Perceptions about Monetary Policy

Bjarne Horst* Hannes Twieling†

September 14, 2026

Abstract

Perceptions about future monetary policy depend on expectations of policy-relevant variables and beliefs about the policy rule itself. Empirically, both violate full-information rational expectations: expectations of key variables deviate from FIRE, and perceived policy rules diverge from actual ones. These deviations generate endogenous monetary surprises, which we decompose into two channels: belief misperceptions, arising from incorrect expectations about policy-relevant variables, and rule misperceptions, arising from incorrect beliefs about the policy rule itself. We develop a model that generates these endogenous surprises and show that it closely matches their empirical counterparts. The model reveals that perceptions about how strongly the central bank reacts to policy-relevant variables govern the sign of inflation's response to a monetary policy shock — weak perceived reactions generate a positive response, resolving the price puzzle. Further, the model resolves the forward guidance puzzle.

JEL Codes: E31, E32, E52
Keywords: Monetary Policy, Subjective Expectations, Perceived Taylor Rule, Belief Misperceptions, Price Puzzle, Forward Guidance Puzzle

---

*University of Mannheim, bjarne.horst@uni-mannheim.de.

†Bank of Canada, hannes.twieling@gmx.net.

‡The authors gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project C02) and by Stiftung Geld und Währung. We thank Klaus Adam, Oliver Pfäuti, Rodrigo Sekkel, Miguel Santana, Christopher Sutherland, Tao Wang, Philipp Wangner for helpful comments and suggestions. We also thank participants of the BBL at the Bank of Canada. The views expressed herein are those of the authors and not necessarily those of the Bank of Canada.

I. Introduction

Empirical evidence suggests that private agents are frequently surprised by central banks' interest rate policies. To forecast the central bank's future decisions, private agents need to forecast the policy-relevant variables on which the central bank bases its policies, and they need to understand the mapping from policy-relevant variables into the policy variable, i.e., the policy rule. Whenever they fail in one of these dimensions, policy decisions deviate from prior beliefs, even when the central bank adheres perfectly to it's policy rule. This raises the following questions: Do agents actually hold correct beliefs about future monetary policy in the data? And if not, what are the economic consequences of their misperceptions about monetary policy?

This paper addresses these questions by studying misperceptions about monetary policy in the data and in theory. First, we show that agents' forecasts about policy-relevant variables violate FIRE and that their perceived monetary policy rules are incorrect. As a result, their perceptions of monetary policy do not align with the actual law of motion, due to two types misperceptions: *belief misperceptions* and *rule misperceptions*.

Second, we show that the presence of these misperceptions generates endogenous *monetary policy surprises*.[1] Intuitively, if agents hold incorrect perceptions about monetary policy, they are surprised by movements in the policy rate whenever the central bank reacts to economic developments, generating surprises that are endogenous. We characterize these surprises within a New Keynesian model, decompose them into belief and rule misperceptions, and show that their dynamics align with their empirical counterparts. Empirically we find that 39% of the variation in forecast errors in the federal funds rate can be explained by misperceptions about beliefs or the rule.

Third, we study the implications of these findings within the model. We show that when the perceived or actual reaction of the central bank to policy-relevant variables is sufficiently weak, the model generates a positive correlation between the policy rate and inflation; for sufficiently strong reactions, this sign flips. The model can therefore theoretically resolve the price puzzle. Finally, we show that forward guidance shocks in our model do not generate a forward guidance puzzle. The strength of forward guidance shock depends on the perceived persistence of monetary policy, the credibility of the central bank, and discounting of future utility. The latter breaks the forward guidance puzzle.

[1] We use the term *monetary surprise* to refer specifically to these endogenous surprises throughout the paper. This is distinct from a *monetary policy shock*, which we define, as is standard in the New Keynesian literature, as an exogenous deviation from a Taylor-type rule. Surprises arise endogenously from agents' misperceptions, whereas shocks are exogenous disturbances to the policy rule itself.

We first examine the violation of FIRE with respect to policy-relevant variables and monetary policy rules. On the former, we use local projections to show that forecast errors of policy relevant variables are non zero in response to observed realizations in policy relevant variables, indicating a violation of FIRE and sluggish updating. On the policy rule side, we show that a perceived rule, estimated on forecast data, does not align with an actual rule estimated on the same set of variables using realized data. A formal test of equality between the two rules is strongly rejected, and rerunning the test on alternative forecast data confirms our findings.

Building on these findings, we construct a New Keynesian model. The setting follows the standard New Keynesian framework, augmented with two additional assumptions. (i) FIRE is violated: we assume that agents update expectations over variables external to their optimization problem using a Bayesian learning framework.[2] Further, agents are internally rational, in that they correctly anticipate their own optimality conditions and budget constraints, not only in the current period, but also when forming expectations about future choices (Adam and Marcet, 2011; Adam et al., 2017, 2025). (ii) Agents' expectations over future nominal rates set by the central bank follow a perceived Taylor rule, which need not coincide with the actual Taylor rule. Under these assumptions, current household consumption depends on expected future real interest rates, which in turn depend on the perceived Taylor rule. Perceptions about the evolution of policy-relevant variables and about the policy rule itself therefore directly affect the household's consumption-savings decision.

As FIRE is violated and the perceived and actual Taylor rules do not coincide, any movement in policy-relevant variables generates a monetary policy surprise on the part of agents. Any fundamental shock, for instance a standard productivity shock, can therefore be decomposed into a surprise component, capturing the endogenous monetary surprise, and a residual, which we refer to as the pure component. The monetary policy surprise can, in turn, be further decomposed into surprises attributable to belief misperceptions and rule misperceptions. We show that the monetary surprise and its two subcomponents inherit the dynamics of a standard monetary policy shock: an unanticipated decrease in the nominal rate leads to an increase in both inflation and output.

We then turn to an empirical investigation of these findings. Using the estimated perceived and actual rules together with the model's decomposition, we construct time series for the belief and rule misperception terms and implement the decomposition within a regression framework.

[2]For instance, for households, prices (inflation) constitute an external variable, while consumption is a choice variable. Households therefore rely on Bayesian learning to update their expectations about inflation.

Specifically, we regress forecast errors of the nominal rate on the belief and rule misperception terms; the residuals of these regressions yield a monetary policy shock. The regression reveals, that 39% of the variation in the forecast errors can be attributed to misperception terms, of which the belief misperception clearly dominates, and 61% to monetary policy shocks. Using the results from this decomposition, we first show that the dynamics of this monetary policy shock match those observed in the model. Second, we perform a variance decomposition on the two misperception terms and find that their relative variance shares qualitatively match those implied by the model. Finally, we directly examine the responses of endogenous monetary policy surprises: we identify productivity shocks in a VAR following Gali (1999) and add nominal rate forecast errors to the system. This setup allows us to identify the forecast error response to a productivity shock mirroring the endogenous monetary policy surprise from our model. We find that the forecast error response aligns with our theoretical finding: in response to a positive productivity shock we observe an expansionary monetary policy surprise.

Finally, we study the implications of belief and rule misperceptions for the conduct of monetary policy. First, we show that the impact response of inflation to a standard monetary policy shock depends on the perceived and actual rules. If the perceived or actual reaction of the central bank is too weak, the monetary policy shock correlates positively with inflation; intuitively, agents do not believe that the central bank will control inflation. If, on the other hand, the reaction is sufficiently strong, the correlation is negative and yields the commonly known response of inflation to a monetary policy shock. Our model can therefore theoretically explain the price puzzle. We test this hypothesis empirically: using our theoretical results together with time-varying estimates of perceived and actual Taylor rule coefficients (Bauer et al., 2024), we derive a time series indicating whether the price puzzle is present. Our findings indicate that a price puzzle is present in roughly 25% of our sample. Furthermore, using the time series in a local projections exercise and interacting it with the monetary policy shock confirms our theoretical results.

We then turn to the impact response to forward guidance shocks in our model. This response depends on agents' beliefs about the future rule deviations announced at a given point in time: higher central bank credibility (i.e., less noise in the signal or announcement) and greater persistence in the perceived rule both increase the inflation response to a forward guidance shock. Further, stronger discounting weakens the effect of interest rate shocks announced to occur further in the future. This resolves the forward guidance puzzle in our model. Finally, in an extension, we show how multiple announcements of the same event, made at several different points in time, can strengthen central bank credibility. Similarly, announcements made several periods ago that have not been repeated by the monetary authority can lose potency due to decay in memory.

Literature. Our findings on FIRE violation for policy relevant variables connect to a large empirical literature documenting the same phenomenon.[3]With respect to perceptions about monetary policy rules, Hamilton et al. (2011); Carvalho and Nechio (2014); Jia et al. (2023) estimate static rules on forecast data, while Bauer et al. (2024); Sutherland (2026) estimate time-varying coefficients for perceived monetary policy rules. This literature is purely empirical: we instead estimate a static rule and embed it within a structural New Keynesian model, providing a theoretical mapping between the model and the empirical findings. On the theoretical side, our paper most closely connects to Caballero and Simsek (2022), who study disagreements between a central bank and markets about the future evolution of productivity. They further provide a theoretical explanation for monetary surprises. In contrast, our paper focuses on misperception in beliefs about policy-relevant variables and the policy rule of monetary policy. The different modeling approach not only allows us to endogenously generate monetary policy surprises, but we can link them to misperceptions and use these findings to resolve the price and forward guidance puzzle. We further provide empirical evidence for our findings. On the theoretical side Kim (2026) studies deviations of the perceived from the actual monetary rule. However, his focus lies on rule misperceptions, while belief misperceptions are not addressed.[4] Sastry (2026) studies disagreement between markets and the central bank in terms of beliefs about fundamentals, the policy rule, and confidence about public signals, and quantifies these channels. In contrast, we embed our misperceptions in a standard New Keynesian framework and derive implectaions ofr monetary policy. With respect to expectations formations we build on the Bayesian learning literature and include internal rationality.[5] We adopt a solution approach proposed by Roschitsch and Twieling (2026). Finally, we connect to the literature documenting the price puzzle (Sims, 1980; Christiano et al., 1999) by providing an alternative theoretical explanation for this phenomenon. In contrast to other theoretical explanations that focus on the cost channel of monetary policy (Ravenna and Walsh, 2006; Rabanal, 2007), or asymmetric information (Tas, 2011), we rely on a standard New Keynesian model with perceived monetary policy rules and verify our findings empirically. These findings also align with Castelnuovo and Surico (2010). We also contribute a solution to the forward guidance puzzle, relating to McKay et al. (2016); Del Negro et al. (2023); Gabaix (2020), among others.

[3]Among others: Mankiw et al. (2003); Coibion and Gorodnichenko (2015); Bordalo et al. (2020); Kohlhas and Walther (2021); Angeletos et al. (2021).

[4]Amodeo (2025) also focuses on disagreement between the central bank and markets, while his contribution is mainly empirical, the paper does include a model, which differs from our approach and is focused solely on disagreement.

[5]See: Adam et al. (2017, 2025); Winkler (2020); Caines and Winkler (2019); Adam and Marcet (2011).

Outline. Section (II) provides empirical evidence for the violation of FIRE in policy-relevant variables, as well as misperceptions about the monetary policy rule. Section (III) introduces the model, and Section (IV) derives endogenous monetary surprises. In Section (V), we connect the model to the empirical findings, while Section (VI) presents the implications. Finally, Section (VII) concludes.

II. Empirical assessment of perceptions about monetary policy

In this section, we empirically study agents' perceptions of monetary policy. We focus on two aspects. First, we examine the formation of expectations about policy-relevant variables. If agents hold incorrect expectations about the future evolution of—for instance—inflation, they will, as a result, also incorrectly forecast the policy rate, which is a function of inflation, even if their perceived rule is correct. Second, we examine perceptions of the rule itself. Conversely, if agents form correct expectations about inflation, they will nonetheless make an error in their policy rate forecast if their perceived rule is incorrect.

II.A Perceptions about policy relevant variables

A large literature studies expectation formation with respect to macroeconomic variables. Most of these studies examine variables relevant to monetary policy decisions, and all conclude that FIRE is violated. We briefly review a subset of this literature. Mankiw et al. (2003) document that forecast errors in inflation are autocorrelated and therefore violate full rationality. Using consensus forecasts, Coibion and Gorodnichenko (2015) regress forecast errors of inflation on their revisions and reach a similar conclusion. Bordalo et al. (2020) focus on forecasts on the individual level, extending the set of variables to include several policy-relevant variables, such as GDP and the unemployment rate, and again reject FIRE. Finally, Angeletos et al. (2021) study the impulse response functions of forecast errors and revisions to different shocks, reaching the same conclusion.

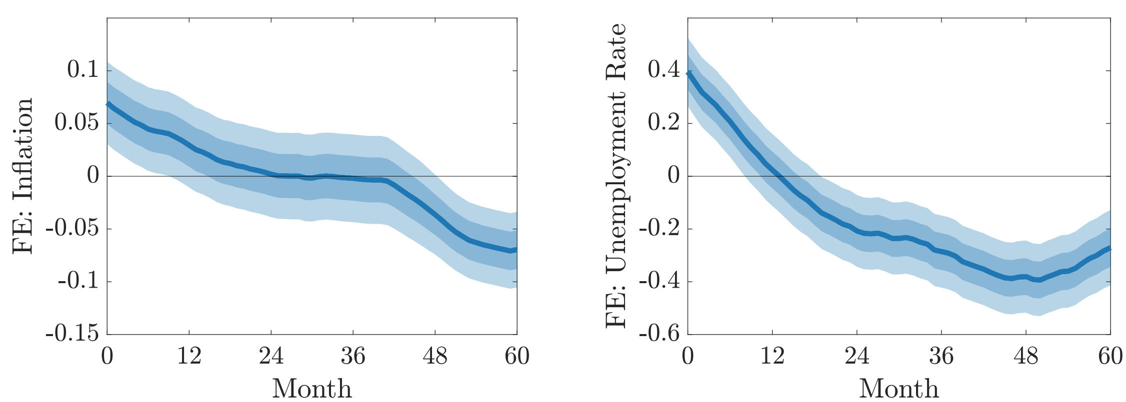
In our context, we focus specifically on policy-relevant variables. A straightforward way to show that FIRE is violated is to study the response of forecast errors in these variables to innovations in observed variables (Adam et al., 2025). We do so using a local projections exercise:

$$FE_{t+12+h} = c_h + \beta_h \times \Delta z_t + u_t$$

The forecast error of changes in a variable z is based on one-year-ahead forecasts and is given by $FE_{t+12} = \Delta z_{h,t+12} - \mathbb{E}_t^{\mathcal{P}} \Delta z_{h,t+12}$, where Δz_t denotes observed changes in the variable z. Under FIRE, the estimates of β_h should be zero, as observed changes are already contained in today's information set and should not systematically affect forecast errors. We run the local projections for annual inflation (z being log CPI) and for annual changes in the unemployment rate. Both are important policy-relevant variables under the dual mandate of the FED. We rely on these variables throughout the paper. The frequency is monthly, and forecast errors are computed from Blue Chip forecasts, which we use throughout the paper as our baseline.

Figure (1) shows the results. In both cases, forecast errors increase significantly in the first month after the innovation, before turning negative toward the end of the horizon. This indicates, first, that FIRE is clearly violated for both variables, and second, that agents update expectations sluggishly. Initially, agents revise their expectations insufficiently, while toward the end of the horizon they overshoot. These dynamics are consistent with extrapolative updating behavior, which we incorporate into our model below.

Figure 1: Forecast error responses to observed innovations



Notes: Forecast error responses of inflation and unemployment changes to their respective observed changes. Confidence Intervals are Newey-West (one std and 95%).

II.B Perceptions about the policy rule

In addition to FIRE violations with respect to policy-relevant variables, agents may also misperceive the policy rule. To test this, we estimate the following regressions for the perceived and actual rules. Each rule specifies the nominal rate as the dependent variable, with inflation, the output gap, and the lagged nominal rate as regressors. The specification is motivated by the original Taylor rule proposed in Taylor (1993).[6] Equation (1) denotes the perceived rule: it includes one-year-ahead expectations of the interest rate (i), inflation (π), and the output gap ($\tilde{y}$), together with the current nominal rate. Perceived coefficients are denoted with a bar above them. Time is measured monthly throughout the paper unless otherwise specified. The actual rule, equation (2), has the same structure but uses realized data in place of forecasts.

$$\mathbb{E}_t^{\mathcal{P}} i_{t+12} = \bar{c} + (1-\bar{\rho})[\bar{\phi}^{\pi}\mathbb{E}_t^{\mathcal{P}}\pi_{t+12} + \bar{\phi}^{\tilde{y}}\mathbb{E}_t^{\mathcal{P}}\tilde{y}_{t+12}] + \bar{\rho} i_t + u_t \tag{1}$$

$$i_t = c + (1-\rho)[\phi^{\pi}\pi_t + \phi^{\tilde{y}}\tilde{y}_t] + \rho i_{t-12} + u_t \tag{2}$$

To assess whether the coefficients of the two rules align, we run a seemingly unrelated regression. For the forecast data, we use Blue Chips forecasts on the FFR, CPI year-on-year inflation, and the unemployment rate as a proxy for the output gap. For the realized data, we use the same data only as realizations. Table (1) shows the results. The estimated coefficients, for both the actual and perceived rule, are consistent with the literature: inflation coefficients exceed one, the coefficient on the unemployment rate is negative and smaller in absolute value than the inflation coefficient, and the coefficients on the lagged nominal rate lie between zero and one. Further, the R^2 is high, indicating that the rules capture a large share of the variation in the nominal rate. The private sector perceives the central bank to be more hawkish than it actually is, manifesting in the larger inflation weight and the lower weight on economic activity. We also conduct a formal test of whether the coefficients align across rules. Specifically, the null hypothesis H_0 tests whether $\bar{\phi}^{\pi} = \phi^{\pi} \wedge \bar{\phi}^{\tilde{y}} = \phi^{\tilde{y}} \wedge \bar{\rho} = \rho$. The test strongly rejects H_0, indicating that the perceived and actual rules differ. In Appendix (A), Table (A.1), we re-estimate the regression using SPF forecasts. Our results remain robust.

[6]One could potentially add forward-looking terms to the policy functions. A mapping into our theoretical model is also possible. We rely on contemporary and backward-looking rules because we can capture a large share of the variation using these rules (see the R^2 of the SUR regression). Further, forward-looking rules for the central bank would require us to switch to a quarterly frequency to allow for central bank forecasts, decreasing the number of observations.

Table 1: Perceived and Actual Policy Rules

	Perceived	*Actual*
ϕ^{π}	2.38	2.05
	(0.44)	(0.45)
$\phi^{\tilde{y}}$	−0.54	−0.83
	(0.19)	(0.27)
ρ	0.80	0.58
	(0.03)	(0.10)
R^2	0.97	
Joint test : Perceived = Actual	0.00	
Frequency	*monthly*	
Sample	1984 – 2020	

Notes: The table reports estimates of the perceived and actual Taylor rules, given by equations (1) & (2). Rules are estimated jointly as a seemingly unrelated regression (SUR). Standard errors are Driscoll-Kraay. *Joint test: Perceived = Actual* reports the p-value of a Wald test of $H_0 : \bar{\phi}^{\pi} = \phi^{\pi} \wedge \bar{\phi}^{\tilde{y}} = \phi^{\tilde{y}} \wedge \bar{\rho} = \rho$.

III. The Model

In this section, we introduce a New Keynesian model that allows for perceptions about monetary policy along two dimensions: beliefs and the policy rule. The setup mirrors the standard New Keynesian model, with two departures: (i) FIRE is violated, and (ii) agents hold a perceived rule for monetary policy that need not coincide with the actual rule.

III.A The Economy

Households. A representative household derives utility from consumption and leisure. The preferences are as follows:

$$\mathbb{E}_0^{\mathcal{P}} \sum_{t=0}^{\infty} \beta^t \left(\frac{c_t^{1-\sigma}}{1-\sigma} - \frac{\chi n_t^{1+\varphi}}{1+\varphi} \right)$$

where $\mathbb{E}_0^{\mathcal{P}}$ is the subjective expectations operator discussed below. c_t and n_t denote consumption of varieties and hours worked, respectively. The household budget constraint is then given by

$$c_t + b_{t+1} = (1+r_t)b_t + \underbrace{w_t n_t + \Sigma_t}_{total\ income = x_t}$$

Σ_t are firm profits and government taxes and transfers; w_t is the real wage; and b_t is a one-period nominal zero-coupon bond. Total income, x_t, is defined as the sum of labor income, profits, and transfers.

Firms. There is a continuum of monopolistically competitive firms that face Rotemberg-type price adjustment costs. The present value of firm profits is given by

$$\mathbb{E}_0^{\mathcal{P}} \sum_{t=0}^{\infty} \beta^t \Lambda_t \Big[P_t(j) y_t(j) - (1-\tau_t^l) W_t n_t(j) - P_t y_t \frac{\kappa}{2} \Big(\frac{P_t(j)}{P_{t-1}(j)} - 1 \Big)^2 \Big]$$

An individual firm j chooses it's price $P_t(j)$, facing demand $y_t(j)$, and using labor input $n_t(j)$. Aggregation of goods varieties follows a CES structure. The real wage is given by $w_t = W_t/P_t$, and τ_t^l denotes a subsidy chosen such that the steady state is efficient, but is subject to a shock generating a cost-push shock. Firms face quadratic costs of adjusting their prices scaled by κ. y_t and P_t denote aggregate production and the price level. Λ_t is the stochastic discount factor implied by the household's Euler equation. Firms face the following demand and production function

$$y_t(j) = \Big(\frac{P_t(j)}{P_t} \Big)^{-\epsilon} y_t, \quad y_t(j) = a_t n_t(j)^{1-\alpha}$$

a_t is aggregate productivity following an AR(1) process.

Monetary policy. The central bank sets the nominal interest rate according to a standard Taylor rule that mirrors equation (2):

$$\widehat{i}_t = (1-\rho)[\phi^{\pi} \widehat{\pi}_t + \phi^{\tilde{y}} \tilde{y}_t] + \rho \widehat{i}_{t-1} + \widehat{\epsilon}_t^{MP} \tag{3}$$

Equation (3) is expressed in log-linear terms where variables with a "head" denote log-linear deviations from its' steady-state. The Taylor rule also includes a standard monetary policy shock $\widehat{\epsilon}_t^{MP}$.

III.B Expectation formation

We model expectations following the Bayesian learning literature (e.g. Adam and Marcet, 2011; Adam et al., 2017; Winkler, 2020).[7] This approach allows for sluggish updating of expectations aligning with our findings in Section (II) and empirical evidence presented in Coibion and

[7] Equivalent expectation updating rules can be derived using rational inattention approaches (Pfäuti, 2023, 2025).

Gorodnichenko (2015); Angeletos et al. (2021).

General setup. Agents, households and firms, are endowed with a set of beliefs, in the form of a probability measure over the full sequence of variables they take as given, henceforth external variables. For households, these are the real rate and total income, $(r_t, x_t)_{t\geq 0}$; for firms, marginal costs, $(mc_t)_{t\geq 0}$.[8] We denote this measure by $\mathcal{P}$. Rational expectations arise as a special case of this setup, in which households' beliefs coincide with the objective, or equivalently the "true" or "equilibrium-implied", distribution of the external variables, $\mathcal{P} = \mathbb{P}$. Although agents may hold beliefs that are generally inconsistent with the equilibrium-implied conditional distribution of the external variables, it is worth emphasizing that agents nonetheless (i) hold a time-consistent set of beliefs and (ii) behave optimally given these beliefs—that is, agents are *internally rational* in the sense of Adam and Marcet (2011). Moreover, it is not common knowledge among agents that all households share identical beliefs and preferences, so agents cannot discover the misspecification of their beliefs, $\mathcal{P} \neq \mathbb{P}$, through inductive reasoning about the structure of the economy. Given the observed path of the external variables up to period t, agents use this information together with $\mathcal{P}$ to form a conditional expectation over the continuation sequence of external variables, denoted $\mathbb{E}_t^{\mathcal{P}}$. $\mathbb{E}_t$ denotes the FIRE expectations operator.

Agents entertain the idea that any external variable z follows a simple state-space model:

$$
\begin{aligned}
\ln \frac{z_{t+1}}{z_t} &= \ln m_{z,t+1} + \ln e_{z,t+1} \\
\ln m_{z,t+1} &= \varrho_z \ln m_{z,t} + \ln v_{z,t+1}, \qquad \varrho_z \in (0,1) \\
\begin{pmatrix} \ln e_{z,t} & \ln v_{z,t} \end{pmatrix}' &\sim \mathcal{N}\left(\begin{pmatrix} -\frac{\sigma_{z,e}^2}{2} & -\frac{\sigma_{v,e}^2}{2} \end{pmatrix}', \begin{pmatrix} \sigma_{z,e}^2 & 0 \\ 0 & \sigma_{v,e}^2 \end{pmatrix} \right)
\end{aligned}
\tag{4}
$$

Hence, agents perceive growth rates of z as the sum of a transitory, and a persistent component. Crucially, $\ln e_{z.t}$ and $\ln v_{z,t}$ are not observable to the agents, rendering $\ln m_{z,t}$ unobservable. Agents apply the optimal Bayesian filter, that is, the Kalman filter, to arrive at the observable system.[9]

Lemma 1 (External variables expectations updating)**.** *Applying the Kalman filter to the state-space*

[8] Total income actually comprises two external variables, wages and profits and transfers, and one internal variable, labor supply. We combine these into a single external variable, as households typically report expectations about their income rather than its individual sub-components. This assumption is innocuous: given expectations over the external variables, the internal variable can always be recovered via the household's labor supply condition, which automatically satisfies internal rationality.

[9] We assume that the agents' prior variance equals the steady-state Kalman variance.

model and log-linearizing around the non-stochastic steady state gives

$$\mathbb{E}_t^{\mathcal{P}} \widehat{z}_{t+s} = \widehat{z}_t + \frac{1-\varrho_z^s}{1-\varrho_z} \varrho_z \widehat{\overline{m}}_{z,t} \tag{5}$$

$$\widehat{\overline{m}}_{z,t} = \varrho_z (1-g_z) \widehat{\overline{m}}_{z,t-1} + g_z \Delta \widehat{z}_t, \tag{6}$$

where $\ln \overline{m}_{z,t} := \mathbb{E}_t^{\mathcal{P}}(\ln m_{z,t})$ *is the posterior mean,* $g_z = \frac{\varrho_z^2 \sigma_z^2 + \sigma_{z,v}^2}{\varrho_z^2 \sigma_z^2 + \sigma_{z,v}^2 + \sigma_{z,e}^2}$ *is the steady-state Kalman filter gain,* $\sigma_z^2 = \frac{1}{2\varrho_z^2}[-(\sigma_{z,v}^2 + (1-\varrho_z^2)\sigma_{z,e}^2) + \sqrt{(\sigma_{z,v}^2 + (1-\varrho_z^2)\sigma_{z,e}^2)^2 + 4\varrho_z^2 \sigma_{z,v}^2 \sigma_{z,e}^2}]$ *is the steady-state Kalman filter uncertainty.*

Proof. See Appendix (B.1) for the application of the Kalman filter. Log linearization around the steady state gives the result. ∎

In short, today's expectations about the future evolution of a variable z depend on the current z and the posterior expectations of its persistent component $\widehat{\overline{m}}_{z,t}$. The latter in turn depends on prior expectations, $\widehat{\overline{m}}_{z,t-1}$, and an updating part, $\Delta \widehat{z}_{i,t}$, which constitutes observed changes in the variable z.

Perceived Taylor Rule. Further we assume that agents, in our model this is only relevant for households, form expectations about monetary policy using a perceived Taylor rule:

$$\mathbb{E}_t^{\mathcal{P}} \widehat{i}_{t+1} = (1-\bar{\rho})[\bar{\phi}^{\pi} \mathbb{E}_t^{\mathcal{P}} \widehat{\pi}_{t+1} + \bar{\phi}^{y} \mathbb{E}_t^{\mathcal{P}} \tilde{y}_{t+1}] + \bar{\rho} \widehat{i}_t \tag{7}$$

Equation (7) mirrors equation (1). As in our empirical setting, the perceived and actual rules share the same structure, though their coefficients may differ. Importantly, note that the expectations of the nominal rate depend, in turn, on expectations of inflation and the output gap. These are external variables from the point of view of the household and therefore evolve according to equations (5) and (6). This formulation makes apparent how expectations about policy-relevant variables and perceptions about the rule jointly shape expectations of the policy rate. Below, we show how this directly affects the household's saving and consumption decisions.

III.C Equilibrium and solution approach

Equilibrium. Before turning to the solution approach, it is important to note that the steady state coincides with that of the standard New Keynesian model. The reasoning is as follows: in the non-stochastic steady state, expectations play no role, and the model is therefore equivalent to the standard New Keynesian model. We adopt the equilibrium concept of internally rational expectations equilibrium, as defined in Adam and Marcet (2011):

Definition 1 (Internally rational expectations equilibrium)**.** *An IREE consists of three bounded stochastic processes: shocks* $(\epsilon_t^j)_{t\geq 0}^{j\in\{a,\mu,MP\}}$*, allocations* $(c_t, b_t, n_t)_{t\geq 0}$*, and prices* $(w_t, i_t, P_t(j), \Sigma_t)_{t\geq 0}$*, such that in all t*

1. *households choose* c, n, b *optimally, given their beliefs* $\mathcal{P}$*,*
2. *firms set prices optimally, given their beliefs* $\mathcal{P}$*,*
3. *the monetary authority acts according to the Taylor rule (3)*
4. *markets for consumption goods and hours clear, given the prices, and the households and firms budget constraints hold.*

Optimal household behavior. In the previous section, we described how expectations about external variables are formed. However, to solve the model, agents must also form expectations about internal variables or variables that correspond to future choices. Ensuring that these expectations satisfy internal rationality requires expressing internal choices as functions of external variables while taking into account agents' optimality conditions and budget constraints. To do so, we rely on the solution approach of Roschitsch and Twieling (2026).

On the household side, an internal variable about which agents form expectations is future consumption. This follows directly from the log-linearized Euler equation:

$$\widehat{c}_t = -\frac{1}{\sigma}\mathbb{E}_t^{\mathcal{P}}\widehat{r}_{t+1} + \mathbb{E}_t^{\mathcal{P}}\widehat{c}_{t+1}$$

Iterating on the Euler equation yields

$$\widehat{c}_t = -\frac{1}{\sigma}\sum_{s\geq 0}\mathbb{E}_t^{\mathcal{P}}\widehat{r}_{t+s+1} + \mathbb{E}_t^{\mathcal{P}}\widehat{c}_{\infty}, \tag{8}$$

where we define $\mathbb{E}_t^{\mathcal{P}}\widehat{c}_{\infty} = \lim_{s\to\infty}\mathbb{E}_t^{\mathcal{P}}\widehat{c}_{t+s}$. In our model, consumption is the only forward-looking internal variable on the household side, so $\mathbb{E}_t^{\mathcal{P}}\widehat{c}_{\infty}$ is the only remaining internal expectation term. It can be shown that $\mathbb{E}_t^{\mathcal{P}}\widehat{c}_{\infty}$ is a function of only external variables, so it suffices to solve for this term to characterize the household's decisions (see Lemma 2). Intuitively, $\mathbb{E}_t^{\mathcal{P}}\widehat{c}_{\infty}$ captures a subjective expectations-driven wealth effect: suppose income is expected to increase; through extrapolation, income is then expected to be higher as $s\to\infty$, which in turn implies an expected increase in wealth. Consequently, expected future consumption, $\mathbb{E}_t^{\mathcal{P}}\widehat{c}_{\infty}$, rises, which raises consumption today via equation (8).[10] Using the solution approach proposed by Roschitsch and Twieling (2026) we

[10]Note that under FIRE, $\mathbb{E}_t\widehat{c}_{\infty} = 0$ always holds, as agents understand that, following a transitory shock, the

derive the following Lemma 2. Note that we explicitly use the households' first-order conditions and budget constraint, implying that decisions are time-consistent and optimal, and thereby internal rationality is satisfied.

Lemma 2 (Subjective expectations wealth effect)**.** *Using the households' first-order conditions and the budget constraint, the subjective expectations wealth effect is given by*

$$\mathbb{E}_t^{\mathcal{P}} \widehat{c}_\infty = \frac{1-\beta}{\beta}\Big(\widehat{b}_{t+1} + \sum_{s=1}^{\infty} \beta^s \big(\frac{1}{\sigma} \sum_{n=s}^{\infty} \mathbb{E}_t^{\mathcal{P}} \widehat{r}_{t+n+1} + \mathbb{E}_t^{\mathcal{P}} \widehat{x}_{t+s}\big)\Big) \tag{9}$$

Proof. See Appendix (B.2). ∎

Optimal firm behavior. On the firm side, we derive the typical Phillips curve structure.[11]

$$\widehat{\pi}_t = \beta \mathbb{E}_t^{\mathcal{P}} \widehat{\pi}_{t+1} + \frac{\epsilon - 1}{\kappa} \Theta \left(\widehat{mc}_t^r + \mu_t\right), \qquad \Theta \equiv \frac{1-\alpha}{1-\alpha+\alpha\epsilon},$$

Marginal costs, $\widehat{mc}_t^r$, and the cost-push shock, $\widehat{\mu}_t$, are given by

$$\widehat{mc}_t^r = \widehat{w}_t - \widehat{a}_t + \alpha \widehat{n}_t, \quad \widehat{\mu}_t = \rho_\mu \widehat{\mu}_{t-1} + \widehat{\varepsilon}_t^\mu$$

Note that, since firms are price setters, $\mathbb{E}_t^{\mathcal{P}} \widehat{\pi}_{t+1}$ constitutes an internal choice from the point of view of the firm. As with households, iterating on the Phillips curve yields

$$\widehat{\pi}_t = \frac{\epsilon - 1}{\kappa} \Theta \mathbb{E}_t^{\mathcal{P}} \sum_{n=0}^{\infty} \beta^n \left(\widehat{mc}_{t+n}^r + \mu_{t+n}\right) \tag{10}$$

This formulation depends only on expectations of external variables, so that current inflation is pinned down entirely by firms' expectations of the discounted sum of future marginal costs and cost-push shocks, without any remaining reference to firms' own future pricing decisions.[12]

III.D Transmission and general representation

Perceptions about monetary policy and household decisions. To understand how perceptions about monetary policy affect household consumption decisions, we first define the Fisher equation

economy will eventually converge back to the steady state.

[11] We assume that firms' expectations of their own price align with their expectations of the future price level. Therefore, all firms set the same price.

[12] Note that $\beta^\infty \mathbb{E}_t^{\mathcal{P}} \widehat{\pi}_\infty = 0$ for standard calibrations of the discount factor. Therefore, the terminal inflation term drops out.

as follows:

$$\mathbb{E}_t^{\mathcal{P}} r_{t+s} = \mathbb{E}_t^{\mathcal{P}} i_{t+s-1} - \mathbb{E}_t^{\mathcal{P}} \pi_{t+s}$$

Substituting the Fisher equation, the perceived rule (7), and the subjective expectations wealth effect (9) into the Euler equation (8), household consumption can be expressed as follows:

$$\widehat{c}_t = -\frac{1}{\sigma}\Big[\frac{1}{1-\bar{\rho}\beta}\widehat{i}_t - \frac{\varrho_\pi(1-\bar{\rho}\beta)-(1-\bar{\rho})\bar{\phi}^\pi\varrho_\pi\beta}{(1-\bar{\rho}\beta)(1-\varrho_\pi\beta)}\widehat{m}_t^\pi + \frac{(1-\bar{\rho})\bar{\phi}^{\tilde{y}}\beta}{(1-\bar{\rho}\beta)(1-\beta)}\Big(\tilde{y}_t + \frac{\varrho_{\tilde{y}}}{1-\beta\varrho_{\tilde{y}}}\widehat{m}_t^{\tilde{y}}\Big)\Big] + \Big(\widehat{x}_t + \frac{\varrho_x}{1-\beta\varrho_x}\widehat{m}_{x,t}\Big) + \frac{1-\beta}{\beta}\widehat{b}_{t+1} \quad (11)$$

Households' current consumption depends on today's nominal rate, the output gap, income, and bond holdings, as well as on posterior expectations about inflation, the output gap, and income. The presence of posterior expectations about inflation and the output gap indicates that expectations of policy-relevant variables directly affect household consumption. Similarly, equation (11) depends on the perceived rule coefficients $(\bar{\phi}^\pi, \bar{\phi}^{\tilde{y}}, \bar{\rho})$, so that perceptions about the policy rule map directly into household consumption. The intuition is straightforward: household consumption depends on the entire expected sequence of future real rates, which, in turn, depends on both the perceived policy rule and expectations of policy-relevant variables.

The role of both channels can be seen directly. Suppose $\varrho_\pi(1-\bar{\rho}\beta) < (1-\bar{\rho})\bar{\phi}^\pi\varrho_\pi\beta$, which holds for sufficiently high $\bar{\phi}^\pi$; in this case, an expected increase in inflation decreases consumption, holding everything else constant. We discuss below what happens when this condition is violated. Similarly, a higher expected output gap decreases consumption, holding everything else constant. On the perceived-rule side, an increase in the perceived rule coefficients $(\bar{\phi}^\pi, \bar{\phi}^{\tilde{y}})$ leads to a more negative reaction of consumption to expected increases in inflation or the output gap.

General representation. Given household consumption equation (11) and the firms Phillips Curve (10) we can derive a general representation of the model in the typical IS & Phillips Curve representations.

Lemma 3 (General New Keynesian Representation)**.** *The IS equation and the Phillips Curve are respectively given by:*

$$\widehat{i}_t = C_\pi\,\widehat{m}_t^\pi - C_{\tilde{y}}\Big(\tilde{y}_t + \tfrac{\rho_{\tilde{y}}}{1-\beta\rho_{\tilde{y}}}\,\widehat{m}_{\tilde{y},t}\Big) + C_x\,\widehat{m}_{y,t} \quad (12)$$

$$\widehat{\pi}_t = \bar{\kappa}\left(\tilde{y}_t + \frac{\beta\varrho_c}{1-\beta\varrho_c}\widehat{m}_{\tilde{y},t}\right) + \kappa_\mu\,\widehat{\mu}_t \quad (13)$$

Proof. See Appendix (B.3). ■

Equations (12) and (13), together with the Taylor rule (3) and the respective expectation updating equations summarize the model.

III.E Calibration

The calibration on the household side with respect to the discount factor, the inverse of the intertemporal elasticity of substitution, and the inverse Frisch elasticity are standard and taken from the literature. The same holds for the elasticity of substitution, the degree of decreasing returns to labor, and price adjustment costs on the firm side.

On the expectations side, we set the perceived autocorrelation of inflation and income, or output, to 0.97 and 0.95, respectively.[13] We set the Kalman gain to 0.25 for both of these variables. This aligns with Pfäuti (2023), who uses an equivalent expectation updating formulation. On the firm side, we assume that the perceived autocorrelation is 0.85 and the Kalman gain is 0.1. Further, we assume that households forecast the output gap by separately forecasting output and the efficient level of output. For the latter, we assume the same perceived autocorrelation as output, but as the efficient level of output is unobservable and therefore hard to forecast, we set the Kalman gain to one tenth of that of output, to reflect the high degree of noise surrounding this variable.

The actual and perceived rule coefficients are estimated as described in Section (II). We transform the coefficients for the lagged interest rate $(\rho, \bar{\rho})$ to monthly frequency, and use Orkun's law to transform the estimates for the unemployment rate to output gap coefficients. All coefficients are described in table (2).

[13] We assume total income and output follow the same updating process.

Table 2: Model calibration

Parameter		Value	Description	Source
Households	φ	1.25	inverse Frisch elasticity	standard
	σ	1.0	inverse of intertemporal EOS	standard
	β	0.998	discount factor	standard for monthly frequency
Production	ϵ	2.0	EOS	standard
	α	0.66	returns to scale	standard
	κ	250	price adjustment costs	standard
Expectations	ϱ^{π}	0.97	autocorrelation of plm	literature
	ϱ^{x}	0.95	autocorrelation of plm	literature
	ϱ^{c}	0.85	autocorrelation of plm	calibrated
	g	0.25	Kalman gain: π, y	literature
	g^{c}	0.1	Kalman gain mc	calibrated
Policy Rule	$\bar{\rho}$	0.98	perceived autocorrelation	estimated
	ρ	0.94	autocorrelation	estimated
	$\bar{\phi}^{\pi}$	2.4	perceived inflation	estimated
	ϕ^{π}	2.1	inflation	estimated
	$\bar{\phi}^{y}$	0.27	perceived output	estimated
	ϕ^{y}	0.42	output	estimated

Notes: The table reports model parameters and the respective calibration. One period in the model is one month.

IV. Endogenous monetary policy surprises

The model above deviated from FIRE in two ways: (i) agents' perceived law of motion of policy-relevant variables does not coincide with the equilibrium dynamics, (ii) perceived and actual monetary policy rules do not align. We now focus on how these features create endogenous monetary policy surprises. We then calibrate the model and generate these surprises in a numerical example.

IV.A Defining endogenous monetary policy surprises

We define a monetary policy surprise as the forecast error of the nominal rate. Subtracting perceived from the actual rule, equations (7) and (3), we get:

$$\begin{aligned}\widehat{i}_t - \mathbb{E}^{\mathcal{P}}_{t-1}\widehat{i}_t = &\underbrace{(1-\rho)\Big[\phi^{\pi}(\widehat{\pi}_t - \mathbb{E}^{\mathcal{P}}_{t-1}\widehat{\pi}_t) + \phi^{\tilde{y}}(\tilde{y}_t - \mathbb{E}^{\mathcal{P}}_{t-1}\tilde{y}_t)\Big]}_{misperception:\ beliefs} \\ &+ \underbrace{\Big[(1-\rho)\phi^{\pi} - (1-\bar{\rho})\bar{\phi}^{\pi}\Big]\mathbb{E}^{\mathcal{P}}_{t-1}\widehat{\pi}_t + \Big[(1-\rho)\phi^{\tilde{y}} - (1-\bar{\rho})\bar{\phi}^{\tilde{y}}\Big]\mathbb{E}^{\mathcal{P}}_{t-1}\tilde{y}_t + (\rho-\bar{\rho})\widehat{i}_{t-1}}_{misperception:\ rule} + \widehat{\epsilon}^{MP}_t \end{aligned} \tag{14}$$

Equation (14) decomposes the forecast error of the nominal rate into belief and rule misperceptions. The former consists of the forecast errors of the policy-relevant variables, inflation and the output gap, scaled by the actual rule coefficients. The latter depends on the difference between the perceived and actual rule coefficients. The final term is the structural monetary policy shock.

Intuitively, forecast errors in the nominal rate can arise from a standard monetary policy shock, as in the standard New Keynesian model, but also from the two misperception terms. This implies that agents are surprised by the central bank's interest rate decisions even when the monetary policy shock is zero. Consequently, whenever belief and rule misperceptions are present, any shock to the model endogenously generates monetary policy surprises. Specifically, a FIRE violation with respect to policy-relevant variables produces forecast errors in these variables, giving rise to belief misperceptions, while a misalignment between the actual and perceived rule gives rise to rule misperceptions.

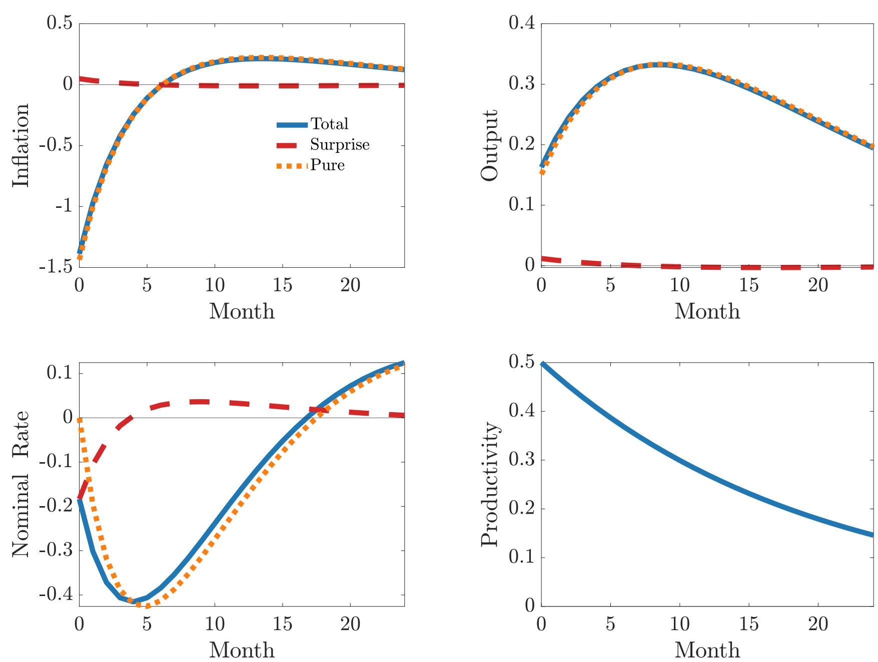
IV.B Numerical example

We now exemplify the emergence of endogenous monetary surprises in response to productivity shocks. Conceptually, we decompose the *total* model IRFs of output and inflation into two components: (i) a monetary *surprise* component, capturing the response of the economy if it had not been hit with a productivity shock, but had faced the same sequence of monetary surprises as those endogenously generated by the TFP shock, and (ii) a *pure* TFP shock component, capturing the response to the TFP shock if households had not been surprises by the monetary response. Formally, for the *surprise* component, we first compute the policy rate forecast error $\widehat{i}_t - \mathbb{E}^{\mathcal{P}}_{t-1}\widehat{i}_t$ induced by the TFP shock. We then feed the sequence of forecast errors into the model, while setting all other shocks, including the TFP shock to zero. For the *pure* TFP component, we simply substract the *surprise* component from the *total* IRF.

We further decompose the monetary surprise into the *belief* and *rule* misperceptions. For both of these components, we use equation (14) to compute the *belief* and the *rule* component of the total monetary suprise and feed the resulting sequences into the model, again setting all other shocks to zero.

Figure (2) plots the decomposition of the *total* response into the *surprise* and *pure* subcomponents. Turning to the *total* response to a productivity shock shown in blue, we see that an increase in productivity increases output and decreases inflation. The central bank responds by decreasing the nominal rate. This is the standard response to a productivity shock in a New Keynesian model. The *surprise* component shows a decrease in the policy rate, implying that agents expected a weaker response by the monetary authority, leading to an increase in inflation and output. These dynamics mirror the responses to a standard expansionary monetary policy shock. The *pure* component follows the total response: as inflation falls, the nominal rate is reduced and simultaneously output increases. Quantitatively, the surprise component is much weaker across all variables, implying that most of the total response is driven by the *pure* component. This does not mean that misperceptions about monetary policy are quantitatively unimportant in general: in Section (V) we show that they explain a large share of monetary surprises, both in the data and the model.

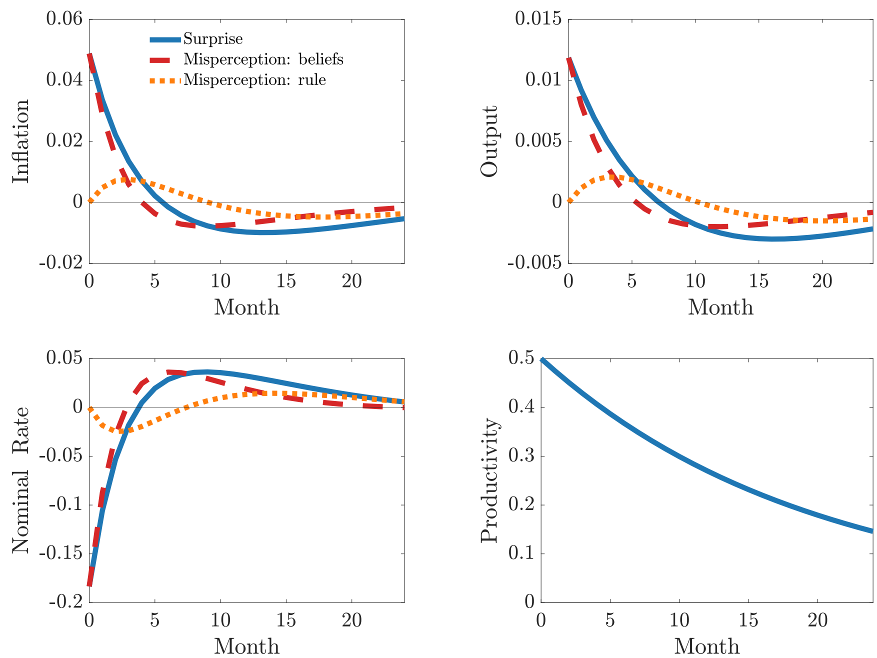
Figure 2: Decomposition of *total* response into *surprise* and *pure* component, productivity shock



Notes: Response to an expansionary productivity shock. Decomposition into monetary surprise and Pure component. The AR(1) coefficient of the shock is set to 0.95.

Decomposing the surprise component into *belief* and *rule* misperceptions, we find that, for both components, agents expected a higher rate, generating an expansionary surprise. This leads to an increase in inflation and output. As with the *surprise* component, the dynamics in both misperception terms mirror the dynamics of a standard expansionary monetary policy shock. Quantitatively, we find that the *belief* misperceptions dominate the *rule* misperceptions. This can be explained by to factors: first, the moderate difference in the estimated coefficients of the perceived and actual rule. Second, the *rule* misperception term is scaled by prior beliefs (see equation 2) and response in expectations is muted and sluggish. In summary, we find that the model can generate endogenous monetary surprises consisting of *belief* and *rule* misperceptions. All components, the overall surprise and the misperception terms, mirror the responses of a standard monetary policy shock. In Appendix (C) we show that the same holds for a cost-push shock.

Figure 3: Decomposition of *surprise* into *belief* and *rule* misperceptions, productivity shock



Notes: Response to an expansionary productivity shock. Decomposition into misperceptions about beliefs of policy-relevant variables and the policy rule. The AR(1) coefficient of the shock is set to 0.95.

V. An emprical assesment of monetary misperceptions

In this section, we empirically assess misperceptions of monetary policy. We proceed in two ways. First, we decompose the forecast error of monetary policy surprises into its subcomponents. Using this decomposition, we identify a monetary policy shock and study its dynamics. We further perform a variance decomposition to assess the share of variation attributable to misperceptions relative to the monetary policy shock. Second, we directly examine the effect of an empirical productivity shock on monetary surprises, tracing out its effect on the nominal rate and inflation.

V.A Decomposing monetary surprises empirically

General setup. We start by applying the theoretical decomposition from equation (14) to the data. To do so, we use our rule estimates from Table 1, together with the Blue Chip forecast data and realized data, to compute time series for each component in equation (14). Specifically, the belief misperception for inflation is given by $b_t^{\pi} = (1-\rho)\phi^{\pi}(\widehat{\pi}_t - \mathbb{E}_{t-12}^{\mathcal{P}}\widehat{\pi}_t)$. $b^{\tilde{y}}$ is the equivalent term for the output gap. The rule misperception terms $r_t^{\pi}, r_t^{\tilde{y}}, r_t^{i}$ are defined in the same manner. Note that, as in Section II, the frequency is monthly, but the forecast horizons are annual. The timing subscripts therefore differ from those in the model counterpart.

With these time series at hand, we can run the decomposition empirically:

$$i_t - \mathbb{E}_{t-12}^{\mathcal{P}} i_t = \alpha + \beta_1 \times b_t^{\pi} + \beta_2 \times b_t^{\tilde{y}} + \beta_3 \times r_t^{\pi} + \beta_4 \times r_t^{\tilde{y}} + \beta_5 \times r_t^{i} + \epsilon_t^{MP} \tag{15}$$

The regression can essentially be understood as a cleaning procedure: it allows us to clean the forecast error of the federal funds rate from belief ($b_t^{\pi}, b_t^{\tilde{y}}$) and rule ($r_t^{\pi}, r_t^{\tilde{y}}, r_t^{i}$) misperceptions. The residual can therefore be interpreted as a deviation from the actual rule, and ϵ_t^{MP} is a standard monetary policy shock. We can now use these results to (i) study the properties of the monetary policy shock, (ii) perform a variance decomposition. Both deliver counterparts to our theoretical model. We report the regression results in Table (A.2).

Monetary policy shock. To study the dynamic responses of the nominal rate and inflation to a monetary policy shock, we rely on local projections. We study the response of a variable y to the monetary policy shock ϵ_t^{MP}, which we identified above.

$$y_{t+h} = c_h + \gamma_h \times \epsilon_t^{MP} + controls_t + u_t$$

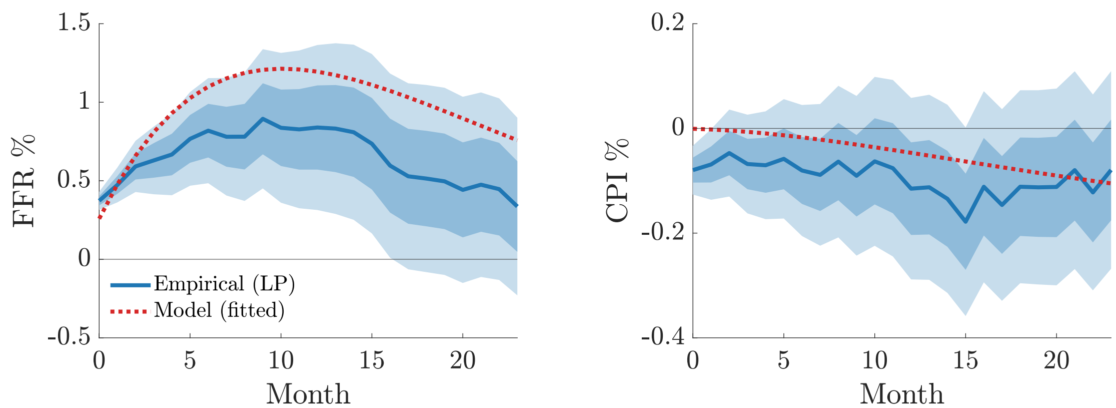
c_h is a constant, γ_h is the coefficient on the monetary policy shock at horizon h after the shock materialized, and $controls_t$ are a set of controls. As controls, we include the FFR, log CPI. log industrial production, the unemployment rate, and log of the Brent oil price index,[14] and as variables of interest, we consider FFR and log CPI. The sample spans 1985 to 2020. To compare the model to the data, we estimate the persistence of the monetary policy shock, a shock to the actual rule (3), by minimizing the sum of squared deviations between the model-implied and empirical impulse

[14]We will use the same set of controls for all local projections throughout the paper,

responses of the policy rate over the first 24 months.[15]

Figure (4) shows the results. First, the empirical IRFs display the standard dynamics of a monetary policy shock. Second, the model matches the targeted FFR response well: it slightly overshoots the empirical estimate but remains within the confidence bands throughout. Third, the untargeted CPI response is also well matched, with the model modestly undershooting at short horizons while, as for the FFR, staying within the estimated confidence intervals. Overall, the model closely reproduces the empirical responses.

Figure 4: Monetary policy shock in the Data and the Model



Notes: Response to a contractionary monetary policy shock. Confidence Intervals are Newey-West (one std and 95%).

Variance decomposition. Using the empirical decomposition in equation (15), we further attribute the total variance of the federal funds rate forecast error to its subcomponents. The first panel of Table (3) decomposes the total variance of the FFR forecast error into $belief$, $rule$, and $shock$ components. The second panel splits the $rule$ component into a part capturing misperceptions in the rule arising from inflation and the output gap ($rule_{\pi,\tilde{y}}$), and a part capturing the backward-looking nominal rate component ($rule_i$). For brevity, the covariance terms are not reported in the table, so that the rows do not necessarily sum to 1.

We find that a significant share of the variation in the forecast error is explained by the monetary policy shock, accounting for roughly 61%. Among the remaining variation attributable to

[15] The empirical shock is identified from one-year-ahead survey forecast errors, normalized to a 100 basis point gap between the realized and twelve-month-ahead-forecast FFR (see equation (15)). We construct the model counterpart analogously: for a given value of ρ_{mp}, we use Lemma (1) to obtain households' s-period-ahead perceived paths of inflation and the output gap, iterate the perceived Taylor rule (7) forward twelve periods to obtain their twelve-month-ahead expectation of the policy rate, and compare it to the rate implied by the actual rule (3). We choose the shock size such that the model-implied twelve-month-ahead forecast error equals 100 basis points.

misperceptions, belief misperceptions clearly dominate, accounting for 32%, while rule misperceptions play a minor role, at 10%. Splitting the rule misperceptions into inflation/output-gap and lagged nominal rate components reveals that the majority of this variation is explained by rule misperceptions in the lagged nominal rate. A likely explanation is the generally sluggish manner in which central banks set rates: central banks typically operate in tightening and easing cycles, adjusting rates slowly, which generates a high degree of autocorrelation in the policy rate. If agents misperceive the degree of this autocorrelation, the resulting forecast errors can be persistent and potentially large.

Table 3: Variance Decomposition of forecast error of the FFR

	$beliefs$	$rule$	$rule_{\pi,\tilde{y}}$	$rule_i$	$shock$
(1)	0.32	0.10			0.61
	(0.12)	(0.07)			(0.11)
(2)	0.32		0.09	0.29	0.61
	(0.12)		(0.22)	(0.07)	(0.11)

Notes: The table reports a variance decomposition based on equation (15). Standard errors are computed by taking into account estimation uncertainty from the SUR estimation of (1) and (2). We employ a block bootstrap using 2000 draws.

To compare the empirical share of misperceptions in beliefs and the rule with their model counterparts, we examine how the variation in total misperceptions is split between belief and rule misperceptions.[16] We further exclude the share of variation generated by the backward-looking component of the rule, $rule_i$. Specifically, we examine how the variation in $beliefs + rule_{\pi,\tilde{y}}$ is distributed across its two components. This yields a measure that is comparable across model and data without relying on overly strong assumptions. The variation in the model is generated by the productivity shock.

Table (4) presents the results: the model captures the qualitative pattern that misperceptions are dominated by beliefs rather than by the rule. However, the model overstates this pattern: The model attributes 85% of the variation to beliefs, compared to 62% in the data. We conclude that, while the model qualitatively matches the higher share to be attributed to belief misperceptions, it overstates the relative importance of belief misperceptions.

[16]Comparing the total variation in the FFR with the model counterpart would require knowledge of the relative sizes of the underlying shocks (i.e., the monetary policy shock relative to the productivity shock). As this lies beyond the scope of the present exercise, we instead focus on the relative shares of misperceptions.

Table 4: Variance decomposition of misperception terms

	Data	*Model*
Beliefs	0.62	0.85
	(0.14)	
Rule	0.38	0.15
	(0.14)	

Notes: The table reports a variance decomposition of the misperception terms omitting the share attributed to the lagged *rule* term. Standard errors are computed by taking into account estimation uncertainty from the SUR estimation of (1) and (2). We employ a block bootstrap using 2000 draws. The model counterpart is computed based on a productivity shock.

V.B Endogenous monetary policy surprise in the data

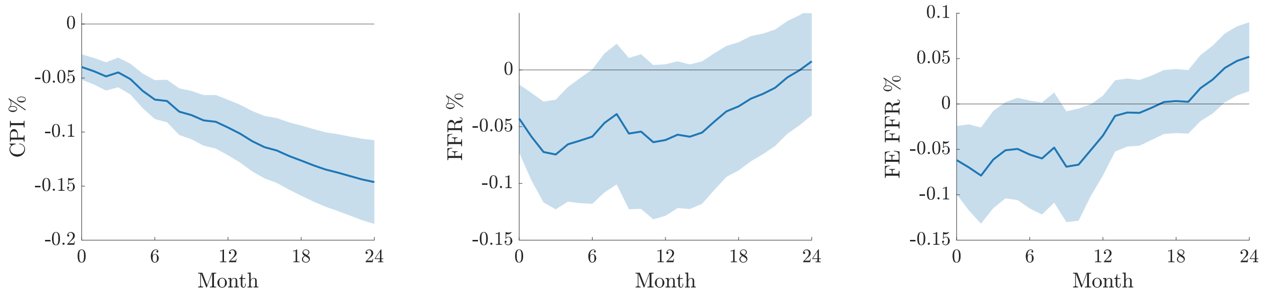
We now directly identify a productivity shock empirically and study the associated monetary surprise component. To do so, we use a VAR and identify the productivity shock following Gali (1999) and Blanchard and Quah (1989). By adding the forecast error of the FFR to the VAR, we trace out the endogenous monetary surprise generated by the productivity shock.

Specifically, we include the following variables in the VAR, ordered as listed: a proxy for productivity, the unemployment rate, log CPI, log oil prices, the FFR forecast error, and the FFR. The frequency is monthly, we use 12 lags, and the sample runs from 1986 to 2019. The proxy for productivity is manufacturing output per hour.[17] The forecast error is constructed as above and is based on Blue Chip forecasts.

Figure (5) shows the results. In line with Gali (1999) and the dynamics of the standard New Keynesian model, an expansionary productivity shock leads to a decline in CPI, to which the central bank responds by reducing rates. The forecast error is negative, indicating that markets underestimate the central bank's reaction. All of these dynamics mirror our findings in Section (IV). Our model therefore also matches the empirical dynamics conditional on a productivity shock.

[17]We construct this series from production and manufacturing, average weekly hours in manufacturing, and manufacturing employment, all obtained from the FRED database.

Figure 5: Empirical productivity shock responses



Notes: Response to an expansionary productivity shock. CIs are 1 std. and have been generated by a bootstrap procedure.

VI. Implications for monetary policy

In this section, we study the implications for monetary policy that arise from our model. Specifically, we show that our model can resolve two prominent puzzles: the price puzzle and the forward guidance puzzle. The former is an empirical puzzle that arises when estimating the response of inflation to a monetary policy shock. It is widely documented that the inflation response correlate positively with the nominal rate response in the short run, while theory would predict the opposite (Sims, 1980; Christiano et al., 1999). We provide a theoretical explanation for these dynamics. On the other hand, the forward guidance puzzle is a theoretical puzzle. Standard New Keynesian models suffer from the property that future announcements about monetary policy, the further ahead this announcement is made, increase the contemporaneous responses in the model. This is add odds with the data (Del Negro et al., 2023).

VI.A Resolving the price puzzle

Characterizing short-run inflation dynamics in the model. We start by characterizing the inflation response to a monetary policy shock arising from the actual rule (3). Given that the price puzzle is only reported in the short-run, we focus on the impact response of inflation to a monetary policy shock. Using equations (12), (13), and the actual rule, we can derive the following Proposition:

Proposition 1. *The impact inflation response to a monetary policy shock is given by*

$$\widehat{\epsilon}_t^{MP} = \Omega \times \widehat{\pi}_t$$

where Ω is a function of $\phi^{\pi}, \phi^{\tilde{y}}, \rho, \bar{\phi}^{\pi}, \bar{\phi}^{\tilde{y}}, \bar{\rho}$. The following holds:

- $\Omega(\phi^{\pi} = \phi^{\tilde{y}} = \rho = \bar{\phi}^{\pi} = \bar{\phi}^{\tilde{y}} = \bar{\rho} = 0) > 0$

- $\partial\Omega/\partial\phi^{\pi}, \partial\Omega/\partial\phi^{\tilde{y}}, \partial\Omega/\partial\bar{\phi}^{\pi}, \partial\Omega/\partial\bar{\phi}^{\tilde{y}} < 0$

Proof. See Appendix (B.4). ■

Proposition (1) makes two points. First, for sufficiently low actual and perceived rule coefficients, the correlation between the monetary policy shock and inflation is positive, and the price puzzle occurs. Second, as the rule coefficients increase, Ω decreases and, for sufficiently large rule coefficients, turns negative. In this case, the correlation between the monetary policy shock and inflation is negative, recovering the standard dynamics.

The intuition is as follows. Agents perceive movements in the nominal rate as being associated with movements in policy-relevant variables. When the perceived rule coefficients are sufficiently low, agents believe the central bank does not lean strongly enough against these dynamics. When the actual coefficients are sufficiently low, the central bank's reaction is truly too weak. Either case can generate a price puzzle. For sufficiently high coefficients, the opposite holds, and the usual dynamics of the standard New Keynesian model emerge.

A similar intuition arises in the standard model: the perceived and actual rule coincide by assumption. If the coefficients are too small, the Taylor principle fails, and multiple equilibria emerge. Among these equilibria are paths that generate a positive correlation between the monetary policy shock and inflation. The key difference between the standard model and ours is that our model is purely backward-looking, which guarantees uniqueness: whenever an equilibrium exists, it must be unique, unlike in the standard model.

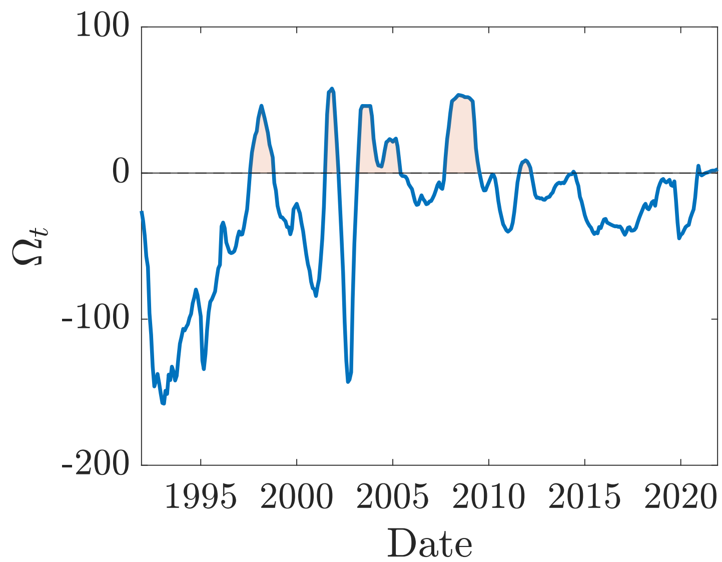
An empirical test for the price puzzle. Given our results above, we now test Proposition 1 in the data. To do so, we rely on time-varying perceived rule estimates from Bauer et al. (2024), who provide estimates for $\bar{\phi}^{\pi}, \bar{\phi}^{\tilde{y}}, \bar{\rho}$ at a monthly frequency. For the actual rule coefficients, we use a rolling-window regression with a seven-year window. Time-varying estimates of Ω_t are then obtained by feeding the rule coefficients into the definition of Ω_t.[18]

Figure (6) shows the results. While Ω_t is negative over the majority of the sample, we find positive values in roughly 25% of the sample, indicating that both standard and price-puzzle dynamics can be generated within our sample. The large values and volatility of Ω_t can be attributed to the volatility of the rule estimates, which is amplified by the definition of Ω_t. For comparison, our static rule estimates yield an Ω value of roughly -3, which falls well within the range of values we find for the time-varying Ω_t. To understand why Ω_t turns positive or negative, we

[18]Before feeding the coefficients into the Ω_t definition, we winsorize and smooth the rule estimates, as the raw estimates are quite volatile.

examine the correlation between 12-month changes in the FFR and Ω_t. The correlation coefficient is negative, at −0.33. Splitting the sample into positive and negative Ω_t values, we find that the mean annual FFR change drops from 0.26 to −0.94 when moving from negative to positive Ω_t values. This indicates that price puzzles are more likely to occur during easing cycles.

Figure 6: Empirical Ω_t estimates



Notes: Ω_t estimates in the data: rule coefficients are winsorized, and a 12-month moving average is applied.

Using our time-varying Ω_t estimate, we can now directly test Proposition (1). We run a local projection exercise studying the response of CPI to a monetary policy shock and add an interaction term of the monetary policy shock with (standardized) Ω_t. Since our proposition is focused on the impact response, we set the horizon for the local projections to $h \in [0, 2]$. We choose the same controls as in the previous local projections, adding Ω_t and the interacted terms. As the number of coefficients estimated is fairly high, we drop log industrial production as a control.[19] The local projection is given by:

$$logCPI_{t+h} = c_h + \beta_{h,1}\epsilon_t^{MP} + \beta_{h,2}(\Omega_t \times \epsilon_t^{MP}) + controls_t + u_{t+h}$$

Our theory would suggest that the coefficient in $\beta_{h,2}$ must be positive: considering a contractionary monetary policy shock, a larger Ω_t should suggest a more positive CPI response, leaning against usual dynamics. To test our theory, we use several monetary policy shocks that have been identified using different procedures. We use our own monetary policy shock identified using the

[19]Exchanging the unemployment rate for log industrial production leads to similar results. Adding both has no major effect on the point estimates, but slightly increases the confidence intervals.

residual of the FFR forecast error regression, equation (15). As a comparison, we also use high-frequency identified monetary policy shocks by Gertler and Karadi (2015), and narrative-identified monetary policy shocks by Aruoba and Drechsel (2024).

Table (5) presents the results. Regarding the point estimates of $\beta_{h,2}$, all coefficients are positive across all horizons and all shocks, indicating that a larger Ω_t can generate a price puzzle: for our own shock, we find a strongly significant coefficient across all horizons. For the Gertler and Karadi (2015) case, the coefficient is also significant across all horizons, while for the Aruoba and Drechsel (2024) shocks we find significant responses at horizons 1 and 2. We also report the estimates for $\beta_{h,1}$: for our own shock and the Gertler and Karadi (2015) shock, we find a negative response indicating that there is no price puzzle when Ω_t is at its mean. For Aruoba and Drechsel (2024), the opposite is the case, which is, however, in line with their own findings.[20] Overall, this provides suggestive evidence that the results of Proposition (1) are present in the data.

Table 5: Local projections: Price Puzzle test

	$\beta_{h,1}$			$\beta_{h,2}$		
	$h=0$	$h=1$	$h=2$	$h=0$	$h=1$	$h=2$
Residual of FFR FE regression	−0.07	−0.13	−0.15	0.05	0.06	0.05
	(0.03)	(0.04)	(0.05)	(0.02)	(0.03)	(0.04)
Gertler & Karadi (2015)	−0.29	−0.53	−0.72	0.22	0.22	0.31
	(0.20)	(0.25)	(0.29)	(0.17)	(0.21)	(0.24)
Aruoba & Drechsel (2024)	0.26	0.09	0.07	0.07	0.31	0.27
	(0.22)	(0.29)	(0.33)	(0.18)	(0.24)	(0.27)
Sample	1992 – 2019					

Notes: Response to a contractionary monetary policy shock. Confidence Intervals are Newey-West (one std).

Discussion. The main theoretical alternative to our explanation is the cost channel of monetary policy (Ravenna and Walsh, 2006). However, Rabanal (2007) finds that this channel is empirically not supported. Our explanation relies instead on the central bank's actual and perceived reaction. We further provide empirical support for our finding. Castelnuovo and Surico (2010) document that price puzzles are more likely when the central bank's actual response to inflation is weak, which maps directly onto the actual-rule coefficient in our Proposition (1). We embed this responsiveness result into a standard New Keynesian framework and extend it to the perceived reaction coefficient. Finally, Tas (2011) shows that the price puzzle can also be explained in a learning model with

[20] In their preferred specification Aruoba and Drechsel (2024) a price puzzle is reported in the inflation response for the first periods.

asymmetric information between the central bank and the public.

VI.B Resolving the forward guidance puzzle

General setup. We now turn to the forward guidance puzzle. To do so, we include expectations about future monetary policy shocks in the perceived rule as follows:

$$\mathbb{E}_t^{\mathcal{P}} i_{t+n} = (1-\bar{\rho})\left[\bar{\phi}^{\pi}\mathbb{E}_t^{\mathcal{P}}\pi_{t+n} + \bar{\phi}^{\tilde{y}}\mathbb{E}_t^{\mathcal{P}}\tilde{y}_{t+n}\right] + \bar{\rho}\mathbb{E}_t^{\mathcal{P}} i_{t+n-1} + \mathbb{E}_t^{\mathcal{P}}\epsilon_{t+n}^{MP} \tag{16}$$

where $\mathbb{E}_t^{\mathcal{P}}\epsilon_{t+n}^{MP}$ denotes an expected deviation from the perceived rule n periods ahead. Agents do not observe future policy shocks but receive signals about them. At date t, the central bank issues a signal $\hat{s}_{t,n}$ about the shock n periods ahead:

$$\hat{s}_{t,n} = \hat{\epsilon}_{t+n}^{MP} + \nu_{t,n}$$

where $\hat{\epsilon}_{t+n}^{MP}$ is the actual future realization of the shock and $\nu_{t,n}$ is a noise component. The shock and the noise component are normally distributed, $\hat{\epsilon}^{MP} \sim N(0,\sigma_\epsilon^2)$ and $\nu \sim N(0,\sigma_\nu^2)$. The realization of the shock n periods later is then given by

$$\hat{\epsilon}_t^{MP} = u_t + \hat{s}_{t-n,n},$$

where u_t denotes the unanticipated component of the shock.[21] Applying Gaussian signal extraction yields

$$\mathbb{E}_t^{\mathcal{P}}\hat{\epsilon}_{t+n}^{MP} = \bar{\lambda}\,\hat{s}_{t,n}, \qquad \bar{\lambda} \in [0,1]. \tag{17}$$

For normally distributed shock and noise components, the posterior mean gives the Bayesian weight $\bar{\lambda} = \sigma_\epsilon^2/(\sigma_\epsilon^2 + \sigma_\nu^2)$, which can be interpreted as central bank credibility. If the central bank is fully credible, the signal contains no noise ($\sigma_\nu^2 = 0 \Rightarrow \bar{\lambda} = 1$), and there is a one-to-one mapping between the signal and the expected shock. If central bank credibility is zero ($\sigma_\nu^2 \to \infty \Rightarrow \bar{\lambda} = 0$), the signal is fully uninformative, and agents disregard it entirely.

Focusing again on the impact response, we can derive the following Proposition:

Proposition 2. *The impact inflation response to a forward guidance shock at horizon n is given by*

$$C_\epsilon\bar{\lambda}\beta^n \times \hat{s}_{t,n} = \Omega \times \hat{\pi}_t$$

[21] This formulation is chosen for illustrative purposes, to make explicit that a realized shock consists of a previously issued signal and an unanticipated component. Substituting yields $u_t = -\nu_{t-n,n}$: the unanticipated component is simply the negative of the noise attached to the earlier signal.

- *No credibility by the Central Bank ($\bar{\lambda} = 0$) implies no response to forward guidance shocks.*

- *Higher perceived persistence in monetary rule ($\partial C_\epsilon / \partial \bar{\rho} > 0$) implies stronger forward guidance response.*

- *Discounting ($\beta \in (0, 1)$) implies weaker effects of forward guidance shock at higher horizons.*

Proof. See Appendix (B.5). ■

Proposition (2) makes three points. The first two concern the general potency of forward guidance shocks. First, if the central bank has no credibility, which arises when the noise component of the signal is too large, forward guidance shocks have no effect at all. Intuitively, for forward guidance to work, markets must believe the signals they receive from the central bank. Second, greater persistence in the perceived rule increases the potency of forward guidance shocks, since higher perceived persistence implies that the effects of a future shock are expected to last longer, and therefore the response of today's inflation is larger. Finally, discounting matters for the potency of forward guidance shocks: shocks further in the future are discounted more heavily, so their effect diminishes with the horizon. This breaks the forward guidance puzzle.

Technically, this result is similar to that obtained by Gabaix (2020), the main difference being that in Gabaix (2020) the result follows directly from excessive discounting in the Euler equation. In our setting, discounting is not excessive, but the result arises because households iterate over their budget constraint when forming beliefs and making decisions. This automatically generates stronger discounting of shocks that materialize at more distant horizons.

Multiple signals. In the previous section, we assumed that the central bank issues one signal at a certain point in time t, for a future deviation from the monetary rule that occurs n periods later. In practice, however, the monetary authority can issue multiple signals at different points in time for the same future shock period n. We now sketch how multiple announcements map into our framework.

As before, the realized shock splits into an unanticipated part, u_t, and the sum of past announcements up to a maximum horizon $\bar{N}$. We assume that the noise in the signal depends on how long ago it was issued: signals issued further in the past carry a higher noise component, reflecting dynamics such as the decay of memory over time. The monetary policy shock and the signal are then given by:

$$\hat{\epsilon}_t^{MP} = u_t + \sum_{n=1}^{\bar{N}} \hat{s}_{t-n,n}, \qquad \hat{s}_{T-h,h} = \hat{\epsilon}_T^{MP} + \hat{v}_{T-h,h},$$

where $\hat{s}_{T-h,h}$ is a signal issued h periods before the target date T, with $\hat{\epsilon}_T^{MP} \sim N(0, \sigma_\epsilon^2)$ and $\hat{v}_{T-h,h} \sim N(0, \sigma_{v,h}^2)$, being independent. Signal noise increases in the horizon at which the signal is issued. The precision of the signal is given by $\tau_h = 1/\sigma_{v,h}^2$ and with a geometric decay we have

$$\sigma_{v,h}^2 = \sigma_v^2\, \delta^{-(h-1)}, \qquad \tau_h = \tau\, \delta^{\,h-1}, \qquad \tau = 1/\sigma_v^2, \qquad \delta \in (0,1]$$

At horizon $h = 1$ and without decay ($\delta = 1$) the formulation nests our previous setup: $\sigma_{v,1}^2 = \sigma_v^2$.

In this setting, announcements about the same event, the future monetary policy shock, may repeat. Let $k_{t,n}$ denote the number of signals received up to t about $\hat{\epsilon}_{t+n}^{MP}$, issued at horizons $n, n+1, \ldots, n + k_{t,n} - 1$. Applying the Gaussian signal extraction yields

$$E_t^P\, \hat{\epsilon}_{t+n}^{MP} = \lambda_n(k_{t,n})\, \bar{s}_{t,n}, \qquad \bar{s}_{t,n} = \frac{\sum_{j \geq 0} \tau_{n+j}\, \hat{s}_{t-j,n+j}}{\bar{\tau}_{t,n}}, \qquad \bar{\tau}_{t,n} = \sum_{j=0}^{k_{t,n}-1} \tau_{n+j}, \tag{18}$$

Compared with the singular-signal case (equation (17)), the multiple-signal case (equation (18)) depends on the number of signals received (k) and their precisions, accounting for the decay (δ). Specifically, the expected future monetary policy shock is a function of the average signals received weighted by their relative precision, $\bar{s}_{t,n}$, and the Bayesian weight $\lambda_n(k)$. The latter is given by $\lambda_n(k) = \sigma_\epsilon^2\, \bar{\tau}_{t,n}/(1 + \sigma_\epsilon^2\, \bar{\tau}_{t,n})$. For simplicity, we drop the subscripts from k. Rewriting the geometric precision as $\bar{\tau}_{t,n} = \tau\, \delta^{n-1} \Lambda(k)$, we can express the Bayesian weight as

$$\lambda_n(k) = \frac{\sigma_\epsilon^2\, \tau\, \delta^{n-1} \Lambda(k)}{1 + \sigma_\epsilon^2\, \tau\, \delta^{n-1} \Lambda(k)}, \qquad with \qquad \Lambda(k) \;\equiv\; \sum_{j=0}^{k-1} \delta^{\,j} \;=\; \begin{cases} \dfrac{1-\delta^k}{1-\delta}, & \delta < 1, \\ k, & \delta = 1, \end{cases}$$

The Bayesian weight itself depends on the number of signals received about a certain shock occurring in n periods from now, and the distance n from today until the event occurs, accounting for the decay in memory. $\Lambda(k)$ is the effective number of signals taking into account decay. The Bayesian weight increases in the number of signals and decreases in the rate of decay in memory. To illustrate this, it is informative to consider the limiting case as $k \to \infty$:

Lemma 4 (Multiple signals, limiting case)**.** *Under infinitely many signals ($k \to \infty$), we have the following Bayesian weight for the case of no decay ($\delta = 1$) in memory and positive decay in memory $\delta \in (0,1)$:*

$$\lambda(k \to \infty) = \frac{\infty \times \sigma_\epsilon^2}{\sigma_v^2 + \infty \times \sigma_\epsilon^2} \to 1, \quad for \quad \delta = 1$$

$$\lambda_n(k \to \infty) = \frac{\sigma_\epsilon^2\, \tau\, \delta^{n-1}}{1 - \delta + \sigma_\epsilon^2\, \tau\, \delta^{n-1}} \in (0,1), \quad for \quad \delta \in (0,1)$$

Proof. Follows directly from the definition of the Bayesian weight. ■

Lemma (4) states that under infinitely many signals and no decay in memory, the central bank can attain full credibility. With positive decay, full credibility is not attainable, and the Bayesian weight is strictly below one. In this case, the horizon n affects credibility.

Given this formulation, one can now use the expected monetary policy shock definition from equations (18), follow the same steps as in Proposition (2), and derive a similar formulation:

$$C_\epsilon \bar{\lambda}_n(k)\beta^n \times \bar{s}_{t,n} = \Omega \ \times \hat{\pi}_t$$

From this formulation, it is apparent that our previous results still hold. Additionally, more signals, an increase in k, improves the potency of forward guidance by increasing the average signal strength ($\bar{s}_{t,n}$) and increasing the credibility ($\lambda_n(k)$). Further, decay in memory weakens the potency of forward guidance shocks through the same channels.

VII. Conclusion

We document that FIRE is violated with respect to policy-relevant variables and that agents have a wrong perceptions about the policy rule. Integrating these properties into a New Keynesian model we show that it can endogenously create monetary policy surprises, which align with empirical evidence. Finally, we show that our model can resolve the price and forward guidance puzzle.

References

Adam, K. and A. Marcet (2011): "Internal rationality, imperfect market knowledge and asset prices," **Journal of Economic Theory**, 146, 1224–1252.

Adam, K., A. Marcet, and J. Beutel (2017): "Stock price booms and expected capital gains," **American Economic Review**, 107, 2352–2408.

Adam, K., O. Pfäuti, T. Reinelt, et al. (2025): "Falling natural rates, rising housing volatility and the optimal inflation target," **Journal of Monetary Economics**, 149.

Amodeo, F. (2025): "Mind the gap: Disagreement and credible monetary policy," **Available at SSRN 6819679**.

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2021): "Imperfect macroeconomic expectations: Evidence and theory," **NBER Macroeconomics Annual**, 35, 1–86.

Aruoba, S. B. and T. Drechsel (2024): "Identifying Monetary Policy Shocks: A Natural Language Approach," NBER Working Paper 32417, National Bureau of Economic Research.

Bauer, M. D., C. E. Pflueger, and A. Sunderam (2024): "Perceptions about monetary policy," **The Quarterly Journal of Economics**, 139, 2227–2278.

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

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction in macroeconomic expectations," **American Economic Review**, 110, 2748–2782.

Caballero, R. J. and A. Simsek (2022): "Monetary policy with opinionated markets," **American Economic Review**, 112, 2353–2392.

Caines, C. and F. Winkler (2019): "Asset price learning and optimal monetary policy," **Available at SSRN 3070761**.

Carvalho, C. and F. Nechio (2014): "Do people understand monetary policy?" **Journal of Monetary Economics**, 66, 108–123.

Castelnuovo, E. and P. Surico (2010): "Monetary policy, inflation expectations and the price puzzle," **The Economic Journal**, 120, 1262–1283.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (1999): "Monetary policy shocks: What have we learned and to what end?" **Handbook of macroeconomics**, 1, 65–148.

COIBION, O. AND Y. GORODNICHENKO (2015): "Information rigidity and the expectations formation process: A simple framework and new facts," **American Economic Review**, 105, 2644–2678.

DEL NEGRO, M., M. P. GIANNONI, AND C. PATTERSON (2023): "The forward guidance puzzle," **Journal of Political Economy Macroeconomics**, 1, 43–79.

GABAIX, X. (2020): "A Behavioral New Keynesian Model," **American Economic Review**, 110, 2271–2327.

GALI, J. (1999): "Technology, employment, and the business cycle: do technology shocks explain aggregate fluctuations?" **American economic review**, 89, 249–271.

GERTLER, M. AND P. KARADI (2015): "Monetary Policy Surprises, Credit Costs, and Economic Activity," **American Economic Journal: Macroeconomics**, 7, 44–76.

HAMILTON, J. D., S. PRUITT, AND S. BORGER (2011): "Estimating the market-perceived monetary policy rule," **American Economic Journal: Macroeconomics**, 3, 1–28.

JIA, P., H. SHEN, AND S. ZHENG (2023): "Monetary policy rules and opinionated markets," **Economics Letters**, 223, 110995.

KIM, H. (2026): "Perceived Monetary Policy Rule: Identification and Macroeconomic Implications," **Available at SSRN**.

KOHLHAS, A. N. AND A. WALTHER (2021): "Asymmetric attention," **American Economic Review**, 111, 2879–2925.

MANKIW, N. G., R. REIS, AND J. WOLFERS (2003): "Disagreement about inflation expectations," **NBER macroeconomics annual**, 18, 209–248.

MCKAY, A., E. NAKAMURA, AND J. STEINSSON (2016): "The power of forward guidance revisited," **American Economic Review**, 106, 3133–3158.

PFÄUTI, O. (2023): "The inflation attention threshold and inflation surges," **arXiv preprint arXiv:2308.09480**, 2.

——— (2025): "Inflation—who cares? Monetary policy in times of low attention," **Journal of Money, Credit and Banking**, 57, 1211–1239.

Rabanal, P. (2007): "Does inflation increase after a monetary policy tightening? Answers based on an estimated DSGE model," **Journal of Economic Dynamics and control**, 31, 906–937.

Ravenna, F. and C. E. Walsh (2006): "Optimal monetary policy with the cost channel," **Journal of Monetary Economics**, 53, 199–216.

Roschitsch, U. and H. Twieling (2026): "United in Booms, Divided in Busts: Regional House Price Cycles and Monetary Policy," .

Sastry, K. A. (2026): "Disagreement about monetary policy," **American Economic Journal: Macroeconomics**, 18, 486–522.

Sims, C. A. (1980): "Macroeconomics and reality," **Econometrica: journal of the Econometric Society**, 1–48.

Sutherland, C. S. (2026): "Two Sides of the Same Coin: Expected Reaction Functions and Perceived Monetary Policy Shocks," **Bank of Canada Staff Analytical Paper**, Staff Working Paper No. 2026.

Tas, B. K. O. (2011): "An explanation for the price puzzle: Asymmetric information and expectation dynamics," **Journal of Macroeconomics**, 33, 259–275.

Taylor, J. B. (1993): "Discretion versus policy rules in practice," in **Carnegie-Rochester conference series on public policy**, Elsevier, vol. 39, 195–214.

Winkler, F. (2020): "The role of learning for asset prices and business cycles," **Journal of Monetary Economics**, 114, 42–58.

Online Appendix

A. Empirical Appendix

A.1 Perceived rule using SPF data

Table (A.1) reports the results of seemingly unrelated regressions on equations (1) and (2) using SPF forecasts instead of Blue Chip data. As the SPF has no forecasts on the FFR, we use the TBILL as a proxy instead. The frequency is quarterly instead of monthly.

Table A.1: Perceived and Actual Policy Rules, SPF

	Perceived (SPF)	*Actual*
ϕ^{π}	0.91	1.82
	(0.50)	(0.50)
$\phi^{\tilde{y}}$	−0.08	−1.15
	(0.30)	(0.40)
ρ	0.92	0.76
	(0.03)	(0.06)
R^2	0.97	
Joint test : Perceived = Actual	0.00	
Frequency	*quarterly*	
Sample	1981 – 2025	

Notes: The table reports estimates of the perceived and actual Taylor rules, given by equations (1) & (2). Rules are estimated jointly as a seemingly unrelated regression (SUR). Standard errors are Driscoll-Kraay. *Joint test: Perceived = Actual* reports the p-value of a Wald test of $H_0 : \bar{\phi}^{\pi} = \phi^{\pi} \wedge \bar{\phi}^{\tilde{y}} = \phi^{\tilde{y}} \wedge \bar{\rho} = \rho$.

A.2 Empirical Decomposition

Table (A.2) shows the estimate from equation (15). Since right-hand-side variables depend on estimates from the SUR estimation, standard errors are estimated using a block bootstrap, accounting for the uncertainty of the estimation from the SUR regressions. From the perspective of our theoretical model, equation (15) is an accounting equation, and all estimated coefficients from the regression should be unity. A joint test on all coefficients being equal to one cannot reject the H_0, supporting our theoretical approach. Combining the belief and rule misperception terms into two regressors, one for each type of misperception, and re-running the regression as well as the joint test increases the p-value of the test from 0.14 to 0.65.

Table A.2: Empirical Decomposition

	β_1	β_2	β_3	β_4	β_5
	0.61	1.52	0.96	0.51	1.04
	(7.87)	(0.72)	(15.06)	(0.70)	(0.59)
$Joint\ test : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 1$			0.14		
R^2			0.39		
Frequency			*monthly*		
Sample			1984 – 2020		

Notes: The table reports estimates of equation (15). Standard errors are computed by taking into account estimation uncertainty from the SUR estimation of (1) and (2). We employ a block bootstrap using 2000 draws. The joint test reports the p-value of the test.

B. Proofs

B.1 Proof Lemma (1)

The agents apply the optimal Bayesian filter, that is, the Kalman filter, to arrive at the observable system:[22]

$$\begin{aligned}\ln\frac{z_{t+1}}{z_t} &= \varrho_z \ln\overline{m}_{z,t} + \ln\widehat{e}_{z,t+1}\\ \ln\overline{m}_{z,t} &= \varrho_z \ln\overline{m}_{z,t-1} - \frac{\sigma^2_{z,e}}{2} + g_z\cdot\left(\ln\widehat{e}_{z,t} + \frac{\sigma^2_{z,e}+\sigma^2_{z,v}}{2}\right),\end{aligned}$$

where $\ln\overline{m}_{z,t} := \mathbb{E}^{\mathcal{P}}_t(\ln m_{z,t})$ is the posterior mean, $g_z = \frac{\varrho_z^2\sigma_z^2+\sigma^2_{z,v}}{\varrho_z^2\sigma_z^2+\sigma^2_{z,v}+\sigma^2_{z,e}}$ is the steady-state Kalman filter gain, $\sigma_z^2 = \frac{1}{2\varrho_z^2}[-(\sigma^2_{z,v}+(1-\varrho_z^2)\sigma^2_{z,e}) + \sqrt{(\sigma^2_{z,v}+(1-\varrho_z^2)\sigma^2_{z,e})^2+4\varrho_z^2\sigma^2_{z,v}\sigma^2_{z,e}}]$ is the steady-state Kalman filter uncertainty, and $\ln\widehat{e}_{z,t}$ is perceived to be a white noise process.

Equation (5) is the result of the following calculations:

$$\mathbb{E}^{\mathcal{P}}_t z_{t+s} = z_t\cdot\mathbb{E}^{\mathcal{P}}_t\frac{z_{t+s}}{z_t} = z_t\cdot\mathbb{E}^{\mathcal{P}}_t\exp\left(\ln z_{t+s}-\ln z_t\right) = z_t\cdot\mathbb{E}^{\mathcal{P}}_t\exp\left(\sum_{n=1}^{s}\Delta\ln z_{t+n}\right)$$

[22] We assume the agents' prior variance equals the steady-state Kalman variance.

$$= z_t \cdot \mathbb{E}_t^{\mathcal{P}} \exp\left(\sum_{n=1}^{s} \ln m_{z,t+n}\right) \cdot \underbrace{\mathbb{E}_t^{\mathcal{P}}\left[\prod_{n=1}^{s} e_{z,t+n}\right]}_{=\prod_n \mathbb{E}_t^{\mathcal{P}} e_{z,t+n}=1}$$

$$= z_t \cdot \mathbb{E}_t^{\mathcal{P}} \exp\left(\sum_{n=1}^{s}\left[\sum_{j=0}^{n-1} \varrho_z^j \ln v_{z,t+n-j} + \varrho_z^n \ln m_{z,t}\right]\right)$$

$$= z_t \cdot \mathbb{E}_t^{\mathcal{P}} \exp\left(\ln m_{z,t} \cdot \sum_{n=1}^{s} \varrho_z^n\right) \cdot \mathbb{E}_t^{\mathcal{P}} \exp\left(\underbrace{\sum_{n=1}^{s}\sum_{j=0}^{n-1} \varrho_z^j \ln v_{z,t+n-j}}_{\sim\mathcal{N}}\right)$$

$$= z_t \cdot \mathbb{E}_t^{\mathcal{P}} \exp\left(\varrho_z \frac{1-\varrho_z^s}{1-\varrho_z} \ln m_{z,t}\right) \cdot \exp(V), V \propto \sigma_{z,v}^2$$

$$\iff \mathbb{E}_t^{\mathcal{P}} z_{t+s} = z_t \cdot \exp\left(\ln \overline{m}_{z,t} \cdot \varrho_z \tfrac{1-\varrho_z^s}{1-\varrho_z} + \frac{1}{2}\sigma^2\left(\varrho_z \tfrac{1-\varrho_z^s}{1-\varrho_z}\right)^2\right) \cdot \exp(V), V \propto \sigma_{z,v}^2$$

B.2 Proof Lemma (2)

The linearized household budget constraint at $t+s$ is given by:

$$c_{ss}\mathbb{E}_t^{\mathcal{P}}\widehat{c}_{t+s} + y_{ss}\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{t+s+1} = \beta^{-1} y_{ss}\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{t+s} + x_{ss}\mathbb{E}_t^{\mathcal{P}}\widehat{x}_{t+s}$$

Without loss of generality, assume bonds are in zero net supply and linearize the bonds around the steady state output. Also, in the standard New Keynesian model, we have that in steady state $c_{ss} = n_{ss} = x_{ss} = y_{ss}$. We therefore arrive at:

$$\mathbb{E}_t^{\mathcal{P}}\widehat{c}_{t+s} + \mathbb{E}_t^{\mathcal{P}}\widehat{b}_{t+s+1} = \beta^{-1}\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{t+s} + \mathbb{E}_t^{\mathcal{P}}\widehat{x}_{t+s} \tag{B.1}$$

Now let $s \to \infty$:

$$\mathbb{E}_t^{\mathcal{P}}\widehat{c}_\infty = \frac{1-\beta}{\beta}\mathbb{E}_t^{\mathcal{P}}\widehat{b}_\infty + \mathbb{E}_t^{\mathcal{P}}\widehat{x}_\infty \tag{B.2}$$

Now using equation (B.1) and substituting equations (8) and (B.2) yields:

$$\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{t+s+1} = \beta^{-1}\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{t+s} - \frac{1-\beta}{\beta}\mathbb{E}_t^{\mathcal{P}}\widehat{b}_\infty + \frac{1}{\sigma}\sum_{n=s}^{\infty}\mathbb{E}_t^{\mathcal{P}}\widehat{r}_{t+n+1} - \mathbb{E}_t^{\mathcal{P}}\widehat{x}_\infty + \mathbb{E}_t^{\mathcal{P}}\widehat{x}_{t+s}$$

For simplicity now define $\widehat{c}_{t+s}^* = \frac{1}{\sigma}\sum_{n=s}^{\infty}\mathbb{E}_t^{\mathcal{P}}\widehat{r}_{t+n+1}$. Backward iteration on the household budget constraint gives:

$$\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{t+s+1} = \beta^{-1}\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{t+s-1} + \frac{\beta-1}{\beta}\mathbb{E}_t^{\mathcal{P}}\widehat{b}_\infty + \widehat{c}_{t+s}^* + \beta^{-1}\widehat{c}_{t+s-1}^* - \mathbb{E}_t^{\mathcal{P}}\widehat{x}_\infty + \mathbb{E}_t^{\mathcal{P}}\widehat{x}_{t+s} + \beta^{-1}\mathbb{E}_t^{\mathcal{P}}\widehat{x}_{t+s-1}$$

$$= \beta^{-s}\widehat{b}_{t+1} - \beta^{-s}(1-\beta^{s})\Big(\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{\infty} + \frac{\beta}{1-\beta}\mathbb{E}_t^{\mathcal{P}}\widehat{x}_{\infty}\Big) + \beta^{-s}\sum_{n=1}^{s}\beta^{n}(\widehat{c}^{*}_{t+n} + \mathbb{E}_t^{\mathcal{P}}\widehat{x}_{t+n})$$

And we arrive at

$$\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{t+s+1} - \mathbb{E}_t^{\mathcal{P}}\widehat{b}_{\infty} = \beta^{-s}\Big(\widehat{b}_{t+1} - \mathbb{E}_t^{\mathcal{P}}\widehat{b}_{\infty} - (1-\beta^{s})\frac{\beta}{1-\beta}\mathbb{E}_t^{\mathcal{P}}\widehat{x}_{\infty} + \sum_{n=1}^{s}\beta^{n}(\widehat{c}^{*}_{t+n} + \mathbb{E}_t^{\mathcal{P}}\widehat{x}_{t+n})\Big)$$

As $s \to \infty$ we have that $\beta^{-s} \to \infty$ and therefore it needs to hold that:

$$\mathbb{E}_t^{\mathcal{P}}\widehat{b}_{\infty} = \widehat{b}_{t+1} + \sum_{n=1}^{\infty}\beta^{n}(\widehat{c}^{*}_{t+n} + \mathbb{E}_t^{\mathcal{P}}\widehat{x}_{t+n}) - \frac{\beta}{1-\beta}\mathbb{E}_t^{\mathcal{P}}\widehat{x}_{\infty} \tag{B.3}$$

And this finally implies:

$$\mathbb{E}_t^{\mathcal{P}}\widehat{c}_{\infty} = \frac{1-\beta}{\beta}\Big(\widehat{b}_{t+1} + \sum_{n=1}^{\infty}\beta^{n}(\widehat{c}^{*}_{t+n} + \mathbb{E}_t^{\mathcal{P}}\widehat{x}_{t+n})\Big) \tag{B.4}$$

B.3 Proof Lemma (3)

The IS equation follows from the Euler equation (11) and imposing market clearing: $\widehat{c}_t = \widehat{x}_t = \widehat{y}_t$. The coefficients are then given by:

$$C_{\pi} = \frac{\varrho_{\pi}\big[(1-\bar{\rho}\beta) - (1-\bar{\rho})\bar{\phi}^{\pi}\beta\big]}{1-\beta\varrho_{\pi}}, \qquad C_{\tilde{y}} = \frac{(1-\bar{\rho})\bar{\phi}_{\tilde{y}}\beta}{1-\beta}, \qquad C_{x} = \frac{\sigma(1-\bar{\rho}\beta)\varrho_{x}}{1-\beta\varrho_{x}}.$$

The Phillips Curve follows from solving for marginal costs expectations, rewriting marginal costs as the output gap, and imposing $\mathbb{E}_t^{\mathcal{P}}\mu_{t+n} = \rho_{\mu}^{n}\mu_t$. To achieve the latter we assume that the subsidy is an observable policy; firms condition on its process directly rather than filtering it through the belief state for marginal costs. Further, the coefficients are given by:

$$\Theta = \frac{1-\alpha}{1-\alpha+\alpha\epsilon}, \qquad \bar{\kappa} = \frac{\epsilon-1}{\kappa(1-\beta)}\lambda_y\Theta, \qquad \kappa_{\mu} = \frac{(\epsilon-1)\Theta}{\kappa(1-\beta\rho_{\mu})}, \tag{B.5}$$

B.4 Proof of Proposition (1)

Substituting the actual rule, and the IS equation into the PC gives:

$$\Omega \equiv \big[C_{\pi}g_{\pi} - (1-\rho)\phi^{\pi}\big] - \Big[C_{\tilde{y}}\Big(1+\frac{\rho_y}{1-\beta\rho_y}g_y\Big) - C_y g_y - (1-\rho_i)\phi^{y}\Big]D^{-1}, \quad with \quad D = \bar{\kappa}\Big(1+g_{mc}\frac{\beta\rho_{mc}}{1-\beta\rho_{mc}}\Big)$$

We also made use of the fact, that a monetary policy shock does not affect the efficient level of output. The results directly follow from the definition of Ω.

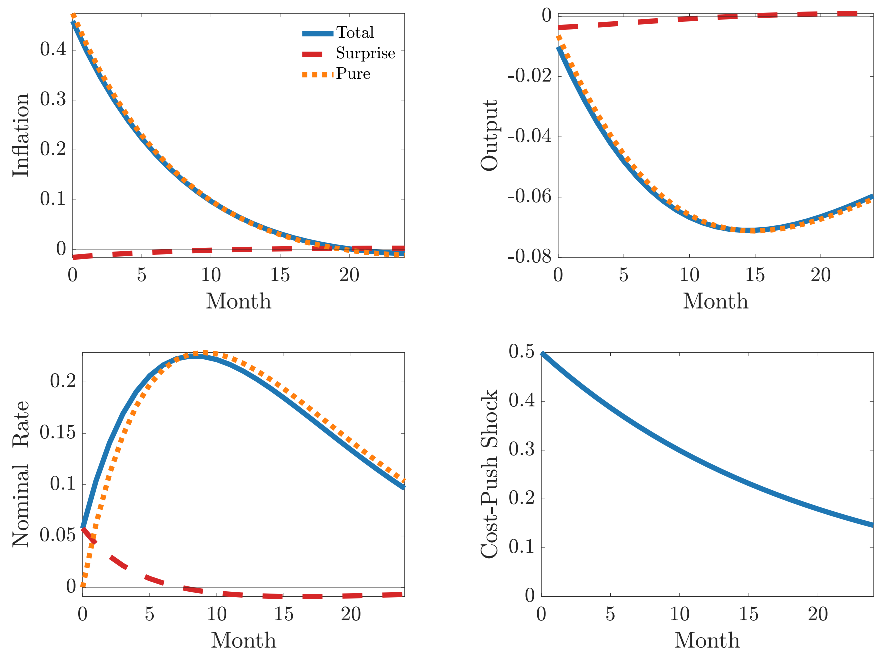
B.5 Proof of Proposition (2)

Results are obtained by first replacing the perceived rule (7) with the perceived rule that includes forward guidance, equation (16). One then re-derives the IS-equation in the same manner as described in Lemma (3). Combining the actual rule, the re-derived IS-equation with forward guidance, and the PC-equation for the impact response, as done in Proposition (1), and substituting the definition of the forward guidance shock ($\mathbb{E}_t^{\mathcal{P}} \hat{\epsilon}_{t+n}^{MP} = \bar{\lambda}\, \hat{s}_{t,n}$) gives the result. Finally, we have that $C_\epsilon = \left(1 + \frac{\bar{\rho}\beta}{1-\bar{\rho}\beta}\right)$.

C. Further Results

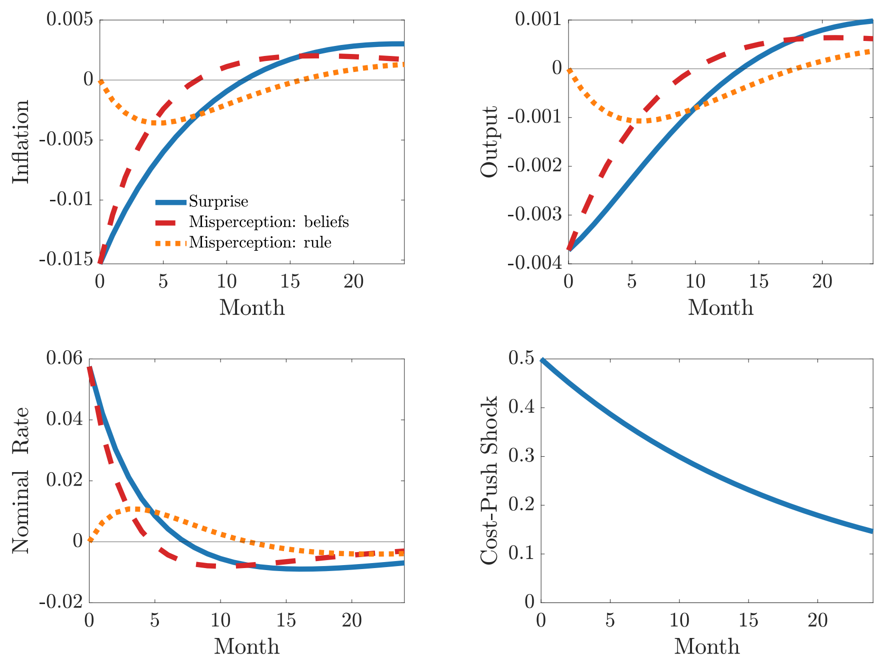
Figures (C.1) and (C.2) repeat the decomposition from the main text for a cost push shock. In this case, the cost-push shock triggers a contractionary surprise in the nominal interest rate. Consequently, inflation and output decline. The same holds for the decomposition into the belief and rule misperception terms.

Figure C.1: Decomposition of Total response into Surprise and Pure component, cost-push shock



Notes: Response to a cost-push shock. Decomposition into monetary surprise and Pure component. The AR(1) coefficient of the shock is set to 0.95.

Figure C.2: Decomposition of Surprise into Belief and Rule misperceptions, cost-push shock



Notes: Response to a cost-push shock. Decomposition into misperceptions about beliefs of policy-relevant variables and the policy rule. The AR(1) coefficient of the shock is set to 0.95.