input is given by the wage index

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

and the demand for the labor of type j is given by the isoelastic labor-demand curve

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

[1]Note that the adjustment cost is a neoclassical cost à la Hayashi (1982), not the investment-adjustment cost common in the DSGE literature. The investment-adjustment cost specification leads to an investment Euler equation with a backward-looking term, whereas investment will have no backward-looking term in the current specification.

[2]This assumption is immaterial under rational expectations. However, this may no longer be the case away from rational expectations: When there is no spot market for capital, investment depends on agents' expectations about the infinite future path of returns to capital; when a spot market exists, investment only depends on agents' expectations of the rental rate of capital and its price in the next period.

Households

Households supply labor, consume the final good, and save in a short-term nominal government bond. Their wages are subjective to nominal rigidities à la Calvo. However, as is common in the literature, I assume that a competitive insurance agency fully insures households against fluctuations in their labor income resulting from nominal frictions. Consequently, the equilibrium labor income of each household is equal to $W_t L_t$, the average labor income in the economy.

Each household takes the labor income and the stream of profits from the ownership of firms as given and chooses consumption and saving in government bonds to maximize the utility function

$$E_{ct}\left[\sum_{s=0}^{\infty}\beta^s\left(\log(C_{t+s})-\varphi\frac{L_{t+s}(j)^{1+\nu}}{1+\nu}\right)\right],$$

subject to a no-Ponzi condition and the nominal budget constraint

$$P_tC_t+T_t+B_t\leq R_{t-1}B_{t-1}+W_tL_t+\Pi_t,$$

where C_t is consumption, T_t denotes lump-sum taxes, B_t is the holding of one-period government bonds, R_t is the gross nominal interest rate, Π_t denotes profits from the ownership of firms, ν is the inverse Frisch elasticity of labor supply, and φ is a constant that determines the steady-state working hours. The operator E_{ct} denotes the time-t forecasts of households about the path $\{L_{t+s}, W_{t+s}, P_{t+s}, T_{t+s}, R_{t+s}, \Pi_{t+s}\}_{s\geq 1}$ of aggregate and idiosyncratic observables that enter their decision problem.

Labor unions

Wages are set by a continuum of labor unions, also indexed by j, each representing a household. Each period, a randomly-selected fraction ξ_w of unions cannot freely set the wage of the household they represent. The nominal wages of those households grow at the rate $\gamma\pi$.[3] The remaining fraction of labor unions sets the optimal wage $W_t(j)$ by maximizing

$$E_{wt}\left[\sum_{s=0}^{\infty}\xi_w^s\beta^s\left(-\varphi\frac{L_{t+s}(j)^{1+\nu}}{1+\nu}+\Lambda_{t+s}(\gamma\pi)^sW_t(j)L_{t+s}(j)\right)\right],$$

[3] Since there is technological progress, absent this assumption, there would be no balanced-growth path without wage dispersion. Note that this is different than the assumption of wage indexation common in the DSGE literature: Wages are not indexed to the current inflation rate but to its steady-state value.

subject to the labor demand curve

$$L_{t+s}(j) = \left(\frac{(\gamma\pi)^s W_t(j)}{W_{t+s}} \right)^{-\frac{1+\lambda_w}{\lambda_w}} L_{t+s},$$

where E_{wt} denotes the time-t forecasts of labor unions about the variables they take as given.

The government

The monetary policy sets the nominal interest rate following a Taylor rule

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R} \right)^{\rho_R} \left(\frac{\pi_t}{\pi} \right)^{(1-\rho_R)\phi_\pi} m_t,$$

where $\pi_t \equiv P_t/P_{t-1}$, and R and π are the steady-state gross nominal interest rate and inflation rate, respectively.[4] [5] m_t is a monetary policy shock that follows the AR(1) process $\log m_t = \rho_m \log m_{t-1} + \varepsilon_{mt}$ with ε_{mt} is i.i.d. $\mathcal{N}(0, \sigma_m^2)$.

Government spending G_t is exogenous. In the baseline specification, I assume that government spending grows at the same rate as GDP, that is, $G_t = g\gamma^t$ for some g. The government finances spending by issuing short-term nominal bonds and levying lump-sum taxes on households. The nominal government budget constraint is given by

$$R_{t-1}B_{t-1} + P_t G_t - T_t = B_t,$$

where T_t denotes nominal taxes. Taxes follow a tax rule that ensures that the real value of public debt B_t/P_t grows at rate γ, the deterministic growth rate of the economy.[6]

[4] In the New Keynesian literature, it is often assumed that the monetary authority responds both to changes in the inflation rate and to changes in the output gap. However, the right notion of the output gap is not clear here: It can be defined relative to the flexible price allocation in which agents re-estimate their models, the one in which agents' models are unchanged, or the rational-expectations flexible-price allocation. I bypass the question of how the output gap ought to be defined by instead assuming that the monetary authority only responds to deviations in the inflation rate.

[5] The steady-state gross nominal interest rate π can also be seen as the central bank's inflation target.

[6] Ricardian equivalence does not necessarily hold when agents use simple models. Therefore, both the timing of taxes and the value of the outstanding public debt might affect the response of the economy to shocks. See also Eusepi and Preston (2018), where the authors use an adaptive learning framework to study the effects of the level of public debt on the transmission of monetary policy.

C Temporary Equilibrium of the Business-Cycle Economy

In this appendix, I list the equations that characterize the log-linearized temporary equilibrium of the business-cycle economy studied in Section 5. These conditions characterize the fully flexible model used in Bayesian estimation; the baseline specification can be obtained by setting the values of parameters ξ_w, g/y, b/y, σ_ψ, σ_g, σ_p, and σ_w equal to zero.

The steady-state values are given by $\rho = \frac{\gamma}{\beta} - (1-\delta)$, $w = \left[\frac{1}{1+\lambda_p}\alpha^\alpha(1-\alpha)^{1-\alpha}\frac{1}{\rho^\alpha}\right]^{\frac{1}{1-\alpha}}$, $\frac{k}{L} = \frac{w}{\rho}\frac{\alpha}{1-\alpha}$, $\frac{F}{L} = \left(\frac{k}{L}\right)^\alpha - \rho\frac{k}{L} - w$, $\frac{y}{L} = \left(\frac{k}{L}\right)^\alpha - \frac{F}{L}$, $\frac{i}{L} = (\gamma - (1-\delta))\frac{k}{L}$, $\frac{c}{L} = \frac{y}{L} - \frac{i}{L} - \frac{g}{y}\frac{y}{L}$, $\frac{x}{L} = \frac{y}{L} - \frac{i}{L}$, and $\frac{\tau}{L} = \left(\frac{g}{y} + \frac{1-\beta}{\beta}\frac{b}{y}\right)\frac{y}{L}$.

The log-linear permanent-income equation is given by

$$\hat{c}_t = \hat{\psi}_t - \hat{R}_t + E_{ct}\left[\sum_{s=1}^{\infty}\beta^s\left(\frac{1-\beta}{\beta}\frac{x}{c}\hat{x}_{t+s} - \frac{1-\beta}{\beta}\frac{\tau}{c}\hat{\tau}_{t+s} - \frac{1-\beta}{\beta}\hat{\psi}_{t+s} - \frac{x-\tau}{c}\hat{R}_{t+s} + \frac{1}{\beta}\frac{x-\tau}{c}\hat{\pi}_{t+s}\right)\right]. \tag{C.1}$$

Households' pre-tax income is given by

$$\hat{x}_t = \frac{y}{x}\hat{y}_t - \frac{i}{x}\hat{i}_t. \tag{C.2}$$

The intertemporal preference shock follows the exogenous process

$$\hat{\psi}_t = \rho_\psi\hat{\psi}_{t-1} + \varepsilon_{\psi t}, \qquad \varepsilon_{\psi t} \sim \mathcal{N}(0, \sigma_\psi^2). \tag{C.3}$$

Investment is given by

$$\begin{aligned}\hat{i}_t =& \hat{k}_t + \frac{1}{\varsigma_k}\left(\hat{\mu}_t - \hat{\psi}_t + \hat{c}_t\right) \\ &+ E_{it}\left[\sum_{s=1}^{\infty}\beta^s\left(\frac{1-\beta}{\varsigma_k\beta}\hat{\psi}_{t+s} - \frac{1-\beta}{\varsigma_k\beta}\hat{c}_{t+s} + \frac{1}{\varsigma_k}\left(\frac{1}{\beta} - \frac{1-\delta}{\gamma}\right)\hat{\rho}_{t+s} + \frac{1}{\varsigma_k}\left(1 - \frac{1-\delta}{\gamma}\right)\hat{\mu}_{t+s}\right)\right],\end{aligned} \tag{C.4}$$

where the investment shock follows the AR(1) process

$$\hat{\mu}_t = \rho_\mu\hat{\mu}_{t-1} + \varepsilon_{\mu t}, \qquad \varepsilon_{\mu t} \sim \mathcal{N}(0, \sigma_\mu^2). \tag{C.5}$$

Capital stock evolves according to

$$\hat{k}_t = \frac{1-\delta}{\gamma}\hat{k}_{t-1} + \left(1 - \frac{1-\delta}{\gamma}\right)\left(\hat{i}_{t-1} + \hat{\mu}_{t-1}\right). \tag{C.6}$$

Government spending follows the exogenous process

$$\hat{g}_t = \rho_g \hat{g}_{t-1} + \varepsilon_{gt}, \qquad \varepsilon_{gt} \sim \mathcal{N}(0, \sigma_g^2), \tag{C.7}$$

and GDP is given by

$$\hat{y}_t = \frac{c}{y}\hat{c}_t + \frac{i}{y}\hat{i}_t + \frac{g}{y}\hat{g}_t. \tag{C.8}$$

Inflation is given by

$$\begin{aligned}\hat{\pi}_t =& \hat{\lambda}_{pt} + \kappa\left(\alpha\hat{\rho}_t + (1-\alpha)\hat{w}_t - \hat{a}_t\right) \\ &+ E_{pt}\left[\sum_{s=1}^{\infty} \xi_p^s \beta^s \left(\tfrac{1-\xi_p}{\xi_p}\hat{\pi}_{t+s} + \hat{\lambda}_{p,t+s} + \kappa\left(\alpha\hat{\rho}_{t+s} + (1-\alpha)\hat{w}_{t+s} - \hat{a}_{t+s}\right)\right)\right],\end{aligned} \tag{C.9}$$

where $\kappa \equiv \frac{(1-\xi_p)(1-\xi_p\beta)}{\xi_p}$ is a constant, TFP follows the exogenous process

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{at}, \qquad \varepsilon_{at} \sim \mathcal{N}(0, \sigma_a^2), \tag{C.10}$$

and the price markup shock follows the exogenous process

$$\hat{\lambda}_{pt} = \rho_p \hat{\lambda}_{p,t-1} + \varepsilon_{pt}, \qquad \varepsilon_{pt} \sim \mathcal{N}(0, \sigma_p^2). \tag{C.11}$$

The real wage is given by

$$\begin{aligned}\hat{w}_t =& \frac{1+\beta}{1+\beta-\xi_w\beta}\left(\hat{\lambda}_{wt} + \kappa_w\hat{\ell}_t\right) + \frac{1}{1+\beta-\xi_w\beta}\left(\hat{w}_{t-1} - \hat{\pi}_t\right) \\ &+ \tfrac{1+\beta}{1+\beta-\xi_w\beta} E_{wt}\left[\sum_{s=1}^{\infty} \xi_w^s \beta^s \left(\tfrac{\nu_w\kappa_w}{1-\xi_w\beta}\hat{\pi}_{t+s} + \hat{\lambda}_{w,t+s} + \kappa_w\hat{\ell}_{t+s} + \nu_w\kappa_w\hat{w}_{t+s}\right)\right],\end{aligned} \tag{C.12}$$

where $\kappa_w \equiv \frac{(1-\xi_w)(1-\xi_w\beta)}{\xi_w\nu_w(1+\beta)}$ is a constant,

$$\hat{\ell}_t = \nu\hat{L}_t + \hat{c}_t - \hat{w}_t, \tag{C.13}$$

and the wage markup shock $\hat{\lambda}_{wt}$ follows the exogenous process

$$\hat{\lambda}_{wt} = \rho_w \hat{\lambda}_{w,t-1} + \varepsilon_{wt}, \qquad \varepsilon_{wt} \sim \mathcal{N}(0, \sigma_w^2). \tag{C.14}$$

Hours are given by

$$\hat{L}_t = \frac{1}{1-\alpha}\left(\frac{y}{y+F}\hat{y}_t - \alpha\hat{k}_t - \hat{a}_t + \left(\frac{\rho k}{y+F} - \alpha\right)\hat{\rho}_t\right), \tag{C.15}$$

and the rental rate of capital by

$$\hat{\rho}_t = \hat{w}_t + \hat{L}_t - \hat{k}_t. \tag{C.16}$$

The nominal interest rate follows the interest rate rule

$$\hat{R}_t = \rho_R \hat{R}_{t-1} + (1-\rho_R)\phi_\pi \hat{\pi}_t + \hat{m}_t, \tag{C.17}$$

where the monetary-policy shock follows the exogenous process

$$\hat{m}_t = \rho_m \hat{m}_{t-1} + \varepsilon_{mt}, \qquad \varepsilon_{mt} \sim \mathcal{N}(0, \sigma_m^2). \tag{C.18}$$

Finally, taxes follow the tax rule

$$\hat{\tau}_t = \frac{g}{\tau}\hat{g}_t + \frac{b}{\beta\tau}\left(\hat{R}_{t-1} - \hat{\pi}_t\right). \tag{C.19}$$

D Bayesian Estimation

The likelihood is based on the measurement equation

$$\begin{pmatrix} \Delta Y_t & \Delta C_t & \Delta I_t & L_t & \pi_t & \Delta w_t & R_t \end{pmatrix}' = \begin{pmatrix} \hat{y}_t & \hat{c}_t & \hat{i}_t & \hat{L}_t & \hat{\pi}_t & \hat{w}_t & \hat{R}_t \end{pmatrix}' + \overline{\zeta},$$

where Δ denotes the temporal difference operator, Y denotes real GDP per capita, C denotes real consumption per capita, I denotes real investment per capita, L denotes hours worked per capita, π denotes the inflation rate, w denotes the real wage index, R denotes the nominal interest rate, and $\overline{\zeta}$ is the vector containing the sample mean of the vector on the left side of the above equation. The vector of means $\overline{\zeta}$ is only informative about level variables that are fixed in the estimation step. Therefore, the likelihood can be constructed using the demeaned values of ΔY_t, ΔC_t, ΔI_t, L_t, π_t, Δw_t, and R_t.

The data are from the Federal Reserve Economic Database (FRED). Tables D.1 and D.2 describe the original data and the transformations used in Bayesian estimation.

Table D.3 presents the prior densities and posterior estimates of model parameters.

Figure D.1 plots the impulse-response functions of endogenous variables to various shocks given the estimated model parameters.

Table D.1. Description of data.

Data	Mnemonic	Frequency	Transform
Real gross domestic product per capita	A939RX0Q048SBEA	Q	—
Share of GDP: personal consumption expenditures: nondurable goods	DNDGREI1Q156NBEA	Q	—
Share of GDP: personal consumption expenditures: services	DSERREI1Q156NBEA	Q	—
Share of GDP: personal consumption expenditures: durable goods	DDURREI1Q156NBEA	Q	—
Share of GDP: gross private domestic investment	A006REI1Q156NBEA	Q	—
Nonfarm business sector: average weekly hours	PRS85006023	Q	—
Civilian employment level	CE16OV	M	EoP
Civilian non-institutional population	CNP16OV	M	EoP
Gross domestic product: implicit price deflator	GDPDEF	Q	—
Non-farm business sector: real hourly compensation for all workers	COMPRNFB	Q	—
Effective federal funds rate	FEDFUNDS	M	Ave

Note: Q: quarterly, M: monthly, EoP: end of period, Ave: quarterly average.

Table D.2. Variables used in Bayesian estimation.

Variable	Definition
Real GDP per capita	$Y = 100 \times \log(\text{A939RX0Q048SBEA})$
Real consumption per capita	$C = 100 \times \log((\text{DNDGREI1Q156NBEA} + \text{DSERREI1Q156NBEA}) \times \text{A939RX0Q048SBEA})$
Real investment per capita	$I = 100 \times \log((\text{DDURREI1Q156NBEA} + \text{A006REI1Q156NBEA}) \times \text{A939RX0Q048SBEA})$
Hours worked	$L = 100 \times \log(\text{PRS85006023} \times \text{CE16OV}/\text{CNP16OV})$
Inflation rate	$\pi = 100 \times \log(\text{GDPDEF}/\text{GDPDEF}(-1))$
Real wage	$w = 100 \times \log(\text{COMPRNFB})$
Interest rate	$R = \text{FEDFUNDS}/4$

Table D.3. Prior densities and posterior estimates.

Coeff.	Description	Prior distribution Distr.	Mean	Std. Dev.	Posterior mode 1d	RE
ν	Inverse Frisch elasticity	G	2.00	0.50	1.91 [1.48, 2.46]	0.56 [0.44, 0.71]
α	Capital share	N	0.30	0.05	0.28 [0.27, 0.29]	0.28 [0.27, 0.29]
λ_p	Steady-state price markup	B	0.15	0.05	0.50 [0.45, 0.55]	0.41 [0.36, 0.47]
λ_w	Steady-state wage markup	B	0.15	0.05	0.11 [0.07, 0.16]	0.21 [0.16, 0.26]
ξ_p	Calvo, prices	B	0.50	0.10	0.77 [0.74, 0.79]	0.65 [0.61, 0.68]
ξ_w	Calvo, wages	B	0.50	0.10	0.77 [0.72, 0.81]	0.15 [0.11, 0.19]
ρ_R	Taylor-rule smoothing	B	0.60	0.20	0.87 [0.83, 0.90]	0.54 [0.48, 0.61]
ϕ_π	Taylor rule, inflation	N	1.50	0.20	1.07 [1.03, 1.10]	1.67 [1.57, 1.77]
ς_k	Capital-adjustment cost	G	4.00	1.00	1.70 [1.47, 1.98]	1.06 [0.87, 1.29]
ρ_a	Technology shock, AR	B	0.60	0.15	0.93 [0.90, 0.95]	0.91 [0.90, 0.93]
ρ_m	Monetary-policy shock, AR	B	0.60	0.15	0.30 [0.24, 0.37]	0.26 [0.20, 0.33]
ρ_g	Government-spending shock, AR	B	0.60	0.15	0.97 [0.95, 0.98]	0.97 [0.96, 0.98]
ρ_p	Price-markup shock, AR	B	0.60	0.15	0.91 [0.87, 0.93]	0.96 [0.95, 0.97]
ρ_w	Wage-markup shock, AR	B	0.60	0.15	0.97 [0.96, 0.98]	0.97 [0.96, 0.98]
ρ_ψ	Preference shock, AR	B	0.60	0.15	0.96 [0.94, 0.97]	0.92 [0.90, 0.94]
ρ_μ	Investment shock, AR	B	0.60	0.15	0.87 [0.85, 0.89]	0.90 [0.88, 0.92]
θ_p	Price-markup shock, MA	B	0.50	0.20	0.41 [0.35, 0.48]	0.22 [0.14, 0.34]
θ_w	Wage-markup shock, MA	B	0.50	0.20	0.64 [0.57, 0.70]	0.32 [0.24, 0.41]
σ_a	Technology shock, SD	IG	0.50	1.00	0.54 [0.51, 0.57]	0.56 [0.53, 0.59]
σ_m	Monetary-policy shock, SD	IG	0.50	1.00	0.22 [0.21, 0.23]	0.31 [0.29, 0.35]
σ_g	Government-spending shock, SD	IG	0.50	1.00	1.53 [1.46, 1.61]	1.52 [1.45, 1.60]
σ_p	Price-markup shock, SD	IG	0.50	1.00	0.26 [0.24, 0.27]	0.23 [0.20, 0.27]
σ_w	Wage-markup shock, SD	IG	0.50	1.00	0.42 [0.39, 0.45]	0.91 [0.72, 1.15]
σ_ψ	Preference shock, SD	IG	0.50	1.00	0.56 [0.53, 0.59]	1.45 [1.21, 1.72]
σ_μ	Investment shock, SD	IG	0.50	1.00	5.74 [4.92, 6.70]	4.23 [3.54, 5.05]

Notes: B: beta, G: gamma, IG: inverse gamma, N: normal. 68 percent HPDIs computed using Laplace's approximation in brackets.

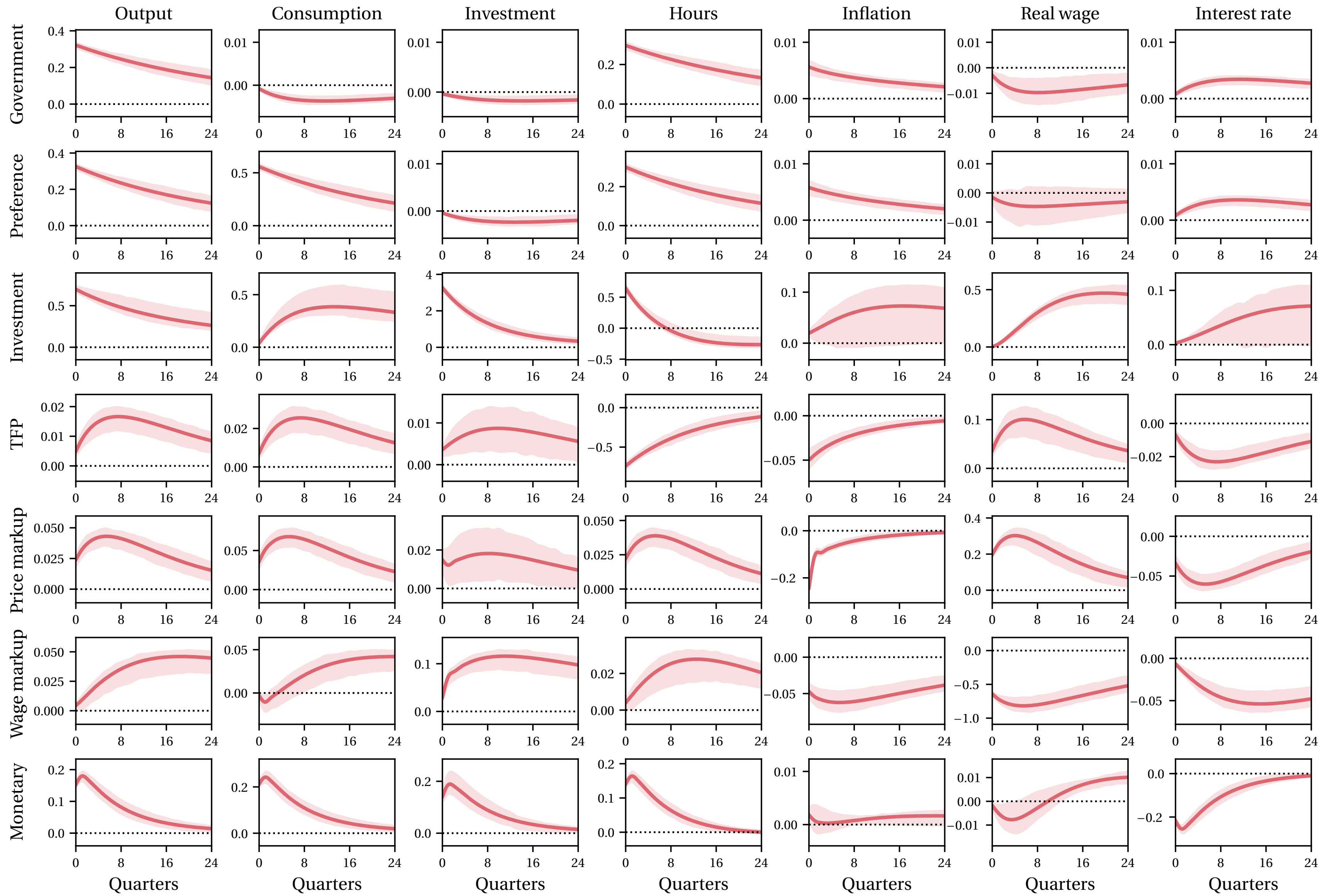

Figure D.1. Impulse-response functions.

Notes: Responses of endogenous variables (columns) to one-standard-deviation shocks (rows). One-dimensional simple models. The solid line represents the posterior mode. Shaded areas are 68 percent HPDIs computed using Laplace's approximation. Output, consumption, investment, hours, and real wage measured in percents; inflation and nominal interest rates measured in percentage points. Shocks are normalized to increase output on impact at the posterior mode.

E Auxiliary Lemmas

Lemma E.1. *For any purely non-deterministic, stationary ergodic, and non-degenerate process with autocorrelation matrices $\{C_l\}_l$, the spectral radii of autocorrelation matrices satisfy $\rho(C_l) \leq 1$ for any l with the inequality strict for $l = 1$.*

Proof. Let λ_l denote an eigenvalue of C_l largest in magnitude and let u_l denote the corresponding eigenvector normalized such that $u_l' u_l = 1$. Define the process $\omega_t^{(l)} \equiv u_l' \Gamma_0^{\frac{-1}{2}} y_t \in \mathbb{R}$. Since y_t is a purely non-deterministic, stationary ergodic, and non-degenerate process, so is $\omega_t^{(l)}$ for any l. I first show that λ_l is the autocorrelation of process $\omega_t^{(l)}$ at lag l. Note that

$$\mathbb{E}[\omega_t^{(l)} \omega_{t-l}^{(l)}] = u_l' \Gamma_0^{\frac{-1}{2}} \mathbb{E}[y_t y_{t-l}'] \Gamma_0^{\frac{-1}{2}} u_l = u_l' \Gamma_0^{\frac{-1}{2}} \Gamma_l \Gamma_0^{\frac{-1}{2}} u_l = u_l' \Gamma_0^{\frac{-1}{2}} \left(\frac{\Gamma_l + \Gamma_l'}{2} \right) \Gamma_0^{\frac{-1}{2}} u_l = u_l' C_l u_l = \lambda_l,$$

where the first, second, and fourth equalities are by definition, the last equality uses the fact that λ_l is the eigenvalue of C_l with eigenvector u_l, normalized such that $u_l' u_l = 1$, and the third equality uses the fact that, since $u_l' \Gamma_0^{\frac{-1}{2}} \Gamma_l \Gamma_0^{\frac{-1}{2}} u_l$ is a scalar and $\Gamma_0^{\frac{-1}{2}}$ is a symmetric matrix,

$$u_l' \Gamma_0^{\frac{-1}{2}} \Gamma_l' \Gamma_0^{\frac{-1}{2}} u_l = \left(u_l' \Gamma_0^{\frac{-1}{2}} \Gamma_l \Gamma_0^{\frac{-1}{2}} u_l \right)' = u_l' \Gamma_0^{\frac{-1}{2}} \Gamma_l \Gamma_0^{\frac{-1}{2}} u_l.$$

On the other hand,

$$\mathbb{E}[\omega_t^{(l)} \omega_t^{(l)}] = u_l' \Gamma_0^{\frac{-1}{2}} \mathbb{E}[y_t y_t'] \Gamma_0^{\frac{-1}{2}} u_l = u_l' \Gamma_0^{\frac{-1}{2}} \Gamma_0 \Gamma_0^{\frac{-1}{2}} u_l = u_l' u_l = 1.$$

Therefore, since $\omega_t^{(l)}$ is purely non-deterministic, stationary ergodic, and non-degenerate,

$$\rho(C_l) = |\lambda_l| = \frac{\mathbb{E}[\omega_t^{(l)} \omega_{t-l}^{(l)}]}{\mathbb{E}[\omega_t^{(l)} \omega_t^{(l)}]} \leq 1.$$

Next, toward a contradiction suppose that $\rho(C_1) = 1$. Then $\omega_t^{(1)}$ is perfectly correlated with $\omega_{t-1}^{(1)}$, and so, with $\omega_{t-l}^{(1)}$ for every l, contradicting the assumption that $\omega_t^{(1)}$ is purely non-deterministic, stationary ergodic, and non-degenerate. □

Lemma E.2. *If $\mathbb{P}$ is purely non-deterministic and stationary ergodic, then so is P^θ for any pseudo-true one-state model θ.*

Proof. Define

$$C(a, \eta) \equiv \sum_{\tau=1}^{\infty} a^\tau \eta^{\tau-1} C_\tau. \tag{E.1}$$

Then

$$\lambda_{\max}(\Omega(a,\eta)) = -\frac{a^2(1-\eta)^2}{1-a^2\eta^2} + \frac{2(1-\eta)(1-a^2\eta)}{1-a^2\eta^2}\lambda_{\max}(C(a,\eta)), \tag{E.2}$$

where $\lambda_{\max}(C(a,\eta))$ denotes the largest eigenvalue of $C(a,\eta)$. To simplify the exposition, I prove the result under the assumption that the largest eigenvalue of $C(a,\eta)$ is simple at the point (a^*,η^*) that maximizes $\lambda_{\max}(\Omega(a,\eta))$.[7] The partial derivatives of $\lambda_{\max}(\Omega(a,\eta))$ with respect to a and η are given by

$$\begin{aligned}
\frac{\partial\lambda_{\max}(\Omega(a,\eta))}{\partial a} &= \frac{-2a(1-\eta)^2}{\left(1-a^2\eta^2\right)^2} - \frac{4a\eta(1-\eta)^2}{\left(1-a^2\eta^2\right)^2}\lambda_{\max}(C) \\
&\quad + \frac{2(1-\eta)(1-a^2\eta)}{1-a^2\eta^2}u'_{\max}(C)\frac{\partial C}{\partial a}u_{\max}(C), && \text{(E.3)}\\
\frac{\partial\lambda_{\max}(\Omega(a,\eta))}{\partial \eta} &= \frac{2a^2(1-\eta)(1-a^2\eta)}{\left(1-a^2\eta^2\right)^2} - \frac{2\left(1+a^4\eta^2+a^2(1-4\eta+\eta^2)\right)}{\left(1-a^2\eta^2\right)^2}\lambda_{\max}(C) \\
&\quad + \frac{2(1-\eta)(1-a^2\eta)}{1-a^2\eta^2}u'_{\max}(C)\frac{\partial C}{\partial \eta}u_{\max}(C), && \text{(E.4)}
\end{aligned}$$

where $u_{\max}(C)$ denotes the eigenvector of C with eigenvalue $\lambda_{\max}(C)$, normalized such that $u'_{\max}(C)u_{\max}(C) = 1$, and

$$\begin{aligned}
\frac{\partial C}{\partial a} &= \sum_{\tau=1}^{\infty}\tau a^{\tau-1}\eta^{\tau-1}C_\tau, \\
\frac{\partial C}{\partial \eta} &= \sum_{\tau=1}^{\infty}(\tau-1)a^{\tau}\eta^{\tau-2}C_\tau.
\end{aligned}$$

Note that

$$\eta u'_{\max}(C)\frac{\partial C}{\partial \eta}u_{\max}(C) + \lambda_{\max}(C) = a u'_{\max}(C)\frac{\partial C}{\partial a}u_{\max}(C) \tag{E.5}$$

for any a and η.

Let a^* and η^* be scalars in the $[-1,1]$ and $[0,1]$ intervals, respectively, that maximize $\lambda_{\max}(\Omega(a,\eta))$. I separately consider the cases $\eta^*=1$ and $\eta^*<1$. If $\eta^*=1$, then $B=0$ in the representation in the proof of Theorem 2, the pseudo-true one-state model is i.i.d., and $A=a^*$ can be chosen arbitrarily to satisfy $|a^*|<1$.[8]

In the rest of the proof, I assume that $\eta^*<1$ and show that this implies $a^*\neq 1$—by a similar argument $a^*\neq -1$. Toward a contradiction, suppose $a^*=1$. Setting $a=1$ in the

[7] The argument can easily be adapted to the case where the largest eigenvalue of $C(a^*,\eta^*)$ is not necessarily simple by replacing the gradient of $\lambda_{\max}(C(a,\eta))$ with its subdifferential and replacing the usual first-order optimality condition with the condition that the zero vector belongs to the subdifferential.

[8] The pseudo-true one-state model then has a zero-state minimal representation.

partial derivatives of $\lambda_{\max}(\Omega(a,\eta))$, I get

$$\left.\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial a}\right|_{a=1} = \frac{2(1-\eta)^2}{\left(1-\eta^2\right)^2}\left[-1-2\eta\lambda_{\max}(C)+(1-\eta^2)u'_{\max}(C)\frac{\partial C}{\partial a}u_{\max}(C)\right],$$

$$\left.\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial \eta}\right|_{a=1} = \frac{2(1-\eta)^2}{\left(1-\eta^2\right)^2}\left[1-2\lambda_{\max}(C)+(1-\eta^2)u'_{\max}(C)\frac{\partial C}{\partial \eta}u_{\max}(C)\right],$$

where $C = C(1,\eta)$ and its partial derivatives are computed at $a = 1$. Multiplying the second equation above by η and subtracting from it the first equation, I get

$$\begin{aligned}
&\eta\left.\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial \eta}\right|_{a=1} - \left.\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial a}\right|_{a=1} \\
&= \frac{2(1-\eta)^2}{\left(1-\eta^2\right)^2}\left[1+\eta+(1-\eta^2)\left(\eta u'_{\max}(C)\frac{\partial C}{\partial \eta}u_{\max}(C) - u'_{\max}(C)\frac{\partial C}{\partial a}u_{\max}(C)\right)\right] \\
&= \frac{2(1-\eta)^2}{\left(1-\eta^2\right)^2}\left[1+\eta-(1-\eta^2)\lambda_{\max}(C)\right],
\end{aligned}$$

where in the second equality I am using identity (E.5). Therefore,

$$\left.\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial a}\right|_{a=1} = \eta\left.\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial \eta}\right|_{a=1} - \frac{2(1-\eta)^2}{\left(1-\eta^2\right)^2}\left(1+\eta-(1-\eta^2)\lambda_{\max}(C(1,\eta))\right).$$

Note that

$$\lambda_{\max}(C(1,\eta)) \le \sum_{\tau=1}^{\infty}\eta^{\tau-1}\lambda_{\max}(C_\tau) < \sum_{\tau=1}^{\infty}\eta^{\tau-1} = \frac{1}{1-\eta},$$

where the second inequality is by Lemma E.1. Therefore,

$$-\frac{2(1-\eta)^2}{\left(1-\eta^2\right)^2}\left(1+\eta-(1-\eta^2)\lambda_{\max}(C(1,\eta))\right) < \frac{2(1-\eta)^2}{\left(1-\eta^2\right)^2}(1+\eta-1-\eta) = 0.$$

On the other hand, by the optimality of $a^* = 1$ and $\eta^* < 1$,

$$\left.\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial \eta}\right|_{a^*=1,\eta=\eta^*} \le 0.$$

Thus,

$$\left.\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial a}\right|_{a^*=1,\eta=\eta^*} < 0,$$

a contradiction to the assumption of optimality of $a^* = 1$ and $\eta^* < 1$. This proves that

$a^* < 1$ and establishes that the one-state model with $a = a^*$ and $\eta = \eta^*$ is purely non-deterministic and stationary ergodic. □

Lemma E.3. *Any Markovian model θ has a representation as in Lemma 2 for which $D'D = I$.*

Proof. Fix a Markovian model θ, and let M, D, and N be as in Lemma 2. By (16), the s-step-ahead forecast under model θ is given by

$$E_t^\theta [y_{t+s}] = N'^{-1} D M^{s-1} \sum_{\tau=0}^{\infty} \left(M\left(I - D'D\right)\right)^\tau M D' N' y_{t-\tau}.$$

Since θ is Markovian and N is invertible, $D\left(M\left(I - D'D\right)\right)^\tau MD' = \mathbf{0}$ for all $\tau \geq 1$. As the first step of the proof, I use this identity and an inductive argument to show that $DM^sD' = (DMD')^s$ for all $s \geq 2$. The following equation establishes the induction base:

$$\mathbf{0} = D\left(M\left(I - D'D\right)\right) MD' = DM^2D' - DMD'DMD'.$$

As the induction hypothesis, suppose $DM^sD' = (DMD')^s$ for some $s \geq 2$. Note that

$$\begin{aligned} D\left(M\left(I - D'D\right)\right)^s MD' &= D\left(M\left(I - D'D\right)\right)^{s-1} M(I - D'D)MD' \\ &= D\left(M\left(I - D'D\right)\right)^{s-1} M^2 D', \end{aligned}$$

where the last equality follows the fact that $D\left(M\left(I - D'D\right)\right)^{s-1} MD' = \mathbf{0}$ for any $s \geq 2$. By a similar argument,

$$D\left(M\left(I - D'D\right)\right)^s MD' = D\left(M\left(I - D'D\right)\right)^{s-2} M^3 D' = \cdots = D\left(M\left(I - D'D\right)\right) M^s D'.$$

Therefore, by the induction hypothesis,

$$D\left(M\left(I - D'D\right)\right)^s MD' = DM^{s+1}D' - DMD'DM^sD' = DM^{s+1}D' - (DMD')^{s+1}.$$

The assumption that $D(M\left(I - D'D\right))^s MD' = \mathbf{0}$ then proves the induction step.

I next find a model $\tilde{\theta}$, represented by matrices $\tilde{M}$, $\tilde{D}$, and $\tilde{N}$, that is observationally equivalent to θ and for which $\tilde{D}'\tilde{D} = I$. Since $D \in \mathbb{R}^{n \times d}$ is a rectangular diagonal matrix and $d \leq n$, $DMD' = \begin{pmatrix} M_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$ for some $d \times d$ matrix M_1. Let $\tilde{M} = M_1$, $\tilde{D} \in \mathbb{R}^{n \times d}$ be the rectangular diagonal matrix with its diagonal elements equal to one, and $\tilde{N} = N$. Then

$\tilde{D}'\tilde{D} = I$. Furthermore,

$$DM^sD' = (DMD')^s = \begin{pmatrix} M_1^s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \tilde{D}\tilde{M}^s\tilde{D}'.$$

By equation (16), the forecasts are identical under the two models:

$$E_t^\theta[y_{t+s}] = N'^{-1}DM^sD'N'y_t = \tilde{N}'^{-1}\tilde{D}\tilde{M}^s\tilde{D}'\tilde{N}'y_t = E_t^{\tilde{\theta}}[y_{t+s}].$$

By equation (17), the unconditional variance of the observable is also identical under the two models:

$$\begin{aligned} \text{Var}^\theta(y) &= N'^{-1}\left(I + \sum_{\tau=1}^{\infty} DM^\tau D'DM'^\tau D'\right)N^{-1} \\ &= \tilde{N}'^{-1}\left(I + \sum_{\tau=1}^{\infty} \tilde{D}\tilde{M}^\tau\tilde{D}'\tilde{D}\tilde{M}'^\tau\tilde{D}'\right)\tilde{N}^{-1} = \text{Var}^{\tilde{\theta}}(y). \end{aligned}$$

On the other hand,

$$E^\theta[y_{t+s}y_t'] = E^\theta[E_t^\theta[y_{t+s}]y_t'] = N'^{-1}DM^sD'N'E^\theta[y_ty_t'] = N'^{-1}DM^sD'N'\text{Var}^\theta(y),$$

and similarly for $E^{\tilde{\theta}}[y_{t+s}y_t']$. Therefore, $E^\theta[y_{t+s}y_t'] = E^{\tilde{\theta}}[y_{t+s}y_t']$ for all s; that is, P^θ and $P^{\tilde{\theta}}$ also have identical autocovariance matrices at all lags. This conclusion, together with the fact that P^θ and $P^{\tilde{\theta}}$ are both zero-mean Gaussian distributions, implies that they are observationally equivalent. □

F Omitted Proofs

Proof of Claim 1. The first-order optimality condition with respect to S is given by

$$S'S - e_1e_1'S'\Omega S = I. \tag{F.1}$$

Multiplying the transpose of the above equation from right by e_1 and from left by S'^{-1}, I get

$$Se_1 - \Omega Se_1 = S'^{-1}e_1. \tag{F.2}$$

On the other hand, multiplying equation (F.1) from left by S and from right by S^{-1}, I get

$$SS' = I + Se_1e_1'S'\Omega. \tag{F.3}$$

By the Sherman–Morrison formula,

$$S'^{-1}S^{-1} = I - \frac{Se_1e_1'S'\Omega}{1 + e_1'S'\Omega Se_1}.$$

Multiplying the above equation from right by Se_1, I get

$$S'^{-1}e_1 = \frac{1}{1 + e_1'S'\Omega Se_1}Se_1. \tag{F.4}$$

Substituting for $S'^{-1}e_1$ from the above equation in (F.2) and rearranging the terms, I get

$$\Omega Se_1 = \frac{e_1'S'\Omega Se_1}{1 + e_1'S'\Omega Se_1}Se_1. \tag{F.5}$$

That is, Se_1 is an eigenvector of Ω. Let λ denote the corresponding eigenvalue and let $u = Se_1/\sqrt{e_1'S'Se_1}$. Then equation (F.5) implies

$$\lambda = \frac{\lambda e_1'S'Se_1}{1 + \lambda e_1'S'Se_1}.$$

I separately consider the cases $\lambda \neq 0$ and $\lambda = 0$. If $\lambda \neq 0$, then $e_1'S'Se_1 = 1/(1-\lambda)$ and $Se_1 = u/\sqrt{1-\lambda}$. Equation (F.4) then implies that $S'^{-1}e_1 = \sqrt{1-\lambda}u$, and equation (F.3) implies that $SS' = I + \frac{\lambda}{1-\lambda}uu'$. If $\lambda = 0$, then equation (F.2) implies that $Se_1 = S'^{-1}e_1$, and so, Se_1 and $S'^{-1}e_1$ are both multiples of u. Furthermore, $e_1'S^{-1}Se_1 = e_1'e_1 = 1$. Therefore, $Se_1 = S'^{-1}e_1 = u$. On the other hand, equation (F.3) implies that $SS' = I$. This completes the proof of the claim. □

Proof of Theorem 3. Setting $M = a$, $D = \sqrt{1-\eta}e_1$, and $N = \Gamma_0^{\frac{-1}{2}}S$ in equation (17), I get

$$\text{Var}^{\theta}(y) = \Gamma_0^{\frac{1}{2}}\left[I + \frac{1}{1-a^2}\left[a^2(1-\eta)^2 - \left(1 - 2a^2\eta + a^2\eta^2\right)\lambda\right]uu'\right]\Gamma_0^{\frac{1}{2}},$$

where a, η, $\lambda = \lambda_{\max}(\Omega(a,\eta))$, and u are as in Theorem 2. Substituting for $\lambda_{\max}(\Omega(a,\eta))$

from equation (E.2) in the above equation, I get

$$\operatorname{Var}^{\theta}(y_t) = \Gamma_0^{\frac{1}{2}} \left[I + \frac{2(1-\eta)(1-a^2\eta)}{(1-a^2)(1-a^2\eta^2)} \left(a^2(1-\eta) - (1-2a^2\eta + a^2\eta^2)\lambda_{\max}(C) \right) uu' \right] \Gamma_0^{\frac{1}{2}}. \tag{F.6}$$

Let a^* and η^* be scalars in the $[-1,1]$ and $[0,1]$ intervals, respectively, that maximize $\lambda_{\max}(\Omega(a,\eta))$. I separately consider the cases $\eta^* = 1$ and $\eta^* < 1$. If $\eta^* = 1$, then the right-hand side of equation (F.6) is equal to Γ_0.

Next suppose $\eta^* < 1$. By the argument in the proof of Lemma E.2, the first-order optimality condition with respect to a must hold with equality at $a = a^*$ and $\eta = \eta^* < 1$. Setting $\partial\lambda_{\max}(\Omega(a,\eta))/\partial a = 0$ in (E.3) and multiplying both sides of the equation by a^*, I get, using (E.5),

$$\begin{aligned} &\frac{2a^{*2}(1-\eta^*)^2}{\left(1-a^{*2}\eta^{*2}\right)^2} + \frac{4a^{*2}\eta^*(1-\eta^*)^2}{\left(1-a^{*2}\eta^{*2}\right)^2}\lambda_{\max}(C) \\ &\quad = \frac{2(1-\eta^*)(1-a^{*2}\eta^*)}{1-a^{*2}\eta^{*2}}\lambda_{\max}(C) + \frac{2(1-\eta^*)(1-a^{*2}\eta^*)}{1-a^{*2}\eta^{*2}}\eta^* u'_{\max}(C)\frac{\partial C}{\partial \eta}u_{\max}(C). \end{aligned} \tag{F.7}$$

Setting $\eta^* = 0$ in the above equation, I get $a^{*2} = \lambda_{\max}(C)$. Setting $a^{*2} = \lambda_{\max}(C)$ in equation (F.6) then establishes that $\operatorname{Var}^{\theta}(y_t) = \Gamma_0$ in the case where $\eta^* = 0$.

Finally, I consider the case where $\eta^* \in (0,1)$. Then additionally the first-order optimality condition with respect to η must hold with equality. Setting $\partial\lambda_{\max}(\Omega(a,\eta))/\partial\eta = 0$ in equation (E.4), multiplying it by η^*, solving for $\eta^* u'_{\max}(C)\frac{\partial C}{\partial\eta}u_{\max}(C)$, and substituting in equation (F.7), I get

$$\begin{aligned} &\frac{2a^{*2}(1-\eta^*)^2}{\left(1-a^{*2}\eta^{*2}\right)^2} + \frac{4a^{*2}\eta^*(1-\eta^*)^2}{\left(1-a^{*2}\eta^{*2}\right)^2}\lambda_{\max}(C) \\ &\quad = \frac{2(1-\eta^*)(1-a^{*2}\eta^*)}{1-a^{*2}\eta^{*2}}\lambda_{\max}(C) - \frac{2a^{*2}\eta^*(1-\eta^*)(1-a^{*2}\eta^*)}{\left(1-a^{*2}\eta^{*2}\right)^2} \\ &\qquad + \frac{2\eta^*\left(1 + a^{*4}\eta^{*2} + a^{*2}(1-4\eta^* + \eta^{*2})\right)}{\left(1-a^{*2}\eta^{*2}\right)^2}\lambda_{\max}(C). \end{aligned}$$

Simplifying the above expression leads to

$$a^{*2}(1-\eta^*) = \left(1 - 2a^{*2}\eta^* + a^{*2}\eta^{*2}\right)\lambda_{\max}(C).$$

Combining the above identity with equation (F.6) implies that $\operatorname{Var}^{\theta}(y_t) = \Gamma_0$ and finishes the proof of the theorem. □

Proof of Proposition 2. By equation (12),

$$\text{Var}_t^\theta(y_{t+1}) = B'\hat{\Sigma}_z B + R,$$

where $\hat{\Sigma}_z$ solves the algebraic Riccati equation (13). The equation can be written as

$$\hat{\Sigma}_z = A\hat{\Sigma}_z^{\frac{1}{2}}\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right)\hat{\Sigma}_z^{\frac{1}{2}}A' + Q. \tag{F.8}$$

Since R is a positive semidefinite matrix, so is $I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}$. Therefore, $\hat{\Sigma}_z \geq Q$, and so, $\text{Var}_t^\theta(y_{t+1}) \geq B'QB + R$. On the other hand,

$$\text{Var}^\theta(y_{t+1}|z_t) = B'\text{Var}^\theta(z_{t+1}|z_t)B + R = B'QB + R,$$

where I am using the assumption that w_t is i.i.d. $\mathcal{N}(0, Q)$, v_t is i.i.d. $\mathcal{N}(0, R)$, and w_t and v_t are independent. This proves the first part of the proposition.

To prove the second part, first assume that $\text{Var}_t^\theta(y_{t+1}) = B'QB + R$. Together with equation (F.8), this implies that

$$B'A\hat{\Sigma}_z^{\frac{1}{2}}\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right)\hat{\Sigma}_z^{\frac{1}{2}}A'B = \mathbf{0}.$$

Since $\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right)$ is a symmetric positive semidefinite matrix, the above equation implies that

$$B'A\hat{\Sigma}_z^{\frac{1}{2}}\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right) = \mathbf{0}.$$

On the other hand, by equation (14), the one-step-ahead forecast under model θ is given by

$$E_t^\theta[y_{t+1}] = B'\sum_{\tau=0}^{\infty}(A - KB')^\tau K y_{t-\tau}.$$

Substituting for K from equation (11), I get

$$\begin{aligned} B'(A - KB') &= B'\left(A - A\hat{\Sigma}_z B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\right) \\ &= B'A\hat{\Sigma}_z^{\frac{1}{2}}\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right)\hat{\Sigma}_z^{\frac{-1}{2}} = \mathbf{0}. \end{aligned}$$

Therefore,

$$E_t^\theta[y_{t+1}] = B' \sum_{\tau=0}^{\infty} (A - KB')^\tau K y_{t-\tau} = B'Ky_t.$$

On the other hand, $\mathrm{Var}_t^\theta(y_{t+1}) = B'\hat{\Sigma}_z B + R$. Under model θ, the mean and variance of y_{t+1} conditional on $\{y_\tau\}_{\tau \le t}$ are both independent of $\{y_\tau\}_{\tau<t}$. Furthermore, P^θ is Gaussian. Therefore, it is Markovian.

Next, suppose P^θ is Markovian. Then by Lemma E.3, model θ has a representation as in Lemma 2 for which $D'D = I$. By equation (18), then

$$\hat{\Sigma}_z^{\frac{1}{2}} B \left(B'\hat{\Sigma}_z B + R\right)^{-1} B'\hat{\Sigma}_z^{\frac{1}{2}} = VD'DV' = I,$$

where the second equality follows the facts that $D'D = I$ and V is orthogonal. Substituting in equation (F.8), I get $\hat{\Sigma}_z = Q$. Therefore, $\mathrm{Var}_t^\theta(y_{t+1}) = B'\hat{\Sigma}_z B + R = B'QB + R$. □

Proof of Theorem 5. By Lemma 2, the agent's model can be represented in terms of matrices M, D, and N. Since the agent is restricted to the set of Markovian models, by Lemma E.3, I can set $D = (I\ \mathbf{0})'$. Let $S \equiv \Gamma_0^{\frac{1}{2}} N$ and $\Gamma \equiv \Gamma_0^{\frac{-1}{2}} \Gamma_1 \Gamma_0^{\frac{-1}{2}}$. The expression for the KLDR in (15) then simplifies to

$$-\frac{1}{2} \log\det(SS') + \frac{1}{2}\mathrm{tr}(S'S) - \mathrm{tr}(MD'S'\Gamma SD) + \frac{1}{2}\mathrm{tr}(MD'S'SDM'),$$

plus a constant that does not depend on (M, S, D). Write $S = (S_1\ S_2)$, where $S_1 \in \mathbb{R}^{n\times d}$ and $S_2 \in \mathbb{R}^{n\times(n-d)}$. The above expression can then be written as

$$-\frac{1}{2}\log\det\left(S_1S_1' + S_2S_2'\right) + \frac{1}{2}\mathrm{tr}\left(S_1'S_1\right) + \frac{1}{2}\mathrm{tr}\left(S_2'S_2\right) - \mathrm{tr}\left(MS_1'\Gamma S_1\right) + \frac{1}{2}\mathrm{tr}\left(MS_1'S_1M'\right).$$

I next optimize the above expression with respect to M, S_1, and S_2. The first-order optimality condition with respect to S_2 is given by $\left(S_1S_1' + S_2S_2'\right)^{-1} S_2 = S_2$, which can be rewritten as $S_1S_1'S_2 + S_2(S_2'S_2 - I) = \mathbf{0}$. Let b_0 be an arbitrary vector in $\mathbb{R}^{n-d}$, $b_1 \equiv S_1'S_2b_0 \in \mathbb{R}^d$, and $b_2 \equiv (S_2'S_2 - I)b_0 \in \mathbb{R}^{n-d}$. The above equation then implies that

$$0 = \left(S_1S_1'S_2 + S_2(S_2'S_2 - I)\right) b_0 = S_1b_1 + S_2b_2 = Sb,$$

where $b \equiv \begin{pmatrix} b_1' & b_2' \end{pmatrix}' \in \mathbb{R}^n$. Since S is an invertible matrix, it must be that $b = 0$. Therefore,

$b_1 = 0$ and $b_2 = 0$. Since b_0 was arbitrary, $S_2'S_2 = I$ and $S_1'S_2 = \mathbf{0}$. On the other hand,

$$\log\det\left(S_1S_1' + S_2S_2'\right) = \log\det(SS') = \log\det(S'S) = \log\det\begin{pmatrix} S_1'S_1 & S_1'S_2 \\ S_2'S_1 & S_2'S_2 \end{pmatrix}.$$

Therefore,

$$\log\det\left(S_1S_1' + S_2S_2'\right) = \log\det\begin{pmatrix} S_1'S_1 & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix} = \log\det(S_1'S_1).$$

The KLDR can thus be written only as a function of M and S_1 as

$$-\frac{1}{2}\log\det\left(S_1'S_1\right) + \frac{1}{2}\operatorname{tr}\left(S_1'S_1\right) - \operatorname{tr}\left(MS_1'\Gamma S_1\right) + \frac{1}{2}\operatorname{tr}\left(MS_1'S_1M'\right), \tag{F.9}$$

plus a constant. The first-order optimality conditions with respect to M and S_1 are then given by

$$-S_1'\Gamma'S_1 + MS_1'S_1 = 0, \tag{F.10}$$

$$-S_1^{\dagger\prime} + S_1 - \Gamma S_1M - \Gamma'S_1M' + S_1M'M = 0. \tag{F.11}$$

Since $S_1'S_1$ is invertible, (F.10) can be solve for M to get $M = S_1'\Gamma'S_1(S_1'S_1)^{-1}$. Substituting in (F.11), I get

$$S_1(S_1'S_1)^{-1} = S_1 - \Gamma S_1S_1'\Gamma'S_1(S_1'S_1)^{-1} - \Gamma'S_1(S_1'S_1)^{-1}S_1'\Gamma S_1 + S_1(S_1'S_1)^{-1}S_1'\Gamma S_1S_1'\Gamma'S_1(S_1'S_1)^{-1}, \tag{F.12}$$

where I am using the fact that $S_1^\dagger = (S_1'S_1)^{-1}S_1'$. Next consider the singular-value decomposition of S_1:

$$S_1 = U\Sigma V', \tag{F.13}$$

where $U \in \mathbb{R}^{n\times n}$ and $V \in \mathbb{R}^{d\times d}$ are orthogonal matrices, and $\Sigma \in \mathbb{R}^{n\times d}$ is a rectangular diagonal matrix. Substituting for S_1 in (F.12) from (F.13) and multiplying the result from left and right by U' and $V\Sigma'$, respectively, I get that $\Sigma(\Sigma'\Sigma)^{-1}\Sigma'$ is equal to

$$\Sigma\Sigma' - X\Sigma\Sigma'X'\Sigma(\Sigma'\Sigma)^{-1}\Sigma' - X'\Sigma(\Sigma'\Sigma)^{-1}\Sigma'X\Sigma\Sigma' + \Sigma(\Sigma'\Sigma)^{-1}\Sigma'X\Sigma\Sigma'X'\Sigma(\Sigma'\Sigma)^{-1}\Sigma', \tag{F.14}$$

where $X \equiv U'\Gamma U$. Note that $\Sigma = \begin{pmatrix} \Sigma_1 \\ \mathbf{0} \end{pmatrix}$ for some diagonal matrix $\Sigma_1 \in \mathbb{R}^{n\times d}$. Moreover, since $S_1'S_1$ is invertible, so is Σ_1. Therefore, $\Sigma(\Sigma'\Sigma)^{-1}\Sigma' = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$ and $\Sigma\Sigma' = \begin{pmatrix} \Sigma_1^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$. Write

$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$, where $X_{11} \in \mathbb{R}^{d\times d}$, $X_{12} \in \mathbb{R}^{d\times(n-d)}$, $X_{21} \in \mathbb{R}^{(n-d)\times d}$, and $X_{22} \in \mathbb{R}^{(n-d)\times(n-d)}$. Equation (F.14) then implies

$$X_{11}'X_{11} = I - \Sigma_1^{-2}, \tag{F.15}$$

$$X_{21}\Sigma_1^2 X_{11}' + X_{12}'X_{11}\Sigma_1^2 = 0. \tag{F.16}$$

These equations fully characterize the set of all (local) extrema of the KLDR.

I next use these equations to show that, as long as either d is equal to one or Γ_1 is symmetric, and for any $i = 1,\ldots,d$, the ith coordinate vector $e_i \in \mathbb{R}^n$ is an eigenvector of $(X+X')/2$ with eigenvalue $e_i'Xe_i$.[9] If $e_i'Xe_i = 0$, then trivially e_i is an eigenvector of $(X+X')/2$ with eigenvalue $e_i'Xe_i = 0$. So in the rest of the proof, I consider the case where $e_i'Xe_i \neq 0$. First, suppose $d = 1$. Then $i = 1$ and $X_{11}' = X_{11} = e_1'Xe_1 \neq 0$. On the other hand, Σ_1 is a non-zero scalar. Equation (F.16) then implies that $X_{21} + X_{12}' = 0$. Therefore,

$$\left(\frac{X+X'}{2}\right)e_1 = \frac{1}{2}\begin{pmatrix} 2X_{11} & X_{12}+X_{21}' \\ X_{21}+X_{12}' & X_{22}+X_{22}' \end{pmatrix}\begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} X_{11} \\ \mathbf{0} \end{pmatrix} = e_1'Xe_1e_1,$$

proving that e_1 is an eigenvector of $(X+X')/2$ with eigenvalue $e_1'Xe_1$. Next, suppose Γ_1 is symmetric. This implies that Γ, and by extension, X are symmetric matrices. Equation (F.15) then implies that X_{11} is a diagonal matrix. Since Σ_1 is also diagonal, it commutes with X_{11}. Equation (F.16) then implies that

$$(X_{21}+X_{12}')X_{11} = 2X_{21}X_{11} = \mathbf{0}, \tag{F.17}$$

where I am using the fact that Σ_1 is non-singular and X is symmetric. But since X_{11} is a diagonal matrix, it can be written as $X_{11} = \sum_{k=1}^{d} e_k'X_{11}e_ke_ke_k'$. Substituting in (F.17), I get $\sum_{k=1}^{d} X_{21}e_k'X_{11}e_ke_ke_k' = \mathbf{0}$. In particular, it must be the case that $X_{21}e_i'X_{11}e_ie_i = 0$. But since $e_i'X_{11}e_i = e_i'Xe_i \neq 0$, it must be that $X_{21}e_i = 0$. Therefore,

$$\left(\frac{X+X'}{2}\right)e_i = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}\begin{pmatrix} e_i \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} X_{11}e_i \\ X_{21}e_i \end{pmatrix} = \begin{pmatrix} e_i'X_{11}e_ie_i \\ 0 \end{pmatrix} = e_i'Xe_ie_i,$$

where the third equality relies on the fact that X_{11} is diagonal. This proves that e_i is an eigenvector of $(X+X')/2$ with eigenvalue $e_i'Xe_ie_i$.

I next show that any matrices M and S_1 that satisfy the first-order optimality conditions

[9] With slight abuse of notation, I use e_i to denote the ith coordinate vector both in $\mathbb{R}^n$ and in $\mathbb{R}^d$. Whether $e_i \in \mathbb{R}^n$ or $e_i \in \mathbb{R}^d$ will be clear from the context.

(F.10) and (F.11) must be of the form

$$M = \sum_{i=1}^{d} a_i v_i v_i', \tag{F.18}$$

$$S_1 = \sum_{i=1}^{d} \frac{1}{\sqrt{1-a_i^2}} u_i v_i', \tag{F.19}$$

where $\{a_i\}_{i=1}^d$ are eigenvalues of C_1, $u_i \in \mathbb{R}^n$ denotes an eigenvector with eigenvalue a_i normalized such that $u_i' u_k = \mathbb{1}_{\{i=k\}}$ for all $i, k \in \{1, \ldots, d\}$, and $\{v_i\}_{i=1}^d$ is an orthonormal basis for $\mathbb{R}^d$. To see this, first note that equation (F.13) can be written as $S_1 = U \Sigma V' = U \sum_{i=1^d} \sigma_i e_i e_i' V'$, where σ_i denotes the ith diagonal element of $\Sigma \in \mathbb{R}^{n \times d}$. I let $u_i \equiv U e_i$ and $v_i \equiv V e_i$. Since U and V are orthogonal matrices, $\{u_i\}_{i=1}^d$ is a set of orthonormal vectors and $\{v_i\}_{i=1}^d$ is an orthonormal basis for $\mathbb{R}^d$. Therefore, to show that S_1 takes the form given in (F.19), I only need to show that u_i is an eigenvector of C_1 with eigenvalue a_i and $\sigma_i = 1/\sqrt{1-a_i^2}$. Note that

$$C_1 u_i = \frac{1}{2}\left(\Gamma + \Gamma'\right) U e_i = \frac{1}{2} U U' \left(\Gamma + \Gamma'\right) U e_i = U \left(\frac{X + X'}{2}\right) e_i = U e_i' X e_i e_i = e_i' X e_i u_i,$$

where the fourth equality uses the fact that e_i is an eigenvector of $(X + X')/2$. Therefore, u_i is an eigenvector of C_1. On the other hand, multiplying equation (F.15) from left and right by e_i' and e_i, respectively, for $i \in \{1, \ldots, d\}$ and using the fact that X_{11} is diagonal, I get $\left(e_i' X_{11} e_i\right)^2 = 1 - \sigma_i^{-2}$. But

$$e_i' X_{11} e_i = e_i' X e_i = e_i' \left(\frac{X + X'}{2}\right) e_i = e_i' U \left(\frac{\Gamma + \Gamma'}{2}\right) U e_i = u_i' C_1 u_i = u_i' a_i u_i = a_i,$$

where a_i denotes the eigenvalue of C_1 with eigenvector u_i. Therefore, $\sigma_i = 1/\sqrt{1-a_i^2}$. Finally, recall that $M = S_1' \Gamma' S_1 (S_1' S_1)^{-1}$. By assumption, either $d = 1$, and so, S_1 is a vector in $\mathbb{R}^n$ or Γ is symmetric. Either way $S_1' \Gamma' S_1 = S_1' (\Gamma + \Gamma') S_1 / 2 = S_1' C_1 S_1$. Therefore,

$$\begin{aligned}
M &= S_1' C_1 S_1 (S_1' S_1)^{-1} \\
&= \left(\sum_{i,k=1}^{d} \frac{1}{\sqrt{1-a_i^2}} v_i u_i' C_1 \frac{1}{\sqrt{1-a_k^2}} u_k v_k'\right)\left(\sum_{i,k=1}^{d} \frac{1}{\sqrt{1-a_i^2}} v_i u_i' \frac{1}{\sqrt{1-a_k^2}} u_k v_k'\right)^{-1} \\
&= \left(\sum_{i=1}^{d} \frac{1}{1-a_i^2} v_i a_i v_i'\right)\left(\sum_{i=1}^{d} \frac{1}{1-a_i^2} v_i v_i'\right)^{-1} = \sum_{i=1}^{d} a_i v_i v_i',
\end{aligned}$$

where I am using the facts that u_i is an eigenvector of C_1 with eigenvalue a_i and that $\{u_i\}_{i=1}^d$ and $\{v_i\}_{i=1}^d$ are orthonormal sets of vectors.

Although any M and S_1 of the forms (F.18) and (F.19) satisfy the necessary optimality conditions, not all such candidates are global minimizers of the KLDR. To find the global optima, I substitute the solutions to the first-order optimality conditions in the KLDR and select the solutions that minimize the KLDR. Multiplying equation (F.11) from left by S_1', I get $I = S_1'S_1 - S_1'\Gamma S_1 M - S_1'\Gamma' S_1 M' + S_1'S_1M'M$. Computing the trace of the above equation and substituting the result in (F.9), I get that (up to an additive constant)

$$\begin{aligned}\text{KLDR}(M, S_1) &= -\frac{1}{2}\log\det\left(S_1'S_1\right) + \frac{1}{2}\operatorname{tr}\left(S_1'S_1\right) - \operatorname{tr}\left(MS_1'\Gamma S_1\right) + \frac{1}{2}\operatorname{tr}\left(MS_1'S_1M'\right)\\ &= -\frac{1}{2}\log\det\left(S_1'S_1\right) + \frac{1}{2}\operatorname{tr}(I).\end{aligned}$$

Therefore, the M and S_1 pairs that minimize the KLDR are the ones that maximize the determinant of $S_1'S_1$. But since $S_1'S_1$ is a symmetric matrix with eigenvalues $\{1/(1-a_i^2)\}_{i=1}^d$, its determinant is equal to $\prod_{i=1}^d \frac{1}{1-a_i^2}$. Therefore, any M and S_1 pair that minimize the KLDR are of the forms (F.18) and (F.19) with $\{a_i\}_{i=1}^d$ the top d eigenvalues of C_1 in magnitude (with the possibility that some of the a_i are equal).

With the expressions for the pseudo-true M and S_1 in hand, I can prove the theorem.

Part (a). The forecasts given a model parameterized by matrices M, D, and N are given by equation (16). Using the definition of $S \equiv \Gamma_0^{\frac{1}{2}}N$ and the fact that $D'D$ can be taken to be identity matrix, I can write equation (16) as follows: $E_t^\theta[y_{t+s}] = \Gamma_0^{\frac{1}{2}}S'^{-1}DM^sD'S'\Gamma_0^{\frac{-1}{2}}y_{t-\tau}$. Note that for any matrix $S = (S_1\ S_2)$ that satisfies the first-order optimality condition with respect to S_2,

$$S^{-1} = \begin{pmatrix}(S_1'S_1)^{-1}S_1'\\ S_2'\end{pmatrix}.$$

Therefore, $S'^{-1} = \begin{pmatrix}S_1(S_1'S_1)^{-1} & S_2\end{pmatrix}$, and so,

$$S'^{-1}D = S_1(S_1'S_1)^{-1}. \tag{F.20}$$

The forecasts can thus be written only in terms of matrices M and S_1 as follows:

$$E_t^\theta[y_{t+s}] = \Gamma_0^{\frac{1}{2}}S_1(S_1'S_1)^{-1}M^sS_1'\Gamma_0^{\frac{-1}{2}}.$$

Substituting for M and S_1 using (F.18) and (F.19) and simplifying the resulting expression, I get $E_t^{\theta}[y_{t+s}] = \Gamma_0^{\frac{1}{2}} \sum_{i=1}^{d} a_i^s u_i u_i' \Gamma_0^{\frac{-1}{2}}$. Letting $p_i \equiv \Gamma_0^{\frac{-1}{2}} u_i$ and $q_i \equiv \Gamma_0^{\frac{1}{2}} u_i$ completes the proof of part (a).

Part (b). Equation (17) gives the variance-covariance matrix under a model parameterized by matrices M, D, and N. Using the definition of S and setting $D'D = I$, equation (17) can be written as follows:

$$\mathrm{Var}^{\theta}(y) = \Gamma_0^{\frac{1}{2}} \left(S'^{-1} S^{-1} + S^{-1\prime} D \sum_{\tau=1}^{\infty} M^{\tau} M'^{\tau} D' S^{-1} \right) \Gamma_0^{\frac{1}{2}}.$$

To prove part (b), I need to show that the terms in parentheses add up to the identity matrix. I start with the first term:

$$S'^{-1} S^{-1} = (SS')^{-1} = (S_1 S_1' + S_2 S_2')^{-1}. \tag{F.21}$$

The fact that $S_2' S_2 = I$ implies that S_2 can be written as $S_2 = \sum_{i=d+1}^{n} u_i w_i'$, where $u_i \in \mathbb{R}^n$ and $w_i \in \mathbb{R}^{n-d}$ for $i = d+1, \ldots, n$, $\{u_i\}_{i=d+1}^{n}$ are orthonormal vectors, and $\{w_i\}_{i=d+1}^{n}$ constitutes an orthonormal basis for $\mathbb{R}^{n-d}$. On the other hand, the fact that $S_1' S_2 = \mathbf{0}$ implies that $u_i' u_k = 0$ for any $i \in \{1, \ldots, d\}$ and $k \in \{d+1, \ldots, n\}$. Therefore, $\{u_i\}_{i=1}^{n}$ constitutes an orthonormal basis for $\mathbb{R}^n$. Substituting for S_1 and S_2 in (F.21), I get

$$S'^{-1} S^{-1} = \left(\sum_{i=1}^{d} \frac{1}{1 - a_i^2} u_i u_i' + \sum_{i=d+1}^{n} u_i u_i' \right)^{-1} = \sum_{i=1}^{d} (1 - a_i^2) u_i u_i' + \sum_{i=d+1}^{n} u_i u_i',$$

where the second equality uses the fact that $\{u_i\}_{i=1}^{n}$ are orthonormal. Next consider the second term:

$$\begin{aligned} S^{-1\prime} D \sum_{\tau=1}^{\infty} M^{\tau} M'^{\tau} D' S^{-1} &= S_1 (S_1' S_1)^{-1} \sum_{\tau=1}^{\infty} M^{\tau} M'^{\tau} (S_1' S_1)^{-1} S_1' \\ &= \sum_{i=1}^{d} \sqrt{1 - a_i^2} u_i v_i' \sum_{\tau=1}^{\infty} \sum_{k=1}^{d} a_k^{2\tau} v_k v_k' \sum_{l=1}^{d} \sqrt{1 - a_l^2} v_l u_l' \\ &= \sum_{i=1}^{d} (1 - a_i^2) u_i u_i' \sum_{\tau=1}^{\infty} a_i^{2\tau} = \sum_{i=1}^{d} a_i^2 u_i u_i', \end{aligned}$$

where the first equality uses (F.20) and the second equality is by (F.18) and (F.19). Putting everything together,

$$\begin{aligned} S'^{-1}S^{-1} + S^{-1\prime}D\sum_{\tau=1}^{\infty} M^{\tau}M'^{\tau}D'S^{-1} &= \sum_{i=1}^{d}(1-a_i^2)u_iu_i' + \sum_{i=d+1}^{n} u_iu_i' + \sum_{i=1}^{d} a_i^2 u_iu_i' \\ &= \sum_{i=1}^{n} u_iu_i' = I, \end{aligned}$$

where the last equality follows the fact that $\{u_i\}_{i=1}^n$ is an orthonormal basis for $\mathbb{R}^n$. □

Proof of Proposition 3. Recall that I have assumed (without loss of generality) that Γ_0 is non-singular. Since C_1 is symmetric, $\{u_i\}_{i=1}^d$ constitutes an orthonormal basis for $\mathbb{R}^n$, and so, $\Gamma_0^{\frac{-1}{2}} y_t$ can be expressed as $\Gamma_0^{\frac{-1}{2}} y_t = \sum_{i=1}^n \omega_{it} u_i$, where $\omega_{it} \equiv u_i'\Gamma_0^{\frac{-1}{2}} y_t$. Therefore,

$$y_t = \Gamma_0^{\frac{1}{2}} \sum_{i=1}^{n} \omega_{it} u_i = \sum_{i=1}^{n} \Gamma_0^{\frac{1}{2}} u_i u_i' \Gamma_0^{\frac{-1}{2}} y_t = \sum_{i=1}^{n} y_t^{(i)} q_i,$$

where the last equality uses the definitions of $y_t^{(i)}$ and q_i.

The lag-one autocovariance of $y_t^{(i)}$ is given by

$$\mathbb{E}\left[y_t^{(i)} y_{t-1}^{(i)}\right] = p_i{}'\mathbb{E}[y_t y_{t-1}]p_i = p_i{}'\Gamma_1 p_i = p_i{}'\left(\frac{\Gamma_1 + \Gamma_1'}{2}\right)p_i = p_i{}'\Gamma_0^{\frac{1}{2}} C_1 \Gamma_0^{\frac{1}{2}} p_i = u_i' C_1 u_i,$$

where the first equality uses the definition $y_t^{(i)}$, and the last equality uses the definition of p_i. The fact that u_i is an eigenvector of C_1 implies $u_i'C_1u_i = a_iu_i'u_i = a_i$, where a_i is the ith largest (in magnitude) eigenvalue of C_1. Moreover, $\mathbb{E}\left[{y_t^{(i)}}^2\right] = p_i{}'\Gamma_0 p_i = u_i'u_i = 1$. Therefore, $\rho_i \equiv \mathbb{E}\left[y_t^{(i)} y_{t-1}^{(i)}\right] / \sqrt{\mathbb{E}\left[{y_t^{(i)}}^2\right]} = a_i$. The proposition follows the fact that a_i is the ith largest eigenvalue of C_1 in magnitude. □

Proof of Proposition 4. I prove the result under the assumption that the top D eigenvalues of the first autocorrelation matrix, C_1, are all distinct. This assumption is true for generic true processes. By Theorem 5(a) (or Theorem 4), the forecasts of an agent who uses a pseudo-true d-state model θ are given by

$$E_t^{\theta}[y_{t+s}] = \sum_{i=1}^{d} a_i{}^s q_i p_i{}' y_t, \tag{F.22}$$

where a_i is the ith largest eigenvalue of C_1, u_i denotes the corresponding eigenvector, normalized to have unit norm, $p_i \equiv \Gamma_0^{\frac{-1}{2}} u_i$, and $q_i \equiv \Gamma_0^{\frac{1}{2}} u_i$. Since the eigenvalues of C_1 are all distinct, the corresponding eigenvectors are unique (up to multiplicative constants). Therefore, all agents use the same values of $\{(a_i, p_i, q_i)\}_i$ to forecast.

Consider agent j who is constrained to models of dimension d_j. The agent's optimal action given her pseudo-true d-state model is given by

$$x_{jt} = E_t^{\theta_j}\left[\sum_{s=1}^{\infty} c'_{js} y_{t+s}\right] = \sum_{s=1}^{\infty} c'_{js} E_t^{\theta_j}\left[y_{t+s}\right] = \sum_{s=1}^{\infty} c'_{js} \sum_{i=1}^{d_j} a_i{}^s q_i p_i{}' y_t = \sum_{i=1}^{d_j} g_{ji} y_t^{(i)},$$

where θ_j denotes agent j's pseudo-true model, $y_t^{(i)} \equiv p_i' y_t$ as before, $g_{ji} \equiv \sum_{s=1}^{\infty} a_i{}^s c'_{js} q_i$ is a constant, which is a finite since $\{c_{js}\}_s$ is absolutely summable. Using vector notation, $x_t \equiv (x_{1t}, \ldots, x_{Jt})' \in \mathbb{R}^J$, I can write the above expression as $x_t = G y_t^{(1:D)}$, where $G \equiv \begin{pmatrix} g'_1 & g'_2 & \cdots & g'_J \end{pmatrix}' \in \mathbb{R}^{J\times D}$, $g_j \equiv \begin{pmatrix} g_{j1} & g_{j2} & \ldots g_{jd_j} & 0 & \ldots & 0 \end{pmatrix} \in \mathbb{R}^{1\times D}$, and $y_t^{(1:D)} \equiv \begin{pmatrix} y_t^{(1)} & y_t^{(2)} & \cdots & y_t^{(D)} \end{pmatrix}' \in \mathbb{R}^D$. □

References

EUSEPI, STEFANO AND BRUCE PRESTON (2018): "Fiscal Foundations of Inflation: Imperfect Knowledge," *American Economic Review*, 108 (9), 2551–89.

HAYASHI, FUMIO (1982): "Tobin's Marginal Q and Average Q: A Neoclassical Interpretation," *Econometrica*, 213–224.