

Estimating Models with Dispersed Information

Leonardo Melosi

Federal Reserve Bank of Chicago

Online Appendix

In Appendix A we show in detail how to derive the law of motion of the price level and of real output (20) and (19) in the main text. Appendix B explains in detail how to solve the ICKM. In Appendix C, we analytically derive the mapping $g_a(\cdot)$ and show how to compute the mapping $g_m(\cdot)$ of Section III.A. The price-setting equations used to simulate firms' prices to obtain the distribution of price changes are derived in Appendix D. In Section III.D, we compute the median of the absolute sizes of price changes from this distribution. In Appendix E, we show how to characterize the profit function (27) in the allocation of attention problem. In Appendix F, we assess the robustness of the MDD-based comparison conducted in Section II.C. In Appendix G, we extend the ICKM by considering an monetary policy rule according to which the monetary authority sets the growth rate of money in response to price fluctuations.

A Law of Motion of Price and Output in the ICKM

The first-order necessary condition¹ of the price-setting problem in the ICKM is:

$$\mathbb{E}_{i,t} \left[\beta Q_t \left(Y_{i,t} - \nu P_{i,t} \left(\frac{P_{i,t}}{P_t} \right)^{-\nu-1} \frac{Y_t}{P_t} + \nu \phi \frac{W_t}{A_{i,t}} \left(\frac{Y_{i,t}}{A_{i,t}} \right)^{\phi-1} \left(\frac{P_{i,t}}{P_t} \right)^{-\nu-1} \frac{Y_t}{P_t} \right) \right] = 0$$

Use the equation (2) in the main text to write:

$$\begin{aligned} \mathbb{E}_{i,t} & \left[\beta Q_t \left(\left(\frac{P_{i,t}}{P_t} \right)^{-\nu} Y_t - \nu P_t^i \left(\frac{P_{i,t}}{P_t} \right)^{-\nu-1} \frac{Y_t}{P_t} \right) \right] + \\ & + \mathbb{E}_{i,t} \left[\beta Q_t \nu \phi^{-1} \frac{W_t}{A_{i,t}} \left(\frac{1}{A_{i,t}} \left(\frac{P_{i,t}}{P_t} \right)^{-\nu} Y_t \right)^{\phi^{-1}-1} \left(\frac{P_{i,t}}{P_t} \right)^{-\nu-1} \frac{Y_t}{P_t} \right] = 0 \end{aligned}$$

¹Note the slight change in notation from the main text. We denote $\mathbb{E}[\cdot | \mathbf{z}_i^t] = \mathbb{E}_{i,t}$.

From the solution to the representative household's problem, the labor supply can be easily shown to be $W_t/P_t = \alpha Y_t H_t^\gamma$. Substituting this result into the equation above yields:

$$\begin{aligned} \mathbb{E}_{i,t} & \left[\beta Q_t (1 - \nu) \left(\frac{P_{i,t}}{P_t} \right)^{-\nu} Y_t \right] + \\ + & \mathbb{E}_{i,t} \left[\beta Q_t \nu \phi^{-1} \frac{\alpha Y_t H_t^\gamma}{A_t e^{\eta_{i,t}}} \left(\frac{1}{A_t e^{\eta_{i,t}}} \left(\frac{P_{i,t}}{P_t} \right)^{-\nu} Y_t \right)^{\phi^{-1}-1} \left(\frac{P_{i,t}}{P_t} \right)^{-\nu-1} Y_t \right] = 0 \end{aligned}$$

Define the stationary variables:

$$y_t \equiv \frac{Y_t}{A_t}, \quad y_{i,t} \equiv \frac{Y_{i,t}}{A_t}, \quad p_{i,t} = \frac{P_{i,t}}{P_t}, \quad h_{i,t} = H_{i,t} \quad (1)$$

With this notation, I can rewrite the price-setting equation as:

$$\begin{aligned} \mathbb{E}_{i,t} & \left[\beta q_t (1 - \nu) p_{i,t}^{-\nu} y_t \right] + \\ + & \mathbb{E}_{i,t} \left[\beta q_t \nu \phi^{-1} \frac{\alpha y_t h_t^\gamma}{e^{\eta_{i,t}}} \left(\frac{1}{e^{\eta_{i,t}}} (p_{i,t})^{-\nu} y_t \right)^{\phi^{-1}-1} (p_{i,t})^{-\nu-1} y_t \right] = 0 \\ \mathbb{E}_{i,t} & \left[\beta q_t p_{i,t}^{-\nu} y_t \left((1 - \nu) + \nu \phi^{-1} \frac{\alpha y_t h_t^\gamma}{e^{\eta_{i,t}}} \left(\frac{1}{e^{\eta_{i,t}}} (p_{i,t})^{-\nu} y_t \right)^{\phi^{-1}-1} (p_{i,t})^{-1} \right) \right] = 0 \end{aligned}$$

It is easy to show that the expression within the round brackets is zero at the deterministic symmetric steady-state. Hence, when one takes the log-linear approximation of the equation above around the deterministic symmetric steady-state, one does not need to care about what is outside those brackets. Hence the price-setting condition can be approximated as follows:

$$0 = \mathbb{E}_{i,t} \left[\gamma \hat{h}_t - [\nu (\phi^{-1} - 1) + 1] \hat{p}_{i,t} + \phi^{-1} (\hat{y}_t - \eta_{i,t}) \right]$$

Note also that from the production function $\hat{h}_{i,t} = \phi^{-1} (\hat{y}_{i,t} - \eta_{i,t})$ and hence² $\hat{h}_t = \phi^{-1} \hat{y}_t$. By substituting, this results into the equation above, one obtains:

$$0 = \mathbb{E}_{i,t} \left[\gamma \phi^{-1} \hat{y}_t + \phi^{-1} (\hat{y}_t - \eta_{i,t}) - [\nu (\phi^{-1} - 1) + 1] \hat{p}_{i,t} \right]$$

and then

$$\begin{aligned} [\nu (\phi^{-1} - 1) + 1] \mathbb{E}_{i,t} \hat{p}_{i,t} & = \mathbb{E}_{i,t} [\gamma \phi^{-1} \hat{y}_t + \phi^{-1} (\hat{y}_t - \eta_{i,t})] \\ \mathbb{E}_{i,t} \hat{p}_{i,t} & = \frac{\mathbb{E}_{i,t} [(\gamma + 1) \phi^{-1} \hat{y}_t - \phi^{-1} \eta_{i,t}]}{[\nu (\phi^{-1} - 1) + 1]} \end{aligned}$$

$$\mathbb{E}_{i,t} \hat{p}_{i,t} = \frac{(\gamma + 1) \phi^{-1}}{\nu (\phi^{-1} - 1) + 1} \mathbb{E}_{i,t} \hat{y}_t - \frac{\phi^{-1}}{\nu (\phi^{-1} - 1) + 1} \mathbb{E}_{i,t} \eta_{i,t}$$

²Log-linearizing $Y_t = \left(\int_0^1 (Y_{i,t})^{\frac{\nu-1}{\nu}} di \right)^{\frac{\nu}{\nu-1}}$ yields $\hat{y}_t = \int \hat{y}_{i,t} di$.

and more compactly, by defining $\lambda \equiv (\gamma + 1) \phi^{-1} / [\nu (\phi^{-1} - 1) + 1]$,

$$\mathbb{E}_{i,t} [\hat{p}_{i,t}] = \lambda \mathbb{E}_{i,t} \hat{y}_t - \frac{\lambda}{\gamma + 1} \mathbb{E}_{i,t} \eta_{i,t}$$

In order to take firm i 's price $P_{i,t}$ out of the expectation operator, I need to recall the definition of the transformed variables in (1) and then write:

$$\mathbb{E}_{i,t} \left[\underbrace{\ln P_{i,t} - \ln P_t}_{\hat{p}_{i,t}} \right] = \lambda \mathbb{E}_{i,t} \left[\underbrace{\ln Y_t - \ln A_t - \ln \bar{y}}_{\hat{y}_t} \right] - \frac{\lambda}{\gamma + 1} \mathbb{E}_{i,t} \eta_{i,t}$$

or equivalently,

$$\ln P_{i,t} = \mathbb{E}_{i,t} [\lambda \ln Y_t + \ln P_t - \lambda \ln A_t] - \frac{\lambda}{\gamma + 1} \mathbb{E}_{i,t} \eta_{i,t} - \lambda \ln \bar{y}$$

Recall equation (9):

$$\ln P_t + \ln Y_t = \ln M_t \Rightarrow \ln Y_t = \ln M_t - \ln P_t$$

and thus,

$$\ln P_{i,t} = \mathbb{E}_{i,t} \left[\lambda (\ln M_t - \ln P_t) + \ln P_t - \lambda \ln A_t - \frac{\lambda}{\gamma + 1} \eta_{i,t} \right] - \lambda \ln \bar{y}$$

and by rearranging:

$$\ln P_{i,t} = \mathbb{E}_{i,t} \left[(1 - \lambda) \ln P_t + \lambda \ln M_t - \lambda \ln A_t - \frac{\lambda}{\gamma + 1} \eta_{i,t} \right] - \lambda \ln \bar{y}$$

This price-setting equation shows that the parameter $1 - \lambda$ controls the strategic complementarity in price-setting (i.e., the extent to which firms want to react to the expected average price $\mathbb{E}_{i,t}(P_t)$). In order to have strategic complementarities in price-setting (i.e., firms want to raise (cut) their prices when the average price goes up (down)), one needs that $\lambda \leq 1$.

If one log-linearizes equation (4) of the main text around the deterministic steady-state, one obtains $\hat{p}_t = \int \hat{p}_{i,t} di$. Hence, by integrating across firms one obtains:

$$\ln P_t = (1 - \lambda) \ln P_{t|t}^{(1)} + \lambda \ln M_{t|t}^{(1)} - \lambda \ln A_{t|t}^{(1)} - \lambda \ln \bar{y}$$

where it can be shown that $\int \mathbb{E}_{i,t}(\eta_{i,t}) di = 0$. From this equation, repeatedly taking the conditional expectation and averaging across firms yield:

$$\ln P_{t|t}^{(j)} = (1 - \lambda) \ln P_{t|t}^{(j+1)} + \lambda \ln M_{t|t}^{(j+1)} - \lambda \ln A_{t|t}^{(j+1)} - \lambda \ln \bar{y}$$

for $j \in \{1, 2, \dots\}$. By repeatedly substituting these results into the average-price equation one obtains:

$$\ln P_t = \sum_{j=0}^{\infty} (1 - \lambda)^j \lambda \ln M_{t|t}^{(j+1)} - (1 - \lambda)^j \lambda \ln A_{t|t}^{(j+1)} - \lambda \ln \bar{y}$$

By recalling that I defined $m_t \equiv \ln M_t - \mu_M t$ and $a_t \equiv \ln A_t - \mu_A t$ and that firms know all the model parameters, I can re-write the equation above as:

$$\ln P_t = \left[\sum_{j=0}^{\infty} (1-\lambda)^j \lambda \left(m_{t|t}^{(j+1)} - a_{t|t}^{(j+1)} \right) \right] - \lambda \ln \bar{y} + \mu_M t - \mu_A t$$

This is measurement equation (20) in the main text.

Furthermore, I can combine equations (20) and (9) in the main text to get:

$$\underbrace{\ln M_t - \ln Y_t}_{\ln P_t} = \left[\sum_{j=0}^{\infty} (1-\lambda)^j \lambda \left(m_{t|t}^{(j+1)} - a_{t|t}^{(j+1)} \right) \right] - \lambda \ln \bar{y} + \mu_M t - \mu_A t$$

and by re-arranging, this yields:

$$\ln Y_t = \left[m_t - \sum_{j=0}^{\infty} (1-\lambda)^j \lambda m_{t|t}^{(j+1)} \right] + \sum_{j=0}^{\infty} (1-\lambda)^j \lambda a_{t|t}^{(j+1)} + \lambda \ln \bar{y} + \mu_A t$$

which is the measurement equation (19) in the main text.

B Solving the ICKM

Define the vector \mathbf{F}_t as

$$\mathbf{F}_t \equiv \sum_{j=1}^{\infty} (1-\lambda)^{j-1} \lambda \mathbf{X}_t^{(j)} \quad (2)$$

$$\text{where } \mathbf{X}_t \equiv [m_t, m_{t-1}, a_t]' \quad (3)$$

Finding an equilibrium for the ICKM requires characterizing the equilibrium law of motion of the finite-dimensional vector \mathbf{F}_t . The transition equations of the ICKM can be shown to be:

$$\hat{y}_t = -\hat{p}_t \quad (4)$$

$$\hat{p}_t = \mathbf{r}' \bar{\mathbf{X}}_t \quad (5)$$

$$\bar{\mathbf{X}}_t = \bar{\mathbf{B}} \bar{\mathbf{X}}_{t-1} + \bar{\mathbf{b}} \mathbf{u}_t \quad (6)$$

where \hat{y}_t and \hat{p}_t denote the log-linear deviations of the stationary output, $y_t \equiv Y_t/A_t$, and price, $p_t \equiv P_t A_t / M_t$, from their deterministic steady-state, respectively, and

$$\begin{aligned} \bar{\mathbf{X}}_t &\equiv \left[\mathbf{X}_t' \ ; \ \mathbf{F}_t' \right]', \quad \mathbf{r} \equiv [-1, 0, 1, 1, 0, -1]' \\ \bar{\mathbf{B}} &\equiv \begin{bmatrix} \mathbf{B}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{G}_{3 \times 3} & \mathbf{H}_{3 \times 3} \end{bmatrix}, \quad \bar{\mathbf{b}} = \left[\mathbf{b}' \ ; \ \mathbf{d}' \right]' \end{aligned} \quad (7)$$

$$\mathbf{B} \equiv \begin{bmatrix} 1 + \rho_m & -\rho_m & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{u}_t = [\varepsilon_{m,t}, \varepsilon_{a,t}]'$$

$$\mathbf{u}_t \stackrel{iid}{\sim} \mathcal{N}(\mathbf{0}, \Sigma_u), \text{ for all } t \text{ and } \Sigma_u = \begin{bmatrix} \sigma_m^2 & 0 \\ 0 & \sigma_a^2 \end{bmatrix}$$

where \mathbf{G} , \mathbf{H} , and \mathbf{d} are matrices that are not known yet. Equation (4) stems from detrending and log-linearizing the market clearing conditions in the money market: $P_t Y_t = M_t$. Equation (5) can be derived by equation (20) of the main text and recalling that

$$\hat{p}_t = \ln P_t + \ln A_t - \ln M_t - \ln \bar{p}$$

and

$$\ln \bar{p} + \ln \bar{y} = 0,$$

because of equation (9) in the main text .

Recall that the signal structure is specified in equations (16) of the main text. Thus, the firms' observation equations can be written as

$$\mathbf{z}_{i,t} = \mathbf{D}\bar{\mathbf{X}}_t + \mathbf{e}_{i,t} \quad (8)$$

where

$$\mathbf{D} \equiv \left[\mathbf{D}_1 \quad : \quad \mathbf{0}_{2 \times 3} \right] \text{ and } \mathbf{D}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{e}_{i,t} \equiv (\varepsilon_{m,i,t}, \varepsilon_{a,i,t})' \quad (9)$$

$$\mathbf{e}_{i,t} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_e), \text{ iid for all } t, \text{ and } i, \boldsymbol{\Sigma}_e = \begin{bmatrix} \tilde{\sigma}_m^2 & 0 \\ 0 & \tilde{\sigma}_a^2 \end{bmatrix} \quad (10)$$

Finding an equilibrium for this economy amounts to characterize the unknown matrices \mathbf{G} , \mathbf{H} , and \mathbf{d} . This requires solving the following fixed point problem. Given the conjectured law of motion (6), optimal firms' behavior must exactly aggregate to the conjectured law of motion (6). As suggested by Woodford (2002), the method of undetermined coefficients can be used to pin down those matrices.

It is easy to see that the firm i 's optimal estimate of the state vector evolves according the so-called *Kalman-filter equation*

$$\bar{\mathbf{X}}_{t|t}(i) = \bar{\mathbf{X}}_{t|t-1}(i) + \mathbf{k} [\mathbf{z}_t(i) - \mathbf{D}\bar{\mathbf{X}}_{t|t-1}(i)] \quad (11)$$

where \mathbf{k} is the 6×2 Kalman gain matrix which is not yet specified. It is easy to show that the one-step-ahead forecast of the state vector is:

$$\bar{\mathbf{X}}_{t|t-1}(i) = \bar{\mathbf{B}}\bar{\mathbf{X}}_{t-1|t-1}(i) \quad (12)$$

I can plug (12) into (11) to get the law of motion for firm i 's estimate of the current state vector

$$\bar{\mathbf{X}}_{t|t}(i) = \bar{\mathbf{B}}\bar{\mathbf{X}}_{t-1|t-1}(i) + \mathbf{k} [\mathbf{z}_t(i) - \mathbf{D}\bar{\mathbf{X}}_{t|t-1}(i)] \quad (13)$$

By integrating (13) over firms one obtains

$$\bar{\mathbf{X}}_{t|t} = \bar{\mathbf{B}}\bar{\mathbf{X}}_{t-1|t-1} + \mathbf{k}\mathbf{D} [\bar{\mathbf{X}}_t - \bar{\mathbf{X}}_{t|t-1}] \quad (14)$$

This result follows from the observing that white noise errors and shocks wash out (i.e. $\int \mathbf{e}_t(i) di = \mathbf{0}$).

By using the transition equation (6) to get rid of $\bar{\mathbf{X}}_t$ in the equation (14) I obtain

$$\bar{\mathbf{X}}_{t|t} = \bar{\mathbf{B}}\bar{\mathbf{X}}_{t-1|t-1} + \mathbf{k}\mathbf{D} [\bar{\mathbf{B}}\bar{\mathbf{X}}_{t-1} + \bar{\mathbf{b}}\mathbf{u}_t - \bar{\mathbf{X}}_{t|t-1}]$$

Then by integrating (12), which yields the average prior forecast (i.e. $\overline{\mathbf{X}}_{t|t-1} = \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1}$), one notices that the above equation can be rewritten as

$$\overline{\mathbf{X}}_{t|t} = \overline{\mathbf{X}}_{t|t-1} + \mathbf{kD} [\overline{\mathbf{B}\mathbf{X}}_{t-1} + \overline{\mathbf{b}\mathbf{u}}_t - \overline{\mathbf{X}}_{t|t-1}]$$

Gathering the common terms yields

$$\overline{\mathbf{X}}_{t|t} = [\mathbf{I} - \mathbf{kD}] \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1} + \mathbf{kD} [\overline{\mathbf{B}\mathbf{X}}_{t-1} + \overline{\mathbf{b}\mathbf{u}}_t] \quad (15)$$

which can be regarded as the law of motion for the *average estimates* of the current state vector.

It is convenient to define the 6×3 vector $\boldsymbol{\varphi}$ such that

$$\boldsymbol{\varphi} \equiv \left[\lambda \cdot \mathbf{I}_3 \ ; \ (1 - \lambda) \cdot \mathbf{I}_3 \right]'$$

Then one can note the following

$$\boldsymbol{\varphi}' \overline{\mathbf{X}}_t^{(1)} = \mathbf{F}_t \quad (16)$$

It is easy to prove that equation (16) is indeed true by working as follows

$$\begin{aligned} \boldsymbol{\varphi}' \overline{\mathbf{X}}_t^{(1)} &= \left[(\lambda) \cdot \mathbf{I}_3 \ ; \ (1 - \lambda) \cdot \mathbf{I}_3 \right] \cdot \begin{bmatrix} \mathbf{X}_t^{(1)} \\ \dots \\ \mathbf{F}_t^{(1)} \end{bmatrix} \\ \boldsymbol{\varphi}' \overline{\mathbf{X}}_t^{(1)} &= \lambda \mathbf{X}_t^{(1)} + (1 - \lambda) \mathbf{F}_t^{(1)} \end{aligned} \quad (17)$$

Let us introduce the following notations:

$$x_{t|t}^{(k-1)} \equiv x_t^{(k)}, \quad \forall k \geq 1; \quad x_t^{(0)} \equiv x_t \quad (18)$$

where x_t is an arbitrary random variable. Hence I can write

$$\boldsymbol{\varphi}' \overline{\mathbf{X}}_t^{(1)} = \lambda \mathbf{X}_{t|t}^{(0)} + (1 - \lambda) \mathbf{F}_{t|t}^{(0)}$$

Moreover, it is easy to derive an equation for $\mathbf{F}_{t|t}$ from equation (2)

$$\mathbf{F}_{t|t}^{(0)} = \sum_{j=1}^{\infty} (1 - \lambda)^{j-1} \lambda \mathbf{X}_{t|t}^{(j)}$$

Combining the last two equations yields

$$\boldsymbol{\varphi}' \overline{\mathbf{X}}_t^{(1)} = \lambda \mathbf{X}_{t|t}^{(0)} + (1 - \lambda) \sum_{j=1}^{\infty} (1 - \lambda)^{j-1} \lambda \mathbf{X}_{t|t}^{(j)}$$

Some easy manipulations lead to

$$\begin{aligned}\varphi' \overline{\mathbf{X}}_t^{(1)} &= (\lambda) \mathbf{X}_{t|t}^{(0)} + \sum_{j=1}^{\infty} (1-\lambda)^j \lambda \mathbf{X}_{t|t}^{(j)} \\ &= \sum_{j=1}^{\infty} (1-\lambda)^{j-1} \lambda \mathbf{X}_{t|t}^{(j-1)}\end{aligned}$$

Now recall equation (18) to finally write

$$\varphi' \overline{\mathbf{X}}_t^{(1)} = \sum_{j=1}^{\infty} (1-\lambda)^{j-1} \lambda \mathbf{X}_t^{(j)}$$

Comparing this equation with the (2) concludes the proof of (16). Now one can plug equation (15) into equation (16) to get

$$\mathbf{F}_t = \left[\varphi' - \tilde{\mathbf{k}}\mathbf{D} \right] \overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1|t-1} + \tilde{\mathbf{k}}\mathbf{D} \left[\overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t \right] \quad (19)$$

where $\tilde{\mathbf{k}} \equiv \varphi' \mathbf{k}$. One can prove the following three facts:

FACT 1

$$\varphi' \overline{\mathbf{B}} = \left[\lambda \mathbf{B} + (1-\lambda) \mathbf{G} : ((1-\lambda)) \mathbf{H} \right]$$

FACT 2

$$\begin{aligned}\mathbf{D}\overline{\mathbf{B}} &= \left[\mathbf{D}_1 \mathbf{B} : \mathbf{0}_{2 \times 3} \right] \\ &= \left[\mathbf{B}^\dagger : \mathbf{0}_{2 \times 3} \right]\end{aligned} \quad (20)$$

where $\mathbf{B}^\dagger \equiv \left[\mathbf{B}'_1 \quad \mathbf{B}'_3 \right]'$ and \mathbf{B}_j stands for the j -th row of \mathbf{B} .

FACT 3

$$\begin{aligned}\mathbf{D}\overline{\mathbf{b}} &= \mathbf{D}_1 \mathbf{b} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}(2)\end{aligned}$$

Then note that the FACT 3 can be used to show that

$$\tilde{\mathbf{k}}\mathbf{D}\overline{\mathbf{b}}\mathbf{u}_t = \tilde{\mathbf{k}}\mathbf{u}_t$$

The FACT 2 allows is to get the following results:

$$\tilde{\mathbf{k}}\mathbf{D}\overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1} = \tilde{\mathbf{k}}\mathbf{B}^\dagger \mathbf{X}_{t-1}$$

and

$$\tilde{\mathbf{k}}\mathbf{D}\overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1|t-1} = \tilde{\mathbf{k}}\mathbf{B}^\dagger \mathbf{X}_{t-1|t-1}$$

Then the FACT 1 can be used in order to prove the following result

$$\varphi' \overline{\mathbf{B}} \mathbf{X}_{t-1|t-1} = \lambda \mathbf{B} \mathbf{X}_{t-1|t-1} + (1 - \lambda) \mathbf{G} \mathbf{X}_{t-1|t-1} + (1 - \lambda) \mathbf{H} \cdot \mathbf{F}_{t-1|t-1}$$

By collecting all these results one can rewrite equation (19) as follows

$$\mathbf{F}_t = \left[\lambda \mathbf{B} + (1 - \lambda) \mathbf{G} - \tilde{\mathbf{k}} \mathbf{B}^\dagger \right] \mathbf{X}_{t-1|t-1} + (1 - \lambda) \mathbf{H} \mathbf{F}_{t-1|t-1} + \tilde{\mathbf{k}} \mathbf{B}^\dagger \mathbf{X}_{t-1} + \tilde{\mathbf{k}} \mathbf{u}_t \quad (21)$$

Next, I will work out the vector \mathbf{F}_{t-1} from $\mathbf{F}_{t-1|t-1}$, since I want to rewrite equation (21) in a form that is comparable to that conjectured in equation (6) so as I can compare our initial guess. One should start from equation (16) to get

$$(1 - \lambda) \cdot \mathbf{F}_{t|t} = \mathbf{F}_t - \lambda \mathbf{X}_{t|t}$$

By lagging the last equation by one period, one gets

$$(1 - \lambda) \cdot \mathbf{F}_{t-1|t-1} = \mathbf{F}_{t-1} - \lambda \mathbf{X}_{t-1|t-1} \quad (22)$$

I can now plug equation (22) into equation (21) to get

$$\begin{aligned} \mathbf{F}_t &= \left[\lambda \mathbf{B} + (1 - \lambda) \mathbf{G} - \tilde{\mathbf{k}} \mathbf{B}^\dagger \right] \mathbf{X}_{t-1|t-1} + \mathbf{H} \left[\mathbf{F}_{t-1} - \lambda \mathbf{X}_{t-1|t-1} \right] + \tilde{\mathbf{k}} \mathbf{B}^\dagger \mathbf{X}_{t-1} + \tilde{\mathbf{k}} \mathbf{u}_t \\ \mathbf{F}_t &= \left[\lambda \mathbf{B} + (1 - \lambda) \mathbf{G} - \tilde{\mathbf{k}} \mathbf{B}^\dagger - \lambda \mathbf{H} \right] \mathbf{X}_{t-1|t-1} + \mathbf{H} \cdot \mathbf{F}_{t-1} + \tilde{\mathbf{k}} \mathbf{B}^\dagger \mathbf{X}_{t-1} + \tilde{\mathbf{k}} \mathbf{u}_t \end{aligned} \quad (23)$$

Now equation (23) has the same form as the bottom rows of equation (6) because $\mathbf{X}_{t-1|t-1}$ does not depend on neither \mathbf{X}_{t-1} nor \mathbf{F}_{t-1} . Thus I can make the following identifications:

$$\mathbf{G} = \tilde{\mathbf{k}} \mathbf{B}^\dagger \quad (24)$$

$$\mathbf{d} = \tilde{\mathbf{k}} \quad (25)$$

and

$$\left[\lambda \mathbf{B} + (1 - \lambda) \mathbf{G} - \tilde{\mathbf{k}} \mathbf{B}^\dagger - \lambda \mathbf{H} \right] \stackrel{!}{=} 0$$

By substituting (24) into the last equation one obtains

$$\begin{aligned} \left[\mathbf{B} - \tilde{\mathbf{k}} \mathbf{B}^\dagger - \mathbf{H} \right] &\stackrel{!}{=} 0 \\ \mathbf{H} &\stackrel{!}{=} \mathbf{B} - \tilde{\mathbf{k}} \mathbf{B}^\dagger \end{aligned} \quad (26)$$

which identifies the matrix \mathbf{H} .

The matrix \mathbf{k} is the steady-state matrix of Kalman gains which is well-known to be equal to

$$\mathbf{k} = \mathbf{P} \mathbf{D}' \left[\mathbf{D} \mathbf{P} \mathbf{D}' + \boldsymbol{\Sigma}_e \right]^{-1} \quad (27)$$

with the matrix \mathbf{P} that solves the following algebraic Riccati equation

$$\mathbf{P} = \overline{\mathbf{B}} \left[\mathbf{P} - \mathbf{P} \mathbf{D}' \left[\mathbf{D} \mathbf{P} \mathbf{D}' + \boldsymbol{\Sigma}_e \right]^{-1} \mathbf{D} \mathbf{P} \right] \overline{\mathbf{B}}' + \overline{\mathbf{b}} \boldsymbol{\Sigma}_u \overline{\mathbf{b}}' \quad (28)$$

and where $\mathbf{B}^\dagger \equiv \left[\mathbf{B}'_1 \quad \mathbf{B}'_3 \right]'$ and \mathbf{B}_j stands for the j -th row of \mathbf{B} .

Since $\bar{\mathbf{B}}$ and $\bar{\mathbf{b}}$ turn out to be function of \mathbf{P} , the ultimate goal is to find out the fixed-point of a larger equation to solve for \mathbf{P} , specified solely in terms of model parameters. Computationally, finding this fixed point turns out to be fast and reliable. This makes the ICKM suitable for estimation.

The loop to numerically find out a REE is the following: given a set of parameter values and a guess for the Kalman-gain matrix \mathbf{k}^0 , one has to characterize the matrices \mathbf{G} , \mathbf{H} , and \mathbf{d} through equations (24)-(26). Then one has to solve the algebraic Riccati equation (28) for \mathbf{P} and obtain a new Kalman-gain matrix \mathbf{k}^* through the equation (27). Then if the new Kalman-gain matrix is sufficiently close to the guess, one has just found the fixed point and stops, otherwise one goes through another loop by using the matrix \mathbf{k}^* as a new guess for the Kalman-gain matrix. Once a fixed point is found, one can use the resulting Kalman-gain matrix to fully characterize the state-space system of the ICKM model described in (6)-(7) through (24)-(28), which combined with the equations (4)-(5) delivers the equilibrium dynamics of the log-deviations of real output and inflation.

C Information Flows

As shown in the main text, the information flow κ_a is measured as follows:

$$\kappa_a \equiv H\left(a_{i,t}|z_{a,i}^{t-1}\right) - H\left(a_{i,t}|z_{a,i}^t\right) \quad (29)$$

Since a_t and $z_{a,i,t}$ are Gaussian, I can write:

$$H\left(a_{i,t}|z_{a,i}^t\right) \equiv \frac{1}{2} \log_2 \left[2\pi e \cdot \text{VAR}\left(a_{i,t}|z_{a,i}^t\right)\right] \quad (30)$$

First, let us focus on the mapping

$$\text{VAR}\left(a_{i,t}|z_{a,i}^t\right) = g\left(\tilde{\sigma}_a, \sigma_a\right)$$

The mapping $g_a(\cdot)$ can be implicitly characterized through the Kalman filter. The standard Kalman-equation for updating conditional variances is:

$$\text{VAR}\left(a_{i,t}|z_{a,i}^t\right) = \text{VAR}\left(a_{i,t}|z_{a,i}^{t-1}\right) - \frac{\text{VAR}\left(a_{i,t}|z_{a,i}^{t-1}\right)^2}{\text{VAR}\left(a_{i,t}|z_{a,i}^{t-1}\right) + \tilde{\sigma}_a^2}$$

One can show that $\text{VAR}\left(a_{i,t}|z_{a,i}^{t-1}\right) = \text{VAR}\left(a_{i,t-1}|z_{a,i}^{t-1}\right) + \sigma_a^2 + \sigma_\eta^2$. Plugging this result into the equation above and some straightforward manipulations yield

$$\text{VAR}\left(a_{i,t}|z_{a,i}^t\right) = \frac{\left[\text{VAR}\left(a_{i,t-1}|z_{a,i}^{t-1}\right) + \sigma_a^2 + \sigma_\eta^2\right] \tilde{\sigma}_a^2}{\text{VAR}\left(a_{i,t-1}|z_{a,i}^{t-1}\right) + \sigma_a^2 + \sigma_\eta^2 + \tilde{\sigma}_a^2}$$

Note that

$$\begin{aligned} \tilde{\sigma}_a^2 &= 0 \implies \text{VAR}\left(a_t|z_{a,i}^t\right) = 0 \\ \tilde{\sigma}_a^2 &\longrightarrow \infty \implies \text{VAR}\left(a_t|z_{a,i}^t\right) = \text{VAR}\left(a_t\right) \longrightarrow \infty \end{aligned}$$

where the last result follows from the fact that a_t follows a random walk. After manipulating a bit I obtain the quadratic equation:

$$VAR(a_{i,t}|z_{a,i}^t)^2 + VAR(a_{i,t}|z_{a,i}^t)(\sigma_a^2 + \sigma_\eta^2) = (\sigma_a^2 + \sigma_\eta^2)\tilde{\sigma}_a^2$$

This admits two solutions. There exists a unique acceptable solution ($VAR(a_t|z_{a,i}^t) \geq 0$) though, that is

$$VAR(a_{i,t}|z_{a,i}^t) = \frac{-(\sigma_a^2 + \sigma_\eta^2) + \sqrt{(\sigma_a^2 + \sigma_\eta^2)^2 + 4(\sigma_a^2 + \sigma_\eta^2)\tilde{\sigma}_a^2}}{2}$$

Note that I can write:

$$\begin{aligned} \sqrt{(\sigma_a^2 + \sigma_\eta^2)^2 + 4(\sigma_a^2 + \sigma_\eta^2)\tilde{\sigma}_a^2} &= 2VAR(a_{i,t}|z_{a,i}^t) + (\sigma_a^2 + \sigma_\eta^2) \\ \tilde{\sigma}_a^2 &= \frac{[2VAR(a_{i,t}|z_{a,i}^t) + (\sigma_a^2 + \sigma_\eta^2)]^2}{4(\sigma_a^2 + \sigma_\eta^2)} - \frac{\sigma_a^2 + \sigma_\eta^2}{4} \end{aligned}$$

and finally,

$$\tilde{\sigma}_a^2 = \frac{[2VAR(a_{i,t}|z_{a,i}^t) + \sigma_a^2]^2}{4\sigma_a^2} - \frac{\sigma_a^2}{4} \quad (31)$$

Now I need to find an expression for $VAR(a_t|z_{a,i}^t)$ in terms of the information flow κ_a and the variance σ_a .

Combining the equations (29) and (30) yields

$$\begin{aligned} \kappa_a &= H(a_t|z_{a,i}^{t-1}) - H(a_t|z_{a,i}^t) \\ \kappa_a &= \frac{1}{2} \log_2 \left(\frac{VAR(a_{i,t}|z_{a,i}^{t-1})}{VAR(a_{i,t}|z_{a,i}^t)} \right) \end{aligned}$$

Since firms observe infinitely many signals, $VAR(a_{i,t}|z_{a,i}^{t-1}) = VAR(a_{i,t}|z_{a,i}^t) + \sigma_a^2 + \sigma_\eta^2$. Hence I obtain:

$$\kappa_a = \frac{1}{2} \log_2 \left(\frac{VAR(a_{i,t}|z_{a,i}^t) + \sigma_a^2 + \sigma_\eta^2}{VAR(a_{i,t}|z_{a,i}^t)} \right)$$

If one inverts this equation, one obtains:

$$VAR(a_{i,t}|z_{a,i}^t) = \frac{\sigma_a^2 + \sigma_\eta^2}{2^{2\kappa_a} - 1} \quad (32)$$

Plugging this result into equation (31) leads to:

$$\kappa_a = \frac{1}{2} \log_2 \left[\frac{1}{\left(\frac{\tilde{\sigma}_a^2}{\sigma_a^2 + \sigma_\eta^2} + \frac{1}{4}\right)^{\frac{1}{2}} - \frac{1}{2}} + 1 \right] \quad (33)$$

This is the mapping g_a in equation (26) of the main text.

An analytical closed-form solution for the mapping g_m in equation (26) of the main text cannot be derived. I computationally approximate this mapping. To do that, I need to compute the conditional entropies $H(m_t|z_{m,i}^{t-1})$ and $H(m_t|z_{m,i}^t)$. Since the state m_t and signals $z_{m,i,t}$ are Gaussian, one can show that the conditional entropy is:

$$H(m_t|z_{1,i}^\tau) = \frac{1}{2} \log_2 [2\pi e \cdot VAR(m_t|z_{1,i}^\tau)] \quad (34)$$

Hence, I have to characterize the conditional variances of $VAR(m_t|z_{1,i}^\tau)$, $\tau \in \{t-1, t\}$. Let us define the variance-covariance matrices:

$$\mathbf{P}_{t|\tau} \equiv \mathbb{E} \left[(\bar{\mathbf{X}}_t - \mathbb{E}(\bar{\mathbf{X}}_t|\mathbf{z}_i^\tau)) (\bar{\mathbf{X}}_t - \mathbb{E}(\bar{\mathbf{X}}_t|\mathbf{z}_i^\tau))' | \mathbf{z}_i^\tau \right]$$

for $\tau \in \{t-1, t\}$, where $\bar{\mathbf{X}}_t \equiv [\mathbf{X}'_t \ : \ \mathbf{F}'_t]'$, $\mathbf{X}_t \equiv [m_t, m_{t-1}, a_t]'$, and $\mathbf{F}_t \equiv \sum_{j=1}^{\infty} (1-\lambda)^{j-1} \lambda \mathbf{X}_t^{(j)}$, as defined in Appendix B. It is easy to see that $VAR(m_t|z_{1,i}^{t-1}) = \mathbf{P}_{t|t-1}[1, 1]$ and $VAR(m_t|z_{1,i}^t) = \mathbf{P}_{t|t}[1, 1]$, where the numbers within square brackets denote the matrix component of interest. The matrix $\mathbf{P}_{t|t-1}$ is nothing but the matrix \mathbf{P} in Appendix B. See equation (28). The matrix $\mathbf{P}_{t|t}$ is defined as:

$$\mathbf{P}_{t|t} \equiv \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{D}' [\mathbf{D} \mathbf{P}_{t|t-1} \mathbf{D}' + \boldsymbol{\Sigma}_e]^{-1} \mathbf{D} \mathbf{P}_{t|t-1} \quad (35)$$

where the matrices \mathbf{D} and $\boldsymbol{\Sigma}_e$ have been defined in (9) and in (10), respectively.

Thus, after one has characterized the fixed point as discussed in Appendix B, one can use the resulting matrix \mathbf{P} and equation (35) to pin down the conditional variances $VAR(m_t|z_{1,i}^\tau)$, for $\tau \in \{t-1, t\}$, the condition entropies $H(m_t|z_{m,i}^\tau)$, for $\tau \in \{t-1, t\}$, through equation (34), and finally the information flow $\kappa_m \equiv H(m_t|z_{m,i}^{t-1}) - H(m_t|z_{m,i}^t)$.

D Simulating firm's prices

Firm's price setting equation is given by

$$\ln P_{i,t} = \mathbb{E}_{i,t} \left[(1-\lambda) \ln P_t + \lambda \ln M_t - \lambda \ln A_t - \frac{\lambda}{\gamma+1} \eta_{i,t} \right] - \lambda \ln \bar{y}$$

Recall that $\ln P_t = \left[\sum_{j=0}^{\infty} (1-\lambda)^j \lambda \left(m_{t|t}^{(j+1)} - a_{t|t}^{(j+1)} \right) \right] - \ln \bar{y} + \mu_M t - \mu_A t$. Then

$$\begin{aligned} \ln P_{i,t} &= \mathbb{E}_{i,t} \left[\sum_{j=0}^{\infty} (1-\lambda)^{j+1} \lambda \left(m_{t|t}^{(j+1)} - a_{t|t}^{(j+1)} \right) \right] - \ln \bar{y} + (1-\lambda) \mu_M t - (1-\lambda) \mu_A t \\ &\quad + \lambda \mathbb{E}_{i,t} \ln M_t - \lambda \mathbb{E}_{i,t} \ln A_t - \frac{\lambda}{\gamma+1} \mathbb{E}_{i,t} \eta_{i,t} \end{aligned}$$

recall that $m_t = \ln M_t - \mu_M t$ and $a_t = \ln A_t - \mu_A t$.

$$\begin{aligned} \ln P_{i,t} &= \sum_{j=0}^{\infty} (1-\lambda)^{j+1} \lambda \left(m_{t|t}^{(j+2)}(i) - a_{t|t}^{(j+2)}(i) \right) - \ln \bar{y} + \mu_M t - \mu_A t \\ &\quad + \lambda m_t^{(1)}(i) - \lambda a_t^{(1)}(i) - \frac{\lambda}{\gamma+1} \mathbb{E}_{i,t} \eta_{i,t} \end{aligned}$$

and finally

$$\ln P_{i,t} = \sum_{j=0}^{\infty} (1-\lambda)^j \lambda \left(m_{t|t}^{(j+1)}(i) - a_{t|t}^{(j+1)}(i) \right) - \frac{\lambda}{\gamma+1} \mathbb{E}_{i,t} \eta_{i,t} - \ln \bar{y} + \mu_M t - \mu_A t$$

So in order to simulate the price setting behavior at the micro level we need to keep track of $\mathbb{E}_{i,t} \eta_{i,t}$. In order to do that we need to set up the following state-space model model for firms, which hierarchical to the one analyzed in Appendix B:

$$\begin{aligned} \tilde{\mathbf{X}}_{i,t} &= \bar{\mathbf{B}} \tilde{\mathbf{X}}_{i,t} + \tilde{\mathbf{b}} \tilde{\mathbf{u}}_{i,t} \\ \mathbf{z}_{i,t} &= \tilde{\mathbf{D}} \tilde{\mathbf{X}}_{i,t} + \mathbf{e}_{i,t} \end{aligned}$$

where

$$\tilde{\mathbf{X}}_{i,t} \equiv \left[\bar{\mathbf{X}}_t' \ ; \ \eta_{i,t} \right]', \quad (36)$$

$$\tilde{\mathbf{B}} \equiv \begin{bmatrix} \bar{\mathbf{B}}_{6 \times 6} & \mathbf{0}_{6 \times 1} \\ \mathbf{0}_{1 \times 6} & \mathbf{0}_{1 \times 1} \end{bmatrix}, \quad \tilde{\mathbf{b}} = \begin{bmatrix} \bar{\mathbf{b}} & \mathbf{0}_{2 \times 1} \\ \mathbf{0}_{1 \times 2} & 1 \end{bmatrix} \vdots \quad (37)$$

$$\begin{aligned} \tilde{\mathbf{u}}_{i,t} &= [\varepsilon_{m,t}, \varepsilon_{a,t}, \eta_{i,t}]' \\ \tilde{\mathbf{u}}_t &\overset{iid}{\sim} \mathcal{N}(\mathbf{0}, \tilde{\Sigma}_u), \text{ for all } t \text{ and } \tilde{\Sigma}_u = \begin{bmatrix} \sigma_m^2 & 0 & 0 \\ 0 & \sigma_a^2 & 0 \\ 0 & 0 & \sigma_\eta^2 \end{bmatrix} \end{aligned}$$

where

$$\tilde{\mathbf{D}} \equiv \begin{bmatrix} \mathbf{D}_{2 \times 6} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{bmatrix} \text{ and } \mathbf{D}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{e}_{i,t} \equiv (\varepsilon_{m,i,t}, \varepsilon_{a,i,t})' \quad (38)$$

$$\mathbf{e}_{i,t} \sim \mathcal{N}(\mathbf{0}, \tilde{\Sigma}_e), \text{ iid for all } t, \text{ and } i, \tilde{\Sigma}_e = \begin{bmatrix} \tilde{\sigma}_m^2 & 0 \\ 0 & \tilde{\sigma}_a^2 \end{bmatrix} \quad (39)$$

This SS model delivers the following laws of motion

$$\tilde{\mathbf{X}}_{t|t}^{(1)}(i) = \bar{\mathbf{B}} \tilde{\mathbf{X}}_{t-1|t-1}^{(1)}(i) + \tilde{\mathbf{k}} \left[\mathbf{z}_t(i) - \mathbf{D} \bar{\mathbf{B}} \tilde{\mathbf{X}}_{t-1|t-1}^{(1)}(i) \right]$$

where

$$\tilde{\mathbf{k}} = \tilde{\mathbf{P}} \tilde{\mathbf{D}}' \left[\tilde{\mathbf{D}} \tilde{\mathbf{P}} \tilde{\mathbf{D}}' + \Sigma_e \right]^{-1} \quad (40)$$

with the matrix \mathbf{P} that solves the following algebraic Riccati equation

$$\tilde{\mathbf{P}} = \tilde{\mathbf{B}} \left[\tilde{\mathbf{P}} - \tilde{\mathbf{P}}\tilde{\mathbf{D}}' \left[\tilde{\mathbf{D}}\tilde{\mathbf{P}}\tilde{\mathbf{D}}' + \tilde{\Sigma}_e \right]^{-1} \tilde{\mathbf{D}}\tilde{\mathbf{P}} \right] \tilde{\mathbf{B}}' + \tilde{\mathbf{b}}\tilde{\Sigma}_u\tilde{\mathbf{b}}' \quad (41)$$

The price equation can then be represented as

$$\ln P_{i,t} = (\lambda, 0, -\lambda, 1 - \lambda, 0, -(1 - \lambda), -\lambda/(\gamma + 1)) \tilde{\mathbf{X}}_{t|t}^{(1)}(i) - \ln \bar{y} + \mu_M t - \mu_A t$$

To see this rewrite the price equation as follows:

$$\ln P_{i,t} = \sum_{j=0}^{\infty} (1 - \lambda)^j \lambda \left(m_{t|t}^{(j+1)}(i) - a_{t|t}^{(j+1)}(i) \right) - \frac{\lambda}{\gamma + 1} \mathbb{E}_{i,t} \eta_{i,t} - \ln \bar{y} + \mu_M t - \mu_A t$$

but note that

$$\begin{aligned} \ln P_{i,t} &= \lambda \left(m_{t|t}^{(1)}(i) - a_{t|t}^{(1)}(i) \right) + \underbrace{\sum_{j=1}^{\infty} (1 - \lambda)^j \lambda \left(m_{t|t}^{(j+1)}(i) - a_{t|t}^{(j+1)}(i) \right)}_{\text{related to } \mathbb{E}_{i,t}(F_t) = F_{t|t}^{(1)}(i) = \sum_{j=1}^{\infty} (1 - \lambda)^{j-1} \lambda X_{t|t}^{(j+1)}(i)} \\ &\quad - \frac{\lambda}{\gamma + 1} \mathbb{E}_{i,t} \eta_{i,t} - \ln \bar{y} + \mu_M t - \mu_A t \end{aligned}$$

E Optimal Allocation of Attention

The objective function in the attention problem (27)-(31) of the main text is defined as

$$\mathbb{E} \left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t \left(\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t \right) \right]$$

where $\hat{\pi}_t(\cdot)$ is the log-quadratic approximation of $Q_t \pi_t$, where Q_t is the stochastic discount factor that is treated as exogenous by firms. Let us define the profit-maximizing price (i.e., the price that solves the log-quadratic price-setting problem under perfect information. $\tilde{\sigma}_m = \tilde{\sigma}_a = 0$) as $\hat{p}_{i,t}^{\diamond}$. It is easy to show that

$$\hat{\pi}_t \left(\hat{p}_{i,t}^{\diamond}, \hat{p}_t, \hat{y}_t, \hat{q}_t \right) - \hat{\pi}_t \left(\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t \right) \propto \left(\hat{p}_{i,t}^{\diamond} - \hat{p}_{i,t}^* \right)^2 \quad (42)$$

up to a constant that is a function of structural parameters.

$$\hat{\pi}_t \left(\hat{p}_{i,t}^{\diamond}, \hat{p}_t, \hat{y}_t, \hat{q}_t \right) - \hat{\pi}_t \left(\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t \right) \propto \left(\ln \left(P_{i,t}^{\diamond} \right) - \ln \left(P_{i,t}^* \right) \right)^2 \quad (43)$$

First note that $\hat{\pi}_t \left(\hat{p}_{i,t}^{\diamond}, \hat{p}_t, \hat{y}_t, \hat{q}_t \right)$ is not affected by the attention problem as the profit-maximizing price is obtained by setting $\tilde{\sigma}_m = \tilde{\sigma}_a = 0$ (i.e., complete information). Hence the objective function

in the attention problem can be rewritten as

$$\begin{aligned} \mathbb{E} \left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t (\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t) \right] &\propto -\mathbb{E} \left[\sum_{t=1}^{\infty} \beta^t \left[\hat{\pi}_t (\hat{p}_{i,t}^\diamond, \hat{p}_t, \hat{y}_t, \hat{q}_t) - \hat{\pi}_t (\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t) \right] \right] \\ &\propto -\mathbb{E} \left[\sum_{t=1}^{\infty} \beta^t (\hat{p}_{i,t}^\diamond - \hat{p}_{i,t}^*)^2 \right] \end{aligned}$$

Since $\hat{p}_{i,t}^\diamond$ and $\hat{p}_{i,t}^*$ are stationary processes, they do not depend on t . Therefore,

$$\mathbb{E} \left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t (\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t) \right] \propto -\mathbb{E} \left[(\hat{p}_{i,t}^\diamond - \hat{p}_{i,t}^*)^2 \right]$$

Recall that in the log-quadratic problem $\hat{p}_{i,t}$ is defined as $\ln(P_{i,t}/P_t)$. Hence,

$$\begin{aligned} \hat{p}_{i,t}^\diamond &= \ln(P_{i,t}^\diamond) - \ln(P_t) \\ \hat{p}_{i,t}^* &= \ln(P_{i,t}^*) - \ln(P_t) \end{aligned}$$

It then follows

$$\mathbb{E} \left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t (\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t) \right] \propto -\mathbb{E} \left[\left(\ln(P_{i,t}^\diamond) - \ln(P_{i,t}^*) \right)^2 \right] \quad (44)$$

The price-setting equations under complete and incomplete information (see appendix A) are

$$\begin{aligned} \ln P_{i,t}^\diamond &= (1 - \lambda) \ln P_t + \lambda \ln M_t - \lambda \ln A_t - \frac{\lambda}{\gamma + 1} \eta_{i,t} - \lambda \ln \bar{y} \\ \ln P_{i,t}^* &= \mathbb{E}_{i,t} \left[(1 - \lambda) \ln P_t + \lambda \ln M_t - \lambda \ln A_t - \frac{\lambda}{\gamma + 1} \eta_{i,t} \right] - \lambda \ln \bar{y} \end{aligned}$$

Thus

$$\begin{aligned} \mathbb{E} \left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t (\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t) \right] &\propto \\ &-\mathbb{E} \left[\left\{ (1 - \lambda) \ln \bar{P}_t + \lambda \ln \bar{M}_t - \lambda \ln \bar{A}_t - \frac{\lambda}{\gamma + 1} \bar{\eta}_{i,t} \right\}^2 \right] \end{aligned}$$

where $\ln \bar{P}_t \equiv \ln P_t - \mathbb{E}_{i,t}(\ln P_t)$, $\ln \bar{M}_t \equiv \ln M_t - \mathbb{E}_{i,t}(\ln M_t)$, $\ln \bar{A}_t \equiv \ln A_t - \mathbb{E}_{i,t}(\ln A_t)$, and $\bar{\eta}_{i,t} = \eta_{i,t} - \mathbb{E}_{i,t}(\eta_{i,t})$. It follows that

$$\begin{aligned} \mathbb{E} \left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t (\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t) \right] &\propto - (1 - \lambda)^2 VAR(\tilde{\mathbf{r}}\tilde{\mathbf{X}}_t|\mathbf{z}_i^t) - \lambda^2 VAR(m_t|\mathbf{z}_i^t) \\ &\quad - \lambda^2 VAR(a_t|\mathbf{z}_i^t) - \left(\frac{\lambda}{\gamma + 1}\right)^2 VAR(\eta_{i,t}|\mathbf{z}_i^t) \\ &\quad - 2(1 - \lambda)\lambda COV(\tilde{\mathbf{r}}\tilde{\mathbf{X}}_t, m_t|\mathbf{z}_i^t) \\ &\quad + 2(1 - \lambda)\lambda COV(\tilde{\mathbf{r}}\tilde{\mathbf{X}}_t, a_t|\mathbf{z}_i^t) \\ &\quad + 2\frac{(1 - \lambda)\lambda}{\gamma + 1} COV(\tilde{\mathbf{r}}\tilde{\mathbf{X}}_t, \eta_{i,t}|\mathbf{z}_i^t) \end{aligned}$$

where $\mathbf{F}_t \equiv \sum_{j=1}^{\infty} (1 - \lambda)^{j-1} \lambda \mathbf{X}_t^{(j)}$, $\mathbf{X}_t \equiv [m_t, m_{t-1}, a_t]'$, and $\tilde{\mathbf{r}} \equiv [1, 0, -1, 1, 0, -1, 0]$. Let us define the matrix $\tilde{\mathbf{P}}_{t|t}$ as

$$\tilde{\mathbf{P}}_{t|t} \equiv \tilde{\mathbf{P}}_{t|t-1} - \tilde{\mathbf{P}}_{t|t-1} \tilde{\mathbf{D}}' \left[\tilde{\mathbf{D}} \tilde{\mathbf{P}}_{t|t-1} \tilde{\mathbf{D}}' + \tilde{\Sigma}_e \right]^{-1} \tilde{\mathbf{D}} \tilde{\mathbf{P}}_{t|t-1} \quad (45)$$

where the matrix $\tilde{\mathbf{P}}_{t|t-1}$ is the matrix \mathbf{P} in equation (41) that is obtained by solving the fixed point as discussed in Appendix D. The matrices $\tilde{\mathbf{D}}$ and $\tilde{\Sigma}_e$ have been defined in (38) and in (39), respectively. The variances and covariances can be obtained from the matrix $\tilde{\mathbf{P}}_{t|t}$:

$$\begin{aligned} VAR(\tilde{\mathbf{r}}\tilde{\mathbf{X}}_t|\mathbf{z}_i^t) &\equiv \tilde{\mathbf{r}}\tilde{\mathbf{P}}_{t|t}\tilde{\mathbf{r}}' \\ VAR(m_t|\mathbf{z}_i^t) &\equiv [1, \mathbf{0}_{1 \times 6}] \tilde{\mathbf{P}}_{t|t} [1, \mathbf{0}_{1 \times 6}]' \\ VAR(a_t|\mathbf{z}_i^t) &\equiv [\mathbf{0}_{1 \times 2}, 1, \mathbf{0}_{1 \times 4}] \tilde{\mathbf{P}}_{t|t} [\mathbf{0}_{1 \times 2}, 1, \mathbf{0}_{1 \times 4}]' \\ VAR(\eta_{i,t}|\mathbf{z}_i^t) &= [\mathbf{0}_{1 \times 6}, 1] \tilde{\mathbf{P}}_{t|t} [\mathbf{0}_{1 \times 6}, 1]' \\ COV(\tilde{\mathbf{r}}\tilde{\mathbf{X}}_t, m_t|\mathbf{z}_i^t) &= \tilde{\mathbf{r}}\tilde{\mathbf{P}}_{t|t} [1, \mathbf{0}_{1 \times 6}]' \\ COV(\tilde{\mathbf{r}}\tilde{\mathbf{X}}_t, a_t|\mathbf{z}_i^t) &= \tilde{\mathbf{r}}\tilde{\mathbf{P}}_{t|t} [\mathbf{0}_{1 \times 2}, 1, \mathbf{0}_{1 \times 4}]' \\ COV(\tilde{\mathbf{r}}\tilde{\mathbf{X}}_t, \eta_{i,t}|\mathbf{z}_i^t) &= \tilde{\mathbf{r}}\tilde{\mathbf{P}}_{t|t} [\mathbf{0}_{1 \times 6}, 1]' \end{aligned}$$

F Robustness Checks

The Marginal Data Density (MDD) associated with a model is defined as the integral of the likelihood with respect to the prior over the parameter space. Therefore, it can be regarded as a measure of conflicting information in the likelihood and in the prior. As the variance of the prior changes, the MDD also varies. Therefore, one may be concerned about the robustness of the figures in Table 3.

To address this concern, I have computed the marginal data density (MDD) of the two structural models using the following prior specifications. The first prior is more dogmatic than that in the paper. The second prior is more loose than that in the paper.³ Finally, the third prior is more

³Note that the prior variance for the standard deviation of monetary shocks σ_m cannot be further reduced

Name	Type	<i>Dogmatic Prior</i>			<i>Loose Prior</i>			<i>Dogmatic Prior for λ</i>		
		Median	5%	95%	Median	5%	95%	Median	5%	95%
ρ_m	\mathcal{B}	0.50	0.33	0.66	0.50	0.08	1.00	0.50	0.17	0.82
$100\sigma_m$	\mathcal{IG}	2.00	1.09	2.36	2.00	0.44	12.77	2.00	0.43	12.81
$100\sigma_a$	\mathcal{IG}	0.70	0.56	0.81	0.70	0.42	1.00	0.70	0.51	0.87
$100\tilde{\sigma}_m$	\mathcal{IG}	5.01	2.73	5.92	5.01	1.58	12.42	5.01	2.12	7.91
$100\tilde{\sigma}_a$	\mathcal{IG}	1.06	0.41	0.89	1.06	0.17	4.80	1.06	0.24	1.87
λ	\mathcal{B}	0.41	0.33	0.49	0.41	0.08	0.73	0.41	0.40	0.42
$100\mu_M$	\mathcal{N}	0.00	-8.20	8.20	0.00	-82.20	82.20	0.00	-41.00	41.00
$100\mu_A$	\mathcal{N}	0.00	-8.20	8.20	0.00	-82.20	82.20	0.00	-41.00	41.00
$100\kappa_{pc}$	\mathcal{G}	12.00	0.02	25.27	12.00	0.00	35.86	12.00	0.00	22.00
ω	\mathcal{B}	0.50	0.33	0.63	0.50	0.00	1.00	0.50	0.08	1.00

Table 1: Prior statistics for the parameters of the ICKM and the Calvo model. Results based on 1,000,000 posterior draws using the Metropolis-Hastings algorithm.

dogmatic only with respect to λ , which affects the strategic complementarity in price setting.

In addition, I consider two other prior specifications: (1) *Dogmatic Prior #2*, whose second moments are the same as those of the *Dogmatic Prior* of Table 1 except for those for the parameters that are specific to the Calvo model (i.e., κ_{pc} and ω). The prior variance for those parameters is the same as that in the paper. (2) *Loose Prior #2*, whose second moments are the same as those of the *Loose Prior* of Table 1 except for those for the parameters that are specific to the Calvo model (i.e., κ_{pc} and ω). The prior variance for those parameters is the same as that in the paper. In a nutshell, *Dogmatic Prior #2* is more dogmatic than the prior in the paper but only regarding the ICKM parameters. Analogously, the *Loose Prior #2* is broader than the prior used in the paper, but only for the ICKM parameters. Table 1 summarizes the prior and posterior statistics associated with the prior distributions used in the exercise.

The MDD associated with the two structural models under all the prior specifications are reported in Table 2. The MDD associated with the ICKM is always larger than the Calvo model under all prior specifications. This means that the ICKM attains a larger posterior probability than the Calvo model under all the prior specifications considered. This result shows the robustness of the finding in the paper to changes in the prior variance. Furthermore, the results in Table 2 suggest that the MDD associated with the Calvo model is very sensitive to the prior variance for its specific parameters (i.e., κ_{pc} and ω), while the MDD associated with the ICKM seems to be fairly insensitive to the prior specification.

G Endogenous Monetary Policy Rule

The central bank sets the stock of money M_t following the rule below:

$$\ln M_t = \phi_p \ln P_t + \ln M_t^*, \quad (46)$$

where P_t is the price level, $\phi_p < 0$ is a parameter that controls how aggressively the central bank adjusts the stock of money to changes in the price level, and $\ln M_t^*$ is the exogenous component

because the prior used in the paper is an inverted gamma distribution whose degree of freedom is 1.

Log Marginal Data Density		
	ICKM	Calvo
<i>Prior Table 1</i>	1737.88	1706.02
<i>Dogmatic Prior</i>	1732.60	1675.08
<i>Loose Prior</i>	1735.44	1685.60
<i>Dogmatic Prior for λ</i>	1735.62	1706.02
<i>Dogmatic Prior #2</i>	1732.60	1715.14
<i>Loose Prior #2</i>	1735.44	1715.83

Table 2: The table reports the log-marginal likelihood for the ICKM and the Calvo model, computed using Geweke’s (1999) harmonic mean estimator.

of the rule that follows a Gaussian process. To simplify the model, assume that there is no firm-specific technology shock (i.e., $\sigma_\eta = 0$). Price setters observe a vector of signals $\mathbf{z}_{i,t} \equiv [z_{m,i,t}, z_{a,i,t}]'$ that is driven by the following observation equations:

$$\mathbf{z}_{i,t} = \begin{bmatrix} \ln M_t^* \\ \ln A_t \end{bmatrix} + \begin{bmatrix} \tilde{\sigma}_m & 0 \\ 0 & \tilde{\sigma}_a \end{bmatrix} \mathbf{e}_{i,t}. \quad (47)$$

The rest of the model is the same as the ICKM presented in Section I of the main text.

The equilibrium conditions in the money market implies that $\ln M_t = \ln P_t + \ln Y_t$. Substituting the money rule into this equilibrium condition yields:

$$(1 - \phi_p) \ln P_t + \ln Y_t = \ln M_t^*. \quad (48)$$

Note that $(1 - \phi_p) \geq 1$. When $\phi_p = 0$, one is back to the situation of the paper in which the central bank does not directly react to inflation and *successfully* controls the nominal output. See equation (9). When the central bank sets its monetary instrument to stabilize the price level ($\phi_p < 0$), then one obtains that a linear combination of price level and real output follows an exogenous process. The more aggressively the central bank targets price stability, that is, the larger $(1 - \phi_p)$ in equation (48), the more the price level responds to shocks.

The optimal price-setting equation can be derived following the steps in Appendix A and reads

$$\ln P_{i,t} = \mathbb{E}_{i,t} [\lambda \ln Y_t + \ln P_t - \lambda \ln A_t] - \lambda \ln \bar{y}. \quad (49)$$

The equation (48) allows one to substitute $\ln Y_t$ out from equation (49):

$$\ln P_{i,t} = \mathbb{E}_{i,t} \{ [1 - \lambda(1 - \phi_p)] \ln P_t + \lambda \ln M_t^* - \lambda \ln A_t \} - \lambda \ln \bar{y}. \quad (50)$$

This price-setting equation shows that the parameter $(1 - \lambda(1 - \phi)) \equiv 1 - (1 - \phi)(\gamma + 1)\phi^{-1} / (\nu(\phi^{-1} - 1) + 1)$ controls the strategic complementarity in price-setting (i.e., the extent to which firms want to react

to the expected average price $\mathbb{E}_{i,t}(P_t)$). Note that the more aggressively the central bank respond to inflation, the smaller the degree of strategic complementarity in price setting. By integrating across firms one obtains:

$$\ln P_t = [1 - \lambda(1 - \phi_p)] \ln P_{t|t}^{(1)} + \lambda \ln M_t^{*(1)} - \lambda \ln A_{t|t}^{(1)} - \lambda \ln \bar{y}, \quad (51)$$

where the fact that $\int \mathbb{E}_{i,t}(\eta_{i,t}) di = 0$ is used. From this equation, repeatedly taking the conditional expectation and averaging across firms yields:

$$\ln P_{t|t}^{(j)} = [1 - \lambda(1 - \phi_p)] \ln P_{t|t}^{(j+1)} + \lambda \ln M_t^{*(j+1)} - \lambda \ln A_{t|t}^{(j+1)} - \lambda \ln \bar{y} \quad (52)$$

for $j \in \{1, 2, \dots\}$. By repeatedly substituting these results into the average-price equation one obtains:

$$\ln P_t = \sum_{j=0}^{\infty} [1 - \lambda(1 - \phi_p)]^j \lambda \left(\ln M_t^{*(j+1)} - \ln A_{t|t}^{(j+1)} \right) - \lambda \ln \bar{y} \quad (53)$$

Define $m_t \equiv \ln M_t^* - \mu_M t$ and $a_t \equiv \ln A_t - \mu_A t$ and rewrite equation (53) as:

$$(1 - \phi_p) \ln P_t = \sum_{j=0}^{\infty} \left(1 - \tilde{\lambda}\right)^j \tilde{\lambda} \left(m_{t|t}^{(j+1)} - a_{t|t}^{(j+1)} \right) - \lambda \ln \bar{y} + \mu_M \cdot t - \mu_A \cdot t \quad (54)$$

where we denote $\tilde{\lambda} \equiv \lambda(1 - \phi_p)$, the average j -th order expectations about the state of monetary policy are defined as $m_{t|t}^{(j)} \equiv \int m_{t|t}^{(j)}(i) di$, where $m_{t|t}^{(j)}(i) \equiv \mathbb{E} \left[m_{t|t}^{(j-1)} | \mathbf{z}_i^t \right]$. Average expectations $a_{t|t}^{(j)}$ about technology are analogously defined.

How does real output respond to a monetary shock in the ICKM where the central bank sets the money stock to respond to price fluctuations? Plugging equation (54) into the equilibrium condition (48) and recalling that $m_t \equiv \ln M_t^* - \mu_M t$, one obtains the following:

$$\ln Y_t = \left[m_t - \sum_{j=0}^{\infty} \left(1 - \tilde{\lambda}\right)^j \tilde{\lambda} m_{t|t}^{(j+1)} \right] + \sum_{j=0}^{\infty} \left(1 - \tilde{\lambda}\right)^j \tilde{\lambda} a_{t|t}^{(j+1)} + \lambda \ln \bar{y} + \mu_A \cdot t \quad (55)$$

First, note that the terms outside the square brackets are not affected by the monetary shocks. Second, note that the summation operator can be regarded as a weighted average of higher-order expectations as long as $(1 - \tilde{\lambda}) \in (0, 1)$. Hence, as shown in Section I.G, a monetary shock produces real effects as long as all average higher-order expectations $m_{t|t}^{(j)}$, any $j \geq 1$, do not fully converge to the stock of money m_t . Yet, unlike in Section I.G, the weights $(1 - \tilde{\lambda})^j \tilde{\lambda}$ are now affected by the policy parameter ϕ_p . It follows that the stronger the central bank's response to price $(1 - \phi_p)$, the larger the weights associated with expectations of relatively lower orders are. Since expectations of lower order are relatively more rapid to adjust (see left plot of Figure 1), the reshuffling of weights due to the feedback in the monetary rule implies that the weighted average $\sum_{j=0}^{\infty} (1 - \tilde{\lambda})^j \tilde{\lambda} m_{t|t}^{(j+1)}$ will take less time to converge to m_t and, hence, have shorter lasting real effects of money.