

Public's Inflation Expectations and Monetary Policy

Leonardo Melosi[†]

London Business School

4 October, 2011

Abstract

This paper studies how public's expectations react to monetary policy decisions. To this end, we introduce a DSGE model in which price-setting firms have to form expectations about the evolution of their nominal marginal costs on the basis of four signals: (1) their idiosyncratic productivity, which is correlated with a persistent aggregate technology shock; (2) last period's output; (3) last period's inflation; (4) the interest rate set by the central bank according to a Taylor rule. Since firms have private information, their expectations are heterogenous. Furthermore, the model features a channel of monetary transmission which is based on affecting firms' expectations with the interest rate. The model is estimated with likelihood methods on a U.S. data set including the Survey of Professional Forecasters as a measure of price setters' expectations. The paper finds that inflation expectations respond positively to monetary shocks because firms interpret the rise of the policy rate as a response of the central bank to a positive demand shock. While the central bank is found to be quite successful in coordinating inflation expectations by maneuvering the policy rate, monetary policy has no effect on the dispersion of expectations about output. Finally, the paper argues that the accommodative approach of the Federal Reserve in the 1970s provides an explanation for why recent VAR studies find that the price puzzle is substantially weaker after that decade.

Keywords: Coordinating public's expectations, price puzzle, heterogenous expectations, higher-order beliefs, Bayesian econometrics, persistent real effects of nominal shocks.

JEL classification: E5, C11, D8.

[†]Correspondence to: Leonardo Melosi, London Business School, Regent's Park, London, NW1 4S, United Kingdom. Email: lmelosi@london.edu. I wish to thank Francesco Bianchi, Harald Cole, Jesus Fernandez-Villaverde, Dirk Krueger, Kristoffer Nimark, Luigi Paciello, Ricardo Reis, Lucrezia Reichlin, Helene Rey, Frank Schorfheide, and Paolo Surico for their very helpful comments and discussion. I thank the participants of the ECB-Bundesbank-University of Frankfurt joint Seminar, London Business School's Workshop on "The Macroeconomics of Incomplete Information: Empirical and Policy Perspectives", the CREI-CEPR Workshop on "Information, Beliefs and Expectations in Macroeconomics" and Bank of England's Macroeconomic seminar, for their helpful comments.

1 Introduction

That monetary policy influences output and inflation by affecting public' expectations has come to a growing consensus among scholars and policy makers¹ in the last twenty years. Woodford (2005) argues that "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. [...] For not only do expectations about policy matter [for effective monetary policy], but, at least under current conditions, very little else matters."

As pointed out by Morris and Shin (2008), market participants take decisions based on variables that are out of direct control of the central bank but they look at central bank's actions for clues about where those variables are headed. In this view, public's expectations are dispersed and can be coordinated by publicly observed action, such as the interest rate set by the central bank. The literature has produced a few theoretical models in which the central bank can coordinate dispersed agents' expectations in self-fulfilling environments (e.g., Morris and Shin, 2003, Hellwig, 2002, Angeletos, Hellwig, and Pavan, 2006, Angeletos and Pavan, 2008, and Hellwig and Veldkamp, 2009). On the one hand, these models are too stylized to conduct a formal econometric investigation about the empirical significance of this new propagation channel (henceforth, *expectation channel*) for monetary policy. On the other hand, the workhorse models (e.g., Christiano, Eichenbaum, and Evans 2005 and Smets and Wouters, 2007) for conducting empirical investigations on how monetary shocks propagate do not feature any explicit role for agent's expectations. A standard assumption in these models is, in fact, that agents have complete information, that is, they observe the complete history of the shocks that have hit the economy. As a consequence, agents share the same expectations about the model variables (e.g., GDP, inflation, interest rate, etc.). Furthermore, agents' expectations correctly forecast the evolution of model variables in the aftermath of an unanticipated monetary policy shock. Therefore, in these models there is no scope for the central bank to manage and coordinate agents' expectations by setting the interest rate or making announcements.

The paper bridges this gap in the literature and empirically investigates to what extent the central bank can affect and coordinate heterogenous agents' expectations. To this end, I introduce a Dynamic Stochastic General Equilibrium (DSGE) model in which price setters have dispersed information about the state variables of the model. More precisely, price-setting firms observe their idiosyncratic productivity (i.e., a *private* signal), last period's real output and inflation as well as the current interest rate, which is set by the central bank

¹See, for instance, Bernanke (2004) and King (2005), and Trichet (2006).

according to a Taylor rule (i.e., a contemporaneous *public* signal). Price setters face costs of price adjustment in the form of Calvo sticky prices (Calvo, 1983). As a result, those price setters who are allowed to optimize their prices need to forecast the evolution of their nominal marginal costs when taking their price-setting decisions. In such an environment, firms face higher-order uncertainty as the optimal price will depend not only on their beliefs about the future price levels, but also on their higher-order beliefs about the price level, that is, their beliefs about other price setters' beliefs about the future price levels, etc. (Townsend 1983a and 1983b). The model is estimated through likelihood methods on a U.S. data set that includes the *Survey of Professional Forecasters* as a measure of price setters' inflation expectations. The data range includes the 1970s, which were characterized by one of the most notorious episodes of inflation and inflation expectations run-up in the recent US economic history.

The paper finds that inflation expectations respond positively to monetary shocks because firms interpret the rise of the policy rate as a response of the central bank to a positive demand shock. Furthermore, the effect of the policy rate set by the Federal Reserve does not always coordinate expectations and, in fact, fosters firms' disagreement. More specifically, the policy rate turns out to reduce firms' disagreement about inflation but has no effect on firms disagreement about output. The explanation for the latter result, it is that while monetary policy is found to coordinate expectations about the level of aggregate technology and preferences, it boosts the dispersion of firms' expectations about monetary shocks, which apparently strongly affect the law of motion of output.

The transmission channel for monetary policy based on affecting price setters' expectations is found to account for three important facts that the VAR literature shows to characterize the dynamic effects of monetary shocks: (*i*) the inflation persistence puzzle (i.e., a slow and delayed fall in inflation in response to a contractionary monetary policy shock)², (*ii*) the price puzzle (i.e., a temporary rise in the price level after a contractionary monetary policy shock, Sims, 1992 and Eichenbaum, 1992)³, (*iii*) the disappearance of the price puzzle: Barth and Ramey (2001), Hansen (2004), Castelnuovo and Surico (2010). and Ravn, Schmitt-Grohe, Uribe, and Uusküla (2010) argue that the price puzzle has become much weaker (or even statistically insignificant) after the 1970s because the U.S. monetary policy has become much more aggressive against inflation than in the past.

To see how the model unravels the three puzzles (*i*)-(iii), it is important to notice that the monetary policy rule (i.e., the Taylor rule) plays a twofold role in the model. First, it is one of the transition equations for the model endogenous state variables (i.e., inflation,

²See Christiano, Eichenbaum, and Evans (1999) and Stock and Watson (2001).

³See Sims (1992) and Eichenbaum (1992).

output, interest rate). Second, this reaction function is the stochastic process driving the public signal of the policy rate. This second role introduces the new transmission channel for monetary policy based on affecting price setters' expectations. The central bank follows a monetary policy rule with feedbacks, that is, it reacts to endogenous variables, namely current inflation and real output growth. When most of the variability of the interest rate stems from these endogenous variables rather than from exogenous variables (i.e., monetary shocks), then changes in the policy rate are more informative about inflation and output than about monetary shocks. In this case, a rising policy rate will be interpreted by price setters as evidence of growing inflation and rising real marginal costs. Consequently, a monetary shock will lead price setters to mistakenly believe, at first, that the central bank is reacting to non-monetary shocks (i.e., technology or preference shocks). The paper will show that this outcome arises when the central bank does not react forcefully enough to inflation. Stronger central bank's reactions to inflation reduce the response of output and inflation to preference and technology shocks and, in turn, cause the policy rate to be less informative about output and inflation. As a result, fewer price setters will confuse a monetary shocks with a technology shock or a preference shock.

The estimated model predicts that a rise in the policy rate due to an unanticipated monetary shock is, at first, interpreted by price setters as a response of the central bank to a positive preference shock. Hence, a monetary tightening raises price setters' expectations about future real marginal costs, which in turn will drag current inflation expectations and hence inflation upwards in the short run. The channel based on price setters' expectations, thus, makes the impact of monetary shocks upon inflation more delayed and persistent. This feature allows the model to capture fact (i). The short-run effects of a monetary tightening upon price setters' expectations are so strong that the model can explain the price puzzle, that is, fact (ii). The model provides a very natural explanation for Fact (iii). When the central bank reacts more aggressively to inflation, firms may be still uncertain about whether the interest rate has increased because of a monetary or a preference shock. Nevertheless, under such a more active monetary regime, firms know that the inflationary effects of a preference or technology shock will be second-order. Thus the model suggests that a switch to a more active monetary policy tends to flip the sign of the response of inflation (and inflation expectations) upon monetary shocks and, hence, may well explain the disappearance of the price puzzle observed after the 1970s.

Finally, the paper performs an econometric evaluation of the channel of monetary transmission based on price-setters' expectations. The paper finds strong empirical support in favor of the channel of monetary transmission based on price-setters' expectations. The presence of the channel helps the model fitting the *Survey of Professional Forecasters*, which

are used as a measure of price setters' expectations.

The paper is organized as follows. After a brief overview of the literature in Section 2, the paper describes the model with heterogenous expectations in Section 3. In that Section, we also describe a model in which the interest rate set by the central bank is not observed and therefore the expectation channel is not active. We introduce this model for studying the empirical relevance and the effects of the expectation channel of monetary transmission. The empirical analysis of the paper is dealt with in Section 4. Section 5 concludes.

2 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.

The paper is related to Del Negro and Eusepi (2010), who perform an econometric evaluation of the extent to which the inflation expectations generated by a prototypical DSGE model are in line with the observed inflation expectations. The main differences with this paper are as follows. First, unlike Del Negro and Eusepi (2010), in our model price setters have heterogeneous and dispersed higher-order expectations as they observe private signals (i.e., firm-specific productivity). Second, the learning mechanism is very different from that in Del Negro and Eusepi (2010), in which agents have to learn the time-varying inflation target from the behavior of the policy rates. In the present paper, price setters have to learn about the history of every shock that have hit the economy. Third, the data range used in this paper differs from the one in Del Negro and Eusepi (2010), who use a data set starting from the early 1980s. In contrast, this paper fits the model to a data set that includes observations for the 1970s.

The model studied in this paper is similar to the model in Nimark (2008), in which 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. The model studied in this paper shares this feature. Nonetheless, Nimark (2008) does not

consider the role of monetary policy in coordinating agents' expectations in that the policy rate conveys redundant information to agents.

The paper is also related to Bianchi (2010) who study how agents' beliefs react to shifts in monetary policy regime and the associated implications for the transmission mechanism of monetary policy. The present paper does not focus on agents' beliefs about policy regime changes. Rather this paper addresses the issue of how public beliefs adjust to monetary policy signals and how such adjustments influence the propagation of monetary impulses. Bianchi and Melosi (2010) develop a DSGE model in which agents have to learn the persistence of the realized monetary regimes. They use this model to study how public expectations and uncertainty reacts to monetary policy decisions and communication.

Lorenzoni (2010) studies optimal monetary policy in a price-setting model where aggregate fluctuations are driven by the private sector's uncertainty about the economy's fundamentals. The main difference from my paper is the imperfect observability of the policy rate. The key mechanism of my paper is based on the fact that the policy rate is a public signal that conveys non-redundant information to price setters. In Lorenzoni's paper, this mechanism is absent and monetary policy rules matter only because they affect agents' incentives to respond to private and public signals.

The paper is also related to the literature that uses incomplete information models for studying the persistence in economic fluctuations (Townsend, 1983a, 1983b; Hellwig, 2002; Adam, 2009; Angeletos and La'O, 2009; Rondina, 2008; and Lorenzoni, 2009) and the propagation of monetary disturbances to real variables and prices (Phelps, 1970; Lucas, 1972; Woodford, 2002; Adam, 2007; Gorodnichenko, 2008; Nimark, 2008; and Lorenzoni, 2010).⁴

3 Models

Section 3.1 introduces the model of interest that features both high-order and heterogeneous expectations as well as the expectation channel of monetary transmission (i.e., the policy rate is a public signal conveying non-redundant information to price setters). In section 3.2, I present the time protocol of the model. Section 3.3 presents the problem of the producers. Section 3.4 presents the problem of households. Section 3.5 presents the price-setting problem of the intermediate-good firms, which have incomplete information. In Section 3.6, central bank's behavior is modeled. Section 3.7 deals with the log-linearization of the model equations. Section 3.8 discusses firm's signal-extraction problem in this log-lineariz(ed) set-up. Section 3.9 sheds light on how the transmission channel based on price

⁴See Mankiw and Reis (2002a, 2002b, 2006, 2007), and Reis (2006a, 2006b, 2009) for models with information frictions that do not feature imperfect common knowledge but can generate sizeable persistence.

setters' expectation works. Section 3.10 introduces a model in which firms do not observe the current interest rate and, hence, the expectation channel is shut down. This model solely deviates from the model of interest in Section 3.1 in that it does not feature the expectation channel of monetary transmission. Introducing this additional model turns out to be useful for evaluating the empirical significance of the expectation channel of monetary transmission.

3.1 The Model with the Expectation Channel (ICKM-EC Model)

The economy is populated by perfectly competitive final-good producers (or, more briefly, the producers), a continuum $(0, 1)$ of intermediate-good firms (or, more briefly, the firms), a continuum $(0, 1)$ of households, a central bank (or central bank), and a government. A Calvo lottery establishes which firms are allowed to re-optimize their prices. The outcome of the Calvo lotteries is assumed to be common knowledge among agents. The central bank supplies money to households so as to control the interest rate at which government's bonds pay out their return. Final goods producers buy intermediate goods from the firms, pack them into a final good to be sold to households and government in a perfectly competitive market. Households consume the final goods, demand money holdings from the central bank and bonds from the government, pay taxes to or receive transfers from the government, and supply labor to the firms.

There are aggregate and idiosyncratic shocks that hit the model economy. The aggregate shocks are: a technology shock, a monetary-policy shock, and a preference shock. All these shocks are orthogonal to each others at all leads and lags. Idiosyncratic shocks include a firm-specific technology, $A_{j,t}$, that determines the level of technology for firm j at time t , and the outcome of the Calvo lottery for price-optimization. The idiosyncratic technology shocks are correlated with the aggregate technology shocks. Since -as we shall show- this model features both imperfect common knowledge (i.e., higher-order uncertainty) and the expectation channel of monetary transmission, we call this model as *ICKM-EC model*.

3.2 The Time Protocol

Any period t is divided into three stages. All actions that are taken in any given stage are simultaneous. At stage 0 $(t, 0)$, shocks realize and the central bank observes the realization of the aggregate shocks and sets the interest rate. At stage 1 $(t, 1)$, firms observe their idiosyncratic productivity, the outcome of the Calvo lottery, and the interest rate set by the central bank and then set their price. At stage 2 $(t, 2)$, households learn the realization of all the shocks in the economy and decide their consumption, $C_{i,t}$, money holdings, $M_{i,t}$, demand for government bonds, $B_{i,t}$ and labor supply, $N_{i,t}$. At this stage, firms learn the

current level of inflation and real output, hire labor, and produce intermediate goods, $Y_{j,t}$ so as to deliver the demanded quantity of their good at the price they have set at the stage 1. Then, intermediate-goods market, final-goods market, money market, bond market, and labor market clear.

3.3 Final-Goods Producers

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

$$Y_t = \left(\int_0^1 (Y_{j,t})^{\frac{\nu-1}{\nu}} dj \right)^{\frac{\nu}{\nu-1}} \quad (1)$$

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, $P_{j,t}$, and output price, P_t , as given. Profit maximization implies that the demand for intermediate goods is:

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

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

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

3.4 Households

Households have perfect information and hence we can use the representative households to solve their problem:

$$\begin{aligned} & \max_{C_{t+s}, B_{t+s}, M_{t+s}, N_{t+s}} \mathbb{E}_t \sum_{s=0}^{\infty} \beta^{t+s} g_{t+s} \left[\ln C_{t+s} + \frac{\chi_m}{1-\gamma_m} \left(\frac{M_{t+s}}{P_{t+s}} \right)^{1-\gamma_m} - \chi_n N_{t+s} \right] \\ & st \\ & P_t C_t + B_t + M_t = W_t N_t + R_{t-1} B_{t-1} + M_{t-1} + \Pi_t + T_t \end{aligned}$$

where β is the deterministic discount factor, g_t denotes a preference shock (or demand shock) that scales up or down the overall period utility, W_t is the (competitive) nominal wage, R_t stands for the interest rate, Π_t are the (equally shared) dividends paid out by the firms, and T_t stands for government transfers. The preference shocks follows an AR process: $\ln g_t =$

$\rho_g \ln g_{t-1} + \sigma_g \varepsilon_{g,t}$ where $\varepsilon_{g,t} \sim \mathcal{N}(0, 1)$.

3.5 Intermediate-Goods Firms' Price-Setting Problem

At the stage 1, intermediate goods firms observe the realization of their idiosyncratic technology shock and the outcome of the Calvo lottery, and set their prices. Firms commit themselves to satisfy any demanded quantity of their intermediate good that will arise at stage 2 at the price they have set at stage 1.

Consider an arbitrary firm j . The real marginal costs for firm j are given by:

$$mc_{j,t} = \frac{W_t}{A_{j,t}P_t} \quad (4)$$

where $A_{j,t}$ is the technology shock that can be decomposed into (1) a trend component γ , (2) a persistent aggregate component, a_t , and (3) a white-noise idiosyncratic component, $\eta_{j,t}^a$. More specifically, we have:

$$A_{j,t} = A_t e^{\sigma_a^j \eta_{j,t}^a} \quad (5)$$

with $\gamma > 1$, and $\eta_{j,t}^a \stackrel{iid}{\sim} \mathcal{N}(0, 1)$, and $A_t = \gamma^t e^{a_t}$ where $a_t = \rho_a a_{t-1} + \sigma_a \varepsilon_{a,t}$ with $\varepsilon_{a,t} \stackrel{iid}{\sim} \mathcal{N}(0, 1)$.

Firms face a Calvo lottery with probability θ of (sub-optimally) adjusting their price to the steady-state (gross) inflation rate, π_* . They set the price of the differentiated good they produce, given the linear technology:

$$Y_{j,t} = A_{j,t} N_{j,t} \quad (6)$$

where $N_{j,t}$ is the amount of labor employed by the firm j at time t .

Let me denote the firm j 's nominal marginal costs at time t as $MC_{j,t} = W_t/A_{j,t}$, firm j 's dividends as $\Pi_{j,t-1}$ and the outcome of the Calvo lottery for firm j at time t as $\mathbb{I}_{j,t}^{Calvo}$. Firm's information set at stage 1 of time t (i.e., when prices are set) $\mathcal{I}_{j,t}$ is defined below:

$$\mathcal{I}_{j,t} \equiv \{R_\tau, \pi_{\tau-1}, Y_{\tau-1}, W_{\tau-1}, \boldsymbol{\theta}_{j,\tau}, \Theta : \tau \leq t\} \quad (7)$$

where $\boldsymbol{\theta}_{j,t} \equiv \{A_{j,t}, P_{j,t}^*, Y_{j,t-1}, MC_{j,t-1}, N_{j,t-1}, \Pi_{j,t-1}, \mathbb{I}_{j,t}^{Calvo}\}$. Note that this information set includes all firm-specific variables that have realized at stage 1 of period t . Moreover, I denote, $\Xi_{t|t+s}$ as the time t value of one unit of the final good in period $t+s$ to the representative

household. Those firms that are allowed to re-optimize their price $P_{j,t}^*$ solve:

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

subject to the firm's specific demand in equation (2), to the production function (6).

3.6 The Central Bank

The central bank sets the interest rate R_t following a Taylor rule of type:

$$R_t = R_{t-1}^{\rho_r} R_t^{*(1-\rho_r)} e^{\sigma_r \eta_{r,t}} \quad (8)$$

where $\eta_{r,t} \stackrel{iid}{\sim} \mathcal{N}(0, 1)$ and the desired nominal interest rate R_t^* is:

$$R_t^* = (r_* \pi_*) \left(\frac{\pi_t}{\pi_*} \right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^*} \right)^{\phi_y} \quad (9)$$

where π_t is the (gross) inflation rate $\ln(P_t/P_{t-1})$, π_* stands for the inflation target, which is assumed to be constant over time, and Y_t^* is the potential output, that is, the output level that realizes if prices were perfectly flexible (i.e., $\theta = 0$).

The central bank observes the contemporaneous realizations of aggregate shocks (i.e., $\varepsilon_{a,t}$, $\eta_{r,t}$, and g_t) and sets the interest rate R_t . The central bank cannot simply tell firms the history of shocks since there is an incentive for the central bank to lie to firms to generate surprise inflation with the aim of pushing output growth above the trend γ .⁵ Unexpected inflation raises output because some prices are sticky. This rise in output has benefits because producers have monopoly power and the unexpected inflation reduces the monopoly distortion. Since there is no commitment device that would back up central bank's words, then any central bank's statements about real output, inflation, and shocks are not deemed as credible by price setters.

The government transfers resources to/from and issue bonds to households. Furthermore, they decide their consumption in terms of final goods. Government spending is denoted by G_t . The government budget constraint is given by:

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

⁵The fact that the central bank sets the interest rate before firms set their prices cannot be considered as a viable commitment device to communicate current inflation and output to firms. The reason is that the Taylor rule makes the interest rate depend on output and inflation only up to a constant (i.e., the monetary shock $\eta_{r,t}$) which is not observed by firms.

Since there is no capital accumulation and the Ricardo equivalence holds, the resource constraint implies $Y_t = C_t$.

3.7 Detrending and Log-linearization

Define the following stationary variables:

$$\begin{aligned} y_t &= \frac{Y_t}{\gamma^t}, \quad c_t = \frac{C_t}{\gamma^t}, \quad y_{j,t} = \frac{Y_{j,t}}{\gamma^t} \\ w_t &= \frac{W_t}{\gamma^t P_t}, \quad a_t = \frac{A_t}{\gamma^t}, \quad R_t = \frac{R_t}{R_*}, \quad mc_{j,t} = \frac{MC_{j,t}}{P_t} \\ \xi_{j,t} &= \gamma^t \Xi_{j,t} \end{aligned}$$

First, I solve firms and households' problems that are described in Sections 3.4 and 3.5 and obtain the consumption Euler equation and a price-setting equation. Second I detrend the non-stationary variables in these equations before log-linearize them around their value at the perfect-information-deterministic steady-state equilibrium.

From the linearized price-setting equation, one can obtain an expression that resembles the New Keynesian Phillips curve, which is reported below (detailed derivations are in Appendix A).

$$\hat{\pi}_t = (1 - \theta)(1 - \beta\theta) \sum_{k=1}^{\infty} (1 - \theta)^k \widehat{mc}_{t|t}^{(k)} + \beta\theta \sum_{k=1}^{\infty} (1 - \theta)^k \widehat{\pi}_{t+1|t}^{(k)} \quad (10)$$

where $\widehat{\pi}_{t+1|t}^{(k)}$ denotes the average k -th order expectations about next period's inflation rate, $\widehat{\pi}_{t+1}$, that is $\widehat{\pi}_{t+1|t}^{(k)} \equiv \underbrace{\int \mathbb{E}_{j,t} \dots \int \mathbb{E}_{j,t}}_k \widehat{\pi}_{t+1}$ and $\widehat{mc}_{t|t}^{(k)}$ denotes the average k -th order expectations about the real aggregate marginal costs, \widehat{mc}_t , which are defined as

$$\widehat{mc}_{t|t}^{(k)} = \hat{y}_{t|t}^{(k)} - a_{t|t}^{(k-1)} \quad (11)$$

any integer $\kappa > 1$. It is easy to show that the log-linearized Euler equation is given by

$$\hat{g}_t - \hat{y}_t = \mathbb{E}_t \hat{g}_{t+1} - \mathbb{E}_t \hat{y}_{t+1} - \mathbb{E}_t \hat{\pi}_{t+1} + \hat{R}_t \quad (12)$$

The Taylor rule can be easily linearized:

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1 - \rho_r) [\phi_\pi \hat{\pi}_t + \phi_y (\hat{y}_t - a_t)] + \sigma_r \eta_{r,t} \quad (13)$$

The parameter set of the model is given by the vector

$$\Theta_{ICKM-EC} = [\theta, \phi_\pi, \phi_y, \beta, \rho_r, \rho_a, \rho_g, \sigma_a, \sigma_a^j, \sigma_r, \sigma_g, \gamma]'$$

3.8 Firms' Signal Extraction Problem and Model Solution

The firms need to form beliefs about the current realization and the future dynamics of their nominal marginal costs given the observable signals in their information set $\mathcal{I}_{j,t}$. Characterizing how firms form such beliefs requires solving a signal extraction problem. It is important to notice that firms need to form expectations on the price level to estimate nominal costs. Hence, they also need to form expectations about what other price-setting firms expect about nominal marginal costs and on what other firms expect that other firms expect and so on (i.e., the so-termed higher-order expectations).

The model can be solved along the lines proposed by Nimark (2008). To solve the model we focus on equilibria where the higher-order expectations about the exogenous state variables, $\varphi_{t|t}^{(0:k)} \equiv [a_{t|t}^{(s)}, \eta_{r,t|t}^{(s)}, g_{t|t}^{(s)} : 0 \leq s \leq k]'$ follow a VAR(1) process:⁶ $\varphi_{t|t}^{(0:k)} = \mathbf{M}\varphi_{t-1|t-1}^{(0:k)} + \mathbf{N}\varepsilon_t$, where $\varepsilon_t \equiv [\varepsilon_{a,t} \quad \eta_{r,t} \quad \varepsilon_{g,t}]'$. Note that we truncate the order of the average expectations at k . Henceforth, we assume that $k = 10$. We found that the results of the paper would not change sensibly by keeping track of an additional order of average expectations. The laws of motion of the three endogenous state variables, which are inflation $\hat{\pi}_t$, real output \hat{y}_t , and the (nominal) interest rate \hat{R}_t , are given by the IS equation (12), the Phillips curve (10), and the Taylor Rule (13). One can use these structural equations to pin down the vectors $\mathbf{v}_0 \equiv [\mathbf{a}'_0, \mathbf{b}'_0, \mathbf{c}'_0]'$ and $\mathbf{v}_1 \equiv [\mathbf{a}'_1, \mathbf{b}'_1, \mathbf{c}'_1]'$ in the equations below:

$$\mathbf{s}_t = \mathbf{v}_0\varphi_{t|t}^{(0:k)} + \mathbf{v}_1\mathbf{s}_{t-1} \tag{14}$$

where $\mathbf{s}_t \equiv [\hat{\pi}_t, \hat{y}_t, \hat{R}_t]'$. Given the matrices \mathbf{M} and \mathbf{N} , the structural equations (11)-(13) can be written as $[\mathbf{v}_0, \mathbf{v}_1]' = f(\mathbf{M}, \mathbf{N}, \mathbf{v}_0, \mathbf{v}_1; \Theta)$. This is a system of non-linear equations in the unknown vectors $\mathbf{v}_0, \mathbf{v}_1$.

For given parameters Θ , take the following steps:

- Set $i = 1$ and guess the matrices $\mathbf{M}^{(i)}$ $\mathbf{N}^{(i)}$

⁶We truncate the state vector of the higher-order expectations at $k < \infty$. When we estimate the model, we set $k = 10$. Nimark (2011) shows that for k bounded but sufficiently high, the approximation of the equilibrium dynamics is very accurate. I find that estimating the model with $\kappa = 11$ would lead to almost identical posterior distributions for model parameters.

- Conditional on $\mathbf{M}^{(i)}$ $\mathbf{N}^{(i)}$, solve the system of nonlinear equations:

$$[\mathbf{v}_0, \mathbf{v}_1]' = f \left(\mathbf{M}^{(i)}, \mathbf{N}^{(i)}, \mathbf{v}_0^{(i)}, \mathbf{v}_1^{(i)}; \Theta \right)$$

where $\mathbf{s}_t = \mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1}$ and $\mathbf{s}_t = [\hat{\pi}_t, \hat{y}_t, \hat{R}_t]'$. See Appendix B.2.

- Use the Kalman equation to work out the mapping $(\mathbf{M}^{(i+1)}, \mathbf{N}^{(i+1)}) = g \left(\mathbf{M}^{(i)}, \mathbf{N}^{(i)}, \mathbf{v}_0^{(i)}, \mathbf{v}_1^{(i)} \right)$. See Appendix C.
- Check that $\|\mathbf{M}^{(i)} - \mathbf{M}^{(i+1)}\| < \varepsilon_m$ and $\|\mathbf{N}^{(i)} - \mathbf{N}^{(i+1)}\| < \varepsilon_n$ with $\varepsilon_m > 0$ and $\varepsilon_n > 0$ small.

3.9 The Expectation Channel of Monetary Transmission

That the central bank can affect price setters' expectations by setting its policy rate is a salient feature of the model developed in the paper. Price setters use the policy rate as a signal both to figure out the shocks that have hit the economy in a given period and affect the evolution of their nominal marginal costs, and to infer potential exogenous deviations from the rule (i.e., non-systematic deviations from the rule $\varepsilon_{r,t}$). It is illustrative to re-write the Taylor rule (13) as follows:

$$\tilde{R}_t = \underbrace{\phi_\pi \hat{\pi}_t + \phi_y \hat{y}_t}_{\text{Endogenous component}} + \underbrace{\sigma_r \eta_{r,t} - \phi_y a_t}_{\text{Exogenous component}} \quad (15)$$

where $\tilde{R}_t \equiv \hat{R}_t - \rho_r \hat{R}_{t-1}$ gathers all the variables in the rule that price setters know at time t . The first term on the right-hand side of the rule gathers all the endogenous variables, while the second term is directly affected by exogenous variables.

When the variability endogenous component is larger than the that of the exogenous component, then a change in the policy rate will be interpreted by price setters as a response of the monetary authority to non-monetary shocks (i.e., technology and preference shocks) that have affected inflation and output. On the contrary, when most of the variability of the interest rate is explained by the exogenous component, then price setters will interpret observed changes in the policy rate as stemming from (i) changes in the potential output (i.e., a_t) and (ii) the non-systematic deviations from the rule (i.e., $\varepsilon_{r,t}$). Whether the endogenous component has a larger variance than the exogenous component has a critical impact on price setters' forecasts about their nominal marginal costs and hence on their price-setting decisions. If price setters, for instance, believe that the policy rate has been increased because a strong positive demand shock has affected the economy, they will forecast growing nominal

marginal costs. On the contrary, if they interpret the rise in the policy rate as an exogenous deviations from the rule, then they will expect a rise in the real interest rate and hence a fall in their nominal marginal costs. Therefore, if the endogenous component of the policy signal is more volatile than the exogenous one, a monetary shock has delayed impacts on inflation and sluggish real effects because price setters learns about this shock only very slowly.

The relative volatility of the endogenous component of the policy signal is influenced by the following events: (a) the central bank becomes more or less proactive to either inflation or output (i.e., the values of the parameters ϕ_π or ϕ_y are changed); (b) a change in the variance of the monetary and technology shock, σ_r ; (c) a change in the variance of preference shocks.

While the effect (c) positively affects the relative volatility of the endogenous component of the policy signal, the direction of the effects (a) and (b) is ambiguous. Consider, first the effect (a). On one hand, a larger degree of proactiveness of monetary policy (i.e., ϕ_π and ϕ_y) will increase the importance of the endogenous component by scaling up its variability. On the other hand, very large values for these policy parameters will end up reducing the unconditional variance of real output and inflation. As far as the effect (b), smaller variability of monetary shocks, σ_r , will increase the variability of the endogenous component relative to the exogenous component. Nonetheless, it will also decrease the variability of output and inflation.

It is finally important to notice that a larger value for the coefficient on inflation, ϕ_π relative to that on the output gap ϕ_y changes the information about the nature of the technology shocks that the policy rate conveys to price setters. Relatively large value for the parameter ϕ_π implies that the central bank will raise its policy rate after a negative technology shock as it cares mostly about inflation stabilization. When ϕ_π is relatively large, a rise in policy rate thus tends to coordinate agent expectations towards a negative technology shocks, which is expected to boost average expectations about real marginal costs and hence inflation. When ϕ_π is relatively small, then a rise of the interest rate tends to coordinate average expectations towards a positive technology shock that reduces inflation. In contrast, a rise of the interest rate will always signal that a positive demand shock may have hit the economy. When price setters interpret the rise in the policy rate as coming from a demand shock, then expected real marginal costs and inflation will go up in the short run.

If price setters interpret a change in the policy rate as mainly coming from central bank's response to a demand shock via the endogenous component in equation (15), the effects of a monetary shock on inflation will be delayed. This is because it will take a while for price setters to realize the correct nature of the shock that has changed the policy rate. By the same token, effects on inflation also tend to be amplified by the expectation channel, if price setters interpret a rise (fall) in the interest rate as coming from the response of the central

bank to a negative (positive) technology shock.

Finally, another important decomposition to learn how the expectation channel works is the variance decomposition of the Taylor rule. If a large portion of the variability of the interest rate is explained by, say, the preference shock, then firms will use the policy signal to learn mainly about the shocks to preferences.

3.10 The Model with No Expectation Channel (ICKM)

Since, in the *ICKM-EC* model discussed in previous Sections, the contemporaneous realization of the policy rate enters the information set of firm j at time t (i.e., $\mathcal{I}_{j,t}$), defined in (7), the policy rate is a public signal conveying information to price setters. This gives rise to what we have called the expectation channel of monetary transmission. To study the effects of this channel, we introduce a model which is exactly as the one presented in the previous sections but the current policy rate R_t does not enter price setters' information set and hence no expectation channel arises. In symbols, firms' information set in the model with no expectation channel is defined as $\tilde{\mathcal{I}}_{j,t} = \mathcal{I}_{j,t}/R_t$.⁷ This model, however, retains the feature of information being dispersed and that of the coordination motive due to the forward-looking behaviors of price setters as in the *ICKM-EC* model. These two features imply that higher-order uncertainty still affects firms' price-setting decisions. Let us call this model with the imperfect-common-knowledge mechanism (*ICKM*) but no expectation channel as *ICKM*. The parameter set of the model is denoted as:

$$\Theta_{ICKM} = [\theta, \phi_\pi, \phi_y, \beta, \rho_r, \rho_a, \rho_g, \sigma_a, \sigma_a^j, \sigma_r, \sigma_g, \gamma]'$$

4 Empirical Analysis

This section contains the quantitative analysis of the model. I combine a prior distribution for the parameter set of the three models with their likelihood function and conduct Bayesian inference and evaluation.

In Section 4.1, I present the data set and the state-space model for the econometrician. In Section 4.2, I discuss the prior distribution for the model parameters. Section 4.3 presents the posterior distribution. In Section 4.4, we conduct an econometric evaluation of the channel for monetary transmission based on price-setters' expectations. Section 4.5 studies how the informative content of the interest rate changes as the inflation coefficient ϕ_π varies. Section 4.6 investigates what the data tells us about the extent to which the central bank

⁷Note that the history of the past interest rates, R^{t-1} , is still assumed to enter the information set $\tilde{\mathcal{I}}_{j,t}$.

can coordinate heterogenous firms' expectations. 4.7 studies the impulse response functions of the observables (i.e., GDP, inflation, Fed Funds rate, and inflation expectations) to an unanticipated monetary shock.

4.1 Econometrician's State-Space model

The data set includes five observable variables: GDP growth rate, inflation, Federal Funds interest rate, one-quarter-ahead inflation expectations, and four-quarter-ahead inflation expectations. The last two observables are obtained from the *Survey of Professional Forecasters* (SPFs). A detailed description of the data set is provided in Table 1. The data set ranges from 1970:3 to 2008:4. The measurement equations are⁸:

$$100 \ln \frac{PGDP_t}{PGDP_{t-1}} = 100 \ln \pi_* + \hat{\pi}_t$$

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

$$100 \cdot FEDRATE_t = \hat{R}_t + 100 \ln R_*$$

$$\ln \left(\frac{PGDP3_t}{PGDP2_t} \right) 100 = \mathbf{1}_1^T \left[\mathbf{v}_0 \mathbf{T}^{(1)} \mathbf{M} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \left(\mathbf{v}_0 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1} \right) \right] + 100 \ln \pi_* + \sigma_{m_1} \varepsilon_t^{m_1}$$

$$\ln \left(\frac{PGDP6_t}{PGDP2_t} \right) 25 = \mathbf{1}_1^T \left(\sum_{l=0}^4 \mathbf{v}_1^{4-l} \mathbf{v}_0 \mathbf{T}^{(1)} \mathbf{M}^l \varphi_{t|t}^{(0:k)} + \mathbf{v}_1^5 \mathbf{s}_{t-1} \right) + 100 \ln \pi_* + \sigma_{m_2} \varepsilon_t^{m_2}$$

where $PGDP2_t$, $PGDP3_t$, $PGDP6_t$ are the *Survey of Professional Forecasters*⁹ about the current, one-quarter-ahead, and four-quarter ahead GDP price index and the vectors $\varphi_t^{(0:k)}$ and \mathbf{s}_t have been defined in Section 3.8. We relate these statistics with the first moment of the distribution of firms' expectations implied by the model. To avoid stochastic singularity, we introduce two i.i.d. measurement errors $\varepsilon_t^{m_1}$ and $\varepsilon_t^{m_2}$, such that $\varepsilon_t^{m_1} \sim \mathcal{N}(0, 1)$ and $\varepsilon_t^{m_2} \sim \mathcal{N}(0, 1)$. Furthermore, these errors are intended to capture the difference between the observed expectations, which are the mean of the thirty professional forecasters' inflation expectations, and their model concepts, $\hat{\pi}_{t+1|t}^{(1)}$ and $\hat{\pi}_{t+4|t}^{(1)}$.

4.2 Priors

The prior medians and the 95% credible intervals are reported in Table 2. In the pre-sample from 1969:1 to 1970:2, I compute the average of real GDP and inflation to get a measure of

⁸Note that the standard deviations of shocks and measurement errors are rescaled by a factor of 100.

⁹Philly Fed Docs on SPFs to be cited here.

$100 \ln \gamma$ and $100 \ln \pi_*$. We centered the priors for the first two parameters at the value of these averages. Note that in steady state the discount factor β depends on the linear trend of real output γ and the steady-state real interest rate R_*/π_* . Hence, I fix the value for this parameter so as the steady-state nominal interest rate $\ln R_*$ matches the low-frequency behavior of $FEDRATE_t$ in the sample.

Note that the only source of private information for price setters is their productivity $a_{j,t}$. Hence, for given σ_a , the standard deviation of the idiosyncratic technology shocks, σ_a^j determines how precise private information about the level of aggregate technology is. Since private information on a_t is the source of expectations' heterogeneity in the model, the signal-to-noise ration σ_a/σ_a^j affects the extent to which firms' expectations are dispersed. The prior for the standard deviation of technology shock, σ_a , is centered at 0.70, which is consistent with the real business cycle literature (e.g., Kydland and Prescott, 1982). The prior median for the standard deviation of the idiosyncratic technology shocks, σ_a^j , is set so that the model, calibrated at the prior medians, matches the average dispersion among professional forecasters' expectations about the current inflation rate in the sample.

The prior distribution puts probability mass to values for the Calvo parameter θ implying that firms adjust their prices about every three quarters. This belief is derived from micro studies on price-setting (Bils and Klenow, 2004, Klenow and Kryvtsov, 2008, Nakamura and Steinsson, 2008, and Klenow and Malin, 2010).

The priors for the autoregressive parameters ρ_a and ρ_g reflect the belief that the corresponding exogenous processes may exhibit sizeable persistence as it is usually observed in the macroeconomic data. Nonetheless, these priors are broad enough to accommodate a wide range of persistence degrees for these exogenous processes.

Priors for the parameters of the Taylor equation (i.e., response to inflation, ϕ_π , response to economic activity, ϕ_y , autoregressive parameter, ρ_r , and the standard deviation of the i.i.d. monetary shock, σ_r) are chosen as follows. The priors for ϕ_π and ϕ_y are centered at 1.60 and .50, respectively, and imply a fairly strong response to inflation and output gap. The prior for the autoregressive parameter, ρ_r is centered at 0.5, conjecturing that past monetary policy decisions have fairly persistent effects on current central bank's decision over the interest rate.

The volatility of the monetary policy shock, σ_r , demand shock, σ_g are informally taken according to the rule proposed by Del Negro and Schorfheide (2008) that the overall variance of endogenous variables is roughly close to that observed in the pre-sample. The prior median for the measurement errors (i.e., σ_{m1} , σ_{m2}) is set so as to match the variance of inflation expectations reported in the *Livingston survey*.

4.3 Posteriors

A closed-form expression for the posterior distribution is not available (Fernandez-Villaverde and Rubio-Ramirez, 2004), but we can approximate the moments of the posterior distributions via the Metropolis-Hastings. 100,000 posterior draws are obtained from the posterior. The posterior moments for the parameters of the model with the expectation channel¹⁰ are reported in Table 3. The posterior median for the Calvo parameter θ implies quite flexible price contracts, which is in line with what found in micro studies (Bils and Klenow, 2004, Klenow and Kryvtsov, 2008, Nakamura and Steinsson, 2008, and Klenow and Malin, 2010). The posterior median for the autoregressive parameters ρ_a and ρ_g is larger than what is conjectured in the prior. In particular, the autoregressive parameter of technology is close to unity, suggesting that the process of technology is almost a unit root.

The posterior median for the inflation coefficient of the Taylor rule, ϕ_π , is smaller than its prior median. The posterior median for the output coefficient, ϕ_y , is quite large if compared to what is usually obtained when perfect-information models are estimated. The posterior median for the variance of the monetary shock σ_r is bigger than that conjectured in the prior by a factor of 3.5. As observed in Section 3.9, the parameters ϕ_π , ϕ_y , and σ_r affect the informative content of the policy rate, which is the (contemporaneous) public signal for price setters. We will discuss the implications of these three estimated parameters values for the transmission of monetary shocks in detail shortly.

The posterior median for the variance of the preference shock σ_g is larger than the other aggregate shocks, such as σ_a and σ_r . As noted in Section 3.9, the larger the variability of the demand shock, the bigger the relative variability of the endogenous component of the policy signal in equation (15). Hence, a large variance for the preference shock is expected to have critical implications for the propagation of monetary shocks via the expectation channel.

The posterior median for the variance of firm-specific technology shock σ_a^j is larger than the that of the aggregate technology shock, suggesting that developments at firm levels are more volatile than those at the macro levels. This is in line with the literature on micro-evidence of price changes (Klenow and Malin, 2010). Furthermore, the implied signal-to-noise ratio σ_a/σ_a^j is smaller than unity, suggesting that idiosyncratic productivity conveys only a modest amount of information about aggregate productivity. Note that the idiosyncratic productivity is the only source of expectation heterogeneity across firms in the model. Therefore, since the posterior median for the signal-to-noise ratio is not extreme, firms' expectations about the level of aggregate technology are expected to be far from similar and

¹⁰The posterior medians and standard deviations for the parameters of the incomplete-information model (ICKM) with no expectation channel are not reported. The function of this model in the paper is only for evaluating the model with the expectation channel.

actually quite different.¹¹ This point is crucial in the paper and it will be thoroughly studied in the subsequent sections.

4.4 Model Evaluation

Bayesian tests rely on computing the marginal data density (MDD). The marginal data density is needed for updating prior probabilities over a given model space. Denote the data set, presented in Section 4.1, as Y . The MDD associated with the ICKM is defined as $P(Y|\mathcal{M}_{ICKM}) = \int \mathcal{L}(Y|\Theta_{ICKM})p(\Theta) d\Theta$, where $\mathcal{L}(Y|\Theta)$ denotes the likelihood function and $p(\Theta)$ is the prior for its parameters, described in Section 4.2.

A Bayesian test of the null hypothesis that the channel based on price setter's expectation is at odds with the data can be performed by comparing the MDDs associated with the model $\mathcal{M}_{ICKM-EC}$ that features the expectation channel of monetary transmission and the model \mathcal{M}_{ICKM} with no expectation channel. Under a 0 – 1 loss function and unitary prior odds, the test rejects the null that the channel based on price-setters' inflation expectations is at odds with the data, if the incomplete information model has a larger posterior probability than the restricted one (Schorfheide, 2000). The posterior probability of the model \mathcal{M}_s , where $s \in \{ICKM-EC, ICKM\}$, is given by:

$$\pi_{T, \mathcal{M}_s} = \frac{\pi_{0, \mathcal{M}_s} \cdot P(Y|\mathcal{M}_s)}{\sum_{s \in \{ICKM-EC, ICKM\}} \pi_{0, \mathcal{M}_s} \cdot P(Y|\mathcal{M}_s)} \quad (16)$$

where π_{0, \mathcal{M}_s} stands for the prior probability of model \mathcal{M}_s . $P(Y|\mathcal{M}_s)$ is the MDD of model \mathcal{M}_s . We use Geweke's harmonic mean estimator (Geweke, 1999) to approximate the two models. Table 4 compares the MDDs of the incomplete information model with that of the restricted model. The former model attains a larger posterior probability and hence the null can be rejected. The null hypothesis cannot be rejected unless the prior probability in favor of the incomplete information model is smaller than $6.55E - 18$. Such a low prior probability suggests that only if one has extremely strong a-priori information against the expectation-based channel, one can conclude that the channel is not supported by the data. This result favors the empirical relevance of the expectation channel.

We also want to assess whether the new channel based on affecting price setters' expectations is consistent with the observed inflation expectations (i.e., the SPFs). We can

¹¹Note that the mapping from the signal-to-noise ratio to the dispersion of firms' expectations is not monotonic. First, as the $\sigma_a/\sigma_a^j \rightarrow \infty$ there is no source of heterogeneity as all firms will have and observe the same productivity. Second, as $\sigma_a/\sigma_a^j \rightarrow 0$ firms will just neglect the private signal as its noise is too big to make this signal worthy being considered for the price-setting problem. In both those cases, firms will expect that the aggregate level of technology equals its unconditional mean, which is the same across firms, hence no expectation heterogeneity.

use the conditional marginal likelihood to make this evaluation. Denote the subset of the data including only the observed expectations as Y^{SPF} and \tilde{Y} denotes the set including the series for the real GDP growth rate, the inflation rate, and the Federal Funds interest rate. The conditional marginal likelihood is defined as (via the Bayes' Theorem):

$$p\left(Y^{SPF}|\tilde{Y}, \mathcal{M}_s\right) = \frac{p\left(Y^{SPF}, \tilde{Y}|\mathcal{M}_s\right)}{p\left(\tilde{Y}|\mathcal{M}_s\right)}, \quad s \in \{ICKM-EC, ICKM\}$$

The numerator is nothing but the MDDs we have already computed for the *ICKM-EC* model¹² and the *ICKM* model. The denominator is the MDD obtained when the two models are estimated to a data set which does not include the SPFs. The conditional marginal likelihood sheds light on models' goodness of fit only relatively to the data on inflation expectations. Table 4 reports the conditional marginal likelihoods for the incomplete information model and the restricted model. The prior probabilities in favor of the hypothesis that the expectation channel improves the fit of the model with respect to the observed inflation expectation has to be smaller $3.098E - 08$ in order for such hypothesis to be rejected. This is strong evidence in favor of the hypothesis that the new channel based on affecting price setters' expectations is consistent with the observed inflation expectations (i.e., the SPFs).

4.5 Informative Content of the Policy Signal

In this Section, we study the variance decomposition of the Taylor rule (15). As pointed out in Section 3.9, the variance decomposition of the interest rate determines its informative content and, hence, how price setters interpret the observed changes in the rate. Table 5 reports the variance decomposition of the interest rate at the posterior medians reported in Table 3. Most of the variability (i.e., 72%) of the policy signal stems from the shocks to preferences $\varepsilon_{g,t}$. Consequently, when firms observe a rise (fall) in the interest rate, they tend to interpret it as a response of the central bank to a positive (negative) preference shock. The reason why the variability of the policy rate is mainly explained by the shocks to preferences has clearly to do with the large posterior median for σ_g reported in Table 3.

Figure 1 shows how the informative content of the policy signal R_t changes as the inflation coefficient ϕ_π varies. For values of the inflation coefficients ranging from 1.1 to 2.5, the policy signal is always more informative about shocks to preferences. As the central bank puts little emphasize on inflation stabilization, the policy rate tends to signal more information about

¹²Recall that $Y^{SPF} \cup \tilde{Y} = Y$, which is the whole data set used for estimating the incomplete information model in Section 4.3

technology shocks and relatively less information about preference shocks and monetary policy shocks.

4.6 Monetary Policy and Heterogenous Expectations

In this Section, we want to evaluate what the data say about the extent to which the central bank can coordinate heterogenous firms' expectations. As pointed out by Morris and Shin (2003b), a salient feature of public signals, such as the policy rate R_t , is to act as a focal point that coordinates heterogenous higher-order expectations. We call this effect as *Morris and Shin effect*. Nonetheless, since signals are endogenous, there is another effect that goes counter to the *Morris and Shin effect* and, in fact, boosts expectation heterogeneity. On the one hand, the public observability of the interest rate reduces the dispersion of expectations about the level of aggregate technology a_t across firms (i.e., the *Morris and Shin effect*). On the other hand, since R_t links the information about the three exogenous variables (i.e., a_t , $\eta_{r,t}$, and g_t), the policy signal provides a channel through which firms' disagreement about the level of technology can spread out to the three exogenous variables.¹³

To assess to the impact of observing the current interest rate upon the dispersion of firms' expectations, we compare the distribution of expectations across firms in the model with the expectation channel (i.e., the ICKM-EC) and in the model in which the expectation channel is not active (i.e., the ICKM). Since these two models are identical except for whether firms observe the current interest rate, this comparison allows us to isolate the effect of observing the policy signal upon firms' disagreement on fundamentals. The comparison is made by setting the parameter values to the posterior medians for the ICKM-EC, reported in Table 3.

The literature of incomplete common knowledge has developed two methods for measuring how dispersed agents' expectations: The cross-sectional dispersion of expectations and the dispersion across orders of average expectations. Let us start with the first measure.¹⁴ Figure 2 reports the distributions of firms' first-order expectations about the three exogenous variables (a_t , $\eta_{r,t}$, and g_t). The solid black line refers to the distribution in the model with the expectation channel (i.e., ICKM-EC) and the red dashed line refers to the distribution implied by the model in which the policy rate is not observed contemporaneously (i.e., ICKM). These distributions are Gaussian and standardized so that their first moments are the same and equal to zero.¹⁵ As far firm's expectations about technology a_t and prefer-

¹³The only way for the central bank to make expectation heterogeneity fade away would be to set the policy rate as a function of the aggregate technology only.

¹⁴See Nimark (2011) for a thorough description on how to compute these two measures. The exact formula used in this paper is discussed in Appendix D.

¹⁵This standardization is made in order to make it easier for the reader to evaluate the role of the expecta-

ences g_t , the *Morris and Shin effect* dominates and observing the contemporaneous interest rate has the effect of reducing the dispersion of firms' expectations. In contrast, observing the contemporaneous policy rate boosts the heterogeneity of firms' expectations about the monetary shock $\eta_{r,t}$. In this respect, the *Morris and Shin effect* is dominated.

What is the effect upon firms' expectations about the endogenous variables, that is, inflation and output? Figures 3 and 4 show the distribution of firms' expectations about inflation and output when firms observe the current policy rate R_t (solid black line) and when they do not it (i.e., red dashed line). The distributions in the Figure are Gaussian and their support is defined in terms of unit of percentage points. These distributions are standardized so that their first moments are the same and equal to zero. Comparing these two densities sheds light on the extent to which monetary policy -as a provider of the contemporaneous public signal- is able to reduce expectations heterogeneity across firms. It is apparent that the expectation channel turns out to have a quite strong effect in coordinating inflation expectations across firms while it fails coordinating expectations about output. As far as firms' expectations about output, the *Morris and Shin effect* is (almost) perfectly offset by the disagreement about the size of monetary shocks triggered by observing the policy rate. The reason why such a disagreement triggered by the observability of the policy rate has a stronger impact on the dispersion of firms' expectations about output has to do with the equilibrium laws of motion obtained when we solve the model at the estimated parameter values. Apparently, the disagreement about monetary shocks (reported in Figure 2) fuels more disagreement on output than on inflation. Finding the exact reason for this result is complicated in a general equilibrium model, such as the one studied in the paper. Nonetheless, one of the most likely culprits is that the model with the expectation channel also features the traditional new Keynesian channel of monetary transmission, which primarily relies on influencing the intertemporal allocation of consumption. This feature turns out to increase the response of output to an unanticipated monetary shock.

To sum up, the estimated model suggests that the expectation channel reduces the cross-sectional standard deviation of inflation expectations by 38%. In contrast, the channel slightly raises the dispersion of firm's expectations about output.

As far as the second measure that focuses on the dispersion across orders of average expectations, Table 6 reports the dispersion of average expectations about the level of aggregate technology a_t , the monetary policy shock $\eta_{r,t}$, the preferences g_t , inflation (deviation from the steady-state inflation rate π_*), output (deviation from the output trend γ), and the

tion channel for coordinating heterogenous agents expectations (i.e. reducing the dispersion of expectations across firms). The first moment are reported in Table 6. The formula used to computed the second moment of the distributions is derived in Appendix D.

nominal interest rate (deviations from its steady-state value $r_*\pi_*$) upon a negative aggregate technology shock. The perfect observability of the interest rate reduces the dispersion of average expectations about technology a_t across orders. However, the across-order dispersion of average expectations about the monetary shock $\eta_{r,t}$ and preferences g_t turn out to be increased by the perfect observability of the policy signal. Note that when firms observe the policy rate, firms' expectations about the interest rate must obviously be equal to the true rate (i.e., $R_{t|t}^{(0)}$) for any $k \geq 1$. Finally note that the columns $k = 1$ report the first moments of the distributions of Figures 2, 3, and 4.

To sum up, we find that the effect of the policy rate set by the Federal Reserve does not always coordinate expectations and, in fact, foster firms' disagreement. More specifically, the policy rate turns out to reduce firms' disagreement about inflation but has basically no effect on firms disagreement about output. The explanation for the latter result, it is that while monetary policy is found to coordinate expectations about the level of aggregate technology and preferences, it boosts the dispersion of firms' expectations about the monetary shocks, which apparently strongly affect the law of motion of output in the model.

4.7 Propagation of Monetary Shocks

The model features two channels of monetary transmission. First, the traditional New Keynesian channel based on affecting the real interest rate and hence the intertemporal allocation of consumption. Second, the expectation channel of monetary transmission, which is based on how price setters changes their beliefs about the fundamentals (i.e., the level of technology a_t , the monetary shocks $\eta_{r,t}$, and households' preferences) as they observe a change in the policy rate.

Figure 5 shows the impulse response functions (and their 95% posterior credible sets in gray) of real GDP, inflation, interest rate, one-quarter-ahead inflation expectations, and four-quarter-ahead inflation expectations to a 25 basis-point rise in the interest rate. The responses are reported as deviations from the balanced-growth path in units of percentage points of annualized rates. Three features of these impulse response functions have to be emphasized. First, inflation does not react much upon a monetary shock, which is quite unusual in a model with no backward-looking component in the Phillips curve and very flexible price contracts, θ (see Table 3). Second, the 95% posterior credible set for the response of inflation upon the shock includes positive values, which are consistent with the price puzzle found by the VAR literature (Sims, 1992). Third, the response of inflation expectations is positive, suggesting that firms expect rising inflation in the aftermath of a monetary shock.

As we shall show, all these three features have to do with the expectation channel, that is, the perfect observability of the policy rate. Price setters interpret a rising policy rate as a response of the central bank to rising inflation and update their inflation expectations accordingly.

To clarify this point, Table 7 reports the response of the average expectations of the exogenous variables up to the second order, evaluated at the posterior medians in Table 3:¹⁶

$$\varphi_{t|t}^{(0:2)} \equiv \left[a_{t|t}^{(s)}, \eta_{r,t|t}^{(s)}, g_{t|t}^{(s)} : 0 \leq s \leq 2 \right]'$$

Average expectations about preference shocks strongly react after a monetary policy shock. While this finding is not surprising given the variance decomposition of the Taylor rule reported in Table 1, it implies that price-setting firms are uncertain about interpreting the observed rise of the interest rate as a response of the central bank to a preference shock that raises inflation and real output or as a monetary policy shock. Since a preference shock tends to raise nominal marginal costs, firms will decrease (or even rise) their price less strongly than what they would have done if they knew that a monetary shock has occurred. As a consequence, inflation rate does not react much to monetary shocks.

Furthermore, consider the following infinite-order moving-average equilibrium representation for the endogenous state variables $\mathbf{s}_t \equiv \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t \right]'$:

$$\mathbf{s}_t = \sum_{s=0}^{\infty} \mathbf{v}_1^s \mathbf{v}_0 \varphi_{t-s|t-s}^{(0:k)} \quad (17)$$

Given this decomposition, it is easy to isolate the cumulative effects of the average higher-order expectations $\varphi_{t|t}^{(0:k)}$ (up to the truncation at order k) about technology a_t , monetary policy $\eta_{r,t}$, and households' preferences, g_t to inflation, output, and interest rate,

¹⁶Recall that by convention the average zero-order expectations correspond to the realizations of the variables themselves, that is $\varphi_{t|t}^{(0)} = \varphi_t \equiv (z_t, \eta_{r,t}, g_t)'$.

$\mathbf{s}_t \equiv [\hat{\pi}_t, \hat{y}_t, \hat{R}_t]'$ after a monetary policy shock:

$$\begin{aligned}
\frac{\partial \mathbf{s}_{t+h}}{\partial \eta_{r,t}} &= \underbrace{\sum_{l=0}^h \left[\frac{\partial \mathbf{s}_{t+h}}{\partial a_{t+l|t+l}^{(0:k)}} \frac{\partial a_{t+l|t+l}^{(0:k)}}{\partial \eta_{r,t}} \right]}_{\equiv \text{cumulative effects of the HOEs about technology}} \\
&+ \underbrace{\sum_{l=0}^h \left[\frac{\partial \mathbf{s}_{t+h}}{\partial \eta_{r,t+l|t+l}^{(0:k)}} \frac{\partial \eta_{r,t+l|t+l}^{(0:k)}}{\partial \eta_{r,t}} \right]}_{\equiv \text{cumulative effects of the HOEs about monetary policy}} \\
&+ \underbrace{\sum_{l=0}^h \left[\frac{\partial \mathbf{s}_{t+h}}{\partial g_{t+l|t+l}^{(0:k)}} \frac{\partial g_{t+l|t+l}^{(0:k)}}{\partial \eta_{r,t}} \right]}_{\equiv \text{cumulative effects of the HOEs about preferences}} \tag{18}
\end{aligned}$$

for $h = 0, 1, \dots, 20$. Note that under perfect information (i.e., if firms observed the history of all shocks), then

$$\frac{\partial a_{t+l|t+l}^{(0:k)}}{\partial \eta_{r,t}} = \frac{\partial g_{t+l|t+l}^{(0:k)}}{\partial \eta_{r,t}} = 0, \text{ all } k \text{ and } l.$$

This means that the effect of the Higher-Order Expectations (HOEs) about the state of technology and that about the state of preferences are equal to zero in the aftermath of a monetary shock. In contrast, under incomplete information, a large component associated with the HOEs about the state of preferences can be interpreted as a situation in which price-setters mistakenly believe that the interest rate has changed as a result of a preference shock, which has inflationary effects. The cumulative effects of the HOEs about the three shocks is reported in Figure 6 that shows why upon a contractionary monetary shock inflation does not react and may even rise: firms mistakenly interpret the rise in the interest rate as a response of the central bank to a positive demand shock. This can be seen by observing that the effects of the HOEs about preferences upon a monetary shock is positive and so large to offset the effects of HOEs about technology and monetary shocks. Such a strong effect of the HOEs about preferences, is due to the fact that firms interpret a rise of the policy rate as central bank's response to a positive preference shock (see Table 7) and pushes the response of inflation into the positive region. Furthermore, the HOEs about the state of preferences boost the persistent adjustment of inflation to a monetary shock in any subsequent period after the initial shock. The response of the component associated with the HOEs about the state of technology to a contractionary monetary policy is relatively small.

Finally note that the HOEs about technology provide a deflationary contribution even though firms expect a negative technology shock (see Table 7). This apparently controversial

result is very illustrative on how the model works. When price setters observe a rise of the policy rate, they believe, to a certain extent, that the central bank responds to a negative technology shock that tends to raise inflation and the output gap. This is captured by a negative sign for $\frac{\partial a_{t|t}^{(0:k)}}{\partial \eta_{r,t}}$. Nonetheless, if the inflation coefficient ϕ_π is sufficiently large, firms are confident on central bank's ability of controlling the inflationary consequences of the expected negative technology shock. To put this in symbols: $\frac{\partial \pi_t}{\partial a_{t|t}^{(0:k)}} > 0$. Thus, the effect of the HOEs about technology is overall $\frac{\partial \pi_t}{\partial a_{t|t}^{(0:k)}} \frac{\partial a_{t|t}^{(0:k)}}{\partial \eta_{r,t}} < 0$. As a result, no inflation comes from technology shocks. Quite clearly, this result does not always hold. As we shall see, expected negative technology shock will lead to rising inflation for responses to inflation ϕ_π that are lower than the posterior median reported in Table 3.

4.7.1 The Price Puzzle and the Disappearance Thereof after the 1970s

In the VAR literature, the price puzzle (i.e., the positive response of inflation immediately after a contractionary monetary shock) has been considered an empirical regularity since Sims (1992) discovered it. Recent studies based on VAR models (Barth and Ramey, 2001, Hanson, 2004, Ravn, Schmitt-Grohe and Uribe, 2010, and Castelnuovo and Surico, 2010)¹⁷ have documented that the price puzzle is strong when pre-1980s data are used and much weaker or even statistically insignificant when post-1970s data are used. As some of these studies have pointed out, the statistical disappearance of the price puzzle has come exactly with the appointment of Paul Volcker who conducted a monetary policy that was notoriously very aggressive against inflation. The model studied in this paper provides a theoretically consistent explanation for why the price puzzle has been disappeared after such a robust change in the U.S. monetary policy regime.

There is a large consensus about the fact that the Fed has conducted a very passive policy in 1970s.¹⁸ Figure 7 shows how the response of inflation upon a monetary shock changes as ϕ_π falls. Recall that the posterior median for the inflation coefficient is 1.42. As the inflation coefficient gets closer to unity, the price puzzle becomes progressively stronger at an increasing pace. A weak response to inflation, which arguably has characterized the U.S. monetary policy in the 1970s, has given rise to the strong price puzzle documented in the VAR literature for the pre-Volcker period.

¹⁷Barth and Ramey (2001) and Hansen (2004) document the statistically disappearance of the price puzzle by using monthly data and zero-restrictions to identify the monetary shock. Castelnuovo and Surico (2010) use sign restrictions as proposed by Uhlig (2005) to identify the shock. Ravn, Schmitt-Grohe, Uribe, and Uusküla (2010) find that after 1970s the price puzzle is substantially weaker.

¹⁸See Judd and Rudebusch (1998), Clarida, Galí, and Gertler (2000), Boivin and Giannoni (2006), Lubik and Schorfheide (2004), and Cogley and Sargent (2005) among many others.

The intuition for this result goes as follows. When the central bank does not react forcefully enough to inflation two things happen. First, a small inflation coefficient implies that the central bank does not react enough to contrast the inflationary consequences of positive preference shocks and negative technology shocks. Thus, *ceteris paribus*, expected preference and technology shocks will boost price setters' inflation expectations and then inflation as the central bank is deemed to be unable to defeat inflationary effects of these shocks. Second, when the inflation coefficient ϕ_π is very low, the variability of inflation and output turn out to be large and, hence, the variance of the endogenous component in the Taylor rule (15) will be relatively more important than the exogenous component. Such a change in the informative content of the policy rate leads firms to interpret the rise of the interest rate as a response of the central bank to a robust preference shock.

As pointed out in the previous section, as the inflation coefficient falls, the inflationary contribution of an expected negative technology shock (the blue bar in Figure 7) goes from negative values (deflationary effects) to positive and quite large values.

In order to convince agents that a rise in the policy rate does not signal higher inflation, the central bank has to react aggressively to inflation. Note that when the central bank is very active, firms may be still uncertain about whether the interest rate has increased because of a monetary or a preference shock, that is $\frac{\partial g_t^{(1:k)}}{\partial \eta_{r,t}}$ may still be large. Nonetheless, under the more active monetary regime, firms know that the inflationary effects of a preference shock are now mitigated by a more aggressive monetary policy towards inflation stabilization. Consequently, the contribution of the HOEs about preference shocks to inflation upon a monetary shock, $\left(\text{i.e., } \frac{\partial \pi_t}{\partial g_t^{(0:k)}} \frac{\partial g_t^{(0:k)}}{\partial \eta_{r,t}} \right)$ will be small. This leads to a negative response of inflation and inflation expectations to monetary shocks.

5 Concluding Remarks

The paper introduces a DSGE model in which price setters observe the interest rate set by the central bank to infer the nature of shocks that have hit their economy. Since there is strategic complementarities in price setting and price setters observe their idiosyncratic productivity, the model features dispersed information and higher-order expectations. In this model, monetary impulses propagate through two channels. First, the traditional new Keynesian channel based on price stickiness and real interest rate is in place. Second, changing the policy rate conveys non-redundant information about inflation and output gap to price setters. This second channel allows the central bank to affect macroeconomic aggregates by affecting price setters' expectations.

The paper, first, fits the model to a data set that includes the *Survey of Professional*

Forecasters as a measure of the price setters' inflation expectations. Second, the paper performs a formal econometric evaluation of this new transmission channel and finds that it is strongly supported by the data. In particular, the presence of the channel based on affecting price setters' beliefs fits well the *Survey of Professional Forecasters*.

After having established the empirical importance of the new channel, the paper turns to study how monetary impulses transmits to GDP and inflation in the model. We find that the expectation channel accounts for the price puzzle (i.e., the positive contemporaneous responses of inflation to monetary shocks) in the 1970s. The model also explains the disappearance of the price puzzle from the 1980s as a result of a more aggressive monetary policy toward inflation.

The paper also finds that inflation expectations respond positively to monetary shocks because firms interpret the rise of the policy rate as a response of the central bank to a positive demand shock. While the central bank is found to be quite successful in coordinating inflation expectations by maneuvering the policy rate, monetary policy has no effect on the dispersion of expectations about output.

In the model, the central bank communicates with price setters only by setting the policy rate. In other words, the central bank is not allowed to vocally communicate to price setters the state of the economy. On a theoretical ground, this feature of the model is justified because the central bank has an incentive to lie and to make surprise inflation so as to raise output and reduce the monopolistic distortion. Consequently, any announcement made by the central bank will not be regarded as truthful by price setters unless a credible commitment device is in place. On a practical ground, it is however well known that market participants react (and sometimes over-react) to central bank's announcements. Empirically assessing how central bank's communication affects the transmission mechanism of monetary impulses is beyond the scope of this paper.

Finally, the paper relies on a number of assumptions that have been made to improve the tractability of the model. Model tractability is essential for conducting reliable econometric inference. For instance, the paper does not study how households' beliefs adjust to new information coming from the central bank. Estimating a DSGE model where both households and firms have incomplete information is left for future research.

References

- ADAM, K. (2007): “Optimal Monetary Policy with Imperfect Common Knowledge,” *Journal of Monetary Economics*, 54(2), 267–301.
- (2009): “Monetary Policy and Aggregate Volatility,” *Journal of Monetary Economics*, 56(S1), S1–S18.
- 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.
- ANGELETOS, G.-M., C. HELLWIG, AND A. PAVAN (2006): “Signaling in a Global Game: Coordination and Policy Traps,” *Journal of Political Economy*, 114(3), 452–484.
- ANGELETOS, G.-M., AND J. LA’O (2009): “Incomplete Information, Higher-Order Beliefs, and Price Inertia,” *Journal of Monetary Economics*, 56(S1), S19–S37.
- ANGELETOS, G.-M., AND A. PAVAN (2008): “Policy with Dispersed Information,” *Journal of the European Economic Association*, 7(1), 11–60.
- BARTH, M., AND V. RAMEY (2001): “The Cost Channel of Monetary Transmission,” *NBER Macroeconomics Annual*, 16, 199–240.
- BERNANKE, B. S. (2004): “The Great Moderation,” Washington, DC: remarks delivered at the Meeting of the Eastern Economic Association.
- BIANCHI, F. (2010): “Regime Switches, Agents’ Beliefs, and Post-World War II U.S. Macroeconomic Dynamics,” Working Papers 10-39, Duke University, Department of Economics.
- BIANCHI, F., AND L. MELOSI (2010): “Inflationary Sentiments and Monetary Policy Communication,” Discussion paper.
- BILS, M., AND P. J. KLENOW (2004): “Some Evidence on the Importance of Sticky Prices,” *Journal of Political Economy*, 112(5), 947–985.
- BOIVIN, J., AND M. P. GIANNONI (2006): “Has Monetary Policy Become More Effective?,” *The Review of Economics and Statistics*, 88(3), 445–462.
- CALVO, G. A. (1983): “Staggered Prices in a Utility Maximizing Framework,” *Journal of Monetary Economics*, 12(3), 383–398.

- CASTELNUOVO, E., AND P. SURICO (2010): “Monetary policy, inflation expectations and the price puzzle,” *Economic Journal*, 120, 1262–1283.
- 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.
- CLARIDA, R., J. GALÍ, AND M. GERTLER (2000): “Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory,” *The Quarterly Journal of Economics*, 115(1), 147–180.
- COGLEY, T., AND T. J. SARGENT (2005): “Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S,” *Review of Economic Dynamics*, 8(2), 262–302.
- DEL NEGRO, M., AND S. EUSEPI (2010): “Modeling Observed Inflation Expectations,” New York FED mimeo.
- DEL NEGRO, M., AND F. SCHORFHEIDE (2008): “Forming priors for DSGE models (and how it affects the assessment of nominal rigidities),” *Journal of Monetary Economics*, 55(7), 1191–1208.
- EICHENBAUM, M. (1992): “Comment on Interpreting the macroeconomic time series facts: the effects of monetary policy,” *European Economic Review*, 36(5), 1001–1011.
- FERNANDEZ-VILLAYERDE, J., AND J. FRANCISCO RUBIO-RAMIREZ (2004): “Comparing dynamic equilibrium models to data: a Bayesian approach,” *Journal of Econometrics*, 123(1), 153–187.
- 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.
- HANSEN, G. D. (2004): “The price puzzle reconsidered,” *Journal of Monetary Economics*, 51, 1385–1413.
- HELLWIG, C. (2002): “Public Announcements, Adjustment Delays and the Business Cycle,” UCLA Mimeo.

- HELLWIG, C., AND L. VELDKAMP (2009): “Knowing What Others Know: Coordination Motives in Information Acquisition,” *Review of Economic Studies*, 76(1), 223–251.
- JUDD, J. P., AND G. D. RUDEBUSCH (1998): “Taylor’s rule and the Fed, 1970-1997,” *Economic Review*, pp. 3–16.
- JUDD, K. L. (1998): *Numerical Methods in Economics*. The MIT Press, Boston.
- KING, M. (2005): “Monetary Policy: Practice Ahead of Theory,” MAIS Lecture.
- KLENOW, P. J., AND O. KRYVTSOV (2008): “State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?,” *Quarterly Journal of Economics*, 123(3), 863–904.
- KLENOW, P. J., AND B. A. MALIN (2010): “Microeconomic Evidence on Price-Setting,” in *Handbook of Monetary Economics*, ed. by B. Friedman, and M. Woodford, vol. 3A. North Holland, New York.
- KYDLAND, F. E., AND E. C. PRESCOTT (1982): “Time To Build and Aggregate Fluctuations,” *Econometrica*, 50(6), 1345–1370.
- LORENZONI, G. (2009): “A Theory of Demand Shock,” *American Economic Review*, 99(5), 2050–2084.
- (2010): “Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information,” *Review of Economic Studies*, 77(1), 305–338.
- LUBIK, T. A., AND F. SCHORFHEIDE (2004): “Testing for Indeterminacy: An Application to U.S. Monetary Policy,” *American Economic Review*, 94(1), 190–217.
- LUCAS, R. E. J. (1972): “Expectations and the Neutrality of Money,” *Journal of Economic Theory*, 4(2), 103–124.
- MANKIW, G. N., AND R. REIS (2002a): “Sticky Information: A Model of Monetary Non-neutrality 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.
- NAKAMURA, E., AND J. STEINSSON (2008): “Five Facts about Prices: A Reevaluation of Menu Cost Models,” *Quarterly Journal of Economics*, 123(4), 1415–1464.
- NIMARK, K. (2008): “Dynamic Pricing and Imperfect Common Knowledge,” *Journal of Monetary Economics*, 55(8), 365–382.
- (2010): “A Low Dimensional Kalman Filter For Systems with Lagged Observables,” Mimeo.
- (2011): “Dynamic higher order expectations,” .
- PHELPS, E. S. (1970): *Microeconomic Foundations of Employment and Inflation Theory*. Macmillan, London.
- RAVN, M., S. SCHMITT-GROHE, M. URIBE, AND L. UUSKÜLA (2010): “Deep habits and the dynamic effects of monetary policy shocks,” *Journal of the Japanese and International Economies*, 24, 236–258.
- 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.
- RONDINA, G. (2008): “Incomplete Information and Informative Pricing,” University of California San Diego, Mimeo.

- SCHORFHEIDE, F. (2000): “Loss Function-Based Evaluation of DSGE Models,” *Journal of Applied Econometrics*, 15(6), 645–670.
- SIMS, C. A. (1992): “Interpreting the macroeconomic time series facts: the effects of monetary policy,” *European Economic Review*, 36(5), 975–1000.
- SMETS, F., AND R. WOUTERS (2007): “Shocks and Frictions in US Business Cycles: A Bayesian 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.
- TRICHET, J.-C. (2006): “The role of money \ddot{U} money and monetary policy in the 21st century,” Speech at the panel intervention at the Fourth ECB Central Banking Conference \S The role of money: money and monetary policy in the twenty-first century \ddot{T} .
- UHLIG, H. (2005): “What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure,” *Journal of Monetary Economics*, 52, 381–419.
- 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.
- (2005): “Central-Bank Communication and Policy Effectiveness,” Federal Reserve Bank of Kansas City Symposium mimeo.

Tables and Figures

Table 1: Observables

Variables	Description	Source
GDP_t	Gross Domestic Product - Quarterly	BEA (GDPC96)
$POP_t^{\geq 16}$	Civilian noninstitutional population - 16 years and over	BLS (LNS1000000)
$PGDP_t$	Consumer Price Index - Averages of Monthly Figures	BLS (CPIAUCSL)
$FEDRATE_t$	Effective Federal Funds Rate - Averages of Daily Figures	Board of Governors (FEDFUNDS)
$PGDP2_t$	Mean of Expectations of current GDP price index	SPFs in mean.xls (PGDP2)
$PGDP3_t$	Mean of Expectations of one-quarter-ahead GDP price index	SPFs in mean.xls (PGDP3)
$PGDP6_t$	Mean of Expectations of one-year-ahead GDP price index	SPFs in mean.xls (PGDP6)

Table 2: Prior Distributions

Name	Support	Density	Median	95% Interval
θ	$[0, 1]$	Beta	0.65	$[0.45, 0.84]$
ϕ_π	\mathbb{R}^+	Gamma	1.60	$[1.03, 2.19]$
ϕ_y	\mathbb{R}^+	Gamma	0.50	$[0.31, 0.70]$
ρ_r	$[0, 1]$	Beta	0.50	$[0.13, 0.87]$
ρ_a	$[0, 1]$	Beta	0.85	$[0.65, 0.99]$
ρ_g	$[0, 1]$	Beta	0.50	$[0.13, 0.87]$
σ_a	\mathbb{R}^+	InvGamma	0.70	$[0.42, 1.22]$
σ_a^j	\mathbb{R}^+	InvGamma	1.4	$[0.93, 2.17]$
σ_r	\mathbb{R}^+	InvGamma	0.10	$[0.06, 0.16]$
σ_g	\mathbb{R}^+	InvGamma	2.50	$[1.52, 4.36]$
σ_{m_1}	\mathbb{R}^+	InvGamma	0.45	$[0.22, 1.10]$
σ_{m_2}	\mathbb{R}^+	InvGamma	0.45	$[0.22, 1.10]$
$\ln \gamma$	\mathbb{R}	Normal	0.00	$[-0.20, 0.20]$
$\ln \pi_*$	\mathbb{R}	Normal	0.00	$[-0.20, 0.20]$

Table 3: Posterior Distributions

Name	95% Interval		
	Median	Lower	Upper
θ	0.52	0.40	0.60
ϕ_π	1.42	1.30	1.54
ϕ_y	0.83	0.60	1.01
ρ_r	0.76	0.72	0.80
ρ_a	0.98	0.96	0.99
ρ_g	0.83	0.78	0.86
σ_a	1.09	0.95	1.22
σ_a^j	1.45	1.21	1.59
σ_r	0.35	0.30	0.39
σ_g	2.40	2.12	2.70
σ_{m_1}	0.23	0.21	0.26
σ_{m_2}	0.22	0.20	0.25
$100 \ln \gamma$	0.31	0.28	0.34
$100 \ln \pi_*$	0.84	0.55	1.13

Table 4: Marginal-Data-Density Comparisons

	$\mathcal{M}_{ICKM-EC}$	\mathcal{M}_{ICKM}
MDD	-403.63	-443.20
Conditional MDD	-15.71	-33.00

Table 5: Information Content of the Public Signal at the posterior medians

	$\varepsilon_{a,t}$	$\eta_{r,t}$	$\varepsilon_{g,t}$
Informative content of R_t	15.08%	12.98%	71.94%

Table 6: Responses of k-th order average expectations about inflation, output and interest rate upon a two-standard-deviation negative aggregate technology shock

	CHANNEL				NO CHANNEL			
	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$\frac{\partial a_{t t}^{(k)}}{\partial \varepsilon_{a,t}}$	-1.09	-0.55	-0.39	-0.35	-1.09	0.39	0.14	0.05
$\frac{\partial \widehat{\pi}_{t t}^{(k)}}{\partial \varepsilon_{a,t}}$	0.338	0.29	0.21	0.15	0.40	0.17	0.11	0.08
$\frac{\partial \widehat{y}_{r,t t}^{(k)}}{\partial \varepsilon_{a,t}}$	-0.60	-0.52	-0.45	-0.40	-0.52	-0.38	-0.27	-0.20
$\frac{\partial R_{t t}^{(k)}}{\partial \varepsilon_{a,t}}$	0.13	0.13	0.13	0.13	0.13	0.10	0.07	0.05

Table 7: Contemporaneous responses of average expectations of order k to a monetary policy shock

	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$\frac{\partial a_{t t}^{(k)}}{\partial \eta_{r,t}}$	0.00	-0.19	-0.26	-0.28
$\frac{\partial \eta_{r,t t}^{(k)}}{\partial \eta_{r,t}}$	0.38	0.22	0.21	0.21
$\frac{\partial g_{t t}^{(k)}}{\partial \eta_{r,t}}$	0.00	1.21	1.18	1.14

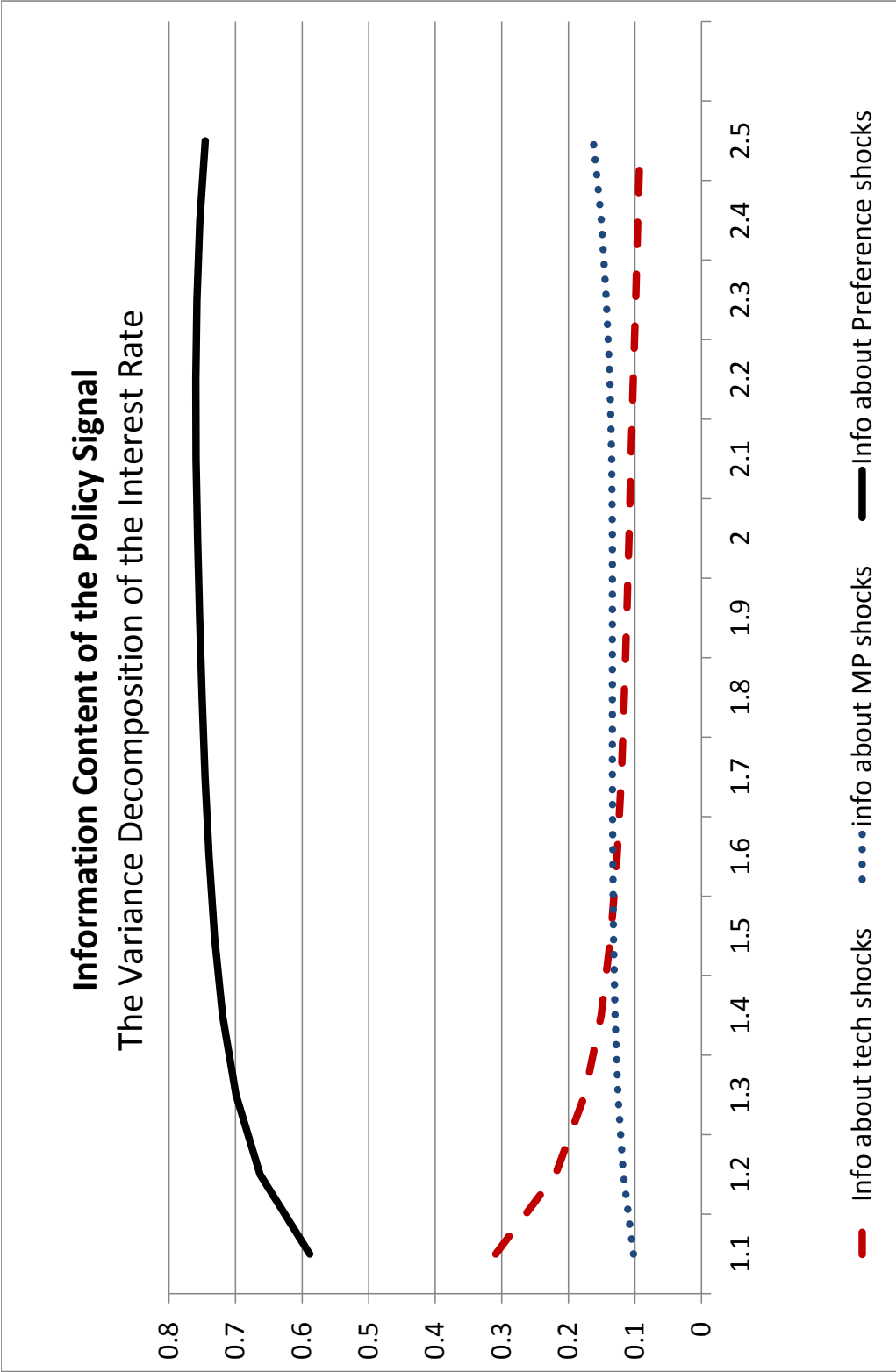


Figure 1: Information Content of the Policy Signal with Respect to the Three Shocks of the Model as the Inflation Coefficient of the Taylor Rule Changes. All parameter but the inflation coefficient are set at the posterior medians

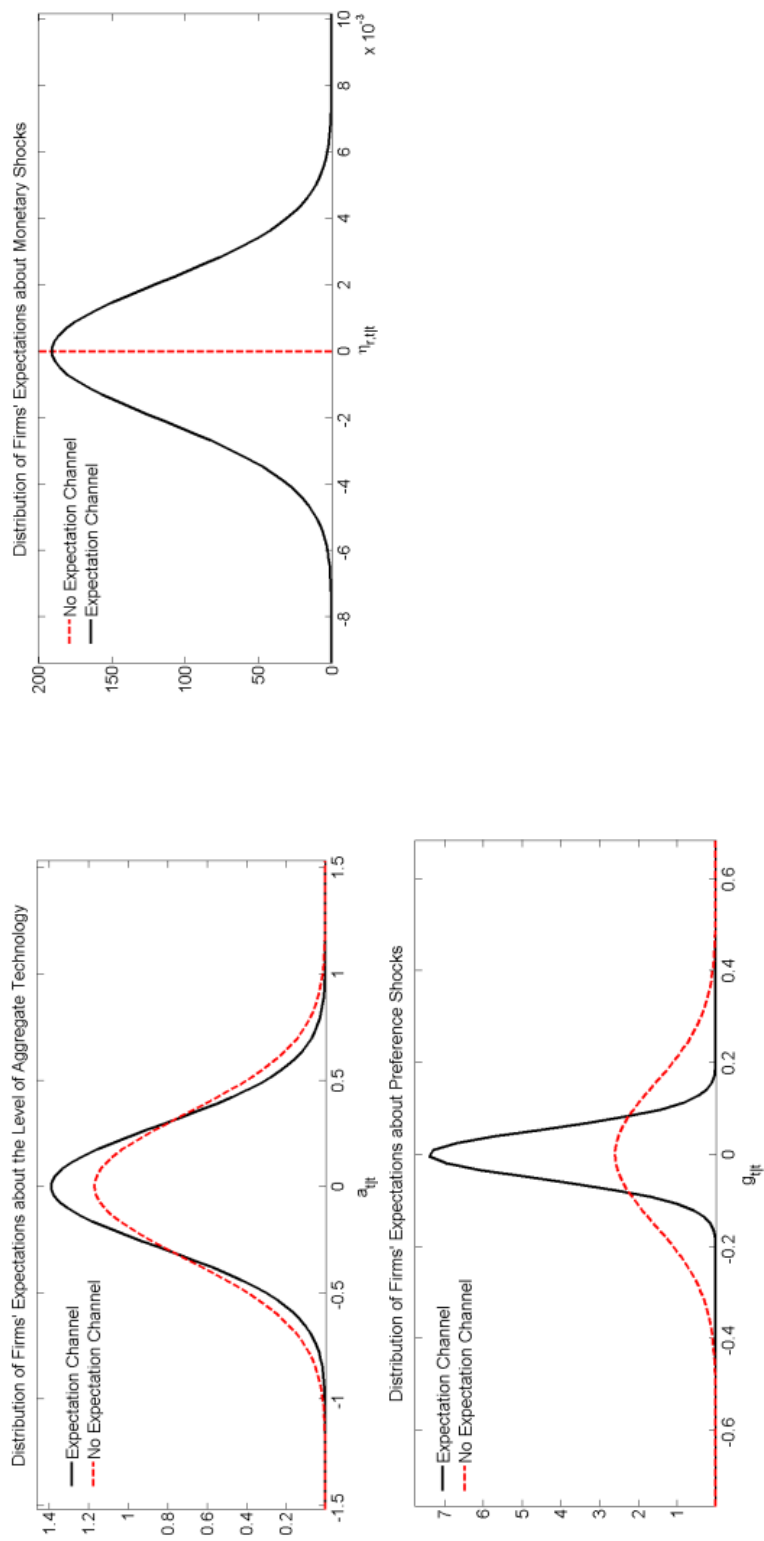


Figure 2: Dispersion of Firms' Expectations about Exogenous State Variables (first moment normalized to be equal to zero). The inflation expectations in the horizontal axis are expressed as units of percentage points. Evaluated at the posterior medians.

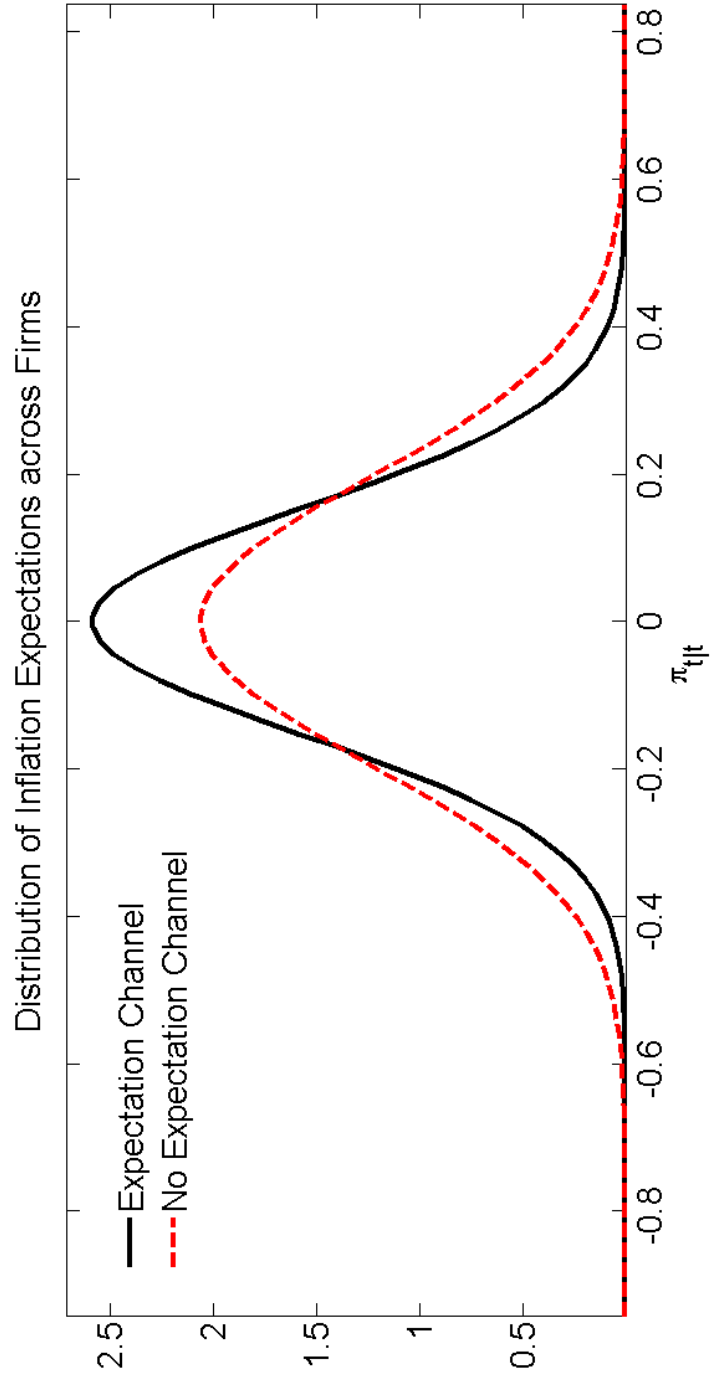


Figure 3: Dispersion of Inflation Expectations across Firms (first moment normalized to be equal to zero). The inflation expectations in the horizontal axis are expressed as units of percentage points. Evaluated at the Posterior Medians

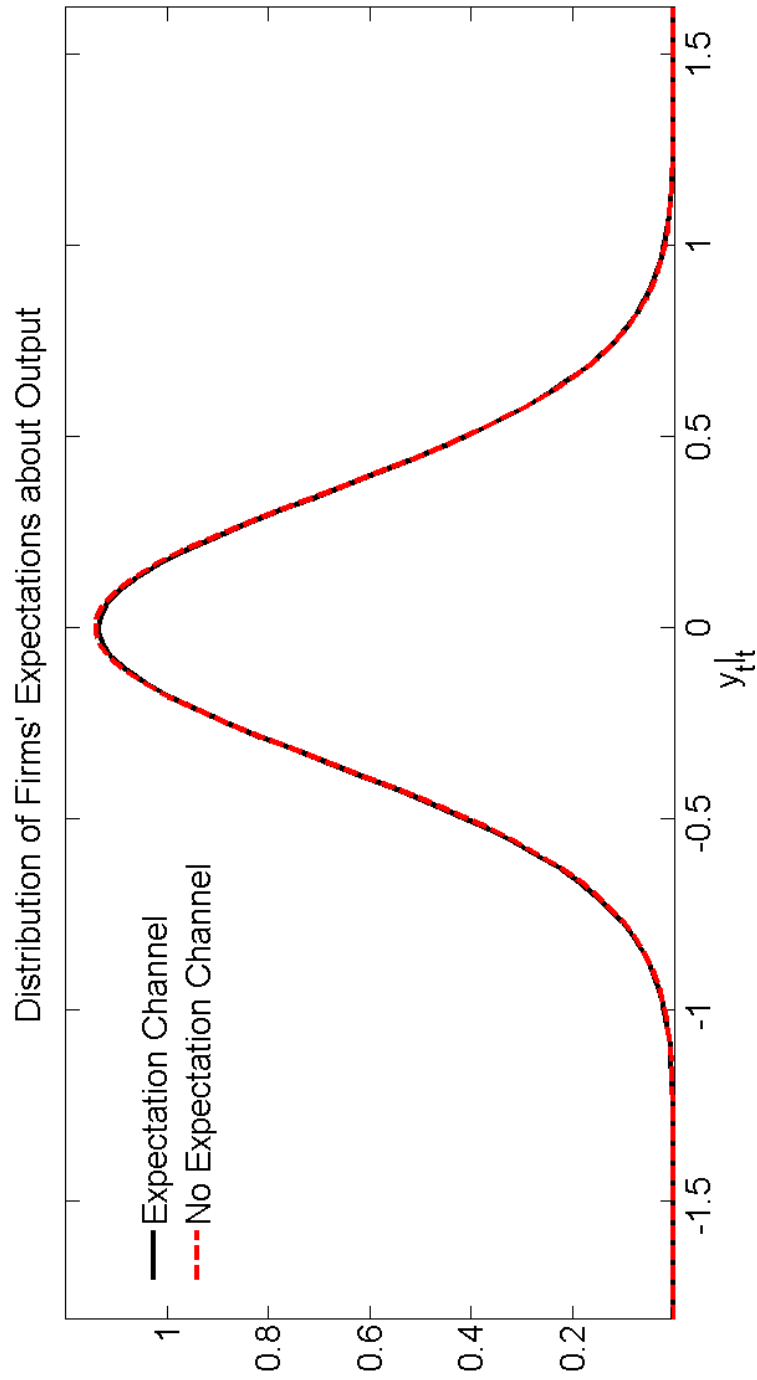


Figure 4: Dispersion of Firms' Expectations about Output (first moment normalized to be equal to zero). The inflation expectations in the horizontal axis are expressed as units of percentage points. Evaluated at the Posterior Medians

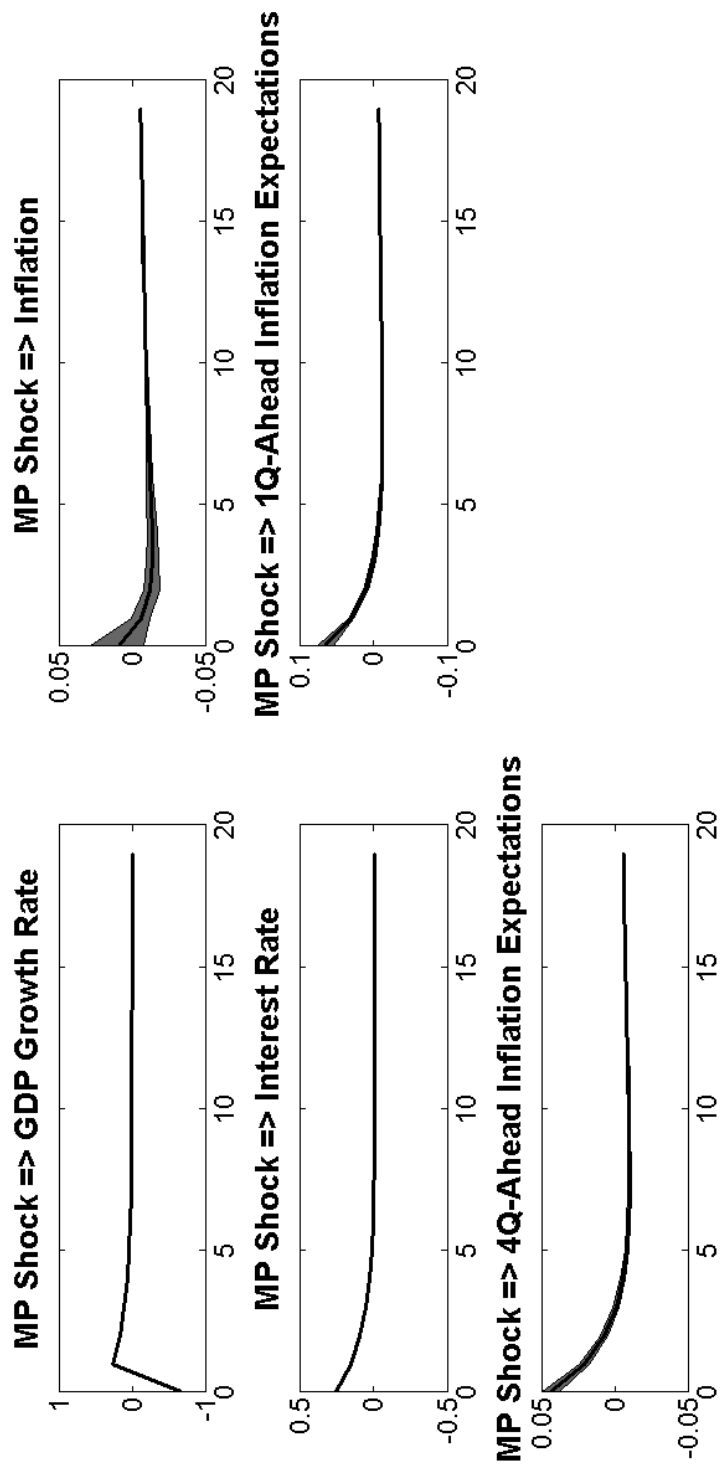


Figure 5: Impulse Response Functions of the Observables to a Monetary Policy Shock.

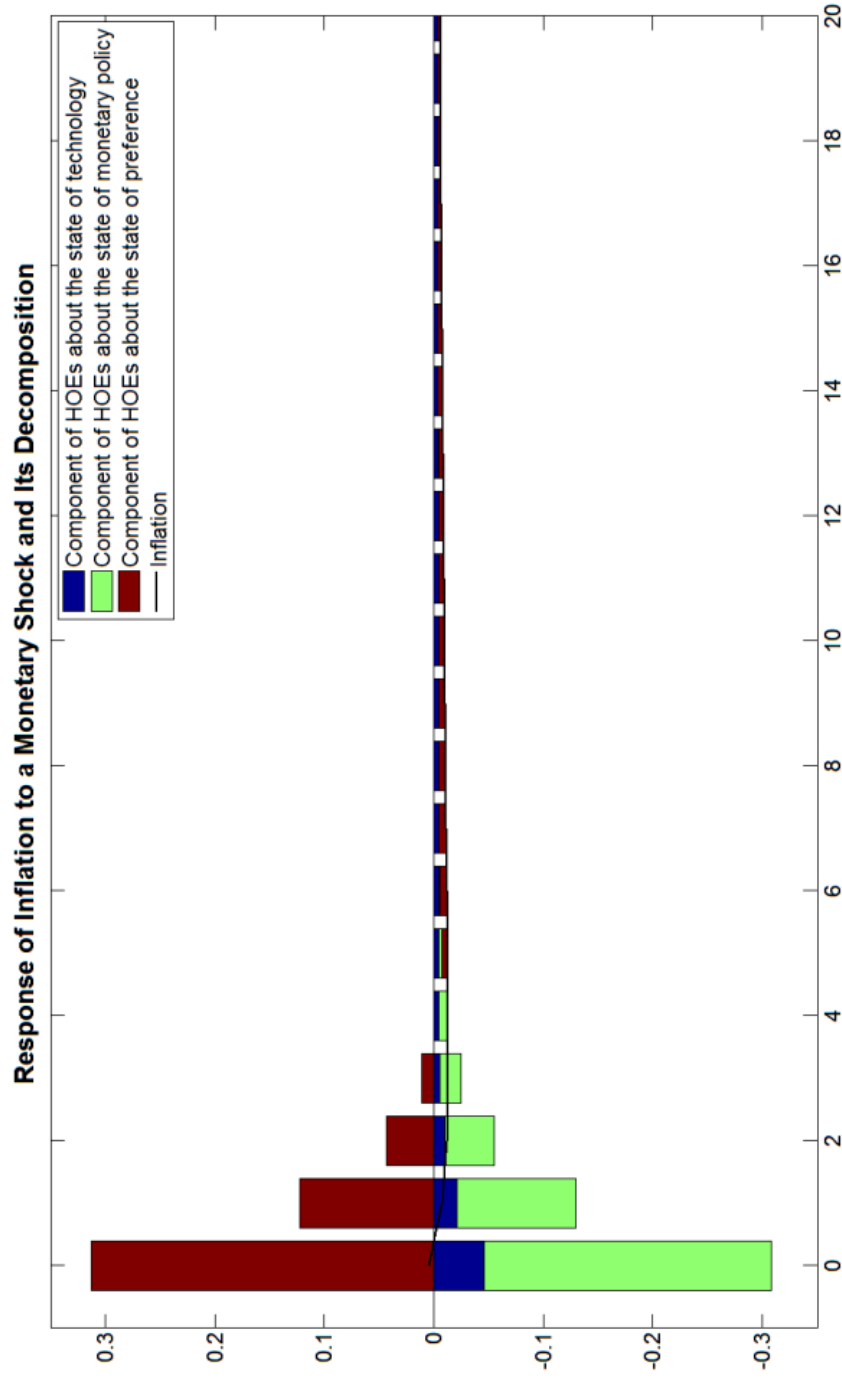


Figure 6: HOE Decomposition of the IRFs of Inflation to a Monetary Policy Shock. Parameters are set to be equal to the posterior medians

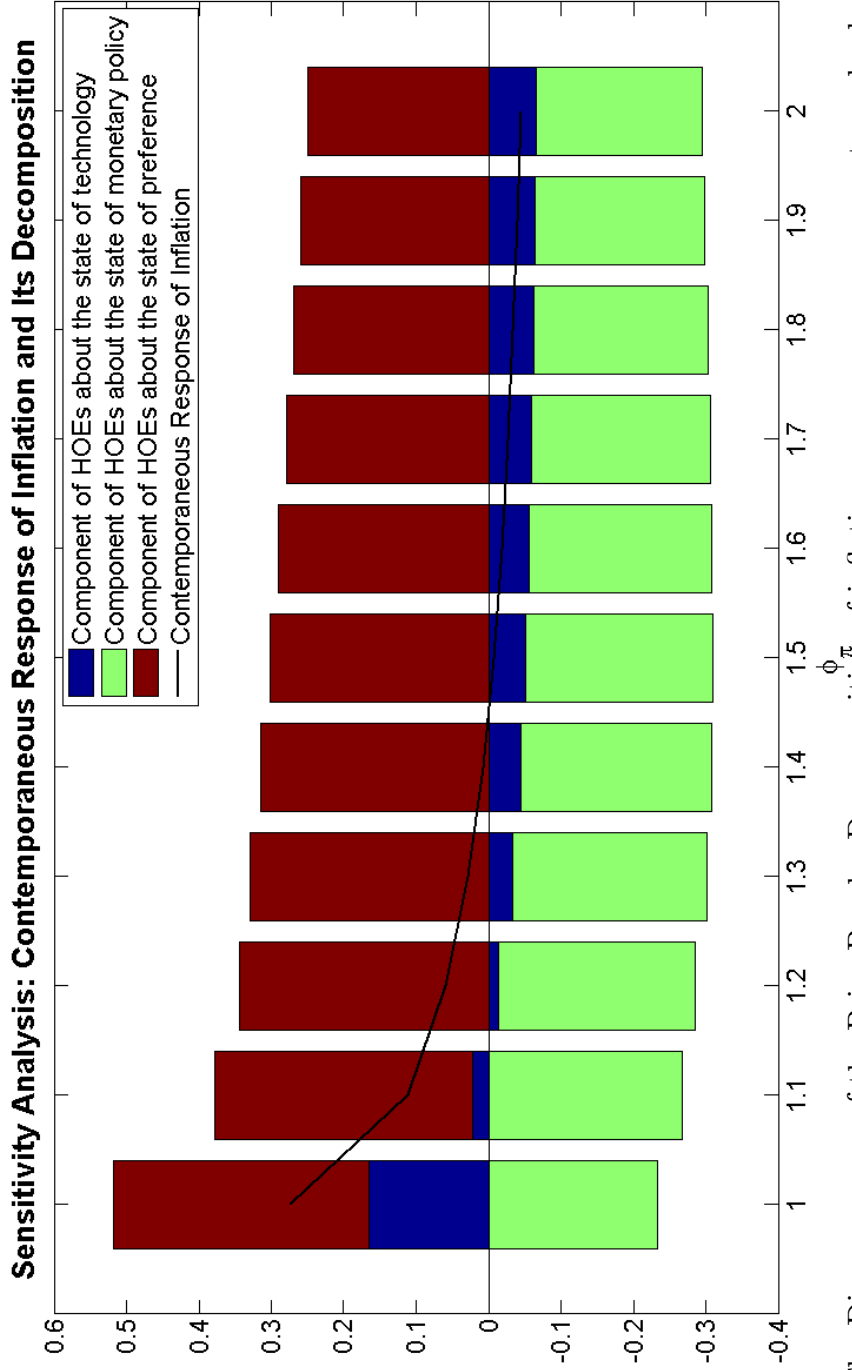


Figure 7: The Disappearance of the Price Puzzle: Decomposition of inflation responses upon a monetary shock as the inflation coefficient in the Taylor rule varies

Appendix

In Section A, I provide derivation of the imperfect-common-knowledge Phillips curve (10). In Section B, I show how to characterize the laws of motion for the three endogenous state variables (i.e., inflation $\hat{\pi}_t$, real output \hat{y}_t and the interest rate \hat{R}_t). In Section C, I characterize the transition equations for the average higher-order expectations about the exogenous state variables.

A The Imperfect Common Knowledge Phillips Curve

The log-linear approximation of the labor supply can be shown to be given by $\hat{c}_t = \hat{w}_t$. Recalling that the resource constraint implies that $\hat{y}_t = \hat{c}_t$, then the labor supply can be written as follows:

$$\hat{y}_t = \hat{w}_t \tag{19}$$

Log-linearizing the equation for the real marginal costs (4) yields

$$\widehat{mc}_{j,t} = \hat{w}_{j,t} - a_t - \eta_{j,t}^a$$

Recall that $(\ln A_{j,t} - \ln \gamma \cdot t) \in \mathcal{I}_{j,t}$ and write:

$$\mathbb{E}_{j,t} \widehat{mc}_{j,t} = \mathbb{E}_{j,t} \underbrace{\hat{w}_{j,t} - a_t - \eta_{j,t}^a}_{\ln A_{j,t} - \ln \gamma \cdot t}$$

where $\mathbb{E}_{j,t}$ are expectations conditioned on firm j 's information set at time t , $\mathcal{I}_{j,t}$, defined in (7). Using the equation (19) for replacing \hat{w}_t yields:

$$\mathbb{E}_{j,t} \widehat{mc}_{j,t} = \mathbb{E}_{j,t} \hat{y}_t - a_t - \eta_{j,t}^a$$

By integrating across firms, we obtain the average expectations on marginal costs:

$$\widehat{mc}_{t|t}^{(1)} = \hat{y}_{t|t}^{(1)} - a_t$$

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 $\hat{\pi}_t = \ln P_t - \ln P_{t-1} - \ln \pi_*$,

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

and then

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

By rearranging:

$$\ln P_t = \theta (\ln P_{t-1} + \ln \pi_*) + (1 - \theta) \int (\ln P_{j,t}^*) dj \quad (20)$$

The price-setting equation is:

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

Define

$$\begin{aligned} y_t &= \frac{Y_t}{\gamma^t}, \quad c_t = \frac{C_t}{\gamma^t}; \quad p_{j,t}^* = \frac{P_{j,t}^*}{P_t}, \quad y_{j,t} = \frac{Y_{j,t}}{\gamma^t} \\ w_t &= \frac{W_t}{\gamma^t P_t}, \quad a_t = \frac{A_t}{\gamma^t}, \quad R_t = \frac{R_t}{R_*}, \quad mc_{j,t} = \frac{MC_{j,t}}{P_t} \\ \xi_{j,t} &= \gamma^t \Xi_{j,t} \end{aligned}$$

Hence, write:

$$\mathbb{E} \left\{ \xi_{j,t} \left[1 - \nu + \nu \frac{mc_{j,t}}{p_{j,t}^*} \right] y_{j,t} + \sum_{s=1}^{\infty} (\beta\theta)^s \xi_{j,t+s} \left[(1 - \nu) \pi_*^s + \nu \frac{mc_{j,t+s}}{p_{j,t}^*} (\prod_{\tau=1}^s \pi_{t+\tau}) \right] y_{j,t+s} | \mathcal{I}_{j,t} \right\} = 0 \quad (21)$$

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. We can write

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

Taking the derivatives yield:

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

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:

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

Recall that $\widehat{p}_{j,t}^* \equiv \ln P_{j,t}^* - \ln P_t$ and cannot be taken out of the expectation operator. We write:

$$\ln P_{j,t}^* = (1 - \beta\theta) \mathbb{E} \left[\widehat{mc}_{j,t} + \frac{1}{1 - \beta\theta} \ln P_t + \sum_{s=1}^{\infty} (\beta\theta)^s \left(\widehat{mc}_{j,t+s} + \sum_{\tau=1}^s \widehat{\pi}_{t+\tau} \right) | \mathcal{I}_{j,t} \right] \quad (22)$$

Rolling this equation one step ahead yields:

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

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

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

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

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

Therefore,

$$\sum_{s=1}^{\infty} (\beta\theta)^s \mathbb{E}[\widehat{m}c_{j,t+s} | \mathcal{I}_{j,t}] = \frac{\beta\theta}{1 - \beta\theta} [\mathbb{E}(\ln P_{j,t+1}^* | \mathcal{I}_{j,t}) - \mathbb{E}(\ln P_{t+1} | \mathcal{I}_{j,t})] - \beta\theta \sum_{s=1}^{\infty} (\beta\theta)^s \sum_{\tau=1}^s \mathbb{E}[\widehat{\pi}_{t+\tau+1} | \mathcal{I}_{j,t}] \quad (23)$$

The equation (22) can be rewritten as:

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

By substituting the result in equation (23) we obtain:

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

Consider the last term:

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

Therefore we can write that

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

Note that

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

Hence,

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

and by simplifying:

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

We substitute this result into the original equation to get:

$$\begin{aligned}
\ln P_{j,t}^* &= (1 - \beta\theta) \left[\mathbb{E} [\widehat{m}c_{j,t} | \mathcal{I}_{j,t}] + \frac{1}{1 - \beta\theta} \mathbb{E} [\ln P_t | \mathcal{I}_{j,t}] \right] \\
&+ \beta\theta \left[\mathbb{E} (\ln P_{j,t+1}^* | \mathcal{I}_{j,t}) - \mathbb{E} (\ln P_{t+1} | \mathcal{I}_{j,t}) \right] - (1 - \beta\theta) \sum_{s=1}^{\infty} (\beta\theta)^{s+1} \sum_{\tau=1}^s \mathbb{E} [\hat{\pi}_{t+\tau+1} | \mathcal{I}_{j,t}] \\
&+ \beta\theta \mathbb{E} [\hat{\pi}_{t+1} | \mathcal{I}_{j,t}] + (1 - \beta\theta) \sum_{s=1}^{\infty} (\beta\theta)^{s+1} \sum_{\tau=1}^s \mathbb{E} [\hat{\pi}_{t+\tau+1} | \mathcal{I}_{j,t}]
\end{aligned} \tag{24}$$

After simplifying we get:

$$\begin{aligned}
\ln P_{j,t}^* &= (1 - \beta\theta) \left[\mathbb{E} [\widehat{m}c_{j,t} | \mathcal{I}_{j,t}] + \frac{1}{1 - \beta\theta} \mathbb{E} [\ln P_t | \mathcal{I}_{j,t}] \right] \\
&+ \beta\theta \left[\mathbb{E} (\ln P_{j,t+1}^* | \mathcal{I}_{j,t}) - \mathbb{E} (\ln P_{t+1} | \mathcal{I}_{j,t}) \right] + \beta\theta \mathbb{E} [\hat{\pi}_{t+1} | \mathcal{I}_{j,t}]
\end{aligned} \tag{25}$$

We can rearrange:

$$\begin{aligned} \ln P_{j,t}^* &= (1 - \beta\theta) \mathbb{E} [\widehat{mc}_{j,t} | \mathcal{I}_{j,t}] + \mathbb{E} [\ln P_t | \mathcal{I}_{j,t}] \\ &\quad + \beta\theta \left[\mathbb{E} (\ln P_{j,t+1}^* | \mathcal{I}_{j,t}) + \mathbb{E} [\widehat{\pi}_{t+1} | \mathcal{I}_{j,t}] - \mathbb{E} (\ln P_{t+1} | \mathcal{I}_{j,t}) \right] \end{aligned} \quad (26)$$

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

$$\begin{aligned} \ln P_{j,t}^* &= (1 - \beta\theta) \cdot \mathbb{E} [\widehat{mc}_{j,t} | \mathcal{I}_{j,t}] + (1 - \beta\theta) \mathbb{E} [\ln P_t | \mathcal{I}_{j,t}] \\ &\quad + \beta\theta \cdot \mathbb{E} (\ln P_{j,t+1}^* | \mathcal{I}_{j,t}) - \beta\theta \ln \pi_* \end{aligned} \quad (27)$$

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

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

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 (27) across firms to obtain an equation for the average reset price:

$$\begin{aligned} \ln P_t^* &= (1 - \beta\theta) \cdot \widehat{mc}_{t|t}^{(1)} + (1 - \beta\theta) \ln P_t^{(1)} \\ &\quad + \beta\theta \ln P_{t+1|t}^{*(1)} - \beta\theta \ln \pi_* \end{aligned} \quad (28)$$

where we use the claim of the proposition above. Keep in mind that the price index equation can be manipulated to get equation (20)

$$\ln P_t = \theta (\ln P_{t-1} + \ln \pi_*) + (1 - \theta) \ln P_t^* \quad (29)$$

Let us plug the equation (28) into the equation (29):

$$\begin{aligned} \ln P_t &= \theta \ln P_{t-1} + (\theta - (1 - \theta) \beta\theta) \ln \pi_* \\ &\quad + (1 - \theta) \left[(1 - \beta\theta) \cdot \widehat{mc}_{t|t}^{(1)} + (1 - \beta\theta) \ln P_t^{(1)} + \beta\theta \ln P_{t+1|t}^{*(1)} \right] \end{aligned} \quad (30)$$

Use the fact that $\ln P_t = \hat{\pi}_t + \ln P_{t-1} + \ln \pi_*$ and from the price index (20):¹⁹

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

Furthermore, the following fact is easy to establish:

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

Applying these three results to equation (30) yields:

$$\begin{aligned} \hat{\pi}_t + \ln P_{t-1} + \ln \pi_* &= \theta \ln P_{t-1} + (\theta - (1-\theta)\beta\theta) \ln \pi_* \\ &+ (1-\theta) \left[(1-\beta\theta) \cdot \widehat{mc}_{t|t}^{(1)} + (1-\beta\theta) \ln P_{t|t}^{(1)} + \beta\theta \left(\frac{\hat{\pi}_{t+1|t}^{(1)}}{1-\theta} + \ln P_{t|t}^{(1)} + \ln \pi_* \right) \right] \end{aligned} \quad (31)$$

and after simplifying:

$$\hat{\pi}_t = (1-\theta)(1-\beta\theta) \cdot \widehat{mc}_{t|t}^{(1)} + (1-\theta) \hat{\pi}_{t|t}^{(1)} + \beta\theta \left(\hat{\pi}_{t+1|t}^{(1)} \right) \quad (32)$$

By repeatedly taking firm j 's expectation and average the resulting equation across firms:

$$\hat{\pi}_{t|t}^{(k)} = (1-\theta)(1-\beta\theta) \cdot \widehat{mc}_{t|t}^{(k)} + (1-\theta) \hat{\pi}_{t|t}^{(k+1)} + \beta\theta \left(\hat{\pi}_{t+1|t}^{(k+1)} \right)$$

Repeatedly substituting these equations for $k \geq 1$ back to equation (32) yields: the imperfect-common-knowledge Phillips curve:

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

B The Laws of Motion for the Endogenous State Variables

In this section I, first, introduce some useful results and, second, characterize the law of motion for the endogenous state variables $(\hat{\pi}_t, \hat{y}_t, \hat{R}_t)$, which are inflation $\hat{\pi}_t$, real output \hat{y}_t , and the (nominal) interest rate \hat{R}_t . It will be shown that this law of motion depends on model parameters and the

¹⁹This last result comes from observing that

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

By using the fact that $\ln P_t = \hat{\pi}_t + \ln P_{t-1} + \ln \pi_*$:

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

Rolling one period forward:

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

and finally by rearranging we get the result in the text.

coefficient matrices, \mathbf{M} and \mathbf{N} , of the transition equation for the average higher-order expectations about the exogenous variables.

B.1 Preliminaries

Recall that the *assumption of common knowledge in rationality* ensures that agents use the actual law of motion of higher-order expectations to forecast the dynamics of the higher-order expectations. The following claims turn out to be useful:

Proposition 1 *If one neglects the effect of average beliefs of order larger than k , then the following is approximately true:*

$$\varphi_{t|t}^{(s:k+s)} = \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k)}$$

where

$$\mathbf{T}^{(s)} \equiv \begin{bmatrix} \mathbf{0}_{3(k-s+1) \times 3s} & \mathbf{I}_{3(k-s+1)} \\ \mathbf{0}_{3s \times 3s} & \mathbf{0}_{3s \times (k+1-s)3} \end{bmatrix}$$

Proof. It is straightforward but help to fix some notation. Since we neglect the average beliefs of order larger than k

$$\varphi_{t|t}^{(s:k+s)} \equiv \begin{bmatrix} \varphi_{t|t}^{(s:k)} \\ \varphi_{t|t}^{(s:k+s)} \end{bmatrix}_{3(k+1) \times 1} = \begin{bmatrix} \varphi_{t|t}^{(s:k)} \\ \mathbf{0}_{3s \times 1} \end{bmatrix}_{3(k+1) \times 1}$$

Note that

$$\varphi_{t|t}^{(s:k+s)} = \begin{bmatrix} \mathbf{0}_{3(k-s+1) \times 3s} & \mathbf{I}_{3(k-s+1)} \\ \mathbf{0}_{3s \times 3s} & \mathbf{0}_{3s \times (k+1-s)3} \end{bmatrix} \underbrace{\begin{bmatrix} \varphi_{t|t}^{(0:s-1)} \\ \varphi_{t|t}^{(s:k)} \end{bmatrix}}_{\varphi_{t|t}^{(0:k)}}$$

■

Proposition 2 $\mathbf{s}_{t|t}^{(s)} = \mathbf{v}_0 \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k+s)} + \mathbf{v}_1 \mathbf{s}_{t-1}$, for any $0 \leq s \leq k$.

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

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

■

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

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

Proposition 3 *The following holds true for any $h \in \{0 \cup \mathbb{N}\}$*

$$\mathbf{s}_{t+h|t}^{(1)} = \sum_{l=0}^h \mathbf{v}_1^{h-l} \mathbf{v}_0 \mathbf{M}^l \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1^{h+1} \mathbf{s}_{t-1}$$

Proof. Consider

$$\mathbf{s}_t = \mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1}$$

Then it is easy to see that by taking agents' expectations and then averaging across them we obtain by the *assumption of common knowledge in rationality*:

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

and by neglecting the contribution of beliefs of order higher than k we can write: $\mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} = \varphi_{t|t}^{(1:k+1)}$. This leads to write:

$$\mathbf{s}_{t|t}^{(1)} = \mathbf{v}_0 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1} \quad (33)$$

Furthermore, consider \mathbf{s}_{t+1} :

$$\mathbf{s}_{t+1} = \mathbf{v}_0 \varphi_{t+1|t+1}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_t$$

By taking agents' expectations and then averaging across them we obtain:

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

Recall result (33) and write:

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

First note that by the assumption of common knowledge in rationality we can write: $\varphi_{t+h|t}^{(1:k+1)} = \mathbf{M}^h \varphi_{t|t}^{(1:k+1)}$. Second, recall that we neglect the contribution of beliefs of order higher than k . These two facts lead us to

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

Consider now \mathbf{s}_{t+2} . By taking agents' expectations and then averaging across them we obtain:

$$\mathbf{s}_{t+2|t}^{(1)} = \mathbf{v}_0 \varphi_{t+2|t}^{(1:k+1)} + \mathbf{v}_1 \mathbf{s}_{t+1|t}^{(1)}$$

and substituting $\mathbf{s}_{t+1|t}^{(1)}$ that we have characterized above yields:

$$\mathbf{s}_{t+2|t}^{(1)} = \mathbf{v}_0 \mathbf{M}^2 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \left\{ \mathbf{v}_0 \mathbf{M} \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \left[\mathbf{v}_0 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1} \right] \right\}$$

Keeping on deriving $\mathbf{s}_{t+h|t}^{(1)}$ for any other $h \in \{0 \cup \mathbb{N}\}$ as shown above leads at the formula in the claim. ■

B.2 The Laws of Motion of the Endogenous State Variables

The laws of motion of the three endogenous state variables, which are inflation $\hat{\pi}_t$, real output \hat{y}_t , and the (nominal) interest rate \hat{R}_t , are given by the IS equation (12), the Phillips curve (10), and the Taylor Rule (13). One can use these structural equations to pin down the vectors $\mathbf{v}_0 \equiv [\mathbf{a}'_0, \mathbf{b}'_0, \mathbf{c}'_0]'$

and $\mathbf{v}_1 \equiv [\mathbf{a}'_1, \mathbf{b}'_1, \mathbf{c}'_1]'$ in the equations below:

$$\mathbf{s}_t = \mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1}$$

where $\mathbf{s}_t \equiv [\hat{\pi}_t, \hat{y}_t, \hat{R}_t]'$.

Let start from the IS equation (12)

$$\begin{aligned} \mathbf{b}_0 \varphi_{t|t}^{(0:k)} + \mathbf{b}_1 \mathbf{s}_{t-1} &= \mathbf{1}_3^T \varphi_{t|t}^{(0:k)} - \mathbf{1}_3^T \mathbf{M} \varphi_{t|t}^{(0:k)} \\ &+ (\mathbf{1}_1^T + \mathbf{1}_2^T) \left[\underbrace{\mathbf{v}_0 \mathbf{M} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 (\mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1})}_{\mathbb{E}_t \mathbf{s}_{t+1}} \right] \\ &- (\mathbf{c}_0 \varphi_{t|t}^{(0:k)} + \mathbf{c}_1 \mathbf{s}_{t-1}) \end{aligned}$$

where we used Proposition 3, $\mathbf{1}_i^T$ ($i \in \mathbb{N}$) is a comfortable row vector of all zero elements except for the i -th element, which is equal to one. One can show that the following condition has to be satisfied by the vectors of coefficients, \mathbf{b}_0 and \mathbf{b}_1 , to ensure that the IS equation (12) holds in equilibrium.

$$\mathbf{b}_0 = \mathbf{1}_3^T + (\mathbf{1}_1^T + \mathbf{1}_2^T) (\mathbf{v}_0 \mathbf{M} + \mathbf{v}_1 \mathbf{v}_0) - \mathbf{1}_3^T \mathbf{M} - \mathbf{c}_0 \quad (34)$$

$$\mathbf{b}_1 = (\mathbf{1}_1^T + \mathbf{1}_2^T) \mathbf{v}_1 \mathbf{v}_1 - \mathbf{c}_1 \quad (35)$$

The Phillips curve (10) can be rewritten as:

$$\begin{aligned} \mathbf{a}_0 \varphi_{t|t}^{(0:k)} + \mathbf{a}_1 \mathbf{s}_{t-1} &= (1 - \theta) (1 - \beta \theta) \sum_{s=0}^{k-1} (1 - \theta)^s \mathbf{1}_2^T \left[\mathbf{v}_0 \mathbf{T}^{(s+1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1} \right] + \\ &- (1 - \theta) (1 - \beta \theta) \sum_{s=0}^{k-1} (1 - \theta)^s \left[\boldsymbol{\gamma}_a^{(s)'} \varphi_{t|t}^{(0:k)} \right] \\ &+ \beta \theta \sum_{s=0}^{k-1} (1 - \theta)^s \mathbf{1}_1^T \left[\mathbf{v}_0 \mathbf{M} \mathbf{T}^{(s+1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 (\mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1}) \right] \end{aligned}$$

where $\boldsymbol{\gamma}_a^{(s)} = [\mathbf{0}_{1 \times 3s}, (1, 1, 0, 0, 0, 0), \mathbf{0}_{1 \times 3(k-s)}]'$. The following restrictions upon vectors of coefficients \mathbf{a}_0 and \mathbf{a}_1 can be derived from the Phillips curve above:

$$\begin{aligned} \mathbf{a}_0 &= (1 - \theta) (1 - \beta \theta) \left[\boldsymbol{\nu} \mathbf{m}_1 - \left(\sum_{s=0}^{k-1} (1 - \theta)^s \boldsymbol{\gamma}_a^{(s)'} \right) \right] \\ &+ \beta \theta \boldsymbol{\nu} \mathbf{m}_2 + \beta \theta \left(\sum_{s=0}^{k-1} (1 - \theta)^s \right) \mathbf{1}_1^T \mathbf{v}_1 \mathbf{v}_0 \end{aligned} \quad (36)$$

$$\mathbf{a}_1 = (1 - \theta) (1 - \beta \theta) \left(\sum_{s=0}^{k-1} (1 - \theta)^s \right) \mathbf{1}_2^T \mathbf{v}_1 + \beta \theta \left(\sum_{s=0}^{k-1} (1 - \theta)^s \right) \mathbf{1}_1^T \mathbf{v}_1 \mathbf{v}_1 \quad (37)$$

where I define:

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

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

Finally, the Taylor rule (13) implies that:

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

It is simple to see that the equation above translates into the following restrictions:

$$\mathbf{c}_0 = (1-\rho_r) [\phi_\pi \mathbf{a}_0 + \phi_y (\mathbf{b}_0 - \mathbf{1}_1^T)] + \mathbf{1}_2^T \quad (38)$$

$$\mathbf{c}_1 = \rho_r \mathbf{1}_3^T + (1-\rho_r) [\phi_\pi \mathbf{a}_1 + \phi_y \mathbf{b}_1] \quad (39)$$

Equations (34)-(39) are a system of non-linear equations in the coefficients $\mathbf{v}_0 \equiv [\mathbf{a}'_0, \mathbf{b}'_0, \mathbf{c}'_0]'$ and $\mathbf{v}_1 \equiv [\mathbf{a}'_1, \mathbf{b}'_1, \mathbf{c}'_1]'$. For any given set of parameter values and matrices \mathbf{M} and \mathbf{N} of coefficients, the solution for this system of equations can be found by using one of the many non-linear equation solvers. This task never turn out to be computationally challenging. I find zeros of a system of non-linear equations through Newton's method, discussed in Judd (1998) (pp. 167-8).²⁰

C Transition Equation of High-Order Expectations

In this section, we show how to derive the law of motion of the average higher-order expectations of the exogenous variables (i.e., $a_t, \varepsilon_{r,t}, g_t$) for given parameter values and vectors of coefficients \mathbf{v}_0 .²¹ We focus on equilibria where HOEs evolve:

$$\varphi_{t|t}^{(0:k)} = \mathbf{M} \varphi_{t-1|t-1}^{(0:k)} + \mathbf{N} \boldsymbol{\varepsilon}_t \quad (40)$$

where $\boldsymbol{\varepsilon}_t \equiv [\varepsilon_{a,t} \quad \eta_{r,t} \quad \varepsilon_{g,t}]'$

A quick inspection of equation (40) reveals that such a law of motion is entirely determined by the matrices \mathbf{M} and \mathbf{N} . Since the model is linear and all shocks are Gaussian, one can pin down these matrices (i.e., solve firms' signal-extraction problem) through the Kalman filter, which requires the specification of firms' state-space model. Denote $\mathbf{X}_t \equiv \varphi_{t|t}^{(0:k)}$, for notational convenience. Firms'

²⁰I find that non-linear solver fails to find a solution when the parameter ϕ_π is too small.

²¹Recall that the vectors \mathbf{v}_0 and \mathbf{v}_1 have been shown to depend on parameters and the matrices \mathbf{M} and \mathbf{N} (see Appendix B). In this Section, these vectors are treated as given.

reduced-form state-space model can be concisely cast as follows:

$$\mathbf{X}_t = \mathbf{M}\mathbf{X}_{t-1} + \mathbf{N}\varepsilon_t \quad (41)$$

$$\mathbf{Z}_t = \mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t} \quad (42)$$

where

$$\begin{aligned} \mathbf{D}_1 &= \begin{bmatrix} \mathbf{d}'_1 & \mathbf{0}_{1 \times 3(k+1)} & \mathbf{0}_{1 \times 3(k+1)} & \mathbf{c}'_0 \end{bmatrix}' \\ \mathbf{D}_2 &= \begin{bmatrix} \mathbf{0}_{1 \times 3(k+1)} & \mathbf{a}'_0 & \mathbf{b}'_0 & \mathbf{0}_{1 \times 3(k+1)} \end{bmatrix}' \end{aligned}$$

with $\mathbf{d}'_1 = [1, \mathbf{0}_{1 \times 3(k+1)-1}]$ and $\mathbf{Q} = [\sigma_a^j, 0, 0, 0]'$.

Solving firms' signal extraction problem requires applying the Kalman filter. Note that firms' observation equations (42) include lagged state variables. This slightly modified the usual Kalman formula for signal extraction. The correct formula has been provided by Nimark (2010). Here we propose a(n alternative) Bayesian derivation of this formula.

The initial period:

In period 0, we start with a prior distribution for the initial state \mathbf{X}_0 . This prior is of the form $\mathbf{X}_{0|0} \sim \mathcal{N}(\mathbf{X}_{0|0}, \mathbf{P}_{0|0})$.

Forecasting:

At $(t-1)^+$, that is after observing Z_{t-1} , the belief about the state vector has the form $\mathbf{X}_t | \mathbf{Z}^{t-1} \sim \mathcal{N}(\mathbf{X}_{t|t-1}, \mathbf{P}_{t|t-1})$ where

$$\begin{aligned} \mathbf{X}_{t|t-1} &= \mathbf{W}\mathbf{X}_{t-1|t-1} \\ \mathbf{P}_{t|t-1} &= \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{W}' + \mathbf{U}\mathbf{U}' \end{aligned} \quad (43)$$

The density

$$\mathbf{Z}_t | \mathbf{Z}^{t-1} \sim \mathcal{N}(\mathbf{Z}_{t|t-1}, \mathbf{F}_{t|t-1})$$

where

$$\mathbf{Z}_{t|t-1} = \mathbf{D}_1\mathbf{X}_{t|t-1} + \mathbf{D}_2\mathbf{X}_{t-1|t-1}$$

and $\mathbf{F}_{t|t-1} = E[\mathbf{Z}_t\mathbf{Z}'_t | \mathbf{Z}^{t-1}]$ and hence:

$$\mathbf{F}_{t|t-1} = E[(\mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t})(\mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t})' | \mathbf{Z}^{t-1}]$$

and by working the product out, one obtains:

$$\begin{aligned} \mathbf{F}_{t|t-1} &= \mathbf{D}_1\mathbf{P}_{t|t-1}\mathbf{D}'_1 + \mathbf{D}_2\mathbf{P}_{t-1|t-1}\mathbf{D}'_2 + \mathbf{Q}\mathbf{Q}' + \\ &\quad + \mathbf{D}_1E[\mathbf{X}_t\mathbf{X}'_{t-1} | \mathbf{Z}^{t-1}]\mathbf{D}'_2 + \mathbf{D}_2E[\mathbf{X}_{t-1}\mathbf{X}'_t | \mathbf{Z}^{t-1}]\mathbf{D}'_1 \end{aligned} \quad (44)$$

Note that

$$\begin{aligned} E[\mathbf{X}_t\mathbf{X}'_{t-1} | \mathbf{Z}^{t-1}] &= E[(\mathbf{W} \cdot \mathbf{X}_{t-1} + \mathbf{U} \cdot \varepsilon_t) \mathbf{X}'_{t-1} | \mathbf{Z}^{t-1}] \\ &= \mathbf{W}\mathbf{P}_{t-1|t-1} \end{aligned}$$

Combining this result with equation (44) yields:

$$\begin{aligned} \mathbf{F}_{t|t-1} &= \mathbf{D}_1 \mathbf{P}_{t|t-1} \mathbf{D}'_1 + \mathbf{D}_2 \mathbf{P}_{t-1|t-1} \mathbf{D}'_2 + \mathbf{Q} \mathbf{Q}' + \\ &\quad + \mathbf{D}_1 \mathbf{W} \mathbf{P}_{t-1|t-1} \mathbf{D}'_2 + \mathbf{D}_2 \mathbf{P}_{t-1|t-1} \mathbf{W}' \mathbf{D}'_1 \end{aligned}$$

by substituting equation (43) into the equation above leads to

$$\begin{aligned} \mathbf{F}_{t|t-1} &= \mathbf{D}_1 \mathbf{W} \mathbf{P}_{t-1|t-1} \mathbf{W}' \mathbf{D}'_1 + \mathbf{D}_2 \mathbf{P}_{t-1|t-1} \mathbf{D}'_2 + \mathbf{Q} \mathbf{Q}' + \\ &\quad + \mathbf{D}_1 \mathbf{W} \mathbf{P}_{t-1|t-1} \mathbf{D}'_2 + \mathbf{D}_2 \mathbf{P}_{t-1|t-1} \mathbf{W}' \mathbf{D}'_1 + \mathbf{D}_1 \mathbf{U} \mathbf{U}' \mathbf{D}'_1 \end{aligned}$$

and finally to:

$$\mathbf{F}_{t|t-1} = (\mathbf{D}_1 \mathbf{W} + \mathbf{D}_2) \mathbf{P}_{t-1|t-1} (\mathbf{D}_1 \mathbf{W} + \mathbf{D}_2)' + \mathbf{Q} \mathbf{Q}' + \mathbf{D}_1 \mathbf{U} \mathbf{U}' \mathbf{D}'_1$$

The joint distribution of \mathbf{X}_t and \mathbf{Z}_t , thus, is

$$\begin{bmatrix} \mathbf{X}_t \\ \mathbf{Z}_t \end{bmatrix} | \mathbf{Z}^{t-1} \sim \mathcal{N} \left(\begin{bmatrix} \mathbf{X}_{t|t-1} \\ \mathbf{Z}_{t|t-1} \end{bmatrix}, \begin{bmatrix} \mathbf{P}_{t|t-1} & \mathbf{P}_{t|t-1} \mathbf{D}'_1 + \mathbf{W} \mathbf{P}_{t-1|t-1} \mathbf{D}'_2 \\ \mathbf{D}_1 \mathbf{P}'_{t|t-1} + \mathbf{D}_2 \mathbf{P}_{t-1|t-1} \mathbf{W}' & \mathbf{F}_{t|t-1} \end{bmatrix} \right)$$

as

$$\begin{aligned} E [\mathbf{X}_t \mathbf{Z}'_t | \mathbf{Z}^{t-1}] &= \mathbf{P}_{t|t-1} \mathbf{D}'_1 + E (\mathbf{X}_t \mathbf{X}'_{t-1} | \mathbf{Z}^{t-1}) \mathbf{D}'_2 \\ &= \mathbf{P}_{t|t-1} \mathbf{D}'_1 + E [(\mathbf{W} \cdot \mathbf{X}_{t-1} + \mathbf{U} \cdot \varepsilon_t) \mathbf{X}'_{t-1} | \mathbf{Z}^{t-1}] \mathbf{D}'_2 \\ &= \mathbf{P}_{t|t-1} \mathbf{D}'_1 + \mathbf{W} \mathbf{P}_{t-1|t-1} \mathbf{D}'_2 \end{aligned}$$

where the second equality follows from using equation (41). Hence, it follows that

$$\mathbf{X}_t | (\mathbf{Z}_t, \mathbf{Z}^{t-1}) = \mathbf{X}_t | \mathbf{Z}^t \sim \mathcal{N} (\mathbf{X}_{t|t}, \mathbf{P}_{t|t})$$

where²²

$$\mathbf{X}_{t|t}(j) = \mathbf{X}_{t|t-1}(j) + [\mathbf{P}_{t|t-1} \mathbf{D}'_1 + \mathbf{W} \mathbf{P}_{t-1|t-1} \mathbf{D}'_2] \mathbf{F}_{t|t-1}^{-1} [\mathbf{Z}_t - \mathbf{Z}_{t|t-1}] \quad (45)$$

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - [\mathbf{P}_{t|t-1} \mathbf{D}'_1 + \mathbf{W} \mathbf{P}_{t-1|t-1} \mathbf{D}'_2] \mathbf{F}_{t|t-1}^{-1} [\mathbf{D}_1 \mathbf{P}'_{t|t-1} + \mathbf{D}_2 \mathbf{P}_{t-1|t-1} \mathbf{W}'] \quad (46)$$

Therefore, combining equation (46) with equation (43) yields:

$$\mathbf{P}_{t+1|t} = \mathbf{W} \left[\mathbf{P}_{t|t-1} - (\mathbf{P}_{t|t-1} \mathbf{D}'_1 + \mathbf{W} \mathbf{P}_{t-1|t-1} \mathbf{D}'_2) \mathbf{F}_{t|t-1}^{-1} (\mathbf{D}_1 \mathbf{P}'_{t|t-1} + \mathbf{D}_2 \mathbf{P}_{t-1|t-1} \mathbf{W}') \right] \mathbf{W}' + \mathbf{U} \mathbf{U}' \quad (47)$$

²²Here we use the following lemma to get the moments in the text. Let the random vector $(x', y')' \sim \mathcal{N}(\mu, \Sigma)$ such that we denote

$$\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

Then the pdf $x|y \sim \mathcal{N}(\mu_{x|y}, \Sigma_{xx|y})$ with

$$\begin{aligned} \mu_{x|y} &= \mu_x + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \mu_y) \\ \Sigma_{xx|y} &= \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx} \end{aligned}$$

Denote the Kalman-gain matrix as $\mathbf{K}_t \equiv [\mathbf{P}_{t|t-1}\mathbf{D}'_1 + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}'_2] \mathbf{F}_{t|t-1}^{-1}$. Recall equation (42) and write the law of motion of the firm j 's first-order beliefs about \mathbf{X}_t as

$$\mathbf{X}_{t|t}(j) = \mathbf{X}_{t|t-1}(j) + \mathbf{K}_t [\mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t} - (\mathbf{D}_1\mathbf{X}_{t|t-1}(j) + \mathbf{D}_2\mathbf{X}_{t-1|t-1}(j))]$$

where we have combined equations (45) and (42). By recalling that $\mathbf{X}_{t|t-1}(j) = \mathbf{W}\mathbf{X}_{t-1|t-1}(j)$, we have:

$$\mathbf{X}_{t|t}(j) = \mathbf{W}\mathbf{X}_{t-1|t-1}(j) + \mathbf{K}_t [\mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t} - (\mathbf{D}_1\mathbf{W}\mathbf{X}_{t-1|t-1}(j) + \mathbf{D}_2\mathbf{X}_{t-1|t-1}(j))]$$

By rearranging one obtains:

$$\mathbf{X}_{t|t}(j) = (\mathbf{W} - \mathbf{K}\mathbf{D}_1\mathbf{W} - \mathbf{K}\mathbf{D}_2)\mathbf{X}_{t-1|t-1}(j) + \mathbf{K}[(\mathbf{D}_1\mathbf{W} + \mathbf{D}_2) \cdot \mathbf{X}_{t-1} + \mathbf{D}_1\mathbf{U} \cdot \boldsymbol{\varepsilon}_t + \mathbf{Q}e_{j,t}] \quad (48)$$

The vector $\mathbf{X}_{t|t}(j)$ contains firm j 's first-order expectations about model's state variables. Integrating across firms yields the law of motion of the average expectation about $\mathbf{X}_{t|t}^{(1)}$:

$$\mathbf{X}_{t|t}^{(1)} = (\mathbf{W} - \mathbf{K}\mathbf{D}_1\mathbf{W} - \mathbf{K}\mathbf{D}_2)\mathbf{X}_{t-1|t-1}^{(1)} + \mathbf{K}[(\mathbf{D}_1\mathbf{W} + \mathbf{D}_2) \cdot \mathbf{X}_{t-1} + \mathbf{D}_1\mathbf{U} \cdot \boldsymbol{\varepsilon}_t]$$

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

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

So by using the assumption of common knowledge in rationality, we can fully characterize the matrices \mathbf{M} and \mathbf{N} :

$$\begin{aligned} \mathbf{M} &= \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3k} \\ \mathbf{0}_{3k \times 3} & (\mathbf{W} - \mathbf{K}\mathbf{D}_1\mathbf{W} - \mathbf{K}\mathbf{D}_2)|_{(1:3k, 1:3k)} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{K}(\mathbf{D}_1\mathbf{W} + \mathbf{D}_2)|_{(1:3k, 1:3(k+1))} \end{bmatrix} \quad (49) \\ \mathbf{N} &= \begin{bmatrix} \mathbf{R}_2 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{K}\mathbf{D}_1\mathbf{U}|_{(1:3k, 1:3)} \end{bmatrix} \quad (50) \end{aligned}$$

where $\cdot|_{(n_1:n_2, m_1:m_2)}$ denotes the submatrix obtained by taking the elements lying between the n_1 -th row and the n_2 -th row and between the m_1 -th column and the m_2 -th column. Note that \mathbf{K} in the above equation denotes the steady-state Kalman gain matrix, which is obtained by iterating the equations (46)-(47) and the equation for the Kalman-gain matrix below:

$$\mathbf{K} = [\mathbf{P}_{t|t-1}\mathbf{D}'_1 + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}'_2] \mathbf{F}_{t|t-1}^{-1}$$

until convergence.

D Cross-Sectional Dispersion of Expectations

The laws of motion of firm's expectation is given by equation (48). The cross-section variance of the (truncated) vector of higher-order beliefs $(\mathbf{X}_{t|t}(j) = \varphi_{t|t}^{(1:k+1)})$ can be computed by solving the

Lyapunov equation below:

$$\mathbb{E} (\mathbf{X}_{t|t} (j) \mathbf{X}_{t|t} (j)') = \mathbf{A} \mathbb{E} (\mathbf{X}_{t|t} (j) \mathbf{X}_{t|t} (j)') \mathbf{A}' + \mathbf{K} \mathbf{Q} \mathbf{Q}' \mathbf{K}'$$

where $\mathbf{A} \equiv (\mathbf{W} - \mathbf{K} \mathbf{D}_1 \mathbf{W} - \mathbf{K} \mathbf{D}_2)$.

Also recall that the endogenous variables \mathbf{s}_t evolves according to equation (14) and hence

$$\mathbf{s}_{t|t} (j) = \mathbf{v}_0 \mathbf{X}_{t|t} (j) + \mathbf{v}_1 \mathbf{s}_{t-1}$$

The cross-sectional variance of the endogenous variables (i.e., output, inflation, and the interest rate) is given by

$$\mathbb{E} [\mathbf{s}_{t|t} (j) \mathbf{s}_{t|t} (j)'] = \mathbf{v}_0 \mathbb{E} [\mathbf{X}_{t|t} (j) \mathbf{X}_{t|t} (j)'] \mathbf{v}_0'$$