A Likelihood Analysis of Models with Information Frictions[1] (Job Market Paper)

Leonardo Melosi[2]

University of Pennsylvania

November 19, 2009

Abstract

I develop and estimate a dynamic general equilibrium model with imperfectly informed firms in the sense of Woodford (2002). The model has two aggregate shocks: a monetary policy shock and a technology shock. Firms observe idiosyncratic noisy signals about these shocks and face strategic complementarities in price setting. In this environment, agents' "forecasting the forecasts of others" can produce realistic dynamics of model variables, with associated highly persistent real effects of monetary shocks and delayed effects of such shocks on inflation. The paper provides a full Bayesian analysis of the model, revealing that it can capture the persistent propagation of monetary shocks only by predicting that firms acquire less information about monetary policy than about technology. To investigate the plausibility of this finding, I augment the model to allow firms to optimally choose how much information to acquire about the two shocks, subject to an information-processing constraint à la Sims (2003). This constraint sets the rate at which firms can substitute pieces of information about the two shocks. I find that, in the estimated model, firms' marginal value of the information about monetary policy shocks is much higher than that about technology shocks. Hence, I argue that the estimated model predicts that 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.

[1] 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, Marco Del Negro, Cristina Fuentes-Albero, Matthieu Darracq Pariès, Max Kryshko, Bartosz Maćkowiak, Kristoffer Nimark, Ricardo Reis, and Mirko Wiederholt for very helpful comments. I would also thank participants at the Midwest Macro Meetings 2008, the Annual Congress of the European Economic Association 2008, the PENN Macro and Econometric Lunches, 2009 North American Summer Meeting of the Econometric Society, SED Annual Meeting 2009, and the 2009 European Meeting of the Econometric Society. Of course, all errors are my own.

[2] Correspondence to: Leonardo Melosi, University of Pennsylvania, Department of Economics, 160 McNeil, 3718 Locust Walk, Philadelphia, PA 19104-6297, USA. Email: lmelosi@sas.upenn.edu.

1 Introduction

A number of influential empirical studies of the U.S. economy have documented that money disturbances have highly persistent real effects and delayed impacts on inflation (Christiano, Eichenbaum, and Evans, 1999, Stock and Watson, 2001). The conventional approach to explaining this evidence relies upon sticky-price models (e.g., Galí and Gertler, 1999, Eichenbaum and Fisher, 2004, Christiano, Eichenbaum, and Evans 2005, and Smets and Wouters, 2007). These models can generally account for the highly persistent effects of monetary shocks only with sufficiently large costs of price adjustment. Such sizable costs imply a frequency of price adjustments that is inconsistent with some evidence from the micro-data on price changes (Bils and Klenow, 2004).[3]

Woodford (2002) proposes an alternative modeling approach: imperfect-common-knowledge models. In his price-setting model, monopolistic producers set their prices under limited information and strategic complementarities. Firms observe idiosyncratic noisy signals regarding the state of monetary policy and solve a signal-extraction problem in order to keep track of the model variables. Since the signal is noisy, firms do not immediately learn of the occurrence of monetary disturbances. As a result, the price level fails to adjust enough to entirely neutralize the real effects of nominal shocks. Moreover, because of the idiosyncratic nature of the signals, in the aftermath of a shock, firms are also uncertain about what other firms know that other firms know...that other firms know about that shock. Owing to strategic complementarities in price-setting, a problem of forecasting the forecasts of others of the type envisioned by Townsend (1983b) arises. This feature of the model has been shown to amplify the persistence in economic fluctuations (Townsend, 1983a, 1983b, Hellwig, 2002, Adam, 2008, Angeletos and La'O, 2008, and Lorenzoni, forthcoming A) and in the propagation of monetary disturbances to real variables and prices (Phelps, 1970, Adam,

[3] A modelling solution that preserves sticky prices and is not in conflict with micro-data on price-setting is developed by Altig, Christiano, Eichenbaum, and Linde, 2005.

2007, Gorodnichenko, 2008, and Lorenzoni, forthcoming B).[4] Moreover, it is worth emphasizing that in this model prices adjust frequently, but move only gradually to their complete information levels. Thus the resolution proposed by Woodford (2002) is appealing in that it can potentially explain sluggish adjustments of macro variables without necessarily being in discord with the micro evidence on price changes.

This paper addresses two questions. The first question is: can a version of the imperfect-common-knowledge model (ICKM) account for the persistent effects of monetary shocks we observe in the data? The answer to this question is yes but with one caveat. To get this answer, the paper proceeds by constructing a dynamic stochastic general equilibrium (DSGE) model with two shocks: a monetary policy shock and an aggregate technology shock. Firms receive one idiosyncratic noisy signal about each of these two shocks and face strategic complementarities in price-setting. The signal-extraction problem and the price-setting problem are similar to Woodford (2002). I estimate the ICKM and a vector autoregressive model (VAR) through Bayesian methods. I consider the impulse response functions (IRFs) implied by this statistical model as an accurate description of the propagation of monetary shocks in the data. From a Bayesian perspective, this conjecture is sensible because the VAR turns out to dominate the ICKM in terms of time series fit (Schorfheide, 2000). I then compare the IRFs of output and inflation to a monetary shock implied by the ICKM to those implied by the VAR. I find that the ICKM successfully captures the sluggish and hump-shaped response of output and inflation to monetary shocks implied by the VAR. Moreover, the estimated signal-to-noise ratio of monetary policy is smaller than that of technology by a factor of six. The reason is that the ICKM generates highly sluggish responses to monetary shocks only if firms acquire so little information about monetary policy.

This finding raises the second question: is it plausible that firms acquire so little information about monetary policy? The answer to this question is no. I reach this conclusion

[4] 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 sizable persistence.

by augmenting the model so as to allow firms to optimally choose the signal-to-noise ratios, subject to a constraint. This constraint is commonly used in the literature of rational inattention (Sims, 2003) and sets the rate at which firms can substitute pieces of information between the two shocks. I find that the firms' marginal value of the information about monetary shocks is much higher than that about technology shocks in the estimated model. Furthermore, when I solve for the optimal signal-to-noise ratios, firms find it optimal to acquire more information about monetary shocks than about technology shocks. These results suggest that the signal-to-noise ratio relative to monetary policy is implausibly small in the estimated ICKM.

This paper departs from Woodford (2002) in several 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 size of information frictions, which strongly influences the persistence in this type of models. Furthermore, Woodford (2002) closes his model by specifying an exogenous stochastic process that drives the nominal output.[5] 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 in Woodford (2002) to solve my model. 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.

Finally, the paper investigates what the imperfect-common-knowledge mechanism of generating persistence adds to or takes away from a more popular mechanism based on

[5] This approach is common in the macroeconomic literature of incomplete information. Examples are Lucas (1972),Mankiw and Reis (2002b), and Reis (2006b).

Calvo sticky prices (Calvo, 1983). To this end, I consider a model (henceforth, Calvo model) that differs from the ICKM in two main respects: (1) firms are perfectly informed, and (2) firms can re-optimize their prices only at random periods, as in Calvo (1983). I estimate the Calvo model through Bayesian techniques. First, I find that, unlike the ICKM and the VAR, the Calvo model fails to generate hump-shaped responses of output and inflation to monetary shocks. Second, the ICKM fits the data moderately better than the Calvo model.

This paper is related to the literature of rational inattention (Sims, 2003, 2006, Luo, 2008, Paciello, 2008, Van Nieuwerburgh and Veldkamp, 2008, Woodford, 2008, Maćkowiak, Moench, and Wiederholt 2009, and Maćkowiak and Wiederholt, 2009a, 2009b). Maćkowiak and Wiederholt (2009b) introduce a model in which firms optimally decide how much attention to pay to aggregate and idiosyncratic conditions, subject to a constraint on information flows. When they calibrate their model to match the average absolute size of price changes observed in the micro-data, they find that nominal shocks have sizable and persistent real effects.

The remainder of the paper proceeds as follows. Section 2 presents the ICKM and the Calvo model. Section 3 shows the Bayesian analysis of these two models and provides an answer to the first question of the paper. I address the second question of the paper in section 4. In section 5, I conclude.

2 The Model Economy

This section is organized as follows. First, I introduce the maintained assumptions of the ICKM. Second, I show the problems of agents in the model. Third, I discuss how to detrend and log-linearize the model around the deterministic steady state equilibrium. Fourth, I analyze the source of persistence in the log-linearized ICKM. Fifth, I briefly discuss the challenges one faces when solving models with imperfect common knowledge. Sixth, I describe how to modify the ICKM so that the information frictions are replaced with nominal rigidities à la Calvo (1983).

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). Households derive utility from consumption and disutility from supplying labor to firms. Furthermore, households face a cash-in-advance (CIA) constraint, requiring that they must have sufficient cash available before they can buy consumption goods. Firms set the prices of their intermediate goods in a monopolistic competitive market. Firms do not bear any cost when they change their prices and do not accumulate capital. Furthermore, there are two shocks: an aggregate technology shock and a monetary policy shock.

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. Third, firms face limitations on the overall amount of information they can process. As in Woodford (2002), information-processing frictions are modelled by assuming that firms do not observe past and current realizations of any model variables. They solely observe signals about the two shocks. 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}. Shocks and signals realize. 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 signals, set prices, hire labor service from households, and then produce. They use the liquidity facilities provided by the financial intermediary at the fixed fee τ so as 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 $M_{t-1} - D_t + W_t H_t$. The CIA constraint requires 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:

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

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:

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

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

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

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

$$\max_{\{C_t,H_t,M_t,D_t\}} \mathbb{E}_t \sum_{s=0}^{\infty} \beta^s \left[\ln C_{t+s} - \alpha \frac{H_{t+s}^{1+\eta}}{1+\eta} \right] \tag{2}$$

such that

$$P_t C_t \leq M_{t-1} - D_t + W_t H_t \tag{3}$$

$$0 \leq D_t \tag{4}$$

$$M_t = (M_{t-1} + W_t H_t - D_t - P_t C_t) + R_t D_t + \Pi_t + \Pi_t^b \tag{5}$$

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

The first constraint is the CIA constraint requiring that the household has to hold money up-front to finance its consumption. The second constraint prevents households from borrowing from the financial intermediary. The third constraint is the Dixit-Stiglitz aggregator of consumption varieties. The fourth constraint is the law of motion of households' money holdings.

2.4 The Financial Intermediary

The financial intermediary solves the trivial static problem:

$$\max_{\{L_t,D_t\}} (1 - R_t) D_t + X_t + \tau \cdot \mathbb{I} \{L_t > 0\} \tag{6}$$

such that

$$L_t \leq X_t + D_t \tag{7}$$

where L_t is the aggregate amount of liquidity supplied to firms $L_t = \int L_{i,t} di$, $X_t = M_{t+1} - M_t$ is the monetary injection, $\mathbb{I}\{\cdot\}$ is an indicator function that has the value one if the statement within curly brackets is true. τ is a fixed fee the intermediary receives from firms.

The financial intermediary lends cash to firms so that they can pay wages before households consume. This timing assumption allows households to use the cash from their current labor income to finance current consumption. This feature of the model makes the labor supply depend only on current variables and substantially simplifies the firms' signal-extraction problem. Replacing the fixed fee τ with an equilibrium interest rate would introduce forward-looking variables in the problem of firms and would unnecessarily complicate the signal-extraction problem.

2.5 The Monetary Authority

The monetary authority lets the money stock M_t grow at rate

$$\Delta \ln M_t = (1 - \rho_m) M_0 + \rho_m \Delta \ln M_{t-1} + \sigma_m \varepsilon_{m,t} \tag{8}$$

with $\varepsilon_{m,t} \backsim \mathcal{N}(0,1)$ and where Δ stands for the first-difference operator, 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 (8) can be interpreted as a simple monetary policy rule without feedbacks. The innovations $\varepsilon_{m,t}$ capture unexpected changes in the growth rate of money. Finally, it is useful to denote:

$$m_t \equiv \ln M_t - M_0 \cdot t \tag{9}$$

Finally, 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 value of intermediate-good firm i's profit conditional on the history of signals observed by firm i 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 numeraire 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 (see section 2.2):

$$Y_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\nu} Y_t \tag{12}$$

$N_{i,t}$ is the labor input demanded by firm i at time t. The production function is

$$Y_{i,t} = A_t N_{i,t}^{\phi} \tag{13}$$

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

$$\ln A_t = A_0 + \ln A_{t-1} + \sigma_a \varepsilon_{a,t} \tag{14}$$

$\varepsilon_{a,t} \backsim \mathcal{N}(0,1)$. The technology shocks, $\varepsilon_{a,t}$, are assumed to be orthogonal to monetary shocks, $\varepsilon_{m,t}$, at all leads and lags. I denote the loans of firm i at time t as $L_{i,t}$. Firms borrow liquidity from the financial intermediary in order to pay their nominal labor costs:

$$L_{i,t} = W_t N_{i,t} \tag{15}$$

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

$$\mathbf{z}_{i,t} = \begin{bmatrix} m_t \\ a_t \end{bmatrix} + \begin{bmatrix} \tilde{\sigma}_m & 0 \\ 0 & \tilde{\sigma}_a \end{bmatrix} \mathbf{e}_{i,t} \tag{16}$$

where $\mathbf{z}_{i,t} \equiv [z_{m,i,t}, z_{a,i,t}]'$, $a_t \equiv \ln A_t - A_0 \cdot t$, $\mathbf{e}_{i,t} \equiv [e_{m,i,t}, e_{a,i,t}]'$ and $\mathbf{e}_{i,t} \overset{iid}{\backsim} \mathcal{N}(\mathbf{0}, \mathbb{I}_2)$. Note that a_t and m_t are 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. After all, firms might learn about the state of monetary policy m_t from observing the signals concerning the state of technology a_t (i.e., $z_{a,i,t}$). I find, however, that relaxing this assumption of orthogonality of signals does not substantially affect the main predictions of the estimated model.

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 their expected current profits (11) subject to equations (12)-(16) 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.

I will log-linearize the price-setting equation around the deterministic steady state to simplify the signal-extraction issues. Furthermore, it is important to emphasize that I assume that at time 0 firms are endowed with an infinite sequence of signals, that is $\mathbf{z}_i^t = \{\mathbf{z}_{i,\tau}\}_{\tau=-\infty}^{t}$. This assumption simplifies the analysis in that firms will have the same Kalman gain matrix in their signal-extraction problem. Furthermore, this matrix can be shown to be time-invariant. This assumption makes the task of solving the model easier.

2.7 Detrending, Log-Linear Approximation

In the two models, the exogenous processes (8) and (14) 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 follows:

$$y_t \equiv \frac{Y_t}{A_t}, \quad p_{i,t} \equiv \frac{P_{i,t}}{P_t} \tag{17}$$

In order to log-linearize the model, I take the following steps. First, I derive the price-setting equation by solving firms' problem (11)-(16). Second, I transform the variables according to the definitions (17). 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 Source of Persistence in the ICKM

Let me introduce some notation. By convention, firm i's expectations of order zero about the state of monetary policy are the state itself, that is, $m_t^{(0)}(i) \equiv m_t$. Firm i's first-order expectations about the state of monetary policy are denoted by $m_{t|t}^{(1)}(i) \equiv \mathbb{E}\left[m_t|\mathcal{I}_t^i\right]$. Average first-order expectations about the state of monetary policy can be computed as follows $m_{t|t}^{(1)} \equiv \int m_{t|t}^{(1)}(i)\,di$. Firm i's second-order expectations are firm i's first-order expectations of the average first-order expectations, or more concisely $m_{t|t}^{(2)}(i) \equiv \mathbb{E}\left[m_{t|t}^{(1)}|\mathcal{I}_t^i\right]$. By rolling this argument forward I obtain the average j-th order expectation, for any $j \geq 0$,

$$m_{t|t}^{(j)} \equiv \int m_{t|t}^{(j)}(i)\,di \tag{18}$$

Moreover, firm i's $(j+1)$-th order expectations about the state of monetary policy, for any $j \geq 0$, are:

$$m_{t|t}^{(j+1)}(i) \equiv \mathbb{E}\left[m_{t|t}^{(j)}|I_t^i\right] \tag{19}$$

The speed of adjustment of variables to a shock is affected by the signal-to-noise ratio associated with that shock and the strategic complementarity in price-setting. The strategic complementarity in price-setting measures the extent to which firms want to react to the expected average price P_t. The degree of strategic complementarity turns out to be determined by $1-\lambda$, where $\lambda \equiv (\eta+1)\,\phi^{-1}/\left[\nu\left(\phi^{-1}-1\right)+1\right]$. See Appendix A.

In Appendix A, the law of motion of price level is:

$$\ln P_t = \left[\sum_{j=0}^{\infty}(1-\lambda)^j\,\lambda\left(m_{t|t}^{(j+1)}-a_{t|t}^{(j+1)}\right)\right]-\ln\bar{y}+M_0t-A_0t \tag{20}$$

where $m_{t|t}^{(j)}$ and $a_{t|t}^{(j)}$ are the average j-th order expectations about the state of monetary policy and technology at time t and $\bar{y}$ is the steady-state value of the detrended output, y_t. From equation (10) and equation (20) and after straightforward manipulations, it is easy to derive the law of motion of real output:

$$\ln Y_t = \left[m_t-\sum_{j=0}^{\infty}(1-\lambda)^j\,\lambda m_{t|t}^{(j+1)}\right]+\sum_{j=0}^{\infty}(1-\lambda)^j\,\lambda a_{t|t}^{(j+1)}-\ln\bar{y}+A_0t \tag{21}$$

Note that both price level and output are affected by weighted averages[6] of the infinite hierarchy of higher-order expectations about the exogenous states.

Equation (21) shows that monetary shocks have real effects as long as they are not fully anticipated by the average higher-order expectations of firms. More specifically, if the realization of m_t is common knowledge among firms, then $m_t^{(j)}=m_t$ for all j and the terms inside the square brackets cancel out. This shows that if monetary policy is common knowledge among firms, money is neutral in the model.

Equations (20)-(21) make it clear that the more sluggishly the weighted averages adjust to shocks, the more persistent the effects of shocks upon price and output are. The sluggishness of the weighted averages to shocks depends on the speed of adjustment of higher-order expectations. Sluggish adjustment of higher-order expectations depends on the signal-to-noise ratios that influence the precision of signals.[7] The more imprecise the signals are, the

[6] I restrict $\lambda \in (0,2)$ so that weights $(1-\lambda)^j\,\lambda$ are absolutely summable.

[7] Since, in the ICKM, firms observe two orthogonal signals, the speed of propagation may differ between the

more sluggishly the average expectations of every order will respond to shocks. Thus, the signal-to-noise ratios are a source of persistence in the model.

The strategic complementarity (i.e. $1 - \lambda$) influences the persistence of output and inflation by affecting the relative weights in the weighted averages of higher-order expectations. More precisely, the larger the strategic complementarity is, the bigger the weights of the average expectations of higher order are. The economic intuition is that the degree of strategic complementarity affects how strongly firms want to react to prices set by other firms. The stronger firms' reaction to other firms' pricing behavior is, the more they care about what other firms think that other firms think...about the exogenous state of the economy. In other words, strategic complementarity is the factor triggering the mechanism of forecasting the forecasts of others.

It is important to emphasize that the signal structure (16) implies that signals provide less and less information about expectations of higher and higher order. Therefore, the higher the order of average expectations, the more sluggishly they will adjust to shocks. Since larger strategic complementarity raises the weights associated with the average expectations of higher order in equations (20)-(21), it boosts the persistence of output and inflation responses to shocks. Thus, for any given degree of information incompleteness, strategic complementarity plays a crucial role in amplifying the persistence in the propagation of shocks.

2.9 Model Solution

When one characterizes rational expectation equilibria (REE) in models with incomplete information, a typical challenge is dealing with an infinite-dimensional state vector (*infinite regress*)[8] (Townsend, 1983b). The reason is that the laws of motion of infinitely many higher-order expectations have to be characterized in order to solve the model. This task is clearly

two shocks. Evidence that macroeconomic variables react at different speed to monetary and to technology shocks is documented in Paciello (2009).

[8] See Nimark (2009) for a thorough explanation of this problem.

unmanageable. In my ICKM, this problem arises when there is strategic complementarity in price-setting (i.e., $1-\lambda > 0$). Yet, here, this issue can be elegantly resolved as in Woodford (2002), since it is possible to re-define the state vector of the model as a weighted average of infinitely many higher-order expectations.[9] This leads to a state space of very small dimension. A detailed description of the method that numerically solves the model is in Appendix B. The solution method turns out to be fast and robust so that I can evaluate the likelihood at several points of the parameter space. This leads to accurate estimates of model parameters.

2.10 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_*. Moreover, as standard in models with sticky prices à la Calvo, I assume that the production function exhibits constant return to scale (i.e., $\phi = 1$). This implies that marginal costs are the same across firms. The problem of the firms that are allowed to re-optimize their prices at time t is:

[9] 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.

$$\max_{P_{i,t}} \mathbb{E}_t \sum_{s=0}^{\infty} \left[\theta_p^s \beta^{t+s} Q_{t+s|t} \left(\pi_*^s P_{i,t} - MC_{t+s}\right) Y_{i,t+s} - \tau_t \mathbb{I}\left\{L_{i,t} > 0\right\}\right] \tag{22}$$

such that

$$MC_t = \frac{W_t}{A_t},\ Y_{i,t+s} = \left(\frac{\pi_*^s P_{i,t}}{P_{t+s}}\right)^{-\nu} Y_{t+s} \tag{23}$$

where $Q_{t+s|t}$ is the marginal utility of a unit of the numeraire at time $t+s$ in terms of the utility of the representative household at time t, 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:

$$P_t^{1-\nu} = \left[\left(1-\theta_p\right) P_t^{*(1-\nu)} + \theta_p \ \left(\pi_* P_{t-1}\right)^{1-\nu}\right] \tag{24}$$

In the Calvo model, the speed of adjustment of variables to shocks is determined by the size of the Calvo parameter θ_p. Log-linearization and solution of the Calvo model is standard and hence omitted. I use the routine *gensys* developed by Sims (2002) to numerically solve this model.

3 Empirical Analysis

I fit the ICKM to observations on output and price level. I place a prior distribution on parameters and conduct Bayesian inference. First, I present the data set, the measurement equations, the prior distributions and the posterior distributions for model parameters. I then conduct a Bayesian evaluation of whether the ICKM provides an accurate description of the propagation mechanism of monetary shocks to output and inflation. To do that, I introduce a largely parameterized VAR model. I conjecture that if the response of output and inflation to monetary shocks implied by the ICKM is similar to the one implied by the VAR, then the ICKM provides an accurate description of the propagation of monetary disturbances. From a Bayesian perspective, this conjecture is sensible as long as the VAR model attains a higher posterior probability than the ICKM, as pointed out in Schorfheide (2000). I verify that this is indeed true by comparing the marginal data densities of the ICKM and the VAR.

Finally, I also estimate the Calvo model and compare the response of output and inflation to monetary policy shocks implied by this model with that of the ICKM. This comparison would allow me to assess what the ICK mechanism of generating persistence adds to or takes away from the more popular mechanism based on Calvo sticky prices.

3.1 The 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

Denote the U.S. per capita real GDP, and the U.S. GDP deflator as $\{Y_t,\ t = 1, 2, ...T\}$, and $\{P_t,\ t = 1, 2, ...T\}$, respectively. The measurement equations are given by equations (20)-(21).

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, 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. Nonetheless, I can estimate the parameter λ that affects the strategic complementarity in price-setting. Furthermore, the parameter, α, and the discount factor, β, drop out when I log-linearize the ICKM[10]. After log-linearization, the set of identifiable parameters in the ICKM is:

$$\Theta_I \equiv (\rho_m, A_0, M_0, \lambda, \sigma_m, \sigma_a, \tilde{\sigma}_m, \tilde{\sigma}_a) \tag{25}$$

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

Since I do not have data on the degree of strategic complementarity[11] and the parameter λ is very crucial for the persistence in the model (see section 2.8), I will set a broad prior for this parameter with the aim of learning its value from the likelihood function.

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 of monetary policy, ρ_m, the standard deviation of the monetary policy shock, σ_m, and the trend M_0 can be estimated by using presample observations of the (detrended) U.S. per capita real GDP and the (detrended) 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 of the standard deviation of the technology shock, σ_a, is centered at 0.007.

[10] See appendices A and B.

[11] There are studies (e.g., Rotemberg and Woodford, 1997) that quantify the degree of strategic complementarity in the U.S. However, they use a data set that is likely to be collinear to the one used in the paper. Using such information to formulate the prior would be controversial.

This value is the standard deviation of the Solow residual and is standard in the real-business cycle literature (Kydland and Prescott, 1986).

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.[12] In relative terms, these prior specifications are chosen so as to make the two signals equally informative about the corresponding exogenous state variables.[13]

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 section 2.8, these parameter values crucially influence the persistence in the model. Note that the prior median for $1-\lambda$ implies no strategic complementarity in price settings. However, note that this prior is quite broad. This implies that this crucial parameter value will be mainly learned from the likelihood. Finally, note that since the processes driving the two exogenous states are different, the prior medians for the signal-to-noise ratios do not have to be the same to make the firms equally informed about the two states.

As far as the log-linearized Calvo model is concerned, the parameter set is:

$$\Theta_C \equiv (\rho_m, A_0, M_0, \sigma_m, \sigma_a, \theta_p, \beta) \tag{26}$$

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. The prior for the Calvo parameter θ_p is centered at 0.67, implying an average duration of price contracts of three quarters. This value is regarded as consistent with the survey evidence discussed in Blinder, Canetti, Lebow, and Rudd (1998). The prior for the discount factor β is fixed so as to match the

[12] We 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).

[13] I quantify the amount of information that signals convey about the two exogenous states as in Sims (2003). The formal definition of this measure is provided in section 4.1.

long-run average real interest rate.

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 coefficient $(1-\lambda)$ controls the degree of strategic complementarity in price-setting. As shown in section 2.8, this coefficient is very important, since it affects the persistence of the impulse response functions (IRFs) of output and price level to shocks. The prior median of strategic complementarity $(1-\lambda)$ was set at 0. Bayesian updating points toward more strategic complementarity in price-setting. This amplifies the persistence in the mechanism of shock propagation for any finite values of the signal-to-noise ratios. Figure 1 compares the prior and the posterior distributions[14] for the strategic complementarity $(1-\lambda)$. It is apparent that the Bayesian updating clearly pushes the strategic complementarity toward a larger value than what is conjectured in the prior. The posterior median of λ is 0.41. This estimate is plausible. This number is consistent with a Frisch labor-supply elasticity, η, of 0.5 (Ríos-Rull, Schorfheide, Fuentes-Albero, Kryshko, and Santaeulàlia-Llopis, 2009), a technology parameter, ϕ, of 0.65 (Cooley and Prescott, 1995), and a mark-up of about 13% (Woodford, 2003 and Rotemberg and Woodford, 1997).

Moreover, the posterior median of the signal-to-noise ratio regarding the state of monetary policy, $\tilde{\sigma}_m/\sigma_m$, is large relative to that associated with the state of technology, $\tilde{\sigma}_a/\sigma_a$. The signal-to-noise ratio concerning the state of monetary policy is smaller by a factor of

[14] They are non-parametric estimates of the prior and posterior distributions based on the draws obtained from the simulator.

six.

As far as the Calvo model is concerned, the posterior median of the Calvo parameter θ_p implies that firms reset their prices about every two years. This frequency of price adjustments is implausible, according to the existing microeconometric analyses on price changes. Nonetheless, this result is not surprising. In fact, it is well-known that small-scale DSGE models with sticky prices à la Calvo can match the persistence of the macro-data only with price contracts of very long duration (Bils and Klenow, 2004). I might fix this problem by setting a tighter prior for the Calvo parameter, but I find that this would seriously undermine the fit of the Calvo model.

3.5 MDD-Based Comparisons

The paper addresses the question of whether the ICKM provides an accurate description of the propagation mechanism of monetary shocks to output and inflation. To do that, I estimate a largely parameterized VAR model and obtain its IRFs of output and inflation to monetary shocks. I then compare these IRFs to those of the estimated ICKM. In this comparison, the VAR IRFs work as a benchmark. From a Bayesian perspective, this comparison is sensible as long as the VAR model attains a higher posterior probability than the ICKM, as pointed out in Schorfheide (2000). In this section, I verify that this is indeed true by comparing the marginal data densities (MDDs) of the ICKM and the VAR (Kass and Raftery, 1995, Schorfheide 2000, and An and Schorfheide, 2007).

Let me denote the ICKM as $\mathcal{M}_I$ and the data used for estimation as $\tilde{Y}$. The MDD of the ICKM, $P\left(\tilde{Y}|\mathcal{M}_I\right)$, is:

$$P\left(\tilde{Y}|\mathcal{M}_I\right) = \int \mathcal{L}\left(\Theta_I|\tilde{Y},\mathcal{M}_I\right) p\left(\Theta_I|\mathcal{M}_I\right) d\Theta_I$$

where $\mathcal{L}(\cdot)$ stands for the likelihood function, and $p(\cdot|\cdot)$ denotes the posterior distribution, and Θ_I is the parameter set of the ICKM, as defined in section 3.3. I use Geweke's harmonic mean estimator (Geweke, 1999) to approximate the MDDs of the ICKM.

I consider a VAR(4):

$$\tilde{\mathbf{Y}}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \tilde{\mathbf{Y}}_{t-1} + \mathbf{\Phi}_2 \tilde{\mathbf{Y}}_{t-2} + \mathbf{\Phi}_3 \tilde{\mathbf{Y}}_{t-3} + \mathbf{\Phi}_4 \tilde{\mathbf{Y}}_{t-4} + \boldsymbol{\epsilon}_t \tag{27}$$

where $\tilde{\mathbf{Y}}_t = [\ln Y_t, \ln P_t]'$ and $\mathbf{\Sigma}_\epsilon \equiv \mathbb{E}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t')$. I fit this VAR(4) 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. Moreover, I obtain $100,000$ posterior draws through Gibbs sampling. To compute the MDD of the VAR model, I apply the method introduced by Chib (1995).

The log of the MDDs of the VAR and that of the ICKM are reported in Table 3. The VAR outperforms the ICKM in fitting the data. This result is not surprising, since the ICKM is very stylized compared to this statistical model. From a Bayesian perspective, this result legitimates the use of the VAR IRFs as a benchmark for studying whether the estimated ICKM can accurately explain the propagation of monetary shocks.

Moreover, I also compute the MDD of the Calvo model and report the result in Table 3. The ICKM has a larger MDD than the Calvo model. This implies that the ICKM fits the data better than the Calvo model. 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 problematic, since MDD-based comparisons penalize models for their number of 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). Note that 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 VAR, the ICKM, and the Calvo model are plotted in Figures 2 and 3, respectively. The size of the shock is normalized so that the reaction of variables upon impact 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 upon output and inflation.

The Calvo model does not seem to be well-suited to accounting for the hump-shaped pattern of the VAR response, whereas the ICKM appears to be 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 arises two quarters after the occurrence of the shock.

Furthermore, the VAR IRF emphasizes the presence of delayed effects of monetary shocks on inflation, which do not seem to be quite captured by the two DSGE models. The IRF of inflation implied by the VAR reaches its peak after four quarters, while, according to the ICKM, this happens after three quarters.

The estimated ICKM - albeit very stylized - successfully captures the persistent and hump-shaped response of output and inflation to monetary shocks implied by the broadly parameterized VAR. This leads me to conclude that the estimated ICKM provides an accurate description of the propagation mechanism of monetary shocks.

4 A Deeper Look at the Source of Persistence in the ICKM

As discussed in section 2.8, in the ICKM, the persistence of monetary shocks depends on the degree of strategic complementarity, $(1-\lambda)$, and the signal-to-noise ratio of monetary policy $\sigma_m/\tilde{\sigma}_m$. The larger the signal-to-noise-ratio is, the faster firms learn about the occurrence of a monetary shock, and then, *ceteris paribus,* the greater the speed of adjustment of variables to monetary shocks is. In section 3.4, I show that the posterior median for the signal-to-noise ratio of monetary policy is smaller than that of technology by a factor of six (see Table 2).

I then ask the following question: is it plausible that firms acquire so little information about monetary policy? To answer this question, I augment the ICKM with a signal-to-noise schedule that is extensively used in the literature of rational inattention (Sims, 2003). In this augmented ICKM, firms are allowed to choose the optimal signal-to-noise ratios concerning monetary policy and technology along that schedule. I calibrate the schedule to include the estimated signal-to-noise ratios of the ICKM. Then I compare the firms' marginal value of information on monetary policy relative to technology at the optimal and at the estimated signal-to-noise ratio. The more similar these marginal values are, the more plausible the estimated signal-to-noise ratios in the ICKM can be regarded under the lens of the rational-inattention theory.

In the next section I show how one can construct such a signal-to-noise schedule in the ICKM.

4.1 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,

$$\kappa \equiv H\left(m_t, a_t | z_{m,i}^{t-1}, z_{a,i}^{t-1}\right) - H\left(m_t, a_t | z_{m,i}^{t}, z_{a,i}^{t}\right) \tag{28}$$

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

and $z_{a,i}^{t}$. 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 $p\left(m_t | z_{1,i}^{\tau}\right)$ is the conditional probability density function of m_t.

Since signals and exogenous states are orthogonal, one can show that equation (28) can be re-written as

$$\kappa = \kappa_m + \kappa_a \tag{29}$$

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

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

Here, 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:

$$\kappa_m = g_m\left(\sigma_m, \tilde{\sigma}_m, \Upsilon\right), \quad \kappa_a = g_a\left(\sigma_a, \tilde{\sigma}_a\right) \tag{30}$$

where Υ is a vector of autocorrelations of m_t. The mapping g_a can be analytically derived, while the mapping g_m can be computationally approximated. See Appendix C.

For any given κ, σ_m, σ_a, and Υ, the signal-to-noise schedule is defined by equations (29) and (30). In other words, the signal-to-noise schedule is defined as 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

In period zero,[15] firms allocate their available attention[16] by solving:

$$\max_{\kappa_m,\kappa_a} \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t\left(\hat{p}_{i,t}^*, \hat{p}_t, \hat{y}_t, \hat{q}_t\right)|\mathbf{z}_i^t\right], \tag{31}$$

st

$$\ln P_{i,t}^* \equiv \mathbb{E}\left[(1-\lambda)\ln P_t + \lambda m_t - \lambda a_t|\mathbf{z}_i^t\right] \tag{32}$$

$$\mathbf{z}_{i,t} = \begin{bmatrix} m_t \\ a_t \end{bmatrix} + \begin{bmatrix} \tilde{\sigma}_m & 0 \\ 0 & \tilde{\sigma}_a \end{bmatrix} \mathbf{e}_{i,t} \tag{33}$$

$$\tilde{\sigma}_m = g_m^{-1}\left(\kappa_m, \sigma_m, \rho_m\right), \quad \tilde{\sigma}_a = g_a^{-1}\left(\kappa_a, \sigma_a\right) \tag{34}$$

$$\kappa_m + \kappa_a = \kappa, \text{ any t} \tag{35}$$

where $\hat{\pi}_t(\cdot)$ is the log-quadratic approximation of $Q_t\pi_t$, where π_t is the period profit function (11), $\hat{p}_{i,t}^* = \ln\left(P_{i,t}^*/P_t\right)$, $\hat{q}_t$ is the log deviations of $q_t = M_t Q_t$ from its value at the deterministic steady state, and $\mathbf{e}_{i,t} \overset{iid}{\backsim} \mathcal{N}(\mathbf{0}, \mathbb{I}_2)$. The model economy is assumed to be at its deterministic steady state in period 0. Moreover, I assume that firms have received an infinite sequence of signals at time 0. Note also that the mappings $g_m^{-1}(\cdot)$ and $g_a^{-1}(\cdot)$ in equation (34) are the inverse of the functions (30). The constraint (35) is the information-processing constraint and sets an upper-bound to the overall amount of information firms can process in every period t.

In this problem, firms decide how to allocate their overall available attention, which is

[15] Firms are not allowed to reconsider the 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' policies. See Maćkowiak and Wiederholt (2009b).

[16] 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 $t = 0$, the objective function of the allocation-of-attention problem can be shown to be the same across firms. See Maćkowiak and Wiederholt (2009b). 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 (35).

quantified by the parameter κ, between observing monetary policy and technology. Solving the allocation-of-attention problem (31)-(35) delivers the optimal allocation of attention (κ_m^*, κ_a^*). Note that when firms decide how to allocate their attention, they are aware that their choice will affect their optimal price-setting policy (32) in any subsequent periods.

The marginal rate of profit is defined as:

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

where Π is the sum of discounted profits:

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

It is very simple to see that the MRP at the optimal allocation of attention (κ_m^*, κ_a^*) is equal to unity. In the estimated ICKM, however, MRP may be different from one. The reason is that the estimated allocation of attention (κ_m, κ_a) 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, one finds that the MRP in the ICKM is 48.20. This number is hugely bigger than unity. In the estimated ICKM, firms are willing to trade more than 48 bits of information about technology to get one bit of information about monetary policy. This number is too big to reconcile itself to the rational-inattention theory. This result leads me to conclude that the estimated ICKM implies that firms acquire implausibly too little information about monetary policy.

4.3 A Robustness Check

By using tools provided by the rational-inattention theory, I find that firms acquire implausibly little information about monetary policy. Now the question is: does the ICKM model really need to make such an implausible prediction to match the persistent adjustment of variables to monetary shocks? To answer this question, I compare the impulse response functions of output and inflation to monetary shocks at the optimal allocation of attention, (κ_m^*, κ_a^*), and at the estimated allocation of attention, (κ_m, κ_a). The goal is to assess to what extent the persistence of output and inflation falls if firms are allowed to optimally choose their allocation of attention as modelled in the problem (31)-(35).

I will first compute the estimated information flows, (κ_m, κ_a), and the estimated overall amount of information processed, κ, in the ICKM. Given the mappings in (30) and the prior (posterior) draws for the parameter of the ICKM, Θ_I, I approximate the moments of the prior (posterior) distribution for the information flows κ_m and κ_a through standard Monte Carlo methods. Table 4 shows the prior and posterior medians for those parameters and their 90% credible intervals in the estimated ICKM. The posterior medians of κ_m and κ_a are 0.10 bits and 0.45 bits, respectively. The posterior median of the overall amount of information processed by firms per quarter, κ, is 0.55 bits.[17] Figure 4 compares the prior and the posterior distributions[18] of the fraction of the overall 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, according to the data, 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 overall attention paid to technology: $\kappa_a / (\kappa_m + \kappa_a)$.

Now I have to solve the problem (31)-(35) for the optimal allocation of attention (κ_m^*, κ_a^*). Yet, I need first to pin down the information-processing constraint (35). To do that, I need to fix one degree of freedom: the size of the parameter κ. I calibrate the value of this

[17] This is obtained by using the prior and posterior draws for κ_m and κ_a as long as equation (29).

[18] They are non-parametric estimates of the prior and posterior distributions based on the draws obtained from the simulator.

parameter by using its estimated value in Table 4, that is $\kappa = 0.55$ bits. I then solve[19] the problem (31)-(35) for the optimal allocation of attention and obtain that κ_m^* is equal to 0.33 and κ_a^* is equal to 0.22. These findings show that the estimated allocation of attention (κ_m, κ_a) (see Table 4) is very different from the optimal one (κ_m^*, κ_a^*). The optimal allocation of attention implies that firms pay more attention to monetary policy than to technology.

Figures 5-6 show the IRFs of output and inflation to a monetary shock implied by the ICKM at the estimated allocation of attention (EAA) and at the optimal allocation of attention (OAA). These figures also show the same IRFs implied by the benchmark VAR, analyzed in section 3. Output and inflation adjust very fast to monetary policy shocks at the optimal allocation of attention. This is not consistent with what is documented by the VAR. Hence, I conclude that the ICKM requires that firms acquire implausibly little information about monetary policy in order to generate the persistent propagation of monetary disturbances that is found in the data.

5 Concluding Remarks

I develop a DSGE model with imperfect common knowledge in the sense of Woodford (2002). The model features two aggregate shocks: a monetary policy shock and a technology shock. I obtain Bayesian estimates for the model parameters. I find that even though the model is very stylized, its impulse response functions of real output and inflation to a monetary policy shock closely match those implied by a largely parameterized VAR. Quite remarkably for such a stylized model, output and inflation react in a hump-shaped and persistent fashion to monetary shocks, as is widely documented by other influential empirical studies (e.g.,

[19] The optimal allocation of attention can be computed in four steps. First, I guess the values of the information flows κ_m and κ_a and use the mappings in (34) to obtain the implied noise variances, $\tilde{\sigma}_m$ and $\tilde{\sigma}_a$. Second, given this guess, I numerically characterize the law of motion of the price level exactly as I do when solving the ICKM (see section 2.9). Third, I numerically solve the problem (31)-(35) to obtain the optimal allocation of attention, k_m^* and κ_a^*. Fourth, I check whether $\|\overrightarrow{\kappa} - \overrightarrow{\kappa}^*\| < \varepsilon$, for vectors $\overrightarrow{\kappa} \equiv (\kappa_m, \kappa_a)'$ and $\overrightarrow{\kappa}^* \equiv (\kappa_m^*, \kappa_a^*)'$, with $\varepsilon > 0$ and small. If this criterion is not satisfied, I do another loop by setting the guess $\overrightarrow{\kappa} = \overrightarrow{\kappa}^*$. Otherwise, STOP.

Christiano *et al.*, 1999).

Nonetheless, I argue that the estimated model predicts that firms acquire little information about monetary policy shocks to an extent that is not plausible. I draw this conclusion from evaluating a simplified rational-inattention model à la Sims (2003). This model is an imperfect-common-knowledge model in which firms are allowed to choose the optimal information flows about the two shocks along a schedule that is commonly used in the literature of rational inattention. I show that the marginal value of information about monetary policy is much higher than that about technology at the point on the schedule predicted by the estimated imperfect-common-knowledge model. Furthermore, I find that the imperfect-common-knowledge model requires that firms acquire implausibly little information about monetary policy to generate the persistent propagation of monetary disturbances observed in the data. This result calls for further research on the substitution rate of information that firms actually face when they allocate their attention.

Yet I believe that it would be wrong to conclude that imperfect-common-knowledge models à la Woodford (2002) have no chance to generate persistent adjustments of model variables under plausible parameterizations. There is margin to modify the standard imperfect-common-knowledge model so as to make its predictions on firms' allocation of attention among shocks plausible and, at the same time, to retain the persistent propagation of monetary shocks. This is also left for future research.

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.

Altig, D., L. J. Christiano, M. Eichenbaum, and J. Linde (2005): "Firm-specific Capital, Nominal Rigidities and the Business Cycle," NBER Working Paper 11034.

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.

Bils, M., and P. J. Klenow (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112(5), 947–985.

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

Blinder, A. S., E. R. Canetti, D. E. Lebow, and J. B. Rudd (1998): *A New Approach to Understanding Price Stickiness*. Russell Sage Foundation, New York.

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 American Statistical Association*, 90(432), 1313–1321.

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.

Eichenbaum, M., and J. D. Fisher (2004): "Evaluating the Calvo Model of Sticky Prices," NBER Working Paper 10617.

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.

Galí, J., and M. Gertler (1999): "Inflation dynamics: A structural econometric analysis," *Journal of Monetary Economics*, 44(2), 195–222.

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.

Kass, R. E., and A. E. Raftery (1995): "Bayes Factors," *Journal of the American Statistical Association*, 430(90), 773–795.

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

Lorenzoni, G. (forthcoming A): "A Theory of Demand Shock," *American Economic Review*.

——— (forthcoming B): "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information," *Review of Economic Studies*.

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 (2009a): "Business Cycle Dynamics under Rational Inattention," Northwestern University Mimeo.

——— (2009b): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99(3), 769–803.

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.

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.

——— (2009): "Does Inflation Adjust Faster to Aggregate Technology Shocks than to Monetary Policy Shocks?," 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. Santaeulìlia-Llopis (2009): "Methods versus Substance: Measuring the Effects of Technology Shocks on Hours," NBER Working paper 15375.

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.

Sims, C. A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1-2), 1–20(21).

——— (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	$[0.18, 0.83]$
A_0	$\mathbb{R}$	Normal	0.00	$[-0.41, 0.41]$
M_0	$\mathbb{R}$	Normal	0.00	$[-0.41, 0.41]$
λ	$\mathbb{R}+$	Gamma	1.00	$[0.24, 1.74]$
$100\sigma_m$	$\mathbb{R}+$	InvGamma	1.6	$[0.44, 12.82]$
$100\sigma_a$	$\mathbb{R}+$	InvGamma	0.7	$[0.51, 0.87]$
$100\tilde{\sigma}_m$	$\mathbb{R}+$	InvGamma	5.02	$[2.11, 7.92]$
$100\tilde{\sigma}_a$	$\mathbb{R}+$	InvGamma	1.07	$[0.24, 1.87]$
θ_p	$[0,1)$	Beta	0.67	$[0.37, 0.99]$
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	0.00	$[-0.50, 0.63]$
$\sigma_m/\tilde{\sigma}_m$	signal-to-noise ratio MP	0.56	$[0.06, 3.64]$
$\sigma_a/\tilde{\sigma}_a$	signal-to-noise ratio tech.	0.95	$[0.10, 1.80]$

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	0.41	$[0.06, 0.77]$	0.80	$[0.13, 1.58]$
$100\sigma_m$	0.88	$[0.81, 0.91]$	0.89	$[0.82, 0.97]$
$100\sigma_a$	0.86	$[0.70, 1.02]$	2.66	$[2.04, 3.36]$
$100\tilde{\sigma}_m$	9.75	$[4.40, 15.01]$	–	–
$100\tilde{\sigma}_a$	1.45	$[0.60, 2.31]$	–	–
θ_p	–	–	0.88	$[0.82, 0.94]$
β	–	–	0.99	$[0.99, 0.99]$
$1 - \lambda$	0.64	$[0.25, 0.93]$	–	–
$\sigma_m/\tilde{\sigma}_m$	0.10	$[0.04, 0.15]$	–	–
$\sigma_a/\tilde{\sigma}_a$	0.63	$[0.33, 0.96]$	–	–

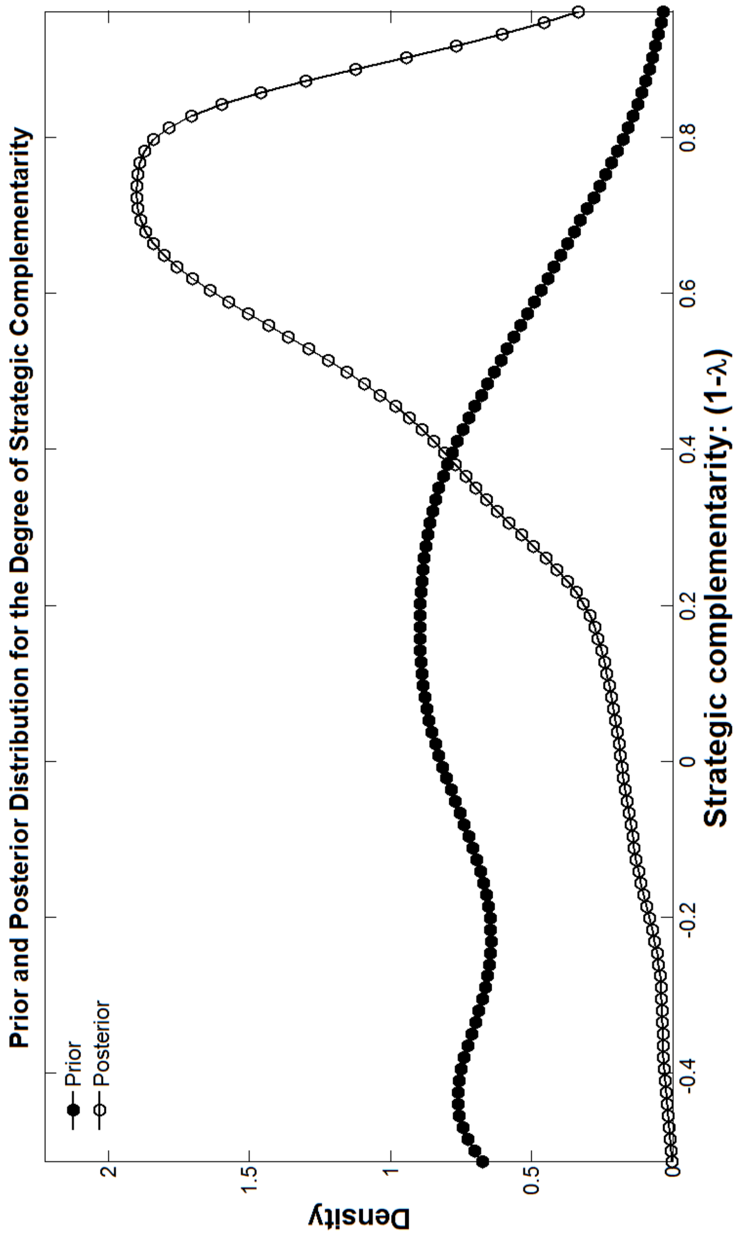
Table 3: Logarithms of Marginal Data Densities (MDDs)

	Models		
	ICKM	Calvo	VAR(4)
log MDD	1547.01	1529.38	1727.04

Table 4: Implied prior and posterior distributions

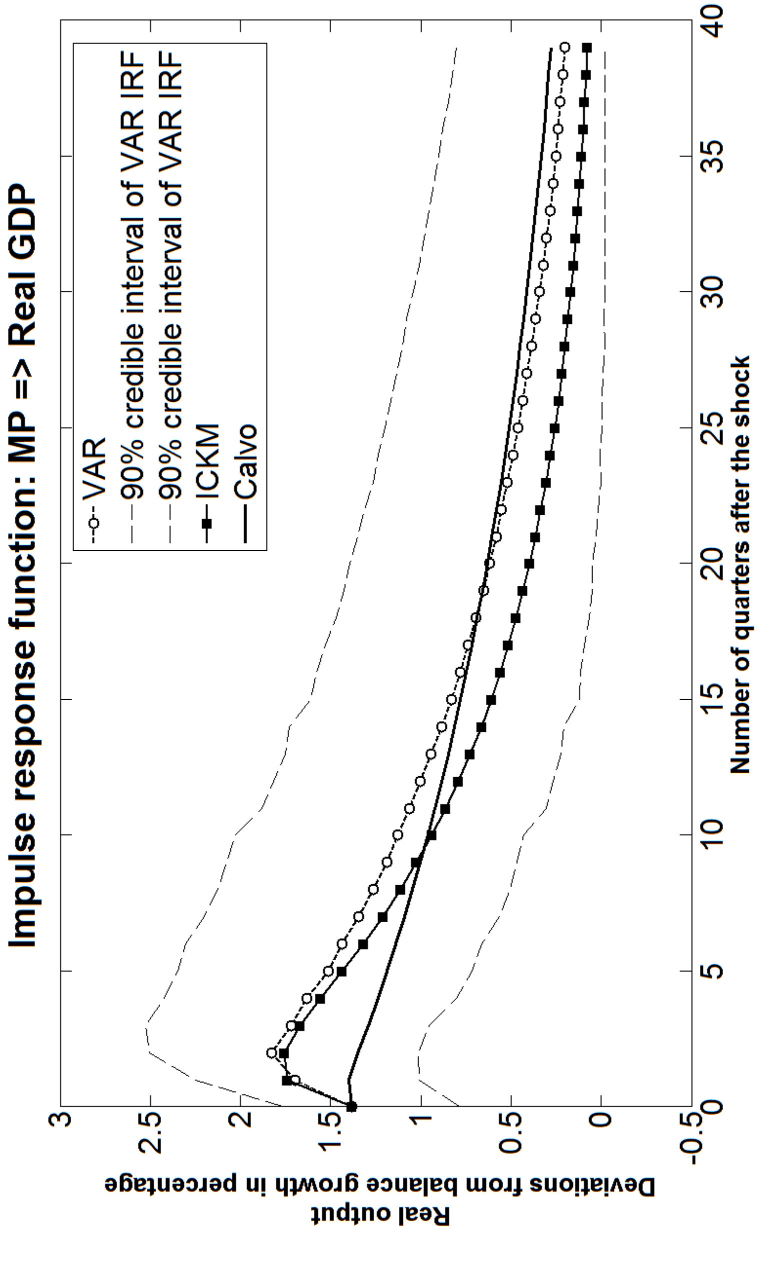
	PRIOR		
Variables	Descriptions	Median	90% Interval
κ_m	information flow MP	0.52	$[0.07, 2.51]$
κ_a	information flow tech.	0.67	$[0.12, 1.23]$
$\kappa = \kappa_m + \kappa_a$	overall level of attention	1.28	$[0.29, 3.20]$
$\frac{\kappa_a}{\kappa_m+\kappa_a}$	allocation of attention to tech.	0.53	$[0.11, 0.82]$

	POSTERIOR		
Variables	Descriptions	Median	90% Interval
κ_m	information flow MP	0.10	$[0.05, 0.18]$
κ_a	information flow tech.	0.45	$[0.22, 0.79]$
$\kappa = \kappa_m + \kappa_a$	overall level of attention	0.55	$[0.27, 0.94]$
$\frac{\kappa_a}{\kappa_m+\kappa_a}$	allocation of attention to tech.	0.82	$[0.76, 0.87]$



Prior and Posterior Distribution for the Degree of Strategic Complementarity
Prior
Posterior
Density
Strategic complementarity: (1-λ)
0
0.5
1
1.5
2
-0.4
-0.2
0
0.2
0.4
0.6
0.8

Figure 1



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
3
2.5
2
1.5
1
0.5
0
-0.5
0
5
10
15
20
25
30
35
40

Figure 2

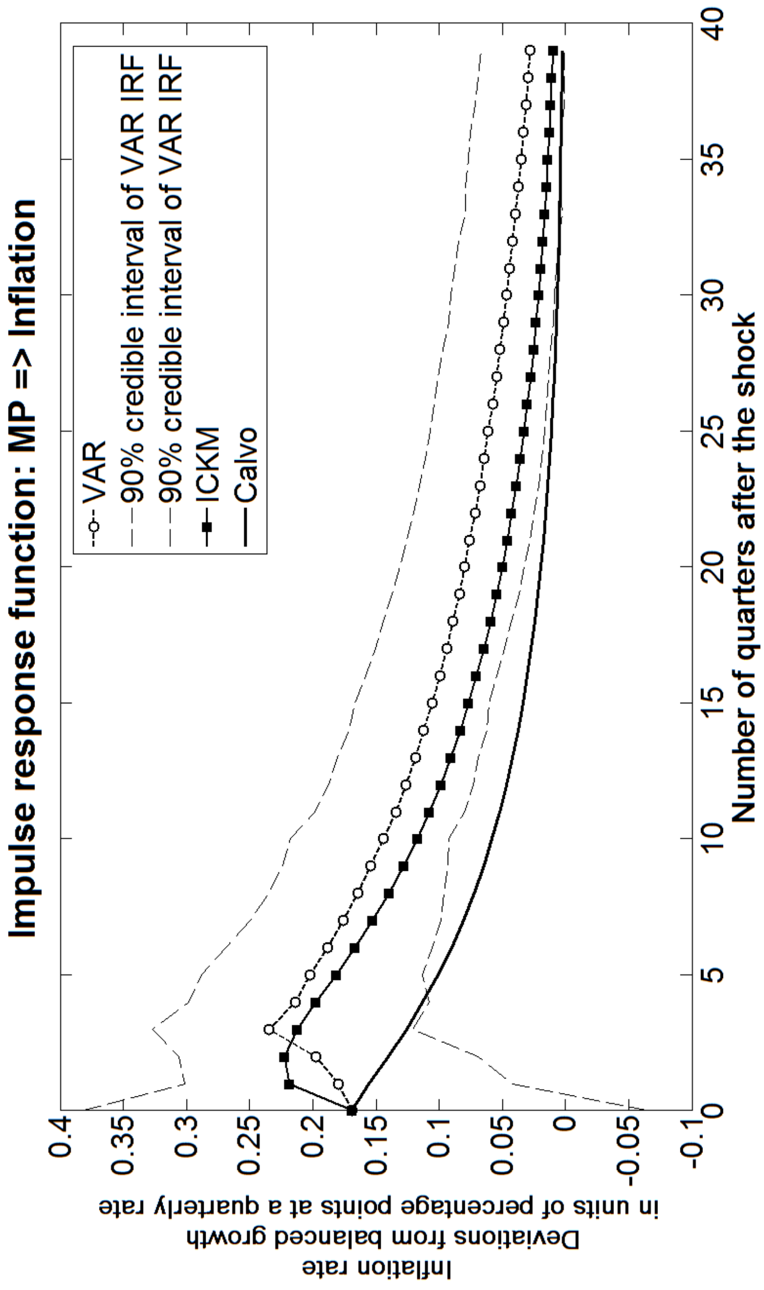
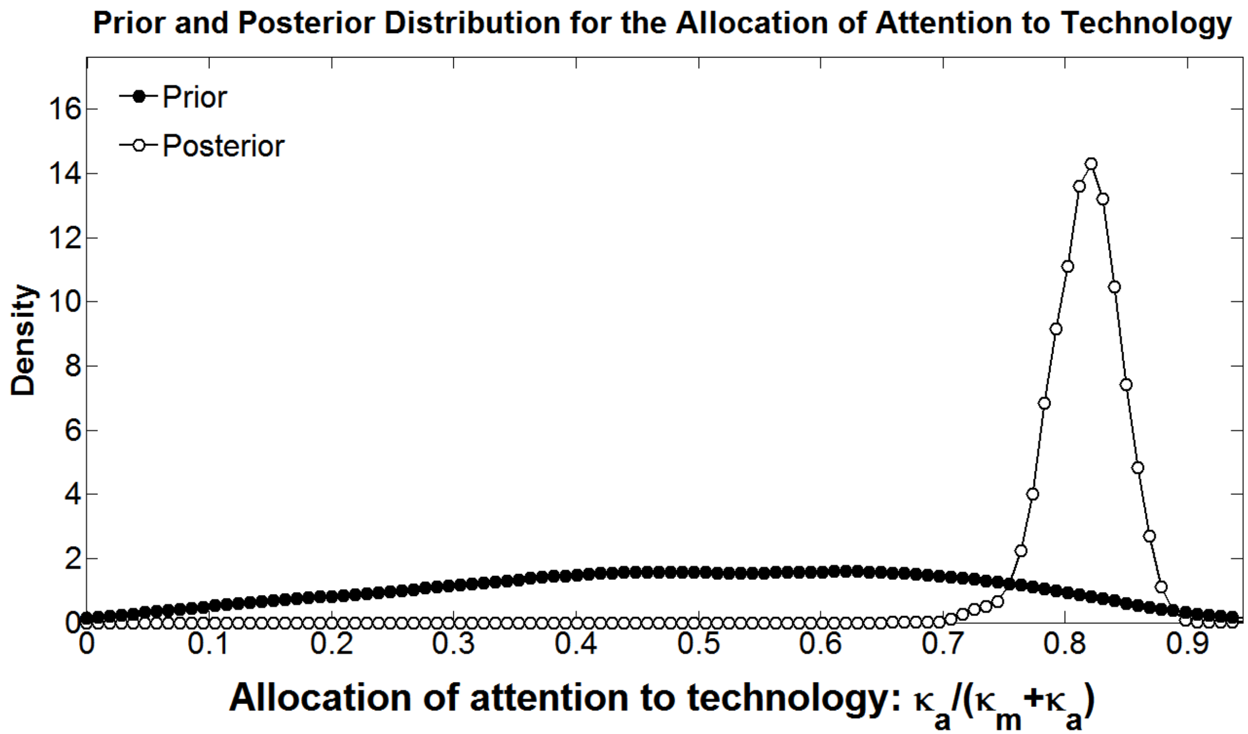

Impulse response function: MP => Inflation
VAR
90% credible interval of VAR IRF
90% credible interval of VAR IRF
ICKM
Calvo
Inflation rate
Deviations from balanced growth
in units of percentage points at a quarterly rate
Number of quarters after the shock
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1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

Figure 3



Prior and Posterior Distribution for the Allocation of Attention to Technology
Prior
Posterior
Density
16
14
12
10
8
6
4
2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
Allocation of attention to technology: κa/(κm+κa)

Figure 4

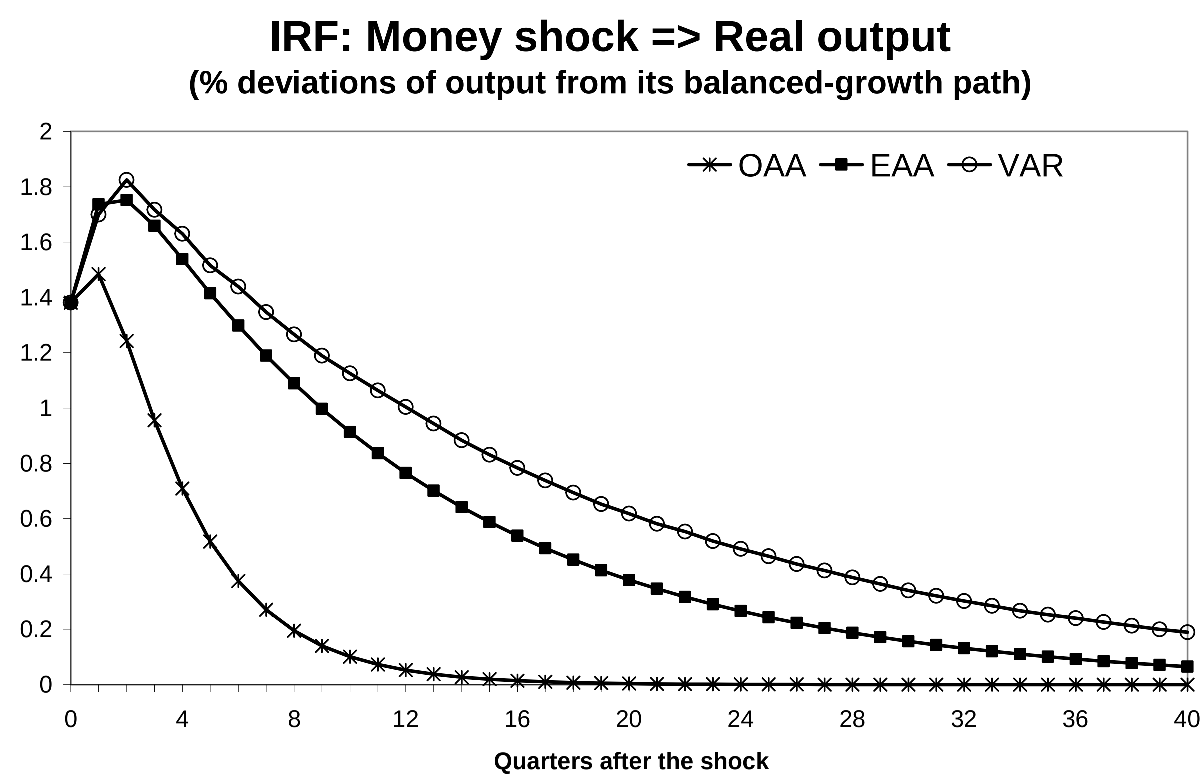

IRF: Money shock => Real output
(% deviations of output from its balanced-growth path)
OAA
EAA
VAR
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
4
8
12
16
20
24
28
32
36
40
Quarters after the shock

Figure 5

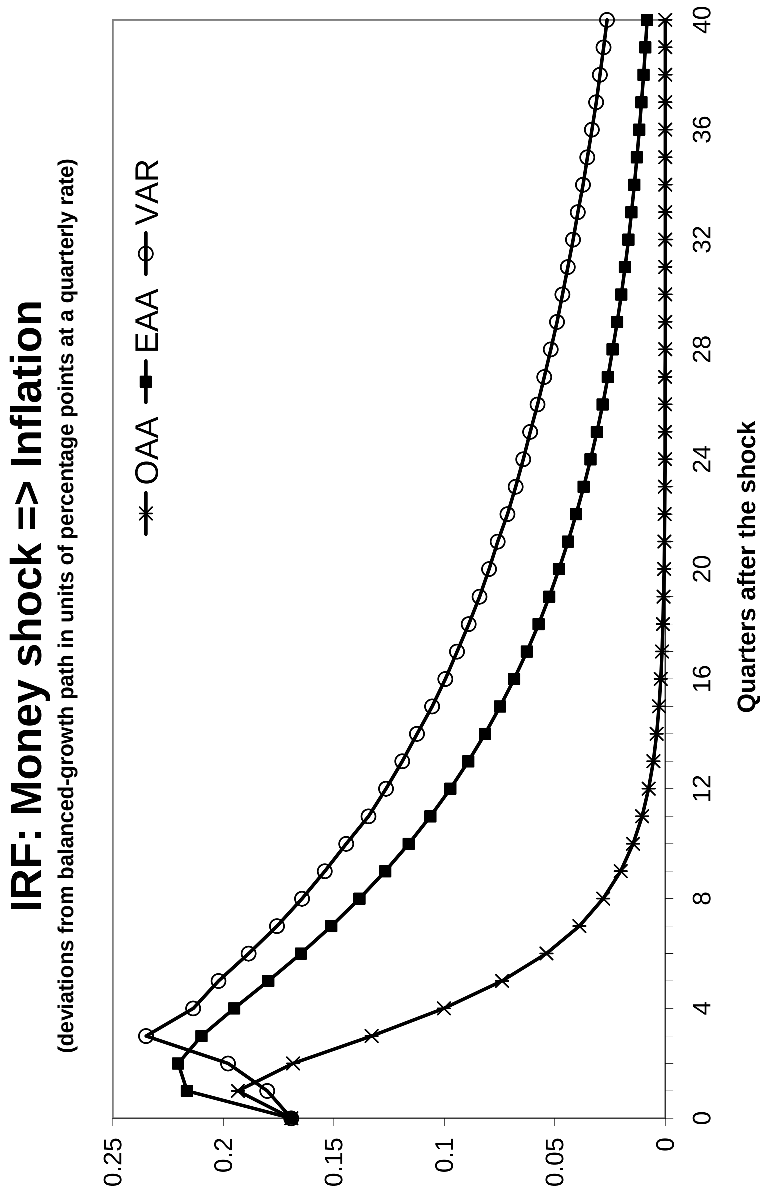

IRF: Money shock => Inflation
(deviations from balanced-growth path in units of percentage points at a quarterly rate)
OAA
EAA
VAR
0.25
0.2
0.15
0.1
0.05
0
0
4
8
12
16
20
24
28
32
36
40
Quarters after the shock

Figure 6

Appendix

A Law of motion of price and output in the ICKM

The first-order necessary condition[20] of the price-setting problem (11)-(16) 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 (12) to write:

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

From the solution to the representative household's problem (2)-(5), 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 yields:

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

Define the stationary variables:

$$y_t\equiv\frac{Y_t}{A_t},\ y_{i,t}\equiv\frac{Y_{i,t}}{A_t}\ ,\ p_{i,t}=\frac{P_{i,t}}{P_t},\ h_t=H_t \tag{36}$$

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[21] $\hat{h}_t=\phi^{-1}\hat{y}_t$. By substituting, this results into the equation above, one obtains:

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

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

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

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

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

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 (36) 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 coefficient $1-\lambda$ 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 (1) 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:

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

This is 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 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

$$\mathbf{F}_t \equiv \sum_{j=1}^{\infty} (1-\lambda)^{j-1} \lambda \mathbf{X}_t^{(j)} \quad (37)$$

$$\text{where } \mathbf{X}_t \equiv [m_t, m_{t-1}, a_t]' \quad (38)$$

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:

$$\widehat{y}_t = \widehat{p}_t \quad (39)$$

$$\widehat{p}_t = \mathbf{r}'\overline{\mathbf{X}}_t \quad (40)$$

$$\overline{\mathbf{X}}_t = \overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t \quad (41)$$

where

$$\overline{\mathbf{X}}_t \equiv \left[\mathbf{X}_t' \vdots \mathbf{F}_t'\right]', \ \mathbf{r} \equiv [-1, 0, 1, 1, 0, -1]'$$

$$\overline{\mathbf{B}} \equiv \begin{bmatrix} \mathbf{B}_{3x3} & \mathbf{0}_{3x3} \\ \mathbf{G}_{3x3} & \mathbf{H}_{3x3} \end{bmatrix}, \quad \overline{\mathbf{b}} = \left[\mathbf{b}\ ' \vdots \mathbf{d}'\right]' \quad (42)$$

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

$$\mathbf{u}_t \overset{iid}{\backsim} \mathcal{N}\left(\mathbf{0}, \mathbf{\Sigma}_u\right), \text{ for all } t \text{ and } \Sigma_u = \begin{bmatrix} \sigma_m^2 & 0 \\ 0 & \sigma_a^2 \end{bmatrix}$$

where $\mathbf{G}$, $\mathbf{H}$, and $\mathbf{d}$ are matrices that are not known yet. Equation (39) 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 (40) 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 (16). Thus, the firms' observation equations are

$$\mathbf{z}_{i,t} = \mathbf{D}\overline{\mathbf{X}}_t + \mathbf{e}_{i,t} \tag{43}$$

where

$$\mathbf{D} \equiv \left[\begin{array}{ccc} \mathbf{D}_1 & \vdots & \mathbf{0}_{2x3} \end{array} \right] \text{ and } \mathbf{D}_1 = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right] \tag{44}$$

$$\mathbf{e}_{i,t} \backsim \mathcal{N}\left(\mathbf{0}, \mathbf{\Sigma}_e\right), iid \text{ for all } t, \text{ and } i, \ \mathbf{\Sigma}_e = \left[\begin{array}{cc} \tilde{\sigma}_m^2 & 0 \\ 0 & \tilde{\sigma}_a^2 \end{array} \right] \tag{45}$$

Finding an equilibrium for this economy amounts to characterize the unknown matrices $\mathbf{G}$, $\mathbf{H}$, and $\mathbf{d}$. This requires solving the following fixed point problem. Given the conjectured law of motion (41), optimal firms' behaviors must exactly aggregate to the conjectured law of motion (41). Like in 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*

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

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:

$$\overline{\mathbf{X}}_{t|t-1}(i) = \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1}(i) \tag{47}$$

I can plug the (47) into the (46) to get the law of motion for firm i's estimate of the current state vector

$$\overline{\mathbf{X}}_{t|t}(i) = \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1}(i) + \mathbf{k}\left[\mathbf{z}_t(i) - \mathbf{D}\overline{\mathbf{X}}_{t|t-1}(i)\right] \tag{48}$$

By integrating the (48) over firms (i.e. $\int \overline{\mathbf{X}}_{t|t}(i)\, di \equiv \overline{\mathbf{X}}_{t|t}$) one gets

$$\overline{\mathbf{X}}_{t|t} = \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1} + \mathbf{k}\mathbf{D}\left[\overline{\mathbf{X}}_t - \overline{\mathbf{X}}_{t|t-1}\right] \tag{49}$$

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 (41) to get rid of $\overline{\mathbf{X}}_t$ in the equation (49) 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 (47), 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

$$\overline{\mathbf{X}}_{t|t} = [\mathbf{I} - \mathbf{kD}]\,\overline{\mathbf{B}\mathbf{X}}_{t-1|t-1} + \mathbf{kD}\left[\overline{\mathbf{B}\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t\right] \tag{50}$$

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

$$\boldsymbol{\varphi}'\overline{\mathbf{X}}_t^{(1)} = \mathbf{F}_t \tag{51}$$

It is easy to prove that equation (51) is indeed true by working as follows

$$\begin{aligned} \boldsymbol{\varphi}'\overline{\mathbf{X}}_t^{(1)} &= \left[(\lambda)\cdot \mathbf{I}_3 \vdots\ (1-\lambda)\cdot \mathbf{I}_3\right] \cdot \begin{bmatrix} \mathbf{X}_t^{(1)} \\ \cdots \\ \mathbf{F}_t^{(1)} \end{bmatrix} \\ \boldsymbol{\varphi}'\overline{\mathbf{X}}_t^{(1)} &= \lambda \mathbf{X}_t^{(1)} + (1-\lambda)\,\mathbf{F}_t^{(1)} \end{aligned} \tag{52}$$

Let me introduce the following notations:

$$x_{t|t}^{(k-1)} \equiv x_t^{(k)}, \quad \forall k \geq 1; \quad x_t^{(0)} \equiv x_t \tag{53}$$

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 (37)

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

Combining the last two equations yields

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

Some easy manipulations lead to

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

Now recall equation (53) to finally write

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

Comparing this equation with the (37) concludes the proof of (51). Now one can plug equation (50) into equation (51) to get

$$\mathbf{F}_t = \left[\varphi' - \widetilde{\mathbf{k}}\mathbf{D}\right]\overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1|t-1} + \widetilde{\mathbf{k}}\mathbf{D}\left[\overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t\right] \tag{54}$$

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

FACT 1

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

FACT 2

$$\begin{aligned}\mathbf{D}\overline{\mathbf{B}} &= \left[\mathbf{D}_1\mathbf{B} \vdots \mathbf{0}_{2x3}\right] \\ &= \left[\mathbf{B}^{\dagger} \vdots \mathbf{0}_{2x3}\right]\end{aligned} \tag{55}$$

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 $\mathbf{B}$.

FACT 3

$$\begin{aligned}\mathbf{D}\overline{\mathbf{b}} &= \mathbf{D}_1\mathbf{b} \\ &= \left[\begin{array}{cc}1 & 0\\ 0 & 1\end{array}\right] = \mathbf{I}\,(2)\end{aligned}$$

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:

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

and

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

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

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

By collecting all these results one can rewrite equation (54) as follows

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

Next, I will work out the vector $\mathbf{F}_{t-1}$ from $\mathbf{F}_{t-1|t-1}$, since I want to rewrite equation (56) in a form that is comparable to that conjectured in equation (41) so as I can compare my initial guess. One should start from equation (51) 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

$$(1-\lambda)\cdot\mathbf{F}_{t-1|t-1} = \mathbf{F}_{t-1} - \lambda\mathbf{X}_{t-1|t-1} \tag{57}$$

I can now plug equation (57) into equation (56) to get

$$\begin{aligned}
\mathbf{F}_t &= \left[\lambda\mathbf{B} + (1-\lambda)\,\mathbf{G} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\right]\mathbf{X}_{t-1|t-1} + \mathbf{H}\left[\mathbf{F}_{t-1} - \lambda\mathbf{X}_{t-1|t-1}\right] + \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\mathbf{X}_{t-1} + \widetilde{\mathbf{k}}\mathbf{u}_t \\
\mathbf{F}_t &= \left[\lambda\mathbf{B} + (1-\lambda)\,\mathbf{G} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} - \lambda\mathbf{H}\right]\mathbf{X}_{t-1|t-1} + \mathbf{H}\cdot\mathbf{F}_{t-1} + \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\mathbf{X}_{t-1} + \widetilde{\mathbf{k}}\mathbf{u}_t
\end{aligned} \tag{58}$$

Now equation (58) has the same form as the bottom rows of equation (41) 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:

$$\mathbf{G} = \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} \tag{59}$$

$$\mathbf{d} = \widetilde{\mathbf{k}} \tag{60}$$

and

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

By substituting (59) into the last equation one obtains

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

$$\mathbf{H} \overset{!}{=} \mathbf{B} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} \tag{61}$$

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

$$\mathbf{k} = \mathbf{P}\mathbf{D}'\left[\mathbf{D}\mathbf{P}\mathbf{D}' + \mathbf{\Sigma}_e\right]^{-1} \tag{62}$$

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

$$\mathbf{P} = \overline{\mathbf{B}}\left[\mathbf{P} - \mathbf{P}\mathbf{D}'\left[\mathbf{D}\mathbf{P}\mathbf{D}' + \mathbf{\Sigma}_e\right]^{-1}\mathbf{D}\mathbf{P}\right]\overline{\mathbf{B}}' + \overline{\mathbf{b}}\mathbf{\Sigma}_u\overline{\mathbf{b}}' \tag{63}$$

and 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 $\mathbf{B}$.

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 (59)-(61). Then one has to solve the algebraic Riccati equation (63) for $\mathbf{P}$ and obtain a new Kalman-gain matrix $\mathbf{k}^*$ through the equation (62). 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 (41)-(42) through (59)-(63), which combined with the equations (39)-(40) 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:

$$\kappa_a \equiv H\left(a_t|z_{a,i}^{t-1}\right) - H\left(a_t|z_{a,i}^{t}\right) \tag{64}$$

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

$$H\left(a_t|z_{a,i}^{t}\right) \equiv \frac{1}{2}\log_2\left[2\pi e \cdot VAR\left(a_t|z_{a,i}^{t}\right)\right] \tag{65}$$

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,

$$\tilde{\sigma}_a^2 = \frac{\left[2VAR\left(a_t|z_{a,i}^t\right) + \sigma_a^2\right]^2}{4\sigma_a^2} - \frac{\sigma_a^2}{4} \tag{66}$$

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 (64) and (65) 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:

$$VAR\left(a_t|z_i^t\right) = \frac{\sigma_a^2}{2^{2\kappa_a} - 1} \tag{67}$$

Plugging this result into equation (66) leads to:

$$\kappa_a = \frac{1}{2}\log_2\left[\frac{1}{\left(\frac{\tilde{\sigma}_a^2}{\sigma_a^2} + \frac{1}{4}\right)^{\frac{1}{2}} - \frac{1}{2}} + 1\right] \tag{68}$$

This is the mapping g_a in equation (30).

An analytical closed-form solution for the mapping g_m in equation (30) 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:

$$H\left(m_t|z_{1,i}^{\tau}\right)=\frac{1}{2}\log_2\left[2\pi e\cdot VAR\left(m_t|z_{1,i}^{\tau}\right)\right] \tag{69}$$

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 (63). The matrix $\mathbf{P}_{t|t}$ is defined as:

$$\mathbf{P}_{t|t}\equiv\mathbf{P}_{t|t-1}-\mathbf{P}_{t|t-1}\mathbf{D}'\left[\mathbf{D}\mathbf{P}_{t|t-1}\mathbf{D}'+\mathbf{\Sigma}_e\right]^{-1}\mathbf{D}\mathbf{P}_{t|t-1} \tag{70}$$

where the matrices $\mathbf{D}$ and $\mathbf{\Sigma}_e$ have been defined in (44) and in (45), respectively.

Thus, after one has characterized the fixed point as discussed in appendix B, one can use the resulting matrix $\mathbf{P}$ and equation (70) 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 (69), and finally the information flow $\kappa_m\equiv H\left(m_t|z_{m,i}^{t-1}\right)-H\left(m_t|z_{m,i}^{t}\right)$.