ESSAYS ON THE MACROECONOMICS OF INCOMPLETE INFORMATION

Leonardo Melosi

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2010

Supervisors of Dissertation:

Frank Schorfheide

Professor of Economics

Co-Supervisor

Francis X. Diebold

Professor of Economics

Graduate Group Chairperson

Dirk Krueger, Professor of Economics

Dissertation Committee

Francis X. Diebold, Professor of Economics
Jesús Fernández-Villaverde, Associate Professor of Economics
Dirk Krueger, Professor of Economics
Frank Schorfheide, Professor of Economics

Essays on the Macroeconomics of Incomplete Information

A Francesca.

ACKNOWLEDGMENTS

I am indebted to Frank Schorfheide, Frank Diebold, Jesús Fernández-Villaverde, and Dirk Krueger for the valuable advice they have provided at each stage of this paper. I am also very grateful to Harold Cole for very helpful comments and guidance. I thank Dario Caldara, Matthieu Darracq Pariès, Marco Del Negro, Cristina Fuentes-Albero, Christian Hellwig, Max Kryshko, Bartosz Maćkowiak, 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, CEMMAP (UCL) Applied Macroeconomics and macroeconometrics Workshop. I also wish to thank the seminar participants at UPF, UCL, LBS, UCLA, Federal Reserve Bank of Richmond, Federal Reserve of Dallas, IMF, Federal Reserve Board of Governors, Bocconi, EIEF, Banca d'Italia, LSE, and Boston University. Of course, all errors are my own.

ABSTRACT

ESSAYS ON THE MACROECONOMICS OF

INCOMPLETE INFORMATION

Leonardo Melosi

Frank Schorfheide and Francis X. Diebold

In the first chapter, 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 further investigate 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. I argue that this finding admits two alternative interpretations. First, firms acquire implausibly too little information about the monetary shock in the estimated model. Second, the rate of substitution implied by the information-processing constraint is inconsistent with the data. In the third chapter, I develop a model where firms have incomplete and dispersed information to study how monetary policy affects agents' beliefs. I estimate the model through Bayesian methods and find that dispersed information has two main implications for monetary policy. First, it reduces the real effects of money. Second, it raises the output loss associated with a monetary policy of disinflation.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

I conduct a formal econometric analysis of dynamic stochastic general equilibrium models where firms have incomplete information. In the first chapter, I develop and estimate a dynamic general equilibrium model with imperfectly informed firms in the sense of Woodford (2002). I provide an econometric investigation of whether the model can be reliably used for studying the propagation of monetary disturbances to macroeconomic variables in the US economy. 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. I provide a full Bayesian analysis of the model, revealing that, although it is highly-stylized, it provides a useful laboratory for studying the effects of monetary shocks. Furthermore, I find that firms acquire much less information about monetary policy than that about technology in the estimated model.

In the second chapter, I investigate the plausibility of the finding that firms acquire much less information about monetary policy than that about technology. To this end, 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 that is widely used in the literature of rational inattention propelled by Sims (1998, 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. This result calls for further research on the substitution rate of information that firms actually face.

The third chapter addresses the following question: does monetary policy play a role in coordinating agents' inflation expectations? To answer this question the essay develops a dynamic stochastic general equilibrium (DSGE) model where firms have heterogenous inflation expectations. In the model, monetary policy can work as a focal point that helps coordinating expectations among firms. I will perform Bayesian estimation and evaluation of the model. The main findings of this chapter are two. First, the presence of imperfect information reduces the real effects of monetary policy. Second, the output loss associated with a monetary policy of disinflation (i.e., a negative shock to central bank's inflation targeting) is larger when one considers imperfect information.

Chapter 2

A Model of Incomplete Common Knowledge

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).[1]

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, 2007, Gorodnichenko, 2008, and Lorenzoni, forthcoming B).[2] Moreover, it is worth emphasizing that in this model prices adjust fre-

[1] 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).

[2] 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

quently, 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 chapter addresses the question: 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 that will be addressed in the second chapter. To get this answer, the essay 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

generate sizable persistence.

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.

Finally, this chapter investigates what the imperfect-common-knowledge mechanism of generating persistence adds to or takes away from a more popular mechanism based on Calvo sticky prices (Calvo, 1983). To this end, I consider a model (henceforth, Calvo model) that differs from the ICKM in 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.

2.1 The Model Economy

This chapter 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).

2.1.1 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.

2.1.2 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

$$P_t = \left(\int (P_{i,t})^{1-\nu} \, di \right)^{\frac{1}{1-\nu}}. \tag{2.1}$$

2.1.3 The Representative Household

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

$$\max_{\{C_t,H_t,M_t,D_t\}} \mathbb{E}_t \sum_{s=0}^{\infty} \beta^s \left[\ln C_{t+s} - \alpha \frac{H_{t+s}^{1+\eta}}{1+\eta} \right] \tag{2.2}$$

such that

$$P_t C_t \leq M_{t-1} - D_t + W_t H_t \tag{2.3}$$

$$0 \leq D_t \tag{2.4}$$

$$M_t = (M_{t-1} + W_t H_t - D_t - P_t C_t) + R_t D_t + \Pi_t + \Pi_t^b \tag{2.5}$$

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.

2.1.4 The Financial Intermediary

The financial intermediary solves the trivial static problem:

$$\max_{\{L_t,D_t\}} (1 - R_t) D_t + X_t + \tau \cdot \mathbb{I}\{L_t > 0\} \tag{2.6}$$

such that

$$L_t \leq X_t + D_t \tag{2.7}$$

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.

2.1.5 The Monetary Authority

The monetary authority lets the money stock M_t grow at rate

$$\Delta \ln M_t = (1 - \rho_m) M_0 + \rho_m \Delta \ln M_{t-1} + \sigma_m \varepsilon_{m,t} \tag{2.8}$$

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 (2.8) 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:

$$m_t \equiv \ln M_t - M_0 \cdot t \tag{2.9}$$

Finally, market clearing for the monetary market requires that:

$$\ln M_t = \ln Y_t + \ln P_t \tag{2.10}$$

2.1.6 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:

$$\mathbb{E}\left[\beta^t Q_t \left(P_{i,t} Y_{i,t} - W_t N_{i,t} - \tau \mathbb{I}\left\{L_{i,t}\right\}\right) | \mathbf{z}_i^t\right] \tag{2.11}$$

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:

$$Y_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\nu} Y_t \tag{2.12}$$

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

$$Y_{i,t} = A_t N_{i,t}^{\phi} \tag{2.13}$$

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

$$\ln A_t = A_0 + \ln A_{t-1} + \sigma_a \varepsilon_{a,t} \tag{2.14}$$

$\varepsilon_{a,t} \backsim \mathcal{N}(0,1)$. 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:

$$L_{i,t} = W_t N_{i,t} \tag{2.15}$$

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

$$\mathbf{z}_{i,t} = \begin{bmatrix} m_t \\ a_t \end{bmatrix} + \begin{bmatrix} \tilde{\sigma}_m & 0 \\ 0 & \tilde{\sigma}_a \end{bmatrix} \mathbf{e}_{i,t} \tag{2.16}$$

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 (2.11) subject to equations (2.12)-(2.16) 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.

2.1.7 Detrending, Log-Linear Approximation

The exogenous processes (2.8) and (2.14) 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:

$$y_t \equiv \frac{Y_t}{A_t}, \quad p_{i,t} \equiv \frac{P_{i,t}}{P_t} \tag{2.17}$$

In order to log-linearize the model, I take the following steps. First, I derive the price-setting equation by solving firms' problem (2.11)-(2.16). Second, I transform the variables according to the definitions (2.17). 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 (2.10).

2.1.8 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 $m_{t|t}^{(1)}(i) \equiv \mathbb{E}[m_t|\mathcal{I}_t^i]$. 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 $j \geq 0$,

$$m_{t|t}^{(j)} \equiv \int m_{t|t}^{(j)}(i)\, di \tag{2.18}$$

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

$$m_{t|t}^{(j+1)}(i) \equiv \mathbb{E}\left[m_{t|t}^{(j)}|I_t^i\right] \tag{2.19}$$

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:

$$\ln P_t = \left[\sum_{j=0}^{\infty} (1-\lambda)^j \lambda \left(m_{t|t}^{(j+1)} - a_{t|t}^{(j+1)}\right)\right] - \ln \bar{y} + M_0 t - A_0 t \qquad (2.20)$$

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 (2.10) and equation (2.20) and after straightforward manipulations, it is easy to derive the law of motion of real output:

$$\ln Y_t = \left[m_t - \sum_{j=0}^{\infty} (1-\lambda)^j \lambda m_{t|t}^{(j+1)}\right] + \sum_{j=0}^{\infty} (1-\lambda)^j \lambda a_{t|t}^{(j+1)} - \ln \bar{y} + A_0 t \quad (2.21)$$

Note that both price level and output are affected by weighted averages[3] of the infinite hierarchy of higher-order expectations about the exogenous states.

Equation (2.21) 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.

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

This shows that if monetary policy is common knowledge among firms, money is neutral in the model.

Equations (2.20)-(2.21) 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.[4] The more imprecise the signals are, 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

[4] Since, in the ICKM, firms observe two orthogonal signals, the speed of propagation may differ between the two shocks. Evidence that macroeconomic variables react at different speed to monetary and to technology shocks is documented in Paciello (2009).

mechanism of forecasting the forecasts of others.

It is important to emphasize that the signal structure (2.16) 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 (2.20)-(2.21), 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.

2.1.9 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*)[5] (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 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 av-

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

erage of infinitely many higher-order expectations.[6] 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.

2.1.10 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_*. The problem of the firms that are allowed to re-optimize their prices at time t is:

[6] 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.

$$\max_{P_{i,t}} \mathbb{E}_t \sum_{s=0}^{\infty} \left[\theta_p^s \beta^{t+s} Q_{t+s|t} \left(\pi_*^s P_{i,t} - MC_{t+s}\right) Y_{i,t+s} - \tau \mathbb{I}\left\{L_{i,t} > 0\right\}\right] \tag{2.22}$$

such that

$$Y_{i,t+s} = \left(\frac{\pi_*^s P_{i,t}}{P_{t+s}}\right)^{-\nu} Y_{t+s} \tag{2.23}$$

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:

$$P_t^{1-\nu} = \left[\left(1-\theta_p\right) P_t^{*(1-\nu)} + \theta_p \ \left(\pi_* P_{t-1}\right)^{1-\nu}\right] \tag{2.24}$$

In the Calvo model, the speed of adjustment of variables to shocks is determined by the size of the Calvo parameter θ_p and the strategic complementarity parameter, λ. I detrend the non-stationary variables and log-linearize the model around the deterministic steady state. I obtain the standard New Keynesian Phillips curve, whose slope, κ_{pc}, depends on a function of parameters: $\kappa_{pc} = \left(1-\theta_p\right)\left(1-\theta_p \beta\right) \lambda / \theta_p$.

2.2 Empirical Analysis

I fit the ICKM to observations on output and price level. I place a prior distribution on parameters and conduct Bayesian inference. 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.

2.2.1 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*).

2.2.2 Measurement Equations

Denote the U.S. per capita real GDP, and the U.S. GDP deflator as Y_t, and P_t, respectively. The measurement equations are given by equations (2.20)-(2.21).

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.

2.2.3 Prior Distributions

Given the observables presented in section 2.2.1, 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[7]. After log-linearization, the set of identifiable parameters in the ICKM is:

$$\Theta_I \equiv (\rho_m, A_0, M_0, \lambda, \sigma_m, \sigma_a, \tilde{\sigma}_m, \tilde{\sigma}_a) \tag{2.25}$$

Table 2.1 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[8] and the parameter λ is very crucial for the persistence in the model (see section 2.1.8), I will set a broad prior for this parameter with the aim of learning its value from the likelihood function. If the Frisch labor-supply elasticity, η, is equal to

[7] See appendices A and B.

[8] 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.

Table 2.1: Prior Distributions

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

0.5 (Ríos-Rull, Schorfheide, Fuentes-Albero, Kryshko, and Santaeulàlia-Llopis, 2009) and the technology parameter, ϕ, is equal to 0.65 (Cooley and Prescott, 1995), then the 90% credible interval for the parameter λ includes mark-ups ranging from 5% to 23%.

Market clearing for the monetary market implies that the stock of money M_t is equal to nominal output. See equation (2.10). 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. 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.[9] In relative terms, these prior specifications are chosen so as to make the two signals equally informative about the corresponding exogenous state variables.[10]

Table 2.2 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 2.1.9, these parameter values crucially influence the persistence in the model. Priors for these parameters are very uninformative as I want to learn their values from the likelihood. I fix the discount factor β is fixed so as to match the long-run average real interest rate.

Table 2.2: Implied Prior Distributions (ICKM)

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

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

$$\Theta_C \equiv (\rho_m, A_0, M_0, \sigma_m, \sigma_a, \kappa_{pc}, \beta) \tag{2.26}$$

[9] 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).

[10] 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.

In Table 2.1 the priors for these parameters are reported. I use the same prior distributions for those parameters that are common to the ICKM. My priors reflect the beliefs that the slope of the Phillips curve, κ_{pc}, ranges from 0.00 to 0.22. This range includes values that are supported by several studies that estimate the slope of the Phillips curve, as surveyed by Schorfheide (2008).

2.2.4 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 2.3.

The coefficient $(1-\lambda)$ controls the degree of strategic complementarity in price-setting. As shown in section 2.1.9, 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 2-1 compares

Table 2.3: Posterior Distributions

	ICKM		Calvo Model	
Name	Median	90% Interval	Median	90% Interval
ρ_m	0.34	$[0.24, 0.45]$	0.24	$[0.15, 0.33]$
$100A_0$	0.45	$[0.36, 0.55]$	0.43	$[0.11, 0.74]$
$100M_0$	1.34	$[1.18, 1.49]$	1.34	$[1.20, 1.48]$
λ	0.32	$[0.13, 0.50]$	1.00	$[0.00, 0.00]$
$100\sigma_m$	0.88	$[0.81, 0.95]$	0.89	$[0.82, 0.97]$
$100\sigma_a$	0.88	$[0.70, 1.04]$	2.66	$[2.04, 3.36]$
$100\tilde{\sigma}_m$	9.04	$[4.97, 12.77]$	–	–
$100\tilde{\sigma}_a$	1.36	$[0.69, 2.02]$	–	–
κ_{pc}	–	–	0.01	$[0.01, 0.02]$
β	–	–	0.99	$[0.99, 0.99]$
$1-\lambda$	0.69	$[0.50, 0.87]$	–	–
$\sigma_m/\tilde{\sigma}_m$	0.10	$[0.06, 0.14]$	–	–
$\sigma_a/\tilde{\sigma}_a$	0.66	$[0.44, 0.94]$	–	–

the prior and the posterior distributions[11] 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 0.32. This estimate is plausible. This number is consistent with a Frisch labor-supply elasticity, η, of 0.5 (Ríos-Rull et al., 2009), a technology parameter, ϕ, of 0.65 (Cooley and Prescott, 1995), and a mark-up of about 9.5% (Woodford, 2003 and Rotemberg and Woodford, 1997).

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

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

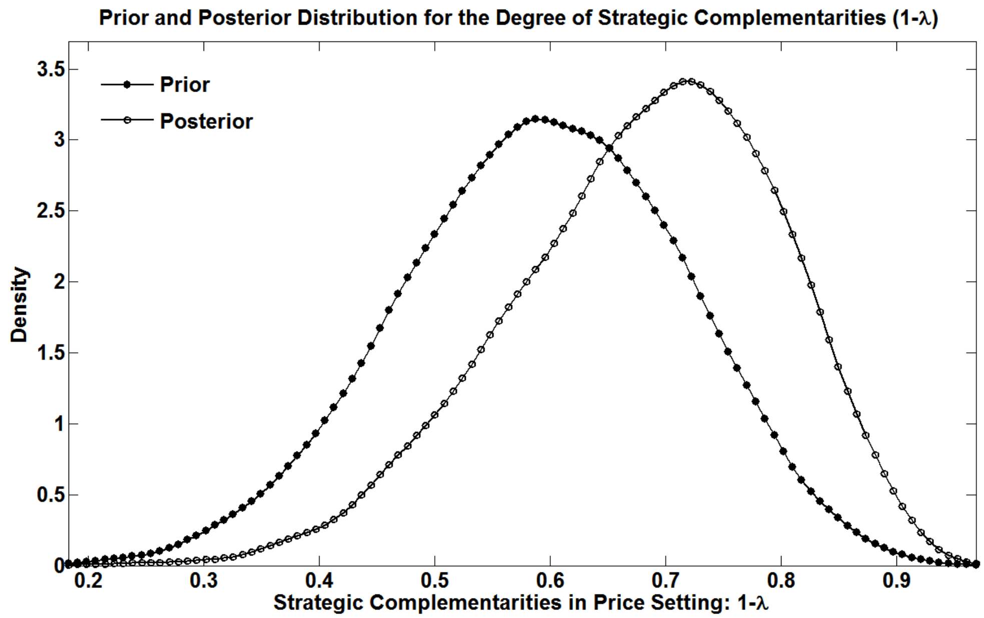

Figure 2-1: Prior and Posterior Distribution for the Degree of Strategic Complementarities

monetary policy is smaller by a factor of six.

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

2.2.5 MDD-Based Comparisons

The essay 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 2.2.3. I use Geweke's harmonic mean estimator (Geweke, 1999) to approximate the MDDs of the ICKM.

I consider a VAR(4):

$$\tilde{\mathbf{Y}}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1\tilde{\mathbf{Y}}_{t-1} + \mathbf{\Phi}_2\tilde{\mathbf{Y}}_{t-2} + \mathbf{\Phi}_3\tilde{\mathbf{Y}}_{t-3} + \mathbf{\Phi}_4\tilde{\mathbf{Y}}_{t-4} + \boldsymbol{\epsilon}_t \tag{2.27}$$

where $\tilde{\mathbf{Y}}_t = [\ln Y_t, \ln P_t]'$ and $\mathbf{\Sigma}_\epsilon \equiv \mathbb{E}\left(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'\right)$. I fit this VAR(4) to the data set presented in section 2.2.1. 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 2.4. 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.

Table 2.4: Logarithms of Marginal Datat Densities (MDDs)

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

Moreover, I also compute the MDD of the Calvo model and report the result in Table 2.4. 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.

2.2.6 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 2-2 and 2-3, 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

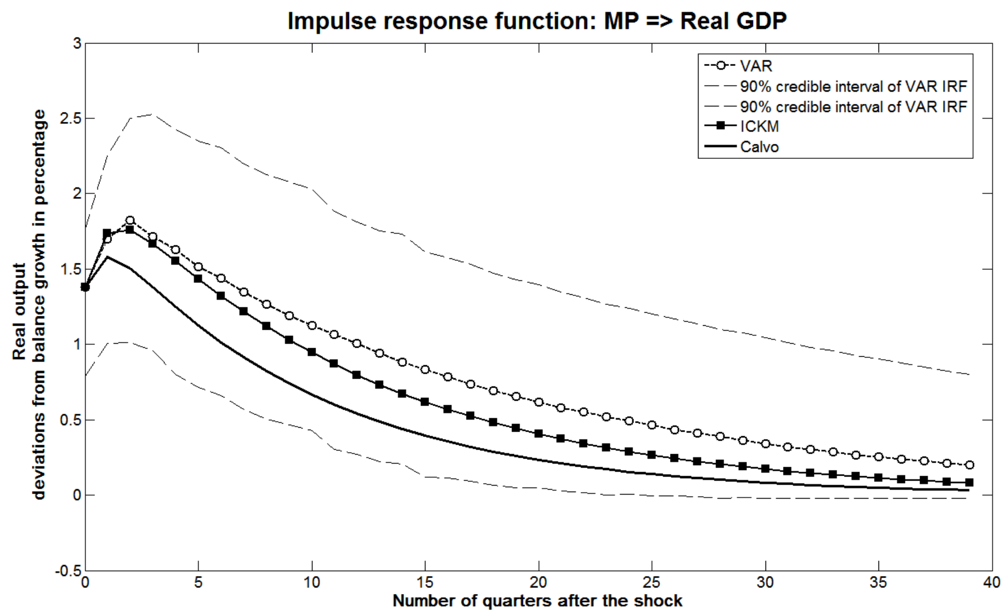

Figure 2-2: Impulse Response Function MP => Real GDP

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.

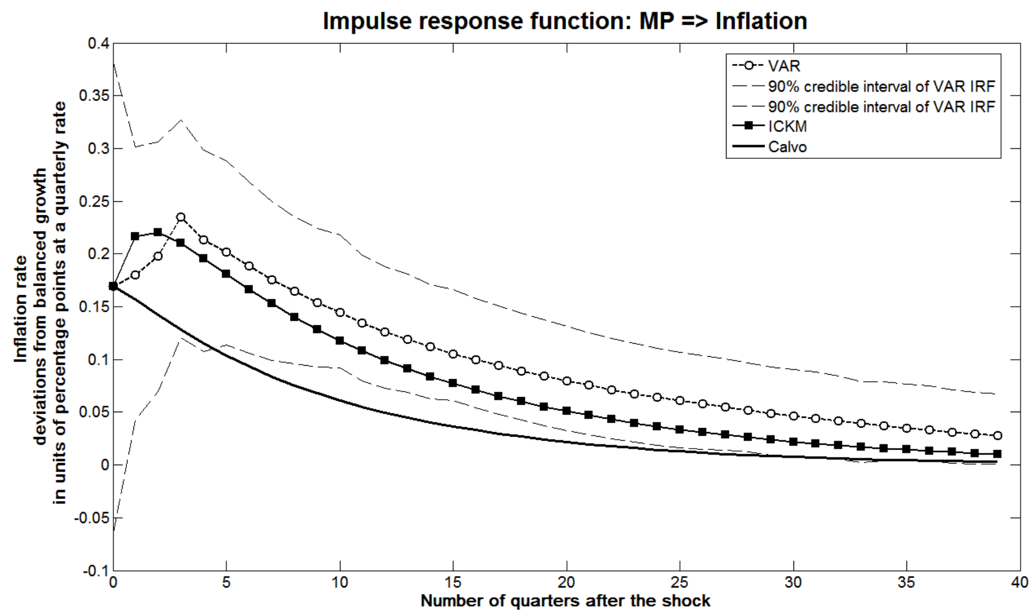

Figure 2-3: Impulse Response Function MP => Inflation

2.3 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., Christiano *et al.*, 1999). Nonetheless, the estimated signal-to-noise ratio concerning monetary policy is smaller than that of technology by a factor of six. This implies that firms have far less information about the state of monetary policy. This begs the question of whether such a finding is plausible or not. In the next chapter, I will shed some light on this issue.

Chapter 3

Optimal Allocation of Attention

In the first chapter, I show that a DSGE model with information frictions á la Woodford (2002) can account for the highly persistent effects of monetary disturbances that are observed in the data. Nonetheless, the estimated signal-to-noise ratio about monetary policy turns out to be smaller than that about technology by a factor of six. The signal-to-noise ratio determines the accuracy of a signal and, hence, how much information the signal conveys to those who receive it. Thus, this finding raises a question: is it plausible that firms acquire so little information about monetary policy? The answer to this question is no. I reach this conclusion by augmenting the model so as to allow firms to optimally choose the signal-to-noise ratios, subject to a constraint that sets an upper-bound to the overall precision of the signals. This constraint is widely used in the literature of rational inattention and is termed information-processing constraint (Sims, 2003 and Maćkowiak and Wiederholt, 2009). It specifies the rate at which firms can substitute pieces of information between the two shocks of the model (i.e., monetary policy shock and technology shock). Hence, the information-processing constraint defines a schedule along which firms can choose the optimal signal to noise ratios. I will call this augmented ICKM rational inattention model.

I find that the firms' marginal value of the information about monetary shocks is much higher than that about technology shocks in the model I estimated in the previous chapter. 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 admit two alternative interpretations. First, the signal-to-noise ratio relative to monetary policy seems to be implausibly small in the estimated ICKM. Second, the likelihood suggests that the rate at which firms are allowed to substitute information between the two shocks in rational inattention models may be wrong.

In the first section, I discuss how to construct the signal-to-noise schedule that is consistent with the information-processing constraint used in the literature of rational inattention. In the second section, firms' problem of allocating their attention is presented. In the third section, I compare the marginal rates of profit in the estimated ICKM. and in the rational inattention model. Finally, I show the results of some robustness check.

3.1 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,

$$\kappa \equiv H\left(m_t, a_t | z_{m,i}^{t-1}, z_{a,i}^{t-1}\right) - H\left(m_t, a_t | z_{m,i}^{t}, z_{a,i}^{t}\right) \tag{3.1}$$

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 (3.1) can be re-written as

$$\kappa = \kappa_m + \kappa_a \tag{3.2}$$

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

$$\kappa_m \equiv H\left(m_t | z_{m,i}^{t-1}\right) - H\left(m_t | z_{m,i}^{t}\right)$$

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

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:

$$\kappa_m = g_m\left(\sigma_m, \tilde{\sigma}_m, \Upsilon\right), \quad \kappa_a = g_a\left(\sigma_a, \tilde{\sigma}_a\right) \tag{3.3}$$

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 (3.2) and (3.3). 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.

3.2 The Optimal Allocation of Attention

In period zero,[1] firms allocate their available attention[2] by solving:

$$\max_{\kappa_m,\kappa_a} \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^t \hat{\pi}_t\left(\hat{p}^*_{i,t}, \hat{p}_t, \hat{y}_t, \hat{q}_t\right)|\mathbf{z}_i^0\right], \tag{3.4}$$

st

$$\ln P^*_{i,t} \equiv \mathbb{E}\left[(1-\lambda)\ln P_t + \lambda m_t - \lambda a_t|\mathbf{z}_i^t\right] \tag{3.5}$$

$$\mathbf{z}_{i,t} = \begin{bmatrix} m_t \\ a_t \end{bmatrix} + \begin{bmatrix} \tilde{\sigma}_m & 0 \\ 0 & \tilde{\sigma}_a \end{bmatrix} \mathbf{e}_{i,t} \tag{3.6}$$

$$\tilde{\sigma}_m = g_m^{-1}\left(\kappa_m, \sigma_m, \rho_m\right), \quad \tilde{\sigma}_a = g_a^{-1}\left(\kappa_a, \sigma_a\right) \tag{3.7}$$

$$\kappa_m + \kappa_a = \kappa, \text{ any t} \tag{3.8}$$

where $\hat{\pi}_t(\cdot)$ is the log-quadratic approximation of $Q_t\pi_t$, where π_t is the period profit function (2.11), $\hat{p}^*_{i,t} = \ln\left(P^*_{i,t}/P_t\right)$, $\hat{q}_t$ is the log deviations of $q_t = M_tQ_t$ from its value at the deterministic steady state, and $\mathbf{e}_{i,t} \overset{iid}{\backsim} \mathcal{N}\left(\mathbf{0}, \mathbb{I}_2\right)$. The

[1]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 (2009).

[2]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 (2009). 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 (3.8).

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 (3.7) are the inverse of the functions (3.3). The constraint (3.8) 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 quantified by the parameter κ, between observing monetary policy and technology. Solving the allocation-of-attention problem (3.4)-(3.8) 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 (3.5) in any subsequent periods.

3.3 Marginal Rate of Profit

The marginal rate of profit is defined as:

$$\textsc{mrp} \equiv \frac{\partial \Pi / \partial \kappa_m}{\partial \Pi / \partial \kappa_a}$$

where Π is the sum of discounted profits:

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

It is very simple to see that the MRP at the optimal allocation of attention (κ_m^*, κ_a^*) is equal to unity. Interestingly, enough, this result does not require to set the parameter values in the rational inattention model. 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 47.19. This number is hugely bigger than unity. In the estimated ICKM, firms are willing to trade more than 47 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.

It is worthwhile noticing that the finding above is affected by the value of the rate of substitution embedded into the information-processing constraint (3.8). The signal-to-noise ratio picked by the likelihood would be consistent with an information-processing constraint of the following type:

$$\kappa_m + \frac{1}{47.19}\kappa_a = \kappa \tag{3.9}$$

Therefore, one could also interpret the finding above as evidence that the rate of substitution implied by the information-processing constraint (3.8) is inconsistent with the data.

3.4 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 (3.4)-(3.8).

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 (3.3) 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 3.1 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 0.41 bits, respectively. The posterior median of the overall amount of information processed by firms per quarter, κ, is 0.51 bits.[3] Figure 3-1 compares the prior and the posterior distributions[4] of the fraction of the overall firms' attention paid to the technology shocks, that is, $\kappa_a / (\kappa_m + \kappa_a)$.

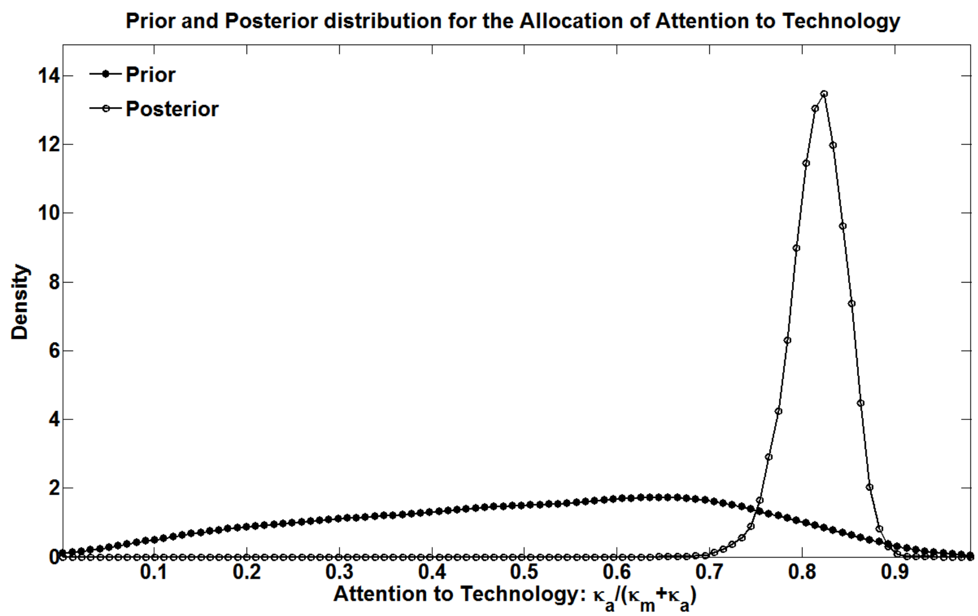

Figure 3-1: Prior and Posterior Distribution for the Allocation of Attention to Technology

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 80%). Hence, according to the data, the adjustment of output and inflation to monetary shocks is rather slow, as confirmed by the IRFs in Figures 2-2 and 2-3.

[3] This is obtained by using the prior and posterior draws for κ_m and κ_a as long as equation (3.2).

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

Furthermore, in Figure 3-1 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: $\kappa_a / (\kappa_m + \kappa_a)$.

Table 3.1: Implied Prior and Posterior Distributions

	Prior		
Variables	Descriptions	Median	90% Interval
κ_m	information flow MP	0.54	$[0.08, 1.87]$
κ_a	information flow tech.	0.66	$[0.12, 1.21]$
$\kappa = \kappa_m + \kappa_a$	overall level of attention	1.31	$[0.40, 2.72]$
$\frac{\kappa_a}{\kappa_m + \kappa_a}$	allocation of attention to tech.	0.53	$[0.15, 0.83]$

	Posterior		
Variables	Descriptions	Median	90% Interval
κ_m	information flow MP	0.10	$[0.06, 0.15]$
κ_a	information flow tech.	0.41	$[0.26, 0.60]$
$\kappa = \kappa_m + \kappa_a$	overall level of attention	0.51	$[0.34, 0.75]$
$\frac{\kappa_a}{\kappa_m + \kappa_a}$	allocation of attention to tech.	0.80	$[0.74, 0.86]$

Now I have to solve the problem (3.4)-(3.8) for the optimal allocation of attention (κ_m^*, κ_a^*). Yet, I need first to pin down the information-processing constraint (3.8). To do that, I need to fix one degree of freedom: the size of the parameter κ. I calibrate the value of this parameter by using its estimated value in Table 3.1, that is $\kappa = 0.51$ bits. I then solve[5] the problem (3.4)-(3.8) for the optimal allocation of attention and obtain that κ_m^* is equal to 0.33

[5] 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 (3.7) 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 (3.4)-(3.8) 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.

and κ_a^* is equal to 0.22. These findings show that the estimated allocation of attention (κ_m, κ_a) (see Table 3.1) 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 3-2 and 3-3 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 2.2.6. 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.

3.5 Conclusion

In this chapter, 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 ratio-

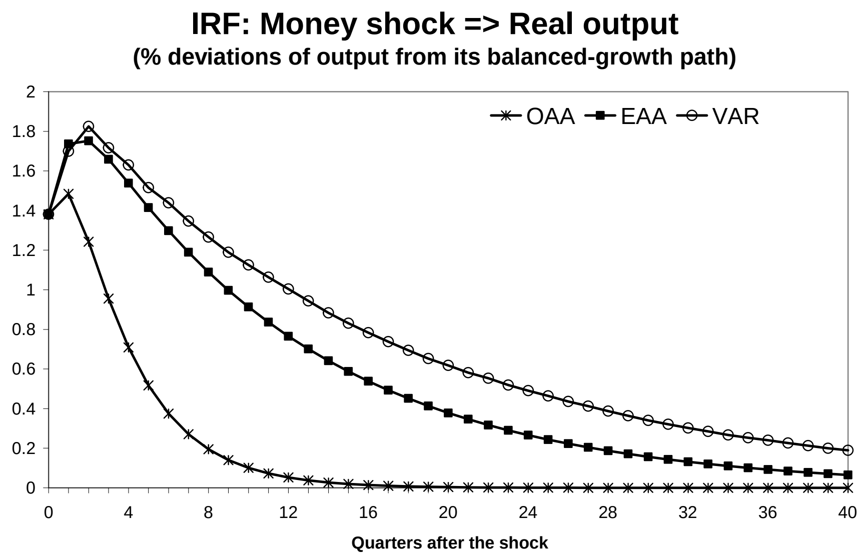

Figure 3-2: IRF: Money shock => Real Output

nal 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. For instance, finding evidence that can help quantifying the relative difficulty of learning about different shocks for firms would be very useful.

Other leading studies based upon rational inattention models (e.g., Maćkowiak and Wiederholt 2009) put restrictions upon the choice set of signal processes so

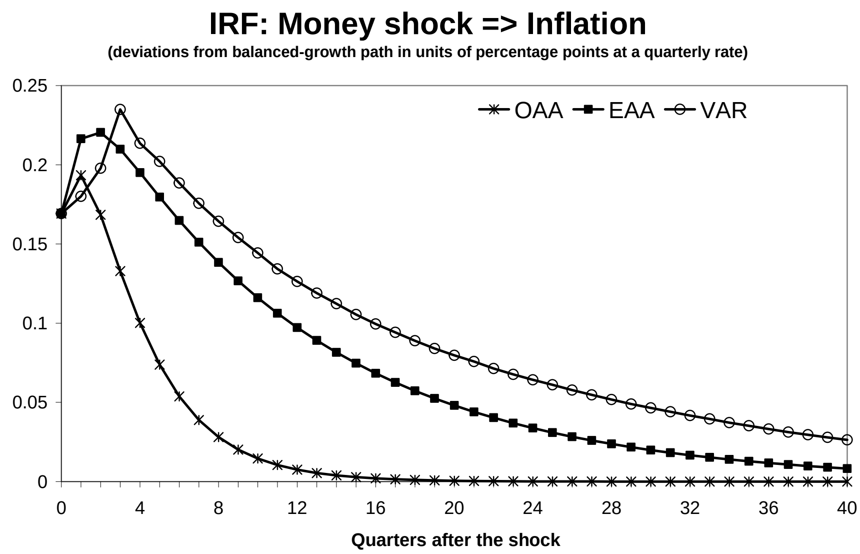

Figure 3-3: IRF: Money shock => Inflation

that the optimal signals are Gaussian and orthogonal to each others. This restriction leads to a rate of substitution that is equal to one. This study suggests that this rate may not be sensible from an empirical perspective.

Finally, it is worthy noticing that in full-fledge rational inattention models (e.g., Sims, 2003, 2006) agents choose the stochastic process of the signal under no parametric restrictions. In such models, optimal signals might not be orthogonal and Gaussian. Hence, the rate of substitution is not said to be equal to one and depends on the nature of the optimal signal. While this may be seen as a possible resolution of the discrepancy between the likelihood estimates and model predictions, this approach has the shortcoming of complicating the economic interpretation of the implied rate of substitution.

Chapter 4

Monetary Policy and Beliefs

That monetary policy influences output and inflation by affecting agents' expectations has come to a growing consensus among scholars and policy makers (Woodford, 2005, Morris and Shin, 2007). Michael Woodford writes:

> *Central banking is not like steering an oil tanker, or even guiding a spacecraft, which follows a trajectory that depends on constantly changing factors, but that does not depend on the vehicle's own expectations about where it is heading. Because the key decisionmakers in an economy are forward-looking, central banks affect the economy as much through their influence on expectations as through any direct, mechanical effects of central bank trading in the market for overnight cash.*

The aim of this chapter is to quantitatively assess the relevance of this new transmission channel of monetary policy that relies on affecting agents' beliefs. I develop a DSGE model where agents take decisions under incomplete and disperse information about aggregate state variables or fundamentals (e.g., inflation and output). The model economy is populated by six classes of agents: households, intermediate goods firms, final goods producers, a monetary authority (or central bank) and a government. There are three aggregate shocks: a technology shock, a monetary-policy shock, and a government spending-shocks. The monetary authority perfectly observes the history of aggregate shocks and sets its monetary policy instrument (i.e., interest rate) by following a Taylor-type reaction function. For tractability, I shall assume that households, final goods producers, and government have perfect information. Each intermediate goods firm lives on an island. No information can be traded among islands. Firms face nominal rigidities á la Calvo: there exists a lottery that establishes which firms are allowed to re-optimize their prices. Those firms that are allowed to re-optimize their prices have to forecast the dynamics of future marginal costs that depend on output gap and inflation. They perform this forecast by observing last period's output and inflation, the current island-specific technology shock, and the current interest rate set by the central bank.

Firms observe the history of island-specific technology shocks but they do not observe any of the aggregate shocks. The island-specific technology shock is correlated with the aggregate technology shock. Hence, the island-specific technology shocks are *private signals* that convey information about the current aggregate technology. As a result of the idiosyncratic nature of these signals, firms' expectations about inflation and output differ across islands.

Firms also know the monetary policy reaction function and perfectly observe the interest rate set by the monetary authority in every period. The salient feature of this model is that the monetary authority holds superior information about the aggregate shocks than firms. In such a setup, the monetary policy instrument works as a *public signal* that conveys novel information to firms about the aggregate shocks and can influence output and inflation by affecting firms' expectations.

An appealing feature of the model is that strategic complementarities in price settings and disperse information lead to the inertial behavior of inflation. In the model, firms face strategic complementarities in price-setting: they find it optimal to raise (cut) their prices when the average price increases (decreases). Private information is introduced into the price setting problem of the firm through the island-specific technology shocks that work as private signals. The optimal price of an individual good depends positively on a firm's own marginal cost and the price chosen by other firms, but individual firms cannot observe the marginal cost of other firms and therefore do not know the current price chosen by other firms with certainty. This set up may be referred to as firms having imperfect common knowledge. As shown in chapter 1, in such an environment, a forecasting-the-forecasts-of-others type of problem arises and generates sluggish responses of output and inflation to nominal disturbances.

Furthermore, the model can be shown to be nested within a standard New

Keynesian DSGE model where firms are perfectly informed. If one assumes that firms' information sets are complete (i.e., firms also observe current output and inflation), the dispersed information model boils down to a prototypical three-equation New Keynesian DSGE model, where monetary disturbances affect output by influencing the intertemporal allocation of consumption. This traditional transmission channel of monetary policy still exists in the dispersed information model. Yet, in the dispersed information model, there is another transmission channel through which monetary impulses affect model variables. This channel relies on the role of interest rate as a public signal. More precisely, changes in the interest rate are interpreted by firms as realizations of a public signal that provides unanticipated information on the dynamic of marginal costs. Quite importantly, the nestedness of the dispersed information model within a standard New Keynesian DSGE model allows me to assess the significance of this new transmission channel by running simple counterfactuals.

4.1 A Brief Overview of the Literature

From a theoretical perspective, the idea that publicly observed policy can coordinate agents' expectations has been recently explored by the literature of global games (Morris and Shin, 2003a). Morris and Shin (2003b) and Amato and Shin (2003, 2006) derive normative implications for incomplete-information settings and focus on the welfare effects of disclosing public information. Hellwig (2002) derives impulse responses to a large range of shocks for an economy with monopolistic competition and incomplete information. These partial equilibrium models, however, are too stylized to be used for empirically assessing central banks' role for coordinating expectations.

My model is built on Nimark (2008) who introduces a model where firms hold private information about the dynamics of their future marginal costs, and face both strategic complementarities in price setting and nominal rigidities. The nice feature of this model is that the supply side of this economy can be analytically worked out and turns out to be characterized by an equation that resembles the standard New-Keynesian Phillips curve. One shortcoming of this model is that the monetary policy framework is too stylized to have a chance to capture the complexity of modern monetary policy practices. Furthermore, the role of monetary policy in coordinating agents' expectations is completely absent in that central bank's actions only convey redundant information to agents.

The model that is presented in this chapter is also related to Lorenzoni (forthcoming), who studies optimal monetary policy in a model where aggregate fluctuations are driven by the private sector's uncertainty about the economy's fundamentals. Information on aggregate productivity is dispersed across agents and there are two aggregate shocks: a standard productivity shock and a noise shock affecting public beliefs about aggregate productivity. The nature of the latter shock is related to public news about technological advances, aggregate statistics, and information reflected in stock market prices and other financial variables. The central bank does not perfectly observe the noise shocks. The central question addressed by Lorenzoni (forthcoming) is whether the monetary policy should accommodate those noise shocks that work as a coordination device of agents' expectations.

4.2 The Model

There is a continuum of islands, indexed by $j \in (0, 1)$. In every island a continuum of households, indexed by $i \in (0, 1)$, lives along with one intermediate goods firm. There is a centralized entity that conducts monetary policy: the central bank. Another centralized entity, the government, carries out the fiscal policy. Perfectly competitive final goods producers also populate the economy. Households consume the final good and supply labor to the firm located in the same island. It is assumed that households cannot ship their labor services to firms that are located in other islands. A Calvo lottery establishes which islands are allowed to re-optimize their prices. The outcome of the Calvo lotteries is common knowledge among agents.

There are aggregate shocks and island specific shocks that hit the model economy. The aggregate shocks are: a technology shock, a monetary-policy shock, and a government-spending shock. The aggregate shocks can be decomposed into a persistent and white noise component[1]. Every of these shocks is orthogonal to others at all leads and lags. Island-specific shocks include the island-specific technology, $A_{j,t}$, that determines the level of technology in the island j at time t, and the outcome of the Calvo lottery for price-optimization. The former shock is correlated with the aggregate technology shock. Both island-specific shocks are orthogonal to each other at all leads and lags.

Any period t is divided into three stages. At stage 0 $(t, 0)$, the market for state-contingent claims opens and closes. At stage 1 $(t, 1)$, island-specific shocks realize, households and firms observe island-specific shocks, and firms set their prices. At stage 2 $(t, 2)$, state-contingent claims are settled and households learn the state of the economy. At this stage, households choose consumption, money holdings, bonds and their labor supplies. The fiscal authority decides how much to consume. Final goods producers demand intermediate goods across islands and use them as inputs to produce the final good to be sold to the households and to the government. Intermediate goods firms hire labor and produce so as to deliver the demanded quantity of their good at the price they set.

The market for state-contingent claims has three important features: first, households can trade in this market only at the stage 0. Second, the claims traded in the market pay out the numeraire at the stage 2. Hence, these claims can be made contingent to the *observed* states at the stage 2. Third, firms are

[1] Monetary shocks are decomposed into an inflation-targeting shock and a white noise shock to the Taylor rule.

not allowed to trade these claims.

The market for state-contingent claims ensures that the wealth distribution across islands is degenerate. Furthermore, this market also plays an informative role in the model, since households learn the state of the economy when they get the proceeds from their investments in this market at the stage 2. Since households are not heterogenous in their information sets at the stage 2 and wealth is the same across islands, then one can use the representative household to solve the problem of households.

4.2.1 Technology and Intermediate Goods Firms

Consider an arbitrary island j. In this island, the firm j produces according to a linear technology:

$$Y_{j,t} = A_{j,t} N_{j,t} \tag{4.1}$$

We assume that labor markets are segmented across islands. The real marginal costs for firm j are given by:

$$mc_{j,t} = \frac{W_{j,t}}{A_{j,t} P_t}$$

where $W_{j,t}$ is the competitive nominal wage in the island j and P_t is the price of the final good. $A_{j,t}$ is the island-specific technology shock that can be decomposed into (1) a trend component A_0, (2) a persistent aggregate component, z_t,

(3) a white-noise aggregate component, $\eta_{a,t}$, and (4) a white-noise idiosyncratic component, $\eta^a_{j,t}$. More specifically, we have:

$$A_{j,t} = A_0^t A_t e^{\eta^a_{j,t}} \tag{4.2}$$

with $A_0 > 1$, and $\eta^a_{j,t} \overset{iid}{\backsim} \mathcal{N}\left(0, \sigma^j_a\right)$, and

$$A_t = e^{z_t + \sigma_a \eta_{a,t}} \tag{4.3}$$

where $\eta_{a,t} \overset{iid}{\backsim} \mathcal{N}\left(0, 1\right)$ and

$$z_t = \rho_z z_{t-1} + \sigma_z \varepsilon_{z,t}$$

with $\varepsilon_{z,t} \overset{iid}{\backsim} \mathcal{N}\left(0, 1\right)$.

Firms face a Calvo lottery with probability θ of not adjusting their prices. After having observed the outcome of the Calvo lottery, intermediate goods firms set the prices in their islands and they commit to satisfy any demanded quantity at that price. Those firms that are allowed to re-optimize their prices solve:

$$\max_{P^*_{j,t}} \mathbb{E}\left[\sum_{s=0}^{\infty} \left(\beta\theta\right)^s \Xi_{j,t+s} \left(\pi^s_* P^*_{j,t} - MC_{j,t+s}\right) Y_{j,t+s} | \mathcal{I}_{j,t}\right]$$

such that

$$Y_{j,t+s} = \left(\frac{\pi_*^s P_{j,t}^*}{P_{t+s}}\right)^{-\nu} Y_{t+s}, \quad Y_{j,t} = A_{j,t} N_{j,t}$$

$$\mathcal{I}_{j,t} = \{R_\tau, P_{\tau-1}, Y_{\tau-1}, \boldsymbol{\theta}_{j,\tau}, \Theta : \tau \leq t\}$$

$$\boldsymbol{\theta}_{j,t} = \left\{A_{j,t}, P_{j,t}^*, Y_{j,t-1}, MC_{j,t-1}, N_{j,t-1}, W_{j,t-1}, \Pi_{j,t-1}\right\}$$

where $MC_{j,t}$ is the nominal marginal cost. $\Xi_{j,t}$ stands for the stochastic discount factor. The firm's specific demand $Y_{j,t}$ is standard and can be easily derived from final goods producers' problem. $\mathcal{I}_{j,t}$ is the information set of firm j at time t.

If one solves the problem of those firms that are allowed to re-optimize its price, one obtains:

$$\mathbb{E}\left[\sum_{s=0}^{\infty} (\beta\theta)^s \, \Xi_{j,t+s} \left[(1-\nu)\,\pi_*^s + \nu \frac{MC_{j,t+s}}{P_{j,t}^*}\right] Y_{j,t+s} | \mathcal{I}_{j,t}\right] = 0$$

The price index is given by

$$P_t^{1-\nu} = \int P_{j,t}^{1-\nu} dj$$

and hence,

$$P_t^{1-\nu} = \theta \left(\pi_* P_{t-1}\right)^{1-\nu} + (1-\theta) \int \left(P_{j,t}^*\right)^{1-\nu} dj$$

4.2.2 Preferences

Consumption $C_{j,t}$, labor $N_{j,t}$, and financial decisions (money $M_{j,t}$, and bonds $B_{j,t}$) are taken at the second stage. We assume that the habit stock is given by the level of technology A_t. This assumption ensures that the economy evolves along a balanced growth path even if the utility function is additively separable in consumption, real money balances, and leisure. Households solve[2]:

$$\max_{C_{j,t},B_{j,t},M_{j,t},N_{j,t}} \frac{(C_{j,t}/A_0^t)^{1-\gamma}-1}{1-\gamma}+\frac{\chi_m}{1-\gamma_m}\left(\frac{M_{j,t}}{P_t}\right)^{1-\gamma_m}-\eta N_{j,t}$$

such that

$$P_tC_{j,t}+B_{j,t}+M_{j,t}=W_{j,t}N_{j,t}+R_{t-1}B_{j,t-1}+M_{j,t-1}+\Pi_{j,t}-T_t+q_t\left(\omega_{j,t}\right)d_{j,t}^*\left(\omega_{j,t}\right)$$

$$\omega_{j,t}=\left\{\varphi_t,A_{j,t},\mathbb{I}_{j,t}^{Calvo}\right\}$$

where R_t is the interest rate paid out by the bond, $\Pi_{j,t}$ is the dividend paid by the firm located in the island j, T_t is a lump-sum tax/transfer, and $d_{j,t}^*\left(\omega_{j,t}\right)$ is the amount of state-contingent claims that pay one unit of the numeraire at stage (t,2) if the state $\omega_{j,t}$ realizes. Note that by observing the price of these claims, q_t, households learn the state variables $\omega_{j,t}$. φ_t is a vector that

[2]Note that the utility function is linear in labor. Relaxing this restriction would make the marginal costs in the island j to be different from the structure: $MC_{j,t}=MC_t+\varepsilon_{j,t}$, where $\varepsilon_{j,t}$ is a linear functions of island specific shocks. This property of the island-specific marginal costs is necessary to derive the imperfect-information Phillips curve as in Nimark (2008).

includes the decomposed fundamentals, that is the technology shock $(z_t, \eta_{a,t})$, monetary-policy shock $(\hat{\pi}^*_{t|t}, \eta_{r,t})$, government spending shocks $(g_t, \eta_{g,t})$.

The consumption Euler equation can be written as:

$$\frac{(C_t/A_0^t)^{-\gamma}}{P_t} = \frac{\beta}{A_0}\mathbb{E}_t\left(\frac{\left(C_{t+1}/A_0^{t+1}\right)^{-\gamma}}{P_{t+1}}\right)R_t \tag{4.4}$$

The demand for money is:

$$\chi_m\left(\frac{M_t}{A_0^{t/\gamma_m}P_t}\right)^{-\gamma_m}\left(\frac{C_t}{A_0^t}\right)^{\gamma} = \left(\frac{R_t-1}{R_t}\right) \tag{4.5}$$

The supply for labor becomes:

$$\eta\left(\frac{C_t}{A_0^t}\right)^{\gamma} = \frac{W_t}{A_0^t P_t} \tag{4.6}$$

4.2.3 Monetary and Fiscal Policy

The monetary policy is modeled by specifying a Taylor-rule type reaction function:

$$\frac{R_t}{R_*} = \left(\frac{R_{t-1}}{R_*}\right)^{\rho_r}\left[\left(\frac{\pi_t}{\pi_t^*}\right)^{\phi_\pi}\left(\frac{Y_t}{Y_t^*}\right)^{\phi_y}\right]^{(1-\rho_r)} e^{\sigma_r \eta_{r,t}} \tag{4.7}$$

where $\eta_{r,t} \overset{iid}{\backsim} \mathcal{N}(0,1)$. The monetary authority sets the interest rate R_t and is perfectly informed (i.e., it observes the contemporaneous realizations of all aggregate variables). π_t^* stands for the inflation target that follows a stationary

AR process:

$$\ln \pi_t^* = (1 - \rho_\pi)\, \pi_* + \rho_\pi \ln \pi_{t-1}^* + \sigma_\pi \varepsilon_{\pi,t}$$

where $\varepsilon_{\pi,t} \overset{iid}{\backsim} \mathcal{N}(0,1)$. We finally define output gap $Y_t^* \equiv A_0^t A_t y_*$ where y_* is the steady-state value of detrended output.

The government transfers resources to/from and issue bonds to households at the stage 2. Furthermore, they decide their consumption of final goods. The government budget constraint is:

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

where government spending G_t is

$$G_t = (1 - 1/\zeta_t)\, Y_t$$

We can decompose the government spending shock ζ_t into:

$$\zeta_t = g_t e^{\sigma_g \eta_{g,t}}$$

where $\varepsilon_{g,t} \backsim \mathcal{N}(0,1)$ and

$$\ln g_t = \left(1 - \rho_g\right) \ln g_0 + \rho_g \ln g_{t-1} + \tilde{\sigma}_g \varepsilon_{g,t}$$

where $\eta_{g,t} \backsim \mathcal{N}(0,1)$. We will denote $\hat{\zeta}_t = \ln(\zeta_t/g_0)$ and $\hat{g}_t = \ln g_t - \ln g_0$.

The resource constraint can be shown to be:

$$Y_t = \zeta_t C_t$$

4.2.4 Detrending

Let us denote the inflation rate as $\pi_t \equiv P_t/P_{t-1}$. Moreover, define:

$$y_t \equiv \frac{Y_t}{A_0^t},\ c_t \equiv \frac{C_t}{A_0^t};\ p_{j,t}^* \equiv \frac{P_{j,t}^*}{P_t},\ y_{j,t} \equiv \frac{Y_{j,t}}{A_0^t}$$

$$w_{j,t} \equiv \frac{W_{j,t}}{A_0^t P_t},\ a_t \equiv \frac{A_t}{A_0^t}, R_t \equiv \frac{R_t}{R_*},\ mc_{j,t} \equiv \frac{MC_{j,t}}{P_t}$$

$$\xi_{j,t} \equiv A_0^t \Xi_{j,t}$$

The Euler equation (4.4) becomes:

$$c_t^{-\gamma} = \frac{\beta}{A_0} \mathbb{E}_t \left(\frac{c_{t+1}^{-\gamma}}{\pi_{t+1}} \right) R_t \tag{4.8}$$

The price setting equation becomes:

$$\begin{aligned} &\mathbb{E}\left[\xi_t \left[(1-\nu) + \nu \frac{mc_{j,t}}{p_{j,t}^*}\right] y_{j,t} | \mathcal{I}_{j,t}\right] + \\ &\mathbb{E}\left[\sum_{s=1}^{\infty} (\beta\theta)^s \xi_{t+s} \left[(1-\nu)\pi_*^s + \nu \frac{mc_{j,t+s}}{p_{j,t}^*} \left(\Pi_{\tau=1}^s \pi_{t+\tau}\right)\right] y_{j,t+s} | \mathcal{I}_{j,t}\right] = 0 \end{aligned} \tag{4.9}$$

The resource constraint becomes:

$$y_t = \zeta_t c_t \tag{4.10}$$

4.2.5 Log-linearization

First, it is easy to show that the resource constraint (4.10) becomes:

$$\hat{y}_t = \hat{\zeta}_t + \hat{c}_t \tag{4.11}$$

Note that the consumption Euler equation becomes:

$$-\gamma \hat{c}_t = -\gamma \mathbb{E}_t \hat{c}_{t+1} - \mathbb{E}_t \hat{\pi}_t + \hat{R}_t$$

By substituting the resource constraint (4.11) we obtain the standard IS equation:

$$\hat{y}_t - \hat{\zeta}_t = \mathbb{E}_t \left(\hat{y}_{t+1} - \hat{\zeta}_{t+1} \right) + \frac{1}{\gamma} \mathbb{E}_t \hat{\pi}_t - \frac{1}{\gamma} \hat{R}_t \tag{4.12}$$

The Phillips curve is worked out in the appendix D:

$$\hat{\pi}_t = (1-\theta)(1-\beta\theta) \sum_{k=0}^{\infty} (1-\theta)^k \widehat{mc}_{t|t}^{(k)} + \beta\theta \sum_{k=0}^{\infty} (1-\theta)^k \hat{\pi}_{t+1|t}^{(k+1)}$$

where $\widehat{mc}_{t|t}^{(k)}$ denotes the average k-th order expectations about the real aggregate marginal costs, $\widehat{mc}_t$. For any $k \in \{1, 2, \ldots\}$ one can show that:

$$\widehat{mc}_{t|t}^{(k)} = \gamma \left(\hat{y}_{t|t}^{(k+1)} - \hat{\zeta}_{t|t}^{(k+1)} \right) - z_{t|t}^{(k)} - \sigma_a \eta_{a,t|t}^{(k)} \tag{4.13}$$

By using this result, we write:

$$\begin{aligned}\hat{\pi}_t = {} & (1-\theta)(1-\beta\theta) \sum_{k=0}^{\infty} (1-\theta)^k \left[\gamma \left(\hat{y}_{t|t}^{(k+1)} - \hat{\zeta}_{t|t}^{(k+1)} \right) - z_{t|t}^{(k)} - \sigma_a \eta_{a,t|t}^{(k)} \right] \\ & + \beta\theta \sum_{k=0}^{\infty} (1-\theta)^k \, \widehat{\pi}_{t+1|t}^{(k+1)} \end{aligned} \tag{4.14}$$

The Taylor rule can be easily linearized:

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1-\rho_r) \left[\phi_\pi \left(\hat{\pi}_t - \hat{\pi}_t^* \right) + \phi_y \left(\hat{y}_t - z_t - \sigma_a \eta_{a,t} \right) \right] + \sigma_r \eta_{r,t} \tag{4.15}$$

where

$$\hat{\pi}_t^* = \rho_\pi \hat{\pi}_{t-1}^* + \sigma_\pi \varepsilon_t^\pi$$

4.2.6 Perfect Information Model (PIM)

A nice feature of the model is that it nests standard three-equation New Keynesian DSGE model. If one assumes that firms perfectly observe current output and inflation: $\mathcal{I}_{j,t}^p = \{R_\tau, P_\tau, Y_\tau, \boldsymbol{\theta}_{j,\tau}, \Theta : \tau \leq t\}$, then the model boils down to:

$$\hat{y}_t - \hat{\zeta}_t = \mathbb{E}_t \left(\hat{y}_{t+1} - \hat{\zeta}_{t+1} \right) + \frac{1}{\gamma} \mathbb{E}_t \hat{\pi}_t - \frac{1}{\gamma} \hat{R}_t \tag{4.16}$$

$$\hat{\pi}_t = \frac{(1-\theta)(1-\beta\theta)}{\theta}\left[\gamma\left(\hat{y}_t - \hat{\zeta}_t\right) - z_t - \sigma_a\eta_{a,t}\right] + \beta\mathbb{E}_t\widehat{\pi}_{t+1} \tag{4.17}$$

$$\hat{R}_t = \rho_r\hat{R}_{t-1} + (1-\rho_r)\left[\phi_\pi\left(\hat{\pi}_t - \hat{\pi}_t^*\right) + \phi_y\left(\hat{y}_t - z_t - \sigma_a\eta_{a,t}\right)\right] + \sigma_r\eta_{r,t} \tag{4.18}$$

4.3 Model Solution

The model is solved through a guess-and-verify strategy. Let's guess the law of motion of the endogenous variables:

$$\hat{\pi}_t = \mathbf{a}_0\varphi_{t|t}^{(0:k)} + a_1\hat{R}_{t-1}$$

$$\hat{y}_t = \mathbf{b}_0\varphi_{t|t}^{(0:k)} + b_1\hat{R}_{t-1}$$

$$\hat{R}_t = \mathbf{c}_0\varphi_{t|t}^{(0:k)} + c_1\hat{R}_{t-1}$$

where $\varphi_{t|t}^{(0:k)}$ is a column vector collecting the average expectations of the exogenous variables up to the k-th order. More specifically,

$$\varphi_{t|t}^{(0:k)} \equiv \left[z_t^{(s)}, \eta_{a,t}^{(s)}, \hat{\pi}_{t|t}^{*(s)}, \eta_{r,t}^{(s)}, g_t^{(s)}, \eta_{g,t}^{(s)} : s \in \{0,1,\ldots,k\}\right]'$$

The vector $\mathbf{s}_t$ includes the endogenous state variables of the model:

$$\mathbf{s}_t = \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t\right]'$$

I will also guess that the vector of higher-order beliefs, $\varphi_{t|t}^{(0:k)}$, and the state variables, $\mathbf{s}_t$, follows a VAR(1). Thus, we conjecture that firms' state-space model is the following[3]:

$$\mathbf{A}\underbrace{\begin{bmatrix} \varphi_{t|t}^{(0:k)} \\ \mathbf{s}_t \\ \varphi_{t-1|t-1}^{(0:k)} \\ \mathbf{s}_{t-1} \end{bmatrix}}_{\mathbf{X}_t} = \mathbf{B}\underbrace{\begin{bmatrix} \varphi_{t-1|t-1}^{(0:k)} \\ \mathbf{s}_{t-1} \\ \varphi_{t-2|t-2}^{(0:k)} \\ \mathbf{s}_{t-2} \end{bmatrix}}_{\mathbf{X}_{t-1}} + \mathbf{C}\cdot\underbrace{\begin{bmatrix} \varepsilon_{z,t} \\ \eta_{a,t} \\ \varepsilon_{\pi,t} \\ \eta_{r,t} \\ \varepsilon_{g,t} \\ \eta_{g,t} \end{bmatrix}}_{\boldsymbol{\varepsilon}_t} \tag{4.19}$$

$$\begin{bmatrix} \ln A_{j,t} - \ln A_0 t \\ \pi_{t-1} - \ln \pi_* \\ \ln Y_{t-1} - \ln y_* - \ln A_0 t \\ \hat{R}_t - \ln \hat{R}_* \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{L} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_1^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_2^T \\ \mathbf{0} & \mathbf{1}_3^T & \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\mathbf{D}} \mathbf{X}_t + \underbrace{\begin{bmatrix} \sigma_a^j \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{Q}} \underbrace{\eta_{j,t}^a}_{\mathbf{e}_{j,t}}.$$

[3]Note that we do not need to include the past island-specific nominal marginal costs into the set of observables as they are just a linear combination of known variables, that is the island-specific nominal wage and island-specific shock. Indeed,

$$\ln MC_{j,t} - \ln A_0 \cdot t = \ln W_{j,t} \underbrace{- z_t - \sigma_a \eta_{a,t} - \eta_{j,t}^a}_{\ln A_{j,t}}$$

The same is true for the output sold by firm j, which is a linear combination of aggregate output and price level.

where $\mathbf{1}_i^T$ is a row vector with the i-th element equal to one and zero elsewhere, and

$$\mathbf{A}=\begin{bmatrix} \mathbb{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{v}_0 & \mathbb{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbb{I} \end{bmatrix}, \ \mathbf{B}=\begin{bmatrix} \mathbf{M} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_2 & \mathbf{0} & \mathbf{0} \\ \mathbb{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{C}=\begin{bmatrix} \mathbf{N} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

where $\mathbb{I}$ is the identity matrix, $\mathbf{L}=\begin{bmatrix} 1, & 1, & \mathbf{0}_{1\times 6k+4} \end{bmatrix}$, $\mathbf{v}_0=[\mathbf{a}_0', \mathbf{b}_0', \mathbf{c}_0']'$, and

$$\mathbf{v}_2=[0_{3\times 2}, \mathbf{v}_1],$$

with $\mathbf{v}_1=[a_1, b_1, c_1]'$.

The full characterization of the matrices $\mathbf{M}$ and $\mathbf{N}$ is obtained through repeated applications of the Kalman filter and is detailed in appendix F. The first six rows of the matrices $\mathbf{M}$ and $\mathbf{N}$ are known:

$$\mathbf{M}_{(1:6,:)} = \begin{bmatrix} \rho_z & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & \rho_\pi & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & \rho_g & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix} ; \ \mathbf{N}_{(1:6,:)} = \begin{bmatrix} \sigma_z & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_a & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\pi & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_r & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{\sigma}_g & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_g \end{bmatrix}$$

Solving the model, hence, requires to computationally find a fixed point over the space of vectors $[\mathbf{v}_0, \mathbf{v}_1]$ and matrices $[\mathbf{M}, \mathbf{N}]$. We verify the guess by using the the three structural equations of the model: the IS equation (4.12), the Phillips curve (4.14) and the Taylor rule (4.15).

In appendix E, we show that the Euler equation (4.12) implies that:

$$\mathbf{b}_0 = \mathbf{1}_5^T + \mathbf{1}_6^T + \mathbf{b}_0\mathbf{M} + b_1\mathbf{c}_0 - \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right)\mathbf{M} + \frac{1}{\gamma}\left(\mathbf{a}_0\mathbf{M} + a_1\mathbf{c}_0\right) - \frac{1}{\gamma}\mathbf{c}_0$$
$$b_1 = b_1 c_1 + \frac{1}{\gamma} a_1 c_1 - \frac{1}{\gamma} c_1$$

The Phillips curve (4.14) can be rewritten as (see appendix E):

$$\mathbf{a}_0 = (1-\theta)(1-\beta\theta)\cdot\left[\boldsymbol{\nu}\mathbf{m}_1 - \left(\sum_{s=0}^{k-1}(1-\theta)^s\left(\gamma\boldsymbol{\gamma}_g^{(s+1)\prime}+\boldsymbol{\gamma}_a^{(s)\prime}\right)\right)\right]$$
$$+\beta\theta\boldsymbol{\nu}\mathbf{m}_2 + \beta\theta\left(\sum_{s=0}^{k-1}(1-\theta)^s\right)\mathbf{1}_1^T\mathbf{v}_1\mathbf{1}_3^T\mathbf{v}_0$$
$$a_1 = (1-\theta)(1-\beta\theta)\left(\sum_{s=0}^{k-1}(1-\theta)^s\right)\gamma\mathbf{1}_2^T\mathbf{v}_1 + \beta\theta\left(\sum_{s=0}^{k-1}(1-\theta)^s\right)\mathbf{1}_1^T\mathbf{v}_1\mathbf{1}_3^T\mathbf{v}_1$$

where

$$\boldsymbol{\gamma}_g^{(s)} = \left[\mathbf{0}_{1\times 6s},(0,0,0,0,1,1),\mathbf{0}_{1\times 6(k-s)}\right]$$

$$\boldsymbol{\gamma}_a^{(s)} = \left[\mathbf{0}_{1\times 6s},(1,1,0,0,0,0),\mathbf{0}_{1\times 6(k-s)}\right]'$$

and

$$\mathbf{m}_1 \equiv \begin{bmatrix} \left[\gamma\mathbf{1}_2^T\mathbf{v}_0\boldsymbol{\tau}^{(1)}\mathbf{T}^{(1)}\right] \\ (1-\theta)\left[\gamma\mathbf{1}_2^T\mathbf{v}_0\boldsymbol{\tau}^{(2)}\mathbf{T}^{(2)}\right] \\ (1-\theta)^2\left[\gamma\mathbf{1}_2^T\mathbf{v}_0\boldsymbol{\tau}^{(3)}\mathbf{T}^{(3)}\right] \\ \vdots \\ (1-\theta)^{k-1}\left[\gamma\mathbf{1}_2^T\mathbf{v}_0\boldsymbol{\tau}^{(k)}\mathbf{T}^{(k)}\right] \end{bmatrix},\ \mathbf{m}_2 \equiv \begin{bmatrix} \left[\mathbf{1}_1^T\mathbf{v}_0\mathbf{M}\boldsymbol{\tau}^{(\mathbf{1})}\mathbf{T}^{(1)}\right] \\ (1-\theta)\left[\mathbf{1}_1^T\mathbf{v}_0\mathbf{M}\boldsymbol{\tau}^{(\mathbf{2})}\mathbf{T}^{(2)}\right] \\ (1-\theta)^2\left[\mathbf{1}_1^T\mathbf{v}_0\mathbf{M}\boldsymbol{\tau}^{(\mathbf{3})}\mathbf{T}^{(3)}\right] \\ \vdots \\ (1-\theta)^{k-1}\left[\mathbf{1}_1^T\mathbf{v}_0\mathbf{M}\boldsymbol{\tau}^{(\mathbf{k})}\mathbf{T}^{(k)}\right] \end{bmatrix},$$

$$\boldsymbol{\nu} = \mathbf{1}_{1\times k}$$

The Taylor rule (4.15) imposes that (see appendix E)

$$\mathbf{c}_0 = (1-\rho_r)\left[\phi_\pi\left(\mathbf{a}_0 - \mathbf{1}_3^T\right) + \phi_y\left(\mathbf{b}_0 - \left(\mathbf{1}_1^T + \mathbf{1}_2^T\right)\right)\right] + \mathbf{1}_4^T$$

$$c_1 = \rho_r + (1-\rho_r)\left[\phi_\pi a_1 + \phi_y b_1\right]$$

As shown in appendix F, the matrices $\mathbf{M}$ and $\mathbf{N}$ are pinned down by the mappings:

$$\mathbf{M} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{6\times 6} & \mathbf{0}_{6\times 6k} \\ \mathbf{0}_{6k\times 6} & \left(\mathbb{I} - \mathbf{KD}\right)\mathbf{W}|_{(1:6k,1:6k)} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{KDW}|_{(1:6k,1:6(k+1))} \end{bmatrix}$$

$$\mathbf{N} = \begin{bmatrix} \mathbf{R}_2 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{KDU}|_{(1:6k,1:6)} \end{bmatrix}$$

where

$$\mathbf{R}_1 = \begin{bmatrix} \rho_z & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & \rho_\pi & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & \rho_g & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}$$

$$\mathbf{R}_2 = \begin{bmatrix} \sigma_z & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_a & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\pi & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_r & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{\sigma}_g & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_g \end{bmatrix}$$

and $\mathbf{K}$ is the steady-state matrix of Kalman gains, which is well-known to be equal to

$$\mathbf{K} = \mathbf{PD}' \left[\mathbf{DPD}' + \mathbf{\Sigma}_e\right]^{-1} \tag{4.20}$$

where $\mathbf{\Sigma}_e = \mathbf{QQ}'$. The variance and covariance matrix $\mathbf{P}$ solves the following algebraic Riccati equation:

$$\mathbf{P} = \mathbf{W} \left[\mathbf{P} - \mathbf{PD}' \left[\mathbf{DPD}' + \mathbf{\Sigma}_e\right]^{-1} \mathbf{DP}\right] \mathbf{W}' + \mathbf{UU}' \tag{4.21}$$

and the matrices $\mathbf{W}$ and $\mathbf{U}$ are the reduced-form matrices of the system (4.19):

$\mathbf{W} = \mathbf{A}^{-1}\mathbf{B},\ \mathbf{U} = \mathbf{A}^{-1}\mathbf{C}.$

4.3.1 Average Higher-Order Beliefs

Once one has solved the dispersed information model, the law of motion of the average higher-order beliefs about the endogenous state variables, $\mathbf{s}_t$, can be characterized. In appendix E. I show that for $0 \leq s \leq k$,

$$\mathbf{s}_{t|t}^{(s)} \equiv \begin{bmatrix} \hat{\pi}_{t|t}^{(s)} \\ \hat{y}_{t|t}^{(s)} \\ \hat{R}_{t|t}^{(s)} \end{bmatrix} = \mathbf{v}_0 \boldsymbol{\tau}^{(s)} \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \hat{R}_{t-1} \tag{4.22}$$

where $\mathbf{T}^{(s)} = \left[\mathbf{0}_{6(k+1-s)\times 6s}, \mathbb{I}_{6(k+1-s)}\right]$, and $\boldsymbol{\tau}^{(s)} = \left[\mathbb{I}_{6(k+1-s)}, \mathbf{0}_{6(k+1-s)\times 6s}\right]'$. The average higher-order beliefs one-step-ahead beliefs follows:

$$\mathbf{s}_{t+1|t}^{(s)} \equiv \begin{bmatrix} \hat{\pi}_{t+1|t}^{(s)} \\ \hat{y}_{t+1|t}^{(s)} \\ \hat{R}_{t+1|t}^{(s)} \end{bmatrix} = \mathbf{v}_0 \mathbf{M} \boldsymbol{\tau}^{(\mathbf{s})} \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{1}_3^T \left(\mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \hat{R}_{t-1} \right) \tag{4.23}$$

4.4 Empirical Analysis

I fit the model to observations of the growth rate of output, inflation, and the interest rate. I place a prior distribution on parameters and conduct Bayesian inference. I solve the dispersed information model by truncating the higher order beliefs at $k = 10$. I approximate the posterior distribution by means of a random-walk Metropolis-Hastings method. In this section, I present the data set, the measurement equations, the prior distributions and the posterior distributions for model parameters. Finally, I will study the transmission of monetary impulses to the macroeconomic aggregates implied by the estimated model.

4.4.1 The Data and the Measurement Equation

The data set ranges from the third quarter of 1954 to the fourth quarter of 2007. Data are quarterly and are displayed in Table 4.1:

Table 4.1: Data

Observables	Description	Source
GDP_t	Gross Domestic Product	BEA
$POP_t^{\geq 16}$	Civilian nonistitutional population, 16 yrs and over	BLS
CPI_t	Consumer Price Index-Averages of Monthly Figures	BLS
$FEDRATE_t$	Effective Fed. Funds Rate-Averages of Daily Figures	Board of Gov.

The measurement equations are:

$$\ln\left(\frac{GDP_t}{POP_t^{\geq 16}}\right) - \ln\left(\frac{GDP_{t-1}}{POP_{t-1}^{\geq 16}}\right) = 100\left(\hat{y}_t - \hat{y}_{t-1} + \ln A_0\right)$$

$$100 \ln \frac{CPI_t}{CPI_{t-1}} = 100\left(\hat{\pi}_t + \ln \pi_*\right)$$

$$100 \ln(1 + FEDRATE_t/100) = 100\left(4\hat{R}_t + \ln R_*\right)$$

4.4.2 Priors

Solving the model takes 3-4 minutes for each parameter draw. So one might not be able to generate a sufficiently large number of draws to accurately approximate the posterior moments. This concern is expected to be more severe as the dimensionality of the parameter space is large. In order to address this concern, I restrict the value of seven out of the sixteen parameters of the model. The parameters whose values are fixed are listed in table 4.2

Table 4.2: Fixed Parameter Values

ϕ_π	ϕ_y	ρ_r	ρ_z	ρ_g	ρ_π	β	γ
1.5	0.125	0.75	0.90	0.95	0.95	0.9971	2

These parameter values are quite standard in the empirical literature on DSGE models (An and Schorfheide, 2007 and Smets and Wouters, 2007). I set the priors for the remaining parameters as displayed by the table 4.3.

Table 4.3: Priors

Name	Range	Density	Median	90% Interval
θ	$[0, 1)$	Beta	0.50	$0.06 - 0.93$
σ_a	$\mathbb{R}+$	Inv.Wishart	0.50	$0.06 - 4.00$
σ_z	$\mathbb{R}+$	Inv.Wishart	0.50	$0.06 - 4.00$
σ_a^j	$\mathbb{R}+$	Inv.Wishart	1.50	$0.85 - 5.75$
σ_π	$\mathbb{R}+$	Inv.Wishart	0.06	$0.02 - 0.12$
σ_r	$\mathbb{R}+$	Inv.Wishart	0.12	$0.05 - 0.24$
$\tilde{\sigma}_g$	$\mathbb{R}+$	Inv.Wishart	0.85	$0.42 - 2.05$
σ_g	$\mathbb{R}+$	Inv.Wishart	1.00	$0.49 - 2.41$

The prior distributions are very diffuse as I do not hold precise information about the decomposition of shocks. I aim at learning these parameter values from the likelihood. The prior for the variance of the idiosyncratic technology shock, σ_a^j is centered so as to make the perfect information model able to replicate the average absolute size of price changes that is documented by the literature on the microdata on price changes (Bils and Klenow, 2004).

4.4.3 Posteriors

Given the priors and the likelihood function implied by the model, a closed-form solution for the posterior distributions for parameters cannot be derived. However, one can 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 $1,500$ draws from the posteriors. The posterior medians and 90% credible sets are shown in Table 4.4.

Table 4.4: Posteriors

Name	Median	90% Interval
θ	0.88	$0.86 - 0.89$
σ_a	0.28	$0.22 - 0.32$
σ_z	0.35	$0.27 - 0.42$
σ_a^j	0.62	$0.55 - 0.73$
σ_π	0.02	$0.02 - 0.03$
σ_r	0.02	$0.02 - 0.03$
σ_g	0.22	$0.18 - 0.25$
$\tilde{\sigma}_g$	0.28	$0.19 - 0.34$

Figure 4-1 reports the recursive means of the posterior draws for the eight parameters displayed in table 4.4.

The recursive means of the posterior draws for most of the parameters in table 4.4 show that the Metropolis-Hastings algorithm converges. Slower convergence is observed for σ_a^j and $\tilde{\sigma}_g$.

The posterior median of the signal-to-noise ratio associated with the obser-

Recursive Means

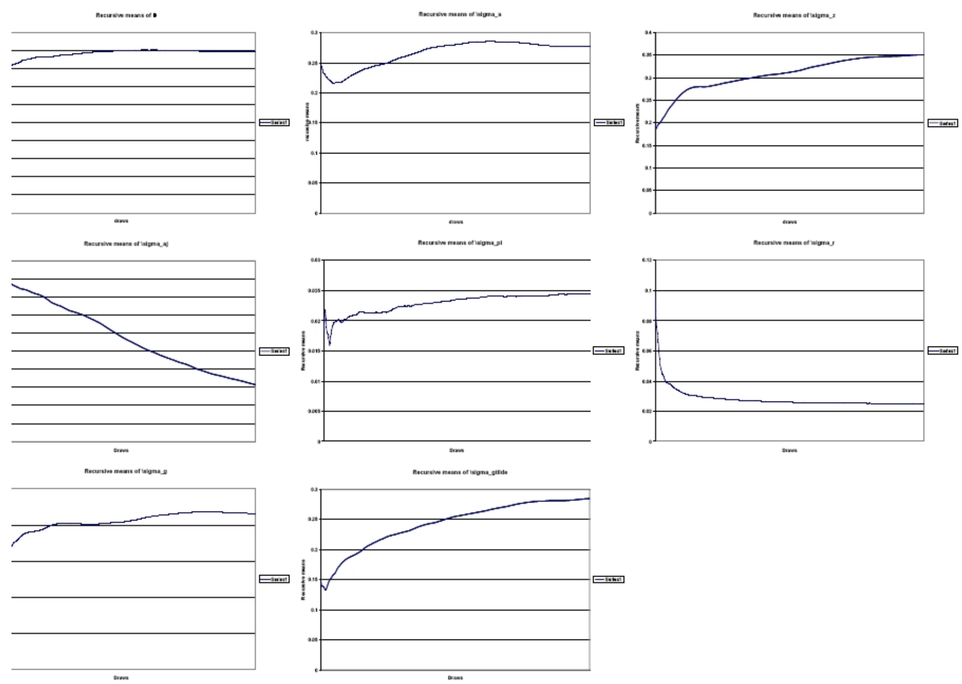

Figure 4-1: Recursive Means

vation equation (4.2) is 1.37. This quantifies the amount of information firms learn about the aggregate level of the technology, A_t, from observing their own productivity shock.

4.4.4 Transmission of Monetary Disturbances

Figure 4-2 shows the responses of real GDP (deviations from the steady state at a quarterly rate), inflation (deviations from the steady state in percentage points at a quarterly rate), and interest rate (percentage deviations from the steady state at a quarterly rate) to a one-standard deviation temporary monetary policy shock $(\eta_{r,t})$. Black squares denote the perfect information model, while the solid lines refer to the dispersed information model. The impulse response functions are computed at the posterior medians displayed in table 4.4.

Impulse Response Functions to a Monetary Policy Shock

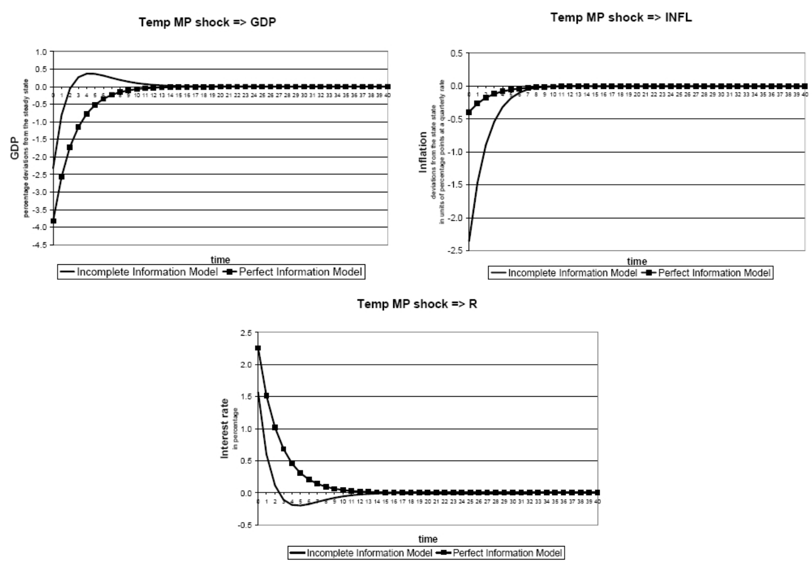

Figure 4-2: Impulse Response Functions to a Monetary Policy Shock

The most striking finding is that monetary policy appears to have weaker real effects when firms are imperfectly informed. One can compute by how much inflation changes owing to a monetary policy shock that reduces output by 1%. The imperfect information model predicts that a monetary policy shock that reduces output by one percent causes inflation to decrease by 1.01 percent. In the perfect information model, the inflation reacts only by 0.10 percent. These numbers show that prices are relatively less sticky in the imperfect information model. Moreover, both these numbers are in line with the survey study conducted by Schorfheide (2008). In the light of that study, the imperfect information model can be regarded as having similar degree of price rigidity as that of the New Keynesian DSGE models with quite flexible prices.

Why are the prices so flexible in the imperfect information model? As standard in New Keynesian models, inflation depends on a current component (CC_t^I) and a forward-looking component (FC_t^I) in the dispersed information model. For the imperfect information model, these two components are defined in equation (4.14). They are:

$$CC_t^i = (1-\theta)(1-\beta\theta)\sum_{k=0}^{\infty}(1-\theta)^k\,\widehat{mc}_{t|t}^{(k)} \tag{4.24}$$

$$FC_t^i = \beta\theta\sum_{k=0}^{\infty}(1-\theta)^k\,\widehat{\pi}_{t+1|t}^{(k+1)} \tag{4.25}$$

with $\hat{\pi}_t = CC^i + FC^i$. In the perfect information model, these two components are defined in equation (4.17):

$$CC_t^p = \frac{(1-\theta)(1-\beta\theta)}{\theta}\widehat{mc}_t \tag{4.26}$$

$$FC_t^p = \beta\mathbb{E}_t\widehat{\pi}_{t+1} \tag{4.27}$$

with $\hat{\pi}_t = CC^p + FC^p$. Table 4.5 shows the value of these components at the posterior medians for the two models.

Table 4.5: Inflation Response Decomposition

Imperfect Information Model (IIM)			
	$100\frac{\partial \hat{\pi}_t}{\partial \eta_{r,t}}$	$100\frac{\partial CC_t^i}{\partial \eta_{r,t}}$	$100\frac{\partial FC_t^i}{\partial \eta_{r,t}}$
Total	-2.3546	-0.5157	-1.8389
In perc.	100.00%	21.90%	78.10%
Perfect Information Model (PIM)			
	$100\frac{\partial \hat{\pi}_t}{\partial \eta_{r,t}}$	$100\frac{\partial CC_t^p}{\partial \eta_{r,t}}$	$100\frac{\partial FC_t^p}{\partial \eta_{r,t}}$
Total	-0.3932	-0.1303	-0.2629
In perc.	100.00%	33.13%	66.87%
Excess Response of Inflation in the IIM			
	$100\Delta\frac{\partial \hat{\pi}_t}{\partial \eta_{r,t}}$	$100\left(\frac{\partial CC_t^i}{\partial \eta_{r,t}} - \frac{\partial CC_t^p}{\partial \eta_{r,t}}\right)$	$100\left(\frac{\partial FC_t^i}{\partial \eta_{r,t}} - \frac{\partial FC_t^p}{\partial \eta_{r,t}}\right)$
Total	-1.9614	-0.3854	-1.5760
In perc.	100.00%	19.65%	80.35%

In both models, monetary disturbances influence inflation mainly through inflation expectations (i.e., the forward looking component). In the imperfect information model, 78% of the inflation adjustment upon the monetary shock stems from the change in inflation expectations. In the perfect information model, this fraction is a bit smaller. In addition, in the imperfect information model, the response of the forward-looking component to monetary shock is about seven times bigger than that in the perfect information model. The excess response of inflation to a monetary shock in the imperfect information model is -1.9614. The overreaction of the inflation expectations in the imperfect information model accounts for 80.35% for the excess responsiveness of the inflation in the imperfect information model.

Why do inflation expectations react so much in the imperfect information model? Since the model nests the perfect information model, the excess responsiveness of inflation in the imperfect information model comes from the information frictions that alter the response of firms beliefs to shocks. In figure 4-3 I plot the impulse response functions of the average first-order beliefs about the exogenous variables to a temporary monetary shocks.

Impulse Response Functions of Higher-Order Beliefs

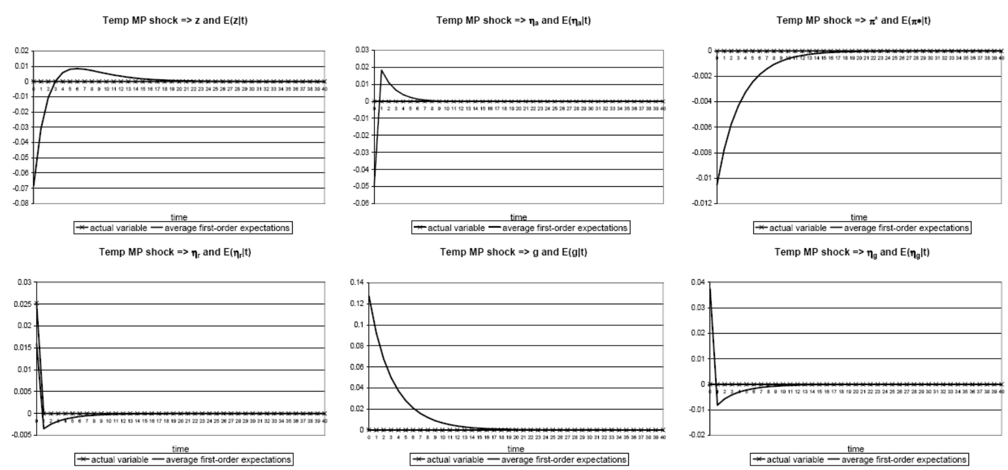

Figure 4-3: Impulse Response Functions of Higher-Order Beliefs

We observe that the average first order expectations about the inflation target, π_t^*, and the persistent government spending shock, g_t, are quite affected by the monetary policy shock. These beliefs deviate from the actual value of the inflation target and that of the persistent government-spending shock.

Upon the monetary disturbance, these beliefs change roughly by about half their conditional standard deviations, σ_g and σ_π.

Which average beliefs are the culprits for the overreaction of the inflation expectations? We denote the contributions of the six classes of average higher-order beliefs $\varphi_{t|t}^{(0:k)}$ to the response of inflation expectations as

$$\left(C_{z,t}^f, C_{a,t}^f, C_{\pi^*,t}^f, C_{r,t}^f, C_{\tilde{g},t}^f, C_{g,t}^f\right).$$

Table 4.6 displays the result of this exercise:

Table 4.6: Decomposition of the Components CC and FC

Current Component CC_t^d						
$100\frac{\partial CC_t^i}{\partial \eta_{r,t}}$	$100\frac{\partial C_{z,t}^c}{\partial \eta_{r,t}}$	$100\frac{\partial C_{a,t}^c}{\partial \eta_{r,t}}$	$100\frac{\partial C_{\pi^*,t}^c}{\partial \eta_{r,t}}$	$100\frac{\partial C_{r,t}^c}{\partial \eta_{r,t}}$	$100\frac{\partial C_{\tilde{g},t}^c}{\partial \eta_{r,t}}$	$100\frac{\partial C_{g,t}^c}{\partial \eta_{r,t}}$
−0.5157	−0.0689	0.0006	−0.1791	−0.0820	−0.1803	−0.0060
100.00%	13.36%	−0.12%	34.73%	15.90%	34.97%	1.16%
Forward-Looking Component FC_t^d						
$100\frac{\partial FC_t^i}{\partial \eta_{r,t}}$	$100\frac{\partial C_{z,t}^f}{\partial \eta_{r,t}}$	$100\frac{\partial C_{a,t}^f}{\partial \eta_{r,t}}$	$100\frac{\partial C_{\pi^*,t}^f}{\partial \eta_{r,t}}$	$100\frac{\partial C_{r,t}^f}{\partial \eta_{r,t}}$	$100\frac{\partial C_{\tilde{g},t}^f}{\partial \eta_{r,t}}$	$100\frac{\partial C_{g,t}^f}{\partial \eta_{r,t}}$
−1.8389	−0.0056	0.0000	−0.7708	−0.2648	−0.7979	0.0002
100.00%	0.30%	0.00%	41.92%	14.40%	43.39%	−0.01%

The table also reports the contributions of the six classes of average higher-order beliefs to the change in the current component, CC_t^i[4]. One can observe

[4] How does one quantify these components? Recall that the current component of inflation, CC_t^i, depends on $\widehat{mc}_{t|t}^{(0:k)}$, and some deep parameters. Yet the hierarchy of average higher-order beliefs about real marginal costs, $\widehat{mc}_{t|t}^{(0:k)}$, depend on the hierarchy of average higher-order beliefs about real output $\widehat{y}_{t|t}^{(0:k)}$, and some exogenous variables: $\hat{\zeta}_{t|t}^{(k+1)}$, $z_{t|t}^{(k)}$, and $\eta_{a,t|t}^{(k)}$. See (4.13) equation. Equation (4.22) pins down the response of the beliefs $\widehat{y}_{t|t}^{(0:k)}$. Equation (4.19)

that the change in the forward-looking component is mainly due to the response of average beliefs about the inflation-target and the persistent government-spending shock. The excess response of inflation in the imperfect information model (table 4.5) is almost entirely explained by the response these two average beliefs.

The results in table 4.6 suggest that upon a temporary monetary shock, $\eta_{r,t}$, firms are unsure about whether the change in the interest rate is due to a an inflation-targeting shock or a government-spending shock. This confusion causes inflation expectations to react very much to a monetary policy shock and reduces the real effects of money.

4.4.5 Costs of Disinflation

A fall in the inflation target, π_t^*, is generally interpreted as a structural change in policy aimed at reducing inflation. Figure 4-4 shows the responses of real GDP (deviations from the steady state at a quarterly rate), inflation (deviations from the steady state in percentage points at a quarterly rate), and interest rate (percentage deviations from the steady state at a quarterly rate) to a structural disinflation policy ($\varepsilon_{\pi,t}$) that reduces inflation by 5% in the long run. Black squares denote the perfect information model, while the solid lines refer to the

determines how the average beliefs $\hat{\zeta}_{t|t}^{(k+1)}$, $z_{t|t}^{(k)}$, and $\eta_{a,t|t}^{(k)}$ react to structural shocks. Equation (4.23) sets the response of FC_t^d, which depends on $\widehat{\pi}_{t|t}^{(0:k)}$, to exogenous state variables, φ_t.

dispersed information model. These impulse response functions are computed at the posterior medians displayed in table 4.4.

Impulse Response Functions to an Inflation-Targeting Shock

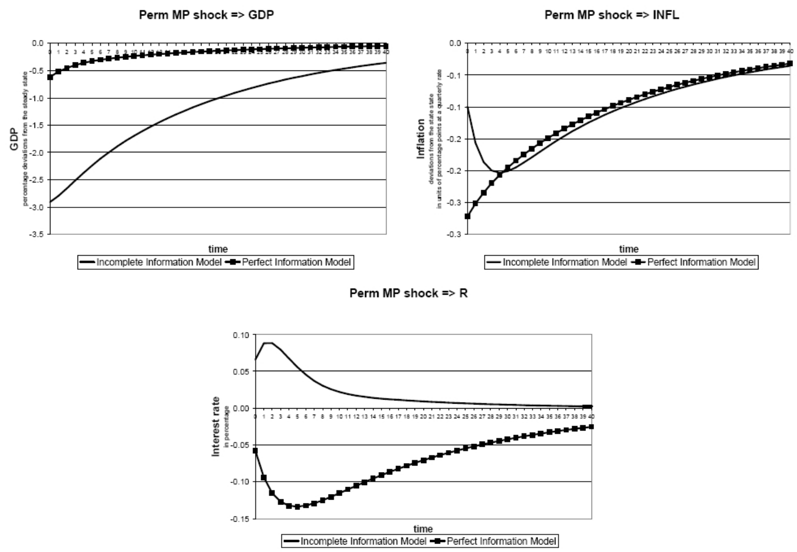

Figure 4-4: Impulse Response Functions to an Inflation-Targeting Shock

Figure 4-4 shows that GDP falls more in the imperfect information model, suggesting that costs of disinflation are larger. The cumulative effect of the policy upon real output is -56.54% for the imperfect information model, and -8.34% for the perfect information model. This leads to conclude that the imperfect information model predicts that the output loss associated to a disinflation policies is bigger than what is estimated by the perfect information model.

The reasons for this finding can be understood by looking at the responses of the average expectations to a disinflation policy, which are shown in figure 4-5.

Impulse Response Functions of Higher-Order Beliefs

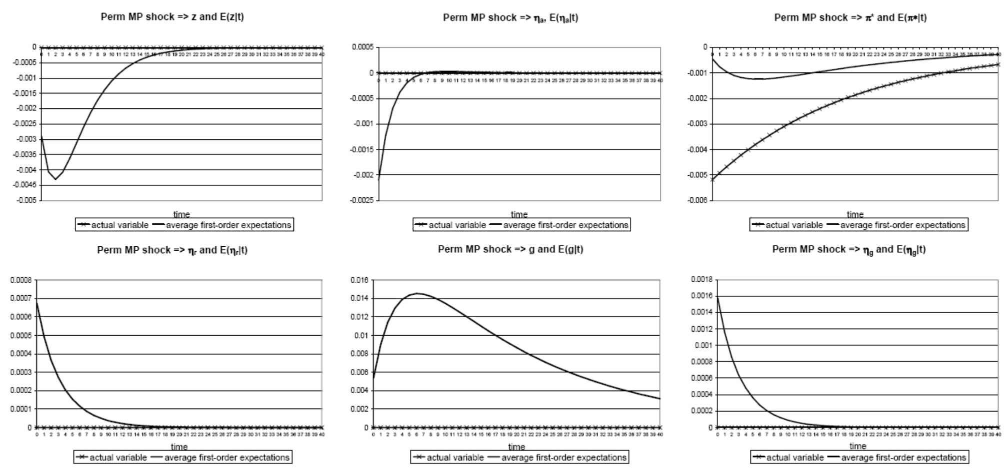

Figure 4-5: Impulse Response Functions of Higher-Order Beliefs

It is clear that when the central bank raises the interest rate to disinflate the economy, firms fail to interpret this as such. One can see this by observing that the average expectations about π_t^* move very weakly upon a persistent monetary policy shock. The plot also shows that firms learn extremely slowly about the nature of the shock in the subsequent periods.

Table 4.7 plots the inflation response decomposition upon the inflation-targeting shock.

Table 4.7: Inflation Response Decomposition

In the perfect information model inflation expectations react very strongly and fully anticipate the disinflation policy. This curbs the cost of disinflation. Inflation expectations react so strongly that the central bank has to cut the interest rate to disinflate (see figure 4-4). In the imperfect information model inflation expectations fail to anticipate the drop in inflation and the central bank has to forcefully raise the interest rate to disinflate the economy. The last row of table 4.7 shows that the lack of adjustment of inflation expectations fully account for the little response of inflation in the imperfect information model.

4.5 Concluding Remarks

This chapter shows that modelling agents' expectations may have a great deal of implications for the transmission mechanism of monetary policy. I obtain two main results. First, the presence of imperfect information reduces the real effects of monetary policy. Second, the output loss associated with a monetary policy of disinflation is larger when one considers imperfect information. Both of these results arise because imperfectly informed firm cannot recognize the exact nature of the shocks. More specifically, upon a monetary policy shock firms are not sure about whether the change in the interest rate is due to a permanent monetary shock or a government-spending shock. This confusion causes inflation to react very much to a monetary policy shock and hence real effects of money are weak. After a disinflation policy, the inflation expectations fail to anticipate the effects of the policy on prices. Hence, this policy turns out to be very costly in terms of output loss.

Future work requires further model validation. In particular, it is key to assess the capability of the model to capture the dynamics of the first moment of the distribution of the survey of professional forecasters. From a Bayesian perspective, this assessment can be done by using the date on expectations as observables and run posterior predictive checks (An and Schorfheide, 2007). The outcome of this check is by no means obvious. Del Negro and Eusepi (2009) show that the imperfect information model in Erceg and Levin (2003) does worse than a perfect information model in fitting observed expectations.

It is important to emphasize that in the imperfect information model different monetary policy rules affect the information content of the monetary policy instrument and hence how monetary disturbances affect macroeconomic variables. A fascinating extension of this project would be characterizing the optimal monetary policy in such a framework.

Bibliography

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.

Amato, J. D., and H. S. Shin (2003): "Public and Private Information In Monetary Policy Models," BIS Working Paper series 138.

——— (2006): "Imperfect Common Knowledge and the Information Value of Prices," *Economic Theory*, 27, 213–241.

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.

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.

Del Negro, M., and S. Eusepi (2009): "Modeling Observed Inflation Expectations," New York FED mimeo.

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.

Erceg, C. J., and A. T. Levin (2003): "Imperfect Credibility and Inflation Persistence," *Journal of Monetary Economics*, 50(4), 915–944.

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.

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

——— (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*.

Maćkowiak, B., and M. Wiederholt (2009): "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.

Morris, S., and H. S. Shin (2003a): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics, the Eighth World Congress*, ed. by M. D. et al. Cambridge University Press, Cambridge.

——— (2003b): "Social Value of Public Information," *American Economic Review*, 92(5), 1521–1534.

——— (2008): "Coordinating Expectations in Monetary Policy," in *Central Banks as Economic Institutions*, ed. by J.P.Touffut. Edward Elgar Publishing, Northampton.

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. (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.

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

Sims, C. A. (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.

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.

——— (2005): "Central-Bank Communication and Policy Effectiveness," Federal Reserve Bank of Kansas City Symposium mimeo.

Appendices

Appendix A Deriving the law of motion of price and output in the ICKM

The first-order necessary condition[5] of the price-setting problem (2.11)-(2.16) in the ICKM is:

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

From the solution to the representative household's problem (2.2)-(2.5), the labor supply can be easily shown to be $W_t/P_t=\alpha Y_t H_t^\eta$. Substituting this result and using the equation (2.12) into the equation above yield:

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

Define the stationary variables:

$$y_t\equiv\frac{Y_t}{A_t},\ y_{i,t}\equiv\frac{Y_{i,t}}{A_t}\ ,\ p_{i,t}=\frac{P_{i,t}}{P_t},\ h_t=H_t \tag{4.28}$$

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

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

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

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

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

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

and then

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

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

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

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 (4.28) 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 (2.10):

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

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

If one log-linearizes equation (2.1) around the deterministic steady-state, one ob-

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

tains $\hat{p}_t = \int \hat{p}_{i,t} di$. 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 (2.20) in the main text. Furthermore, I can combine equations (2.20) and (2.10) 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 (2.21) in the main text.

Appendix 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 (2.20)-(2.21). Define the vector $\mathbf{F}_t$ as

$$\mathbf{F}_t \equiv \sum_{j=1}^{\infty} (1-\lambda)^{j-1} \lambda \mathbf{X}_t^{(j)} \tag{4.29}$$

$$\text{where } \mathbf{X}_t \equiv [m_t, m_{t-1}, a_t]' \tag{4.30}$$

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

$$\widehat{y}_t = \widehat{p}_t \tag{4.31}$$

$$\widehat{p}_t = \mathbf{r}'\overline{\mathbf{X}}_t \tag{4.32}$$

$$\overline{\mathbf{X}}_t = \overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t \tag{4.33}$$

where

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

$$\overline{\mathbf{B}} \equiv \begin{bmatrix} \mathbf{B}_{3x3} & \mathbf{0}_{3x3} \\ \mathbf{G}_{3x3} & \mathbf{H}_{3x3} \end{bmatrix}, \quad \overline{\mathbf{b}} = \left[\mathbf{b}\, ' \vdots \; \mathbf{d}'\right]' \tag{4.34}$$

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

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

where $\mathbf{G}$, $\mathbf{H}$, and $\mathbf{d}$ are matrices that are not known yet. Equation (4.31) stems from the log-linearized version of equation (2.10), 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 (4.32) can be derived by equation (2.20) 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 (2.10).

Recall that the signal structure is specified in equations (2.16). Thus, the firms' observation equations are

$$\mathbf{z}_{i,t} = \mathbf{D}\overline{\mathbf{X}}_t + \mathbf{e}_{i,t} \tag{4.35}$$

where

$$\mathbf{D} \equiv \left[\begin{array}{ccc} \mathbf{D}_1 & \vdots & \mathbf{0}_{2x3} \end{array} \right] \text{ and } \mathbf{D}_1 = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right] \tag{4.36}$$

$$\mathbf{e}_{i,t} \backsim \mathcal{N}\left(\mathbf{0}, \mathbf{\Sigma}_e\right), iid \text{ for all } t, \text{ and } i, \;\; \mathbf{\Sigma}_e = \left[\begin{array}{cc} \tilde{\sigma}_m^2 & 0 \\ 0 & \tilde{\sigma}_a^2 \end{array} \right] \tag{4.37}$$

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 (4.33), optimal firms' behaviors must exactly aggregate to the conjectured law of motion (4.33). 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*

$$\overline{\mathbf{X}}_{t|t}\left(i\right) = \overline{\mathbf{X}}_{t|t-1}\left(i\right) + \mathbf{k}\left[\mathbf{z}_t\left(i\right) - \mathbf{D}\overline{\mathbf{X}}_{t|t-1}\left(i\right)\right] \tag{4.38}$$

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:

$$\overline{\mathbf{X}}_{t|t-1}\left(i\right) = \overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1|t-1}\left(i\right) \tag{4.39}$$

I can plug the (4.39) into the (4.38) to get the law of motion for firm i's estimate of the current state vector

$$\overline{\mathbf{X}}_{t|t}\left(i\right) = \overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1|t-1}\left(i\right) + \mathbf{k}\left[\mathbf{z}_t\left(i\right) - \mathbf{D}\overline{\mathbf{X}}_{t|t-1}\left(i\right)\right] \tag{4.40}$$

By integrating the (4.40) over firms (i.e. $\int \overline{\mathbf{X}}_{t|t}\left(i\right) di \equiv \overline{\mathbf{X}}_{t|t}$) one gets

$$\overline{\mathbf{X}}_{t|t} = \overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1|t-1} + \mathbf{kD}\left[\overline{\mathbf{X}}_t - \overline{\mathbf{X}}_{t|t-1}\right] \tag{4.41}$$

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

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

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

By using the transition equation (4.33) to get rid of $\overline{\mathbf{X}}_t$ in the equation (4.41) I obtain

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

Then by integrating the (4.39), which yields the average prior forecast (i.e. $\overline{\mathbf{X}}_{t|t-1} = \overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1|t-1}$), one notices that the above equation can be rewritten as

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

Gathering the common terms yields

$$\overline{\mathbf{X}}_{t|t} = [\mathbf{I} - \mathbf{kD}]\, \overline{\mathbf{B}\mathbf{X}}_{t-1|t-1} + \mathbf{kD}\left[\overline{\mathbf{B}\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t\right] \tag{4.42}$$

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

$$\boldsymbol{\varphi}'\overline{\mathbf{X}}_t^{(1)} = \mathbf{F}_t \tag{4.43}$$

It is easy to prove that equation (4.43) is indeed true by working as follows

$$\begin{aligned}\boldsymbol{\varphi}'\overline{\mathbf{X}}_t^{(1)} &= \left[(\lambda) \cdot \mathbf{I}_3 \,\vdots\, (1-\lambda) \cdot \mathbf{I}_3\right] \cdot \begin{bmatrix} \mathbf{X}_t^{(1)} \\ \cdots \\ \mathbf{F}_t^{(1)} \end{bmatrix} \\ \boldsymbol{\varphi}'\overline{\mathbf{X}}_t^{(1)} &= \lambda \mathbf{X}_t^{(1)} + (1-\lambda)\, \mathbf{F}_t^{(1)}\end{aligned} \tag{4.44}$$

Let me introduce the following notations:

$$x_{t|t}^{(k-1)} \equiv x_t^{(k)}, \quad \forall k \geq 1; \quad x_t^{(0)} \equiv x_t \tag{4.45}$$

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 (4.29)

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

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

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

Now recall equation (4.45) to finally write

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

Comparing this equation with the (4.29) concludes the proof of (4.43). Now one can plug equation (4.42) into equation (4.43) to get

$$\mathbf{F}_t = \left[\boldsymbol{\varphi}' - \widetilde{\mathbf{k}}\mathbf{D}\right]\overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1|t-1} + \widetilde{\mathbf{k}}\mathbf{D}\left[\overline{\mathbf{B}}\overline{\mathbf{X}}_{t-1} + \overline{\mathbf{b}}\mathbf{u}_t\right] \tag{4.46}$$

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

FACT 1

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

FACT 2

$$\mathbf{D}\overline{\mathbf{B}} = \left[\mathbf{D}_1\mathbf{B} \vdots\; \mathbf{0}_{2x3}\right] \tag{4.47}$$
$$= \left[\mathbf{B}^{\dagger} \vdots\; \mathbf{0}_{2x3}\right]$$

where $\mathbf{B}^{\dagger} \equiv \begin{bmatrix} \mathbf{B}_1' & \mathbf{B}_3' \end{bmatrix}'$ and $\mathbf{B}_j$ stands for the j-th row of $\mathbf{B}$.

FACT 3

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

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

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

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

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

By collecting all these results one can rewrite equation (4.46) as follows

$$\mathbf{F}_t = \left[\lambda\mathbf{B} + (1-\lambda)\,\mathbf{G} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\right]\mathbf{X}_{t-1|t-1} + (1-\lambda)\,\mathbf{H}\mathbf{F}_{t-1|t-1} + \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\mathbf{X}_{t-1} + \widetilde{\mathbf{k}}\mathbf{u}_t \quad (4.48)$$

Next, I will work out the vector $\mathbf{F}_{t-1}$ from $\mathbf{F}_{t-1|t-1}$, since I want to rewrite equation (4.48) in a form that is comparable to that conjectured in equation (4.33) so as I can compare my initial guess. One should start from equation (4.43) to get

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

By lagging the last equation by one period, one gets

$$(1-\lambda)\cdot\mathbf{F}_{t-1|t-1} = \mathbf{F}_{t-1} - \lambda\mathbf{X}_{t-1|t-1} \quad (4.49)$$

I can now plug equation (4.49) into equation (4.48) to get

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

$$\mathbf{F}_t = \left[\lambda\mathbf{B} + (1-\lambda)\,\mathbf{G} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} - \lambda\mathbf{H}\right]\mathbf{X}_{t-1|t-1} + \mathbf{H}\cdot\mathbf{F}_{t-1} + \widetilde{\mathbf{k}}\mathbf{B}^{\dagger}\mathbf{X}_{t-1} + \widetilde{\mathbf{k}}\mathbf{u}_t \quad (4.50)$$

Now equation (4.50) has the same form as the bottom rows of equation (4.33) 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:

$$\mathbf{G} = \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} \quad (4.51)$$

$$\mathbf{d} = \widetilde{\mathbf{k}} \quad (4.52)$$

and

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

By substituting (4.51) into the last equation one obtains

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

$$\mathbf{H} \stackrel{!}{=} \mathbf{B} - \widetilde{\mathbf{k}}\mathbf{B}^{\dagger} \quad (4.53)$$

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

$$\mathbf{k} = \mathbf{P}\mathbf{D}' \left[\mathbf{D}\mathbf{P}\mathbf{D}' + \mathbf{\Sigma}_e\right]^{-1} \tag{4.54}$$

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

$$\mathbf{P} = \overline{\mathbf{B}}\left[\mathbf{P} - \mathbf{P}\mathbf{D}'\left[\mathbf{D}\mathbf{P}\mathbf{D}' + \mathbf{\Sigma}_e\right]^{-1}\mathbf{D}\mathbf{P}\right]\overline{\mathbf{B}}' + \overline{\mathbf{b}}\mathbf{\Sigma}_u\overline{\mathbf{b}}' \tag{4.55}$$

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

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

The loop to numerically find out a REE is the following: given a set of parameter values and a guess for the Kalman-gain matrix $\mathbf{k}^0$, one has to characterize the matrices $\mathbf{G}$, $\mathbf{H}$, and $\mathbf{d}$ through equations (4.51)-(4.53). Then one has to solve the algebraic Riccati equation (4.55) for $\mathbf{P}$ and obtain a new Kalman-gain matrix $\mathbf{k}^*$ through the equation (4.54). 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 (4.33)-(4.34) through (4.51)-(4.55), which combined with the equations (4.31)-(4.32) delivers the equilibrium dynamics of the log-deviations of real output and inflation.

Appendix C Information flows

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

$$\kappa_a \equiv H\left(a_t|z_{a,i}^{t-1}\right) - H\left(a_t|z_{a,i}^{t}\right) \tag{4.56}$$

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

$$H\left(a_t|z_{a,i}^{t}\right) \equiv \frac{1}{2}\log_2\left[2\pi e \cdot VAR\left(a_t|z_{a,i}^{t}\right)\right] \tag{4.57}$$

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 $g_a\left(.\right)$ 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

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

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:

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

and finally,

$$\tilde{\sigma}_a^2 = \frac{\left[2VAR\left(a_t|z_{a,i}^{t}\right) + \sigma_a^2\right]^2}{4\sigma_a^2} - \frac{\sigma_a^2}{4} \tag{4.58}$$

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 (4.56) and (4.57) yields

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

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:

$$VAR\left(a_t|z_i^t\right) = \frac{\sigma_a^2}{2^{2\kappa_a}-1} \tag{4.59}$$

Plugging this result into equation (4.58) leads to:

$$\kappa_a = \frac{1}{2}\log_2\left[\frac{1}{\left(\frac{\tilde{\sigma}_a^2}{\sigma_a^2}+\frac{1}{4}\right)^{\frac{1}{2}}-\frac{1}{2}}+1\right] \tag{4.60}$$

This is the mapping g_a in equation (3.3).

An analytical closed-form solution for the mapping g_m in equation (3.3) 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:

$$H\left(m_t|z_{1,i}^{\tau}\right) = \frac{1}{2}\log_2\left[2\pi e\cdot VAR\left(m_t|z_{1,i}^{\tau}\right)\right] \tag{4.61}$$

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

$$\mathbf{P}_{t|t} \equiv \mathbf{P}_{t|t-1}-\mathbf{P}_{t|t-1}\mathbf{D}'\left[\mathbf{D}\mathbf{P}_{t|t-1}\mathbf{D}'+\mathbf{\Sigma}_e\right]^{-1}\mathbf{D}\mathbf{P}_{t|t-1} \tag{4.62}$$

where the matrices $\mathbf{D}$ and $\mathbf{\Sigma}_e$ have been defined in (4.36) and in (4.37), respectively.

Thus, after one has characterized the fixed point as discussed in appendix B, one can use the resulting matrix $\mathbf{P}$ and equation (4.62) to pin down the conditional variances $VAR\left(m_t|z^{\tau}_{1,i}\right)$, for $\tau \in \{t-1,t\}$, the condition entropies $H\left(m_t|z^{\tau}_{m,i}\right)$, for $\tau \in \{t-1,t\}$, through equation (4.61), and finally the information flow $\kappa_m \equiv H\left(m_t|z^{t-1}_{m,i}\right) - H\left(m_t|z^{t}_{m,i}\right)$.

Appendix D The Imperfect Common Knowledge Phillips Curve

The linearized price index can be written as:

$$0 = -\theta\hat{\pi}_t + (1-\theta)\int \hat{p}^*_{j,t}dj$$

By rearranging:

$$\int \hat{p}^*_{j,t}dj = \frac{\theta}{1-\theta}\hat{\pi}_t$$

Recall that we defined $\hat{p}^*_{j,t} = \ln P^*_{j,t} - \ln P_t$ and $\hat{\pi}_t = \ln P_t - \ln P_{t-1} - \ln \pi_*$,

$$\int \ln P^*_{j,t}dj - \ln P_t = \frac{\theta}{1-\theta}\left(\ln P_t - \ln P_{t-1} - \ln \pi_*\right)$$

and then

$$\int \ln P^*_{j,t}dj = \frac{1}{1-\theta}\ln P_t - \frac{\theta}{1-\theta}\left(\ln P_{t-1} + \ln \pi_*\right)$$

By rearranging:

$$\ln P_t = \theta\left(\ln P_{t-1} + \ln \pi_*\right) + (1-\theta)\int \left(\ln P^*_{j,t}\right)dj \tag{4.63}$$

The price-setting equation is:

$$\begin{aligned} &\mathbb{E}\left[\xi_t\left[(1-\nu)+\nu\frac{mc_{j,t}}{p^*_{j,t}}\right]y_{j,t}|\mathcal{I}_{j,t}\right] + \\ + \quad &\mathbb{E}\left[\sum_{s=1}^{\infty}(\beta\theta)^s\xi_{t+s}\left[(1-\nu)\pi^s_* + \nu\frac{mc_{j,t+s}}{p^*_{j,t}}\left(\Pi^s_{\tau=1}\pi_{t+\tau}\right)\right]y_{j,t+s}|\mathcal{I}_{j,t}\right] = 0 \end{aligned}$$

We can write

$$\begin{aligned} \mathbb{E} \quad &\left(\xi_t\left[1-\nu+\nu mc_{j,*}e^{\widehat{mc}_{j,t}-\widehat{p}^*_{j,t}}\right]y_{j,t}|\mathcal{I}_{j,t}\right) + \\ + \quad \mathbb{E}&\left(\sum_{s=1}^{\infty}(\beta\theta)^s\xi_{t+s}\left[(1-\nu)\pi^s_* + \nu mc_{j,*}e^{\widehat{mc}_{j,t+s}-\widehat{p}^*_{j,t}+\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau}}\right]y_{j,t+s}|\mathcal{I}_{j,t}\right) = 0 \end{aligned}$$

where the variables denoted with $\widehat{\ }$ are the log-linear deviations from the steady state.

First realize that the square brackets are equal to zero at the steady state and hence we do not care about the terms outside them. Taking the derivatives yields:

$$\mathbb{E}\left[\left[\nu mc_{j,*}\left(\widehat{mc}_{j,t}-\widehat{p}^{*}_{j,t}\right)\right]+\sum_{s=1}^{\infty}\left(\beta\theta\right)^{s}\left[\nu mc_{j,*}\left(\widehat{mc}_{j,t+s}-\widehat{p}^{*}_{j,t}+\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau}\right)\right]|\mathcal{I}_{j,t}\right]=0$$

We can take the term $\widehat{p}_{j,t}$ out of the sum operator in the second term and gather the common term to obtain:

$$\mathbb{E}\left[\nu mc_{j,*}\widehat{mc}_{j,t}-\nu mc_{j,*}\frac{1}{1-\beta\theta}\widehat{p}^{*}_{j,t}+\nu mc_{j,*}\sum_{s=1}^{\infty}\left(\beta\theta\right)^{s}\left(\widehat{mc}_{j,t+s}+\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau}\right)|\mathcal{I}_{j,t}\right]=0$$

Recall that $\widehat{p}^{*}_{j,t}=\ln P^{*}_{j,t}-\ln P_{t}$ and cannot be taken out of the expectation operator. We need to take two steps to do that. First, we write:

$$\begin{aligned}&\mathbb{E}\left[\left(\nu mc_{j,*}\widehat{mc}_{j,t}-\nu mc_{j,*}\frac{1}{1-\beta\theta}\left(\ln P^{*}_{j,t}-\ln P_{t}\right)|\mathcal{I}_{j,t}\right)\right]+\\ +\quad&\mathbb{E}\left[\nu mc_{j,*}\sum_{s=1}^{\infty}\left(\beta\theta\right)^{s}\left(\widehat{mc}_{j,t+s}+\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau}\right)|\mathcal{I}_{j,t}\right]=0\end{aligned}$$

and then

$$\begin{aligned}\frac{\nu mc_{j,*}}{1-\beta\theta}\ln P^{*}_{j,t}\quad=\quad&\mathbb{E}\left[\left(\nu mc_{j,*}\widehat{mc}_{j,t}+\nu mc_{j,*}\frac{1}{1-\beta\theta}\ln P_{t}\right)|\mathcal{I}_{j,t}\right]+\\&+\mathbb{E}\left[\nu mc_{j,*}\sum_{s=1}^{\infty}\left(\beta\theta\right)^{s}\left(\widehat{mc}_{j,t+s}+\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau}\right)|\mathcal{I}_{j,t}\right]\end{aligned}$$

By simplifying the common coefficients:

$$\ln P^{*}_{j,t}=\left(1-\beta\theta\right)\mathbb{E}\left[\widehat{mc}_{j,t}+\frac{1}{1-\beta\theta}\ln P_{t}+\sum_{s=1}^{\infty}\left(\beta\theta\right)^{s}\left(\widehat{mc}_{j,t+s}+\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau}\right)|\mathcal{I}_{j,t}\right]\tag{4.64}$$

Rolling this equation one step ahead yields:

$$\begin{aligned}\ln P^{*}_{j,t+1}\quad=\quad&\left(1-\beta\theta\right)\mathbb{E}\left[\widehat{mc}_{j,t+1}+\frac{1}{1-\beta\theta}\ln P_{t+1}|\mathcal{I}_{j,t}\right]+\\&+\mathbb{E}\left[\sum_{s=1}^{\infty}\left(\beta\theta\right)^{s}\left(\widehat{mc}_{j,t+s+1}+\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau+1}\right)|\mathcal{I}_{j,t}\right]\end{aligned}$$

Take firm j's conditional expectation at time t on both sides and apply the law of iterated expectations:

$$\begin{aligned}\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right) &= (1-\beta\theta)\,\mathbb{E}\left[\widehat{mc}_{j,t+1}+\frac{1}{1-\beta\theta}\ln P_{t+1}|\mathcal{I}_{j,t}\right]+\\ &\quad +\mathbb{E}\left[\sum_{s=1}^{\infty}(\beta\theta)^s\left(\widehat{mc}_{j,t+s+1}+\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau+1}\right)|\mathcal{I}_{j,t}\right]\end{aligned}$$

We can take $\widehat{mc}_{j,t+1}$ inside the sum operator and write:

$$\begin{aligned}\mathbb{E}\quad&\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right) = (1-\beta\theta)\,\mathbb{E}\left[\frac{1}{1-\beta\theta}\ln P_{t+1}|\mathcal{I}_{j,t}\right]+\\ &+\mathbb{E}\left[\frac{1}{\beta\theta}\sum_{s=1}^{\infty}(\beta\theta)^s\,\widehat{mc}_{j,t+s}+\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{s=1}^{\infty}(\beta\theta)^s\,\mathbb{E}\left[\widehat{mc}_{j,t+s}|\mathcal{I}_{j,t}\right] &= \frac{\beta\theta}{1-\beta\theta}\left[\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right)-\mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right]+\\ &\quad -\beta\theta\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\end{aligned}\tag{4.65}$$

The equation (4.64) can be rewritten as:

$$\begin{aligned}\ln P^*_{j,t} = (1-\beta\theta)&\left\{\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right]+\frac{1}{1-\beta\theta}\mathbb{E}\left[\ln P_t|\mathcal{I}_{j,t}\right]\right\}+\\ &+(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\,\mathbb{E}\left[\widehat{mc}_{j,t+s}|\mathcal{I}_{j,t}\right]+\\ &+(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]\end{aligned}\tag{4.66}$$

By substituting the result in equation (4.65) we obtain:

$$\begin{aligned}\ln P^*_{j,t} = (1-\beta\theta)&\left[\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right]+\frac{1}{1-\beta\theta}\mathbb{E}\left[\ln P_t|\mathcal{I}_{j,t}\right]\right]+\\ &+\beta\theta\left[\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right)-\mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right]+\\ &-(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]+\\ &+(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]\end{aligned}\tag{4.67}$$

Consider the last term:

$$(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]=(1-\beta\theta)\,\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]+$$
$$+\,(1-\beta\theta)\sum_{s=2}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]$$

and then

$$(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]=(1-\beta\theta)\,\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]+$$
$$+\,(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\left(\sum_{\tau=1}^{s}\left[\left(\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\right)\right]+\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]\right)$$

Therefore we can write that

$$(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]=(1-\beta\theta)\,\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]+$$
$$+\,(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]+$$
$$+\,(1-\beta\theta)\left(\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\right)\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]$$

Note that

$$\left(\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\right)=\frac{(\beta\theta)^2}{1-\beta\theta}$$

Hence,

$$(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]=(1-\beta\theta)\,\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]+$$
$$+\,(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]+$$
$$+\,(\beta\theta)^2\,\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]$$

and by simplifying:

$$(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]=\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]+$$
$$+\,(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]$$

We substitute this result into the original equation to get:

$$\begin{aligned}\ln P^*_{j,t} &= (1-\beta\theta)\left[\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right] + \frac{1}{1-\beta\theta}\mathbb{E}\left[\ln P_t|\mathcal{I}_{j,t}\right]\right] + \\ &+ \beta\theta\left[\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right) - \mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right] + \\ &- (1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right] + \\ &+ \beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right] + \\ &+ (1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\end{aligned} \tag{4.68}$$

After simplifying we get:

$$\begin{aligned}\ln P^*_{j,t} &= (1-\beta\theta)\left[\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right] + \frac{1}{1-\beta\theta}\mathbb{E}\left[\ln P_t|\mathcal{I}_{j,t}\right]\right] \\ &+ \beta\theta\left[\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right) - \mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right] + \beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]\end{aligned} \tag{4.69}$$

We can rearrange:

$$\begin{aligned}\ln P^*_{j,t} &= (1-\beta\theta)\,\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right] + \mathbb{E}\left[\ln P_t|\mathcal{I}_{j,t}\right] \\ &+ \beta\theta\left[\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right) + \mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right] - \mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right]\end{aligned} \tag{4.70}$$

Note that by definition $\hat{\pi}_{t+1} \equiv \ln P_{t+1} - \ln P_t - \ln \pi_*$. Hence we can show that

$$\begin{aligned}\ln P^*_{j,t} &= (1-\beta\theta)\cdot\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right] + (1-\beta\theta)\,\mathbb{E}\left[\ln P_t|\mathcal{I}_{j,t}\right] \\ &+ \beta\theta\cdot\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right) - \beta\theta\ln\pi_*\end{aligned} \tag{4.71}$$

We denote the firm j's average k-th order expectation about an arbitrary variable $\hat{x}_t$ as

$$\mathbb{E}^{(k)}\left(\hat{x}_t|\mathcal{I}_{j,t}\right) \equiv \int \mathbb{E}\left(\int \mathbb{E}\left(\ldots\left(\int \mathbb{E}\left(\hat{x}_t|\mathcal{I}_{j,t}\right)dj\right)\ldots|\mathcal{I}_{j,t}\right)dj|\mathcal{I}_{j,t}\right)dj$$

where expectations and integration across firms are taken k times.

Let us denote the **average reset price** as $\ln P^*_t = \int \ln P^*_{j,t} dj$. We can integrate equation (4.71) across firms to obtain an equation for the average reset price:

$$\begin{aligned}\ln P^*_t &= (1-\beta\theta)\cdot\widehat{mc}^{(0)}_{t|t} + (1-\beta\theta)\ln P^{(1)}_{t|t} \\ &+ \beta\theta\ln P^{*(1)}_{t+1|t} - \beta\theta\ln\pi_*\end{aligned} \tag{4.72}$$

where we use the claim of the proposition above. Keep in mind that the price index equation can be manipulated to get equation (4.63)

$$\ln P_t = \theta\left(\ln P_{t-1} + \ln \pi_*\right) + (1-\theta)\ln P_t^* \tag{4.73}$$

Let us plug the equation (4.72) into the equation (4.73):

$$\begin{aligned} \ln P_t = {} & \theta \ln P_{t-1} + \left(\theta - (1-\theta)\beta\theta\right)\ln \pi_* \\ & + (1-\theta)\left[(1-\beta\theta)\cdot \widehat{mc}_{t|t}^{(0)} + (1-\beta\theta)\ln P_{t|t}^{(1)} + \beta\theta \ln P_{t+1|t}^{*(1)}\right] \end{aligned} \tag{4.74}$$

It is easy to show that for any $k \in \{1, 2, \dots\}$:

$$\begin{aligned} \ln P_{t|t}^{(k)} = {} & \theta \ln P_{t-1} + \left(\theta - (1-\theta)\beta\theta\right)\ln \pi_* \\ & + (1-\theta)\left[(1-\beta\theta)\cdot \widehat{mc}_{t|t}^{(k+1)} + (1-\beta\theta)\ln P_{t|t}^{(k+1)} + \beta\theta \ln P_{t+1|t}^{*(k+1)}\right] \end{aligned} \tag{4.75}$$

Substituting the equation (4.75) associated with $k = 1$ into the equation (4.74):

$$\begin{aligned} \ln P_t = {} & \theta \ln P_{t-1} + \left(\theta - (1-\theta)\beta\theta\right)\ln \pi_* \\ & + (1-\theta)\left[(1-\beta\theta)\cdot \widehat{mc}_{t|t}^{(1)} + \beta\theta \ln P_{t+1|t}^{*(1)}\right] \\ & + (1-\theta)(1-\beta\theta)\left[\theta \ln P_{t-1} + \left(\theta - (1-\theta)\beta\theta\right)\ln \pi_*\right] \\ & + (1-\theta)(1-\beta\theta)\left[(1-\theta)\left[(1-\beta\theta)\cdot \widehat{mc}_{t|t}^{(2)} + \beta\theta \ln P_{t+1|t}^{*(2)}\right]\right] \\ & + (1-\theta)^2(1-\beta\theta)^2 \ln P_{t|t}^{(2)} \end{aligned} \tag{4.76}$$

Plugging the equation (4.75) associated with $k = 2$ into the equation (4.76):

$$\begin{aligned} \ln P_t = {} & \theta \ln P_{t-1} + \left(\theta - (1-\theta)\beta\theta\right)\ln \pi_* \\ & + (1-\theta)\left[(1-\beta\theta)\cdot \widehat{mc}_{t|t}^{(0)} + \beta\theta \ln P_{t+1|t}^{*(1)}\right] \\ & + (1-\theta)(1-\beta\theta)\left[\theta \ln P_{t-1} + \left(\theta - (1-\theta)\beta\theta\right)\ln \pi_*\right] \\ & + (1-\theta)^2(1-\beta\theta)\left[(1-\beta\theta)\cdot \widehat{mc}_{t|t}^{(1)} + \beta\theta \ln P_{t+1|t}^{*(1)}\right] \\ & + (1-\theta)^2(1-\beta\theta)^2\left[\theta \ln P_{t-1} + \left(\theta - (1-\theta)\beta\theta\right)\ln \pi_*\right] \\ & + (1-\theta)^3(1-\beta\theta)^2\left[(1-\beta\theta)\cdot \widehat{mc}_{t}^{(2)} + \beta\theta \ln P_{t+1|t}^{*(2)}\right] \\ & + (1-\theta)^3(1-\beta\theta)^3 \ln P_{t|t}^{(3)} \end{aligned}$$

Keeping on substituting the equation (4.75) associated with larger and larger k up to infinity into the resulting equation yields:

$$
\begin{aligned}
\ln P_t = &\frac{\ln P_{t-1} + (1-(1-\theta)\,\beta)\ln \pi_*}{(1+(1-\theta)\,\beta)} \\
&+ (1-\theta)\,(1-\beta\theta)\sum_{k=0}^{\infty}(1-\theta)^k\,(1-\beta\theta)^k\,\widehat{mc}_{t|t}^{(k)} \\
&+ (1-\theta)\,\beta\theta\sum_{k=0}^{\infty}(1-\theta)^k\,(1-\beta\theta)^k\ln P_{t+1|t}^{*(k+1)}
\end{aligned}
$$

Use the fact that $\ln P_t = \widehat{\pi}_t + \ln P_{t-1} + \ln \pi_*$ and from the price index (4.63):

$$\ln P_t = \theta\,(\ln P_{t-1} + \ln \pi_*) + (1-\theta)\ln P_t^* \tag{4.77}$$

$$\widehat{\pi}_t + \ln P_{t-1} + \ln \pi_* = \theta\,(\ln P_{t-1} + \ln \pi_*) + (1-\theta)\ln P_t^* \tag{4.78}$$

$$\widehat{\pi}_{t+1} = (\theta-1)\,(\ln P_t + \ln \pi_*) + (1-\theta)\ln P_{t+1}^* \tag{4.79}$$

$$\ln P_{t+1}^* = \frac{\hat{\pi}_{t+1}}{1-\theta} + \ln P_t + \ln \pi_*$$

By using these results we obtain:

$$
\begin{aligned}
\widehat{\pi}_t + \ln P_{t-1} + \ln \pi_* = &\frac{\ln P_{t-1} + (1-(1-\theta)\,\beta)\ln \pi_*}{(1+(1-\theta)\,\beta)} \\
&+ (1-\theta)\,(1-\beta\theta)\sum_{k=0}^{\infty}(1-\theta)^k\,(1-\beta\theta)^k\,\widehat{mc}_{t|t}^{(k)} \\
&+ (1-\theta)\,\beta\theta\sum_{k=0}^{\infty}(1-\theta)^k\,(1-\beta\theta)^k\ln P_{t|t}^{(k+1)} \\
&+ \beta\theta\sum_{k=0}^{\infty}(1-\theta)^k\,(1-\beta\theta)^k\,\hat{\pi}_{t+1|t}^{(k+1)} \\
&+ \underbrace{\frac{(1-\theta)\,\beta\theta}{1-(1-\theta)\,(1-\beta\theta)}}_{\frac{(1-\theta)\beta}{1+(1-\theta)\beta}}\ln \pi_*
\end{aligned}
$$

By simplifying and rearranging:

$$\begin{aligned}
\widehat{\pi}_t = & \frac{-\left(1-\theta\right)\beta}{1+\left(1-\theta\right)\beta}\left(\ln P_{t-1}+\ln\pi_*\right) \\
& +\left(1-\theta\right)\left(1-\beta\theta\right)\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\beta\theta\right)^k\widehat{mc}_{t|t}^{(k)} \\
& +\left(1-\theta\right)\beta\theta\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\beta\theta\right)^k\ln P_{t|t}^{(k+1)} \\
& +\beta\theta\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\beta\theta\right)^k\hat{\pi}_{t+1|t}^{(k+1)}
\end{aligned}$$

Let us focus on the first term on the rhs:

$$\begin{aligned}
\frac{-\left(1-\theta\right)\beta\theta}{\theta\left(1+\left(1-\theta\right)\beta\right)}\left(\ln P_{t-1}+\ln\pi_*\right) &= \frac{-\left(1-\theta\right)\beta\theta}{1-\left(1-\theta\right)\left(1-\theta\beta\right)}\left(\ln P_{t-1}+\ln\pi_*\right) \\
&= -\left(1-\theta\right)\beta\theta\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\theta\beta\right)^k\left(\ln P_{t-1}+\ln\pi_*\right)
\end{aligned}$$

By using this result in the equation above we can write:

$$\begin{aligned}
\widehat{\pi}_t = & \left(1-\theta\right)\left(1-\beta\theta\right)\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\beta\theta\right)^k\widehat{mc}_{t|t}^{(k)} \\
& +\left(1-\theta\right)\beta\theta\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\beta\theta\right)^k\left(\underbrace{\ln P_{t|t}^{(k+1)}-\ln P_{t-1}-\ln\pi_*}_{\pi_{t|t}^{(k+1)}}\right) \\
& +\beta\theta\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\beta\theta\right)^k\hat{\pi}_{t+1|t}^{(k+1)}
\end{aligned}$$

and hence:

$$\begin{aligned}
\widehat{\pi}_t+\ln P_{t-1}+\ln\pi_* = & \left(1-\theta\right)\left(1-\beta\theta\right)\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\beta\theta\right)^k\widehat{mc}_{t|t}^{(k)} \\
& +\left(1-\theta\right)\beta\theta\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\beta\theta\right)^k\pi_{t|t}^{(k+1)} \\
& +\beta\theta\sum_{k=0}^{\infty}\left(1-\theta\right)^k\left(1-\beta\theta\right)^k\hat{\pi}_{t+1|t}^{(k+1)}
\end{aligned}$$

We use the assumption of common knowledge in rationality to obtain:

$$\hat{\pi}_t = (1-\beta\theta)(1-\theta)\sum_{k=0}^{\infty}(1-\theta)^k\,\widehat{mc}_t^{(k)} + \beta\theta\sum_{k=0}^{\infty}(1-\theta)^k\,\hat{\pi}_{t+1|t}^{(k+1)}.$$

Appendix E Verifying the Guess

The transition equations of the model can be written as:

$$\hat{y}_t - \hat{\zeta}_t = \mathbb{E}_t\left(\hat{y}_{t+1} - \hat{\zeta}_{t+1}\right) + \frac{1}{\gamma}\mathbb{E}_t\hat{\pi}_t - \frac{1}{\gamma}\hat{R}_t \tag{4.80}$$

$$\begin{aligned}\hat{\pi}_t = {} & (1-\theta)(1-\beta\theta)\sum_{k=0}^{\infty}(1-\theta)^k\left[\gamma\left(\hat{y}_{t|t}^{(k+1)} - \hat{\zeta}_{t|t}^{(k+1)}\right) - z_{t|t}^{(k)} - \sigma_a\eta_{a,t|t}^{(k)}\right] \\ & + \beta\theta\sum_{k=0}^{\infty}(1-\theta)^k\,\hat{\pi}_{t+1|t}^{(k+1)}\end{aligned} \tag{4.81}$$

$$\hat{R}_t = \rho_r\hat{R}_{t-1} + (1-\rho_r)\left[\phi_\pi\left(\hat{\pi}_t - \hat{\pi}_t^*\right) + \phi_y\left(\hat{y}_t - z_t - \sigma_a\eta_{a,t}\right)\right] + \sigma_r\eta_{r,t} \tag{4.82}$$

To solve the model we need the following assumption (see also Nimark, 2007).

Assumption of common knowledge in rationality: It is true that

$$\underbrace{\int\mathbb{E}_{j,t}\ldots\int\mathbb{E}_{j,t}}_{s}\varphi_{t+h|t+h}^{(0:k)} = \mathbf{M}^h\varphi_{t|t}^{(s:k+s)}$$

.

This assumption implies that agents use the actual law of motion of higher-order beliefs to forecast the dynamics of the higher-order beliefs. The following claims turn out to be useful:

CLAIM: $\mathbf{s}_{t|t}^{(s)} = \mathbf{v}_0\boldsymbol{\tau}^{(s)}\mathbf{T}^{(s)}\varphi_{t|t}^{(0:k)} + \mathbf{v}_1\hat{R}_{t-1}$, where

$$\begin{aligned}\boldsymbol{\tau}^{(s)} &= \left[\mathbb{I}_{6(k+1-s)}, \mathbf{0}_{6(k+1-s)\times 6s}\right]' \\ \mathbf{T}^{(s)} &= \left[\mathbf{0}_{6(k+1-s)\times 6s}, \mathbb{I}_{6(k+1-s)}\right]\end{aligned}$$

and for any $0 \leq s \leq k$.

Proof. We conjectured that $\mathbf{s}_t = \mathbf{v}_0\varphi_{t|t}^{(0:k)} + \mathbf{v}_1\hat{R}_{t-1}$.

$$\mathbf{s}_{t|t}^{(s)} = \mathbf{v}_0\varphi_{t|t}^{(s:k+s)} + \mathbf{v}_1\hat{R}_{t-1}$$

Since we truncate beliefs after the k-th order we have that

$$\mathbf{s}_{t|t}^{(s)} = \mathbf{v}_0 \boldsymbol{\tau}^{(s)} \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \hat{R}_{t-1}, \text{for any } 0 \leq s \leq k$$

■

CLAIM: $\mathbf{s}_{t+1|t}^{(s)} = \mathbf{v}_0 \mathbf{M} \boldsymbol{\tau}^{(\mathbf{s})} \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k+s)} + \mathbf{v}_1 \mathbf{1}_3^T \mathbf{s}_t$, where

$$\begin{aligned} \boldsymbol{\tau}^{(s)} &= \left[\mathbb{I}_{6(k+1-s)}, \mathbf{0}_{6(k+1-s)\times 6s}\right]' \\ \mathbf{T}^{(s)} &= \left[0_{6(k+1)\times 6s}, \mathbb{I}_{6(k+1)}\right] \end{aligned}$$

and for any $0 \leq s \leq k$.

Proof. We conjectured that $\mathbf{s}_{t+1} = \mathbf{v}_0 \varphi_{t+1|t+1}^{(0:k)} + \mathbf{v}_1 \hat{R}_t$. Common knowledge in rationality implies that

$$\mathbf{s}_{t+1|t}^{(s)} = \mathbf{v}_0 \mathbf{M} \varphi_{t|t}^{(s:k+s)} + \mathbf{v}_1 \hat{R}_{t|t}^{(s)}$$

Since $\hat{R}_t \in \mathcal{I}_{j,t}$, all j, then we can write:

$$\mathbf{s}_{t+1|t}^{(s)} = \mathbf{v}_0 \mathbf{M} \varphi_{t|t}^{(s:k+s)} + \mathbf{v}_1 \mathbf{1}_3^T \left(\mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \hat{R}_{t-1}\right)$$

Since we truncate beliefs after the k-th order we have that

$$\begin{aligned} \mathbf{s}_{t+1|t}^{(s)} &= \mathbf{v}_0 \mathbf{M} \boldsymbol{\tau}^{(\mathbf{s})} \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{1}_3^T \left(\mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \hat{R}_{t-1}\right) \\ &= \mathbf{v}_0 \mathbf{M} \boldsymbol{\tau}^{(\mathbf{s})} \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{1}_3^T \mathbf{s}_t \end{aligned}$$

■

The Euler equation (4.80) implies:

$$\begin{aligned} \mathbf{b}_0 \varphi_{t|t}^{(0:k)} + b_1 \hat{R}_{t-1} - \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right) \varphi_{t|t}^{(0:k)} = {} & \mathbf{b}_0 \mathbf{M} \varphi_{t|t}^{(0:\infty)} + b_1 \left(\mathbf{c}_0 \varphi_{t|t}^{(0:k)} + c_1 \hat{R}_{t-1}\right) \\ & - \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right) \mathbf{M} \varphi_{t|t}^{(0:\infty)} \\ & + \frac{1}{\gamma} \left[\mathbf{a}_0 \mathbf{M} \varphi_{t|t}^{(0:\infty)} + a_1 \left(\mathbf{c}_0 \varphi_{t|t}^{(0:k)} + c_1 \hat{R}_{t-1}\right)\right] \\ & - \frac{1}{\gamma} \left(\mathbf{c}_0 \varphi_{t|t}^{(0:k)} + c_1 \hat{R}_{t-1}\right) \end{aligned}$$

and hence,

$$\begin{aligned} & \mathbf{b}_0 = \mathbf{1}_5^T + \mathbf{1}_6^T + \mathbf{b}_0 \mathbf{M} + b_1 \mathbf{c}_0 - \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right) \mathbf{M} + \frac{1}{\gamma} \left(\mathbf{a}_0 \mathbf{M} + a_1 \mathbf{c}_0\right) - \frac{1}{\gamma} \mathbf{c}_0 \\ & b_1 = b_1 c_1 + \frac{1}{\gamma} a_1 c_1 - \frac{1}{\gamma} c_1 \end{aligned}$$

The Phillips curve (4.81) can be rewritten as:

$$\begin{aligned}
&\mathbf{a}_0\varphi_{t|t}^{(0:k)}+\mathbf{a}_1\hat{R}_{t-1}= \\
&\quad=(1-\theta)(1-\beta\theta)\gamma\sum_{s=0}^{k-1}(1-\theta)^s\,\mathbf{1}_2^T\left[\mathbf{v}_0\boldsymbol{\tau}^{(s+1)}\mathbf{T}^{(s+1)}\varphi_{t|t}^{(0:k)}+\mathbf{v}_1\hat{R}_{t-1}\right]+ \\
&\quad+(1-\theta)(1-\beta\theta)\sum_{s=0}^{k-1}(1-\theta)^s\left[-\left(\gamma\boldsymbol{\gamma}_g^{(s+1)\prime}+\boldsymbol{\gamma}_a^{(s)\prime}\right)\varphi_{t|t}^{(0:k)}\right]+ \\
&\quad+\beta\theta\sum_{s=0}^{k-1}(1-\theta)^s\,\mathbf{1}_1^T\left[\mathbf{v}_0\mathbf{M}\boldsymbol{\tau}^{(s+1)}\mathbf{T}^{(s+1)}\varphi_{t|t}^{(0:k)}+\mathbf{v}_1\mathbf{1}_3^T\left(\mathbf{v}_0\varphi_{t|t}^{(0:k)}+\mathbf{v}_1\hat{R}_{t-1}\right)\right]
\end{aligned}$$

where

$$\boldsymbol{\gamma}_g^{(s)}=\left[\mathbf{0}_{1\times 6s},(0,0,0,0,1,1),\mathbf{0}_{1\times 6(k-s)}\right]$$

and

$$\boldsymbol{\gamma}_a^{(s)}=\left[\mathbf{0}_{1\times 6s},(1,1,0,0,0,0),\mathbf{0}_{1\times 6(k-s)}\right]'.$$

Therefore we obtain:

$$\begin{aligned}
\mathbf{a}_0&=(1-\theta)(1-\beta\theta)\cdot\left[\boldsymbol{\nu}\mathbf{m}_1-\left(\sum_{s=0}^{k-1}(1-\theta)^s\left(\gamma\boldsymbol{\gamma}_g^{(s+1)\prime}+\boldsymbol{\gamma}_a^{(s)\prime}\right)\right)\right] \\
&\quad+\beta\theta\boldsymbol{\nu}\mathbf{m}_2+\beta\theta\left(\sum_{s=0}^{k-1}(1-\theta)^s\right)\mathbf{1}_1^T\mathbf{v}_1\mathbf{1}_3^T\mathbf{v}_0 \\
a_1&=(1-\theta)(1-\beta\theta)\left(\sum_{s=0}^{k-1}(1-\theta)^s\right)\gamma\mathbf{1}_2^T\mathbf{v}_1+\beta\theta\left(\sum_{s=0}^{k-1}(1-\theta)^s\right)\mathbf{1}_1^T\mathbf{v}_1\mathbf{1}_3^T\mathbf{v}_1
\end{aligned}$$

where

$$\mathbf{m}_1\equiv\begin{bmatrix}\left[\gamma\mathbf{1}_2^T\mathbf{v}_0\boldsymbol{\tau}^{(1)}\mathbf{T}^{(1)}\right]\\(1-\theta)\left[\gamma\mathbf{1}_2^T\mathbf{v}_0\boldsymbol{\tau}^{(2)}\mathbf{T}^{(2)}\right]\\(1-\theta)^2\left[\gamma\mathbf{1}_2^T\mathbf{v}_0\boldsymbol{\tau}^{(3)}\mathbf{T}^{(3)}\right]\\\vdots\\(1-\theta)^{k-1}\left[\gamma\mathbf{1}_2^T\mathbf{v}_0\boldsymbol{\tau}^{(k)}\mathbf{T}^{(k)}\right]\end{bmatrix},$$

$$\mathbf{m}_2\equiv\begin{bmatrix}\left[\mathbf{1}_1^T\mathbf{v}_0\mathbf{M}\boldsymbol{\tau}^{(\mathbf{1})}\mathbf{T}^{(1)}\right]\\(1-\theta)\left[\mathbf{1}_1^T\mathbf{v}_0\mathbf{M}\boldsymbol{\tau}^{(\mathbf{2})}\mathbf{T}^{(2)}\right]\\(1-\theta)^2\left[\mathbf{1}_1^T\mathbf{v}_0\mathbf{M}\boldsymbol{\tau}^{(\mathbf{3})}\mathbf{T}^{(3)}\right]\\\vdots\\(1-\theta)^{k-1}\left[\mathbf{1}_1^T\mathbf{v}_0\mathbf{M}\boldsymbol{\tau}^{(\mathbf{k})}\mathbf{T}^{(k)}\right]\end{bmatrix},$$

$$\boldsymbol{\nu}=\mathbf{1}_{1\times k}$$

Finally the Taylor rule (4.82):

$$
\begin{aligned}
\mathbf{c}_0\varphi_{t|t}^{(0:k)} + c_1\hat{R}_{t-1} = {} & \rho_r\hat{R}_{t-1} + \\
& + (1-\rho_r)\,\phi_\pi\left(\mathbf{a}_0\varphi_{t|t}^{(0:k)} + a_1\hat{R}_{t-1} - \mathbf{1}_3^T\varphi_{t|t}^{(0:k)}\right) \\
& + (1-\rho_r)\,\phi_y\left(\mathbf{b}_0\varphi_{t|t}^{(0:k)} + b_1\hat{R}_{t-1} - \left(\mathbf{1}_1^T + \mathbf{1}_2^T\right)\varphi_{t|t}^{(0:k)}\right) \\
& + \mathbf{1}_4^T\varphi_{t|t}^{(0:k)}
\end{aligned}
$$

and then

$$
\begin{aligned}
\mathbf{c}_0 &= (1-\rho_r)\,\phi_\pi\left(\mathbf{a}_0 - \mathbf{1}_3^T\right) \\
&\quad + (1-\rho_r)\,\phi_y\left(\mathbf{b}_0 - \left(\mathbf{1}_1^T + \mathbf{1}_2^T\right)\right) + \mathbf{1}_4^T \\
c_1 &= \rho_r + (1-\rho_r)\left[\phi_\pi a_1 + \phi_y b_1\right]
\end{aligned}
$$

Appendix F The Law of Motion of the Average Beliefs

This section shows how to derive the law of motion of the average expectation. This step is required in order to solve the model. Firms' reduced-form state-space model is

$$\mathbf{X}_t = \mathbf{W}\cdot\mathbf{X}_{t-1} + \mathbf{U}\cdot\boldsymbol{\varepsilon}_t \tag{4.83}$$

The average expectations of $\mathbf{X}_t$ follows

$$\mathbf{X}_{t|t}^{(1)} = \left(\mathbb{I} - \mathbf{KD}\right)\mathbf{W}\mathbf{X}_{t-1|t-1}^{(1)} + \mathbf{KDW}\mathbf{X}_{t-1} + \mathbf{KDU}\boldsymbol{\varepsilon}_t$$

where the steady-state matrix of Kalman gains is well-known to be equal to

$$\mathbf{K} = \mathbf{PD}'\left[\mathbf{DPD}' + \boldsymbol{\Sigma}_e\right]^{-1} \tag{4.84}$$

where $\boldsymbol{\Sigma}_e = \mathbf{QQ}'$. The variance and covariance matrix $\mathbf{P}$ solves the following algebraic Riccati equation:

$$\mathbf{P} = \mathbf{W}\left[\mathbf{P} - \mathbf{PD}'\left[\mathbf{DPD}' + \boldsymbol{\Sigma}_e\right]^{-1}\mathbf{DP}\right]\mathbf{W}' + \mathbf{UU}' \tag{4.85}$$

Note that $\varphi_{t|t}^{(0:\infty)} = \left[\varphi_t, \varphi_{t|t}^{(1:\infty)}\right]'$ and that:

$$\varphi_t = \underbrace{\begin{bmatrix} \rho_z & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & \rho_\pi & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & \rho_g & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}}_{\mathbf{R}_1} \varphi_{t-1|t-1}^{(0:k)} + \underbrace{\begin{bmatrix} \sigma_z & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_a & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\pi & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_r & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{\sigma}_g & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_g \end{bmatrix}}_{\mathbf{R}_2} \cdot \boldsymbol{\varepsilon}_t$$

So we can fully characterize the matrices $\mathbf{M}$ and $\mathbf{N}$:

$$\mathbf{M} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{6\times 6} & \mathbf{0}_{6\times 6k} \\ \mathbf{0}_{6k\times 6} & (\mathbb{I} - \mathbf{KD})\,\mathbf{W}|_{(1:6k,1:6k)} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{KDW}|_{(1:6k,1:6(k+1))} \end{bmatrix}$$

$$\mathbf{N} = \begin{bmatrix} \mathbf{R}_2 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{KDU}|_{(1:6k,1:6)} \end{bmatrix}$$