A Likelihood Analysis of Models with Information Frictions*

Leonardo Melosi†

London Business School

June 22, 2010

Abstract

The paper estimates a dynamic stochastic general equilibrium model where incomplete information is modeled as in Woodford (2002). I provide a formal econometric analysis of the model, revealing that it accounts well for the highly persistent propagation of monetary policy shocks observed in the data. I use the theory of rational inattention proposed by Sims (2003) to argue that incomplete-information models à la Woodford (2002) can explain such a persistence only if firms acquire implausibly too little information about monetary policy.

Keywords: Imperfect common knowledge; forecasting the forecasts of others; rational inattention; Bayesian econometrics; persistent real effects of nominal shocks.

JEL classification: E3, E5, C32, D8.

---

*I am indebted to Frank Schorfheide, Frank Diebold, Jesús Fernández-Villaverde, and Dirk Krueger for the valuable advice they have provided at each stage of this paper. I am also very grateful to Harold Cole for very helpful comments and guidance. I thank Dario Caldara, Matthieu Darracq Pariès, Marco Del Negro, Cristina Fuentes-Albero, Christian Hellwig, Max Kryshko, Bartosz Maćkowiak, Guido Menzio, Kristoffer Nimark, Ricardo Reis, and Mirko Wiederholt for very helpful comments. I would also like to thank participants at the Midwest Macro Meetings 2008, the Annual Congress of the European Economic Association 2008, the PENN Macro and Metric Lunches, 2009 North American Summer Meeting of the Econometric Society, SED Annual Meeting 2009, and the 2009 European Meeting of the Econometric Society, CEMMAP-UCL Applied Macroeconomics and Macroeconometrics Workshop. I also wish to thank the seminar participants at UPF, UCL, LBS, UCLA, Federal Reserve Bank of Richmond, Federal Reserve Bank of Dallas, IMF, Federal Reserve Board of Governors, Bocconi, EIEF, Banca d'Italia, LSE, and Boston University.

†Correspondence to: Leonardo Melosi, London Business School, Regent's Park, London, NW1 4S, United ingdom. Email: lmelosi@london.edu.

1 Introduction

A number of influential empirical studies of the U.S. economy have documented that monetary disturbances have highly persistent real effects and delayed impacts on inflation (Christiano, Eichenbaum, and Evans, 1999, Stock and Watson, 2001). This paper performs an econometric analysis to assess whether a model where firms have incomplete information in the sense of Woodford (2002) can account for such empirical evidence. The model has two exogenous state variables: the state of monetary policy (i.e., the stock of money) and the state of technology (i.e., total factor productivity). Price-setting firms observe one idiosyncratic noisy signal about each exogenous state variable. Moreover, they face strategic complementarities in price-setting. In this environment, firms find it optimal to set their prices by reacting not only to their beliefs about the exogenous state variables but also to their higher-order beliefs - i.e., their beliefs about other firms' beliefs, their beliefs about other firms' beliefs about other firms' beliefs, and so on (Townsend, 1983a, 1983b). Although firms do not face any cost of price adjustment, monetary disturbances have real effects in this model as a result of the inertial response of firms' beliefs. Furthermore, the degree of strategic complementarities amplifies the persistence in the propagation of monetary disturbances through the sluggish adjustment of firms' higher-order beliefs. By following the terminology introduced by Woodford (2002), I call this model the imperfect-common-knowledge model (ICKM).

This paper addresses the following question: can the ICKM account for the persistent effects of monetary shocks observed in the data? The answer to this question is yes. To get this answer, I estimate the ICKM and a vector autoregressive model (VAR) through Bayesian methods. I show that from a Bayesian perspective, the impulse response functions implied by the VAR can be considered an accurate description of the propagation of monetary shocks in the data. I find that the estimated ICKM successfully captures the sluggish and hump-shaped response of output and inflation to monetary shocks implied by the VAR. Furthermore, the paper compares the imperfect-common-knowledge mechanism of generating price inertia with a more popular mechanism based on Calvo sticky prices (Calvo, 1983). To this end, I consider a model (henceforth, Calvo model) that differs from the ICKM in only two respects: (1) firms are perfectly informed and (2) firms can re-optimize their prices only at random periods, as in Calvo (1983). I conduct a Bayesian comparison of the ICKM and the Calvo model and find that the imperfect-common-knowledge mechanism is better suited to explaining the highly persistent effects of monetary shocks observed in the data.

Finally, the paper emphasizes that the signal-to-noise ratio of monetary policy is smaller than that of technology by a factor of six in the estimated ICKM, implying that firms acquire little information about the state of monetary policy. This finding raises the following question: is it plausible that firms acquire so little information about monetary policy? The answer to this question is no. I reach this conclusion by constructing a simple rational inattention model (RIM) that is nested within the ICKM. While in the ICKM the signal-to-noise ratios are taken as given by firms, in the RIM these ratios are optimally chosen by firms, subject to an information-processing constraint as in Sims (2003) and Maćkowiak and Wiederholt (2009). This constraints sets an upper-bound to the overall precision of signals. I calibrate the RIM by using the estimated parameters of the ICKM and solve for the optimal signal-to-noise ratios. Two main results emerge. First, unlike in the ICKM, firms find it optimal to acquire more information about monetary policy than about technology. Second, I find that the marginal value of information about monetary policy is much smaller than that in the ICKM. These results suggest that the signal-to-noise ratio relative to monetary policy is implausibly small in the ICKM. Moreover, when I calibrate the ICKM by using plausible signal-to-noise ratios (i.e., the optimal signal-to-noise ratios predicted by the RIM), the model fails to generate the persistent propagation of monetary disturbances observed in the data. I argue that this last finding indicates that the ICKM can match the persistence observed in the data only if firms acquire implausibly too little information about monetary policy.

Claiming that firms have incomplete information about the state of monetary policy in the ICKM or in the RIM is by no means a statement about the availability of information on monetary policy. In fact, information about the state of monetary policy and any other model variable is assumed to be publicly available to every firm in these models. Information frictions arise from the assumption that firms cannot attend to all available information because they have limited capability of processing information (Woodford, 2002, Sims, 2003). Therefore, firms do not acquire all available information about the state of monetary policy, even though they can get information on the stock of money and the interest rate at no cost.

This paper departs from Woodford (2002) in the following respects. My empirical strategy is likelihood-based, while Woodford (2002) calibrates the parameters of his model. This is the first paper that obtains likelihood-based estimates for the parameters of an ICKM à la Woodford (2002). This empirical approach is motivated by the aim of countering the lack of empirical guidance in selecting the variance of signal noise, which strongly influences the persistence in this type of model. Furthermore, Woodford (2002) closes his model by specifying an exogenous stochastic process that drives the nominal output. I develop a micro-founded demand side of the economy so that the resulting general equilibrium model is isomorphic to Woodford's partial equilibrium model. This feature is desirable as it allows me to apply the same method as that in Woodford (2002) to solve the ICKM. This solution method is fast and robust. Hence, I can evaluate the likelihood at several points of the parameter space and get accurate estimates of parameters. Finally, Woodford's model has one rather than two shocks. Having an additional shock allows me to get around the problem of stochastic singularity when I evaluate the likelihood.

This paper is related to the literature of rational inattention that was pioneered by Sims (1998, 2003, 2006). Rational inattention provides a positive theory about how much information agents acquire on each variable in a given model, subject to an information-processing constraint. The fascinating paper by Maćkowiak and Wiederholt (2009) lays down a tractable framework for integrating the theory of rational inattention into a simple linear-quadratic macroeconomic model with constant velocity of money. In recent years, scholars have applied this theory to several fields in macroeconomics (Luo, 2008, Paciello, 2008, Van Nieuwerburgh and Veldkamp, 2009, Woodford, 2008, Maćkowiak, Moench, and Wiederholt 2009, and Maćkowiak and Wiederholt, 2010).

The paper is also related to the literature that uses incomplete-information models for studying the persistence in economic fluctuations (Townsend, 1983a, 1983b; Hellwig, 2002; Adam, 2009; Angeletos and La'O, 2009; Rondina, 2008; and Lorenzoni, 2009) and the propagation of monetary disturbances to real variables and prices (Phelps, 1970; Lucas, 1972; Woodford, 2002; Adam, 2007; Gorodnichenko, 2008; Nimark, 2008; and Lorenzoni, 2010).[1].

The remainder of the paper proceeds as follows. Section 2 presents the ICKM and

[1]See Mankiw and Reis (2002a, 2002b, 2006, 2007), and Reis (2006a, 2006b, 2009) for models with information frictions that do not feature imperfect common knowledge but can generate sizeable persistence.

the Calvo model. Section 3 conducts Bayesian estimation and evaluation of these two models. In Section 4, I evaluate the plausibility of the estimates by using the RIM that is nested within the ICKM. Section 5 concludes.

2 The Model Economy

This section is organized as follows. In Section 2.1, I introduce the maintained assumptions of the ICKM. In Sections 2.2-2.6, I show the problems of agents in the ICKM. In Section 2.7, I briefly explain how I detrend and log-linearize the model around the deterministic steady state equilibrium. Section 2.8 briefly describes how to solve the ICKM. In Section 2.9, I introduce the Calvo model.

2.1 Maintained Assumptions

The economy is populated by perfectly competitive final-good producers (or, more briefly, producers), households, a financial intermediary, a monetary authority (or central bank), and a continuum $(0, 1)$ of intermediate-good firms (or, more briefly, firms). The model has two exogenous state variables: the state of monetary policy (i.e., the stock of money) and the state of technology (i.e., total factor productivity). Monetary shocks affect the stock of money that the central bank injects into the economy.

The information structure of the model can be summarized as follows. First, all information is publicly available to every agent. Second, firms cannot attend perfectly to all available information. This last feature is modeled as suggested by Woodford (2002): firms do not observe past and current realizations of any model variables and solely observe signals about the two exogenous state variables. For tractability, it is assumed that the other agents (i.e., final-good producers, households, the financial intermediary, and the monetary authority) perfectly observe the past and the current realizations of all the model variables.

At the beginning of period t, the households inherit the entire money stock of the economy, M_{t-1}. Monetary shocks, technology shocks, and signals are realized. Households decide how much money D_t to deposit at the financial intermediary after observing current-period innovations to technology and monetary shocks. These deposits yield interest at a rate of $R_t - 1$. The financial intermediary receives households' deposits and a monetary injection from the monetary authority, which it lends to firms at a fixed fee τ. The firms observe their signals, set their prices, hire labor service from households, and then produce. They use the liquidity facilities provided by the financial intermediary at the fixed fee τ to pay wages $W_t H_t$, where W_t is the nominal hourly wage, and H_t is hours worked. Households' cash balance increases to $M_{t-1} - D_t + W_t H_t$. Households face a cash-in-advance (CIA) constraint, requiring that households pay for all consumption purchases with the accumulated cash balances. Firms sell their goods to producers that integrate them into a final good that they sell to households. Firms also pay back their loans, $L_{i,t}$. Finally, households receive back their deposits inclusive of interest and dividends from both firms, Π_t, and the financial intermediary, Π_t^b.

2.2 Final-Good Producers

The representative final-good producer combines a continuum of intermediate goods, $Y_{i,t}$, by using the technology:

$$Y_t = \left(\int_0^1 (Y_{i,t})^{\frac{\nu-1}{\nu}} \, di \right)^{\frac{\nu}{\nu-1}} \tag{1}$$

where Y_t is the amount of the final good produced at time t, the parameter ν represents the elasticity of demand for each intermediate good and is assumed to be strictly larger than one. The producer takes the input prices, $P_{i,t}$, and output price, P_t, as given. Profit maximization implies that the demand for intermediate goods is:

$$Y_{i,t} = \left(\frac{P_{i,t}}{P_t} \right)^{-\nu} Y_t \tag{2}$$

where the competitive price of the final good, P_t, is given by

$$P_t = \left(\int (P_{i,t})^{1-\nu} \, di \right)^{\frac{1}{1-\nu}}. \tag{3}$$

2.3 The Representative Household

The representative household derives utility from consuming the final good, C_t, and disutility from hours worked, H_t, and maximizes

$$\max_{\{C_t,H_t,M_t\geq 0,D_t\geq 0\}} \mathbb{E}_t \sum_{s=0}^{\infty} \beta^s \left[\ln C_{t+s} - \alpha \frac{H_{t+s}^{1+\eta}}{1+\eta} \right] \tag{4}$$

such that

$$P_t C_t \leq M_{t-1} - D_t + W_t H_t \tag{5}$$

$$M_t = (M_{t-1} + W_t H_t - D_t - P_t C_t) + R_t D_t + \Pi_t + \Pi_t^b \tag{6}$$

where β is the discount factor, α is a parameter that affects the marginal utility of leisure, and $\eta > 0$ is the Frisch labor elasticity.

The first constraint is the CIA constraint, requiring that households hold money up-front to finance their consumption.[2] The second constraint is the law of motion of households' cash. It is also assumed that households cannot borrow from the financial intermediary.

2.4 The Financial Intermediary

The financial intermediary solves the trivial static problem in every period t:

$$\max_{\{L_t,D_t\}} (1 - R_t) D_t + X_t + \tau \cdot \mathbb{I}\{L_t > 0\} \tag{7}$$

such that

$$L_t \leq X_t + D_t \tag{8}$$

where L_t is the aggregate amount of liquidity supplied to firms, $X_t = M_{t+1} - M_t$ is the monetary injection, $\mathbb{I}\{\cdot\}$ is an indicator function that equals one if the statement within curly brackets is true. τ is a fixed fee that the intermediary receives from firms.[3] The equilibrium in the market for loans ensures that $L_t \equiv \int L_{i,t} di$, where $L_{i,t}$

[2] I assume that households can use their current nominal labor income to finance current consumption purchases. Including the nominal labor income in the CIA constraints prevents labor supply from depending on forward-looking variables and hence simplifies the signal-extraction problem faced by the intermediate good firms (see Section 2.6). This specification of the CIA constraint has been widely used in the literature of limited participation. See Christiano (1991), Christiano and Eichenbaum (1992), and Nason and Cogley (1994).

[3] As in the limited participation literature (Christiano, 1991, Christiano and Eichenbaum, 1992, and Nason and Cogley, 1994), the financial intermediary lends cash to firms so that they can pay wages before households consume. However, in this literature the intermediary receives an interest

stands for the liquidity borrowed by firm i at time t.

2.5 The Monetary Authority

The monetary authority sets the money stock M_t to grow at rate

$$\Delta \ln M_t = (1 - \rho_m) M_0 + \rho_m \Delta \ln M_{t-1} + \sigma_m \varepsilon_{m,t} \tag{9}$$

with $\varepsilon_{m,t} \backsim \mathcal{N}(0,1)$ and where Δ stands for the first-difference operator and the degree of smoothness in conducting monetary policy ρ_m is such that $\rho_m \in [0,1)$. M_0 is a parameter that represents the long-run average growth rate of money. Equation (9) can be interpreted as an empirical monetary policy rule without feedbacks. The monetary shock $\varepsilon_{m,t}$ captures unexpected changes in the growth rate of money in every period t. Market clearing for the monetary market requires that:

$$\ln M_t = \ln Y_t + \ln P_t \tag{10}$$

2.6 Intermediate-Good Firms

The expected firm i's period-t profit (as valued by households) conditional on the history of signals observed by that firm at time t, $\mathbf{z}_i^t$, is given by

$$\mathbb{E}\left[\beta^t Q_t \left(P_{i,t} Y_{i,t} - W_t N_{i,t} - \tau \mathbb{I}\left\{L_{i,t}\right\}\right) | \mathbf{z}_i^t\right] \tag{11}$$

where Q_t is the time 0 value of one unit of the final good in period t to the representative household. $Y_{i,t}$ is the amount of intermediate goods i demanded by the final-good producers at time t. $N_{i,t}$ is the labor input demanded by firm i at time t. The production function is

$$Y_{i,t} = A_t N_{i,t}^{\phi} \tag{12}$$

where $\phi \in (0,1)$ and A_t is the level of technology that follows an exogenous process:

$$\ln A_t = A_0 + \ln A_{t-1} + \sigma_a \varepsilon_{a,t} \tag{13}$$

---

rate on loans. Replacing the fixed fee τ with an equilibrium interest rate would introduce forward-looking variables in the problem of firms and would substantially complicate their signal-extraction problem.

$\varepsilon_{a,t} \backsim \mathcal{N}(0,1)$. The technology shocks, $\varepsilon_{a,t}$, are assumed to be orthogonal to monetary shocks, $\varepsilon_{m,t}$, at all leads and lags. Firms borrow liquidity at the fixed cost τ from the financial intermediary to pay their nominal labor costs:

$$L_{i,t} = W_t N_{i,t} \tag{14}$$

Firms are charged with a fixed fee τ for this service. Similar to Woodford (2002), firm i's signals are defined as:

$$\mathbf{z}_{i,t} = \begin{bmatrix} m_t \\ a_t \end{bmatrix} + \begin{bmatrix} \tilde{\sigma}_m & 0 \\ 0 & \tilde{\sigma}_a \end{bmatrix} \mathbf{e}_{i,t} \tag{15}$$

where $\mathbf{z}_{i,t} \equiv [z_{m,i,t}, z_{a,i,t}]'$, $m_t \equiv \ln M_t - M_0 \cdot t$, $a_t \equiv \ln A_t - A_0 \cdot t$, $\mathbf{e}_{i,t} \equiv [e_{m,i,t}, e_{a,i,t}]'$ and $\mathbf{e}_{i,t} \overset{iid}{\backsim} \mathcal{N}(\mathbf{0}, \mathbb{I}_2)$. Note that a_t and m_t represent the exogenous state variables of the model and the signal noises $e_{m,i,t}$ and $e_{a,i,t}$ are assumed to be *iid* across firms and time. Furthermore, I assume that the two signals are orthogonal. This may be considered a strong assumption. For instance, it is reasonable to think that firms might learn something about the state of monetary policy, m_t, from observing the signal concerning the state of technology $z_{a,i,t}$. Nonetheless, I find that relaxing the assumption of orthogonality of signals would not change the main results of the paper.

In every period t, firms observe the history of their signals, $\mathbf{z}_i^t$, and choose their prices, $P_{i,t}$, so as to maximize the objective function (11) subject to equations (2) and (12)-(15) and by taking the stochastic discount factor, Q_t, and the nominal wage, W_t, as exogenous. The equilibrium laws of motion of all model variables are assumed to be common knowledge among firms.

As I shall explain in more details in Section 2.7, I log-linearize the price-setting equation around the deterministic steady state so that the transition equation of average price is linear. Working with a log-linear model enhances the tractability of firms' signal extraction problem, which can thus be solved analytically through the Kalman filter. Furthermore, it is important to emphasize that I assume that firms are endowed with an infinite sequence of signals at time 0, that is, $\mathbf{z}_i^t = \{\mathbf{z}_{i,\tau}\}_{\tau=-\infty}^t$. This assumption implies that all firms have the same Kalman gain matrix in their signal-extraction problem. Furthermore, this matrix can be shown to be time-invariant. These properties of the Kalman gain matrix make the task of solving the model easier.

2.7 Detrending, Log-Linear Approximation

The exogenous processes (9) and (13) induce both a deterministic and a stochastic trend to all endogenous variables, except labor. I will detrend the non-stationary variables before log-linearizing the models. It is useful to define the stationary variables as:

$$y_t \equiv \frac{Y_t}{A_t}, \quad y_{i,t} \equiv \frac{Y_{i,t}}{A_t}, \quad p_{i,t} \equiv \frac{P_{i,t}}{P_t} \tag{16}$$

In order to log-linearize the model, I take the following steps. First, I derive the price-setting equation by solving the problem of firms described in Section 2.6. Second, I transform the variables in the price-setting equation according to the definitions (16). Third, I log-linearize the resulting price-setting equation around the deterministic steady state. Fourth, I aggregate the log-linearized price-setting equation across firms and obtain the law of motion of price level. Fifth, the law of motion of real output can be easily obtained from combining the law of motion of price level and equation (10).

2.8 Model Solution

When one characterizes rational expectation equilibria (REE) in models where agents have private information and react to endogenous variables, a typical challenge is dealing with an infinite-dimensional state vector (also known as infinite regress)[4] (Townsend, 1983b). The reason is that the laws of motion of average expectations of infinitely many orders have to be characterized in order to solve these models. This task is clearly unmanageable.

In the ICKM, this problem arises when there is strategic complementarity in price setting. Yet the issue can be elegantly resolved as in Woodford (2002), since it is possible to write the state vector of the model in terms of weighted averages of infinitely many average higher-order expectations.[5] This leads to a state space of

[4] See Nimark (2009) for a thorough explanation of this challenge.

[5] Different methods have been developed to solve dynamic models with incomplete information. Following Townsend (1983b), the customary approach of solving this class of models is to assume that the realizations of states at some arbitrary distant point in the past are perfectly revealed. Rondina and Walker (2009) have challenged this approach by showing that such a truncation reveals the entire history of the realizations of states to agents, regardless of the point of truncation. See Nimark (2008) for a truncation-based method that preserves the recursive structure.

very small dimension.[6] The solution method turns out to be so fast and robust that I can evaluate the likelihood at several points of the parameter space. This leads to accurate estimates of model parameters.

2.9 The Calvo Model

In the Calvo model all agents (i.e., final-good producers, households, the financial intermediary, the monetary authority, the intermediate-good firms) perfectly observe the past and current realizations of the model variables. Moreover, the price charged by each intermediate-good firm is re-optimized only at random periods. The key (simplifying) assumption is that the probability that a given firm will optimally adjust its price within a particular period is independent of the state of the model, the current price charged, and how long ago it was last re-optimized. Specifically, only a fraction $(1-\theta_p)$ of firms re-optimize their prices, while the remaining θ_p fraction adjusts them to the steady-state inflation π_*. The problem of the firms that are allowed to re-optimize their prices at time t is:

$$\max_{P_{i,t}} \mathbb{E}_t \sum_{s=0}^{\infty} \left[\theta_p^s \beta^{t+s} Q_{t+s|t} \left(\pi_*^s P_{i,t} - MC_{t+s}\right) Y_{i,t+s} - \tau \mathbb{I}\left\{L_{i,t} > 0\right\}\right] \tag{17}$$

such that

$$Y_{i,t+s} = \left(\frac{\pi_*^s P_{i,t}}{P_{t+s}}\right)^{-\nu} Y_{t+s} \tag{18}$$

where $Q_{t+s|t}$ is the time t value of one unit of the final good in period $t+s$ to the representative household, π_* is the steady-state (gross) inflation rate, and MC_{t+s} stands for the nominal marginal costs in period $t+s$. The price level is given by:

$$P_t^{1-\nu} = \left[\left(1-\theta_p\right) P_t^{*(1-\nu)} + \theta_p \ \left(\pi_* P_{t-1}\right)^{1-\nu}\right] \tag{19}$$

I detrend the non-stationary variables and log-linearize the model around the deterministic steady state. I obtain the standard New Keynesian Phillips curve, whose slope, κ_{pc}, depends on a function of parameters: $\kappa_{pc} = \left(1-\theta_p\right)\left(1-\theta_p\beta\right)\lambda/\theta_p$, where λ is the strategic-complementarity parameter. The strategic-complementarity parameter, λ, is a function of the Frisch labor-supply elasticity, η, the technology

[6] A detailed description of the method that numerically solves the model is described in appendix B.

parameter, ϕ, and the demand elasticity, ν.

3 Empirical Analysis

This section contains the econometric analysis of the ICKM and the Calvo model. I combine a prior distribution with the likelihood function derived from the models and conduct Bayesian inference. In Section 3.1, I present the data set. Section 3.2 introduces the measurement equations. In Section 3.3, I discuss the prior distribution for model parameters. Section 3.4 presents the posterior distribution. In Section 3.5, I introduce a VAR model and formally show that this statistical model can be used as a valid benchmark for Bayesian model evaluation. In Section 3.6, the impulse response functions (IRFs) of real GDP and inflation to a monetary shock are analyzed for the ICKM, the Calvo model, and the VAR.

3.1 Data

The data are quarterly and range from the third quarter of 1954 to the fourth quarter of 2005. I use the U.S. per capita real GDP and the U.S. GDP deflator from Haver Analytics (Haver mnemonics are in italics). Per capita real GDP is obtained by dividing the nominal GDP (*GDP*) by the population 16 years and older (*LN16N*) and deflating using the chained-price GDP deflator (*JGDP*). The GDP deflator is given by the appropriate series (*JGDP*).

3.2 Measurement Equations

The data on the U.S. GDP deflator and on the U.S. per capita real GDP are denoted as $\{P_t,\ t = 1, 2, ...T\}$ and $\{Y_t,\ t = 1, 2, ...T\}$, respectively. In the log-linearized ICKM, the measurement equation for the GDP price deflator is[7]

$$\ln P_t = \left[\sum_{j=0}^{\infty} (1-\lambda)^j \lambda \left(m_{t|t}^{(j+1)} - a_{t|t}^{(j+1)} \right) \right] - \ln \bar{y} + M_0 \cdot t - A_0 \cdot t \tag{20}$$

where $\lambda \equiv (\eta + 1)\, \phi^{-1} / \left[\nu \left(\phi^{-1} - 1 \right) + 1 \right]$, and $\bar{y}$ is the steady-state value of the detrended real output, y_t. The variables $m_{t|t}^{(j)}$ and $a_{t|t}^{(j)}$ are the average expectations

[7] Equation (20) is the law of motion of the price level in the log-linearized ICKM. Detailed derivations are in appendix A.

of order j about the state of monetary policy, m_t, and the state of technology, a_t, respectively. The average j-th order expectations about the state of monetary policy are defined[8] 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 about technology are analogously defined. The parameter $(1-\lambda)$ can be shown to quantify the degree of strategic complementarities in price setting, which is defined as the partial derivative of firm i's optimal price with respect to its first-order expectations about the average price.

In the log-linearized ICKM, the measurement equation for the real GDP is[9]

$$\ln Y_t = \left[m_t - \sum_{j=0}^{\infty} (1-\lambda)^j \lambda m_{t|t}^{(j+1)}\right] + \sum_{j=0}^{\infty} (1-\lambda)^j \lambda a_{t|t}^{(j+1)} - \ln \bar{y} + A_0 \cdot t \qquad (21)$$

The measurement equations of the Calvo model are standard and hence omitted.

The Kalman filter can be used to evaluate the likelihood function of the models. Yet, the filter must be initialized and a distribution for the state vector in period $t = 0$ has to be specified. As far as the vector of stationary state variables is concerned, I use their unconditional distributions. I cannot initialize the vector of non-stationary state variables (i.e. m_t, and a_t) in the same manner, since their unconditional variance is not defined. I follow the approach introduced by Chang, Doh, and Schorfheide (2007), who propose to factorize the initial distribution as $p(\mathbf{s}_{1,t})\, p(\mathbf{s}_{2,t})$, where $\mathbf{s}_{1,t}$ and $\mathbf{s}_{2,t}$ are the vector of stationary and non-stationary variables, respectively. They suggest setting the first component $p(\mathbf{s}_{1,t})$ equal to the unconditional distribution of $\mathbf{s}_{1,t}$, whereas the second component $p(\mathbf{s}_{2,t})$ is absorbed into the specification of the prior.

3.3 Prior Distributions

Given the observables presented in Section 3.1, it is easy to show that the Frisch labor elasticity, η, the demand elasticity, ν, and the technology parameter, ϕ, cannot be separately identified in the log-linearized ICKM. Furthermore, the parameter, α, and the discount factor, β, drop out when one log-linearizes the model. After the

[8] By convention, the average zero-order expectations equals the variable itself: $m_{t|t}^{(0)} = m_t$.

[9] Equation (21) is the law of motion of real output in the log-linearized ICKM. Detailed derivations are in appendix A.

log-linearization, the set of identifiable parameters in the log-linearized ICKM is:

$$\Theta_I \equiv (\rho_m, A_0, M_0, \lambda, \sigma_m, \sigma_a, \tilde{\sigma}_m, \tilde{\sigma}_a) \tag{22}$$

Table 1a reports the prior medians and 90% credible intervals for the parameters of the ICKM.

The parameter λ determines the strategic complementarity in the ICKM. It can be shown that this parameter is a function of the Frisch labor-supply elasticity, η, the technology parameter, ϕ, and the demand elasticity, ν. If the Frisch labor-supply elasticity, η, is equal to 0.5 (Ríos-Rull, Schorfheide, Fuentes-Albero, Kryshko, and Santaeulàlia-Llopis, 2009) and the technology parameter, ϕ, is equal to 0.65 (Cooley and Prescott, 1995), then the prior for the parameter λ reflects the beliefs that the (net) mark-up $(\nu - 1)^{-1}$ in the U.S. ranges from 5% to 23%.[10]

Market clearing for the monetary market implies that the stock of money M_t is equal to nominal output. See equation (10). Hence, the autoregressive parameter, ρ_m, and the standard deviation of the monetary shock, σ_m can be estimated by using presample observations of the U.S. real GDP and the U.S. GDP deflator. This presample data set is obtained from Haver Analytics and ranges from the first quarter of 1949 to the second quarter of 1954.

The prior for the standard deviation of the technology shock, σ_a, is centered at 0.007. This value is standard in the real-business cycle literature (Kydland and Prescott, 1982).

In absolute terms, I set the priors for standard deviations of signal noise, $\tilde{\sigma}_m$, and $\tilde{\sigma}_a$, so as to ensure that signals are quite informative about the business-cycle-frequency variations of model variables.[11] In relative terms, these prior specifications are chosen so as to make the two signals equally informative about the corresponding exogenous state variables.[12]

[10] There are studies (e.g., Rotemberg and Woodford, 1997) that quantify the degree of strategic complementarity in the U.S. Unfortunately, they rely on data sets that are very likely to be collinear to the one used in this paper. Using the results of these studies to formulate priors for parameters would thus be controversial.

[11] I achieve that by setting the prior medians of the coherences between the process of the state variables, in first difference, and their corresponding signals such that these are not smaller than 0.50 at business-cycle frequencies (3-5 years). The coherence ranges from 0 to 1 and measures the degree to which two stationary stochastic processes are jointly influenced by cycles of a given frequency (Hamilton, 1994).

[12] I quantify the amount of information that signals convey about the two exogenous states as in

Table 1b presents the implied prior distributions for the strategic complementarity, $1-\lambda$, and the signal-to-noise ratios, $\sigma_m/\tilde{\sigma}_m$ and $\sigma_a/\tilde{\sigma}_a$. As discussed in Woodford (2002), these parameter values crucially influence the persistence in the ICKM.

As far as the log-linearized Calvo model is concerned, the parameter λ cannot be identified. I can only identify the slope of the New Keynesian Phillips curve, κ_{pc}. Furthermore, the parameter, α, is cancelled when one log-linearizes the model. In the log-linearized Calvo model, the set of identifiable parameters is:

$$\Theta_C \equiv (\rho_m, A_0, M_0, \sigma_m, \sigma_a, \kappa_{pc}, \beta) \tag{23}$$

In Table 1a the priors for these parameters are reported. I use the same prior distributions for those parameters that are common to the ICKM. Furthermore, priors reflect the belief that the slope of the Phillips curve, κ_{pc}, ranges from 0.00 to 0.22. This range includes values that are supported by several studies that estimate the slope of the Phillips curve, as surveyed by Schorfheide (2008).

3.4 Posterior Distributions

Given the priors and the likelihood functions implied by the ICKM and the Calvo model, a closed-form solution for the posterior distributions for parameters cannot be derived. However, I am able to evaluate the posteriors numerically through the random-walk Metropolis-Hastings algorithm. How these procedures apply to macro DSGE models is exhaustively documented by An and Schorfheide (2007). I generate $1,000,000$ draws from the posteriors. The posterior medians and 90% credible sets are shown in Table 2.

The posterior median of the strategic-complementarity parameter, λ, is 0.32. This estimate is plausible. This number is consistent with a Frisch labor-supply elasticity, η, of 0.5 (Ríos-Rull et al., 2009), a technology parameter, ϕ, of 0.65 (Cooley and Prescott, 1995), and a mark-up of about 9.5% (Woodford, 2003 and Rotemberg and Woodford, 1997). The parameter $(1-\lambda)$ controls the degree of strategic complementarity in price-setting. The prior median for this parameter was set at 0.59 (see Table 1b). The Bayesian updating pushes the degree of strategic complementarities toward a larger value than what is conjectured in the prior. As pointed out by Woodford (2002), a larger degree of strategic complementarities amplifies the persistence in the

Sims (2003). The formal definition of this measure is provided in Section 4.1.

mechanism of shock propagation, for any finite value of the signal-to-noise ratios. Figure 1 compares the prior and the posterior distributions for the degree of strategic complementarities.

The signal-to-noise ratio concerning the state of monetary policy, $\sigma_m/\tilde{\sigma}_m$, is smaller than that of technology, $\sigma_a/\tilde{\sigma}_a$, by a factor of six. This number affects the relative speed of adjustment of variables to the two shocks. Furthermore, the Bayesian updating raises the difference between the two signal-to-noise ratios (compare Table 1b with Table 2). The likelihood thus indicates that firms acquire little information about the state of monetary policy.

As far as the Calvo model is concerned, the posterior median of the slope of the Phillips curve, κ_{pc}, is 0.012. The 90% posterior credible set ranges from 0.006 to 0.019. This number is in line with previous studies as surveyed by Schorfheide (2008).

3.5 MDD-Based Comparisons

I fit a VAR with four lags to the data set presented in Section 3.1. The Minnesota random walk prior (Doan, Litterman, and Sims, 1984) is implemented in order to obtain a prior distribution for the VAR parameters. I obtain $100,000$ posterior draws through Gibbs sampling. I then compute the IRFs of output and inflation to monetary shocks in the VAR and compare these IRFs with those implied by the two structural DSGE models. In this comparison, the IRFs implied by the VAR are used as the benchmark. From a Bayesian perspective, this comparison is sensible only if the VAR attains a posterior probability larger than those of the DSGE models, as pointed out in Schorfheide (2000).

Let me denote the ICKM, the Calvo model, and the VAR as $\mathcal{M}_I$, $\mathcal{M}_C$, and $\mathcal{M}_V$, respectively. The posterior probability of the model $\mathcal{M}_s$, where $s \in \{I, C, V\}$, is given by:

$$\pi_{T,\mathcal{M}_s} = \frac{\pi_{0,\mathcal{M}_s} \cdot P\left(\tilde{Y}|\mathcal{M}_s\right)}{\sum_{s\in\{I,C,V\}} \pi_{0,\mathcal{M}_s} \cdot P\left(\tilde{Y}|\mathcal{M}_s\right)}$$

where $\pi_{0,\mathcal{M}_s}$ stands for the prior probability of model $\mathcal{M}_s$ and $\tilde{Y}$ denotes the data set that is used for estimation. $P\left(\tilde{Y}|\mathcal{M}_s\right)$ is the so-called marginal data density (MDD) of model $\mathcal{M}_s$. The MDD of a model $\mathcal{M}_s$ is defined as $P\left(\tilde{Y}|\mathcal{M}_s\right) \equiv \int \mathcal{L}\left(\Theta_s|\tilde{Y}, \mathcal{M}_s\right) p\left(\Theta_s|\mathcal{M}_s\right) d\Theta_s$, where Θ_s denotes the vector of parameter in the model $\mathcal{M}_s$, $\mathcal{L}(\cdot)$ stands for the likelihood function, and $p(\Theta_s|\mathcal{M}_s)$ is the prior distribution.

As standard, the prior probabilities, $\pi_{0,\mathcal{M}_s}$, are assumed to be the same across models, that is $\pi_{0,\mathcal{M}_s} = 1/3$, all $s \in \{I, C, V\}$. Therefore, the model that attains the largest posterior probability is the one with the highest MDD. I use Geweke's harmonic mean estimator (Geweke, 1999) to approximate the MDDs of the ICKM and the Calvo model. To compute the MDD of the VAR, I apply the method introduced by Chib (1995).

The log of the MDDs of the three models are reported in Table 3. The VAR has the largest posterior probability. This result is not surprising, since both the ICKM and the Calvo model are very stylized. From a Bayesian perspective, this result legitimates the use of the IRFs implied by the VAR as a benchmark for studying whether the ICKM can accurately explain the propagation of monetary shocks observed in the data.

Moreover, the ICKM has a posterior probability larger than that in the Calvo model. More precisely, the prior probability of the Calvo model, $\pi_{0,\mathcal{M}_C}$, must be bigger than that of the ICKM, $\pi_{0,\mathcal{M}_I}$, by a factor of 2.46E+08 in order for the former model to attain larger posterior probabilities than the latter. From this result, it follows that the ICKM is better than the Calvo model in approximating the true probability distribution of the data-generating process under the Kullback-Leibler distance (Fernández-Villaverde and Rubio-Ramírez, 2004). It is important to emphasize that the fact that the Calvo model has one parameter less than the ICKM is not worrisome, since the MDDs penalize models for the number of their parameters.

3.6 IRF-Based Comparisons

In order to identify the monetary shock in the VAR, I use the restriction that monetary policy has no long-run real effects (e.g., Blanchard and Quah, 1989). This identification scheme is consistent with both the ICKM and the Calvo model.

The IRFs of real output and inflation to a monetary shock implied by the ICKM, the Calvo model and the VAR are plotted in Figures 2 and 3, respectively. The size of the shock is normalized so that the reaction of variables upon the shock is the same in all models. As also found by other studies (e.g., Christiano *et al.*, 2005), the VAR-based IRFs document highly persistent and hump-shaped effects of monetary disturbances on output and inflation.

The Calvo model does not seem to be well-suited for matching the hump-shaped pattern of the VAR response, whereas the ICKM appears to be far more successful in this respect. Moreover, it is worthwhile noticing that the IRF of real output implied by the ICKM peaks three quarters after the occurrence of the shock, exactly as suggested by the benchmark VAR. On the contrary, the Calvo model predicts that the largest response of real output occurs two quarters after the shock. These results thus suggest that the ICKM - albeit very stylized - provides an accurate description of the propagation mechanism of monetary shocks.

4 A Simple Rational Inattention Model

In Section 3.4, I show that the posterior median for the signal-to-noise ratio concerning the state of monetary policy is smaller than that of technology by a factor of six (see Table 2). This result implies that firms acquire little information about the rate of monetary policy.

I then ask the following question: is it plausible that firms acquire so little information about monetary policy? To answer this question, I construct a simple rational inattention model (RIM), which is nested within the ICKM. In the RIM, firms choose the optimal signal-to-noise ratios concerning monetary policy and technology along a schedule that is extensively used in the literature of rational inattention. This problem is called optimal allocation of attention.

In Section 4.1, I show how to construct the signal-to-noise schedule in the RIM. In Section 4.2, I formally describe the problem that firms face to optimally allocate their attention in the RIM. Section 4.3 compares the marginal values of information on monetary policy in the RIM with that in the estimated ICKM. In Section 4.4, I compare the optimal allocation of attention with the estimated one in the ICKM. In Section 4.5, the paper assesses whether the ICKM can account for the highly persistent effects of money shocks when the allocation of attention is plausibly calibrated.

4.1 The Signal-to-Noise Schedule

Rational-inattention models rely on an information-theoretic measure to quantify the amount of processed information, as proposed by Sims (2003). This measure quantifies the reduction of uncertainty that occurs after having observed the last realization of signals. More formally,

$$\kappa \equiv H\left(m_t, a_t | z_{m,i}^{t-1}, z_{a,i}^{t-1}\right) - H\left(m_t, a_t | z_{m,i}^{t}, z_{a,i}^{t}\right) \tag{24}$$

where $H(\cdot)$ denotes the conditional entropy, which measures the uncertainty about a random variable, and $z_{m,i}^t$ and $z_{a,i}^t$ stand for the history of the two signals observed by firm i at time t. The conditional entropy is defined as

$$H\left(m_t, a_t | z_{m,i}^{\tau}, z_{a,i}^{\tau}\right) = \int\int \log_2\left[p\left(m_t a_t | z_{m,i}^{\tau}, z_{a,i}^{\tau}\right)\right] p\left(m_t a_t | z_{m,i}^{\tau}, z_{a,i}^{\tau}\right) dm_t da_t$$

where $\tau \in \{0, 1, \ldots\}$ and $p\left(m_t a_t | z_{m,i}^{\tau}, z_{a,i}^{\tau}\right)$ is the conditional probability density function of the two exogenous state variables, m_t and a_t.

Since both signals and the two exogenous state variables are orthogonal, one can show that equation (24) can be re-written as

$$\kappa = \kappa_m + \kappa_a \tag{25}$$

where κ_m and κ_a stand for the information flows regarding monetary policy and technology, respectively, and are defined as:

$$\begin{aligned} \kappa_m &\equiv H\left(m_t | z_{m,i}^{t-1}\right) - H\left(m_t | z_{m,i}^{t}\right) \\ \kappa_a &\equiv H\left(a_t | z_{a,i}^{t-1}\right) - H\left(a_t | z_{a,i}^{t}\right) \end{aligned}$$

The unit of measurement of the information flows κ, κ_m, κ_a is the bit.

To define the signal-to-noise schedule, let me introduce the mappings g_m and g_a that link the signal-to-noise ratios and the information flows as follows:

$$\kappa_m = g_m\left(\sigma_m, \tilde{\sigma}_m, \rho_m\right), \quad \kappa_a = g_a\left(\sigma_a, \tilde{\sigma}_a\right) \tag{26}$$

The mapping g_a can be analytically derived, while the mapping g_m can be only computationally approximated.[13]

For any given κ, σ_m, σ_a, and ρ_m, the signal-to-noise schedule is implicitly defined by equations (25) and (26). In other words, the signal-to-noise schedule is defined as

[13] The derivations are contained in appendix, C.

a set of pairs of signal-to-noise ratios $\{\sigma_m/\tilde{\sigma}_m, \sigma_a/\tilde{\sigma}_a\}$ that imply the same overall amount of processed information, κ.

4.2 The Optimal Allocation of Attention

The RIM shares all the features of the log-linearized ICKM, as detailed in Section 2.2 to Section 2.7. In the RIM, unlike in the ICKM, firms decide how to allocate their available attention, κ, between observing monetary policy, κ_m, and technology, κ_a. In period zero,[14] firms allocate their available attention, κ,[15] by solving:

$$\max_{\kappa_m,\kappa_a} \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t \left(\hat{p}^*_{i,t}, \hat{p}_t, \hat{y}_t, \hat{q}_t\right) | \mathbf{z}_i^0\right], \tag{27}$$

subject to:

$$\ln P^*_{i,t} \equiv \mathbb{E}\left[(1-\lambda)\ln P_t + \lambda m_t - \lambda a_t | \mathbf{z}_i^t\right] \tag{28}$$

$$\mathbf{z}_{i,t} = \begin{bmatrix} m_t \\ a_t \end{bmatrix} + \begin{bmatrix} \tilde{\sigma}_m & 0 \\ 0 & \tilde{\sigma}_a \end{bmatrix} \mathbf{e}_{i,t} \tag{29}$$

$$\tilde{\sigma}_m = g_m^{-1}(\kappa_m, \sigma_m, \rho_m), \quad \tilde{\sigma}_a = g_a^{-1}(\kappa_a, \sigma_a) \tag{30}$$

$$\kappa_m + \kappa_a = \kappa, \text{ any } t \tag{31}$$

where $\hat{\pi}_t(\cdot)$ is the log-quadratic approximation of $Q_t\pi_t$, where π_t is the profit function within round brackets in expression (11), $\hat{p}^*_{i,t} \equiv \ln\left(P^*_{i,t}/P_t\right)$, and $\hat{q}_t$ is the log deviations of $q_t \equiv M_t Q_t$ from its value at the deterministic steady state. When firms decide how to allocate their limited attention, they are aware that their choice will affect their optimal price-setting policy as given by equation (28) in any subsequent periods. The signal process (29) is the same as that in the ICKM. See equation (15). Note

[14] The model economy is assumed to be at its deterministic steady state in period 0. Moreover, firms are not allowed to reconsider their allocation of attention in any period after $t = 0$. Since firms' period profit function is quadratic and all shocks are Gaussian, it can be shown that this assumption does not give rise to a problem of time inconsistency of firms' attention policies.

[15] Since [1] the period profit function is quadratic, [2] all shocks are Gaussian and [3] firms are assumed to have received an infinite sequence of signals at time $t = 0$ (see Section 2.6), the objective function of the allocation-of-attention problem can be shown to be the same across firms. Thus, every firm will find it optimal to choose the same allocation of attention, (κ^*_m, κ^*_a). These three conditions are also sufficient to obtain that the information flows, κ_m and κ_a, do not vary over time in the information-processing constraint (31).

also that the mappings $g_m^{-1}(\cdot)$ and $g_a^{-1}(\cdot)$ in equation (30) are the inverse of the functions in (26). The constraint (31) is the information-processing constraint that was derived in Section 4.1. This constraint sets an upper-bound to the overall amount of information that firms can process in every period t. It is worth emphasizing that the information-processing constraint implies that the rate of substitution between pieces of information about the two state variables is one.[16] Solving the allocation-of-attention problem (27)-(31) delivers the optimal allocation of attention (κ_m^*, κ_a^*).

4.3 Marginal Value of Information

The marginal value of information on monetary policy can be quantified by computing the marginal rate of profit, defined as

$$\text{MRP} \equiv \frac{\partial \Pi / \partial \kappa_m}{\partial \Pi / \partial \kappa_a}$$

where Π is the sum of discounted profits in the attention problem.

$$\Pi \equiv \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t \left(\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t\right) |\mathbf{z}_i^0\right] \tag{32}$$

RP measures how many bits of information about technology firms are willing to forgo at most for one bit of information about monetary policy.

The MRP at the optimal allocation of attention (κ_m^*, κ_a^*) is equal to unity as the rate of substitution implied by the information-processing constraint (31) is one. In the estimated ICKM, however, MRP may be different from one. The reason is that the estimated allocation of attention in the ICKM[17] may differ from the optimal one (κ_m^*, κ_a^*). In fact, when one calibrates the parameters of the ICKM by using the posterior medians displayed in Table 2, one finds that the MRP is 46.43.[18] Hence, in the estimated ICKM the marginal value of information on monetary policy is much

[16] In some rational inattention model (e.g., Sims, 2003, 2006), agents choose the stochastic process of the signal under no parametric restrictions. In such models, optimal signals might not be orthogonal and Gaussian. Hence, the rate of substitution is determined endogenously, depending on the nature of the optimal signal. See Maćkowiak and Wiederholt (2009) for a rational inattention model where the rate of substitution is exogenously specified as in this paper.

[17] The estimated allocation of attention in the ICKM is defined as the posterior medians of the information flows κ_m and κ_a, which are defined in (26).

[18] The MRP in the estimated ICKM is obtained by numerically computing the following partial

bigger than in the RIM, implying that firms acquire implausibly too little information about monetary policy in the estimated ICKM.

4.4 Information Flows

One can compute the optimal allocation of attention (k_m^*, κ_a^*) in the RIM as well as the moments of the posterior distributions for the information flows, κ_m and κ_a, in the ICKM. I implement Algorithm 1 and Algorithm 2 that are detailed in the appendix F to perform this task.

Table 4 shows the prior and posterior medians for those parameters and their 90% credible intervals in the estimated ICKM. The posterior medians of the information flows κ_m and κ_a are 0.11 bit and 0.46 bit, respectively. The posterior median of the overall amount of information processed by firms in any quarter, κ, is 0.57 bits. Figure 4 compares the prior and the posterior distributions of the fraction of firms' attention paid to the technology shocks, that is, $\kappa_a / (\kappa_m + \kappa_a)$. This graphical comparison emphasizes that, starting from a very agnostic prior for the allocation of attention, the posterior distribution attributes a large portion of firms' attention to technology (the posterior median is about 82%). Hence, the likelihood indicates that the adjustment of output and inflation to monetary shocks is rather slow, as confirmed by the IRFs in Figures 2 and 3. Furthermore, in Figure 4 the posterior appears to be far tighter than the prior, suggesting that the data are quite informative about the proportion of attention paid to technology.

The optimal allocation of attention is computed by taking the median values of the draws for κ_m^*, κ_a^*, and $\kappa_a^* / (\kappa_m^* + \kappa_a^*)$ obtained by applying Algorithm 1, which is detailed in the appendix F. Results are reported in the bottom panel of Table 4. The optimal allocation of attention is very different from the estimated one (κ_m, κ_a). Unlike in the estimated ICKM, firms pay more attention to monetary policy than to

derivatives at the posterior medians in Table 2:

$$\frac{\partial \Pi}{\partial \tilde{\sigma}_m}; \ \frac{\partial \tilde{\sigma}_m}{\partial \kappa_m}, \ \frac{\partial \Pi}{\partial \tilde{\sigma}_a}; \ \frac{\partial \tilde{\sigma}_a}{\partial \kappa_a}$$

These can be calculated by using the profit function (32) and the mappings g_a^{-1} and g_m^{-1} in the constraint (30). The MRP can then be computed by using the fact that

$$\text{MRP} = \frac{\frac{\partial \Pi}{\partial \tilde{\sigma}_m} / \frac{\partial \tilde{\sigma}_m}{\partial \kappa_m}}{\frac{\partial \Pi}{\partial \tilde{\sigma}_a} / \frac{\partial \tilde{\sigma}_a}{\partial \kappa_a}}.$$

technology in the RIM. While the posterior median for the allocation of attention to technology, $\kappa_a/(\kappa_m+\kappa_a)$, is 82%, the optimal fraction of attention paid to technology, $\kappa_a^*/(\kappa_m^*+\kappa_a^*)$, is 40%. Therefore, according to the theory of rational inattention, firms acquire implausibly too little information about monetary policy.

4.5 Persistence under Plausible Calibration

Firms acquire little information about monetary policy in the estimated ICKM. This finding raises the following questions: why does the likelihood choose this allocation of attention in the ICKM? Is this necessary to match the persistent adjustment of variables to monetary shocks? To answer these questions, I compare the IRFs of output and inflation to monetary shocks in the estimated ICKM and in the RIM. The goal is to assess to what extent the persistence in the IRFs falls if firms optimally choose their allocation of attention. This is by no means an obvious question, since in the RIM the speed of adjustment of variables to shocks does not depend only on the signal-to-noise ratios but also on the degree of strategic complementarities in price setting.

Figures 5-6 show the IRFs of output and inflation to a monetary shock in the estimated ICKM and in the RIM. The latter model can be interpreted as the ICKM where the allocation of attention is plausibly calibrated. These figures also report the IRFs implied by the benchmark VAR, introduced in Section 3.5. Output and inflation adjust very fast to monetary shocks in the model where firms' allocation of attention is plausible. This is not consistent with what is documented by the benchmark VAR. When the allocation of attention in the ICKM is plausibly calibrated, the responses of output and inflation to monetary shocks lie outside the 90% credible set implied by the VAR. The ICKM can thus explain the highly persistent effects of money observed in the data only if firms acquire implausibly too little information about monetary policy.

5 Concluding Remarks

I perform a formal Bayesian evaluation of an ICKM à la Woodford (2002). The model features two aggregate shocks: a monetary shock and a technology shock. I obtain Bayesian estimates for the model parameters and find that the model successfully matches the persistent effects of nominal shocks observed in the data. Nonetheless, I argue that the model can account for the highly persistent effects of monetary shocks only if firms acquire implausibly too little information about these shocks. I reach this conclusion by establishing two results. First, in the estimated ICKM, firms acquire a suboptimal amount of information about monetary shocks. Second, if firms were allowed to optimally decide how much information to acquire about monetary shocks, the model would fail to generate the sluggish adjustment of output and inflation to nominal shocks observed in the data.

As in Woodford (2002), the ICKM features a very stylized monetary policy framework, implying that the central bank controls the stock of money. It turns out that the standard deviation of monetary shocks is estimated to be about four times bigger than in DSGE models where the central bank follows an interest-rate rule (e.g., Smets and Wouters, 2007). It is important to emphasize that if the monetary policy framework leads to overestimating the variance of monetary shocks, then the marginal value of information on monetary shocks in the ICKM is overstated too. The reason is that, *ceteris paribus*, the larger the variance of the monetary shock, the more information firms find it optimal to acquire about the state of monetary policy. The underlying intuition is simple: firms find it optimal to acquire more information about that shock that is harder to keep track of. Therefore, the misspecification of the monetary policy framework in the ICKM might be the culprit for the divergence of the marginal value of information in the estimated ICKM and in the RIM. Assessing whether this conjecture is true would thus be a crucial step for reliably using models of imperfect-common-knowledge à la Woodford (2002) for monetary policy analysis.

References

Adam, K. (2007): "Optimal Monetary Policy with Imperfect Common Knowledge," *Journal of Monetary Economics*, 54(2), 267–301.

——— (2009): "Monetary Policy and Aggregate Volatility," *Journal of Monetary Economics*, 56(S1), S1–S18.

An, S., and F. Schorfheide (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26(2-4), 113–172.

Angeletos, G.-M., and J. La'O (2009): "Incomplete Information, Higher-Order Beliefs, and Price Inertia," *Journal of Monetary Economics*, 56(S1), S19–S37.

Blanchard, O., and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79(4), 655–673.

Calvo, G. A. (1983): "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics*, 12(3), 383–398.

Chang, Y., T. Doh, and F. Schorfheide (2007): "Non-stationary Hours in a DSGE Model," *Journal of Money, Credit, and Banking*, 39(6), 1357–1373.

Chib, S. (1995): "Marginal Likelihood From the Gibbs Output," *Journal of the American Statistical Association*, 90(432), 1313–1321.

Christiano, L. J. (1991): "Modeling the Liquidity Effect of a Money Shock," *Federal Reserve Bank of Minneapolis*, Winter, 1–45.

Christiano, L. J., and M. Eichenbaum (1992): "Liquidity Effects and the Monetary Transmission Mechanism," *American Economic Review*, 82(2), 346–353.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): "Monetary Policy Shocks: What Have We Learned and To What End?," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1A, pp. 65–148. North Holland, New York.

——— (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Cooley, T. F., and E. C. Prescott (1995): "Economic Growth and Business Cycles," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley. Princeton University Press, Princeton.

Doan, T., R. Litterman, and C. Sims (1984): "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3(1), 1–100.

Fernández-Villaverde, J., and J. F. Rubio-Ramírez (2004): "Comparing Dynamic Equilibrium Economies to Data: a Bayesian Approach," *Journal of Econometrics*, 123(3), 153–187.

——— (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74(4), 1059–1087.

Geweke, J. (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communications," *Econometric Reviews*, 18(1), 1–127.

Gorodnichenko, Y. (2008): "Endogenous Information, Menu Costs, and Inflation Persistence," NBER Working paper 14184.

Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press, Princeton.

Hellwig, C. (2002): "Public Announcements, Adjustment Delays and the Business Cycle," UCLA Mimeo.

Kydland, F. E., and E. C. Prescott (1982): "Time To Build and Aggregate Fluctuations," *Econometrica*, 50(6), 1345–1370.

Lorenzoni, G. (2009): "A Theory of Demand Shock," *American Economic Review*, 99(5), 2050–2084.

——— (2010): "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information," *Review of Economic Studies*, 77(1), 305–338.

Lucas, R. E. J. (1972): "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4(2), 103–124.

Luo, Y. (2008): "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics*, 11(2), 366–385.

Maćkowiak, B., E. Moench, and M. Wiederholt (2009): "Sectoral Price Data and Models of Price Settings," *Journal of Monetary Economics*, 56(S)(S1), 78–99.

Maćkowiak, B., and M. Wiederholt (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99(3), 769–803.

——— (2010): "Business Cycle Dynamics under Rational Inattention," CEPR Discussion Paper No. 7691.

Mankiw, G. N., and R. Reis (2002a): "Sticky Information: A Model of Monetary Nonneutrality and Structural Slumps," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, pp. 64–86. Princeton University Press, Princeton.

——— (2002b): "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117(4), 1295–1328.

——— (2006): "Pervasive Stickiness," *American Economic Review*, 96(2), 164–169.

——— (2007): "Sticky Information in General Equilibrium," *Journal of the European Economic Association*, 5(2-3), 603–613.

Nason, J. M., and T. Cogley (1994): "Testing the Implications of Long-Run Neutrality for Monetary Business Cycle Models," *Journal of Applied Econometrics*, (9), S37–S70.

Nimark, K. (2008): "Dynamic Pricing and Imperfect Common Knowledge," *Journal of Monetary Economics*, 55(8), 365–382.

——— (2009): "Dynamic Higher Order Expectations," Mimeo.

Paciello, L. (2008): "Monetary Policy Activism and Price Responsiveness to Aggregate Shocks under Rational Inattention," EIEF Mimeo.

PHELPS, E. S. (1970): *Microeconomic Foundations of Employment and Inflation Theory*. Macmillan, London.

REIS, R. (2006a): "Inattentive Consumers," *Journal of Monetary Economics*, 53(8), 1761–1800.

——— (2006b): "Inattentive Producers," *Review of Economic Studies*, 73(3), 793–821.

——— (2009): "A Sticky-Information General-Equilibrium Model for Policy Analysis," NBER Working paper 14732.

RÍOS RULL, J. V., F. SCHORFHEIDE, C. FUENTES-ALBERO, M. KRYSHKO, AND R. SANTAEULÀLIA-LLOPIS (2009): "Methods versus Substance: Measuring the Effects of Technology Shocks on Hours," NBER Working paper 15375.

RONDINA, G. (2008): "Incomplete Information and Informative Pricing," University of California San Diego, Mimeo.

RONDINA, G., AND T. B. WALKER (2009): "Information Equilibria in Dynamic Economies," Mimeo.

ROTEMBERG, J. J., AND M. WOODFORD (1997): "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomics Annual*, 12, 297–346.

SCHORFHEIDE, F. (2000): "Loss Function-Based Evaluation of DSGE Models," *Journal of Applied Econometrics*, 15(6), 645–670.

——— (2008): "DSGE Model-Based Estimation of the New Keynesian Phillips Curve," *Economic Quarterly*, 94(4), 397–433.

SIMS, C. A. (1998): "Stickiness," *Carnegie-Rochester Conference Series on Public Policy*, 49, 317–356.

——— (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3), 665–690.

——— (2006): "Rational Inattention: Beyond the Linear Quadratic Case," *American Economic Review*, 96(2), 158–163.

SMETS, F., AND R. WOUTERS (2007): "Shocks and Frictions in US Business Cycles: A Bayeasian DSGE Approach," *American Economic Review*, 97(3), 586–606.

STOCK, J. H., AND M. W. WATSON (2001): "Vector Autoregressions," *Journal of Economic Perspectives*, 15(4), 101–115.

TOWNSEND, R. M. (1983a): "Equilibrium Theory with Learning and Disparate Information," in *Individual Forecasting and Aggregate Outcomes*, ed. by R. Frydman, and E. S. Phelps, pp. 169–197. Cambridge University Press, Cambridge.

——— (1983b): "Forecasting the Forecasts of Others," *Journal of Political Economy*, 91(4), 546–588.

VAN NIEUWERBURGH, S., AND L. VELDKAMP (2009): "Information Immobility and the Home Bias Puzzle," *Journal of Finance*, 64, 1187–1215.

WOODFORD, M. (2002): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, pp. 25–58. Princeton University Press, Princeton.

——— (2003): *Interest and Prices*. Princeton University Press, Princeton.

——— (2008): "Information-Constrained State-Dependent Pricing," NBER Working paper 14620.

Tables and Figures

Table 1a: Prior distributions

Name	Range	Density	Median	90% Interval
ρ_m	$[0,1)$	Beta	0.50	$[0.17, 0.0.82]$
A_0	$\mathbb{R}$	Normal	0.00	$[-0.41, 0.41]$
M_0	$\mathbb{R}$	Normal	0.00	$[-0.41, 0.41]$
λ	$[0,1)$	Beta	0.41	$[0.21, 0.60]$
$100\sigma_m$	$\mathbb{R}+$	InvGamma	2.0	$[0.43, 12.87]$
$100\sigma_a$	$\mathbb{R}+$	InvGamma	0.7	$[0.51, 0.87]$
$100\tilde{\sigma}_m$	$\mathbb{R}+$	InvGamma	5.01	$[2.12, 7.91]$
$100\tilde{\sigma}_a$	$\mathbb{R}+$	InvGamma	1.06	$[0.24, 1.87]$
κ_{pc}	$\mathbb{R}+$	Gamma	0.12	$[0.00, 0.22]$
β	$[0,1)$	Beta	0.99	$[0.98, 0.99]$

Table 1b: Implied prior distributions (ICKM)

Name		ICKM	
		Median	90% Interval
$1-\lambda$	strategic complementarity	0.59	$[0.40, 0.79]$
$\sigma_m/\tilde{\sigma}_m$	signal-to-noise ratio MP	0.53	$[0.06, 3.15]$
$\sigma_a/\tilde{\sigma}_a$	signal-to-noise ratio tech.	0.95	$[0.17, 1.88]$

Table 2: Posterior distributions

	ICKM		Calvo Model	
Name	Median	90% Interval	Median	90% Interval
ρ_m	0.34	$[0.24, 0.45]$	0.24	$[0.15, 0.33]$
$100A_0$	0.45	$[0.36, 0.55]$	0.43	$[0.11, 0.74]$
$100M_0$	1.34	$[1.18, 1.49]$	1.34	$[1.20, 1.48]$
λ	0.32	$[0.13, 0.50]$	–	–
$100\sigma_m$	0.88	$[0.81, 0.95]$	0.89	$[0.82, 0.97]$
$100\sigma_a$	0.88	$[0.70, 1.04]$	2.66	$[2.04, 3.36]$
$100\tilde{\sigma}_m$	9.04	$[4.97, 12.77]$	–	–
$100\tilde{\sigma}_a$	1.36	$[0.69, 2.02]$	–	–
κ_{pc}	–	–	0.01	$[0.01, 0.02]$
β	–	–	0.99	$[0.99, 0.99]$
$1-\lambda$	0.69	$[0.50, 0.87]$	–	–
$\sigma_m/\tilde{\sigma}_m$	0.10	$[0.06, 0.14]$	–	–
$\sigma_a/\tilde{\sigma}_a$	0.66	$[0.44, 0.94]$	–	–

Table 3: Logarithms of Marginal Data Densities (MDDs)

	Models		
	ICKM	Calvo	VAR(4)
log MDD	1548.70	1529.38	1727.04

Table 4: Implied prior and posterior distributions

	Prior (ICKM)		
Variables	Descriptions	Median	90% Interval
κ_m	information flow MP	0.54	$[0.07, 1.87]$
κ_a	information flow tech.	0.66	$[0.12, 1.21]$
$\kappa = \kappa_m + \kappa_a$	overall level of attention	1.31	$[0.40, 2.73]$
$\frac{\kappa_a}{\kappa_m+\kappa_a}$	allocation of attention to tech.	0.53	$[0.15, 0.83]$
	Posterior (ICKM)		
Variables	Descriptions	Median	90% Interval
κ_m	information flow MP	0.11	$[0.07, 0.16]$
κ_a	information flow tech.	0.46	$[0.33, 0.72]$
$\kappa = \kappa_m + \kappa_a$	overall level of attention	0.57	$[0.41, 0.86]$
$\frac{\kappa_a}{\kappa_m+\kappa_a}$	allocation of attention to tech.	0.82	$[0.76, 0.86]$
	Optimal Attention		
Variables	Descriptions	Median	90% Interval
κ_m^*	optimal information flow MP	0.34	$[0.24, 0.52]$
κ_a^*	optimal information flow tech.	0.23	$[0.17, 0.33]$
$\kappa = \kappa_m + \kappa_a$	overall level of attention	0.57	$[0.41, 0.86]$
$\frac{\kappa_a}{\kappa_m+\kappa_a}$	allocation of attention to tech.	0.40	$[0.35, 0.46]$

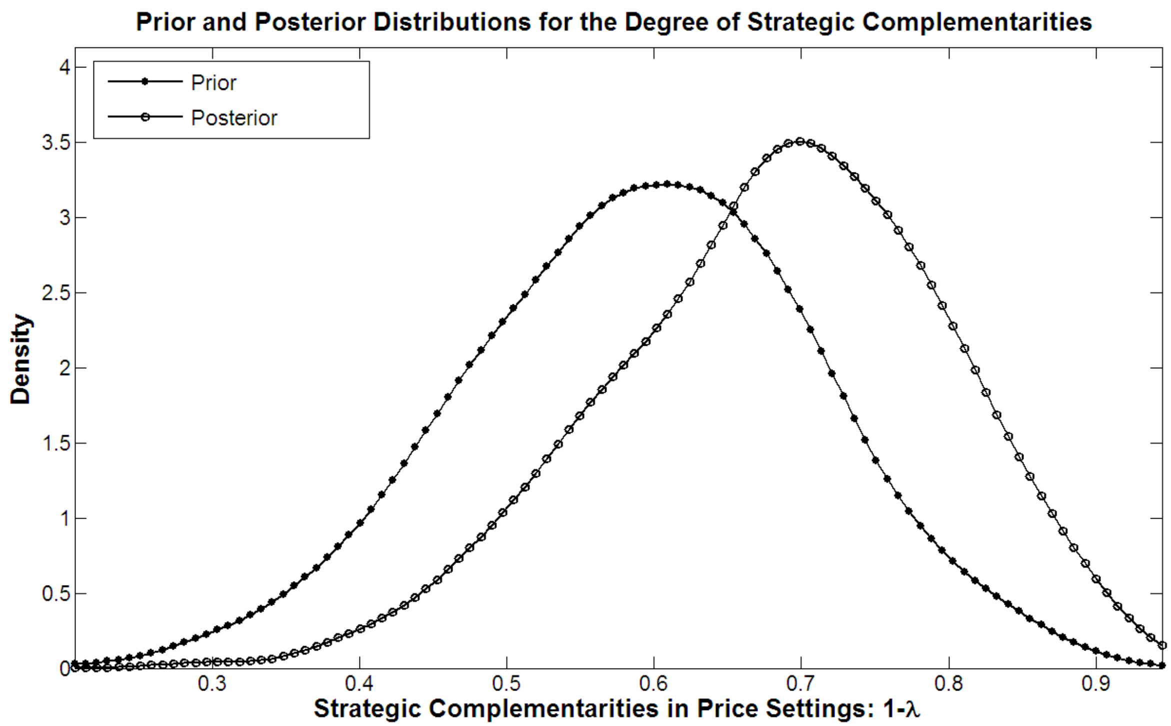
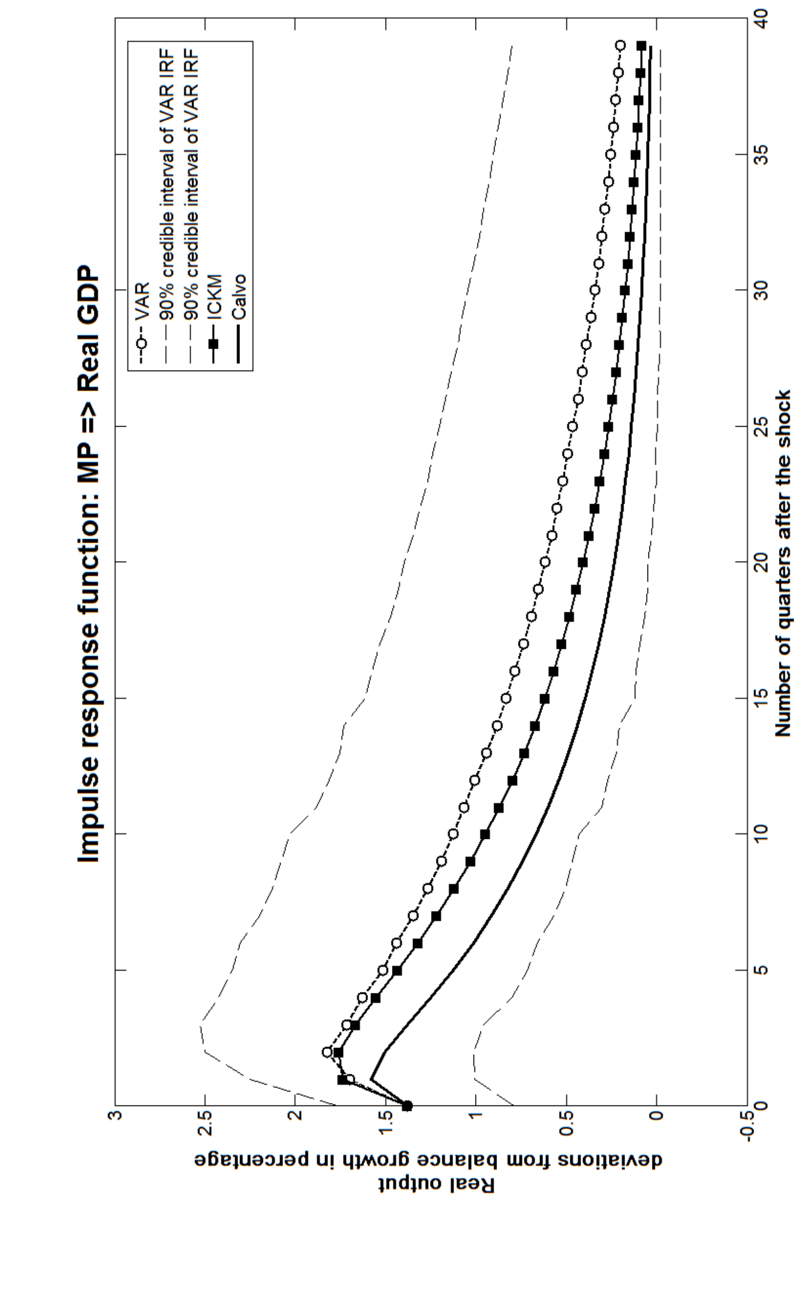


Figure 1

Non-parametric estimates of the prior and posterior distributions based on the draws obtained from the Metropolis-Hastings simulator.



Impulse response function: MP => Real GDP
VAR
90% credible interval of VAR IRF
90% credible interval of VAR IRF
ICKM
Calvo
Real output
deviations from balance growth in percentage
Number of quarters after the shock

Figure 2

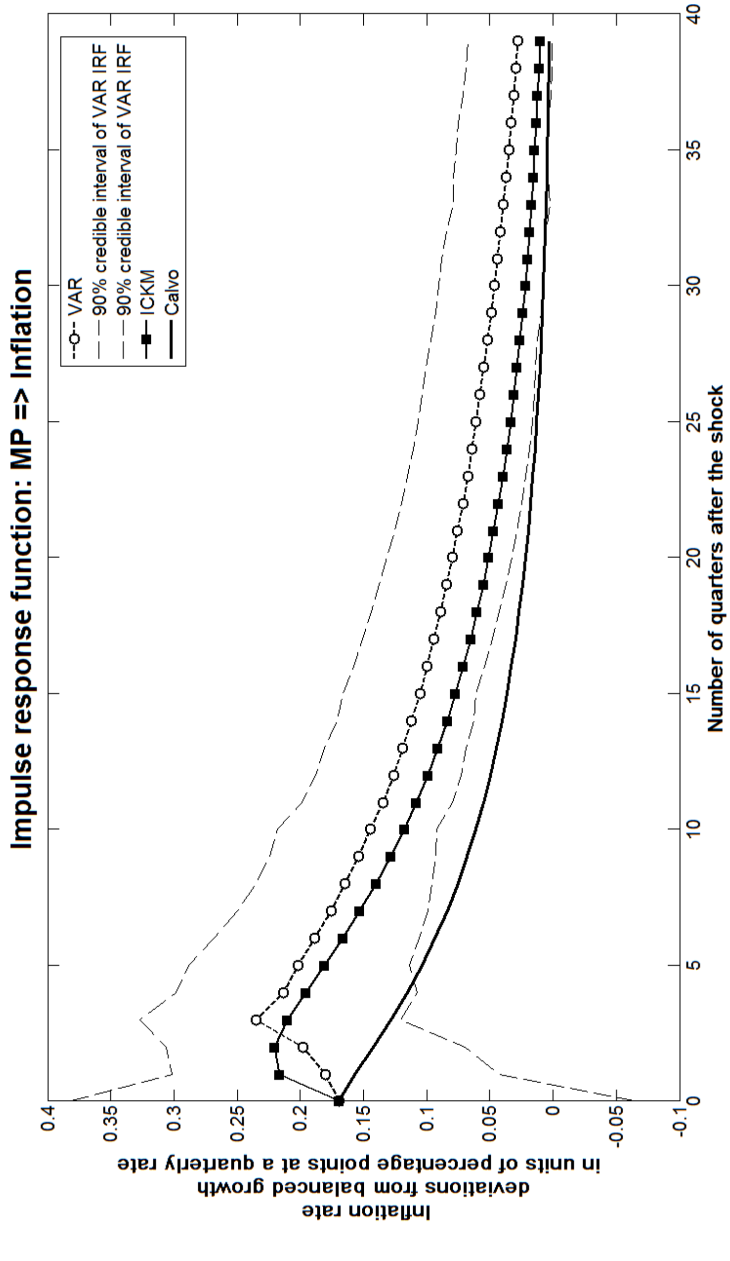

Impulse response function: MP => Inflation
VAR
90% credible interval of VAR IRF
90% credible interval of VAR IRF
ICKM
Calvo
Number of quarters after the shock
Inflation rate
deviations from balanced growth
in units of percentage points at a quarterly rate

Figure 3

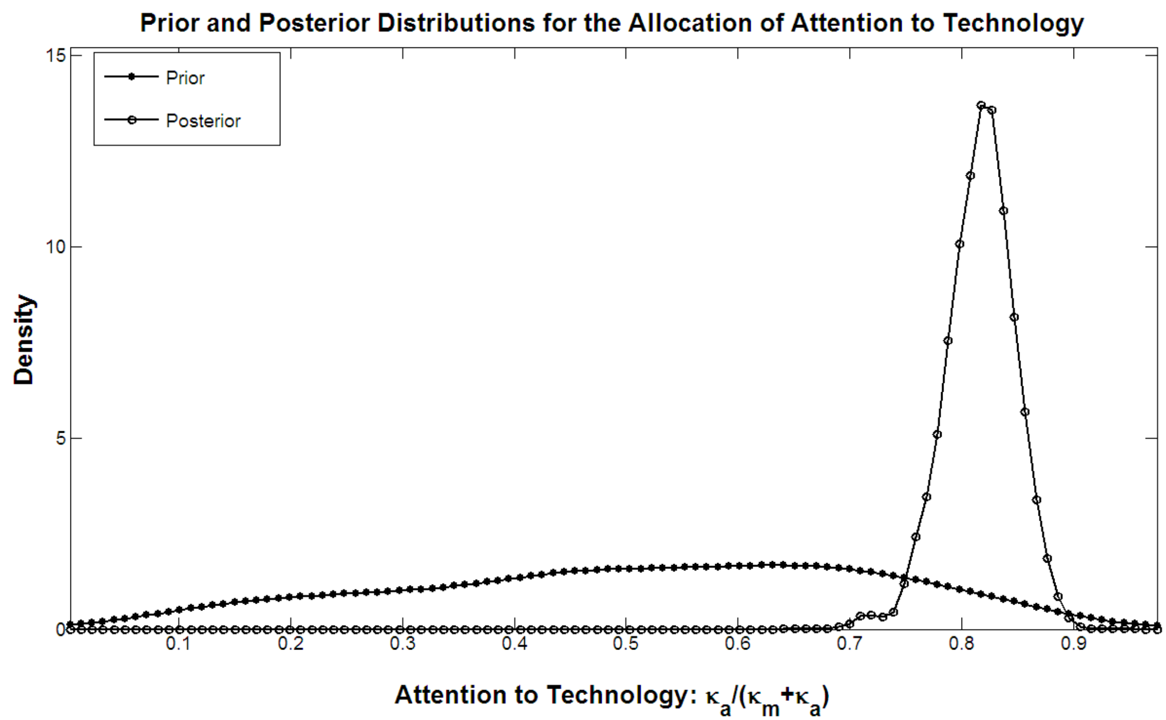


Figure 4

Non-parametric estimates of the prior and posterior distributions based on the draws obtained from the Metropolis-Hastings simulator.

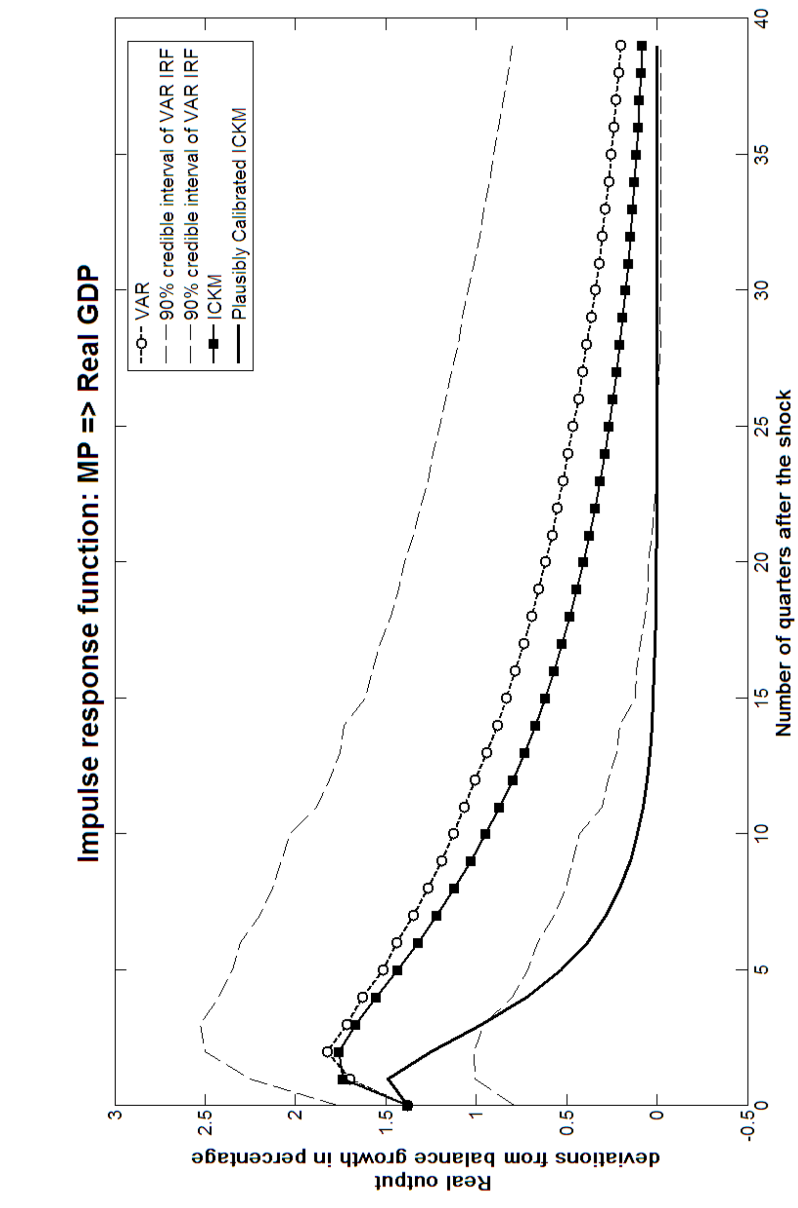

Impulse response function: MP => Real GDP
VAR
90% credible interval of VAR IRF
90% credible interval of VAR IRF
ICKM
Plausibly Calibrated ICKM
Real output
deviations from balance growth in percentage
Number of quarters after the shock
3
2.5
2
1.5
1
0.5
0
-0.5
0
5
10
15
20
25
30
35
40

Figure 5

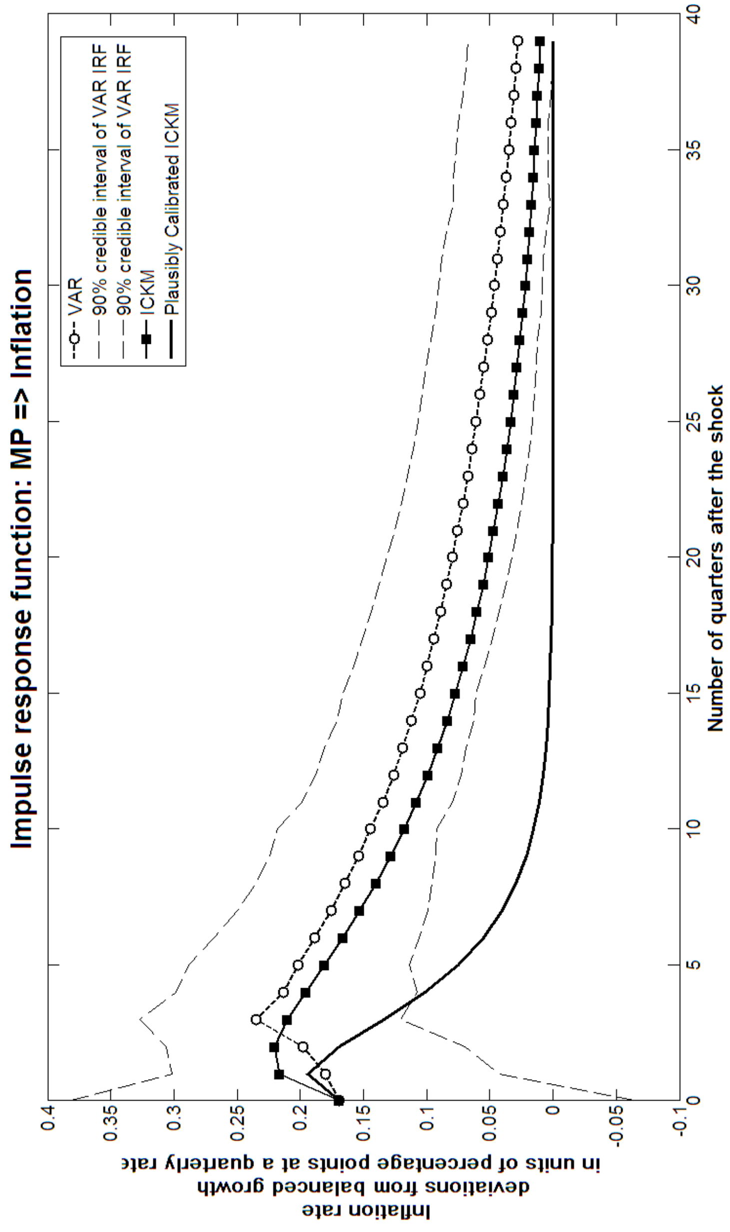

Impulse response function: MP => Inflation
VAR
90% credible interval of VAR IRF
90% credible interval of VAR IRF
ICKM
Plausibly Calibrated ICKM
Inflation rate
deviations from balanced growth
in units of percentage points at a quarterly rate
Number of quarters after the shock
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
0
5
10
15
20
25
30
35
40

Figure 6

Appendix

In appendix A I show in detail how to derive the law of motion of price level and real output as described in Section 2.7. Appendix B explains in much more details than in Section 2.8 how to solve the ICKM. In appendix C, I analytically derive the mapping $g_a(\cdot)$ and show how to compute the mapping $g_m(\cdot)$. In appendix D, I show how to characterize the profit function (27) in the allocation of attention problem. In appendix E, I show how to solve the RIM by exploiting the nestedness with the ICKM. In appendix F, I present two algorithms that can be used to compute the optimal allocation of attention (κ_m^*, κ_a^*) in the RIM as well as the moments of the posterior distributions for the allocation of attention (κ_m, κ_a) in the estimated ICKM.

A Law of Motion of Price and Output in the ICKM

The first-order necessary condition[19] of the price-setting problem in the ICKM is:

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

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

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

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

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

yields:

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

Define the stationary variables:

$$y_t \equiv \frac{Y_t}{A_t}, \; y_{i,t} \equiv \frac{Y_{i,t}}{A_t} \;, \; p_{i,t} = \frac{P_{i,t}}{P_t}, \; h_t = H_t \tag{33}$$

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

$$(1-\nu)\,\mathbb{E}_{i,t} \left[\beta Q_t Y_t p_{i,t}^{-\nu} \left(1 + \nu\phi^{-1} \alpha y_t h_t^{\eta} \left(p_{i,t}^{-\nu} y_t \right)^{\phi^{-1}-1} p_{i,t}^{-1} \right) \right] = 0$$

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:

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

Note also that from the production function $\hat{h}_{i,t} = \phi^{-1}\hat{y}_{i,t}$ and hence[20] $\hat{h}_t = \phi^{-1}\hat{y}_t$. By substituting, this results into the equation above, one obtains:

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

and then

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

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

$$\mathbb{E}_{i,t}\left[\hat{p}_{i,t}\right] = \lambda \mathbb{E}_{i,t}\left[\hat{y}_t\right]$$

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

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 (33) and then write:

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

or equivalently,

$$\ln P_{i,t}=\mathbb{E}_{i,t}\left[\lambda\ln Y_t+\ln P_t-\lambda\ln A_t\right]-\lambda\ln\bar{y}$$

Recall equation (10):

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

and thus,

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

and by rearranging:

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

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 $\mathbb{E}_{i,t}\left(P_t\right)$). 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 (3) around the deterministic steady-state, one obtains $\hat{p}_t=\int\hat{p}_{i,t}di$. Hence, by integrating across firms one obtains:

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

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

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

By recalling that I defined $m_t \equiv \ln M_t - M_0 t$ and $a_t \equiv \ln A_t - A_0 t$ and that firms know all the model parameters, I can re-write the equation above as:

$$\ln P_t = \left[\sum_{j=0}^{\infty} (1-\lambda)^j \lambda \left(m_{t|t}^{(j+1)} - a_{t|t}^{(j+1)}\right)\right] - \ln \bar{y} + M_0 t - A_0 t \tag{34}$$

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

Furthermore, I can combine equations (20) and (10) to get:

$$\underbrace{\ln M_t - \ln Y_t}_{\ln P_t} = \left[\sum_{j=0}^{\infty} (1-\lambda)^j \lambda \left(m_{t|t}^{(j+1)} - a_{t|t}^{(j+1)}\right)\right] - \ln \bar{y} + M_0 t - A_0 t$$

and by re-arranging, this yields:

$$\ln Y_t = \left[m_t - \sum_{j=0}^{\infty} (1-\lambda)^j \lambda m_{t|t}^{(j+1)}\right] + \sum_{j=0}^{\infty} (1-\lambda)^j \lambda a_{t|t}^{(j+1)} - \ln \bar{y} + A_0 t$$

which is the measurement equation (21) in the main text.

B Solving the ICKM

In general, finding an equilibrium in models with incomplete informations requires characterizing infinitely many equilibrium laws of motion, which is absolutely unmanageable. In the present model, this issue can be elegantly resolved as in Woodford (2002). More specifically, I need only to keep track of a specific linear combination of average expectations, appearing in equations (20)-(21). Define the vector $\mathbf{F}_t$ as

$$\mathbf{F}_t \equiv \sum_{j=1}^{\infty} (1-\lambda)^{j-1} \lambda \mathbf{X}_t^{(j)} \tag{35}$$

$$\text{where } \mathbf{X}_t \equiv [m_t, m_{t-1}, a_t]' \tag{36}$$

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:

$$\widehat{y}_t = \widehat{p}_t \tag{37}$$

$$\hat{p}_t = \mathbf{r}'\overline{\mathbf{X}}_t \tag{38}$$

$$\overline{\mathbf{X}}_t = \overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t \tag{39}$$

where

$$\overline{\mathbf{X}}_t \equiv \left[\mathbf{X}_t' \;\vdots\; \mathbf{F}_t'\right]', \; \mathbf{r} \equiv [-1,0,1,1,0,-1]'$$
$$\overline{\mathbf{B}} \equiv \begin{bmatrix} \mathbf{B}_{3x3} & \mathbf{0}_{3x3} \\ \mathbf{G}_{3x3} & \mathbf{H}_{3x3} \end{bmatrix}, \quad \overline{\mathbf{b}} = \left[\mathbf{b}\;' \vdots\; \mathbf{d}'\right]' \tag{40}$$

$$\mathbf{B} \equiv \begin{bmatrix} 1+\rho_m & -\rho_m & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{u}_t = [\varepsilon_{m,t}, \varepsilon_{a,t}]'$$
$$\mathbf{u}_t \overset{iid}{\backsim} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_u), \text{ for all } t \text{ and } \Sigma_u = \begin{bmatrix} \sigma_m^2 & 0 \\ 0 & \sigma_a^2 \end{bmatrix}$$

where $\mathbf{G}$, $\mathbf{H}$, and $\mathbf{d}$ are matrices that are not known yet. Equation (37) stems from the log-linearized version of equation (10), where I defined the log-linear deviations of the stationary output, y_t, and price, p_t, from their deterministic steady-state, as $\hat{y}_t$ and $\hat{p}_t$, respectively. Equation (38) can be derived by equation (20) by simply adding $\ln A_t - \ln M_t - \ln \bar{p}$ to both sides of this equation and by 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 (10).

Recall that the signal structure is specified in equations (15). Thus, the firms' observation equations are

$$\mathbf{z}_{i,t} = \mathbf{D}\overline{\mathbf{X}}_t + \mathbf{e}_{i,t} \tag{41}$$

where

$$\mathbf{D} \equiv \begin{bmatrix} \mathbf{D}_1 & \vdots & \mathbf{0}_{2x3} \end{bmatrix} \text{ and } \mathbf{D}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{42}$$

$$\mathbf{e}_{i,t} \backsim \mathcal{N}\left(\mathbf{0}, \mathbf{\Sigma}_e\right), iid \text{ for all } t, \text{ and } i, \ \mathbf{\Sigma}_e = \begin{bmatrix} \tilde{\sigma}_m^2 & 0 \\ 0 & \tilde{\sigma}_a^2 \end{bmatrix} \tag{43}$$

Finding an equilibrium for this economy amounts to characterize the unknown matrices **G**, **H**, and **d**. This requires solving the following fixed point problem. Given the conjectured law of motion (39), optimal firms' behaviors must exactly aggregate to the conjectured law of motion (39). 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-termed *kalman-filter equation*

$$\overline{\mathbf{X}}_{t|t}(i) = \overline{\mathbf{X}}_{t|t-1}(i) + \mathbf{k}\left[\mathbf{z}_t(i) - \mathbf{D}\overline{\mathbf{X}}_{t|t-1}(i)\right] \tag{44}$$

where $\mathbf{k}$ is the $6x2$ Kalman gain matrix which is not yet specified. It is easy to show that the one-step-ahead forecast of the state vector is:

$$\overline{\mathbf{X}}_{t|t-1}(i) = \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1}(i) \tag{45}$$

I can plug the (45) into the (44) to get the law of motion for firm i's estimate of the current state vector

$$\overline{\mathbf{X}}_{t|t}(i) = \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1}(i) + \mathbf{k}\left[\mathbf{z}_t(i) - \mathbf{D}\overline{\mathbf{X}}_{t|t-1}(i)\right] \tag{46}$$

By integrating the (46) over firms (i.e. $\int \overline{\mathbf{X}}_{t|t}(i)\,di \equiv \overline{\mathbf{X}}_{t|t}$) one gets

$$\overline{\mathbf{X}}_{t|t} = \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1} + \mathbf{k}\mathbf{D}\left[\overline{\mathbf{X}}_t - \overline{\mathbf{X}}_{t|t-1}\right] \tag{47}$$

This result follows from the observing that on aggregate the signal noise washes out (i.e. $\int \mathbf{e}_t(i)\,di = \mathbf{0}$) and hence

$$\int \mathbf{z}_t(i)\,di = \mathbf{D}\overline{\mathbf{X}}_t + \int \mathbf{e}_t(i)\,di$$

$$\int \mathbf{z}_t(i)\,di = \mathbf{D}\overline{\mathbf{X}}_t$$

By using the transition equation (39) to get rid of $\overline{\mathbf{X}}_t$ in the equation (47) I obtain

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

Then by integrating the (45), 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

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

Gathering the common terms yields

$$\overline{\mathbf{X}}_{t|t} = [\mathbf{I} - \mathbf{kD}]\,\overline{\mathbf{B}\mathbf{X}}_{t-1|t-1} + \mathbf{kD}\left[\overline{\mathbf{B}\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t\right] \tag{48}$$

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

It is convenient to define the $6x3$ vector $\boldsymbol{\varphi}$ such that

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

Then one can note the following

$$\boldsymbol{\varphi}'\overline{\mathbf{X}}_t^{(1)} = \mathbf{F}_t \tag{49}$$

It is easy to prove that equation (49) is indeed true by working as follows

$$\begin{aligned} \boldsymbol{\varphi}'\overline{\mathbf{X}}_t^{(1)} &= \left[(\lambda)\cdot \mathbf{I}_3 \vdots (1-\lambda)\cdot \mathbf{I}_3\right] \cdot \begin{bmatrix} \mathbf{X}_t^{(1)} \\ \cdots \\ \mathbf{F}_t^{(1)} \end{bmatrix} \\ \boldsymbol{\varphi}'\overline{\mathbf{X}}_t^{(1)} &= \lambda \mathbf{X}_t^{(1)} + (1-\lambda)\,\mathbf{F}_t^{(1)} \end{aligned} \tag{50}$$

Let me introduce the following notations:

$$x_{t|t}^{(k-1)} \equiv x_t^{(k)}, \qquad \forall k \geq 1; \quad x_t^{(0)} \equiv x_t \tag{51}$$

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 (35)

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

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

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

Now recall equation (51) to finally write

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

Comparing this equation with the (35) concludes the proof of (49). Now one can plug equation (48) into equation (49) to get

$$\mathbf{F}_t = \left[\boldsymbol{\varphi}' - \widetilde{\mathbf{k}}\mathbf{D}\right] \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1} + \widetilde{\mathbf{k}}\mathbf{D}\left[\overline{\mathbf{B}\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t\right] \tag{52}$$

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

FACT 1

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

FACT 2

$$\begin{aligned} \mathbf{D}\overline{\mathbf{B}} &= \left[\mathbf{D}_1\mathbf{B} \vdots\; \mathbf{0}_{2x3}\right] \\ &= \left[\mathbf{B}^{\dagger} \vdots\; \mathbf{0}_{2x3}\right] \end{aligned} \tag{53}$$

where $\mathbf{B}^{\dagger} \equiv \begin{bmatrix} \mathbf{B}_1' & \mathbf{B}_3' \end{bmatrix}'$ 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

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

The FACT 2 allows is to get the following results:

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

and

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

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

$$\varphi'\overline{\mathbf{BX}}_{t-1|t-1} = \lambda\mathbf{BX}_{t-1|t-1} + (1-\lambda)\,\mathbf{GX}_{t-1|t-1} + \;(1-\lambda)\,\mathbf{H}\cdot\mathbf{F}_{t-1|t-1}$$

By collecting all these results one can rewrite equation (52) as follows

$$\mathbf{F}_t = \left[\lambda\mathbf{B} + (1-\lambda)\,\mathbf{G} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\right]\mathbf{X}_{t-1|t-1} + (1-\lambda)\,\mathbf{HF}_{t-1|t-1} + \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\mathbf{X}_{t-1} + \widetilde{\mathbf{k}}\mathbf{u}_t \tag{54}$$

Next, I will work out the vector $\mathbf{F}_{t-1}$ from $\mathbf{F}_{t-1|t-1}$, since I want to rewrite equation

(54) in a form that is comparable to that conjectured in equation (39) so as I can compare my initial guess. One should start from equation (49) 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

$$(1-\lambda)\cdot \mathbf{F}_{t-1|t-1} = \mathbf{F}_{t-1} - \lambda \mathbf{X}_{t-1|t-1} \tag{55}$$

I can now plug equation (55) into equation (54) to get

$$\begin{aligned} \mathbf{F}_t &= \left[\lambda \mathbf{B} + (1-\lambda)\,\mathbf{G} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\right]\mathbf{X}_{t-1|t-1} + \mathbf{H}\left[\mathbf{F}_{t-1} - \lambda \mathbf{X}_{t-1|t-1}\right] + \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\mathbf{X}_{t-1} + \widetilde{\mathbf{k}}\mathbf{u}_t \\ \mathbf{F}_t &= \left[\lambda \mathbf{B} + (1-\lambda)\,\mathbf{G} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} - \lambda \mathbf{H}\right]\mathbf{X}_{t-1|t-1} + \mathbf{H}\cdot \mathbf{F}_{t-1} + \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\mathbf{X}_{t-1} + \widetilde{\mathbf{k}}\mathbf{u}_t \qquad (56) \end{aligned}$$

Now equation (56) has the same form as the bottom rows of equation (39) 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:

$$\mathbf{G} = \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} \tag{57}$$

$$\mathbf{d} = \widetilde{\mathbf{k}} \tag{58}$$

and

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

By substituting (57) into the last equation one obtains

$$\left[\mathbf{B} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} - \mathbf{H}\right] \stackrel{!}{=} 0$$

$$\mathbf{H} \stackrel{!}{=} \mathbf{B} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} \tag{59}$$

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

$$\mathbf{k} = \mathbf{P}\mathbf{D}'\left[\mathbf{D}\mathbf{P}\mathbf{D}' + \mathbf{\Sigma}_e\right]^{-1} \tag{60}$$

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

$$\mathbf{P} = \overline{\mathbf{B}} \left[\mathbf{P} - \mathbf{P}\mathbf{D}' \left[\mathbf{D}\mathbf{P}\mathbf{D}' + \boldsymbol{\Sigma}_e \right]^{-1} \mathbf{D}\mathbf{P} \right] \overline{\mathbf{B}}' + \overline{\mathbf{b}} \boldsymbol{\Sigma}_u \overline{\mathbf{b}}' \tag{61}$$

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

Since $\overline{\mathbf{B}}$ and $\overline{\mathbf{b}}$ 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 (57)-(59). Then one has to solve the algebraic Riccati equation (61) for $\mathbf{P}$ and obtain a new Kalman-gain matrix $\mathbf{k}^*$ through the equation (60). 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 (39)-(40) through (57)-(61), which combined with the equations (37)-(38) 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:

$$\kappa_a \equiv H\left(a_t | z_{a,i}^{t-1}\right) - H\left(a_t | z_{a,i}^{t}\right) \tag{62}$$

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

$$H\left(a_t | z_{a,i}^{t}\right) \equiv \frac{1}{2} \log_2 \left[2\pi e \cdot VAR\left(a_t | z_{a,i}^{t}\right) \right] \tag{63}$$

First, let me focus on the mapping

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

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

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

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

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

Note that

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

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

$$VAR\left(a_t|z_{a,i}^t\right)^2 + VAR\left(a_t|z_{a,i}^t\right)\sigma_a^2 = \sigma_a^2\tilde{\sigma}_a^2$$

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

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

Note that I can write:

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

and finally,

$$\tilde{\sigma}_a^2 = \frac{\left[2VAR\left(a_t|z_{a,i}^t\right) + \sigma_a^2\right]^2}{4\sigma_a^2} - \frac{\sigma_a^2}{4} \tag{64}$$

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

Combining the equations (62) and (63) yields

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

Since firms observe infinitely many signals, $VAR\left(a_t|z_i^{t-1}\right) = VAR\left(a_t|z_i^{t}\right) + \sigma_a$. Hence I obtain:

$$\kappa_a = \frac{1}{2}\log_2\left(\frac{VAR\left(a_t|z_i^{t}\right) + \sigma_a^2}{VAR\left(a_t|z_i^{t}\right)}\right)$$

If one inverts this equation, one obtains:

$$VAR\left(a_t|z_i^{t}\right) = \frac{\sigma_a^2}{2^{2\kappa_a} - 1} \tag{65}$$

Plugging this result into equation (64) leads to:

$$\kappa_a = \frac{1}{2}\log_2\left[\frac{1}{\left(\frac{\tilde{\sigma}_a^2}{\sigma_a^2} + \frac{1}{4}\right)^{\frac{1}{2}} - \frac{1}{2}} + 1\right] \tag{66}$$

This is the mapping g_a in equation (26).

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

$$H\left(m_t|z_{1,i}^{\tau}\right) = \frac{1}{2}\log_2\left[2\pi e \cdot VAR\left(m_t|z_{1,i}^{\tau}\right)\right] \tag{67}$$

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

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

for $\tau \in \{t-1, t\}$, where $\overline{\mathbf{X}}_t \equiv \left[\mathbf{X}_t' \,\vdots\, \mathbf{F}_t'\right]'$, $\mathbf{X}_t \equiv [m_t, m_{t-1}, a_t]'$, 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 $VAR\left(m_t|z_{1,i}^{t-1}\right) = \mathbf{P}_{t|t-1}[1,1]$ and

$VAR\left(m_t|z_{1,i}^t\right) = \mathbf{P}_{t|t}[1,1]$, 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 (61). The matrix $\mathbf{P}_{t|t}$ is defined as:

$$\mathbf{P}_{t|t} \equiv \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{D}'\left[\mathbf{D}\mathbf{P}_{t|t-1}\mathbf{D}' + \mathbf{\Sigma}_e\right]^{-1}\mathbf{D}\mathbf{P}_{t|t-1} \tag{68}$$

where the matrices $\mathbf{D}$ and $\mathbf{\Sigma}_e$ have been defined in (42) and in (43), respectively.

Thus, after one has characterized the fixed point as discussed in Appendix B, one can use the resulting matrix $\mathbf{P}$ and equation (68) to pin down the conditional variances $VAR\left(m_t|z_{1,i}^\tau\right)$, for $\tau \in \{t-1, t\}$, the condition entropies $H\left(m_t|z_{m,i}^\tau\right)$, for $\tau \in \{t-1, t\}$, through equation (67), and finally the information flow $\kappa_m \equiv H\left(m_t|z_{m,i}^{t-1}\right) - H\left(m_t|z_{m,i}^t\right)$.

D Optimal Allocation of Attention

The objective function in the attention problem (27)-(31) 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 $\hat{p}_{i,t}^{\diamondsuit}$. It is easy to show that

$$\hat{\pi}_t\left(\hat{p}_{i,t}^{\diamondsuit}, \hat{p}_t, \hat{y}_t, \hat{q}_t\right) - \hat{\pi}_t\left(\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t\right) \propto \left(\hat{p}_{i,t}^{\diamondsuit} - \hat{p}_{i,t}^*\right)^2 \tag{69}$$

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

$$\hat{\pi}_t\left(\hat{p}_{i,t}^{\diamondsuit}, \hat{p}_t, \hat{y}_t, \hat{q}_t\right) - \hat{\pi}_t\left(\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t\right) \propto \left(\ln\left(P_{i,t}^{\diamondsuit}\right) - \ln\left(P_{i,t}^*\right)\right)^2 \tag{70}$$

First note that $\hat{\pi}_t\left(\hat{p}_{i,t}^{\diamondsuit}, \hat{p}_t, \hat{y}_t, \hat{q}_t\right)$ 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

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

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

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

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

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

It then follows

$$\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^t\hat{\pi}_t\left(\hat{p}_{i,t}^{*},\hat{p}_t,\hat{y}_t,\hat{q}_t\right)\right]\propto-\mathbb{E}\left[\left(\ln\left(P_{i,t}^{\diamondsuit}\right)-\ln\left(P_{i,t}^{*}\right)\right)^2\right]\tag{71}$$

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

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

Thus

$$\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^t\hat{\pi}_t\left(\hat{p}_{i,t}^{*},\hat{p}_t,\hat{y}_t,\hat{q}_t\right)\right]\propto-\mathbb{E}\left[\{(1-\lambda)\left[\ln P_t-\mathbb{E}_{i,t}\left(\ln P_t\right)\right]+\lambda\left[\ln M_t-\mathbb{E}_{i,t}\left(\ln M_t\right)\right]+\lambda\left[\ln A_t-\mathbb{E}_{i,t}\right.\right.$$

or equivalently,

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

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

$$\mathbf{P}_{t|t}\equiv\mathbf{P}_{t|t-1}-\mathbf{P}_{t|t-1}\mathbf{D}'\left[\mathbf{D}\mathbf{P}_{t|t-1}\mathbf{D}'+\mathbf{\Sigma}_e\right]^{-1}\mathbf{D}\mathbf{P}_{t|t-1}\tag{72}$$

where the matrix $\mathbf{P}_{t|t-1}$ is the matrix $\mathbf{P}$ in equation (61) that is obtained by solving the fixed point as discussed in Appendix B. The matrices $\mathbf{D}$ and $\mathbf{\Sigma}_e$ have been defined in (42) and in (43), respectively. The variances and covariances can be obtained from the matrix $\mathbf{P}_{t|t}$:

$$
\begin{aligned}
VAR\left(\tilde{\mathbf{r}}\mathbf{F}_t|\mathbf{z}_i^t\right) &\equiv [\mathbf{0}_{1\times 3}, \tilde{\mathbf{r}}]\,\mathbf{P}_{t|t}\,[\mathbf{0}_{1\times 3}, \tilde{\mathbf{r}}]' \\
VAR\left(m_t|\mathbf{z}_i^t\right) &\equiv [1, \mathbf{0}_{1\times 5}]\,\mathbf{P}_{t|t}\,[1, \mathbf{0}_{1\times 5}]' \\
VAR\left(a_t|\mathbf{z}_i^t\right) &\equiv [\mathbf{0}_{1\times 2}, 1, \mathbf{0}_{1\times 3}]\,\mathbf{P}_{t|t}\,[\mathbf{0}_{1\times 2}, 1, \mathbf{0}_{1\times 3}]' \\
COV\left(\tilde{\mathbf{r}}\mathbf{F}_t, m_t|\mathbf{z}_i^t\right) &= \tilde{\mathbf{r}}\mathbf{P}_{t|t}\,[1, \mathbf{0}_{1\times 5}]' \\
COV\left(\tilde{\mathbf{r}}\mathbf{F}_t, a_t|\mathbf{z}_i^t\right) &= \tilde{\mathbf{r}}\mathbf{P}_{t|t}\,[\mathbf{0}_{1\times 2}, 1, \mathbf{0}_{1\times 3}]'
\end{aligned}
$$

E Solving the RIM

The RIM can be solved through the following steps:

0. Set $i = 1$. Guess the information flows $\kappa_m^{(i)}, \kappa_a^{(i)}$.

1. Obtained the signal noises, $\tilde{\sigma}_m^{(i)}$ and $\tilde{\sigma}_a^{(i)}$ that are implied by the guessed information flows $\left(\kappa_m^{(i)}, \kappa_a^{(i)}\right)$ through the mappings in (30)

2. Find the variance and covariance matrix of the state variables, $\mathbf{P}$, that solves the fixed point, as detailed in Appendix B. Denote this matrix, $\mathbf{P}^{(i)}$.

3. Use the matrix $\mathbf{P}^{(i)}$ to characterize the profit function, as shown in Appendix D. Solve the attention problem (27)-(31) and obtain $\kappa_m^{*(i)}$ and $\kappa_a^{*(i)}$.

4. Compare the guessed information flows $\kappa^{(i)} \equiv \left(\kappa_m^{(i)}, \kappa_a^{(i)}\right)$ with the optimal ones $\kappa^{*(i)} \equiv \left(\kappa_m^{*(i)}, \kappa_a^{*(i)}\right)$. If $\left|\kappa^{(i)} - \kappa^{*(i)}\right| < \varepsilon$ with $\varepsilon > 0$ and small, then STOP. Otherwise, guess $\kappa^{(i+1)} = \kappa^{*(i+1)}$, set $i = i + 1$, and GO TO STEP 1.

F Codes for Solving the RIM

Algorithm 1: Approximation of the Posterior (Prior) Distribution of Information Flows

STEP 0 Set i=1.

STEP 1 Take the i-th posterior (prior) draw for the parameter of the ICKM, $\Theta_I^{(i)}$ from the Metropolis-Hastings posterior (prior) simulator.

STEP 2 Compute the information flows in the ICKM, $\kappa_m^{(i)}$ and $\kappa_a^{(i)}$, by using the mappings in (26).

STEP 3 Given $\kappa_m^{(i)}$ and $\kappa_a^{(i)}$ and equation (26), compute the overall attention in the ICKM $\kappa^{(i)}$.

STEP 4 Solve for the optimal allocation of attention through the Algorithm 2 below. Obtain the optimal allocation of attention $\left(\kappa^{*(i)}_m, \kappa^{*(i)}_a\right)$ and compute the optimal fraction of attention allocated to technology, $\kappa^{*(i)}_a \Big/ \left(\kappa^{*(i)}_m + \kappa^{*(i)}_a\right)$.

STEP 5 If the indicator i equals the number of posterior (prior) draws, STOP. Otherwise, set i=i+1,000 and GO TO STEP 1.

Algorithm 2: Characterizing the Optimal Allocation of Attention

STEP 1 Guess the values of the information flows $\vec{\kappa} \equiv (\kappa_m, \kappa_a)$ and use the mappings in (30) to obtain the implied noise variances $\tilde{\sigma}_m$ and $\tilde{\sigma}_a$.

STEP 2 Given this guess, numerically characterize the law of motion of the price level by applying the method that solves the ICKM (see Section 2.8.)

STEP 3 Given the law of motion of price level, numerically solve the problem (27)-(31) to obtain the optimal allocation of attention, $\vec{\kappa}^* \equiv (\kappa^*{}_m, \kappa^*{}_a)$.

STEP 4 Check whether $\left\|\vec{\kappa} - \vec{\kappa}^*\right\| < \varepsilon$, with $\varepsilon > 0$ and small. If this criterion is not satisfied, guess that $\vec{\kappa} = \vec{\kappa}^*$ and GO TO STEP 1. Otherwise, STOP.