# Public's Inflation Expectations and Monetary Policy

VERY PRELIMINARY

Leonardo Melosi†

London Business School

14 May, 2011

**Abstract**

The paper develops a DSGE model to study the propagation of monetary impulses through their influence on inflation expectations. Price setters do not observe the history of the shocks hitting the economy. Since price setters face costs of price adjustment, they have to form expectations about the evolution of their nominal marginal costs. To forecast their costs, they rely on 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. The model is estimated through likelihood methods on a U.S. data set including the Survey of Professional Forecasters as a measure of price setters' expectations. The paper performs three formal econometric tests on the model, leading to the following results. First, the channel of monetary transmission based on price-setters' expectations is found to be strongly supported by the data. Second, the model is found to fit very well the observed inflation expectations. Third, information frictions are found to substantially improve the fit of the model.

**Keywords:** Coordinating public's expectations, price puzzle, higher order beliefs, heterogenous 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 Harald Cole, Matthieu Darracq Pariès, Jesus Fernandez-Villaverde, Dirk Krueger, Kristoffer Nimark, Luigi Paciello, Ricardo Reis, Massimo Rostagno, 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 for their helpful comments. The hospitality of the European Central Bank during part of this project is gratefully acknowledged.

# 1 Introduction

That monetary policy influences output and inflation by affecting public' expectations has come to a growing consensus among scholars and policy makers[1] in the last twenty years. Woodford (2005) writes:

> *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.*

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 stylized models where the central bank can coordinate dispersed agents' expectations in self-fulfilling environments (e.g., Morris and Shin, 2003 and Hellwig, 2002). On one hand, these models are too stylized to conduct a formal econometric investigation about the empirical significance of this new propagation channel for monetary policy. On the other hand, the workhorse models (e.g., Christiano, Eichenbaum, and Evans 2005 and Smets and Wouters, 2007 ) for evaluating 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 such expectations by setting the interest rate or making announcements.

The paper bridges the gap in the literature by constructing and estimating a Dynamic Stochastic General Equilibrium (DSGE) model in which monetary policy can influence the macroeconomic aggregates (e.g., GDP and inflation) by affecting and coordinating dispersed agents' expectations. The model features technology, monetary, and preference shocks. Price setters observe their idiosyncratic technology shocks (i.e., a *private* signal), which conveys information about the aggregate productivity. Furthermore, the price-setters observe last period's real output and inflation as well as the interest rate, which is set by the central bank

[1] See, for instance, Bernanke (2004) and King (2005), and **?**.

according to a Taylor rule (i.e., a *public* signal). This signal provides information about the current inflation and output to the price setters. Price setters face cost of price adjustment in the form of Calvo sticky prices (Calvo, 1983). Hence, 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, hence, face higher-order uncertainty as the optimal price will depend not only on their beliefs about the future price levels, but also on 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 episode of substantial rise in inflation and inflation expectations in recent US economic history.

The main contribution of the paper is to perform an econometric evaluation of the channel of monetary transmission based on price-setters' expectations. The main findings are the following. First, the paper finds strong empirical support in favor of the channel of monetary transmission based on price-setters' expectations. Second, the presence of the channel helps the model fitting the *Survey of Professional Forecasters*, which are used as a measure of price setters' expectations in the model. Third, the information frictions substantially improve the fit of the model. This last result is not in line with what found by Del Negro and Eusepi (2010), who estimate an imperfect information model à la Erceg and Levin (2003) to a data set including the *Survey of Professional Forecasters* as a measure of agents' beliefs. Two main reasons explain why that paper obtain findings which differ from those discussed in this paper. (*i*) The learning mechanism in this paper is 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 every shock that have hit the economy. Furthermore, in our model price setters have dispersed higher-order expectations as they observe private signals (i.e., firm-specific technology shocks). (*ii*), The data range used in this paper differs from the one in Del Negro and Eusepi (2010). They use a data set starting from the early 1980s, while our paper also considers the 1970s, which was characterized by high inflation.

Furthermore, the transmission channel for monetary policy based on affecting price setters' expectations is found to be empirically relevant, accounting 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

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

price level after a contractionary monetary policy shock)[3], (*iii*) the disappearance of the price puzzle after the 1970s.[4]

To see how the model provides an explanation for 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, 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. In such a scenario, fewer price setters will confuse a monetary shocks with a technology shock or a preference shock.

The estimated model predicts that, in the 1970s, 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 the 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, the fact (*ii*). When the paper estimates the model on a subsample that does not include data on 1970s, the price puzzle disappears, that is, the fact (*iii*). The last result is explained by the well-known switch to a monetary policy regime in which the Federal Reserve became more aggressive against inflation.

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

[4]See Barth and Ramey (2001), Hansen (2004), Castelnuovo and Surico (2010). and Ravn, Schmitt-Grohe, Uribe, and Uusküla (2010).

## 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.

This paper is closely related to Nimark (2008), who introduces a model where firms hold private information about the dynamics of their future marginal costs, and face both strategic complementarities in price setting and nominal rigidities. The nice feature of this model is that the supply side of this economy can be analytically worked out and turns out to be characterized by an equation that resembles the standard New-Keynesian Phillips curve. Nonetheless, unlike the present paper the role of monetary policy in coordinating agents' expectations is absent in that central bank's actions only convey 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, but rather on how public beliefs adjust to monetary policy shocks and how such adjustment affects the propagation of these shocks.

Lorenzoni (2010), who 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 in this paper 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).[5]

[5] See Mankiw and Reis (2002a, 2002b, 2006, 2007), and Reis (2006a, 2006b, 2009) for models with infor-

# 3 The 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. The technology shocks have a persistent and white noise component. All these shocks are orthogonal to the others at all leads and lags. Idiosyncratic shocks include the firm-specific technology, $A_{j,t}$, that determines the level of technology of 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. Both of the idiosyncratic shocks are orthogonal to each other at all leads and lags.

In section 3.1, I present the time protocol of the model. Section 3.2 presents the problem of the producers. Section 3.4 presents the price-setting problem of the intermediate-goods firms. Section 3.3 presents the problem of households. In Section 3.5, Central bank's behavior is modeled. Section 3.6 deals with the log-linearization of the model equations. Section 3.7 discusses firm's signal-extraction problem within this log-lineariz(ed) set-up. Section 3.8sheds light on how the transmission channel based on price setters' expectation works. The case of perfect information is discussed in Section 3.9.

## 3.1 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 the realization of the idiosyncratic shocks (i.e., the firm-specific technology shock and the outcome of the

mation frictions that do not feature imperfect common knowledge but can generate sizeable persistence.

Calvo lottery) and set their prices.[6] 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. Hence, intermediate-goods market, final-goods market, money market, bond market, and labor market clear.

## 3.2 Final-Goods Producers

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

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

where $Y_t$ is the amount of the final good produced at time $t$, the parameter $\nu$ 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:

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

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

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

## 3.3 Households

Households solve:

$$\max_{C_{i,t+s}, B_{i,t+s}, M_{i,t+s}, N_{i,t+s}} \mathbb{E}_t \sum_{s=0}^{\infty} \beta^{t+s} g_{t+s} \left[ \ln C_{i,t+s} + \frac{\chi_m}{1-\gamma_m} \left( \frac{M_{i,t+s}}{P_{t+s}} \right)^{1-\gamma_m} - \chi_n N_{i,t+s} \right]$$

$st$

$$P_t C_{i,t} + B_{i,t} + M_{i,t} = W_t N_{i,t} + R_{t-1} B_{i,t-1} + M_{i,t-1} + \Pi_{i,t} + T_t$$

[6] Those firms that are not allowed to re-optimize their prices reset them according to the steady-state inflation rate.

where$\beta$ 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, $\Pi_{i,t}$ are the 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 $\varepsilon_{g,t} \backsim \mathcal{N}(0,1)$. We will denote $\hat{g}_t = \ln g_t$. Since the dividends from firms, $\Pi_{i,t}$, are assumed to be equally shared among households, households turn out to face an identical problem. Thus, we can derive households' policy function as they were chosen by a representative agent.

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:

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

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

## 3.4 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. The firms commit themselves to satisfy any demanded quantity of their intermediate good that will arise in the stage 2 at the price they have set in the stage 1. Consider an arbitrary firm $j$. The real marginal costs for firm $j$ are given by:

$$mc_{j,t} = \frac{W_t}{A_{j,t} P_t} \tag{4}$$

where $A_{j,t}$ is the technology shock that can be decomposed into (1) a trend component $A_0$, (2) a persistent aggregate component, $z_t$, (3) a white-noise aggregate component, $\eta_{a,t}$,and (4) a white-noise idiosyncratic component, $\eta^a_{j,t}$. More specifically, we have:

$$A_{j,t} = A_t e^{\eta^a_{j,t}} \tag{5}$$

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

$$A_t = A_0^t e^{z_t + \sigma_a \eta_{a,t}} \tag{6}$$

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

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

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

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

$$Y_{j,t} = A_{j,t} N_{j,t} \tag{7}$$

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 $\mathbb{I}^{Calvo}_{j,t}$. Firm's information set $\mathcal{I}_{j,t}$ is defined below:

$$\mathcal{I}_{j,t} \equiv \{R_\tau, \pi_{\tau-1}, Y_{\tau-1}, W_{\tau-1}, \boldsymbol{\theta}_{j,\tau}, \Theta : \tau \leq t\} \tag{8}$$

where $\boldsymbol{\theta}_{j,t} \equiv \left\{A_{j,t}, P^*_{j,t}, Y_{j,t-1}, MC_{j,t-1}, N_{j,t-1}, \Pi_{j,t-1}, \mathbb{I}^{Calvo}_{j,t}\right\}$. 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:

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

subject to the firm's specific demand in equation (2), to the production function (7).

## 3.5 The Central Bank

The central bank sets the interest rate $R_t$ following a Taylor rule of type[7]:

$$R_t = R^{\rho_r}_{t-1} R^{*(1-\rho_r)}_t e^{\sigma_r \eta_{r,t}} \tag{9}$$

[7] We follow Aruoba and Schorfheide (2011) in specifying the monetary-policy reaction function.

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

$$R_t^* = (r_* \pi_*) \left( \frac{\pi_t}{\pi_*} \right)^{\phi_\pi} \left( \frac{Y_t / Y_{t-1}}{A_0} \right)^{\phi_y} \tag{10}$$

where $A_0$ denotes the average growth rate of real output, $\pi_t$ the inflation rate $\ln(P_t/P_{t-1})$, and $\pi^*$ stands for the inflation target, which is assumed to be constant over time.

The central bank observes the contemporaneous realizations of the aggregate shocks and sets the interest rate $R_t$. The central bank cannot simply tell firms the contemporaneous realizations of the state variables (i.e., shocks, real output, and inflation) 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 $A_0$.[8] 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 not a commitment device that would back up central bank's words, then any central bank's statements about the real output, inflation, and shocks are not deemed as credible by price setters. If there exists an exogenous commitment device that allows the central bank to commit herself to her statements, then the model would boil down to a canonical (perfect-information) three-equations New Keynesian model (e.g., Rabanal and Rubio-Ramirez, 2005). We will consider this perfect-information case as a benchmark in Section 3.9.

## 3.6 Detrending and Log-linearization

Define the following stationary variables:

$$\begin{aligned} y_t &= \frac{Y_t}{A_0^t}, \; c_t = \frac{C_t}{A_0^t}; \; p_{j,t}^* = \frac{P_{j,t}^*}{P_t}, \; y_{j,t} = \frac{Y_{j,t}}{A_0^t} \\ w_t &= \frac{W_t}{A_0^t P_t}, \; a_t = \frac{A_t}{A_0^t}, R_t = \frac{R_t}{R_*}, \; mc_{j,t} = \frac{MC_{j,t}}{P_t} \\ \xi_{j,t} &= A_0^t \Xi_{j,t} \end{aligned}$$

First, I solve firms and households' problems that are described in Sections 3.4 and 3.3 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 the perfect-information-deterministic steady-state equilibrium.

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

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

$$\hat{\pi}_t = (1-\theta)(1-\beta\theta)\sum_{k=0}^{\infty}(1-\theta)^k \widehat{mc}_{t|t}^{(k)} + \beta\theta\sum_{k=0}^{\infty}(1-\theta)^k \widehat{\pi}_{t+1|t}^{(k+1)} \tag{11}$$

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 $\widehat{\pi}_{t+1|t}^{(k)} \equiv \underbrace{\int \mathbb{E}_{j,t} \ldots \int \mathbb{E}_{j,t}}_{k} \widehat{\pi}_{t+1}$ 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

$$\widehat{mc}_{t|t}^{(k)} = \hat{y}_{t|t}^{(k+1)} - z_{t|t}^{(k)} - \sigma_a \eta_{a,t|t}^{(k)} \tag{12}$$

It is easy to show that the log-linearized Euler equation is given by

$$\widehat{g}_t - \widehat{y}_t = \mathbb{E}_t \widehat{g}_{t+1} - \mathbb{E}_t \widehat{y}_{t+1} - \mathbb{E}_t \hat{\pi}_{t+1} + \hat{R}_t \tag{13}$$

The Taylor rule can be easily linearized:

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1-\rho_r)\left[\phi_\pi \hat{\pi}_t + \phi_y (\hat{y}_t - \hat{y}_{t-1})\right] + \sigma_r \eta_{r,t} \tag{14}$$

The parameter set of the model is given by the vector

$$\Theta = \left[\theta, \phi_\pi, \phi_y, \rho_r, \rho_z, \rho_g, \sigma_a, \sigma_z, \sigma_a^j, \sigma_a^j, \sigma_r, \sigma_g\right]'$$

## 3.7 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 needs to form expectations on the price level to estimate nominal costs. Hence, they also need to form expectations 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, $\varphi_{t|t}^{(0:k)} \equiv \left[z_{t|t}^{(s)}, \eta_{a,t|t}^{(s)}, \eta_{r,t|t}^{(s)}, \eta_{g,t|t}^{(s)} : 0 \leq s \leq k\right]'$ follow a VAR(1) process:[9] $\varphi_{t|t}^{(0:k)} = \mathbf{M}\varphi_{t-1|t-1}^{(0:k)} + \mathbf{N}\boldsymbol{\varepsilon}_t$, where $\boldsymbol{\varepsilon}_t \equiv \begin{bmatrix} \varepsilon_{z,t} & \eta_{a,t} & \varepsilon_{\pi,t} & \eta_{r,t} & \varepsilon_{g,t} & \eta_{g,t} \end{bmatrix}'$. 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 (13), the Phillips curve (11), and the Taylor Rule (14). 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 $\mathbf{s}_t \equiv \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t\right]'$. Given the matrices $\mathbf{M}$ and $\mathbf{N}$, the structural equations (12)-(14) can be written as $[\mathbf{v}_0, \mathbf{v}_1]' = f\left(\mathbf{M}, \mathbf{N}, \mathbf{v_0}, \mathbf{v_1}; \Theta\right)$. This is a system of non-linear equations in the unknown vectors $\mathbf{v}_0, \mathbf{v}_1$.

For given parameters $\Theta$, take the following steps:

- Set $i = 1$ and guess the matrices $\mathbf{M}^{(i)}$ $\mathbf{N}^{(i)}$
- Conditional on $\mathbf{M}^{(i)}$ $\mathbf{N}^{(i)}$, solve the system of nonlinear equations:

  $$[\mathbf{v}_0, \mathbf{v}_1]' = f\left(\mathbf{M}^{(i)}\ \mathbf{N}^{(i)}, \mathbf{v}_\mathbf{0}^{(\mathbf{i})}, \mathbf{v}_\mathbf{1}^{(\mathbf{i})}; \Theta\right)$$

  where $\mathbf{s}_t = \mathbf{v}_0\varphi_{t|t}^{(0:k)} + \mathbf{v}_1\mathbf{s}_{t-1}$and $\mathbf{s}_t = \left[\widehat{\pi}_t, \widehat{y}_t, \widehat{R}_t\right]'$. See Appendix B.2.
- Use the Kalman equation to work out the mapping $\left(\mathbf{M}^{(i+1)}, \mathbf{N}^{(i+1)}\right) = g\left(\mathbf{M}^{(i)}, \mathbf{N}^{(i)}, \mathbf{v}_0^{(i)}, \mathbf{v}_1^{(i)}\right)$. See Appendix C
- Check that $\left\|\mathbf{M}^{(i)} - \mathbf{M}^{(i+1)}\right\| < \varepsilon_m$ and $\left\|\mathbf{N}^{(i)} - \mathbf{N}^{(i+1)}\right\| < \varepsilon_n$ with $\varepsilon_m > 0$ and $\varepsilon_n > 0$ small.

## 3.8 The Channel Based on Affecting Firms' Expectations

The central bank can affect price setters' expectations by setting its policy rate is a salient feature of the model developed in the paper. This new transmission channel arises for the following three reasons. First, price setters have to form expectations about the evolution of their nominal marginal costs in order to set their prices because they face cost of price

[9] We truncate the state vector of the higher-order expectations at $k < \infty$. When we estimate the model, we set $k = 10$.

adjustment. Second, price setters perfectly observe the current policy rate and this is common knowledge among them. Third, the central bank follows a monetary policy rule with feedbacks.

Price setters use the policy rate as a signal both to figure out the current and future level[10] of the endogenous variables (i.e., inflation and real output), which 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 $\eta_{r,t}$). It is illustrative to re-write the Taylor rule (14) as follows:

$$\widetilde{R}_t = \underbrace{(1-\rho_r)\,\phi_\pi\hat{\pi}_t + (1-\rho_r)\,\phi_y\hat{y}_t}_{\textbf{Endogenous component of the signal}} + \underbrace{\sigma_r\eta_{r,t}}_{\textbf{Exogenous component of the signal}} \tag{15}$$

where $\widetilde{R}_t \equiv \hat{R}_t - \rho_r\hat{R}_{t-1} + (1-\rho_r)\,\phi_y\hat{y}_{t-1}$ 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 affected only 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 the non-systematic deviations from the rule, $\eta_{r,t}$. These two interpretations have a drastically different impact on price setters' forecast 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 estimate rising nominal wages and growing 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 due to their initial misinterpretation.

The relative volatility of the endogenous component of the policy signal increases when the following occurs: (a) the central bank adopts a very proactive policy to either inflation or output (i.e., the values of the parameters $\phi_\pi$ or $\phi_y$ go up); (b) the size of exogenous deviations from the rule are smaller on average (i.e., the variance parameters, $\sigma_r$ is smaller);

[10] They care about future level of such a variables because they face cost of price adjustement.

(c) a change in non-policy parameters raising the unconditional variance of real output and inflation (e.g., the variance of technology or preference shocks gets larger).

Note that the direction of the effects (a) and (b) upon the importance of the endogenous component of the policy signal is ambiguous and varies with the values of other deep parameters. In fact, rising the degree of proactiveness of monetary policy or reducing the variability of monetary shocks will produce two opposite effects on the relative importance of the endogenous component. On one hand, they will increase the importance of the endogenous component by scaling up its variability or decreasing the variability of the exogenous component. On the other hand, very large values of these policy parameters will end up reducing the unconditional variance of real output or inflation.

It is finally important to notice that larger value for the coefficient on inflation, $\phi_\pi$ relative to that on output growth $\phi_y$ changes the information about the nature of the technology shocks that the policy rate conveys to price setters. Relatively large value for the parameter $\phi_\pi$ implies that the central bank will raise its policy rate after a negative technology shock as it cares mostly about inflation stability. When $\phi_\pi$ 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. On the contrary, when $\phi_\pi$ is relatively small, then a rise of the interest rate tends to coordinate average expectations towards a positive technology shock that reduces inflation. 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 form 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. Effects on inflation tend to be amplified by the channel based on price setters' expectations, if price setters interpret a rise (fall) in the interest rate as coming from the response of the central bank to a positive (negative) technology shock.

### 3.9 The Case of Perfect Information

If the central bank could commit to her announcements, then she could reveal the state vector $\varphi_{t|t}^{(0:\infty)}$ to the firms. If such a commitment device existed, then price-setting firms would perfectly informed and the model would boil down into a canonical three-equation new-Keynesian model with Calvo sticky-prices (e.g., Rabanal and Rubio-Ramirez, 2005). More specifically, the Phillips curve equation (11) would become $\hat{\pi}_t = \kappa_{pc}\widehat{mc}_t + \beta\mathbb{E}_t\widehat{\pi}_{t+1}$, where $\kappa_{pc} \equiv (1-\theta)(1-\theta\beta)/\theta$ and the real marginal costs $\widehat{mc}_t = \hat{y}_t - z_t - \sigma_a \eta_{a,t}$. The IS equation and the Taylor rule would be the same as in the incomplete information model. See equation (13) and (14). In this perfect information model the monetary shock propagates by affecting the intertemporal allocation of consumption. Real effects of money emerge as a result of price-stickiness as opposed to the sluggish adjustments of firms' expectations. In this model monetary shocks propagate through the channel based on the real interest rate, which is very typical to new Keynesian model. See Section 3.9. We call this standard New-Keynesian model as perfect information model or model with perfect commitment.

# 4 Empirical Analysis

This section contains the quantitative analysis of the model. I combine a prior distribution for the parameter set $\Theta$ with the likelihood function derived from the model and conduct Bayesian estimation of the parameters.

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 and the variance decomposition. In Section 4.4, we conduct an econometric evaluation of the channel for monetary transmission based on price-setters' expectations. Section 4.5 studies the impulse response functions to monetary policy shocks.

## 4.1 Econometrician's State-Space model

Denote the matrices $\mathbf{W} \equiv \mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{U} \equiv \mathbf{A}^{-1}\mathbf{C}$, where the characterization of the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ has been discussed in Section C. The transition equations to be used for evaluating the likelihood function implied by the model are:

$$\begin{bmatrix} \varphi_{t|t}^{(0:k)} \\ \mathbf{s}_t \end{bmatrix} = \mathbf{W} \cdot \begin{bmatrix} \varphi_{t-1|t-1}^{(0:k)} \\ \mathbf{s}_{t-1} \end{bmatrix} + \mathbf{U} \cdot \boldsymbol{\varepsilon}_t$$

where $\varphi_t \equiv \left[z_{t|t}, \eta_{a,t|t}, \eta_{r,t|t}, \eta_{g,t|t}\right]'$ and $\mathbf{s}_t \equiv \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t\right]'$.

Two data sets are used for estimation. Both data sets include 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 first data set $\mathcal{S}_1$ ranges from 1970:3 to 2008:4 while the second one $\mathcal{S}_2$ spans a period where the Federal Reserve has responded more aggressively towards inflation, that is 1983:2-2008:4. The paper wants to study how public's inflation expectations have reacted to unanticipated monetary shocks under different central bank's responsiveness to inflation rate. It has been widely documented that the Federal Reserve has been much more aggressive in fighting inflation during our second sample $\mathcal{S}_2$ than during the 1970s. Given that only few observations for 1970s are available, one cannot reliably estimate the model to a data set that ranges from 1970:3 to 1983:2. Given the larger variability of the observable variables during the 1970s, the posterior for model parameters when the full sample $\mathcal{S}_1$ is used, is likely to be highly affected the observations in the 1970s. The measurement equations are[11]:

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

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

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

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

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

where $PGDP2_t$, $PGDP3_t$, $PGDP6_t$ are the *Survey of Professional Forecasters*[12] about the current, one-quarter-ahead, and four-quarter ahead GDP price index. We relate these statistics with the first moment of the distribution of firms' expectations implied by the model. $\varepsilon_t^{m_1}$ and $\varepsilon_t^{m_2}$ are two i.i.d. measurement errors, such that $\varepsilon_t^{m_1} \backsim \mathcal{N}(0,1)$ and $\varepsilon_t^{m_2} \backsim \mathcal{N}(0,1)$. 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 $100 \ln A_0$ and $100 \ln \pi_*$. We centered the priors for the first two parameters at the value of these averages. Note that the discount factor $\beta$ depends on the linear trend of real output $A_0$ and the steady-state real interest rate $R_*/\pi_*$. Hence, I fix the value for this parameter so as the steady-state nominal interest rate $\ln R_*$ matches the low-frequency behavior of

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

[12] Philly Fed Docs on SPFs to be cited here.

$100 \cdot FEDRATE_t$ in the sample.

The prior median for the standard deviation of the idiosyncratic technology shocks, $\sigma_a^j$, is set so that the model, calibrated at the prior medians, matches the median absolute size of the U.S. finished goods producer price changes reported by Nakamura and Steinsson (2008). The priors for the standard deviations of the technology shocks, $\sigma_z$ and $\sigma_a$, reflect the belief that most of the variance of the technology shocks is explained by the persistent component. This belief is consistent with what is the standard practice in the real business cycle literature (e.g., Kydland and Prescott, 1982).

The volatility of the monetary policy shock, $\sigma_r$, is very crucial as it affects not only the variability of the non-systematic component of monetary policy but also how much informative the interest rate $R_t$ is regarding the current level of inflation and output growth. The prior for this parameter is centered at a smaller value than that for the volatility of the technology and preference shocks in order to reflect the belief that the interest rate is informative for firms. Priors for the remaining standard deviation of the i.i.d. shocks (i.e., $\sigma_g$, $\sigma_{m1}$, $\sigma_{m2}$) 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 priors for the autoregressive parameters $\rho_z$ and $\rho_g$ 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, $\phi_\pi$, response to economic activity, $\phi_y$, autoregressive parameter, $\rho_r$, and the standard deviation of the i.i.d. monetary shock, $\sigma_r$) are chosen as follows. The priors for $\phi_\pi$ and $\phi_y$ are centered at 2.00 and .25, respectively, and imply a fairly strong response to inflation and a moderate response to output growth. The prior for the autoregressive parameter, $\rho_r$ is centered to 0.8, conjecturing that past monetary policy decisions have fairly persistent effects on current central bank's decision over the interest rate. The prior specification for the standard deviation of monetary shocks, $\sigma_r$, have been discussed above.

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

## 4.3 Posteriors

A closed-form expression for the posterior distribution is not available, but we can approximate the moments of the posterior distributions via the Metropolis-Hastings. $40,000$ posterior draws are obtained from the posterior. The posterior moments of parameters are reported in Table 3 when estimation is performed on the two samples $\mathcal{S}_1$and $\mathcal{S}_2$. Let me first describe the results for the full sample $\mathcal{S}_1$. Bayesian updating raises the volatility of the idiosyncratic component of technology shocks relative to that of the persistent component. The posterior median for the variance of the white-noise component of the technology shocks $\sigma_a$ is larger than that for the idiosyncratic technology shock $\sigma_a^j$. These numbers imply a quite large signal-to-noise ratio relative to white-noise technology shocks $\eta_{a,t}$, suggesting that firms learn fairly quickly about these shocks. Nonetheless, firms find it harder to learn about the persistent component of technology $z_t$ because the posterior median of $\sigma_z$ is smaller relative to the posterior median of $\sigma_a^j$. Furthermore, the preference shocks $\varepsilon_{g,t}$ is estimated to be much more volatile than the white-noise shock $\varepsilon_{g,t}$.

The posterior medians for the Taylor rule's parameters are by and large in line with what is often found by empirical studies that rely on perfect information models. Recall that the Taylor rule has two roles in the model. First, it describes central bank's reaction function. Second, it is an observation equation in firms' signal extraction problem. It is important to notice that the only source of information on preference shocks to firms is provided by the interest rate set by the central bank. Nonetheless, the posterior medians for the variance of the white-noise component of the preference shocks is relatively large. These parameter values tend to induce price-setters to believe that change in the interest rate has to be imputed to preference shocks $\varepsilon_{g,t}$. The likelihood function perhaps selects these parameter values so as to allow the model to account for the sluggish adjustments of the observables variables, such as real GDP and inflation, to monetary policy shocks, which has been documented by Christiano, Eichenbaum, and Evans (2005). Furthermore, the Bayesian updating shrinks the response to inflation $\phi_\pi$ and raises the response to output $\phi_y$, making the dynamics of the interest rate less informative about the current inflation and more informative about current output, relatively to what was conjectured in the prior.

As far as the posterior distribution obtained from sample $\mathcal{S}_2$ is concerned, the following four facts are important to emphasize. First, Fed's response to inflation, $\phi_\pi$, has increased in the subsample $\mathcal{S}_2$. Second, all shocks have smaller variance in the subsample $\mathcal{S}_2$. Third, in both samples the variance of the preference shocks is larger than both that of the white-noise ones and that of monetary policy shock, implying that firms mainly interpret changes in the interest rate as central bank's responses to preference shocks. Fourth, in both samples the variance of the white-noise technology shocks is larger than both that of the persistent shocks and that of the firm-specific shocks. This last result implies that firms find it harder to learn about persistent technology shocks relative to the white-noise ones.

### 4.4 Model Evaluation

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

We want to assess the empirical relevance of the transmission channel based on public's inflation expectations. To this end, we compute the posterior probability of a model where the interest rate set by the central bank is not observed by price setters. The only difference between such a *restricted* model and the incomplete information model described in Section 3 is, hence, firms' information set (8). In the restricted model firms' information set does not include the policy rate, $R_t$. Denote this restricted model as $\mathcal{M}_R$. 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 incomplete information model $\mathcal{M}_I$ of Section 3 and the restricted one $\mathcal{M}_R$. 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 $s \in \{I, R\}$, is given by:

$$\pi_{T,\mathcal{M}_s} = \frac{\pi_{0,\mathcal{M}_s} \cdot P(Y|\mathcal{M}_s)}{\sum_{s\in\{I,R\}} \pi_{0,\mathcal{M}_s} \cdot P(Y|\mathcal{M}_s)} \tag{16}$$

where $\pi_{0,\mathcal{M}_s}$ 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 $1.30E-31$. 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 new transmission channel based on affecting public's inflation expectations.

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 use the conditional marginal likelihood to make this evaluation. Denote the subset of the data including only the observed expectations as $Y^{SPF}$ and $\widetilde{Y}$ 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):

$$p\left(Y^{SPF}|\widetilde{Y},\mathcal{M}_s\right)=\frac{p\left(Y^{SPF},\widetilde{Y}|\mathcal{M}_s\right)}{p\left(\widetilde{Y}|\mathcal{M}_s\right)},\ s\in\{I,R\}$$

The numerator is nothing but the MDDs we have already computed for the incomplete information model[13] and the restricted model. The denominator is the MDD of the model $\mathcal{M}_s$ 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 $1.06E-43$ 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).

Finally, the paper evaluates to what extent adding information frictions helps fit the data. To this end, we compare the MDD associated with the incomplete information model and with the perfect information model $\mathcal{M}_{PIM}$, described in Section 3.9. Recall that the perfect information model boils down into a canonical three-equation New Keynesian DSGE model. Table 4 shows that the incomplete information model fits the data better than the perfect information model. This conclusion can be overturned only if the prior probability in favor of the incomplete information model is lower than $4.00E-30$.

## 4.5 Propagation of Monetary Policy Shocks

Figure 1 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. These impulse response functions are obtained when the model is fitted to the full sample (i.e.,

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

sample $\mathcal{S}_1$). The following three facts emerge: ($a$) the response of GDP and inflation to a monetary shock is very persistent; ($b$) the response of inflation and inflation expectations to a monetary shock is delayed and very sluggish, suggesting a great extent of price rigidity; ($c$) the initial response of inflation and inflation expectations to a contractionary monetary shock is positive. Fact ($c$) has been uncovered by Sims (1992) and dubbed as price puzzle by Eichenbaum (1992). In the subsequent sections we will discuss in details the explanation provided by the model for these three empirical facts that are common in the empirical literature based on VAR models.

### 4.5.1 Sluggish Adjustment of Output and Delayed Impact on Inflation

Facts ($a$) and ($b$) mainly stems from the new transmission channel based on inflation expectations. Figure 2 plots the 95% credible interval (the grey area reported in Figure 1) as well as the impulse response functions of the five observables to a monetary shock when price setters can observe the vector $\varphi_{t|t}^{(0:\infty)}$ errorless and this is common knowledge.[14] When this happens the incomplete information model boils down into a standard perfect-information New Keynesian model where nominal frictions are solely generated by price stickiness à la Calvo. In this model monetary shocks propagate through the channel based on the real interest rate, which is very typical to new Keynesian model. Figure 2 allows one to single out the transmission channel based on affecting price setters' expectations from the traditional one based on influencing the intertemporal allocation of consumption. It is apparent that the new channel that takes into account the responses of price setters' inflation expectations produces a great amount of persistence in the adjustment of inflation to a monetary shock.

Such a persistent pattern accords well with what is commonly found by the VAR literature[15]. It should be noted that this finding has been obtained by introducing information frictions into an otherwise very simple three-equation New Keynesian model. Within the class of New Keynesian DSGE models with perfect information, only large-scale DSGE models, such as Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007), with many shocks and frictions can reproduce sluggish and persistent response of macroeconomic variables to monetary disturbances. As shown in Figure, the baseline three-equation sticky-price model with perfect information fails generating such a pattern.

[14] The impulse response functions are computed at the posterior medians for the parameters of the incomplete information model estimated on a sample including observations over the 1970s.

[15] See, for instance, Christiano, Eichenbaum, and Evans (1999) and Stock and Watson (2001).

#### 4.5.2 The Price Puzzle

To investigate the fact $(c)$, Figure 3 plots the IRFs of the observables to a monetary policy shock when price setters do not observe the policy rate (i.e., the new channel based on the perfect observability of the interest rate is shut down).[16] Figure 3 clearly shows that the positive response of prices to a monetary shock stems from the 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. This is how the incomplete information model explains the price puzzle.

Figure 4 plots the response of the vector collecting the average expectations of the exogenous variables up to the second order:[17]

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

Average expectations about a preference shocks strongly react after a monetary policy shock. This 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 would raise nominal marginal costs, firms will decrease their price less forcefully than what they would have done if they knew that a monetary shock has occurred (i.e., the case of perfect information). The lenient monetary policy against inflation conducted during the 1970s exacerbates the issue by increasing the inflationary effects of preference shocks. As a consequence, inflation rate does not react much to monetary shocks. Moreover, it is important to mention that firms hold strong beliefs that rise in the interest rate might be due to a negative transient technology shocks, $\eta_{a,t}$.

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

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

Given this decomposition, the exact (cumulative) contribution of the average higher-order expectations $\varphi_{t|t}^{(0:k)}$ (up to the order $k$) about the state of technology $\left(z_t, \eta_{a,t}\right)$, the state of monetary policy $\left(\pi_t^*, \eta_{r,t}\right)$, and the state of households' preferences $\left(g_t, \eta_{g,t}\right)$ to the responses

[16] These impulse responses are obtained by calibrating the parameters of the incomplete information model at the posterior medians and by removing the policy rate $R_t$ from price setters' information set (8).

[17] Recall that by convention the average zero-order expectations correspond to the realizations of the variables themselves, that is $\varphi_{t|t}^{(0)} = \varphi_t \equiv \left( z_t, \eta_{a,t}, \hat{\pi}_{t|t}^*, \eta_{r,t}, g_t, \eta_{g,t} \right)'$.

of the endogenous variables, $\mathbf{s}_t \equiv \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t\right]'$, owing to a monetary policy shocks can be quantified as follows:

$$
\begin{aligned}
\frac{\partial \mathbf{s}_{t+h}}{\partial \eta_{r,t}} = & \underbrace{\sum_{l=0}^{h}\left[\frac{\partial \mathbf{s}_{t+h}}{\partial z^{(0:k)}_{t+l|t+l}}\frac{\partial z^{(0:k)}_{t+l|t+l}}{\partial \eta_{r,t}} + \frac{\partial \mathbf{s}_{t+h}}{\partial \eta^{(0:k)}_{a,t+l|t+l}}\frac{\partial \eta^{(0:k)}_{a,t+l|t+l}}{\partial \eta_{r,t}}\right]}_{\equiv \text{ Component associated with the HOEs about the state of technology}} \\
& + \underbrace{\sum_{l=0}^{h}\left[\frac{\partial \mathbf{s}_{t+h}}{\partial \eta^{(0:k)}_{r,t+l|t+l}}\frac{\partial \eta^{(0:k)}_{r,t+l|t+l}}{\partial \eta_{r,t}}\right]}_{\equiv \text{ Component associated with the HOEs about the state of monetary policy}} \\
& + \underbrace{\sum_{l=0}^{h}\left[\frac{\partial \mathbf{s}_{t+h}}{\partial g^{(0:k)}_{t+l|t+l}}\frac{\partial g^{(0:k)}_{t+l|t+l}}{\partial \eta_{r,t}}\right]}_{\equiv \text{ Component associated with the HOEs about the state of preferences}} \qquad (18)
\end{aligned}
$$

for $h = 0, 1, ..., 20$. Clearly, under perfect information the component associated with the Higher-Order Expectations (HOEs) about the state of technology and the component associated with the HOEs regarding the state of preferences are equal to zero in the aftermath of a monetary shock. For instance, under incomplete information, a large component associated with the HOEs about the state of preferences can be interpreted as a situation where price-setters mistakenly believe that the interest rate has changed as a result of a preference shock, which has inflationary effects (perhaps because of a monetary policy not tough enough against inflation).

Figure 5 shows why upon a contractionary monetary shock inflation rises: 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 response of the component associated with the HOEs about the state of preferences is positive and very large upon a monetary shock. This pattern partially offsets the contemporaneous response of the component associated with the HOEs about the state of monetary policy, which pushes the response of inflation into the negative 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.

### 4.5.3 The Disappearance of the Price Puzzle after the 1970s

Recent studies based on VAR models[18] have documented that the price puzzle is statistically relevant in the 1970s only. I compute the impulse response functions to a contractionary monetary policy shock when I estimate the model not including the 1970s (i.e., data set $\mathcal{S}_2$). Posterior medians for this subsample are reported in Table 3 and were discussed in Section 4.3. Figure 6 reports the impulse response functions (and their 95% posterior credible sets) to a 25 bps rise of the policy rate after 1970s. The main finding is that excluding the data of the 1970s from the sample leads the (average first-order) inflation expectations to respond negatively to monetary shocks upon impact. From 1980s on, the Federal Reserve, hence, seems to have successfully affirmed a better control over the initial response of price setters' inflation expectations. Four-step ahead inflation expectations never lie in the positive region (i.e., values for inflation expectations that are larger than their pre-shock level).

This finding is line with other VAR studies but raises the question of what can explain the disappearance of the positive initial response of inflation to monetary shock. A quick comparison of Figure 7 with Figure 5 shows that the average higher order expectations about the monetary component react more strongly upon the monetary than what it was when one includes the 1970s in the data range. Furthermore, after 1970s price setters consider, on average, less likely that the rise in the policy rate has been triggered by a positive preference shocks. These two effects causes inflation to respond negatively upon a monetary shock.

# 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 then 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

[18] Price puzzle is found to be statistically insignificant after 1970s by several empirical studies based on VAR models. See for instance, Barth and Ramey (2001), Hansen (2004), Castelnuovo and Surico (2010). and Ravn, Schmitt-Grohe, Uribe, and Uusküla (2010). Barth and Ramey (2001) and Hansen (2004) document the disappearence of the price puzzle by using monthly data and zero-restrictions to identify the monetary shock. Hansen (2004) corroborates this result by using several proxies for the expected future inflation. 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 afetr 1970s the price puzzle is substantially weaker. These scholars argue that the mitigation of the price puzzle observed after the 1970s is due to structural changes in the conduct of monetary policy.

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*. The paper also finds evidence that information frictions significantly contribute to improve the fit of the model.

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 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. In the model, for instance, 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 what it knows. In the present model this assumption is justified by the presence of nominal frictions that create an incentive for the central bank to make surprise inflation. 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. Nonetheless, it is well known that market participants (over-)react to central bank's announcements. Empirically assessing how central bank's communication affects the transmission mechanism of monetary impulses is left for future research. Furthermore, 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 would be a major step forward for understanding the channel of monetary transmission based on agents' expectations.

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

ARUOBA, S. B., AND F. SCHORFHEIDE (2011): "Sticky Prices versus Monetary Frictions: An Estimation of Policy Trade-Offs," *American Economic Journal: Macroeconomics*, 3(1), 60–90.

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," Mimeo Duke University.

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

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, 1262Ű1283.

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.

Del Negro, M., and S. Eusepi (2010): "Fitting Observed Inflation Expectations," FRB of New York Staff Report No. 476.

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.

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

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.

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.

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 Nonneutrality and Structural Slumps," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, pp. 64–86. Princeton University Press, Princeton.

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

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

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

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

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

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

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.

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

Rabanal, P., and J. F. Rubio-Ramírez (2005): "Comparing New Keynesian models of the business cycle: A Bayesian approach," *Journal of Monetary Economics*, 52, 1151–1166.

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 Bayeasian DSGE Approach," *American Economic Review*, 97(3), 586–606.

Stock, J. H., and M. W. Watson (2001): "Vector Autoregressions," *Journal of Economic Perspectives*, 15(4), 101–115.

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

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

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 nonistitutional population - 16 years and over | BLS (LNS10000000) |
| $PGDP_t$ | 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 Expections of current GDP price index | SPFs in mean.xls (PGDP2) |
| $PGDP3_t$ | Mean of Expections 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 |
|---|---|---|---|---|
| $\theta$ | $[0,1]$ | Beta | 0.65 | $[0.55, 0.93]$ |
| $\phi_\pi$ | $\mathbb{R}+$ | Gamma | 1.80 | $[0.89, 2.80]$ |
| $\phi_y$ | $\mathbb{R}+$ | Gamma | 0.25 | $[0.07, 0.45]$ |
| $\rho_r$ | $[0,1]$ | Beta | 0.80 | $[0.60, 0.97]$ |
| $\rho_z$ | $[0,1]$ | Beta | 0.85 | $[0.65, 0.99]$ |
| $\rho_g$ | $[0,1]$ | Beta | 0.85 | $[0.65, 0.99]$ |
| $\sigma_a$ | $\mathbb{R}+$ | InvGamma | 0.10 | $[0.05, 0.24]$ |
| $\sigma_z$ | $\mathbb{R}+$ | InvGamma | 0.70 | $[0.51, 0.97]$ |
| $\sigma_a^j$ | $\mathbb{R}+$ | InvGamma | 2.50 | $[1.00, 4.85]$ |
| $\sigma_r$ | $\mathbb{R}+$ | InvGamma | 0.10 | $[0.05, 0.25]$ |
| $\sigma_g$ | $\mathbb{R}+$ | InvGamma | 0.25 | $[0.10, 0.48]$ |
| $\sigma_{m_1}$ | $\mathbb{R}+$ | InvGamma | 0.25 | $[0.10, 0.48]$ |
| $\sigma_{m_2}$ | $\mathbb{R}+$ | InvGamma | 0.25 | $[0.10, 0.48]$ |
| $\ln A_0$ | $\mathbb{R}$ | Normal | 0.00 | $[-0.20, 0.20]$ |
| $\ln \pi_*$ | $\mathbb{R}$ | Normal | 0.00 | $[-0.20, 0.20]$ |

Table 3: Posterior Distributions

| Name | Full Sample ($\mathcal{S}_1$) | | Sub-sample ($\mathcal{S}_2$) | |
|---|---|---|---|---|
| | Median | 95% Interval | Median | 95% Interval |
| $\theta$ | 0.82 | $[0.79, 0.84]$ | 0.85 | $[0.83, 0.86]$ |
| $\phi_\pi$ | 1.42 | $[1.40, 1.46]$ | 1.81 | $[1.64, 2.08]$ |
| $\phi_y$ | 0.48 | $[0.40, 0.56]$ | 0.67 | $[0.56, 0.75]$ |
| $\rho_r$ | 0.74 | $[0.70, 0.76]$ | 0.83 | $[0.80, 0.85]$ |
| $\rho_z$ | 0.99 | $[0.98, 0.99]$ | 0.99 | $[0.99, 1.00]$ |
| $\rho_g$ | 0.77 | $[0.75, 0.79]$ | 0.79 | $[0.76, 0.81]$ |
| $\sigma_a$ | 5.98 | $[4.16, 7.85]$ | 5.65 | $[5.05, 6.60]$ |
| $\sigma_z$ | 1.29 | $[1.05, 1.54]$ | 0.98 | $[0.80, 1.16]$ |
| $\sigma_a^j$ | 4.33 | $[3.36, 5.41]$ | 3.07 | $[2.71, 3.38]$ |
| $\sigma_r$ | 0.25 | $[0.22, 0.28]$ | 0.13 | $[0.11, 0.15]$ |
| $\sigma_g$ | 3.70 | $[3.20, 4.15]$ | 2.46 | $[2.21, 2.73]$ |
| $\sigma_{m_1}$ | 0.14 | $[0.13, 0.16]$ | 0.13 | $[0.12, 0.14]$ |
| $\sigma_{m_2}$ | 0.13 | $[0.12, 0.15]$ | 0.12 | $[0.11, 0.13]$ |
| $100 \ln A_0$ | 0.28 | $[0.25, 0.31]$ | 0.34 | $[0.27, 0.42]$ |
| $100 \ln \pi_*$ | 0.37 | $[0.03, 0.74]$ | 1.09 | $[0.62, 1.61]$ |

Table 4: Marginal-Data-Density Comparisons

| | $\mathcal{M}_I$ | $\mathcal{M}_R$ | $\mathcal{M}_{PIM}$ |
|---|---|---|---|
| MDD | -205.23 | -276.36 | -272.94 |
| Conditional MDD | 158.14 | 59.19 | NA |

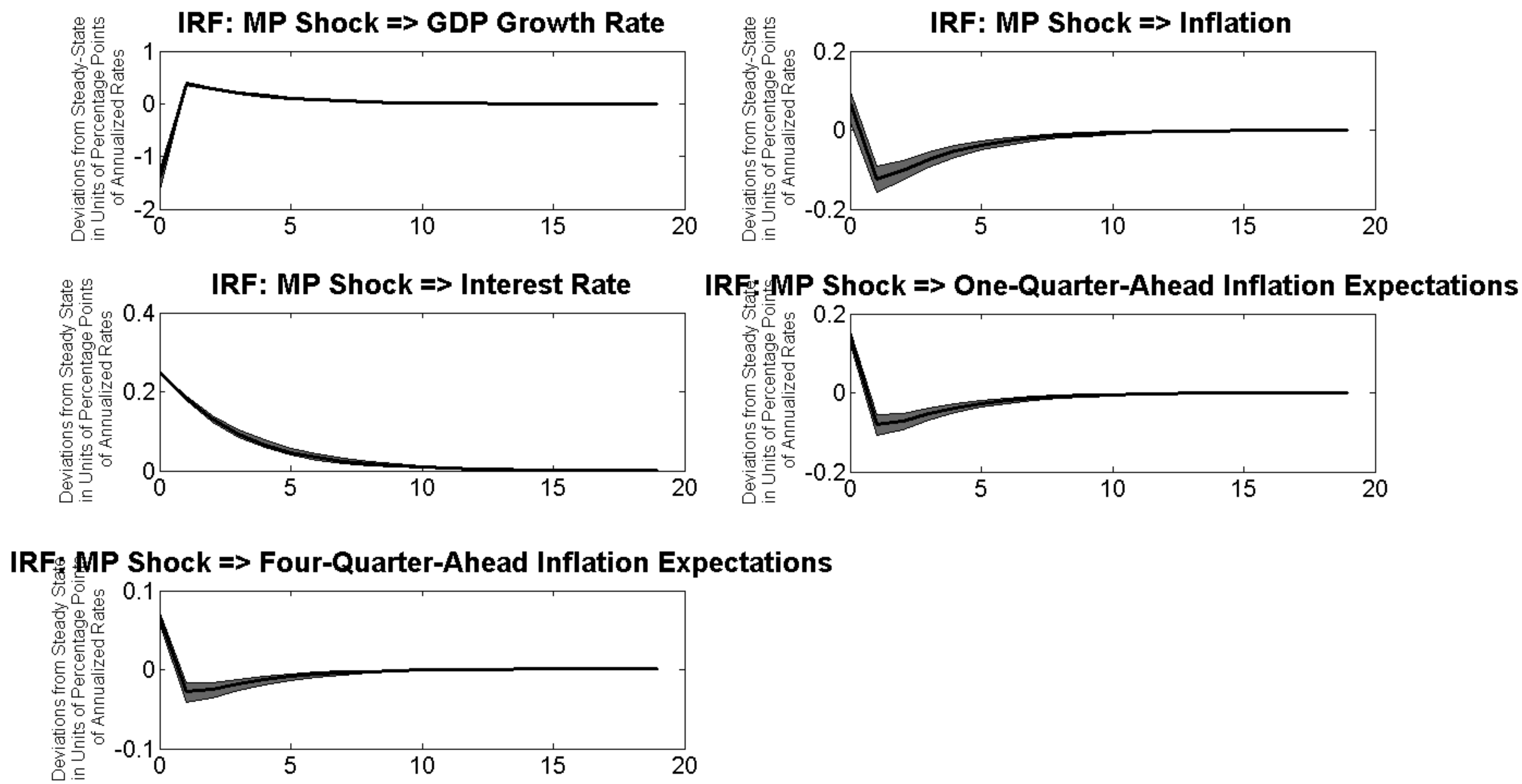


Figure 1: Impulse Response Functions of the Observables to a Monetary Policy Shock. Sample including the 1970s

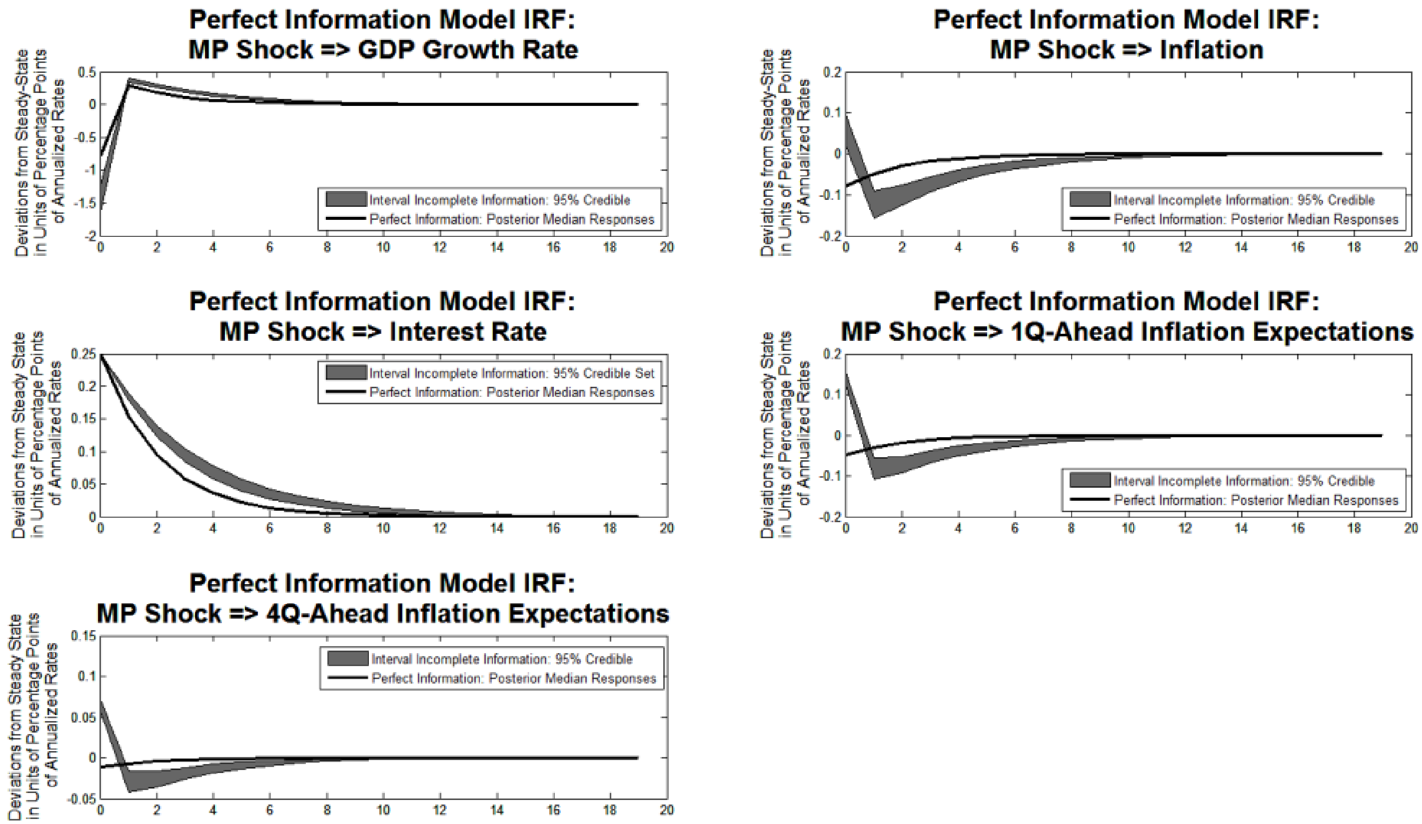


Figure 2: Impulse Response Functions of the Observables to a Monetary Policy Shock: Perfect Information Model (PIM) vs. Incomplete Information Model (IIM).

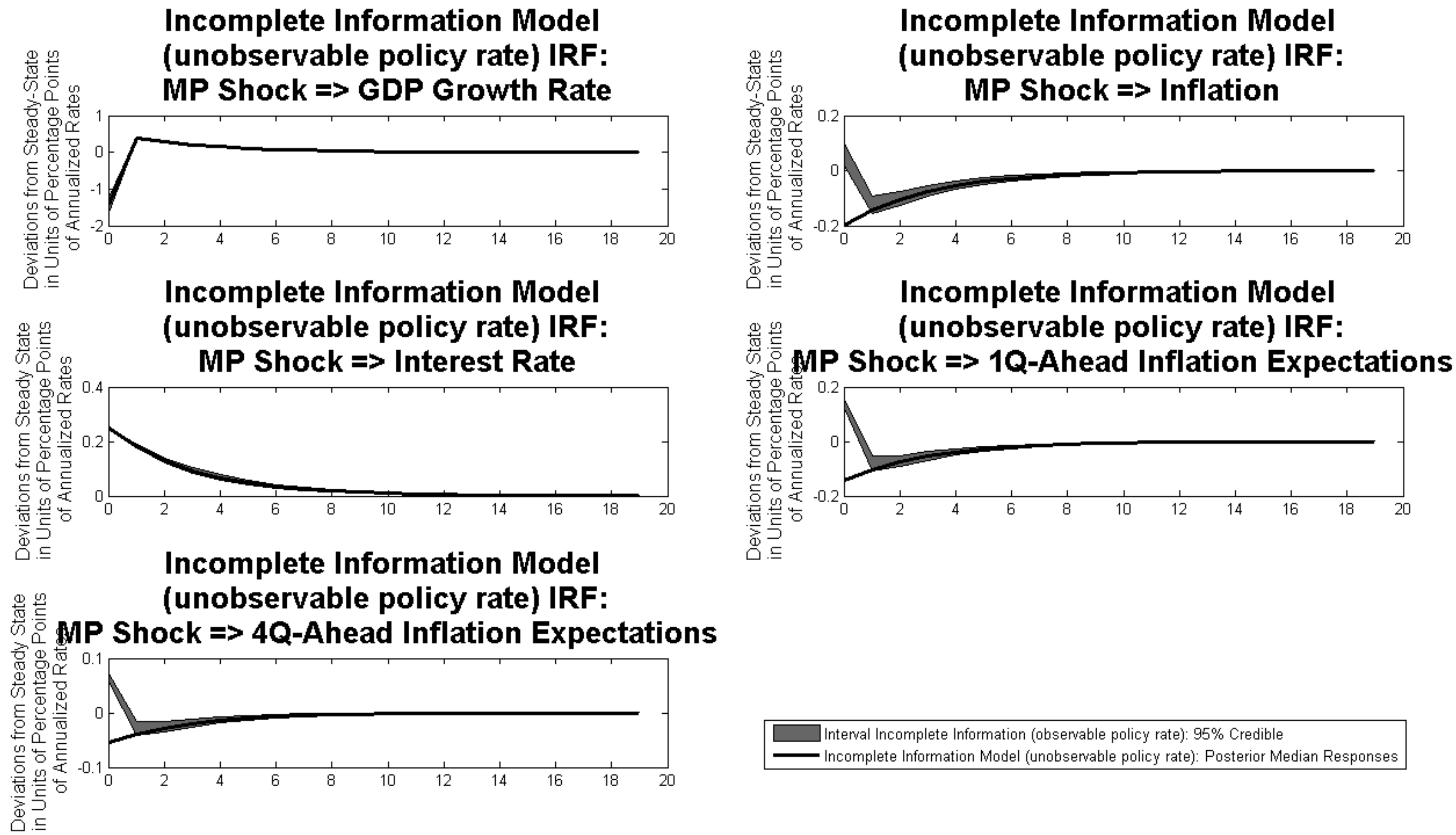


Figure 3: Impulse Response Functions of the Observables to a Monetary Policy Shock: The Effects of the Expectation Channel vs. Incomplete Information Model (IIM). Sample including the 1970s.

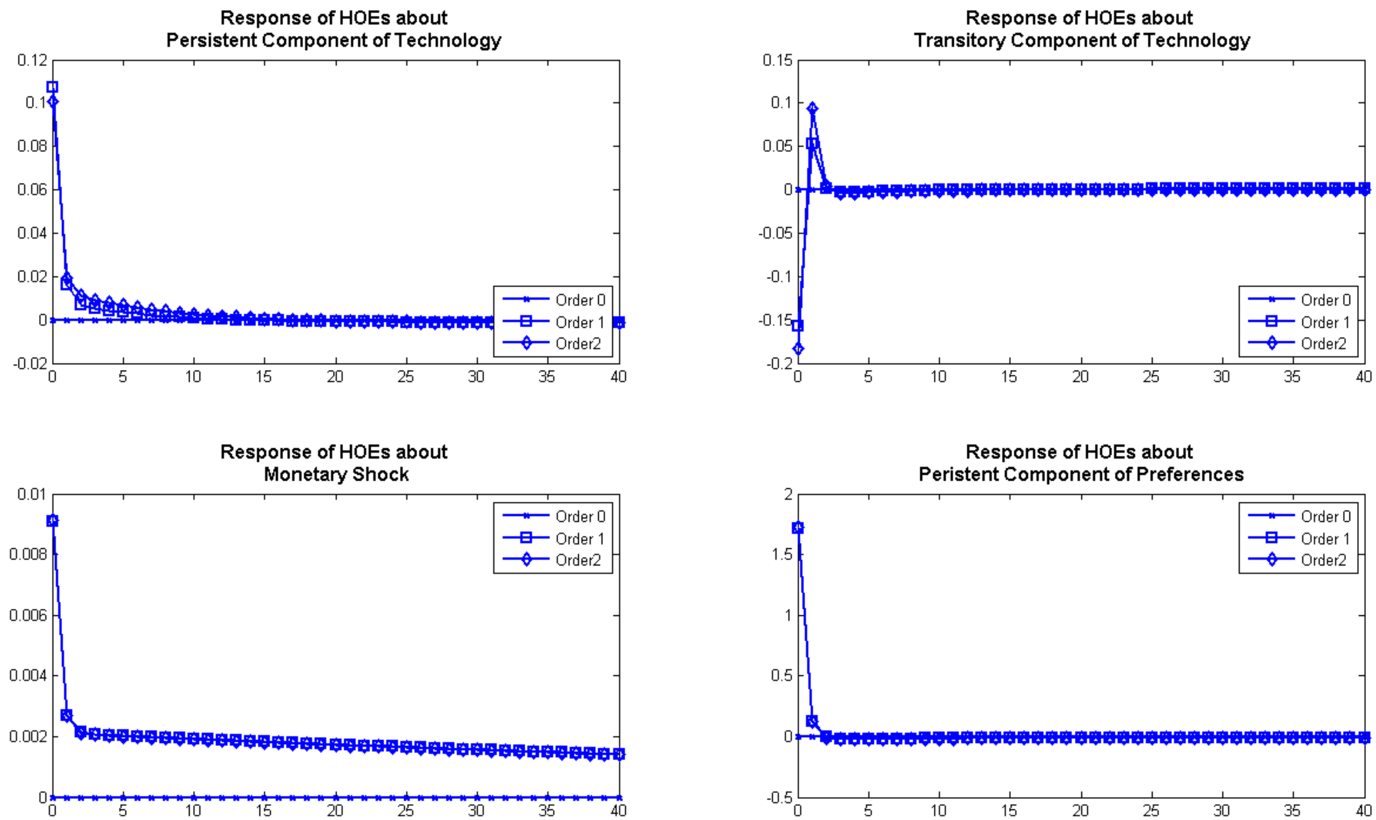


Figure 4: Impulse Response Functions of the High-Order-Expectations about Exogenous Variables to a Monetary Policy Shock. Sample including the 1970s

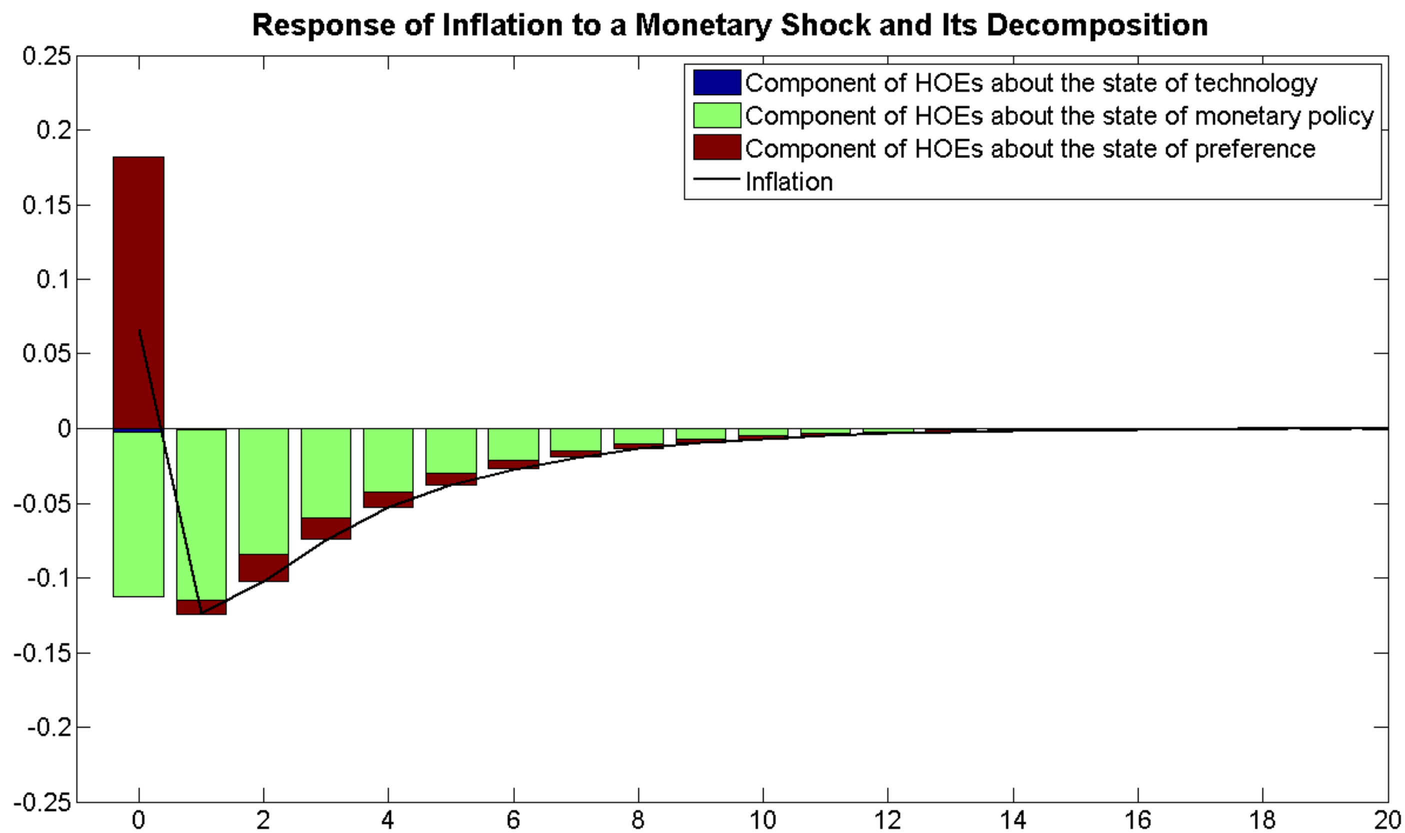


Figure 5: HOE Decomposition of the IRFs of Inflation to a Monetary Policy Shock. Sample including the 1970s

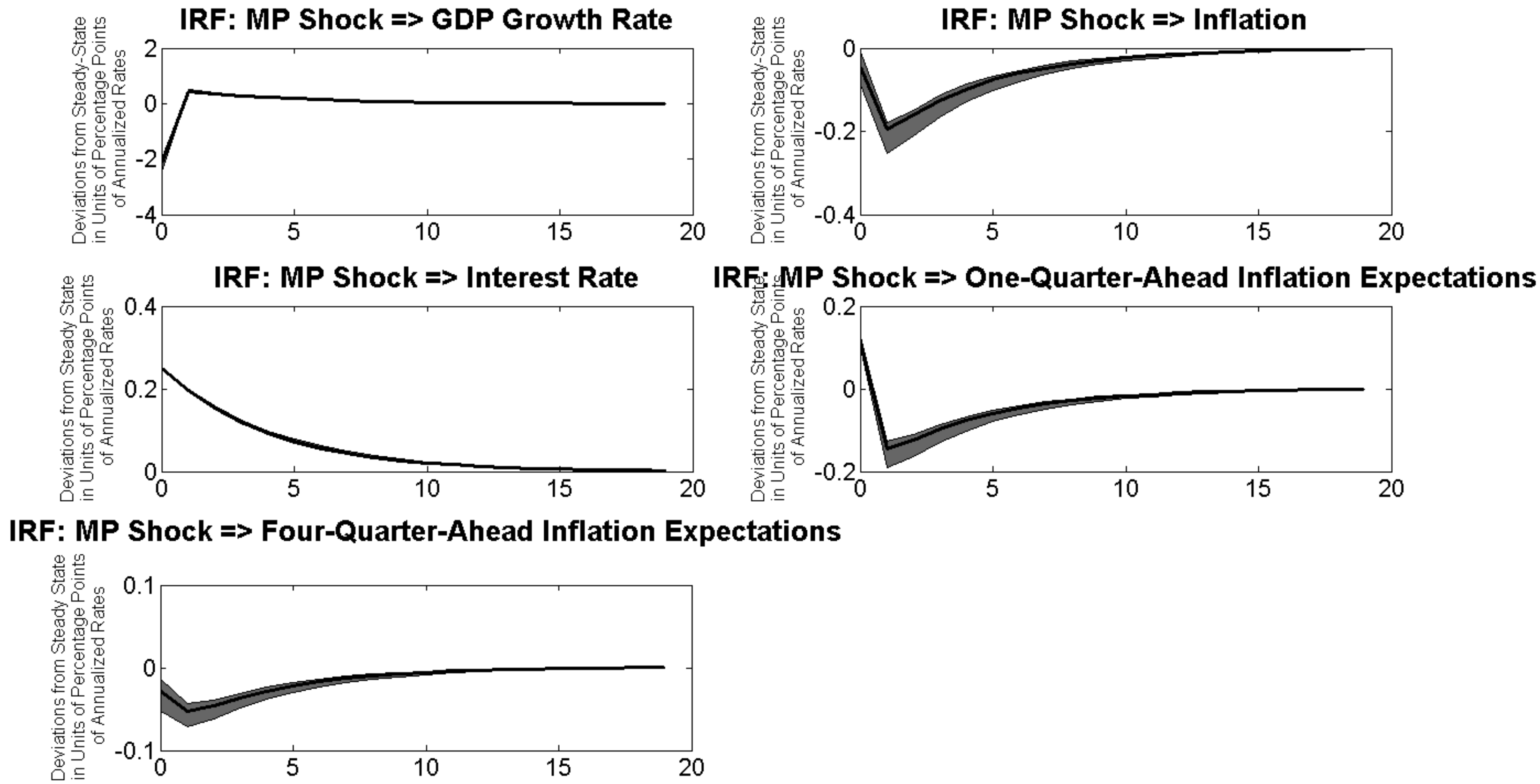


Figure 6: Impulse Response Functions of the Observables to a Monetary Policy Shock in the Sub-samples not including the 1970s

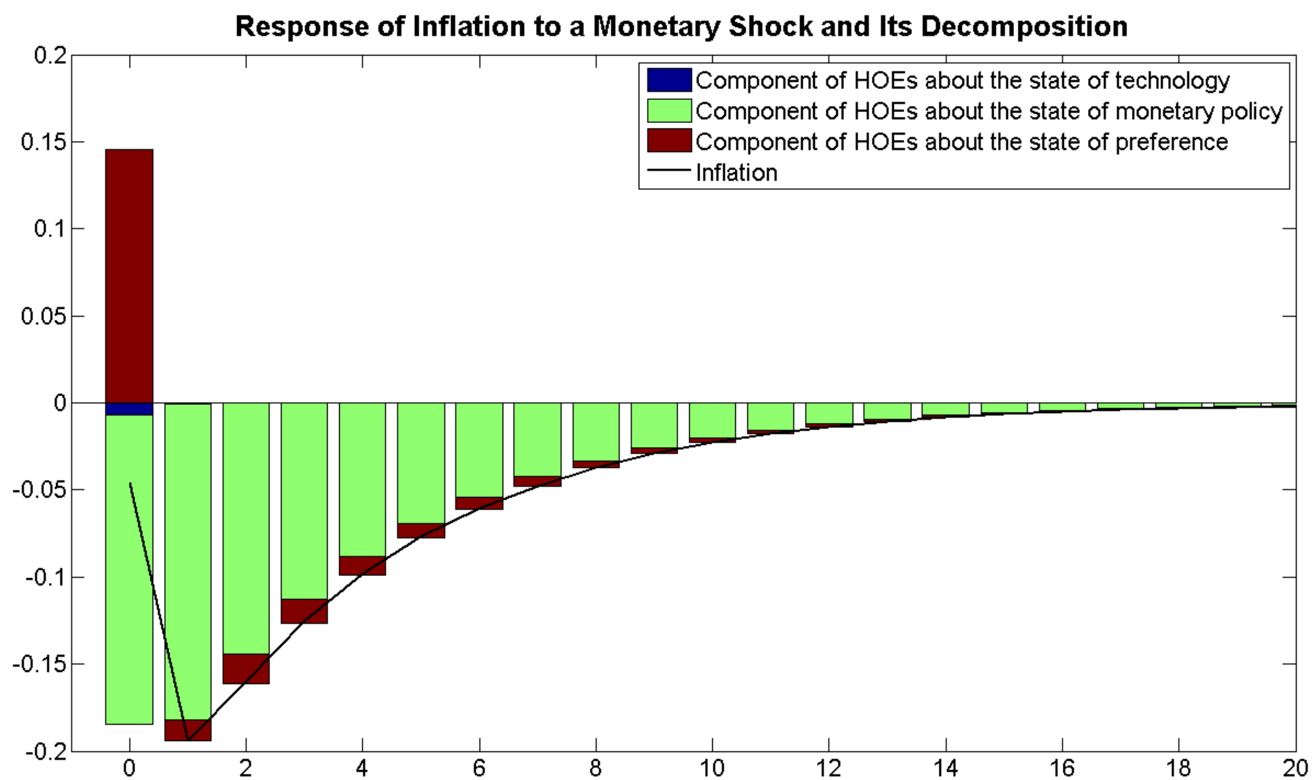


Figure 7: HOE Decomposition of the IRFs of Inflation to a Monetary Policy Shock. Sample not including the 1970s

# Appendix

In Section A, I provide derivation of the imperfect-common-knowledge Phillips curve (11). 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 $\widehat{c}_t = \widehat{w}_t$. 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

$$\widehat{mc}_{j,t} = \hat{w}_{j,t} - z_t - \sigma_a \eta_{a,t} - \eta^a_{j,t}$$

Recall that $(\ln A_{j,t} - \ln A_0 \cdot \gamma t) \in \mathcal{I}_{j,t}$ and write:

$$\mathbb{E}_{j,t}\widehat{mc}_{j,t} = \mathbb{E}_{j,t}\hat{w}_{j,t} \underbrace{- z_t - \sigma_a \eta_{a,t} - \eta^a_{j,t}}_{\ln A_{j,t} - \ln A_0 \cdot \gamma t}$$

where $\mathbb{E}_{j,t}$ are expectations conditioned on firm $j$'s information set at time $t$, $\mathcal{I}_{j,t}$, defined in (8). Using the equation (19) for replacing $\hat{w}_t$ yields:

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

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

$$\widehat{mc}^{(0)}_{t|t} = \hat{y}^{(1)}_{t|t} - z_t - \sigma_a \eta_{a,t}$$

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

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

and then

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

By rearranging:

$$\ln P_t = \theta \left(\ln P_{t-1} + \ln \pi_*\right) + (1-\theta) \int \left(\ln P^*_{j,t}\right) dj \tag{20}$$

The price-setting equation is:

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

Define

$$\begin{aligned}
y_t &= \frac{Y_t}{A_0^t},\ c_t = \frac{C_t}{A_0^t};\ p^*_{j,t} = \frac{P^*_{j,t}}{P_t},\ y_{j,t} = \frac{Y_{j,t}}{A_0^t} \\
w_t &= \frac{W_t}{A_0^t P_t},\ a_t = \frac{A_t}{A_0^t}, R_t = \frac{R_t}{R_*},\ mc_{j,t} = \frac{MC_{j,t}}{P_t} \\
\xi_{j,t} &= A_0^t \Xi_{j,t}
\end{aligned}$$

Hence, write:

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

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

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

Taking the derivatives yield:

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

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

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

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:

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

Rolling this equation one step ahead yields:

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

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

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

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

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

Therefore,

$$\sum_{s=1}^{\infty}(\beta\theta)^s\,\mathbb{E}\left[\widehat{mc}_{j,t+s}|\mathcal{I}_{j,t}\right] = \frac{\beta\theta}{1-\beta\theta}\left[\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right) - \mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right] - \beta\theta\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right] \tag{23}$$

The equation (22) can be rewritten as:

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

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

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

Consider the last term:

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

Therefore we can write that

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

Note that

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

Hence,

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

and by simplifying:

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

We substitute this result into the original equation to get:

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

After simplifying we get:

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

We can rearrange:

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

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

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

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

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

where expectations and integration across firms are taken $k$ times.

Let us denote the **average reset price** as $\ln P_t^* = \int \ln P_{j,t}^* dj$. We can integrate equation (27) across firms to obtain an equation for the average reset price:

$$\begin{aligned}\ln P_t^{*} &= (1-\beta\theta)\cdot\widehat{mc}_{t|t}^{(0)} + (1-\beta\theta)\ln P_{t|t}^{(1)} \\ &\quad +\beta\theta\ln P_{t+1|t}^{*(1)} - \beta\theta\ln\pi_* \end{aligned} \tag{28}$$

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

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

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

$$\begin{aligned}\ln P_t &= \theta\ln P_{t-1} + \left(\theta - (1-\theta)\,\beta\theta\right)\ln\pi_* \\ &\quad +(1-\theta)\left[(1-\beta\theta)\cdot\widehat{mc}_{t|t}^{(0)} + (1-\beta\theta)\ln P_{t|t}^{(1)} + \beta\theta\ln P_{t+1|t}^{*(1)}\right] \end{aligned} \tag{30}$$

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

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

Applying these three results to equation (30) yields:

$$\begin{aligned} \widehat{\pi}_t + \ln P_{t-1} + \ln \pi_* &= \theta \ln P_{t-1} + (\theta - (1-\theta)\,\beta\theta) \ln \pi_* \\ &+ (1-\theta) \left[ (1-\beta\theta) \cdot \widehat{mc}_{t|t}^{(0)} + (1-\beta\theta) \ln P_{t|t}^{(1)} + \beta\theta \left( \frac{\hat{\pi}_{t+1|t}^{(1)}}{1-\theta} + \ln P_{t|t}^{(1)} + \ln \pi_* \right) \right] \end{aligned} \tag{31}$$

and after simplifying:

$$\widehat{\pi}_t = (1-\theta)(1-\beta\theta) \cdot \widehat{mc}_{t|t}^{(0)} + (1-\theta)\,\widehat{\pi}_{t|t}^{(1)} + \beta\theta \left( \hat{\pi}_{t+1|t}^{(1)} \right) \tag{32}$$

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

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

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

$$\widehat{\pi}_t = (1-\theta)(1-\beta\theta) \sum_{k=0}^{\infty} (1-\theta)^k\, \widehat{mc}_{t|t}^{(k)} + \beta\theta \sum_{k=0}^{\infty} (1-\theta)^k\, \hat{\pi}_{t+1|t}^{(k+1)} \tag{33}$$

# 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 $(\widehat{\pi}_t, \widehat{y}_t, \widehat{R}_t)$, 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

[19] This 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 $\ln P_t = \widehat{\pi}_t + \ln P_{t-1} + \ln \pi_*$:

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

Rolling one period forward:

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

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

coefficient matrices, **M** and **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*

$$\mathbf{T}^{(s)} \equiv \begin{bmatrix} \mathbf{0}_{6(k-s+1)\times 6s} & \mathbf{I}_{6(k-s+1)} \\ \mathbf{0}_{6s\times 6s} & \mathbf{0}_{6s\times(k+1-s)6} \end{bmatrix}$$

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

$$\varphi_{t|t}^{(s:k+s)} \equiv \begin{bmatrix} \varphi_{t|t}^{(s:k)} \\ \varphi_{t|t}^{(s:k+s)} \end{bmatrix}_{6(k+1)\times 1} = \begin{bmatrix} \varphi_{t|t}^{(s:k)} \\ \mathbf{0}_{6s\times 1} \end{bmatrix}_{6(k+1)\times 1}$$

Note that

$$\varphi_{t|t}^{(s:k+s)} = \begin{bmatrix} \mathbf{0}_{6(k-s+1)\times 6s} & \mathbf{I}_{6(k-s+1)} \\ \mathbf{0}_{6s\times 6s} & \mathbf{0}_{6s\times(k+1-s)6} \end{bmatrix} \underbrace{\begin{bmatrix} \varphi_{t|t}^{(0:s-1)} \\ \varphi_{t|t}^{(s:k)} \end{bmatrix}}_{\varphi_{t|t}^{(0:k)}}$$

■

**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:

$$\mathbf{s}_{t|t}^{(1)} = \mathbf{v}_0 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1} \tag{34}$$

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 (34) 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 (13), the Phillips curve (11), and the Taylor Rule (14). One can use these structural equations to pin down the vectors $\mathbf{v}_0 \equiv \left[\mathbf{a}_0', \mathbf{b}_0', \mathbf{c}_0'\right]'$

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

Let start from the IS equation (13)

$$\begin{aligned}
\mathbf{b}_0 \varphi_{t|t}^{(0:k)} + \mathbf{b}_1 \mathbf{s}_{t-1} &= \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right) \varphi_{t|t}^{(0:k)} - \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right) \mathbf{M} \varphi_{t|t}^{(0:k)} \\
&+ \left(\mathbf{1}_1^T + \mathbf{1}_2^T\right) \left[\underbrace{\mathbf{v}_0 \mathbf{M} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \left(\mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1}\right)}_{\mathbb{E}_t \mathbf{s}_{t+1}}\right] \\
&- \left(\mathbf{c}_0 \varphi_{t|t}^{(0:k)} + \mathbf{c}_1 \mathbf{s}_{t-1}\right)
\end{aligned}$$

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 (13) holds in equilibrium.

$$\mathbf{b}_0 = \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right) + \left(\mathbf{1}_1^T + \mathbf{1}_2^T\right) \left(\mathbf{v}_0 \mathbf{M} + \mathbf{v}_1 \mathbf{v}_0\right) - \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right) \mathbf{M} - \mathbf{c}_0 \quad (35)$$

$$\mathbf{b}_1 = \left(\mathbf{1}_1^T + \mathbf{1}_2^T\right) \mathbf{v}_1 \mathbf{v}_1 - \mathbf{c}_1 \quad (36)$$

The Phillips curve (11) can be rewritten as:

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

where $\boldsymbol{\gamma}_a^{(s)} = \left[\mathbf{0}_{1\times 6s}, (1,1,0,0,0,0), \mathbf{0}_{1\times 6(k-s)}\right]'$. The following restrictions upon vectors of coefficients $\mathbf{a}_0$ and $\mathbf{a}_1$ can be derived from the Phillips curve above:

$$\begin{aligned}
\mathbf{a}_0 &= (1-\theta)(1-\beta\theta) \left[\boldsymbol{\nu} \mathbf{m}_1 - \left(\sum_{s=0}^{k-1} (1-\theta)^s \boldsymbol{\gamma}_a^{(s)\prime}\right)\right] \\
&+ \beta\theta \boldsymbol{\nu} \mathbf{m}_2 + \beta\theta \left(\sum_{s=0}^{k-1} (1-\theta)^s\right) \mathbf{1}_1^T \mathbf{v}_1 \mathbf{v}_0 \quad (37)
\end{aligned}$$

$$\mathbf{a}_1 = (1-\theta)(1-\beta\theta) \left(\sum_{s=0}^{k-1} (1-\theta)^s\right) \mathbf{1}_2^T \mathbf{v}_1 + \beta\theta \left(\sum_{s=0}^{k-1} (1-\theta)^s\right) \mathbf{1}_1^T \mathbf{v}_1 \mathbf{v}_1 \quad (38)$$

where I define:

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

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

Finally, the Taylor rule (14) implies that:

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

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

$$\mathbf{c}_0 = (1-\rho_r)\left[\phi_\pi \mathbf{a}_0 + (1-\phi_\pi)\,\mathbf{1}_3^T + \phi_y \mathbf{b}_0\right] + \mathbf{1}_4^T \tag{39}$$

$$\mathbf{c}_1 = \rho_r \mathbf{1}_3^T + (1-\rho_r)\left[\phi_\pi \mathbf{a}_1 + \phi_y\left(\mathbf{b}_1 - \mathbf{1}_2^T\right)\right] \tag{40}$$

Equations (35)-(40) are a system of non-linear equations in the coefficients $\mathbf{v}_0 \equiv \left[\mathbf{a}_0', \mathbf{b}_0', \mathbf{c}_0'\right]'$ and $\mathbf{v}_1 \equiv \left[\mathbf{a}_1', \mathbf{b}_1', \mathbf{c}_1'\right]'$. 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., $z_t$, $\eta_{a,t}$, $\hat{\pi}^*_{t|t}$, $\eta_{r,t}$, $g_t$, $\eta_{g,t}$) for given parameter values and vectors of coefficients $\mathbf{v}_0$ and $\mathbf{v}_1$.[21] We focus on equilibria where HOEs evolve:

$$\varphi_{t|t}^{(0:k)} = \mathbf{M}\varphi_{t-1|t-1}^{(0:k)} + \mathbf{N}\boldsymbol{\varepsilon}_t \tag{41}$$

where $\boldsymbol{\varepsilon}_t \equiv \begin{bmatrix} \varepsilon_{z,t} & \eta_{a,t} & \eta_{r,t} & \varepsilon_{g,t} \end{bmatrix}'$

A quick inspection of equation (41) 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. Firms' reduced-form state-space model can be concisely cast as follows:

$$\mathbf{X}_t = \mathbf{W} \cdot \mathbf{X}_{t-1} + \mathbf{U} \cdot \boldsymbol{\varepsilon}_t \tag{42}$$

[20] I find that non-linear solver fails to find a solution when the parameter $\phi_\pi$ is too small.

[21] Recall 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.

$$\mathbf{Z}_t = \mathbf{D}_1 \mathbf{X}_t + \mathbf{D}_2 \mathbf{X}_{t-1} + \mathbf{Q} e_{j,t} \tag{43}$$

where $\mathbf{X}_t = \left[\varphi_{t|t}^{(0:k)\prime}, \mathbf{s}_t'\right]$, $\mathbf{s}_t \equiv \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t\right]'$, $\mathbf{W} \equiv \mathbf{A}^{-1}\mathbf{B}$, $\mathbf{U} \equiv \mathbf{A}^{-1}\mathbf{C}$,

$$\begin{aligned} \mathbf{D}_1 &= \left[ \begin{array}{cccc} \mathbf{d}_1' & \mathbf{0}_{1\times(k+1)\times 6+3} & \mathbf{0}_{1\times(k+1)\times 6+3} & \mathbf{d}_4' \end{array} \right]' \\ \mathbf{D}_2 &= \left[ \begin{array}{cccc} \mathbf{0}_{1\times(k+1)\times 6+3} & \mathbf{d}_2' & \mathbf{d}_3' & \mathbf{0}_{1\times(k+1)\times 6+3} \end{array} \right]'. \end{aligned}$$

where $\mathbf{d}_1' = \left[1, 1, \mathbf{0}_{1\times(k+1)\times 6+1}\right]$, $\mathbf{d}_2' = \left[\mathbf{0}_{1\times(k+1)\times 6}, 1, 0, 0\right]$, $\mathbf{d}_3' = \left[\mathbf{0}_{1\times(k+1)\times 6+1}, 1, 0\right]$, $\mathbf{d}_4' = \left[\mathbf{0}_{1\times(k+1)\times 6+2}, 1\right]$ and $\mathbf{Q} = \left[\sigma_a^j, 0, 0, 0\right]'$. Recall that we defined

$$\mathbf{A} \equiv \left[ \begin{array}{cc} \mathbb{I} & \mathbf{0} \\ -\mathbf{v}_0 & \mathbb{I} \end{array} \right], \ \mathbf{B} \equiv \left[ \begin{array}{cc} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_1 \end{array} \right], \ \mathbf{C} \equiv \left[ \begin{array}{c} \mathbf{N} \\ \mathbf{0} \end{array} \right]$$

where the matrices $\mathbf{v}_0$ and $\mathbf{v}_1$ have some known values. Note that since three signals in $\mathbf{Z}_t$ are endogenous, the state vector $\mathbf{X}_t$ must include the endogenous state variables $\mathbf{s}_t \equiv \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t\right]'$.

Solving firms' signal extraction problem requires applying the Kalman filter. Note that firms' observation equations (43) 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 $\mathbf{X}_{0|0} \backsim \mathcal{N}\left(\mathbf{X}_{0|0}, \mathbf{P}_{0|0}\right)$.

**Forecasting:**

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

$$\begin{aligned} \mathbf{X}_{t|t-1} &= \mathbf{W}\mathbf{X}_{t-1|t-1} \\ \mathbf{P}_{t|t-1} &= \mathbf{W}\mathbf{P}_{\mathbf{t-1|t-1}}\mathbf{W}' + \mathbf{U}\mathbf{U}' \end{aligned} \tag{44}$$

The density

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

where

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

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

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

and by working the product out, one obtains:

$$\begin{aligned} \mathbf{F}_{t|t-1} = \ & \mathbf{D}_1\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{D}_2\mathbf{P}_{t-1|t-1}\mathbf{D}_2' + \mathbf{Q}\mathbf{Q}' + \\ & + \mathbf{D}_1 E\left[\mathbf{X}_t\mathbf{X}_{t-1}'|\mathbf{Z}^{t-1}\right]\mathbf{D}_2' + \mathbf{D}_2 E\left[\mathbf{X}_{t-1}\mathbf{X}_t'|\mathbf{Z}^{t-1}\right]\mathbf{D}_1' \end{aligned} \tag{45}$$

Note that

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

Combining this result with equation (45) yields:

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

by substituting equation (44) into the equation above leads to

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

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

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

as

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

where the second equality follows from using equation (42). Hence, it follows that

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

where[22]

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

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

Therefore, combining equation (47) with equation (44) yields:

$$\mathbf{P}_{t+1|t} = \mathbf{W}\left[\mathbf{P}_{t|t-1} - \left(\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2'\right)\mathbf{F}_{t|t-1}^{-1}\left(\mathbf{D}_1\mathbf{P}_{t|t-1}' + \mathbf{D}_2\mathbf{P}_{t-1|t-1}\mathbf{W}'\right)\right]\mathbf{W}' + \mathbf{U}\mathbf{U}' \quad (48)$$

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

$$\mu = \left[\begin{array}{c}\mu_x\\ \mu_y\end{array}\right] \text{ and } \Sigma = \left[\begin{array}{cc}\Sigma_{xx} & \Sigma_{xy}\\ \Sigma_{yx} & \Sigma_{yy}\end{array}\right]$$

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

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

Recall equation (43) and write the law of motion of the firm j's first-order beliefs about $\mathbf{X}_t$ as

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

where we have combined equations (46) and (43). By recalling that $\mathbf{X}_{t|t-1}(j) = \mathbf{W}\mathbf{X}_{t-1|t-1}(j)$, we have:

$$\begin{aligned}\mathbf{X}_{t|t} &= \mathbf{W}\mathbf{X}_{t-1|t-1} + \\ &+ \underbrace{\left[\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2'\right]\mathbf{F}_{t|t-1}^{-1}}_{\mathbf{K}}\left[\mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t} - \left(\mathbf{D}_1\mathbf{W}\mathbf{X}_{t-1|t-1} + \mathbf{D}_2\mathbf{X}_{t-1|t-1}\right)\right]\end{aligned}$$

By rearranging one obtains:

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

The vector $\mathbf{X}_{t|t}$ 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 $\mathbf{X}_t$:

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

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

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

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

$$\mathbf{M} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{6\times 6} & \mathbf{0}_{6\times 6k} \\ \mathbf{0}_{6k\times 6} & \left(\mathbf{W} - \mathbf{K}\mathbf{D}_1\mathbf{W} - \mathbf{K}\mathbf{D}_2\right)|_{(1:6k,1:6k)} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{K}\left(\mathbf{D}_1\mathbf{W} + \mathbf{D}_2\right)|_{(1:6k,1:6(k+1))} \end{bmatrix} \quad (49)$$

$$\mathbf{N} = \begin{bmatrix} \mathbf{R}_2 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{K}\mathbf{D}_1\mathbf{U}|_{(1:6k,1:6)} \end{bmatrix} \quad (50)$$

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 (47)-(48) and the equation for the Kalman-gain matrix below:

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

until convergence.