Federal Reserve Bank of Cleveland Working Paper Series

Mis-specified Forecasts and Myopia in an Estimated New Keynesian Model

Ina Hajdini

Working Paper No. 22-03R

March 2023

Suggested citation: Hajdini, Ina. 2023. "Mis-specified Forecasts and Myopia in an Estimated New Keynesian Model." Working Paper No. 22-03R. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202203r

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Mis-specified Forecasts and Myopia in an Estimated New Keynesian Model*

Ina Hajdini†

March 1, 2023

Abstract

The paper considers a New Keynesian framework in which agents form expectations based on a combination of autoregressive mis-specified forecasts and myopia. The proposed expectations formation process is shown to be consistent with *all three empirical facts* on consensus inflation forecasts. However, while mis-specified forecasts can be both sufficient and necessary to match all three facts, myopia *alone* is neither. The paper then derives the general equilibrium solution consistent with the proposed expectations formation process and estimates the model with likelihood-based Bayesian methods, yielding three novel results: (i) macroeconomic data strongly prefer a combination of autoregressive mis-specified forecasting rules - of the VAR(1) or AR(1) type - and myopia over other alternatives; (ii) no strong evidence is found in favor of VAR(1) forecasts over simple AR(1) rules; and (iii) frictions such as habit in consumption, which are typically necessary for models with full-information rational expectations, are significantly less important, because the proposed expectations generate substantial internal persistence and amplification to exogenous shocks. Simulated inflation expectations data from the estimated general equilibrium model reflect the three empirical facts on forecasting data.

JEL Classification: C11; C53; D84; E10; E30; E50; E70

Keywords: Mis-specified Forecasts; Myopia; Survey of Professional Forecasters; Bayesian Estimation; Internal Propagation.

---

*I am deeply grateful to my Ph.D. advisors, Marco Airaudo and André Kurmann, for their advice and constructive criticism throughout this project. I thank Enghin Atalay, Marco Del Negro, Keshav Dogra, Thomas Eisenbach, Xavier Gabaix, Bart Hobijn, Ryota Iijima, Edward Knotek, Lars Ljungqvist, Thomas Lubik, Karel Mertens, Tristan Potter, Benjamin Pugsley, Robert Rich, Raphael Schoenle, Ricardo Serrano-Padial, seminar participants at Aix-Marseille School of Economics, Copenhagen Business School, Cornivus University of Budapest, Duke Kunshan University, Federal Reserve Bank of Cleveland, Federal Reserve Bank of Dallas, Federal Reserve Bank of New York, Federal Reserve Board, Örebro University, Osaka University, Özyeğin University, Universitat Autonoma de Barcelona, University of Ottawa, Vanguard, conference participants at the 11th ifo Conference on Macro and Survey Data, International Association for Applied Econometrics Annual Conference, International Conference on Computing in Economics and Finance, Expectations in Dynamic Macroeconomic Models, for comments and suggestions. The views expressed herein are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

†Federal Reserve Bank of Cleveland. Email: Ina.Hajdini@clev.frb.org.

1 Introduction

The full-information rational expectations (FIRE) assumption in macroeconomics postulates that agents understand the true underlying model of the economy and consequently have full knowledge of the equilibrium probability distribution of economic variables. This assumption is the workhorse of modern macro work and has brought to the field much discipline and important insights. However, it contradicts the ample evidence that agents, due to cognitive limitations or information acquiring costs, often resort to simple non-model-based forecasting rules (mis-specified forecasts) *and* do not appropriately take into account future payoffs/quantities (myopia).[1] To date, the literature has not incorporated both departures from FIRE in an equilibrium macro framework and has not formally tested them with macroeconomic data.

The present paper addresses this gap in the literature and makes its first contribution by *jointly* introducing mis-specified forecasting rules and myopia in a New Keynesian framework. The second contribution is to derive the consistent expectations equilibrium for the inflation process and test its three implications with evidence on inflation consensus forecasting data. The third contribution is to develop the full general equilibrium solution while allowing agents to perpetually learn about the equilibrium, and estimate it on US macroeconomic data with likelihood-based Bayesian methods.[2] The key novel result is that the expectations formation process characterized by a combination of autoregressive mis-specified forecasting rules and myopia is consistent with the evidence on consensus forecasting data *and* is strongly preferred by US macroeconomics data as shown by the likelihood-based Bayesian estimation of the full New Keynesian model.[3] Furthermore, the paper

[1]See for instance, Tversky and Kahneman (1973, 1974), Adam (2007), Hommes (2013); Hommes et al. (2019), Petersen (2015), Malmendier and Nagel (2016), and Ganong and Noel (2019), among others. See *Related Literature* for more details.

[2]The literature has long shown that agents tend to focus mostly on recent observations; that is, they rely on perpetual or constant-gain learning. For instance, Fuster, Laibson, and Mendel (2010) argue that "actual people's forecasts place too much weight on recent changes," Malmendier and Nagel (2016) find significant micro evidence in favor of constant-gain learning, and Tversky and Kahneman (1973, 1974) provide theoretical considerations.

[3]I prove that the proposed expectations formation process is consistent with *all three empirical facts* about inflation consensus forecasting data in the US, namely, that forecasts are positively predicted by ex-ante forecast revisions; that is, there is under-reaction to new information at the time of forecast (Coibion and Gorodnichenko (2015)); that forecasters over-react to information at the time of forecast (Kohlhas and Walther (2021)); and that following a one-time shock, the impulse responses of forecast errors initially under-shoot but then over-shoot (Angeletos, Huo, and Sastry (2021)). The three papers above use various sources of forecasting data to validate the three facts. In particular, Angeletos, Huo, and Sastry (2021) use forecasting data from the US SPF, Blue Chip, and Michigan Survey of Consumers (MSC); Coibion and Gorodnichenko (2015) rely on the same data sets, among many others, with the caveat that the term structure of the MSC data is not particularly fit for constructing ex-ante forecasting revisions; Kohlhas and Walther (2021) use data from the US SPF, Euro Area SPF, Livingston Survey, and MSC.

brings forward the new result that while mis-specified forecasting rules can be both sufficient and necessary to match the evidence on forecasting data, myopia *alone* is neither. This outcome is important as, differently from other papers in the literature, it highlights that a *unique* deviation from FIRE can be sufficient to mirror the evidence.

Agents in the private sector are assumed to be homogeneous, but endowed with imperfect common knowledge about each other's economic problems, shocks, and expectations formation processes. Since agents do not understand their uniformity, they are not aware of the true model governing the macroeconomy. As a result, they form forecasts about the endogenous variables based on mis-specified perceived laws of motion, i.e., rules that are structurally different from the minimum state variable solution granted under FIRE. In particular, motivated by evidence in the literature (see footnote 1), I assume that the perceived laws of motion are of an autoregressive nature. To model myopia, I build on the idea of cognitive discounting in Gabaix (2020), where the private sector has difficulty understanding events that are far in the future. As agents try to form expectations about the far future, they shrink their autoregressive forecasts - and consecutively, expectations about aggregate variables many periods ahead - toward the steady state of the economy. As in Gabaix (2020), the private sector is globally patient with respect to the variables' steady-state equilibrium, but is myopic with respect to their deviations from the steady state. Differently from Gabaix (2020), where *well-specified* forecasting rules are myopically adjusted, in the present paper such an adjustment is applied to *mis-specified* forecasting rules.[4]

Once myopia is combined with the autoregressive forecasts, the parameters of the forecasting rules are pinned down by the solution concept of a consistent expectations (CE) equilibrium, as defined in Hommes and Sorger (1998) and Hommes and Zhu (2014). A first-order CE equilibrium arises when the perceived unconditional mean and first-order autocorrelation coefficient/matrix of the endogenous variable(s) coincides with the same moments as implied by the data-generating process, that is, the actual law of motion, of the endogenous variable(s).

To assess the relevance of the proposed expectations formation process, I start off with a partial equilibrium New Keynesian pricing problem, where monopolistically competitive firms that face exogenous marginal costs maximize their present discounted value of real profits. Motivated by the work of Preston (2005) and Eusepi and Preston (2018), I model the implied optimal pricing rule of each firm to be of an infinite horizon nature, and in the presence of mis-specified forecasting

[4]Well-specified forecasting rules are such that they share a common structure with the minimum state variable solution under FIRE.

rules, it cannot be reduced to the standard Phillips curve.[5] I show that along the CE equilibrium path, the inflation persistence is *at least* as high as the inertia of the marginal cost.[6] Importantly, the degree of myopia interacts with the equilibrium outcomes, and specifically, a lower degree of myopia or, equivalently, more forward-lookingness in firms' decisions induces higher inflation persistence in equilibrium.

I then derive three testable implications for forecasting errors along the CE equilibrium path. First, I prove that, consistent with the evidence presented in Angeletos, Huo, and Sastry (2021), a combination of mis-specified forecasts and myopia delivers late over-shooting of forecast errors following a one-time shock to the marginal cost if and only if there is sufficient over-extrapolation, that is, if the equilibrium persistence of inflation sufficiently exceeds that of the marginal cost. Furthermore, the analysis proves that mis-specified forecasts alone, that is, a *unique* departure from FIRE, is sufficient to replicate delayed over-shooting.[7] On the other hand, Angeletos, Huo, and Sastry (2021) show that late over-shooting is matched if and only there is *both* over-extrapolation and a sufficiently large information friction. Second, I prove that a combination of mis-specified forecasts with myopia reflects the empirical fact that ex-post forecasting errors are positively predicted by ex-ante forecast revisions but are negatively predicted by current inflation realizations, as found in Coibion and Gorodnichenko (2015) and Kohlhas and Walther (2021), respectively. The presence of myopia generally facilitates matching the evidence on under-reaction to forecast revisions by slowing down the update of forecasts as new information becomes available to forecasters. On the other hand and similarly with results on delayed over-shooting, mis-specified forecasts alone can be sufficient to replicate over-reaction to current information.

Overall, the theoretical analysis proves that the proposed expectations formation process, that is, a combination of autoregressive mis-specified forecasts with myopia, is consistent with all three empirical facts on consensus inflation forecasting behavior.[8] However, I show that while mis-specified forecasts can be both sufficient and necessary to match the evidence, myopia *alone* is neither.

[5]However, when firms use well-specified forecasting rules, as in, e.g., Gabaix (2020), instead, the implied Phillips curve coincides with the behavioral one in Gabaix (2020) if there is myopia and the standard Phillips curve under FIRE if there is no myopia.

[6]When firms rely on well-specified forecasting rules, the persistence of inflation matches that of the marginal costs.

[7]I further show that well-specified forecasts combined with or absent myopia as in Gabaix (2020) or FIRE, respectively, do not match late over-shooting.

[8]While the current work focuses on a representative agents model and, thus, aggregate/consensus forecasting data, Gabaix (2020) shows that cognitive discounting (myopia) can be microfounded through noisy signals.

Evidence in favor of the proposed expectations formation process for inflation represents a natural motivation to embed such an assumption into a full New Keynesian model with habit in consumption and inflation indexation, similar to the one in Milani (2006).[9] Bayesian estimation of the full general equilibrium New Keynesian model on US macroeconomic data from 1966:Q1 to 2018:Q3 yields the following outcomes. First, macroeconomic data strongly prefer the model whose expectations formation process is a combination of autoregressive mis-specified forecasts and myopia over the other aforementioned alternatives. Second, while VAR(1) forecasting rules combined with myopia fit the data slightly better than the specification with AR(1) forecasts and myopia, the evidence is *not* strongly favoring one over the other. Furthermore, VAR(1) forecasts do not provide, on average, additional information in terms of forecasting. Throughout the rest of the paper, I consider both specifications. Third, compared to a case of well-specified forecasts, frictions such as habit in consumption are significantly less important when mis-specified forecasts are combined with myopia. In particular, the presence of autoregressive forecasts strengthens the internal propagative features of the model by inducing excess persistence and volatility. Therefore, the proposed expectations formation process will often deliver aggregate variable responses to demand, cost-push, and monetary shocks that are more persistent and volatile, relative to a case of well-specified forecasts. Myopia, on the other hand, puts downward pressure on fluctuations, and as a result, the impulse response functions of aggregates can be more amplified when myopia is absent versus when it is present.

Finally, using the estimated posterior distribution to discipline the parameters of the two model specifications that best fit macroeconomic data, I simulate inflation annual forecasting data from the proposed expectations formation process and show that simulated inflation annual forecast errors support the three empirical facts on consensus forecasts as described in the preceding paragraphs.

Related Literature

The literature has shown that various assumptions on the expectations formation processes can

[9]Similar to the partial equilibrium setting, the implied optimal consumption and pricing rules of households and firms are of an infinite horizon nature and in the presence of mis-specified forecasting rules, they cannot be reduced to the standard one-period-ahead Euler equation and Phillips curve, respectively. If well-specified forecasting rules are used, the implied Euler equation and Phillips curve coincide with the behavioral ones in Gabaix (2020) if there is myopia and the standard FIRE ones if myopia is absent.

be consistent with forecasting data. For instance, Coibion and Gorodnichenko (2015) show that the predictability of ex-post forecast errors by ex-ante forecast revisions is consistent *only* with the assumptions of sticky information as in Mankiw and Reis (2002) and Reis (2006), or noisy information as in, e.g., Woodford (2003a), Sims (2003), and Maćkowiak and Wiederholt (2009).[10] However, sticky or noisy information alone cannot match forecasters' over-reaction to recent events (Kohlhas and Walther (2021)) or forecast errors' delayed over-shooting (Angeletos, Huo, and Sastry (2021)). Kohlhas and Walther (2021) rationalize the fact that consensus forecasts under-react to new information but over-react to recent events through a theory of asymmetric attention to procyclical variables. The current paper resonates with the findings in Angeletos, Huo, and Sastry (2021) with two differences. First, I show that the empirical evidence can be matched even with a *unique* deviation from FIRE, that is, by relying on autoregressive mis-specified forecasts only, whereas Angeletos, Huo, and Sastry (2021) show that *both* over-extrapolation and information friction are necessary. Second, in the current paper, over-extrapolation is an equilibrium outcome of agents relying on a mis-specified forecasting rule whose parameters are nonetheless disciplined. In Angeletos, Huo, and Sastry (2021), the distance between the perceived and actual persistence of shocks is disciplined from the impulse response functions of forecast errors.

The paper contributes additional evidence to a rich body of literature that validates the usage of simple forecasting processes by the private sector (e.g., Tversky and Kahneman (1973, 1974), Adam (2007), Hommes (2013), Greenwood and Shleifer (2014), Petersen (2015), Malmendier and Nagel (2016), and Hommes et al. (2019).[11] The paper is also related to a series of papers that discuss the analytical implications of mis-specified forecasting rules, as in Hommes and Sorger (1998), Fuster, Laibson, and Mendel (2010); Fuster, Hebert, and Laibson (2012), Hommes and Zhu (2014), Airaudo and Hajdini (2021), and Branch, McGough, and Zhu (2022), among others. In particular, the paper relies on the solution concept of a first-order CE equilibrium, developed by Hommes and Sorger (1998) and Hommes and Zhu (2014).

The paper shares a common idea with Gabaix (2020) about myopia being an excess discounting

[10] See Coibion, Gorodnichenko, and Kamdar (2018) as well for a review.

[11] Experimental evidence in Adam (2007), Hommes (2013), Petersen (2015), and Hommes et al. (2019), among others, shows that agents are commonly not model-based rational and that they tend to use simple forecasting rules. Using MSC micro data on inflation expectations, Malmendier and Nagel (2016) show that expectations are history dependent. From a psychological standpoint, Tversky and Kahneman (1973, 1974) argue that when trying to solve complex problems, people tend to employ a limited set of heuristics. Moreover, simpler processes generate on average smaller out-of-sample forecasting errors compared to AR(p) for $p > 1$ or VARs, especially for inflation series (see, for example, Atkeson and Ohanian (2001) and Stock and Watson (2007)).

of future deviations from the steady state, with the crucial difference that the present paper relies on mis-specified forecasts instead of well-specified ones. Evidence of myopia enhancing model fit that is presented in the current work further contributes to recent developments in the empirical literature in favor of myopic agents (see, for instance, Ganong and Noel (2019), who show that the only model that could rationalize household behavior given a predictable decrease in income in the data was one with myopic/short-sighted agents).

Furthermore, this work shares common insights with the literature that posits that deviations from FIRE amplify the propagative features of models and thus can explain aggregate data better. For instance, papers by Milani (2006, 2007), Slobodyan and Wouters (2012a), and Hommes et al. (2019) show that imperfect common knowledge explains observed persistence better relative to FIRE. More recent papers, such as Ilut and Saijo (2021) and Bianchi, Ilut, and Saijo (2022), rely on a local projection estimation of empirical impulse response functions to show that deviations from FIRE improve the propagative mechanism of New Keynesian models.[12] While a model set in a FIRE framework with a rich set of frictions as in Smets and Wouters (2003, 2007) can fit the data pretty well, this paper shows that a combination of autoregressive mis-specified forecasts and myopia is powerful in replicating the characteristics of business cycle fluctuations, with a diminished need for mechanical frictions.

Finally, the paper is related to that body of literature that estimates general equilibrium New Keynesian models free of the FIRE assumption, as in, for instance, Del Negro and Eusepi (2011), Slobodyan and Wouters (2012a,b), Ormeño and Molnár (2015), Rychalovska (2016), Gaus and Gibbs (2018), and Cole and Milani (2019).

The rest of the paper is organized as follows. Section 2 describes the expectations formation process in a New Keynesian pricing problem. Section 3 derives a number of implications about inflation forecast errors and tests them with evidence from consensus inflation forecasting data. Section 4 nests the expectations formation process in a full New Keynesian model and presents the main Bayesian estimation results accompanied by a series of implications. Section 5 re-evaluates the three empirical facts about consensus forecasting errors with inflation expectations

[12] Even though the current paper is fundamentally distinctive from Angeletos and Huo (2021), the empirical evidence presented here stands in favor of their analytical result that myopia and "anchoring of the current outcome to the past outcome" can be a substitute for mechanical persistence. Angeletos and Huo (2021) prove the equivalence between a FIRE model with incomplete information and another FIRE model with myopia along with "anchoring of the current income to the past outcome, as if there was habit." In contrast, in this paper, backward-looking components are an attribute of autoregressive mis-specified forecasting rules due to imperfect common knowledge, whereas myopia is realized through an adjustment process to mis-specified forecasting rules.

data simulated from the estimated general equilibrium model. Section 6 concludes.

2 Mis-specified Forecasts and Myopia

In what follows, I integrate a combination of mis-specified autoregressive forecasting rules and myopia into a partial equilibrium New Keynesian pricing problem and solve for the CE equilibrium. Apart from describing the expectations formation process, this section builds the foundation for deriving a number of testable implications for inflation forecasting data in the succeeding section. The rationale for focusing on a pricing problem and therefore expectations about inflation, instead of other macroeconomic variables, is due to their availability in survey data as well as due to their particular importance for macroeconomics. Moreover, since testing implications of various expectations assumptions on inflation forecasting data is the benchmark in the literature, I can naturally compare the present paper's expectations process with other alternatives.

2.1 New Keynesian Pricing

Following Woodford (2003b) and Galí (2008), I assume the economy is populated by a continuum of monopolistically competitive firms, $j \in [0,1]$. Each firm produces a differentiated good, but faces the same isoelastic demand schedule

$$y_{jt} = \left(\frac{P_{jt}}{P_t}\right)^{-\zeta} y_t \tag{1}$$

where $\zeta > 1$ is the elasticity of substitution among the differentiated goods, P_{jt} is the price set by the j^{th} firm, P_t is the aggregate price level, and y_t is the aggregate output level. The pricing problem is subject to Calvo price stickiness: each period firms cannot adjust their price with some constant probability $\alpha \in (0,1)$. Every firm then chooses its current optimal price P^*_{jt} that will maximize its present discounted value of real profits

$$\max_{P^*_{jt}} \widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h Q_{t+h} \left(\frac{P^*_{jt}}{P_t} y_{j,t+h} - mc_{t+h} y_{j,t+h} \right) \tag{2}$$

where $\widetilde{\mathbb{E}}_{jt}$ is a generic subjective expectations operator that satisfies the law of iterative expectations and standard probability rules; Q_{t+h} is a generic stochastic discount factor; $\hat{mc}_t$ is the marginal cost; $\hat{\pi}_t$ is inflation; $\beta \in (0,1)$ is a predetermined discount factor. The log-linearized first-order condition of each firm's pricing problem is given by[13]

$$\hat{p}^*_{jt} = \widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h \left((1-\alpha\beta)\hat{mc}_{t+h} + \alpha\beta\hat{\pi}_{t+h+1}\right) \tag{3}$$

where $\hat{p}^*_{jt} = log(P^*_{jt}/P_t)$ is the log-linear optimal price in deviation from the aggregate price $\hat{P}_t$. The marginal cost is exogenous and it evolves according to

$$\hat{mc}_t = \rho\hat{mc}_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2_\varepsilon) \tag{4}$$

with $\rho \in (0,1)$. Each firm faces the same problem as stated in (2), is subject to the same marginal cost shock in (4), and is endowed with the same beliefs about the future evolution of inflation and marginal costs; hence, they will all choose the same optimal price $\hat{p}^*_t$ in (3). However, firms are endowed with *imperfect common knowledge*, which impedes them from understanding their homogeneity; that is, each individual firm is not aware that every other firm relies on the same optimal pricing rule in (3). As shown in the subsequent subsections, this implies that the optimal pricing rule in (3) cannot be used to make inferences about future deviations of inflation from its steady state. Consequently, firms do not understand the true structure of the law of motion for inflation, and once aggregated, (3) will not produce the standard Phillips curve. For simplicity purposes only, I assume that firms understand that marginal costs evolve according to (4).[14] Since firms are assumed to be homogeneous, from now on, I drop the subscript j.

2.2 Myopia

Let $\widetilde{\mathbb{E}}^\star_t \hat{\pi}_{t+h}$ denote the forecast about future inflation - in deviation from its steady state - prior to myopic adjustment. $\widetilde{\mathbb{E}}^\star_t$ could be associated with a well-specified forecasting rule that would be the minimum state variable solution under FIRE, or it could otherwise be linked to a mis-specified forecasting rule such that forecasts are formed based on a rule that is structurally different from the minimum state variable solution under FIRE. To model myopia, I build on the idea of cognitive discounting in Gabaix (2020), where (well- or mis-specified) forecasts about future inflation and marginal cost in deviations from their steady-state values are discounted by a cognitive discount

[13] See Appendix A for more details.

[14] This assumption can be easily relaxed, without altering the main results of the paper.

factor, $n \in (0, 1]$. The parameter n defines the degree of myopia, such that a higher (respectively, lower) n relates to firms being more (respectively, less) forward-looking with respect to future fluctuations around the steady-state equilibrium. In particular,

$$\widetilde{\mathbb{E}}_t \begin{bmatrix} \hat{\pi}_{t+h} \\ \hat{mc}_{t+h} \end{bmatrix} = n^h \widetilde{\mathbb{E}}_t^{\star} \begin{bmatrix} \hat{\pi}_{t+h} \\ \hat{mc}_{t+h} \end{bmatrix} \tag{5}$$

As myopia increases, i.e., n decreases, the expected value of $\hat{\pi}_{t+h}$ gets closer to 0, or the expected value of future inflation approaches its steady state. Moreover, for $n \in [0, 1)$, as the forecasting horizon h increases, the myopic adjustment becomes more severe. For $n = 1$ myopia is shut down. Substituting for $\widetilde{\mathbb{E}}_t \begin{bmatrix} \hat{\pi}_{t+h} & \hat{mc}_{t+h} \end{bmatrix}'$ in (3), the optimal pricing decision becomes

$$\hat{p}_t^* = \widetilde{\mathbb{E}}_t^{\star} \sum_{h=0}^{\infty} (\alpha\beta n)^h \left((1-\alpha\beta)\hat{mc}_{t+h} + \alpha\beta n \hat{\pi}_{t+h+1}\right) \tag{6}$$

Differently from Gabaix (2020), who assumes that the myopic adjustment happens to the well-specified forecasting rule about future deviations of inflation from its target, the present paper assumes instead that the myopic adjustment occurs with respect to a mis-specified forecast about inflation. I describe the structure of mis-specified forecasts in what follows.

2.3 Mis-specified Forecasts

As mentioned earlier, firms are assumed to understand the process of the exogenous disturbances they are subject to; therefore, absent myopia, they correctly forecast the marginal cost,

$$\widetilde{\mathbb{E}}_t^{\star} \hat{mc}_{t+h} = \rho^h \hat{mc}_t \tag{7}$$

On the other hand, due to imperfect common knowledge, firms do not understand that every other firm in the economy faces the same optimal pricing rule as in (3). As a consequence, they do not use the aggregated version of (6) to make inferences about $\widetilde{\mathbb{E}}_t^{\star} \hat{\pi}_{t+h}$. Leveraging on a large body of evidence showing that economic agents form forecasts based on simple autoregressive rules (see for instance, Adam (2007), Hommes and Zhu (2014), and Malmendier and Nagel (2016), among others), I assume that inflation forecasts are based on an AR(1) process,[15]

$$\hat{\pi}_t = \delta + \gamma(\hat{\pi}_{t-1} - \delta) + \epsilon_t \tag{8}$$

where $\delta \in \mathbb{R}$ is the perceived unconditional mean of inflation, $\gamma \in (-1, 1)$ is the perceived unconditional first-order autocorrelation of inflation, and ϵ_t is perceived to follow a white noise process. The value of ϵ_t is unknown when firms forecast future inflation, therefore

$$\widetilde{\mathbb{E}}_t^\star \hat{\pi}_{t+h} = \delta(1 - \gamma^{h+1}) + \gamma^{h+1}\hat{\pi}_{t-1} \tag{9}$$

As shown in the following section, the pair (δ, γ) will be pinned down using the solution concept of a CE equilibrium. In other words, the only "free" behavioral parameter is n, which defines the degree of myopic adjustment to mis-specified forecasts.

2.4 Consistent Expectations Equilibrium

Aggregating equation (6) delivers an expression for inflation,

$$\hat{\pi}_t = \widetilde{\mathbb{E}}_t^\star \sum_{h=0}^{\infty} (\alpha\beta n)^h \left(\kappa \hat{mc}_{t+h} + \beta n(1-\alpha)\hat{\pi}_{t+h+1}\right) \tag{10}$$

where $\kappa = (1 - \alpha\beta)(1 - \alpha)/\alpha$. To reiterate, equation (10) cannot reduce to the standard one-step-ahead Phillips curve, because each firm - unaware that all the other firms set their optimal price according to (6) - cannot deduce that the dynamic equation for inflation is given by (10). Consequently, they do not form expectations about future inflation based on the expression for inflation in (10).

Remark 1: Equation (10) reduces to the Phillips curve of the behavioral New Keynesian model as in Gabaix (2020), $\hat{\pi}_t = \kappa\hat{mc}_t + \beta n \mathbb{E}_t \hat{\pi}_{t+1}$, *if* firms understand their homogeneity. In that case, $\widetilde{\mathbb{E}}_t^*$ is associated with a well-specified forecasting rule; that is, the structure of the forecasting rule is the same as the minimum state variable solution under FIRE. Therefore, firms can use their own optimal condition in (6) to form expectations about inflation, and the infinite-horizon New

[15] Note that (8) is structurally different from the minimum state variable solution of the partial equilibrium pricing model. Specifically, that would be $\hat{\pi}_t = \frac{(1-\alpha)(1-\alpha\beta)}{\alpha(1-\beta\rho)}\hat{mc}_t$.

Keynesian Phillips curve reduces to the one-step-ahead Phillips curve.[16]

Remark 2: From *Remark 1*, it follows that equation (10) reduces to the standard Phillips curve, $\hat{\pi}_t = \kappa \hat{mc}_t + \beta \mathbb{E}_t \hat{\pi}_{t+1}$, *only if* the expectations operator $\widetilde{\mathbb{E}}_t^*$ is associated with a well-specified forecasting rule *and* there is no myopia ($n = 1$).

Substituting for $\widetilde{\mathbb{E}}_t^\star \hat{mc}_{t+h}$ and $\widetilde{\mathbb{E}}_t^\star \hat{\pi}_{t+h}$ in (10) delivers the actual law of motion for inflation:

$$\hat{\pi}_t = \beta n \delta \left(\frac{1-\alpha}{1-\alpha\beta} - \frac{(1-\alpha)\gamma^2}{1-\alpha\beta n\gamma} \right) + \frac{\kappa}{1-\alpha\beta\rho n} \hat{mc}_t + \frac{\beta n (1-\alpha)\gamma^2}{1-\alpha\beta n\gamma} \hat{\pi}_{t-1} \tag{11}$$

Firms believe that the equation in (8) is a valid perceived law of motion for inflation if and only if its parameters, which represent the perceived unconditional mean (δ) and first-order autocorrelation (γ), are consistent with the same moments from the data-generating process for inflation in (11). Coefficients δ and γ in equilibrium are pinned down through the solution concept of a CE equilibrium, as defined by Hommes and Zhu (2014):

Definition 1 *A pair (δ^*, γ^*), where δ^* and γ^* are real numbers with $\gamma \in (-1, 1)$, is a first-order consistent expectations equilibrium if the stationary stochastic process defined by (11) has unconditional mean δ^** and *unconditional first-order autocorrelation coefficient γ^*.*

Along the FIRE equilibrium, firms would be matching the perceived *distribution* of inflation with its actual/realized *distribution*. Along the CE equilibrium, however, firms are only matching certain perceived unconditional *moments* of the distribution with the actual unconditional *moments*.

Proposition 1 *Let the data-generating process for inflation be described by equation (11). Then, there exists a unique consistent expectations equilibrium (δ^*, γ^*), where $\delta^* = 0$ and $\gamma^* \in [\rho, 1)$.*

[16] Specifically, let $\widetilde{\mathbb{E}}_t^\star = \mathbb{E}_t$ be the RE operator. From (10),

$$\mathbb{E}_t \hat{\pi}_{t+1} = \kappa \mathbb{E}_t \hat{mc}_{t+1} + n\beta(1-\alpha)\mathbb{E}_t \hat{\pi}_{t+2} + \mathbb{E}_t \sum_{h=1}^{\infty} (\alpha\beta n)^h \left(\kappa \hat{mc}_{t+h} + \beta n (1-\alpha) \hat{\pi}_{t+h+1} \right)$$

Hence,

$$\hat{\pi}_t = \kappa \hat{mc}_t + \beta n (1-\alpha) \mathbb{E}_t \hat{\pi}_{t+1} + \underbrace{\mathbb{E}_t \sum_{h=1}^{\infty} (\alpha\beta n)^h \left(\kappa \hat{mc}_{t+h} + \beta n (1-\alpha) \hat{\pi}_{t+h+1} \right)}_{\alpha\beta n \mathbb{E}_t \hat{\pi}_{t+1}} = \kappa \hat{mc}_t + \beta n \mathbb{E}_t \hat{\pi}_{t+1}$$

Proof. See Appendix C.1. ■

Proposition 1 shows that in the partial equilibrium pricing problem, the CE equilibrium exists, it is unique, and that, importantly, it generates a higher inflation persistence relative to the case of well-specified forecasting rules (with or without myopia), i.e., $\gamma^* \geq \rho$. Given the CE equilibrium (δ^*, γ^*) as described in Proposition 1, the mis-specified forecast and actual law of motion of inflation along the CE equilibrium path are, respectively

$$\widetilde{\mathbb{E}}_t^\star \hat{\pi}_{t+h} = (\gamma^*)^{h+1} \hat{\pi}_{t-1} \tag{12}$$

$$\hat{\pi}_t = \underbrace{\frac{\kappa}{1 - \alpha\beta\rho n}}_{a} \hat{mc}_t + \underbrace{\frac{\beta n (1-\alpha)(\gamma^*)^2}{1 - \alpha\beta n (\gamma^*)}}_{b} \hat{\pi}_{t-1} \tag{13}$$

Corollary 1 *Consider the actual law of motion for inflation along the CE equilibrium defined by (13). Then, the following statements are true.*

i) Higher price stickiness (higher α) leads to lower γ^* in equilibrium.

ii) A higher degree of myopia (lower n) leads to lower γ^* in equilibrium.

Proof. See Appendix C.2. ■

Corollary 1 shows that the price stickiness and degree of myopia play an important role in the occurrence of endogenous over-extrapolation. Specifically, as prices become stickier, the dependence of current inflation on backward-looking expectations drops, thus leading to lower inflation persistence in equilibrium. On the other hand, as firms become more forward-looking with respect to future fluctuations of inflation around its steady-state value, the persistence of inflation well exceeds the inertia of the marginal cost. Furthermore, the actual law of motion for inflation in equilibrium in (13) resembles the one that would be derived under FIRE in a setting with inflation indexation/backward-looking pricing (even though these features are missing in the firm's problem presented in this section).

Note that the CE solution for inflation in (13) differs structurally from the one where the forecasting rules remain well-specified (with or without myopia), which describes inflation as a linear function of the marginal cost shock only, i.e.,

$$\hat{\pi}_t = \frac{\kappa}{1 - \beta\rho n} \hat{mc}_t \tag{14}$$

When firms are appropriately forward-looking, i.e., $n = 1$, the inflation solution in (14) is the one implied under full-information RE. On the other hand, when firms are absolutely myopic toward the future ($n = 0$), the CE solution for inflation in (13) collapses to the one in (14).

3 Forecasting Data Evidence and Theory

To assess the empirical relevance of the proposed expectations formation process in Section 2, I first briefly describe the evidence on the behavior of consensus inflation forecasting errors from both professional forecasters and private agents. Building on the New Keynesian pricing problem in Section 2, I then derive three testable implications for forecasting errors when agents rely on a mis-specified forecasting rule and myopia to form inflation expectations.

I prove that a combination of mis-specified forecasts and myopia is consistent with the empirical evidence on consensus inflation forecasts. However, I show that while mis-specified forecasting rules are necessary *and* sufficient to match the empirical evidence along the CE equilibrium, myopia is neither. Importantly, these findings suggest that a *unique* deviation from FIRE, that is, usage of autoregressive mis-specified forecasts along a CE equilibrium, is more than enough to match all three empirical facts on forecast error behavior.

3.1 Brief Overview of the Evidence

Fact 1: Angeletos, Huo, and Sastry (2021) have brought forward evidence that the impulse response function (IRF) of annual inflation forecasting errors, following a supply or demand shock, is initially positive but turns negative at some later point in time.[17] The authors define this phenomenon as *delayed over-shooting*.[18]

Fact 2: Coibion and Gorodnichenko (2015) consider regressing ex-post annual forecast errors on ex-ante forecast revisions, that is

$$\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+h} = c + K_h(\widetilde{\mathbb{E}}_t \hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_{t-1} \hat{\pi}_{t+h}) + error_{t+h} \tag{15}$$

where $h = 4$ is the annual forecast horizon, c is a constant term, and K_h measures forecasters'

[17] Angeletos, Huo, and Sastry (2021) rely on forecasting data from the US SPF, Blue Chip, and the Michigan Survey of Consumers.

[18] I refer the reader to Section 5 in Angeletos, Huo, and Sastry (2021) for a description of their IRF estimation methodologies, as well as Figures 3 and 4 in their paper for a visualization of delayed over-shooting.

reaction to *new* information at the time of forecast. The authors find that K_4 is significantly positive, implying that forecasters under-react to new information at the time of forecast.[19] For instance, Table 1 exhibits the estimates of K_4 for US SPF forecasting data when inflation is measured as the growth rate of the GDP deflator and of the consumer price index (CPI). The table shows that forecasters generally under-react to new information.[20]

	GDP deflator				*CPI*	
	1970:2-2020:1		1982:3-2020:1		1982:3-2020:1	
Revision, K_4	1.01		0.11		0.72	
	(0.16)		(0.18)		(0.18)	
Current, M_4		0.05		-0.15		-0.11
		(0.03)		(0.04)		(0.05)

Table 1: Estimates of regression coefficients in (15) and (16) for inflation using US SPF annual forecasting data. The estimates of K_4 and M_4 when inflation is measured as the growth rate of the CPI are borrowed from Kohlhas and Walther (2021) (see Table C.7 in their online appendix). All regressions include a constant term. Standard errors are given in parenthesis.

Fact 3: Kohlhas and Walther (2021) instead consider regressing ex-post annual inflation forecast errors on inflation realized at the time of forecast, that is,

$$\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+h} = c + M_h \hat{\pi}_t + error_{t+h} \tag{16}$$

where M_h measures forecasters' reaction to information about inflation at the time of forecast. Kohlhas and Walther (2021) find that M_h is significantly negative for annual forecast errors ($h = 4$), implying that forecasters over-react to inflation realized at the time of forecast.[21] Table 1 shows that the estimate of M_4 for US SPF forecasting data is negative when considering inflation forecasting data from the end of 1982 onward.[22]

[19] Coibion and Gorodnichenko (2015) rely on inflation forecasting data from the US SPF, MSC, Livingston Survey, and inflation expectations from the Cleveland Fed based on the method developed in Haubrich, Pennacchi, and Ritchken (2008).

[20] When inflation is measured as the growth rate of the GDP deflator, the K_4 estimate for US SPF forecasting data loses significance after 1982; see footnote 22 for more details.

[21] Kohlhas and Walther (2021) rely on inflation forecasting data from the US SPF, MSC, Livingston Survey, and Euro Area SPF.

[22] See Hajdini and Kurmann (2022) for a detailed discussion regarding the instability of the estimates of K_4 and M_4.

3.2 Facts 1, 2, and 3: Linking Theory with Evidence

I now analyze the implications of mis-specified forecasts and myopia for the behavior of forecast errors along the CE equilibrium path.

Starting with delayed over-shooting, Proposition 2 shows that forecasters initially under-shoot and then over-shoot when there is sufficiently high endogenous over-extrapolation, that is, when $\gamma^* >> \rho$. Note that for $n = 1$, delayed over-shooting is always satisfied, since the condition in Proposition 2 translates into $\gamma^* > \rho$, and Proposition 1 proved that to always be the case. On the contrary, if $n = 0$, the condition for late over-reaction fails to hold, regardless of how the model is parameterized.

Proposition 2 *Let $\mathbb{I}_{k,h}$ be the impulse response function of the h-period-ahead forecasting error at period $(t+k)$ for $k \in \{0, 1, 2, ...\}$ with respect to a one-time shock to the marginal cost ε_t, i.e.,*

$$\mathbb{I}_{k,h} = \frac{\partial(\hat{\pi}_{t+k} - \widetilde{\mathbb{E}}_{t+k-h}\hat{\pi}_{t+k})}{\partial \varepsilon_t} \tag{17}$$

Let the expectations formation process be a combination of autoregressive mis-specified forecasting rules and myopia. Then, delayed over-shooting occurs if $\rho^{h+1} < n^h(\gamma^*)^{h+1}$.

Proof. See Appendix C.3. ■

Figure 1 visualizes the results of Proposition 2 for annual forecast errors ($h = 4$) along the CE equilibrium. Importantly, the figure also speaks to the intuition for why sufficiently high over-extrapolation leads to late over-shooting. Given the backward-looking nature of the expectations formation process, following a one-time shock to the marginal cost, expectations pick up with a lag. Consequently, the momentum in the response of inflation will be reflected in expectations at a later point in time, after the response of inflation has started dissipating (note the difference in the timing when the blue and red lines reach their peak in panels (b) and (c) of Figure 1). The more persistent inflation is in equilibrium; that is, the more over-extrapolation there is, the more amplified the forecasts and hence the higher is the likelihood they exceed ex-post realized inflation as its response approaches 0.

Next, Proposition 3 provides expressions for the model-implied K_h and M_h, for any forecasting horizon $h > 0$. As shown in the proposition, K_h and M_h are composed of positive and negative components, and the signs of both moments depend on the underlying parameters of the model.

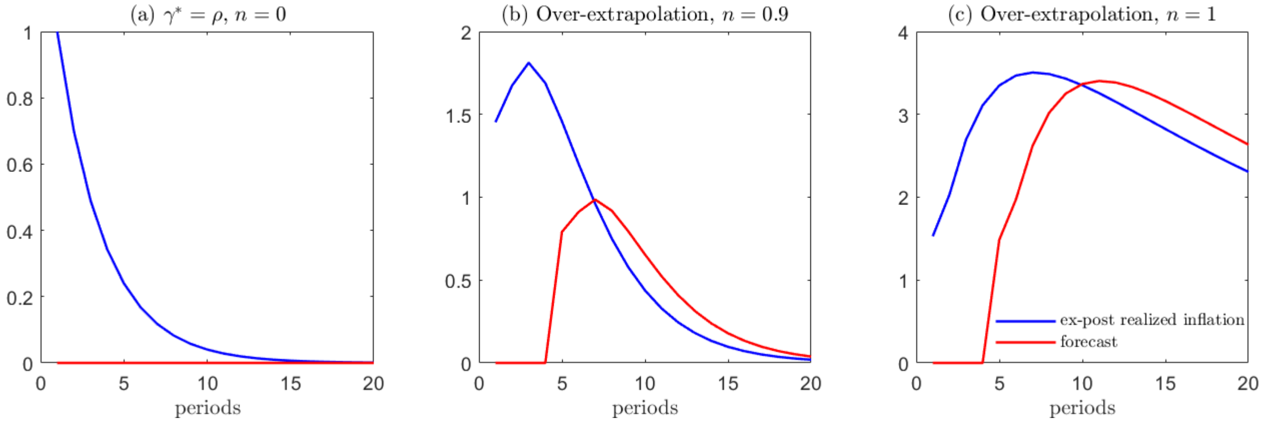


Figure 1: Evolution of the IRFs of the annual forecasts and ex-post realized inflation for various values of n. Parameterization: $\alpha = 0.5$, $\beta = 0.99$, $\kappa = 1$, $\rho = 0.7$, $\sigma = \sqrt{5}$. The implied equilibrium first-order autocorrelation coefficients of inflation are, respectively from left to right, $\gamma^* = 0.7$, $\gamma^* = 0.904$, and $\gamma^* = 0.993$.

Proposition 3 *Let the data-generating process for inflation be described by (13), with the expectations formation process being a combination of mis-specified forecasts and myopia. Then, the model-implied K_h and M_h, for any $h > 0$, are, respectively, given by*

$$\begin{aligned} K_h &= K_h^+ + K_h^- \\ M_h &= M_h^+ + M_h^- \end{aligned} \tag{18}$$

where $K_h^+ = \frac{\rho^h(1-b^2)(1-n\rho\gamma^*)\left(\rho + b\sum_{j=0}^{h-1}\left(\frac{b}{\rho}\right)^j\right) + b^{h+1}(\rho(b-n\gamma^*)+1-nb\gamma^*)}{n^h(\gamma^*)^{h+1}(1+n^2(\gamma^*)^2-2n(\gamma^*)^2)(1+\rho b)} \geq 0$; $K_h^- = -\frac{\rho(b-n\gamma^*)+1-nb\gamma^*}{(1+n^2(\gamma^*)^2-2n(\gamma^*)^2)(1+\rho b)} \leq 0$; $M_h^+ = \frac{\rho^h}{1+.\rho b}\left[\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j - b^2\sum_{j=0}^{h-2}\left(\frac{b}{\rho}\right)^j\right] \geq 0$; $M_h^- = -\frac{n^h(\gamma^*)^{h+1}(b+\rho)}{1+\rho b} \leq 0$.

Proof. See Appendix C.4. ∎

Together, Propositions 2 and 3 show that there can exist parameterizations of the model, including the degree of myopia, for which a combination of the three facts is matched.

Since it is impossible to analytically derive parametric spaces for which all three pieces of empirical evidence are matched, I resort to numerical methods. In particular, I investigate the parametric region of myopia, n, and marginal cost inertia, ρ, for which the aforementioned empirical facts are matched. Figure 2 exhibits the regions where, in equilibrium, there is delayed over-shooting (area to the right of the dashed red curve); under-reaction to ex-ante forecast revisions (areas in white and dark gray); and over-reaction to current inflation (areas in white and light gray).

I now turn to the parametric region of myopia, n, and shock inertia, ρ, for which all three

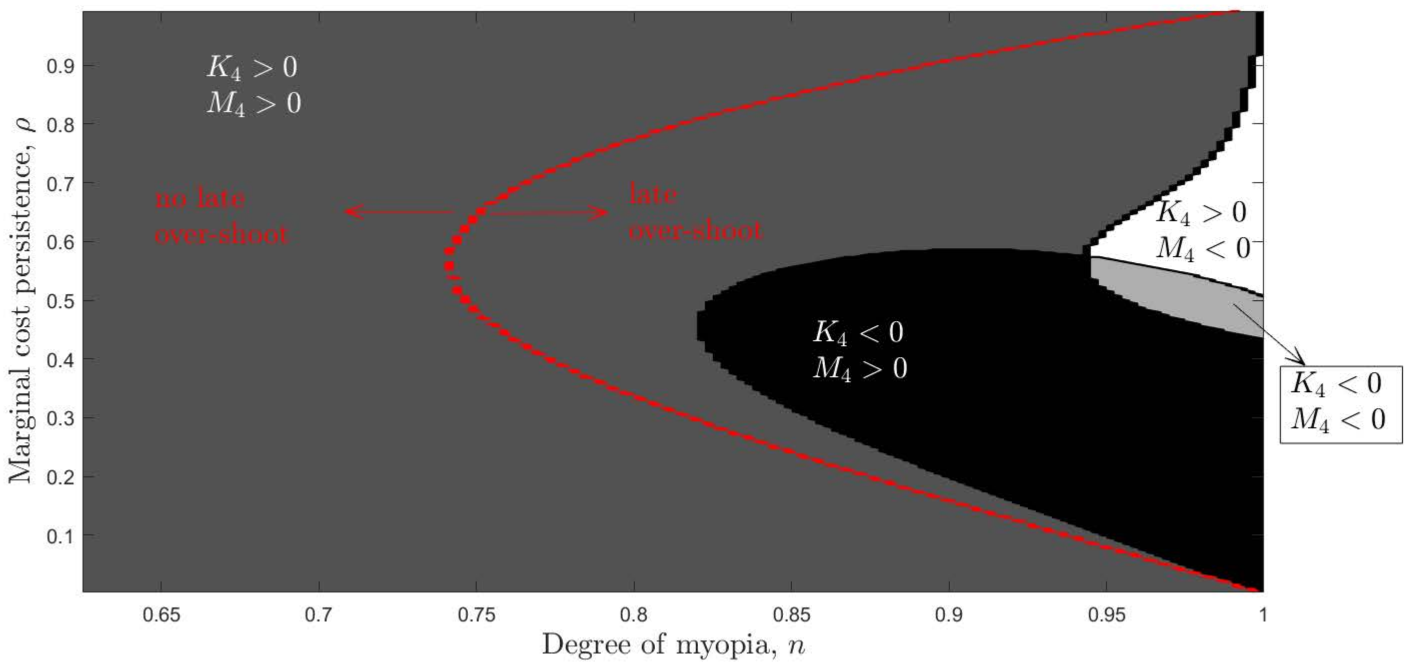


Figure 2: Regions of delayed over-shooting, under-reaction to ex-ante forecast revisions, and over-reaction to current realizations. Delayed over-shooting: region to the right of the dashed red curve. Under-reaction to ex-ante forecast revisions: regions in white and dark gray. Over-reaction to current inflation: regions in white and light gray. Forecasting horizon: $h = 4$. The area for which $n < 0.625$ is truncated for better visibility, but it is a region of no late over-shooting and characterized by $K_4 > 0$, $M_4 > 0$. Parameterization: $\alpha = 0.5$, $\beta = 0.99$, $\kappa = 1$, $\sigma = \sqrt{5}$.

aforementioned empirical facts for annual forecast errors ($h = 4$) are matched. Figure 2 exhibits the regions where, in equilibrium, there is delayed over-shooting (area to the right of the dashed red curve), under-reaction to ex-ante forecast revisions (areas in white and dark gray), and over-reaction to current inflation (areas in white and light gray). The model parameters are set as follows: $\alpha = 0.5$, $\beta = 0.99$, $\kappa = 1$, and $\sigma = \sqrt{5}$.

As shown in Proposition 2, the condition for delayed over-shooting necessitates sufficiently high endogenous over-extrapolation, which, as exhibited in Figure 2, is guaranteed to occur for relatively low degrees of myopia. In line with a similar reasoning, over-reaction to current inflation also requires sufficiently low myopia. On the contrary, under-reaction to ex-ante forecast revisions occurs for a much wider range of (n, ρ) pairs, and for $n < 0.8$ it is guaranteed for any value of the shock persistence. The rationale is that a higher degree of myopia puts downward pressure on forecasts and hence enables a positive correlation between forecast errors and ex-ante revisions. Finally, depending on the parameterization of the model and as hinted by Proposition 3, K_4 and M_4 can take both positive and negative values.

Figure 3 exhibits a sensitivity analysis of the numerical results presented in Figure 2 for different levels of price stickiness. As the degree of price stickiness (α) increases from panel (a) to panel (c), the regions change as follows. First, the dashed red curve shifts to the right, since, as shown in Corollary 1, higher price stickiness puts downward pressure on the equilibrium inflation persistence γ^*. The latter implies that lower degrees of myopia are necessary to grant sufficient endogenous over-extrapolation that would give rise to delayed over-shooting. Second, in line with a similar reasoning, the black region shrinks to the right and the white region shifts upward.

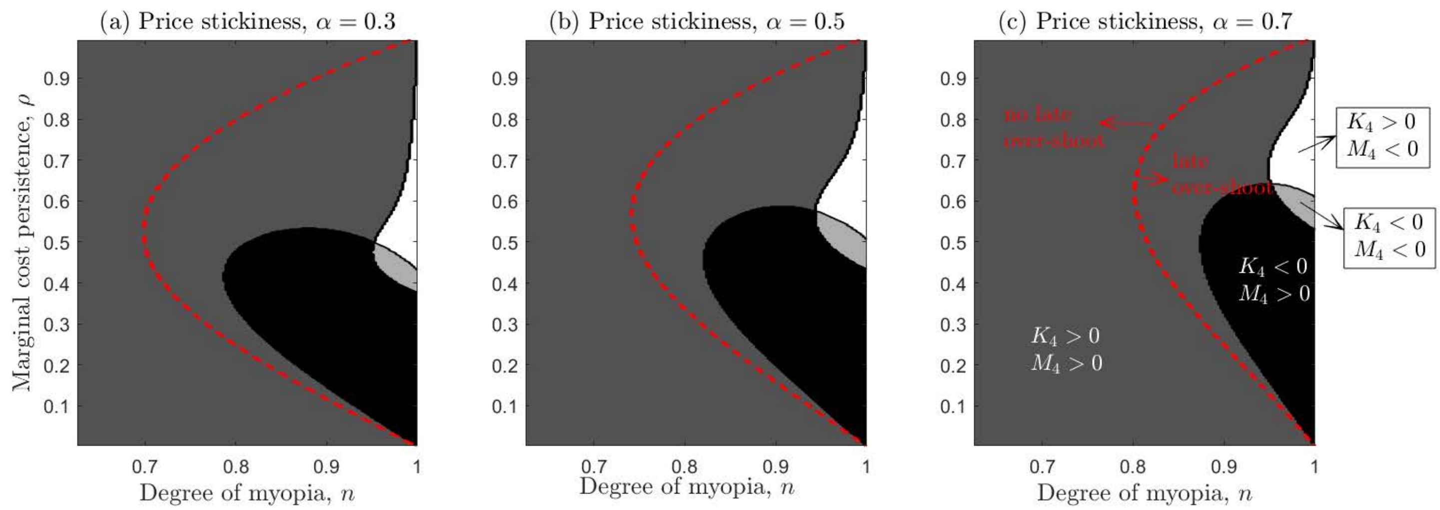


Figure 3: Sensitivity analysis with respect to price stickiness. Interpretation of regions, the forecasting horizon, and parameterization of the model are the same as in Figure 2.

Next, I derive results for the cases when i) there is extreme myopia; and ii) there is no myopia. More specifically, Corollary 2 shows that in the presence of extreme myopia, that is, when agents' forecast for inflation approaches the steady state in every period ($n \to 0$), the behavior of forecasting errors is consistent *only* with Fact 2.

Corollary 2 *Suppose there is extreme myopia, that is, $n \to 0$, and that the marginal cost has some persistence, that is, $\rho > 0$. Then,* only *Fact 2 is matched along the CE equilibrium for any parameterizations of the model, with $lim_{n \to 0} K_h = \infty > 0$.*

Proof. Follows directly from Propositions 2 and 3. ■

On the contrary, Corollary 3 proves that when there is no myopia, that is, for $n = 1$, there exist parameterizations of the model for which *all* three facts are matched. Note that in the case of no myopia, delayed over-shooting is always satisfied.

Corollary 3 *Suppose there is no myopia, i.e., $n = 1$. Then, Facts 1, 2, and 3 are matched for parameterizations of the model that satisfy the following two inequalities along the CE equilibrium:*

$$\underline{\gamma} < (\gamma^*)^{h+1} < \bar{\gamma} \tag{19}$$

where $\underline{\gamma} = \frac{\rho^{h+1} - b^{h+1} - \rho^2 b^2 (\rho^{h-1} - b^{h-1})}{\rho^2 - b^2}$ *and* $\bar{\gamma} = \rho^{h+1} + \rho^h b + ... + \rho b^h + b^{h+1}$.

Proof. See Appendix C.5 ■

For instance, applying the result in Corollary 3, one can quickly show that the empirical evidence is matched for $\rho \to 1$. To see this, note that $\lim_{\rho \to 1} \gamma^* = 1$, while $\gamma^* > \rho$, and $\lim_{\rho \to 1} \bar{\gamma} > 1$. Furthermore, $\lim_{\rho \to 1} \frac{\underline{\gamma}}{(\gamma^*)^{h+1}} = \left(\frac{\rho}{\gamma^*}\right)^{h+1} < 1$.

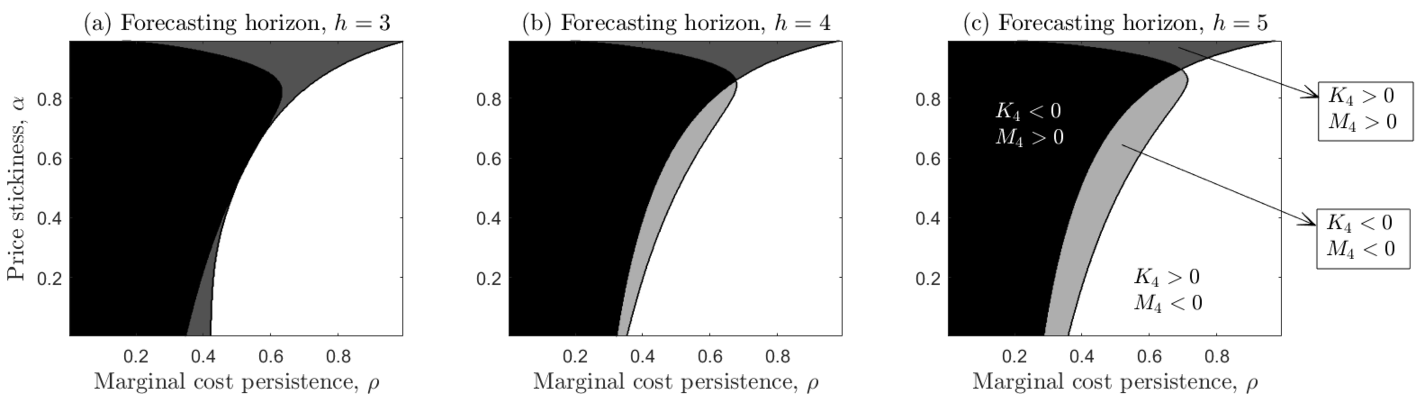


Figure 4: Sensitivity analysis with respect to the forecasting horizon, where there is no myopia, that is, $n = 1$. Interpretation of regions, the forecasting horizon, and the rest of the model parameterization are the same as in Figure 2.

Elaborating more on the case of no myopia, Figure 4 visualizes the parametric space (ρ, α) across various forecasting horizons for which different facts are matched. The colors in the figure are interpreted as in Figure 2. The white region where all three facts are matched is substantial, regardless of the forecasting horizon. For example, for any degree of price stickiness, there exists a range of values for the marginal cost persistence for which forecasts exhibit delayed over-shooting, under-reaction to new information, and over-reaction to current inflation.

Last, I focus on the implications of a well-specified forecasting rule for the occurrence of delayed over-shooting, as well as K_h and M_h. In particular, Proposition 4 shows that, regardless of the degree of myopia, a well-specified forecast can *only* match Fact 2.[23]

[23]It is clear from Proposition 4 that if one sets $n = 1$, that is, if one imposes FIRE, none of the facts would be matched in the current setting.

Proposition 4 *Suppose the forecasting rules are well-specified, such that* $\widetilde{\mathbb{E}}_t\hat{\pi}_{t+h} = n^h\mathbb{E}_t\hat{\pi}_{t+h}$*, where* $\mathbb{E}_t$ *is the RE operator. Then, the following statements are true:*

1. $\mathbb{I}_{k,h} \geq 0$ *for any* $k \geq 0$*; there is no delayed over-shooting.*

2. $K_h = \frac{(1-n\rho^2)(1-n^h)}{n^h(1+n^2\rho^2-2n\rho^2)} \geq 0$*; there is always under-reaction to ex-ante forecast revisions.*

3. $M_h = \rho^h(1-n^h) \geq 0$*; there is always under-reaction to current inflation realizations.*

Proof. See Appendix C.6. ∎

Taken together, Corollary 3 and Proposition 4 highlight that a mis-specified forecasting rule of the form considered here is *necessary* and *sufficient* for the behavior of forecasting errors to be consistent with the empirical evidence. By contrast, myopia is neither a necessary nor a sufficient condition. In fact, while sufficiently low degrees of myopia in the presence of mis-specified forecasts allow for the three facts to be matched, too much myopia does not. Importantly, the analysis in this section shows that, to match all three facts, *only* one deviation from FIRE is sufficient.

Relation to other expectations formation processes. The present paper echoes the findings of Angeletos, Huo, and Sastry (2021), but it is different in two dimensions. First, the current paper shows that all three empirical facts can be matched even with a *unique* deviation from FIRE, that is, through relying on a mis-specified forecasting rule whose mean and first-order autocorrelation are consistent with the actual law of motion. Angeletos, Huo, and Sastry (2021) show that all three facts can be matched for *combinations* of noisy information and over-extrapolation. Second, in the current paper, over-extrapolation is an equilibrium outcome of agents relying on simple forecasting rules that are *structurally* different from the actual law of motion. While the structure of the mis-specified forecasting rule is - in a sense - exogenously chosen, its parameters are disciplined through the CE equilibrium concept.[24] In Angeletos, Huo, and Sastry (2021), the shock's perceived and actual persistence are pinned down from empirically matching the impulse response functions of forecasting errors.

In Kohlhas and Walther (2021) forecasters' under-reaction to ex-ante forecast revisions but over-reaction to recent events are rationalized through a theory of asymmetric attention to pro-cyclical variables. Nevertheless, as mentioned in Angeletos, Huo, and Sastry (2021), asymmetric

[24] As shown in the paper, in the current setting, a CE equilibrium is guaranteed to exist for any realistic parameterization of the model. However, one can think of other mis-specified forecasting rules that do not yield a CE equilibrium for given realistic parameterizations, in which case that particular structure of forecasting rule would never be used.

attention cannot match delayed over-shooting. Finally, Coibion and Gorodnichenko (2015) show that equation (19) is consistent with the assumptions of sticky information (see, e.g., Mankiw and Reis (2002) and Reis (2006)), and noisy information (see, e.g., Woodford (2003a), Sims (2003), and Maćkowiak and Wiederholt (2009)). However, as shown by Angeletos, Huo, and Sastry (2021), such informational frictions alone are not sufficient to be consistent with all three pieces of evidence in the literature. Specifically, some degree of over-extrapolation is necessary to fit the results of forecasters' over-reaction to recent events in Kohlhas and Walther (2021) and delayed over-shooting in Angeletos, Huo, and Sastry (2021).

4 General Equilibrium Model

Given the evidence presented in the previous section, I nest the proposed expectations formation process, namely, a combination of mis-specified forecasts and myopia, into an otherwise baseline New Keynesian DSGE model with habit formation in consumption and inflation indexation. Bayesian estimation of the model on US aggregate data seeks to mainly reveal i) the preferred forecasting process; ii) the estimated value of the degree of myopia; and iii) the relative role of mis-specified forecasting rules and myopia for macroeconomic fluctuations.

4.1 Basics

The model is fairly standard; hence, I delegate all details to Appendix B.

Households. There is a continuum of identical households, $i \in [0, 1]$, that are unaware of each other's homogeneity. They consume from a set of differentiated goods, supply labor, and invest in riskless one-period bonds.[25] The consumption bundle of each household over the set of differentiated goods, $j \in [0, 1]$, is determined by the Dixit-Stiglitz aggregator. The optimal demand of the i^{th} household for the j^{th} good is given by

$$c_{it}(j) = \left(\frac{P_{jt}}{P_t} \right)^{-\zeta} c_{it} \tag{20}$$

where P_{jt} is the price of the j^{th} good and $\zeta > 1$ is the elasticity of substitution among the differentiated goods. Each period, the household receives labor income and dividends from the

[25] Bonds are assumed to be in zero net supply.

monopolistically competitive firms, and it maximizes its expected lifetime utility with respect to the deviation of current consumption from a stock of internal habits in consumption, labor supply, and bonds, subject to a budget constraint. The problem each household faces is

$$\max_{c_{it},H_{it},B_{it}} \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \beta^h \xi_{t+h} \left(\frac{(c_{it} - \eta c_{i,t-1})^{1-\sigma}}{1-\sigma} - \psi \frac{H_{it}^{1+\varphi}}{1+\varphi} \right) \tag{21}$$

s.t.

$$\frac{R_{t-1}}{\pi_t} b_{i,t-1} = b_{it} - w_t H_{it} - d_{it} + c_{it} \tag{22}$$

where $\beta \in (0,1)$ is the discount factor; $0 \leq \eta < 1$ measures the degree of habit in consumption; σ is the inverse intertemporal elasticity of substitution; $\widetilde{\mathbb{E}}_{it}$ is the expectations operator described in the previous section; ξ_t is a preference shock; H_{it} is labor supply; R_{t-1} is the gross return on the past period's real bond choice $b_{i,t-1}$; w_t is the real wage and d_{it} denotes real dividends from firms.

Solving the household's optimization problem, log-linearizing around the steady-state equilibrium, and applying the myopic adjustment process delivers

$$\begin{aligned} \hat{x}_t = {} & \frac{\eta}{1+n\eta\upsilon} \hat{x}_{t-1} + n \frac{\upsilon - n\beta\eta(1-\beta)(1-\eta)}{1+n\eta\upsilon} \widetilde{\mathbb{E}}_t^{\star} \hat{x}_{t+1} + \frac{\beta n^2 (1-\beta)(1-\eta)(1-n\beta\eta)}{1+n\eta\upsilon} \widetilde{\mathbb{E}}_t^{\star} \sum_{h=0}^{\infty} (\beta n)^h \hat{x}_{t+h+2} \\ & - \widetilde{\mathbb{E}}_t^{\star} \sum_{h=0}^{\infty} (\beta n)^h \frac{1-\beta\eta}{\sigma(1+n\eta\upsilon)} \left(\hat{R}_{t+h} - \hat{\pi}_{t+h+1} - \hat{e}_{t+h} \right) \end{aligned} \tag{23}$$

where $\hat{x}_t$ is the output gap and $\upsilon = (1 - \beta + \beta\eta)$. The variable $\hat{e}_t$ is a demand shock assumed to follow an AR(1) process

$$\hat{e}_t = \rho_e \hat{e}_{t-1} + \sigma_e \varepsilon_t^e,\ \varepsilon_t^e \sim \mathcal{N}(0,1) \tag{24}$$

with $\rho_e \in [0,1)$.

Firms. The problem of the monopolistically competitive firms is similar to what was described in Section 2, with the major difference that the marginal cost is now endogenous. The j^{th} firm combines an exogenously given technology, z_t, with labor input, h_{jt}, to produce output y_{jt} as follows

$$y_{jt} = z_t h_{jt}^{a_h} \tag{25}$$

where $a_h \in (0,1]$. As in Section 2, firms cannot adjust their price each period with probability $\alpha \in [0,1)$. However, if the j^{th} firm cannot optimize its price in period t, it can still adjust the price according to the following indexation rule (see Christiano, Eichenbaum, and Evans (2005)),

$$P_{jt} = P_{j,t-1}\pi_{t-1}^{\rho_\pi} \tag{26}$$

where $0 \leq \rho_\pi < 1$ measures the degree of indexation to past inflation. Each firm will choose its optimal price P_{jt}^* that maximizes the present discounted value of real profits, i.e.,

$$\max_{P_{jt}^*} \widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h Q_{t+h} \left(\frac{P_{jt}^*}{P_{t+h}} \left(\frac{P_{t+h-1}}{P_{t-1}} \right)^{\rho_\pi} y_{j,t+h} - w_{t+h} h_{j,t+h} \right) \tag{27}$$

s.t.

$$y_{jt} = \left(\frac{P_{jt}^*}{P_t} \right)^{-\zeta} y_t \tag{28}$$

where Q_{t+h} is a generic stochastic discount factor. The aggregate price level is linked to the aggregate optimal price level P_t^* as described below

$$P_t = \left[\alpha \left(P_{t-1} \left(\frac{P_{t-1}}{P_{t-2}}^{\rho_\pi} \right) \right)^{1-\zeta} + (1-\alpha)(P_t^*)^{1-\zeta} \right]^{\frac{1}{1-\zeta}} \tag{29}$$

After solving the firm's optimization problem, log-linearizing around the steady-state equilibrium, and applying the myopic adjustment process, I derive the aggregate supply as follows[26]

$$\begin{aligned} \hat{\pi}_t = &\frac{1}{1+n\beta\rho_\pi(1-\alpha)}(\rho_\pi\hat{\pi}_{t-1} - \kappa\eta\tau\hat{x}_{t-1}) + \frac{\kappa(\omega + \tau(1 - n\eta\beta(\alpha-\eta))}{1+n\beta\rho_\pi(1-\alpha)}\hat{x}_t + \frac{1}{1-\alpha\beta n\rho_u}\hat{u}_t \\ &+ \frac{n\beta}{1+n\beta\rho_\pi(1-\alpha)}\widetilde{\mathbb{E}}_t^\star \sum_{h=0}^{\infty} (\alpha\beta n)^h \left((1-\alpha)(1-\alpha\beta n\rho_\pi)\hat{\pi}_{t+h+1} + \kappa(\alpha\omega + \tau(\alpha-\eta)(1-\alpha\beta n\eta)\hat{x}_{t+h+1}\right) \end{aligned} \tag{30}$$

where $\tau = \frac{\sigma}{1-\beta\eta}$, $\kappa = \frac{(1-\alpha)(1-\alpha\beta)}{\alpha(1+\omega\zeta)}$; $\omega = \frac{1+\varphi-a_h}{a_h}$. The variable $\hat{u}_t$ is a cost-push shock assumed to follow an AR(1) process,

$$\hat{u}_t = \rho_u\hat{u}_{t-1} + \sigma_u\varepsilon_t^u,\ \varepsilon_t^u \sim \mathcal{N}(0,1) \tag{31}$$

with $\rho_u \in [0,1)$.

Monetary policy. The central bank controls nominal interest rates through a standard Taylor rule that reacts to deviations of inflation from its target $\bar{\pi}$ and deviations of the output gap x_t from its steady state, while smoothing the interest rate path with some degree $\rho_r \in [0,1)$. The

[26] I assume that households and firms share the same degree of myopia, n. This is an assumption that can be easily relaxed.

log-linearized policy rule is

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1-\rho_r)\phi_\pi \hat{\pi}_t + (1-\rho_r)\phi_x \hat{x}_t + \sigma_v \varepsilon_t^v,\ \varepsilon_t^v \sim \mathcal{N}(0,1) \tag{32}$$

Model in matrix form. Let $\Theta = \{\alpha, \beta, n, \sigma, \kappa, \eta, \rho_\pi, \omega, \phi_\pi, \phi_x, \rho_r, \rho_e, \rho_u, \sigma_e, \sigma_u, \sigma_v\}$. Then the model can be compactly written in matrix form as

$$A_0(\Theta)S_t = A_1(\Theta)S_{t-1} + \widetilde{\mathbb{E}}_t^\star \sum_{h=0}^{\infty} (F(\Theta))^h A_2(\Theta) S_{t+h+1} + B(\Theta)\mathcal{E}_t \tag{33}$$

where $S_t = \begin{bmatrix} \hat{x}_t & \hat{\pi}_t & \hat{R}_t & \hat{e}_t & \hat{u}_t \end{bmatrix}'$ is the state vector; $\mathcal{E}_t = \begin{bmatrix} \varepsilon_t^e & \varepsilon_t^u & \varepsilon_t^v \end{bmatrix}'$ is the exogenous shock vector; and A_0, A_1, A_2, B and F are coefficient matrices. The aggregate economy in (33) nests two model specifications, namely, i) $\boldsymbol{n \in (0,1)}$: this is the novel specification of the paper, where a realistic value for n is provided through Bayesian inference in Section 4.3; and ii) $\boldsymbol{n = 1}$: in this case, agents exhibit no myopia at all.[27]

4.2 SAC Learning

Households and firms *learn* to use autoregressive forecasting rules to form expectations about future endogenous variables, that is, the output gap, inflation, and nominal interest rates, nested in $S_{1:3,t}$

$$S_{1:3,t} = \boldsymbol{\delta}_{t-1} + \boldsymbol{\gamma}_{t-1}(S_{1:3,t-1} - \boldsymbol{\delta}_{t-1}) + \boldsymbol{\epsilon}_t \tag{34}$$

where $\boldsymbol{\delta}_{t-1}$ is the mean of the $S_{1:3,1:t-1}$ series; $\boldsymbol{\gamma}_{t-1}$ represents the first-order correlation matrix between the $S_{1:3,0:t-2}$ and $S_{1:3,1:t-1}$ series; and $\boldsymbol{\epsilon}_t$ is a white noise process. The formulation in (34) nests commonly used forecasting rules, such as AR(1) and VAR(1) processes, for which I will estimate the model. The forecast of $S_{1:3,t+h}$ conditional on information about $S_{1:3,t-1}$, available at the beginning of period t is[28]

$$\widetilde{\mathbb{E}}_t^\star S_{1:3,t+h} = \boldsymbol{\delta}_{t-1} + (\boldsymbol{\gamma}_{t-1})^{h+1}(S_{1:3,t-1} - \boldsymbol{\delta}_{t-1}) \tag{35}$$

[27] Preston (2005) and Milani (2007) have used a model similar to (33) with $n = 1$ to investigate implications of adaptive learning in an infinite horizon learning setting. See Eusepi and Preston (2018) as well for an extensive review.

[28] An extensive analysis of the implications of learning for forecast error behavior can be provided by the author upon request.

Households and firms update their forecasting rules using sample autocorrelation coefficient (SAC) learning. This procedure was first introduced in economics by Hommes and Sorger (1998) and it relies on the Yule-Walker equations combined with sample estimates of autocorrelation coefficients; that is, $\boldsymbol{\delta}_t$ and $\boldsymbol{\gamma}_t$ are recursively updated according to

$$\begin{aligned} \boldsymbol{\delta}_t &= \boldsymbol{\delta}_{t-1} + \iota(S_{1:3,t} - \boldsymbol{\delta}_{t-1}) \\ \boldsymbol{\gamma}_t &= \boldsymbol{\gamma}_{t-1} + \iota\left((S_{1:3,t} - \boldsymbol{\delta}_{t-1})(S_{1:3,t-1} - \boldsymbol{\delta}_{t-1})' - \boldsymbol{\gamma}_{t-1}(S_{1:3,t} - \boldsymbol{\delta}_{t-1})(S_{1:3,t} - \boldsymbol{\delta}_{t-1})'\right)\boldsymbol{\eta}_t^{-1} \\ \boldsymbol{\eta}_t &= \boldsymbol{\eta}_{t-1} + \iota\left((S_{1:3,t} - \boldsymbol{\delta}_{t-1})(S_{1:3,t} - \boldsymbol{\delta}_{t-1})' - \boldsymbol{\eta}_{t-1}\right) \end{aligned} \tag{36}$$

where $\boldsymbol{\eta}_t$ is the second moment matrix and ι is the gain parameter that nests the two types of learning. With constant-gain learning, $\iota = \bar{\iota}$ is a constant parameter and it mimics a situation where a rolling window of data with length approximately equal to $\frac{1}{\bar{\iota}}$ is used to revise moments. With decreasing-gain learning, on the other hand, $\iota = \frac{1}{t+1}$ and all available historical data are used to update. The former approach is preferred because it has been universally found to improve empirical fit and the literature has shown that agents focus on recent observations when updating forecasting rules.[29]

4.3 Bayesian Estimation

Incorporating (34) into (33) yields the state-space representation of the model, described by

$$S_t = C_0(\Theta, \boldsymbol{\gamma}_{t-1})\boldsymbol{\Delta}_{t-1} + C_1(\Theta, \boldsymbol{\gamma}_{t-1})S_{t-1} + C_2(\Theta)\mathcal{E}_t \tag{37}$$

$$Y_t - \bar{Y} = PS_t \tag{38}$$

together with the dynamic system in (36), where $\boldsymbol{\Delta}_t = \begin{bmatrix} \boldsymbol{\delta}_t' & \mathbf{0}_{1x2} \end{bmatrix}'$, $Y_t = \begin{bmatrix} x_t^{obs} & \pi_t^{obs} & R_t^{obs} \end{bmatrix}'$ is the vector of observables, P is a matrix mapping model variables to the observables, and $\bar{Y}$ is a vector containing the observables' mean. I use quarterly data on real GDP, real potential GDP as reported by the US Congressional Budget Office, the GDP deflator, and the federal funds rate from 1968 to 2018, extracted from the Federal Reserve Economic Data (FRED). The output gap

[29] See for example, Del Negro and Eusepi (2011), Ormeño and Molnár (2015), Rychalovska (2016), Cole and Milani (2019), and Gaus and Gibbs (2018), among many others. Furthermore, Fuster, Laibson, and Mendel (2010) claim that "actual people's forecasts place too much weight on recent changes." Malmendier and Nagel (2016) find significant micro evidence in favor of constant-gain learning. See Tversky and Kahneman (1973, 1974) as well for theoretical considerations. Additionally, the evolution of beliefs under decreasing-gain learning depends on the length of the data, whereas constant-gain learning is robust to it.

is measured as the log difference between real GDP and potential GDP.[30] I refer the reader to Appendix D for more details on data preparation. The state-space form of the model in (36)-(38) is a Gaussian system; hence, I evaluate the likelihood function using the Kalman filter. The posterior distribution then is computed as

$$p(\Theta \mid Y_{1:T}) \propto p(Y_{1:T} \mid \Theta)p(\Theta) \tag{39}$$

where $p(Y_{1:T} \mid \Theta)$ is the data likelihood and $p(\Theta)$ the prior distribution of parameters. I use the Metropolis-Hastings algorithm to generate two blocks with 360,000 draws each and discard the first 60000 draws from the posterior distribution. In terms of the initiation of beliefs, I evaluate moments of the pre-sample data from 1960 to 1965 and use them as the initial learning parameters, $\boldsymbol{\delta}_0$ and $\boldsymbol{\gamma}_0$, for the Kalman filter procedure.[31]

As commonly done in the literature, I fix the discount factor $\beta = 0.99$. Furthermore, given that both the Phillips curve slope and the Calvo parameter, α, appear in the Phillips curve, and that the former depends on α in addition to other parameters, I cannot estimate the two simultaneously. Instead, I will fix the Phillips curve slope, κ, to 0.0015, which is the value estimated in Giannoni and Woodford (2004) for the flexible wages case. For most of the parameters, I set priors commonly used in the literature, as in, for instance, Milani (2007), Smets and Wouters (2007), and Herbst and Schorfheide (2016). Priors are given in Table 2.

The prior for n follows a beta distribution with mean 0.5 and standard deviation of 0.2.[32] In Appendix D.2, I present results for when the prior of n is instead assumed to follow a uniform distribution with mean 0.5 and standard deviation $1/\sqrt{12}$, and the results remain largely unchanged. The learning gain parameter for the mis-specified forecasting rules follows a gamma distribution prior with mean 0.035 and standard deviation 0.015.

The inverse intertemporal elasticity of substitution (IES) coefficient, σ, follows a gamma dis-

[30] Bayesian inference when the HP-filtered series of output are used as a measure of potential output produces similar results. Estimates can be provided by the author upon request.

[31] Forecasting rules that rely on past aggregate variables *only* have a slight advantage over rules that include shocks as well. Since beliefs are tied to moments from the data, the natural choice is to match initial beliefs to pre-sample data moments. This makes the estimation's vulnerability to initial beliefs - commonly faced in models with forecasting rules that depend on shocks - disappear.

To give an idea of the different approaches used to generate initial beliefs when the forecasting process depends on shocks, Milani (2007) estimates initial conditions on pre-sample data; Milani (2007) treats initial beliefs as parameters and estimates them along with the model's structural parameters; and Slobodyan and Wouters (2012a,b) initiate beliefs at the implied moments of the FIRE solution, apart from the other two aforementioned methods.

[32] Ilabaca, Meggiorini, and Milani (2020) use a prior with mean 0.8 and standard deviation 0.15.

Parameters		pdf	mean	standard deviation
Calvo parameter	α	$\mathcal{B}$	0.5	0.2
Degree of myopia	n	$\mathcal{B}$	0.5	0.2
Inverse IES coefficient	σ	$\mathcal{G}$	2	0.5
Habit in consumption	η	$\mathcal{B}$	0.5	0.2
Inflation indexation	ρ_π	$\mathcal{B}$	0.5	0.2
Elasticity mc	ω	$\mathcal{N}$	0.8975	0.4
Feedback to output gap	ϕ_x	$\mathcal{N}$	0.5	0.25
Feedback to inflation	ϕ_π	$\mathcal{N}$	1.5	0.25
Interest rate smoothing	ρ_r	$\mathcal{B}$	0.5	0.2
Demand shock autocorr.	ρ_e	$\mathcal{B}$	0.5	0.2
Supply shock autocorr.	ρ_u	$\mathcal{B}$	0.5	0.2
Demand shock std.	σ_e	$\mathcal{IG}$	0.1	2
Supply shock std.	σ_u	$\mathcal{IG}$	0.1	2
Monetary shock std.	σ_v	$\mathcal{IG}$	0.1	2
Gain parameter	$\bar{\iota}$	$\mathcal{G}$	0.035	0.015

Table 2: Priors

tribution with mean 2. Habit in consumption and inflation indexation parameters follow a beta distribution with mean 0.5 and standard deviation of 0.2. The Calvo parameter, α, is assumed to follow a beta prior with mean 0.5 and standard deviation 0.2. Elasticity of inflation with respect to marginal cost follows a normal prior with mean 0.8975 and standard deviation 0.4.

Policy reaction coefficients toward deviations of inflation and the output gap from their steady-state values are normally distributed with mean 1.5 and 0.5, respectively, and they share the same standard deviation of 0.25. The interest rate smoothing parameter follows a beta distribution with mean 0.5 and standard deviation 0.2. The autocorrelation coefficient of all shocks follows a beta distribution with mean 0.5 and standard deviation 0.2. The standard deviation of all shocks follows an inverse gamma distribution with mean 0.1 and standard deviation 2.

4.3.1 Posterior Distribution

Table 3 reports characteristics of the posterior distribution under FIRE, well-specified forecasts and myopia, constant-gain SAC learning with AR(1) forecasting rules combined with myopia and absent it.[33] Data fit is judged based on the evaluation of the log marginal data likelihood,

[33] Posterior distributions are generally well-behaved. I rely on the method proposed by Brooks and Gelman (1998) to analyze convergence statistics. Figures exhibiting the evolution of convergence statistics are reported in Appendix D.2.

computed through the modified harmonic mean method in Geweke (1999).[34]

In terms of data fit, I set the expectations formation process with well-specified forecasting rules and myopia be the benchmark one, and compare the other alternatives to that benchmark. The values in parentheses in the last row of Table 3 report the Bayes factor value of the model specification relative to the benchmark specification. The log of the Bayes factor for the model with mis-specified forecasts and myopia is higher than 3, and according to Kass and Raftery (1995), a factor magnitude whose natural log is higher than 3 denotes strong evidence in favor of the model with superior fit. On the other hand, the log of the Bayes factor for the model with FIRE as well as the one with AR(1) forecasting rules absent myopia is negative. Thus, both model specifications fit the data worse than the benchmark and, as a result, worse than mis-specified forecasts combined with myopia. Therefore, the data fit analysis showed that i) the model where households and firms combine learning of AR(1) mis-specified forecasting rules with myopia (column 4 in Table 3) fits the data best; and that ii) generally, the presence of myopia ensures a better fit of the US macroeconomic data for both specifications with well- and mis-specified forecasting rules.

While myopia is neither a necessary nor a sufficient condition to match the empirical evidence on forecasting data, the estimated marginal data densities exhibited in Table 3 show that a *combination* of myopia with mis-specified forecasting rules improves the fit of the macroeconomic data, well beyond the data fit that is granted by the specifications without myopia. Consequently, if one is interested in an expectations formation process that is consistent with consensus forecasting data evidence *only*, then a mis-specified forecasting rule alone is sufficient. However, if one is interested in an expectations formation process that is *both* consistent with the consensus forecasting data evidence *and* fits the macroeconomic data well, then a *combination* of mis-specified forecasting rules and myopia should be preferred over the two departures separately.

The posterior mean estimate of the parameter capturing the degree of myopia, n, is significantly different from 1, showing evidence in favor of considerable cognitive discounting of expected future fluctuations in the US economy. The posterior mean of n is around 0.4 for the model with well-

[34]Bayes' theorem implies that $p(Y_t) = \int p(Y_t \mid \Theta)p(\Theta)d\Theta$, which is impossible to compute analytically. The modified harmonic mean (MHM) method of Geweke (1999) is evaluated using the posterior distribution draws,

$$p(Y_t) \approx \left[\frac{1}{M - M_0} \sum_{m=M_0+1}^{M} \frac{f(\Theta^m)}{p(Y_t|\Theta^m)p(\Theta^m)} \right]^{-1}$$

where M is the total number of draws, M_0 is the number of discarded draws, and $f(.)$ is the density of a truncated normal distribution.

specified forecasting rules, and 0.7 for the model with AR(1) mis-specified forecasts. The posterior mean for the learning gain parameter $\bar{\iota}$ is estimated to be 0.07 in both the presence and absence of myopia, which further implies that a rolling window of about 13 quarters is used to recursively update the parameters of the forecasting process.

The model estimated under the assumption of well-specified forecasts, with or without myopia, requires significantly higher degrees of habit in consumption as compared to the model with mis-specified forecasts and myopia. In particular, the posterior mean of η is 0.9 and 0.8 in columns (1) and (2), respectively. By contrast, in the case of mis-specified forecasts the posterior mean of η varies between 0.4 and 0.5. On the other hand, the models with well-specified forecasts require high degrees of inflation indexation of about 0.9 at the posterior mean, whereas the model specifications with mis-specified forecasts necessitate a somewhat lower degree of inflation indexation of about 0.8.

The posterior mean of the inverse elasticity of intertemporal substitution, σ, is estimated to be around 2 for all specifications. The Calvo parameter cannot be identified in the model specifications with well-specified forecasting rules. In the presence of mis-specified forecasts, the posterior mean of α is estimated to be around 0.4 in cases of both myopia and no myopia. The implied expected price duration for both specifications is about 5 months on average, which is in accordance with findings in Bils and Klenow (2004) that, for most goods, prices change on average once every 4 months. The posterior mean of ω varies in the range between 0.8 and 1 depending on the model specification.

Policy parameters are generally robust across specifications, where the posterior mean estimates of the reaction to the output gap is around 0.4, the reaction to inflation varies between 1.4 and 1.5, and the interest rate smoothing parameter is estimated to be close to 0.9.

The posterior mean of the demand shock autocorrelation in the model with mis-specified forecasts and myopia is about 0.7 and it is slightly lower than its estimate of about 0.8 in the model with mis-specified forecasts and no myopia. Furthermore, the demand shock is less persistent under FIRE ($\rho_e \approx 0.6$ at the posterior mean) relative to the specification in column (4). On the other hand, demand shock innovations are at least two times less volatile under mis-specified forecasts than under well-specified forecasts: the posterior mean estimates for demand innovations vary between 1.3 and 2.3 for the case of mis-specified forecasts, but they increase to values above 5 for well-specified forecasts. The persistence and standard deviation of the supply shock are generally robust across specifications, with ρ_u varying between 0.05 and 0.1, and σ_u taking values between 0.2 and 0.3. Likewise, the standard deviation of the monetary shock is generally robust across models at an estimated posterior mode of 0.2.

	Well-specified Forecasting Rules						**AR(1) Mis-specified Forecasting Rules**					
	(1)			(2)			(3)			(4)		
	no myopia, $n = 1$			**myopia, $n \in (0, 1)$**			**no myopia, $n = 1$**			**myopia, $n \in (0, 1)$**		
Parameters	mean	5%	95%	mean	5%	95%	mean	5%	95%	mean	5%	95%
Calvo parameter, α	-	-	-	-	-	-	0.37	0.12	0.66	0.43	0.15	0.73
Degree of myopia, n	-	-	-	0.39	0.17	0.61	-	-	-	0.65	0.40	0.84
Inverse IES coefficient, σ	2.41	1.44	3.29	1.99	1.16	2.78	2.18	1.45	3.05	2.23	1.41	3.19
Habit in consumption, η	0.90	0.83	0.98	0.78	0.56	0.92	0.39	0.25	0.56	0.53	0.31	0.72
Inflation indexation, ρ_π	0.90	0.84	0.95	0.87	0.81	0.92	0.77	0.53	0.92	0.80	0.65	0.92
Elasticity mc, ω	0.83	0.20	1.43	0.94	0.31	1.57	0.98	0.34	1.62	0.99	0.36	1.63
Feedback to output gap, ϕ_x	0.36	0.22	0.49	0.42	0.22	0.60	0.41	0.25	0.63	0.43	0.25	0.66
Feedback to inflation, ϕ_π	1.43	1.16	1.70	1.46	1.13	1.77	1.45	1.15	1.78	1.45	1.14	1.77
Interest rate smoothing, ρ_r	0.88	0.85	0.91	0.92	0.89	0.95	0.92	0.88	0.94	0.92	0.89	0.95
Demand shock autocorr., ρ_e	0.64	0.50	0.77	0.53	0.30	0.82	0.83	0.72	0.92	0.72	0.55	0.88
Supply shock autocorr., ρ_u	0.06	0.01	0.10	0.05	0.01	0.09	0.11	0.03	0.27	0.09	0.02	0.18
Demand shock std., σ_e	5.25	1.64	9.57	5.19	1.75	8.23	1.25	0.54	1.86	2.34	1.01	3.79
Supply shock std., σ_u	0.26	0.23	0.28	0.16	0.10	0.22	0.33	0.28	0.39	0.33	0.28	0.38
Monetary shock std., σ_v	0.21	0.19	0.23	0.21	0.19	0.22	0.21	0.19	0.22	0.21	0.19	0.22
Learning gain, $\bar{\iota}$	-	-	-	-	-	-	0.07	0.04	0.10	0.07	0.04	0.11
Log marg. data dens.												
Modified Harmonic Mean	-272.593			-258.487			-276.482			-255.186*		
Bayes factor	($e^{-17.41}$)			($e^{-3.30}$)			($e^{-21.30}$)			(e^{0})		

Table 3: Posterior distribution of the model for various assumptions on the expectations formation process. Values in parentheses denote the Bayes factor of the model relative to the benchmark specification with well-specified forecasts and myopia. The asterisk denotes strong evidence in favor of the model relative to the benchmark one.

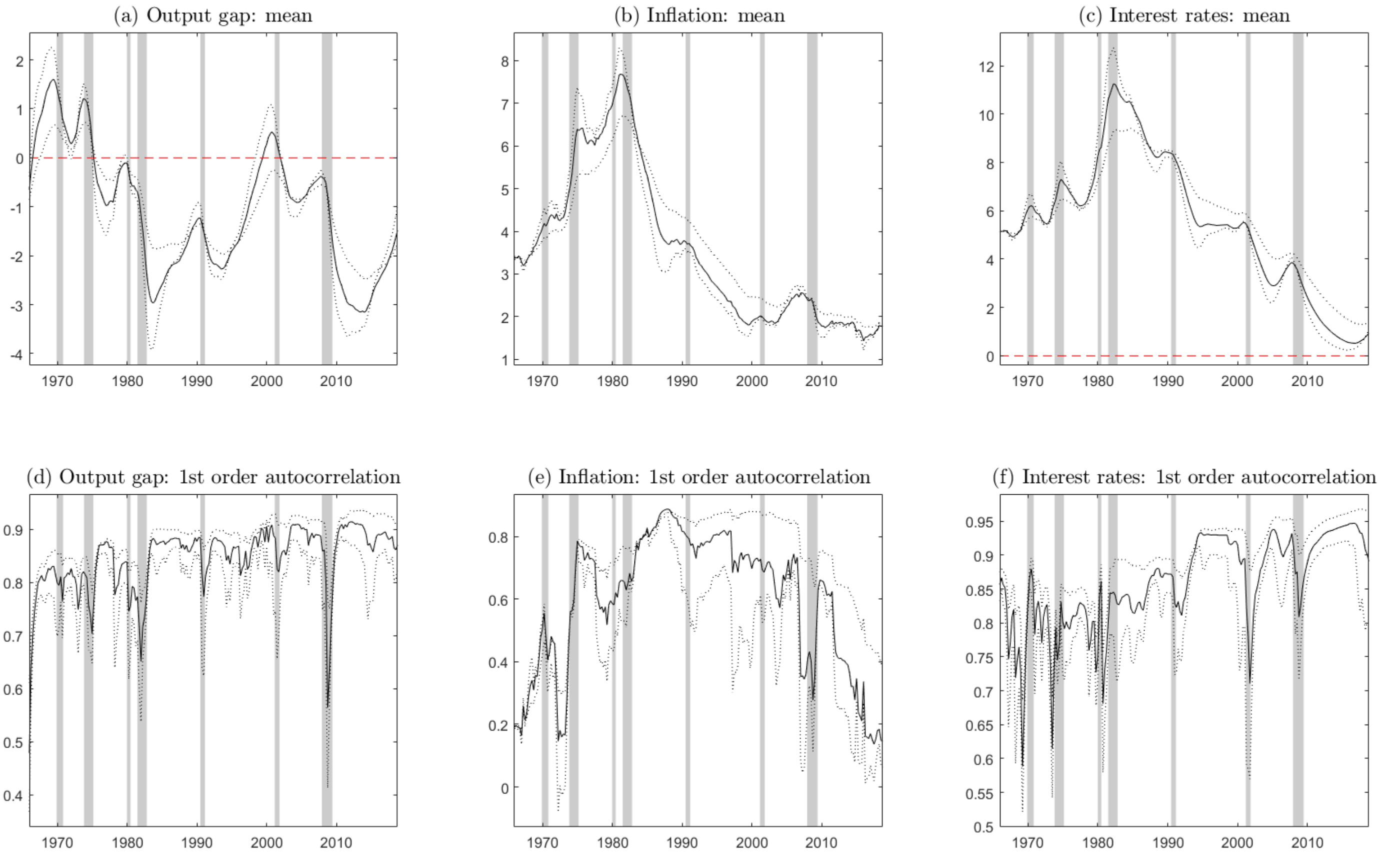


Figure 5: Evolution of the AR(1) forecast coefficients in the model with myopia. The black and dotted curves plot perceived moments for structural parameters set at their estimated posterior mean and 90 percent highest posterior density, respectively. Gray areas indicate recessionary periods as reported by the National Bureau of Economic Research. The dashed red lines indicate the x axis.

Figure 5 plots the historical evolution of the mis-specified forecast coefficients in the model with mis-specified forecasts and myopia, with parameters set at their posterior mean and 90 percent highest posterior density values. As shown in Figure 5, recessionary periods, indicated by the shaded gray areas, have been historically associated with a decrease in the perceived mean and first-order autocorrelation of the output gap. On the other hand, there is a shift in the way agents perceive moments of inflation and nominal interest rates during recessions. Particularly, before the early 1980s, recessions were associated with increasing beliefs about the mean and first-order autocorrelation of inflation and nominal rates. On the contrary, during and after the Great Moderation, economic turmoils are characterized by a decrease in beliefs about the mean and first-order autocorrelation of inflation and nominal interest rates. Therefore, the well-documented contrast between the US macroeconomy during the 1970s and the Great Moderation period is similarly mirrored in agents' perceptions about moments of inflation and nominal interest rates.[35] Another interesting observation from Figure 5 is that the implied beliefs about the annualized mean of inflation from the aftermath of the Great Recession up until the end of 2018 have been particularly steady at 2 percent.

4.3.2 VAR(1) Forecasting Rules

To investigate the model's performance when agents learn to use more sophisticated, yet misspecified, forecasting rules, I re-estimate the model with VAR(1) forecasting rules with and without myopia under constant-gain learning. The characteristics of the posterior distribution of parameters are exhibited in the third and fourth columns of Table 4 jointly with the AR(1) forecasting rule counterpart for easier comparison.

To judge model fit, I set the expectations formation process with AR(1) forecasts and myopia to be the benchmark specification, and compare the other alternatives to it. The values in parenthesis in the last row of Table 4 report the value of the Bayes factor for the model specification relative to the benchmark. The natural log of the Bayes factor for the model with VAR(1) forecasts and myopia is 2.3 units higher than the benchmark. According to Kass and Raftery (1995), this would be positive, but not strong, evidence in favor of VAR(1) forecasts and myopia relative to a combination of AR(1) forecasts and myopia. On the other hand, both specifications with VAR(1) or AR(1) forecasts absent myopia perform significantly worse than the analogue cases with myopia.

Parameter estimates under a combination of VAR(1) forecasts with myopia are very much similar to the ones under an AR(1) forecasting rule with myopia. On the other hand, the specification of VAR(1) forecasts without myopia exhibits some notable differences when compared with the other specifications. First, the Calvo parameter is estimated to be more than twice as high, and the inverse elasticity of substitution is estimated to be twice as high. Similarly, the estimated posterior mean of the supply shock is about 7 times higher than that of the other specifications, whereas the estimated standard deviation of supply innovations is about 4 times smaller. These results imply that, generally, the VAR(1) forecast specification with no myopia requires frictions of larger magnitudes to fit the macroeconomic data.

Finally, Figure 6 plots the evolution of the estimated agents' beliefs when they engage in

[35] See, for instance, Bianchi (2013) and references therein, for a discussion on the differences between the two periods.

constant-gain learning of a VAR(1) forecasting process in the specification with myopia. The perceived first-order correlation between any two distinct aggregate variables is estimated to fluctuate around 0, whereas the perceived first-order autocorrelation fluctuates around a strictly positive value. Therefore, using more elaborate forecasting rules, such as VAR(1), will not add, on average, any significant information to households and firms in terms of forecasting, and it will not strongly enhance the model's fit of the data.

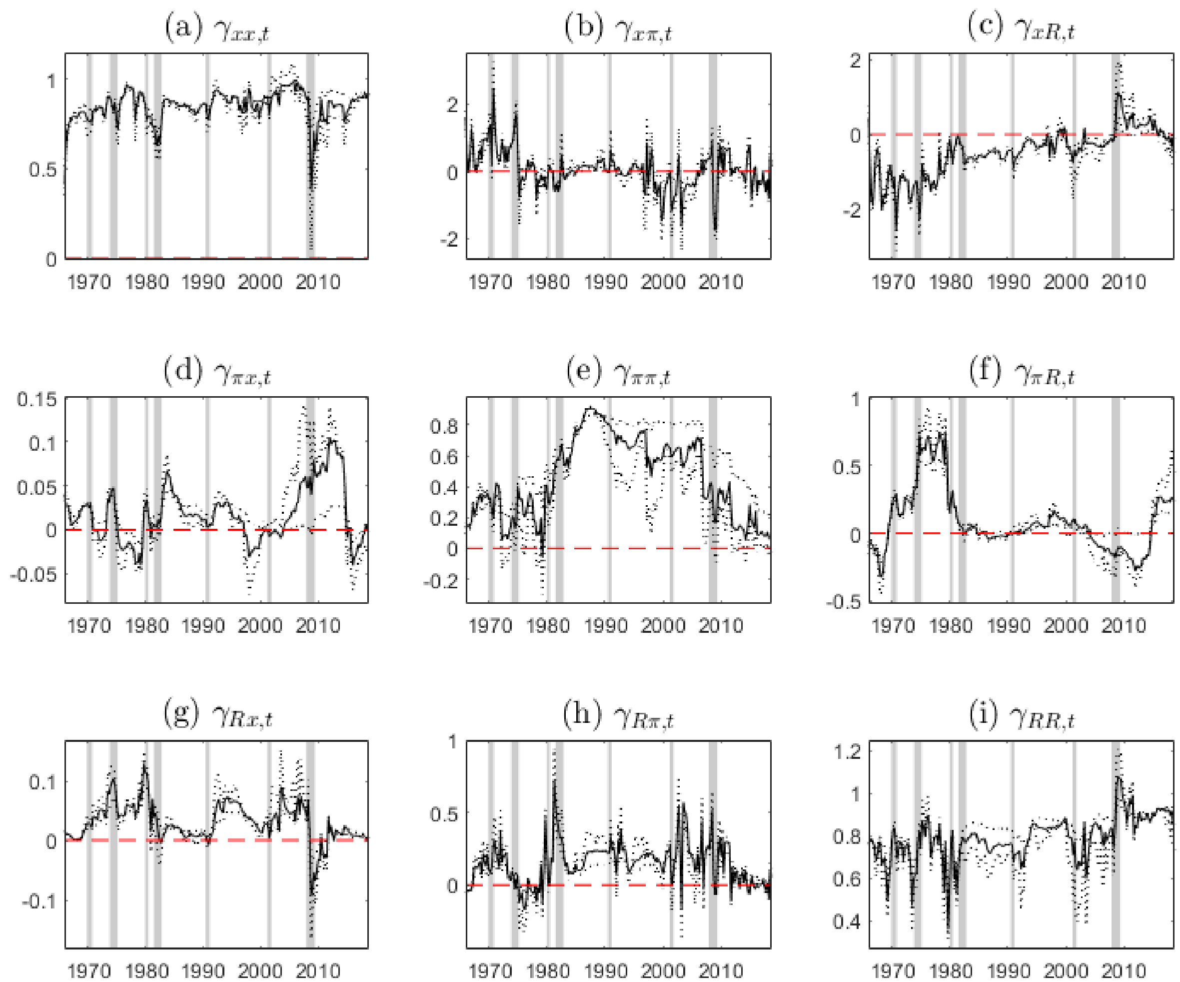


Figure 6: Evolution of the VAR(1) forecast coefficients in the model with myopia. The black and dotted curves plot implied beliefs for structural parameters set at their estimated posterior mean and 90 percent highest posterior density, respectively. Gray areas indicate recessionary periods as reported by the National Bureau of Economic Research. The dashed red lines indicate the x axis.

	AR(1) Mis-specified Forecasting Rules						VAR(1) Mis-specified Forecasting Rules					
	(1)			(2)			(3)			(4)		
	no myopia, $n = 1$			**myopia, $n \in (0, 1)$**			**no myopia, $n = 1$**			**myopia, $n \in (0, 1)$**		
Parameters	mean	5%	95%	mean	5%	95%	mean	5%	95%	mean	5%	95%
Calvo parameter, α	0.37	0.12	0.66	0.43	0.15	0.73	0.97	0.93	0.99	0.37	0.12	0.66
Degree of myopia, n	-	-	-	0.65	0.40	0.84	-	-	-	0.70	0.46	0.88
Inverse IES coefficient, σ	2.18	1.45	3.05	2.23	1.41	3.19	4.08	3.03	5.21	2.18	1.45	3.05
Habit in consumption, η	0.39	0.25	0.56	0.53	0.31	0.72	0.35	0.24	0.54	0.39	0.25	0.56
Inflation indexation, ρ_π	0.77	0.53	0.92	0.80	0.65	0.92	0.08	0.02	0.16	0.77	0.53	0.92
Elasticity mc, ω	0.98	0.34	1.62	0.99	0.36	1.63	0.94	0.32	1.57	0.98	0.34	1.62
Feedback to output gap, ϕ_x	0.41	0.25	0.63	0.43	0.25	0.66	0.40	0.24	0.63	0.41	0.25	0.63
Feedback to inflation, ϕ_π	1.45	1.15	1.78	1.45	1.14	1.77	1.45	1.14	1.78	1.45	1.15	1.78
Interest rate smoothing, ρ_r	0.92	0.88	0.94	0.92	0.89	0.95	0.92	0.88	0.95	0.92	0.88	0.94
Demand shock autocorr., ρ_e	0.83	0.72	0.92	0.72	0.55	0.88	0.97	0.94	0.99	0.83	0.72	0.92
Supply shock autocorr., ρ_u	0.11	0.03	0.27	0.09	0.02	0.18	0.74	0.62	0.84	0.11	0.03	0.27
Demand shock std., σ_e	1.25	0.54	1.86	2.34	1.01	3.79	0.22	0.09	0.52	1.25	0.54	1.86
Supply shock std., σ_u	0.33	0.28	0.39	0.33	0.28	0.38	0.08	0.05	0.11	0.33	0.28	0.39
Monetary shock std., σ_v	0.21	0.19	0.22	0.21	0.19	0.22	0.21	0.19	0.23	0.21	0.19	0.22
Learning gain, $\bar{\iota}$	0.07	0.04	0.10	0.07	0.04	0.11	0.03	0.01	0.05	0.07	0.04	0.10
Log marg. data dens.												
Modified Harmonic Mean	-276.482			-255.186			-277.3623			-252.885		
Bayes factor	($e^{-21.30}$)			(e^{0})			($e^{-22.18}$)			($e^{2.30}$)		

Table 4: Posterior distribution of the model for AR(1) and VAR(1) mis-specified forecasting rules. Values in parentheses denote the Bayes factor of the model relative to the benchmark specification with AR(1) mis-specified forecasts and myopia.

4.3.3 Impulse Response Functions

Computing IRFs under SAC learning is slightly complicated because the response of aggregates to any shock depends on the perceived first-order autocorrelation prior to the economy being shocked. Furthermore, as the shock hits the economy the perceived mean and first-order correlation, which enter the model's solution non-linearly, evolve jointly with the aggregates. To make the IRFs comparable across time periods and assumptions on expectations, I assume that the economy prior to the shock is at its steady state.[36]

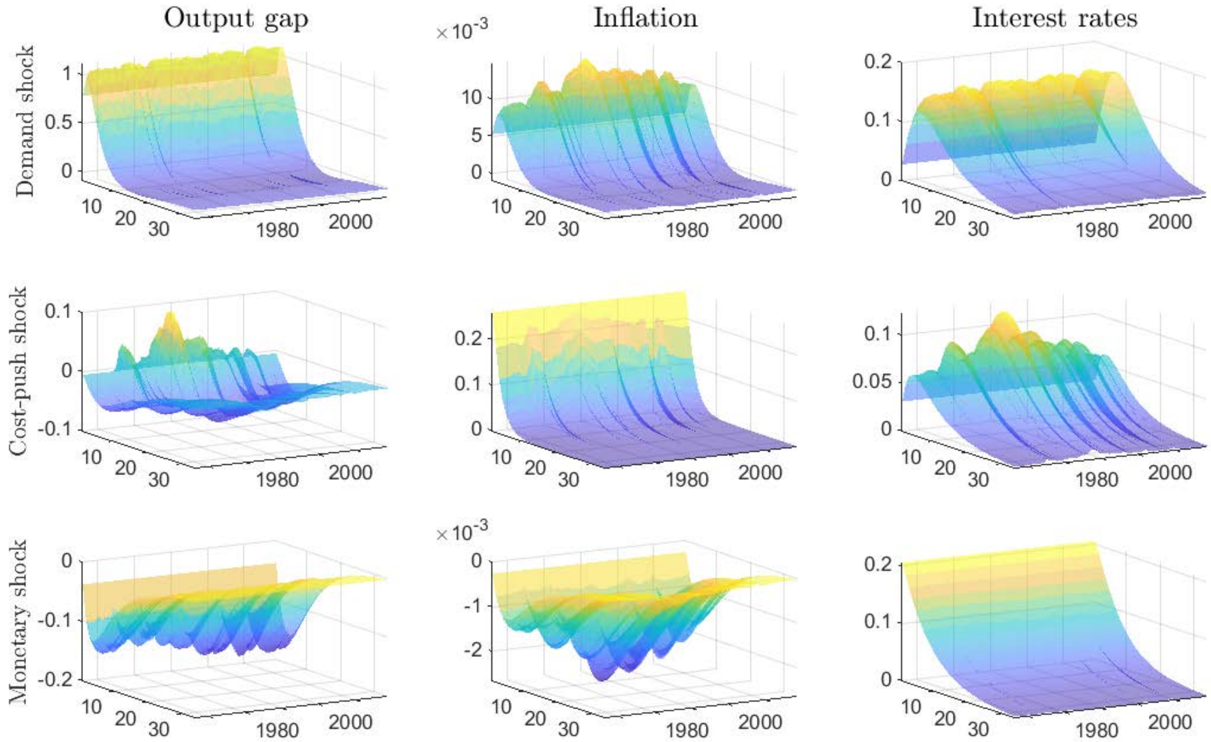
Figure 7 plots the three-dimensional IRFs of the output gap, inflation, and nominal interest rates to a one standard deviation demand, cost-push, and monetary shock, when the expectations formation process is characterized by mis-specified - AR(1) in panel (a) and VAR(1) in panel (b) - forecasts and myopia. The model, including the standard deviation of the various shocks, is calibrated to the estimated posterior mean as shown in Tables 3 and 4. Both panels show that the response of aggregates to various fundamental shocks depends on the perceived first-order correlation coefficients prior to the shock. It is interesting to note that the response of inflation to monetary shocks is always negative under AR(1) forecasts, whereas under VAR(1) forecasts, the sign of the response depends on the time period. To better observe such differences, Figure 8 projects the three-dimensional IRFs of Figure 7 on the [response - time period] plane. Overall, relative to AR(1) forecasts, VAR(1) forecasts introduce more volatility in the sign of the response of inflation to demand and monetary shocks.

To compare the IRFs across different assumptions on expectations, Figure 9 plots the average impulse responses for the output gap, inflation, and nominal interest rates. The standard deviation of all three shocks is normalized to 1, while the rest of the parameters are set at their estimated posterior mean as reported in Tables 3 and 4. As shown in panel (a), mis-specified forecasts generally induce more sufficient internal persistence and amplification to exogenous shocks as compared to well-specified forecasts. Moreover, when compared to AR(1) forecasts, VAR(1) forecasting rules often times generate a more persistent and intense response of aggregates, particularly to demand and monetary shocks. As exhibited in panel (b), this outcome is more pronounced when myopia is absent (note that the estimated standard deviation of the demand shock in the case of VAR(1) forecasts absent myopia is much smaller than 1). Turning to myopia, its presence tends

[36] The prior for the perceived mean of the aggregates is set to 0, whereas the priors for the perceived first-order correlation coefficients are set to their estimated values as plotted in Figures 5 and 6.

to mute fluctuations around the steady state, and therefore, the response of aggregates to various shocks is amplified when there is no myopia as compared to cases when mis-specified forecasts are myopically adjusted.

(a) *AR(1) mis-specified forecasts and myopia*



(b) *VAR(1) mis-specified forecasts and myopia*

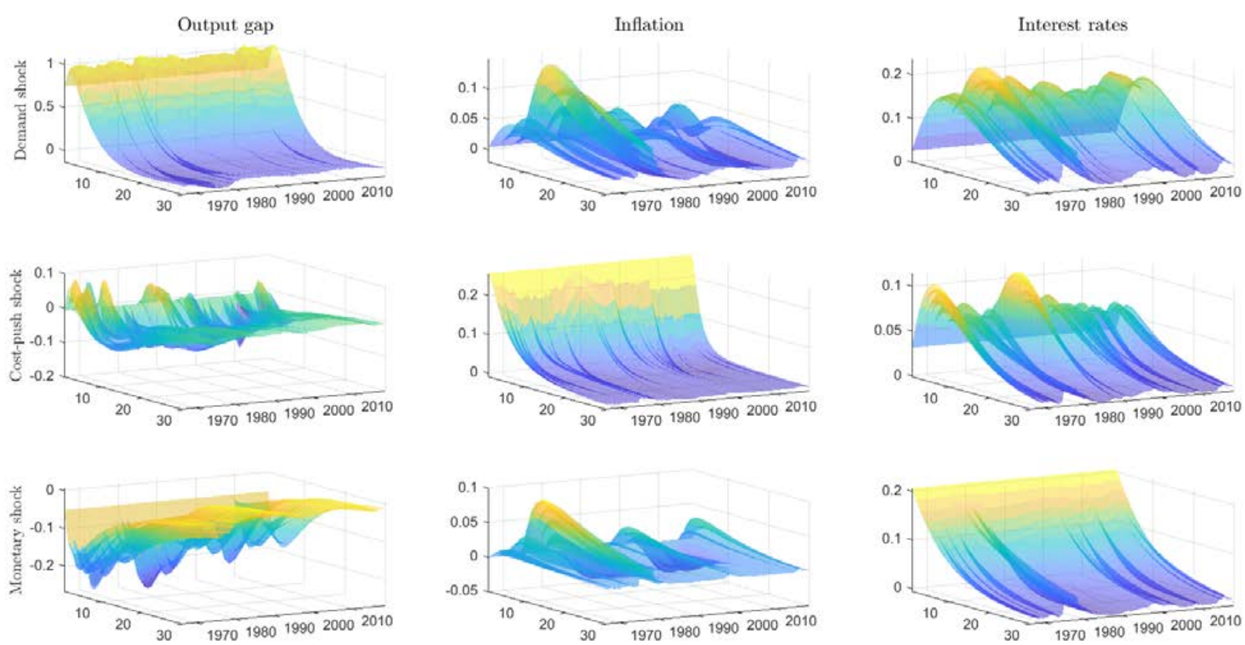
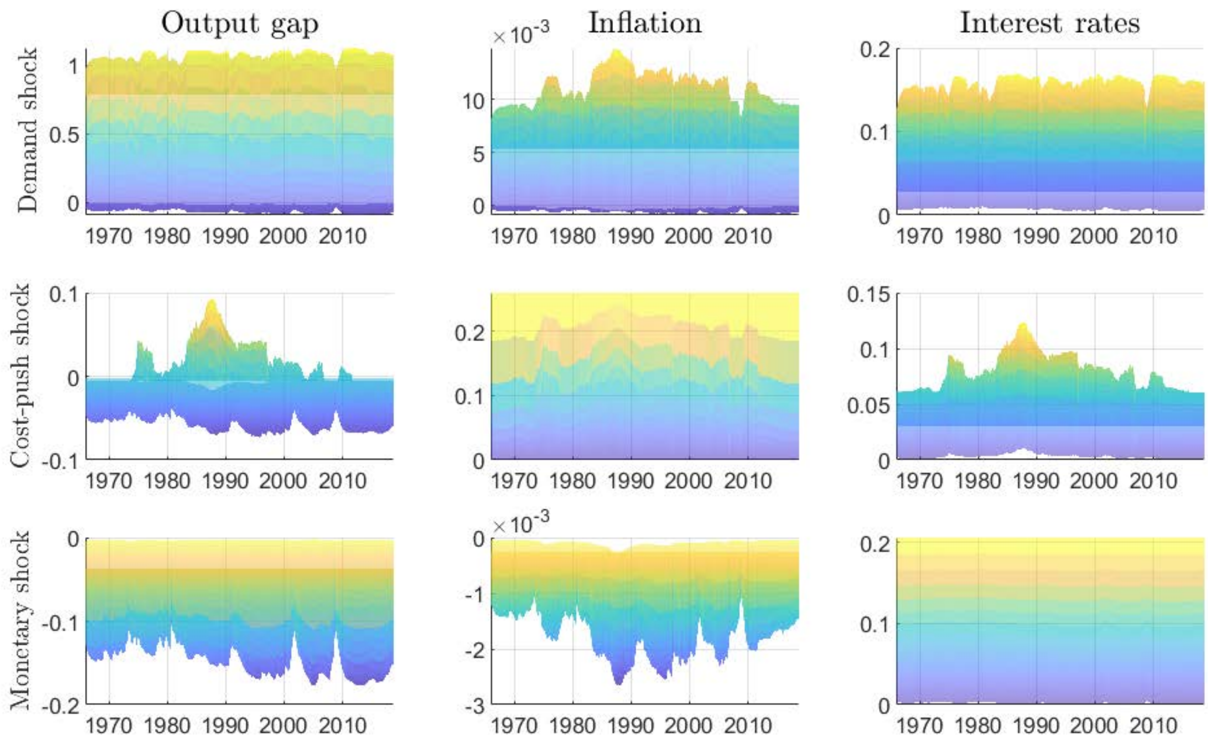


Figure 7: Three-dimensional impulse response functions to a one standard deviation positive demand, cost-push, and monetary shock for the model with mis-specified forecasts and myopia. Panel (a): AR(1) forecasts; panel (b): VAR(1) forecasts. Parameters are set at their estimated posterior mean as shown in column (4) of Tables 3 and 4. X axis: periods of response; y axis: estimation time periods; z axis: impulse response function.

(a) *AR(1) mis-specified forecasts and myopia*



(b) *VAR(1) mis-specified forecasts and myopia*

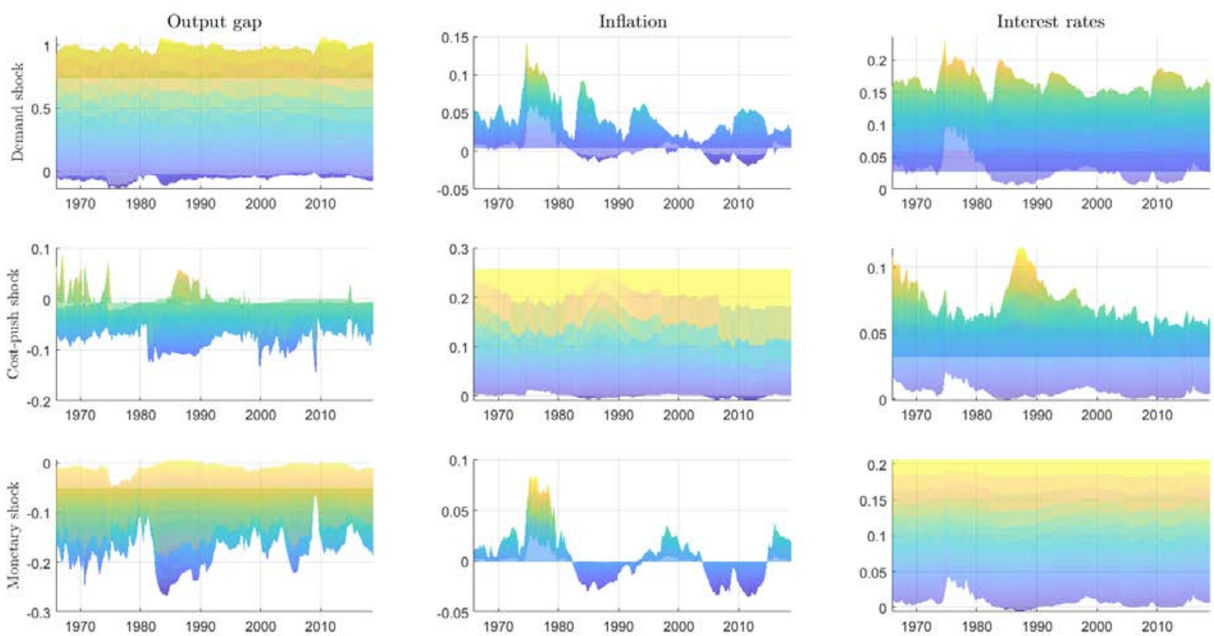
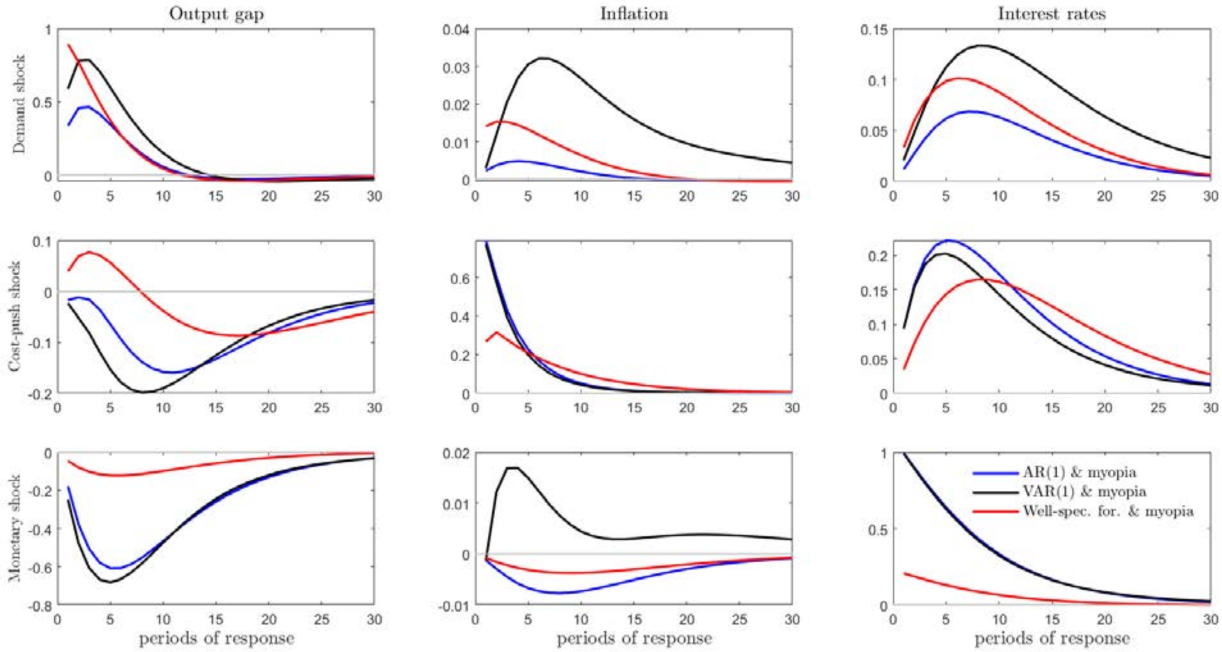


Figure 8: Impulse response functions of Figure 7 projected on the (response periods,time) plane. Panel (a): AR(1) forecasts; panel (b): VAR(1) forecasts.

(a) *Myopia: Mis-specified versus well-specified forecasts*



(b) *Mis-specified forecasts: myopia versus no myopia*

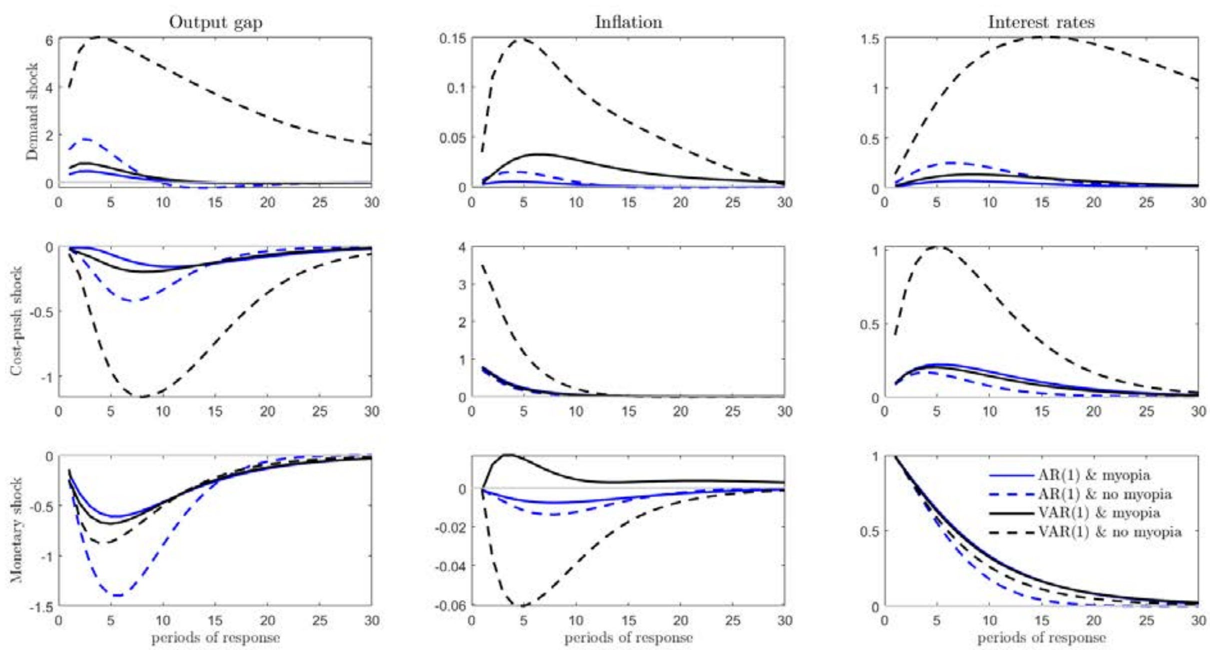


Figure 9: Average impulse response functions to a one standard deviation positive demand, cost-push and monetary shock. In blue: AR(1) mis-specified forecasts; in black: AR(1) mis-specified forecasts; in red: well-specified forecasts. Panel (a): comparison between mis-specified and well-specified forecasts when combined with myopia; panel (b): comparison between mis-specified forecasts with myopia and without myopia. The standard deviation of all shocks is normalized at 1, and all the other parameters are set at their estimated posterior mean as shown in Tables 3 and 4.

5 Implications for Forecasting Errors

In this final section, I revisit the three empirical facts about inflation consensus forecasting errors analyzed in Section 3 when autoregressive mis-specified forecasts of the AR(1) and VAR(1) type are combined with myopia, that is, for both expectations formation processes that best fit the data as shown in the previous section.[37]

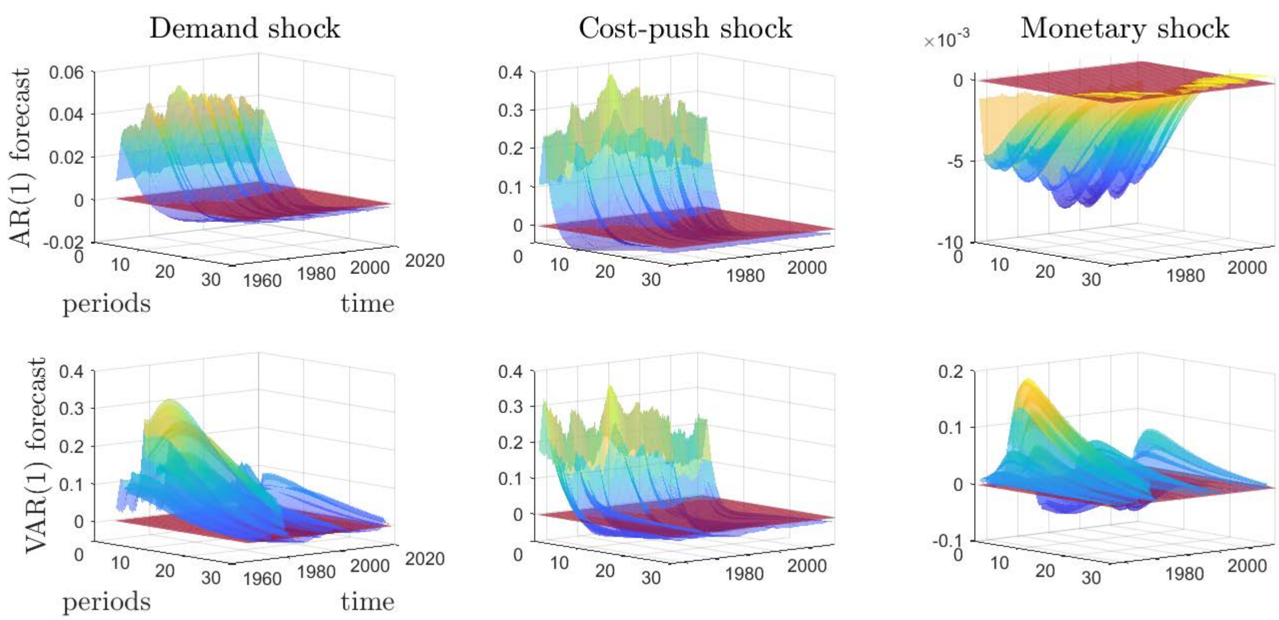


Figure 10: Three-dimensional impulse response functions of annual inflation forecasting errors to a one standard deviation positive demand, cost-push, and monetary shock in the model with mis-specified forecasts and myopia. Top three panels: AR(1) mis-specified forecasts; bottom three pabels: VAR(1) mis-specified forecasts. Model parameters are set equal to the posterior mean as shown in Tables 3 and 4. In dark red: (response, time period) plane.

Delayed over-shooting. Setting the model parameters to the posterior mean as found in column (4) of Tables 3 and 4, I simulate the response of annual inflation forecast errors to a demand, cost-push, and monetary shock over the estimation period of 1966:Q1 - 2018:Q3. Figure 10 plots the three-dimensional IRFs with the top three panels exhibiting the responses for AR(1) forecasting rules, and the bottom three panels showing the responses for VAR(1) forecasting rules. Focusing first on the top three panels, it is clear that, with AR(1) forecast rules, the response of inflation forecast errors to all three shocks switches sign, exhibiting delayed over-shooting. On the other hand, with VAR(1) forecast rules, the response of inflation forecast errors switches sign in the

[37]See Appendix D.3 for an analysis of the behavior of simulated inflation forecast error data when the expectations formation process is characterized by mis-specified forecasts and no myopia.

majority of time periods *only* in response to a cost-push shock. Therefore, VAR(1) mis-specified forecasts combined with myopia do not always guarantee delayed over-shooting.

Figure 11 projects the impulse response functions of Figure 10 on the (response, time period) plane. With AR(1) forecasts, delayed over-shooting occurs irrespective of the shock forecast errors are subjected to and regardless of the initial beliefs about the mean and first-order autocorrelation of inflation. With VAR(1) forecasts, the response of inflation forecast errors to a cost-push shock almost always exhibits delayed over-shooting. On the other hand, the response to a demand shock over-shoots late only during the mid-1980s and 2010s, so it does depend on the initial beliefs about moments of inflation. In contrast, the response of inflation forecast errors to a monetary shock is either always positive or always negative, showing no signs of delayed over-shooting.

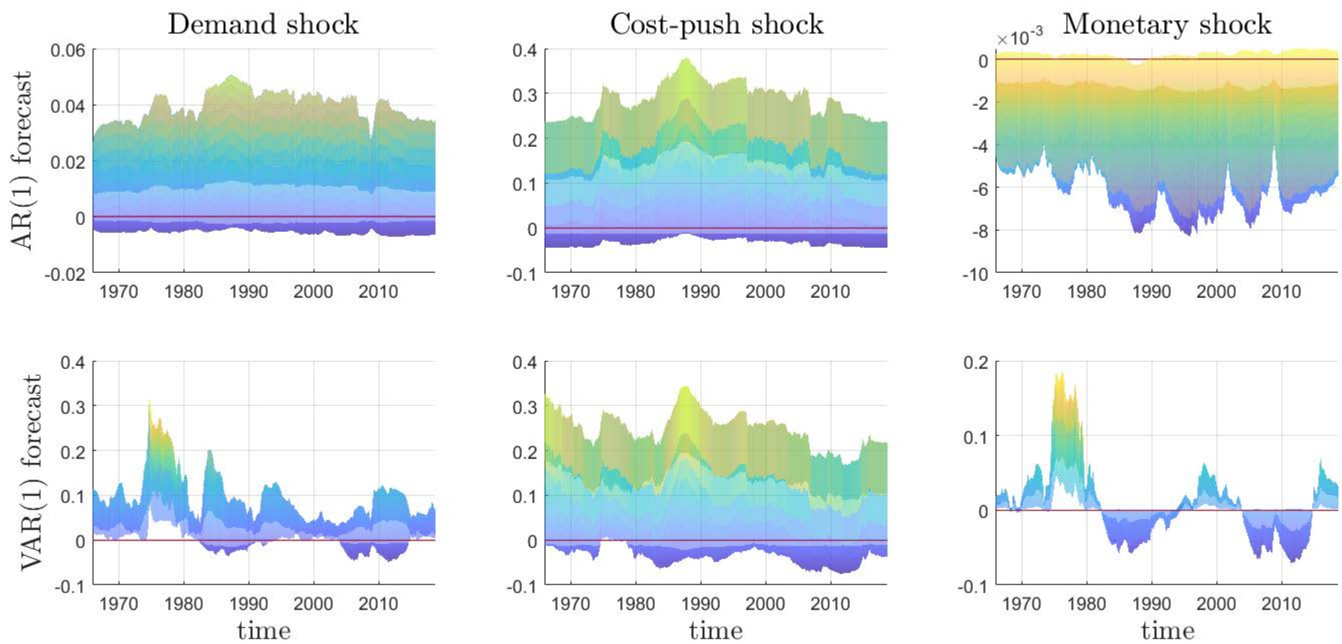


Figure 11: Impulse response functions of annual inflation forecasting errors to a one standard deviation positive demand, cost-push, and monetary shock in the model with mis-specified forecasts and myopia, projected on the (response, time period) plane. Top three panels: AR(1) mis-specified forecasts; bottom three panels: VAR(1) mis-specified forecasts. Model parameters are set equal to the posterior mean as shown in column (4) of Tables 3 and 4. In red: x axis.

Under-reaction to ex-ante forecast revisions, over-reaction to current inflation. I estimate the two regressions in (15) and (16) with simulated annual forecasting data for inflation, when the expectations formation process is a combination of myopia and autoregressive mis-specified forecasts of the AR(1) or VAR(1) type. Table 5 presents the results: panel A shows the estimates of K_4 and M_4 over the full sample from 1967:Q1 through 2018:Q3, whereas panel B exhibits estimates of M_4 after the early 1980s, that is, starting from 1981:Q3 through 2018:Q3.

Panel A: 1967:Q1 - 2018:Q3

	AR(1) and myopia			*VAR(1) and myopia*		
	5%	mean	95%	5%	mean	95%
Revision	8.75	4.01	1.95	6.98	3.38	1.60
	(0.54)	(0.33)	(0.28)	(0.45)	(0.31)	(0.28)
Current	0.75	0.57	0.32	0.72	0.53	0.29
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)

Panel B: 1981:Q3 - 2018:Q3

	AR(1) and myopia			*VAR(1) and myopia*		
	5%	mean	95%	5%	mean	95%
Current	0.41	0.19	-0.11	0.36	0.09	-0.23
	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)	(0.05)

Table 5: Estimates of regressions in (15) and (16) on simulated forecasting data. Model parameters are set at the posterior 5^{th} percentile, the mean, and the 95^{th} percentile of the distribution for the model with mis-specified forecasts and myopia as documented in column (4) of Tables 3 and 4. All regressions include a constant term, and standard errors are given in parenthesis. Panel A: estimates of K_4 (revision) and M_4 (current) over the full sample for AR(1) and VAR(1) mis-specified forecasts. Panel B: estimates of M_4 (current) since the early 1980s for AR(1) and VAR(1) mis-specified forecasts.

Consistent with the evidence presented in Section 3, annual inflation forecast errors depend positively on ex-ante inflation forecast revisions, regardless of the model parameterization and of the structure of the mis-specified forecasting rule. Similarly, current realizations are positively correlated with ex-post forecast errors over the full sample. On the other hand, when the sample starts from the early 1980s, the sign of the correlation between ex-post forecast errors and current inflation realizations depends on the parameterization of the model. In particular, simulated forecast data over-react to current inflation for parameterizations that are closer to the 95^{th} percentile of the posterior distribution. This outcome applies to both AR(1) and VAR(1) mis-specified forecasts, and it is consistent with the analysis in Section 3 showing that low degrees of myopia are needed for over-reaction to occur.

6 Concluding Remarks

The present paper combines two of the most prominent deviations from FIRE, namely, mis-specified forecasts and myopia in a unified New Keynesian framework that is amenable to macroeconomic data. The first part of the paper focuses on a partial equilibrium pricing problem, derives a number of implications, and tests them with evidence from consensus inflation forecasting data. The second part of the paper embeds the same departures from FIRE in a full New Keynesian model with habit in consumption and inflation indexation, derives the general equilibrium solution under sample autocorrelation coefficient learning, and estimates the model using Bayesian methods.

The paper underscores a number of novel results. First, it shows that a combination of autoregressive mis-specified forecasts and myopia is consistent with consensus inflation forecasts' i) delayed over-shooting; ii) under-reaction to ex-ante forecast revisions; and iii) over-reaction to recent events. The paper further proves that while mis-specified forecasts are both sufficient and necessary to match all three facts, myopia alone is neither. Second, the general equilibrium empirical analysis reveals that the best fitting expectations formation process for both households and firms is characterized by a combination of autoregressive mis-specified forecasts and myopia. Third, no strong evidence is found in favor of more elaborate VAR(1) forecast rules over simple AR(1) forecasts. Fourth, autoregressive mis-specified forecasts in the presence of myopia generate substantial internal persistence and amplification to exogenous shocks. Finally, the paper comes full circle and shows that simulated inflation forecast data from the estimated general equilibrium mirror the three empirical facts on inflation forecasting data.

The current paper lays solid ground in service to future research. A salient feature of forecasting data that has not been accounted for in the present paper is heterogeneity. Therefore, one potential extension of the current work would be to allow for heterogeneity in the degree of myopia, as well as in the structure of forecasting rules.

Appendix (For Online Publication)

A Partial Equilibrium New Keynesian Pricing Problem

Firms face nominal rigidities a la Calvo: they cannot reset the price with probability $\alpha \in (0, 1)$ each period. Every firm seeks to maximize the present discounted value of real profits, i.e.,

$$\max_{P_{jt}^*} \widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h Q_{t+h} \left(\frac{P_{jt}^*}{P_{t+h}} y_{j,t+h} - mc_{t+h} y_{j,t+h} \right) \tag{A.1}$$

where Q_t is a generic stochastic discount factor; P_{jt}^* is the optimal price set by the j^{th} firm; P_t is the aggregate price level; y_{jt} is the demand for the j^{th} firm's good; mc_t is the marginal cost; $\beta \in (0, 1)$ is a deterministic discount factor. The demand each firm faces and the aggregate price level are given by

$$y_{jt} = \left(\frac{P_{jt}^*}{P_t} \right)^{-\zeta} y_t \qquad\qquad P_t = \left[\int_{j=0}^{1} P_{jt}^{1-\zeta} \right]^{\frac{1}{\zeta-1}} \tag{A.2}$$

where $\zeta > 1$ is the elasticity of substitution among the differentiated goods. Substituting for $y_{j,t+h}$ into (A.1), we have

$$\max_{P_{jt}^*} \widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h Q_{t+h} \left(\left(\frac{P_{jt}^*}{P_{t+h}} \right)^{1-\zeta} - \left(\frac{P_{jt}^*}{P_{t+h}} \right)^{-\zeta} mc_{t+h} \right) y_{t+h} \tag{A.3}$$

The first-order condition with respect to P_{jt}^* is

$$\frac{P_{jt}^*}{P_t} = \frac{\zeta}{\zeta - 1} \frac{\widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h Q_{t+h} y_{t+h} mc_{t+h} \pi_{t,t+h}^{\zeta}}{\widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h Q_{t+h} y_{t+h} \pi_{t,t+h}^{\zeta-1}} \tag{A.4}$$

where $\pi_{t,t+h} = \frac{P_{t+h}}{P_t} = \prod_{l=0}^{h} \pi_{t+l}$. Due to Calvo pricing, the aggregate price level in (A.2) can be rewritten as

$$P_t = \left[\alpha P_{t-1}^{1-\zeta} + (1-\alpha)(P_t^*)^{1-\zeta} \right]^{\frac{1}{\zeta-1}} \tag{A.5}$$

Assume that the steady state for inflation is $\bar{\pi} = 1$. From (A.5), we have that in the steady state, $P^*/P = 1$. Then, from the optimality condition in (A.4) it follows that in the steady-state equilibrium $\bar{mc} = \frac{\zeta-1}{\zeta}$. Log-linearizing the first-order condition around steady-state values and dropping the subscript j, since every firm has the same optimality condition, we have

$$\hat{p}_t^* = \widetilde{\mathbb{E}}_t \sum_{h=0}^{\infty} (\alpha\beta)^h \left((1-\alpha\beta)\hat{mc}_{t+h} + \alpha\beta\hat{\pi}_{t+h+1}\right) \tag{A.6}$$

where $\hat{p}_t^* = \hat{P}_t^* - \hat{P}_t$ and $\hat{\pi}_{t+1} = \frac{P_{t+1}}{P_t}$ is inflation in period $(t+1)$.

B DSGE Model

Households. There is a continuum of identical households, $i \in [0,1]$, that consume from a set of differentiated goods, supply labor, and invest in riskless one-period bonds. First, households solve for the optimal allocation of consumption across differentiated goods, produced by monopolistically competitive firms $j \in [0,1]$, i.e.,

$$\min_{c_{it}(j)} \int_{j=0}^{1} P_{jt} c_{it}(j) dj$$

s.t.

$$c_{it} = \left[\int_{j=0}^{1} c_{it}(j)^{\frac{\zeta-1}{\zeta}} dj\right]^{\frac{\zeta}{\zeta-1}} \tag{B.1}$$

and

$$P_t = \left[\int_{j=0}^{1} P_{jt}^{1-\zeta}\right]^{\frac{1}{1-\zeta}} \tag{B.2}$$

where ζ is the elasticity of substitution among the differentiated goods. The corresponding Lagrangian is

$$\mathcal{L}_{it} = \min_{c_{it}(j)} \int_{j=0}^{1} P_{jt} c_{it}(j) dj + \chi_{it} \left(c_{it} - \left[\int_{j=0}^{1} (c_{it}(j))^{\frac{\zeta-1}{\zeta}} dj\right]^{\frac{\zeta}{\zeta-1}} \right)$$

where χ_{it} is the Lagrangian multiplier for the Dixit-Stiglitz consumption aggregator in (B.1). The first-order condition is

$$c_{it}(j) = \left(\frac{\chi_{it}}{P_{jt}}\right)^{\zeta} c_{it} \tag{B.3}$$

Plugging the expression for $c_{it}(j)$ above into (B.1) and rearranging terms,

$$\chi_{it} = \left[\int_{j=0}^{1} P_{jt}^{1-\zeta} dj\right]^{\frac{1}{1-\zeta}}$$

This implies further that

$$c_{it}(j) = \left(\frac{P_{jt}}{P_t}\right)^{-\zeta} c_{it} \tag{B.4}$$

Equation (B.4) defines the optimal demand of the i^{th} household for the j^{th} good. The intertemporal problem for the household is to

$$\max_{c_{it},H_{it},B_{it}} \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \beta^h \xi_{t+h} \left(\frac{(c_{i,t+h} - \eta c_{i,t+h-1})^{1-\sigma}}{1-\sigma} - \psi \frac{H_{i,t+h}^{1+\varphi}}{1+\varphi} \right)$$

with budget constraint satisfying

$$R_{t-1}B_{i,t-1} = B_{it} - W_t H_{it} - \int_{j=0}^{1} D_{it}(j)dj + \int_{j=0}^{1} P_{jt}c_{it}(j)dj$$

where H_{it} is labor supply; R_{t-1} gross return on nominal bond choice, $B_{i,t-1}$; W_t nominal wage; $D_{it}(j)$ nominal dividends from the j^{th} firm; and ξ_t a preference shock. Households internalize their optimal demand for good j into their intertemporal maximization problem, therefore

$$\int_{j=0}^{1} P_{jt}c_{it}(j)dj = P_t c_{it}$$

The budget constraint can be rewritten as

$$R_{t-1}B_{i,t-1} = B_{it} - W_t H_{it} - D_{it} + P_t c_{it} \tag{B.5}$$

where $\int_{j=0}^{1} D_{it}(j)dj = D_{it}$. The first-order conditions (FOC) with respect to consumption, bonds, and hours, respectively, are

$$\xi_t(c_{it} - \eta c_{i,t-1})^{-\sigma} - \beta\eta\widetilde{\mathbb{E}}_{it}\xi_{t+1}(c_{i,t+1} - \eta c_{it})^{-\sigma} = \lambda_{it} \tag{B.6}$$

$$\lambda_{it} = \beta\widetilde{\mathbb{E}}_{it}R_t\frac{\lambda_{i,t+1}}{\pi_{t+1}} \tag{B.7}$$

$$\psi\xi_t H_{it}^{\varphi} = \lambda_{it}w_t \tag{B.8}$$

where $w_t = \frac{W_t}{P_t}$ is the real wage.

Firms. There is a continuum of household-owned monopolistically competitive firms, $j \in [0, 1]$, that optimize with respect to price and labor demand. The production technology of each firm is

$$y_{jt} = z_t h_{jt}^{a_h} \tag{B.9}$$

where z_t and h_{jt} are a technology shock and labor demand, respectively, and $0 < a_h \leq 1$. The price optimization problem is subject to Calvo price stickiness as in Appendix A. Differently from Appendix A, if firms cannot optimize the price they can still adjust prices according to

$$P_{j,t+h} = P_{j,t+h-1}(\pi_{t+h-1})^{\rho_\pi} = P_{jt}\left(\frac{P_{t+h-1}}{P_{t-1}}^{1-\zeta}\right)^{\rho_\pi} \tag{B.10}$$

where $0 \leq \rho_\pi < 1$. Given the price aggregator in (B.2) and the nominal rigidities firms face, we have

$$P_t = \left[\alpha\left(P_{t-1}\left(\frac{P_{t-1}}{P_{t-2}}\right)^{\rho_\pi}\right)^{1-\zeta} + (1-\alpha)(P_t^*)\right]^{\frac{1}{1-\zeta}} \tag{B.11}$$

Each firm chooses the optimal price that will maximize the present discounted value of real profits such that the demand for its good is satisfied, and then hire the optimal amount of labor hours that will minimize production costs. Using backward induction, I solve the cost minimization problem first,

$$\mathcal{L}_{jt} = \min_{h_{jt}} w_t h_{jt} + mc_{jt}(y_{jt} - z_t h_{jt}^{a_h}) \tag{B.12}$$

where mc_{jt} is the real marginal cost of production. The FOC with respect to labor reads

$$mc_{jt} = \frac{w_t}{a_h z_t h_{jt}^{a_h - 1}} \tag{B.13}$$

The intermediate firms' problem is

$$\max_{P_{jt}^*} \widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h Q_{t+h}\left(\frac{P_{jt}^*}{P_{t+h}}\left(\frac{P_{t+h-1}}{P_{t-1}}\right)^{\rho_\pi} y_{j,t+h} - w_{t+h}h_{j,t+h}\right) \tag{B.14}$$

Aggregating $c_{it}(j)$ across households in (B.3), we have that the demand faced by the j^{th} firm is

$$y_{jt} = \left(\frac{P_{jt}^*}{P_t}\right)^{-\zeta} y_t \tag{B.15}$$

Substituting for y_{jt} and w_t in the pricing problem becomes

$$\max_{P_{jt}^*} \widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h Q_{t+h} y_{t+h}\left(\left(\frac{P_{jt}^*}{P_{t+h}}\right)^{1-\zeta}\left(\frac{P_{t+h-1}}{P_{t-1}}\right)^{\rho_\pi(1-\zeta)} - a_h mc_{j,t+h}\left(\frac{P_{jt}^*}{P_{t+h}}\right)^{-\zeta}\left(\frac{P_{t+h-1}}{P_{t-1}}\right)^{-\rho_\pi\zeta}\right) \tag{B.16}$$

The first-order condition with respect to P_{jt}^* reads

$$\widetilde{\mathbb{E}}_{jt}\sum_{h=0}^{\infty}(\alpha\beta)^h Q_{t+h}\pi_{t-1,t+h-1}^{-\rho_\pi\zeta}P_{t+h}^{\zeta-1}y_{t+h}\left(a_h\zeta mc_{j,t+h}P_{t+h}-(\zeta-1)P_{jt}^*\pi_{t-1,t+h-1}^{\rho_\pi}\right)=0 \tag{B.17}$$

Monetary Policy. The central bank controls nominal interest rates through a Taylor rule that reacts to inflation and output gap deviations from their steady-state values, with some interest rate smoothing, i.e.,

$$\frac{R_t}{\bar{R}}=\left(\frac{R_{t-1}}{\bar{R}}\right)^{\rho_r}\left(\frac{\pi_t}{\bar{\pi}}\right)^{(1-\rho_r)\phi_\pi}\left(\frac{x_t}{\bar{x}}\right)^{(1-\rho_r)\phi_x}e^{\sigma_v\varepsilon_t^v},\ \varepsilon_t^v\sim\mathcal{N}(0,1) \tag{B.18}$$

where x_t is the output gap; $\bar{\pi}$ and $\bar{x}$ denote the inflation target and output gap steady-state value, respectively; $\rho_r\in[0,1)$.

Steady-state Equilibrium. I calculate steady-state values:

$$\bar{\xi}=1 \qquad \bar{z}=1 \qquad \bar{v}=1 \tag{B.19}$$

$$\bar{\pi}=\beta\bar{R}=1 \tag{B.20}$$

$$\bar{\lambda}=\bar{y}^{-\sigma}(1-\beta\eta) \tag{B.21}$$

$$\bar{w}=\frac{\psi}{1-\beta\eta}(\bar{H})^{\varphi}(\bar{C})^{\sigma} \tag{B.22}$$

$$\bar{d}=\bar{C}-\frac{1-\beta}{\beta}\bar{b}-\bar{w}\bar{H} \tag{B.23}$$

$$\bar{y}=\bar{h}^{a_h} \tag{B.24}$$

$$\bar{mc}=\frac{\zeta-1}{a_h\zeta} \tag{B.25}$$

where $\bar{b}=\frac{\bar{B}}{\bar{P}}$ and $\bar{d}=\frac{\bar{D}}{\bar{P}}$ denote steady-state bond holdings and dividends in real terms.

B.1 Log-linearized Model

Households. Log-linearizing (B.6) and (B.7) around steady states generates

$$\hat{c}c_{it}=\widetilde{\mathbb{E}}_{it}\hat{c}c_{i,t+1}-\frac{1-\beta\eta}{\sigma}\widetilde{\mathbb{E}}_{it}(\hat{R}_t-\hat{\pi}_{t+1})+\frac{1}{\sigma}\widetilde{\mathbb{E}}_{it}\left(\hat{g}_t-\hat{g}_{t+1}\right) \tag{B.26}$$

where $\hat{c}c_{it} = \hat{c}_{it} - \eta\hat{c}_{i,t-1} - \beta\eta\widetilde{\mathbb{E}}_{it}(\hat{c}_{i,t+1} - \hat{c}_{it})$ and $\hat{g}_t = \hat{\xi}_t - \beta\eta\hat{\xi}_{t+1}$. One can make inferences about $\widetilde{\mathbb{E}}_{it}\hat{c}c_{i,t+1}$ by iterating the Euler equation above, i.e.,

$$\hat{c}c_{i,t+1} = \widetilde{\mathbb{E}}_{i,t+1}\hat{c}c_{i,t+2} - \frac{1-\beta\eta}{\sigma}\widetilde{\mathbb{E}}_{i,t+1}(\hat{R}_{t+1} - \hat{\pi}_{t+2}) + \frac{1}{\sigma}\widetilde{\mathbb{E}}_{i,t+1}(\hat{g}_{t+1} - \hat{g}_{t+2})$$

So,

$$\begin{aligned}\widetilde{\mathbb{E}}_{it}\hat{c}c_{i,t+1} &= \widetilde{\mathbb{E}}_{it}\widetilde{\mathbb{E}}_{i,t+1}\hat{c}c_{i,t+2} - \frac{1-\beta\eta}{\sigma}\widetilde{\mathbb{E}}_{it}\widetilde{\mathbb{E}}_{i,t+1}(\hat{R}_{t+1} - \hat{\pi}_{t+2}) - \frac{1}{\sigma}\widetilde{\mathbb{E}}_{it}\widetilde{\mathbb{E}}_{i,t+1}(\hat{g}_{t+1} - \hat{g}_{t+2}) \\ &= \widetilde{\mathbb{E}}_{it}\hat{c}c_{i,t+2} - \frac{1-\beta\eta}{\sigma}\widetilde{\mathbb{E}}_{it}(\hat{R}_{t+1} - \hat{\pi}_{t+2}) - \frac{1}{\sigma}\widetilde{\mathbb{E}}_{it}(\hat{g}_{t+1} - \hat{g}_{t+2})\end{aligned}$$

where the second equality is an application of the law of iterative expectations. Plugging expectations into the log-linear individual Euler equation, we get

$$\hat{c}c_{it} = \widetilde{\mathbb{E}}_{it}\hat{c}c_{i,t+2} - \frac{1-\beta\eta}{\sigma}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{t+1}(\hat{R}_{t+h} - \hat{\pi}_{t+h+1}) + \frac{1}{\sigma}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{t+1}(\hat{g}_{it+h} - \hat{g}_{t+h+1})$$

Similarly, the h-periods-ahead forwardly iterated Euler equation reads

$$\hat{c}c_{it} = \widetilde{\mathbb{E}}_{it}\hat{c}c_{i,t+h} - \frac{1-\beta\eta}{\sigma}\widetilde{\mathbb{E}}_{it}\sum_{l=0}^{h-1}(\hat{R}_{t+l} - \hat{\pi}_{t+l+1}) + \frac{1}{\sigma}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{t+k-1}(\hat{g}_{t+h} - \hat{g}_{t+h+1}) \tag{B.27}$$

It is worth noting that if households knew that everyone is subject to the same preference shocks, and that they all have the same preferences over consumption and labor, then they would know that in the infinite future, consumption is expected to be at its steady state, implying that $\lim_{h\to\infty}\widetilde{\mathbb{E}}_{it}\hat{c}c_{i,t+h} = 0$. This would further imply that households would use the one-step-ahead Euler equation, as under RE. However, households have imperfect knowledge about the rest of the population, and one needs to combine (B.27) with the infinitely forward iterated household's budget constraint in (B.5):

$$\begin{aligned}
B_{i,t-1} &= \frac{B_{it}}{R_{t-1}} - \frac{W_t H_{it}}{R_{t-1}} - \frac{D_{it}}{R_{t-1}} + \frac{P_t c_{it}}{R_{t-1}} \\
&= \widetilde{\mathbb{E}}_{it} RR_{t-1,t} B_{i,t+1} - \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{t+1} RR_{t-1,t+h}(W_{t+h}H_{i,t+h} + D_{i,t+h}) + \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{t+1} RR_{t-1,t+h} P_{t+h} c_{i,t+h} \\
&= \dots \\
&= \lim_{\tau \to \infty} \widetilde{\mathbb{E}}_{it} RR_{t-1,t+h} B_{i,t+h+1} - \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} RR_{t-1,t+h}(W_{t+h}H_{i,t+h} + D_{i,t+h}) + \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} RR_{t-1,t+h} P_{t+h} c_{i,t+h} \\
&= \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} F_{t-1,t+h} P_{t+h} c_{i,t+h} - \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} RR_{t-1,t+h}(W_{t+h}H_{i,t+h} + D_{i,t+h})
\end{aligned}$$

where $RR_{t-1,t+h} = \prod_{l=t-1}^{t+h} \frac{1}{R_l}$. To get the last equality I impose the appropriate no-Ponzi constraint, i.e., $\lim_{h\to\infty} \widetilde{\mathbb{E}}_{it} RR_{t-1,t+h} B_{i,t+h+1} = 0$. To write everything in real terms, I divide by P_{t-1} and get

$$b_{i,t-1} = \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} RR_{t-1,t+h} \pi_{t-1,t+h} c_{i,t+h} - \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} RR_{t-1,t+h} \pi_{t-1,T}(w_{t+h}H_{i,t+h} + d_{i,t+h}) \tag{B.28}$$

The log-linearized version of the iterated budget constraint is:

$$\begin{aligned}
\bar{b}\hat{b}_{i,t-1} = {} & \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \bar{RR}_{t-1,t+h} \bar{\pi}_{t-1,t+h} \bar{c}(\hat{RR}_{t-1,t+h} + \hat{\pi}_{t-1,T} + \hat{c}_{i,t+h}) \\
& - \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \bar{RR}_{t-1,t+h} \bar{\pi}_{t-1,t+h} \bar{w}\bar{H}(\hat{RR}_{t-1,t+h} + \hat{\pi}_{t-1,t+h} + \hat{w}_{t+h} + \hat{H}_{i,t+h}) \\
& - \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \bar{RR}_{t-1,t+h} \bar{\pi}_{t-1,t+h} \bar{d}(\hat{RR}_{t-1,t+h} + \hat{\pi}_{t-1,t+h} + \hat{d}_{i,t+h})
\end{aligned}$$

Using (B.21), $\bar{RR}_{t-1,t+h}\bar{\pi}_{t-1,t+h} = \frac{\bar{\pi}^{h+1}}{\bar{R}^{h+1}} = \beta^{h+1}$. Substituting for $\bar{RR}_{t-1,t+h}\bar{\pi}_{t-1,t+h}$ and optimal labor supply, the final log-linearized iterated budget constraint is

$$\begin{aligned}
\bar{b}\hat{b}_{i,t-1} = {} & \left(\bar{c} + \bar{w}\bar{H}\frac{\sigma}{\varphi}\right) \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \beta^{h+1} \hat{c}_{i,t+h} - \bar{w}\bar{H}\frac{1+\varphi}{\varphi} \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \beta^{h+1} \hat{w}_{t+h} - \bar{d}\widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \beta^{h+1} \hat{d}_{i,t+h} \\
& + (\bar{c} - \bar{w}\bar{H} - \bar{d}) \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \beta^{h+1} (\hat{RR}_{t-1,t+h} + \hat{\pi}_{t-1,t+h})
\end{aligned} \tag{B.29}$$

Next, recall that $\hat{c}c_{it} = \hat{c}_{it} - \eta\hat{c}_{i,t-1} - \beta\eta\widetilde{\mathbb{E}}_{it}(\hat{c}_{i,t+1} - \eta\hat{c}_{it})$, from which it follows that

$$\hat{c}_{it} = \hat{c}c_{it} + \eta\hat{c}_{i,t-1} + \beta\eta\widetilde{\mathbb{E}}_{it}(\hat{c}_{i,t+1} + \eta\hat{c}_{it}) \tag{B.30}$$

Substituting for $\hat{c}_{i,t+h}$ into (B.29), I rewrite the intertermporal budget constraint as

$$\begin{aligned}\bar{b}\hat{b}_{i,t-1} = &\left(\bar{c} + \bar{w}\bar{H}\frac{\sigma}{\varphi}\right)\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h+1}\left(\hat{c}c_{i,t+h} + \eta\hat{c}_{i,t+h-1} + \beta\eta(\hat{c}_{i,t+h+1} - \hat{c}_{i,t+h})\right)\\ &- \bar{w}\bar{H}\frac{1+\varphi}{\varphi}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h+1}\hat{w}_{t+h} - \bar{d}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h+1}\hat{d}_{i,t+h}\\ &+ (\bar{c} - \bar{w}\bar{H} - \bar{d})\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h+1}(\hat{R}R_{t-1,t+h} + \hat{\pi}_{t-1,t+h})\end{aligned} \tag{B.31}$$

From (B.27), one can isolate $\widetilde{\mathbb{E}}_{it}\hat{c}c_{i,t+h}$ and substitute for it into (B.29):

$$\begin{aligned}\bar{b}\hat{b}_{i,t-1} = &\beta(\bar{c} + \bar{w}\bar{H}\frac{\sigma}{\varphi})\left(\frac{1}{1-\beta}\hat{c}c_{it} + \eta\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}(\hat{c}_{i,t+h-1} + \beta(\hat{c}_{i,t+h+1} - \eta\hat{c}_{i,t+h}))\right)\\ &+ \frac{\beta(\bar{c} + \frac{\sigma}{\varphi})}{\sigma(1-\beta)}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h}\left((1-\beta\eta)(\hat{R}_{t+h} - \hat{\pi}_{t+h+1}) - (\hat{g}_{t+h} - \hat{g}_{t+h+1})\right)\\ &- \bar{w}\bar{H}\frac{1+\varphi}{\varphi}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h+1}\hat{w}_{t+h} - \bar{d}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h+1}\hat{d}_{i,t+h}\\ &+ (\bar{c} - \bar{w}\bar{H} - \bar{d})\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h+1}(\hat{R}R_{t-1,t+h} + \hat{\pi}_{t-1,t+h})\end{aligned}$$

Isolating $\hat{c}c_{it}$, one retrieves the individual demand in terms of $\hat{c}c_{it}$,

$$\begin{aligned}\hat{c}c_{it} = &\frac{\bar{b}(1-\beta)}{\beta(\bar{c} + \bar{w}\bar{H}\frac{\sigma}{\varphi})}\hat{b}_{i,t-1} - \eta(1-\beta)\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h}\left(\hat{c}_{i,t+h-1} + \beta(\hat{c}_{i,t+h+1} - \eta\hat{c}_{i,t+h})\right)\\ &+ \frac{1-\beta}{\bar{c} + \bar{w}\bar{H}\frac{\sigma}{\varphi}}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h}\left(\frac{\bar{w}\bar{H}(1+\varphi)}{\varphi}\hat{w}_{t+h} + \bar{d}\hat{d}_{i,t+h}\right) - \frac{\beta(1-\beta\eta)}{\sigma}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h}(\hat{R}_{t+h} - \hat{\pi}_{t+h+1})\\ &- \frac{\beta}{\sigma}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h}(\hat{g}_{t+h} - \hat{g}_{t+h+1}) - \frac{\beta(1-\beta)(\bar{c} - \bar{w}\bar{H} - \bar{d})}{\bar{c} + \bar{w}\bar{H}\frac{\sigma}{\varphi}}\widetilde{\mathbb{E}}_{it}\sum_{h=0}^{\infty}\beta^{h}(\hat{R}R_{t-1,t+h} + \hat{\pi}_{t-1,t+h})\end{aligned} \tag{B.32}$$

Let $w_c = \frac{\bar{w}\bar{H}(1+\varphi)}{\varphi(\bar{c}+\bar{w}\bar{H}\frac{\sigma}{\varphi})}$ and $d_c = \frac{\bar{d}}{\bar{c}+\bar{w}\bar{H}\frac{\sigma}{\varphi}}$. Define

$$\hat{\hat{y}}_t = \hat{y}_t - \eta\hat{y}_{t-1} \tag{B.33}$$

Aggregating equation (B.32), imposing market clearing conditions such that $\hat{c}_t = \hat{y}_t = w_c\hat{w}_t + d_c\hat{d}_t$, $\hat{c}c_t = \hat{y}y_t$, $(\bar{c} - \bar{w}\bar{H} - \bar{d}) = 0$, and $\hat{b}_t = 0$ (since households are homogeneous) one gets

$$\hat{\tilde{y}}_t = (1-\beta+\beta\eta)\widetilde{\mathbb{E}}_t\hat{\tilde{y}}_{t+1} + \widetilde{\mathbb{E}}_t\sum_{h=0}^{\infty}\beta^h\left[\beta(1-\eta)(1-\beta)\widetilde{\mathbb{E}}_t\hat{\tilde{y}}_{t+h+2} - \frac{1-\beta\eta}{\sigma}(\hat{R}_{t+h} - \hat{\pi}_{t+h+1}) - \frac{1}{\sigma}(\hat{g}_{t+h} - \hat{g}_{t+h+1})\right] \tag{B.34}$$

Let $\hat{x}_t = \hat{y}_t - \hat{y}_t^n$ be the output gap with $\hat{y}_t^n$ being the potential level of output. Further, let $\hat{\tilde{x}}_t = \hat{x}_t - \eta\hat{x}_{t-1}$. Rewriting equation (B.34) in terms of the output gap yields the aggregate demand equation:

$$\hat{\tilde{x}}_t = (1-\beta+\beta\eta)\widetilde{\mathbb{E}}_t\hat{\tilde{x}}_{t+1} + \widetilde{\mathbb{E}}_t\sum_{h=0}^{\infty}\beta^h\left(\beta\left(1-\beta\right)\left(1-\eta\right)\hat{\tilde{x}}_{t+h+2} - \frac{1-\beta\eta}{\sigma}\left(\hat{R}_{t+h} - \hat{\pi}_{t+h+1} - \hat{e}_{t+h}\right)\right) \tag{B.35}$$

and $\hat{e}_t = \frac{\sigma}{1-\beta\eta}\left((\hat{y}_{t+1}^n - \eta\hat{y}_t^n - \hat{g}_{t+1}) - (\hat{y}_t^n - \eta\hat{y}_{t-1}^n - \hat{g}_t)\right)$ is such that

$$\hat{e}_t = \rho_e\hat{e}_{t-1} + \sigma_e\varepsilon_t^e,\ \varepsilon_t^e \sim \mathcal{N}(0,1) \tag{B.36}$$

Applying the myopic adjustment to (B.35), the aggregate demand is rewritten as

$$\hat{\tilde{x}}_t = n(1-\beta+\beta\eta)\widetilde{\mathbb{E}}_t^{\star}\hat{\tilde{x}}_{t+1} + \widetilde{\mathbb{E}}_t^{\star}\sum_{h=0}^{\infty}(\beta n)^h\left(n^2\beta\left(1-\beta\right)\left(1-\eta\right)\hat{\tilde{x}}_{t+h+2} - \frac{1-\beta\eta}{\sigma}\left(\hat{R}_{t+h} - \hat{\pi}_{t+h+1} - \hat{e}_{t+h}\right)\right) \tag{B.37}$$

mis-specified, substituting for $\hat{\tilde{x}}_t = (\hat{x}_t - \eta\hat{x}_{t-1})$ delivers

$$\begin{aligned}\hat{x}_t = {} & \frac{\eta}{1+n\eta\upsilon}\hat{x}_{t-1} + n\frac{\upsilon - n\beta\eta(1-\beta)(1-\eta)}{1+n\eta\upsilon}\widetilde{\mathbb{E}}_t^{\star}\hat{x}_{t+1} + \frac{\beta n^2(1-\beta)(1-\eta)(1-n\beta\eta)}{1+n\eta\upsilon}\widetilde{\mathbb{E}}_t^{\star}\sum_{h=0}^{\infty}(\beta n)^h\hat{x}_{t+h+2} \\ & - \widetilde{\mathbb{E}}_t^{\star}\sum_{h=0}^{\infty}(\beta n)^h\frac{1-\beta\eta}{\sigma(1+n\eta\upsilon)}\left(\hat{R}_{t+h} - \hat{\pi}_{t+h+1} - \hat{e}_{t+h}\right)\end{aligned} \tag{B.38}$$

where $\upsilon = (1-\beta+\beta\eta)$.

Firms. Log-linearizing firms' optimal price condition, we get,

$$\hat{P}_{jt}^* - \hat{P}_t = \widetilde{\mathbb{E}}_{jt}\sum_{h=0}^{\infty}(\alpha\beta)^h\left[(1-\alpha\beta)\hat{mc}_{j,t+h} + \alpha\beta(\hat{\pi}_{t+h+1} - \rho_\pi\hat{\pi}_{t+h-1})\right] \tag{B.39}$$

Define $\hat{p}^*_{jt} = \hat{P}^*_{jt} - \hat{P}_t$. The marginal cost of the j^{th} firm is given by

$$\hat{mc}_{j,t+h} = \hat{w}_t + \frac{1}{a_h}\hat{z}_t + \frac{1-a_h}{a_h}\hat{y}_t - \zeta\frac{1-a_h}{a_h}\hat{p}^*_{jt} \tag{B.40}$$

$$\left(1+\zeta\frac{1-a_h}{a_h}\right)\hat{p}^*_{jt} = \widetilde{\mathbb{E}}_{jt}\sum_{h=0}^{\infty}(\alpha\beta)^h\left[(1-\alpha\beta)\left(\hat{w}_{t+h} + \frac{1}{a_h}\hat{z}_{t+h} + \frac{1-a_h}{a_h}\hat{y}_{t+h}\right) + \alpha\beta(\hat{\pi}_{t+h+1} - \rho_\pi\hat{\pi}_{t+h-1})\right] \tag{B.41}$$

From (B.10), $\hat{p}^*_{jt} = \hat{P}^*_{jt} - \hat{P}_t = \frac{\alpha}{1-\alpha}(\hat{\pi}_t - \rho_\pi\hat{\pi}_{t-1}) = \frac{\alpha}{1-\alpha}\hat{\tilde{\pi}}_t$. Since all firms face the same optimal pricing condition above, I drop the subscript j. Define $\hat{u}_t$ to be a supply shock that captures deviations of the empirical output gap from the theoretically relevant gap, assumed to follow an AR(1) process

$$\hat{u}_t = \rho_u\hat{u}_{t-1} + \sigma_u\varepsilon^u_t,\ \varepsilon^u_t \sim \mathcal{N}(0,1) \tag{B.42}$$

Then, the aggregated optimal pricing rule can be written as

$$\hat{\tilde{\pi}}_t = \kappa\left(\omega\hat{x}_t + \frac{\sigma}{1-\eta\beta}\hat{\tilde{x}}_t\right) + \widetilde{\mathbb{E}}_t\sum_{h=0}^{\infty}(\alpha\beta)^h\left(\kappa\alpha\beta\left(\omega\hat{x}_{t+h+1} + \frac{\sigma\beta(\alpha-\eta)}{\alpha(1-\eta\beta)}\hat{\tilde{x}}_{t+h+1}\right) + \beta(1-\alpha)\hat{\tilde{\pi}}_{t+h+1} + \hat{u}_{t+h}\right) \tag{B.43}$$

where $\kappa = \frac{(1-\alpha)(1-\alpha\beta)}{\alpha(a_h+\zeta(1-a_h))}$ and $\omega = (1+\varphi-a_h)$. Applying the myopic adjustment yields

$$\begin{aligned}\hat{\tilde{\pi}}_t &= \kappa\left(\omega\hat{x}_t + \frac{\sigma}{1-\eta\beta}\hat{\tilde{x}}_t\right) + \widetilde{\mathbb{E}}^\star_t\sum_{h=0}^{\infty}(\alpha\beta n)^h\left(\kappa\alpha\beta n\left(\omega\hat{x}_{t+h+1} + \frac{\sigma(\alpha-\eta)}{\alpha(1-\eta\beta)}\hat{\tilde{x}}_{t+h+1}\right)\right)\\ &+ \widetilde{\mathbb{E}}^\star_t\sum_{h=0}^{\infty}(\alpha\beta n)^h\left(n\beta(1-\alpha)\hat{\tilde{\pi}}_{t+h+1} + \hat{u}_{t+h}\right)\end{aligned} \tag{B.44}$$

Substituting for $\hat{\tilde{\pi}}_t = \hat{\pi}_t - \rho_\pi\hat{\pi}_{t-1}$ and $\hat{\tilde{x}}_t = \hat{x}_t - \eta\hat{x}_{t-1}$,

$$\begin{aligned}\hat{\pi}_t &= \frac{1}{1+n\beta\rho_\pi(1-\alpha)}(\rho_\pi\hat{\pi}_{t-1} - \kappa\eta\tau\hat{x}_{t-1}) + \frac{\kappa(\omega+\tau(1-n\eta\beta(\alpha-\eta))}{1+n\beta\rho_\pi(1-\alpha)}\hat{x}_t + \frac{1}{1-\alpha\beta n\rho_u}\hat{u}_t\\ &+ \frac{n\beta}{1+n\beta\rho_\pi(1-\alpha)}\widetilde{\mathbb{E}}^\star_t\sum_{h=0}^{\infty}(\alpha\beta n)^h\left((1-\alpha)(1-\alpha\beta n\rho_\pi)\hat{\pi}_{t+h+1} + \kappa(\alpha\omega + \tau(\alpha-\eta)(1-\alpha\beta n\eta)\hat{x}_{t+h+1}\right)\end{aligned} \tag{B.45}$$

where $\tau = \frac{\sigma}{1-\beta\eta}$.

Monetary Policy. The log-linearized version of the policy rule is

$$\hat{R}_t = \rho_r\hat{R}_{t-1} + (1-\rho_r)\phi_\pi\hat{\pi}_t + (1-\rho_r)\phi_x\hat{x}_t + \sigma_v\varepsilon_t \tag{B.46}$$

Model in Matrix Form. The aggregate economy model in matrix form is described by

$$A_0(\Theta)S_t = A_1(\Theta)S_{t-1} + A_{02}(\Theta)\widetilde{\mathbb{E}}_t^\star S_{t+1} + \widetilde{\mathbb{E}}_t^\star \sum_{h=0}^{\infty} F^h A_{12}(\Theta)S_{t+h+2} + B(\Theta)\mathcal{E}_t \tag{B.47}$$

where $S_t = \begin{bmatrix}\hat{x}_t & \hat{\pi}_t & \hat{R}_t & \hat{e}_t & \hat{u}_t\end{bmatrix}'$; $\mathcal{E}_t = \begin{bmatrix}\varepsilon_t^e & \varepsilon_t^u & \varepsilon_t^v\end{bmatrix}'$; $\Theta = \{\alpha, \beta, n, \sigma, \kappa, \eta, \rho_\pi, \omega, \phi_\pi, \phi_x, \rho_r, \rho_e, \rho_u, \sigma_e, \sigma_u, \sigma_v\}$, F is a zero matrix, with only the first two diagonal entries equal to βn and $\alpha\beta n$, respectively. Using results from the previous subsection, the perceived law of motion (PLM) in matrix form can be written as

$$S_t = \underbrace{\mathbf{\Delta}_{t-1} + \mathbf{\Gamma}_{t-1}(S_{t-1} - \mathbf{\Delta}_{t-1})}_{\text{PLM for aggregate endo var's}} + \underbrace{HS_{t-1}}_{\text{PLM for the shocks}} + \widetilde{\boldsymbol{\epsilon}}_t \tag{B.48}$$

where $\boldsymbol{\delta}_t = \begin{bmatrix}\boldsymbol{\delta}_t' & \mathbf{0}_{1x2}\end{bmatrix}'$; $\mathbf{\Gamma}_t = \begin{bmatrix}\boldsymbol{\gamma}_t & \mathbf{0}_{3x2} \\ \mathbf{0}_{2x3} & \mathbf{0}_{2x2}\end{bmatrix}$; H is a diagonal matrix with diagonal equal to $\begin{bmatrix}\mathbf{0}_{1x3} & \rho_e & \rho_u\end{bmatrix}'$; $\widetilde{\boldsymbol{\epsilon}}_t = \begin{bmatrix}\boldsymbol{\epsilon}_t' & \sigma_e\varepsilon_t^e & \sigma_u\varepsilon_t^u\end{bmatrix}'$. The forecast of the state vector $h \geq 1$ periods ahead is described by

$$\widetilde{\mathbb{E}}_t^\star S_{t+h} = \underbrace{\mathbf{\Delta}_{t-1} + \mathbf{\Gamma}_{t-1}^{\tau-t+1}(S_{t-1} - \mathbf{\Delta}_{t-1})}_{\text{forecast of endo var's}} + \underbrace{H^h S_t}_{\text{forecast of shocks}} \tag{B.49}$$

Plugging (B.49) into (B.47), we get the actual law of motion:

$$\widetilde{A}_0(\Theta)S_t = \widetilde{A}_1(\Theta)\mathbf{\Delta}_{t-1} + \widetilde{A}_2(\Theta, \mathbf{\Gamma}_{t-1})S_{t-1} + B\mathcal{E}_t \tag{B.50}$$

where

$$\widetilde{A}_0 = A_0 - A_{02}H - \left(\sum_{h=0}^{\infty} F^h A_{12} H^h\right) H$$

$$\widetilde{A}_1 = A_{02}(I - \Gamma_{t-1}^2)\Delta_{t-1} + \sum_{h=0}^{\infty} F^h A_{12} - \left(\sum_{h=0}^{\infty} F^h A_{12}\mathbf{\Gamma}_{t-1}^h\right)\mathbf{\Gamma}_{t-1}^2$$

$$\widetilde{A}_2 = A_1 + A_{02}\mathbf{\Gamma}_{t-1}^2 + \left(\sum_{h=0}^{\infty} F^h A_{12}\mathbf{\Gamma}_{t-1}^h\right)\mathbf{\Gamma}_{t-1}^3$$

The infinite sums are defined as follows

$$\sum_{h=0}^{\infty} F^h = (I - F)^{-1}$$

$$vec\left(\sum_{h=0}^{\infty} F^h A_{12} H^h\right) = (I - H \otimes F)^{-1} A_{12}(:)$$

$$\begin{aligned} vec\left(\sum_{h=0}^{\infty} F^h A_{12} \mathbf{\Gamma}_{t-1}^h\right) &= vec(A_{12} + F A_{12} \mathbf{\Gamma}_{t-1} + F^2 A_{12} \mathbf{\Gamma}_{t-1}^2 + ...) \\ &= (I \otimes I + \mathbf{\Gamma}'_{t-1} \otimes F + (\mathbf{\Gamma}'_{t-1})^2 \otimes F^2 + ...) \\ &= (I - \mathbf{\Gamma}'_{t-1} \otimes F)^{-1} A_{12}(:) \end{aligned}$$

The last equality uses the Kronecker product property that $(\mathbf{\Gamma}'_{t-1} \otimes F)(\mathbf{\Gamma}'_{t-1} \otimes F) = (\mathbf{\Gamma}'_{t-1})^2 \otimes F^2$.

B.2 Aggregate Demand and Supply under Well-specified Forecasting Rules

In this subsection, I derive the equilibrium conditions when $\widetilde{\mathbb{E}}_t^\star$ is associated with well-specified forecasting rules. Consider the aggregate demand

$$\hat{\tilde{x}}_t = n\upsilon \widetilde{\mathbb{E}}_t^\star \hat{\tilde{x}}_{t+1} + \widetilde{\mathbb{E}}_t^\star \sum_{h=0}^{\infty} (\beta n)^h \left(n^2 \beta (1-\beta)(1-\eta) \hat{\tilde{x}}_{t+h+2} - \frac{1-\beta\eta}{\sigma} \left(\hat{R}_{t+h} - \hat{\pi}_{t+h+1} - \hat{e}_{t+h} \right) \right) \tag{B.51}$$

Then,

$$\widetilde{\mathbb{E}}_t^\star \hat{\tilde{x}}_{t+1} = n\upsilon \widetilde{\mathbb{E}}_t^\star \hat{\tilde{x}}_{t+2} + \widetilde{\mathbb{E}}_t^\star \sum_{h=0}^{\infty} (\beta n)^h \left(n^2 \beta \left(\upsilon - \eta \right) \hat{\tilde{x}}_{t+h+3} - \frac{1-\beta\eta}{\sigma} \left(\hat{R}_{t+h+1} - \hat{\pi}_{t+h+2} \right) - \hat{e}_{t+h} \right) \Bigg) \tag{B.52}$$

from which

$$\widetilde{\mathbb{E}}_t^\star \sum_{h=0}^{\infty} (\beta n)^h \left(n^2 \beta (\upsilon - \eta) \hat{\tilde{x}}_{t+h+3} - \frac{1-\beta\eta}{\sigma} \left(\hat{R}_{t+h+1} - \hat{\pi}_{t+h+2} - \hat{e}_{t+h} \right) \right) = \widetilde{\mathbb{E}}_t^\star \hat{\tilde{x}}_{t+1} - n\upsilon \widetilde{\mathbb{E}}_t^\star \hat{\tilde{x}}_{t+2} \tag{B.53}$$

Substituting for the expression in the left-hand side in the equation above into the original aggregate demand in (B.51) and setting $\widetilde{\mathbb{E}}_t^\star \equiv \mathbb{E}_t$, we have

$$\hat{\tilde{x}}_t = n \mathbb{E}_t \left((1+\beta\eta) \hat{\tilde{x}}_{t+1} - n\beta\eta \hat{\tilde{x}}_{t+2} \right) - \frac{\sigma}{1-\beta\eta} (\hat{R}_t - \mathbb{E}_t \hat{\pi}_{t+1}) + \frac{\sigma}{1-\beta\eta} \hat{e}_t \tag{B.54}$$

If $n = 1$, then the equation above coincides with the standard Euler equation derived under FIRE.

Similarly, consider the aggregate supply,

$$\begin{aligned}\hat{\tilde{\pi}}_t &= \kappa\left(\omega\hat{x}_t + \tau\hat{\tilde{x}}_t\right) + \widetilde{\mathbb{E}}^\star_t\sum_{h=0}^{\infty}(\alpha\beta n)^h\left(\kappa\alpha\beta n\left(\omega\hat{x}_{t+h+1} + \frac{\tau(\alpha-\eta)}{\alpha}\hat{\tilde{x}}_{t+h+1}\right)\right)\\ &\quad + \widetilde{\mathbb{E}}^\star_t\sum_{h=0}^{\infty}(\alpha\beta n)^h\left(n\beta(1-\alpha)\hat{\tilde{\pi}}_{t+h+1} + \hat{u}_{t+h}\right)\end{aligned} \tag{B.55}$$

Hence,

$$\begin{aligned}\widetilde{\mathbb{E}}^\star_t\hat{\tilde{\pi}}_{t+1} &= \kappa\left(\widetilde{\mathbb{E}}^\star_t\left(\tau\hat{\tilde{x}}_{t+1} + \omega\sum_{h=0}^{\infty}(\alpha\beta n)^h\hat{x}_{t+h+1}\right) + \beta n\tau(\alpha-\eta)\widetilde{\mathbb{E}}^\star_t\sum_{h=0}^{\infty}(\alpha\beta n)^h\hat{\tilde{x}}_{t+h+2}\right)\\ &\quad + \widetilde{\mathbb{E}}^\star_t\sum_{h=0}^{\infty}(\alpha\beta n)^h\left(n\beta(1-\alpha)\hat{\tilde{\pi}}_{t+h+2} + \hat{u}_{t+h}\right)\end{aligned} \tag{B.56}$$

Isolating $\kappa\omega\widetilde{\mathbb{E}}^\star_t\sum_{h=0}^{\infty}(\alpha\beta n)^h\hat{x}_{t+h+1}$ from (B.56), substituting for it into (B.55), and setting $\widetilde{\mathbb{E}}^\star_t \equiv \mathbb{E}_t$, we have

$$\hat{\tilde{\pi}}_t = \kappa\omega\hat{x}_t + \tau\mathbb{E}_t(\hat{\tilde{x}}_t - \beta\eta n\hat{\tilde{x}}_{t+1}) + n\beta\mathbb{E}_t\hat{\tilde{\pi}}_{t+1} + \hat{u}_t \tag{B.57}$$

If $n = 1$, then the equation above coincides with the standard Phillips curve derived under FIRE.

C Proofs

C.1 Proposition 1

Let the data-generating process for inflation be given by $\hat{\pi}_t = a\hat{mc}_t + b\hat{\pi}_{t-1}$, where $a = \frac{\kappa}{1-\alpha\beta\rho n}$ and $b = \frac{\beta n(1-\alpha)}{1-\alpha\beta n\gamma^*}(\gamma^*)^2$. Then, one can show that

$$F(\gamma) = \frac{\mathbb{E}(\hat{\pi}_t\hat{\pi}_{t-1})}{\mathbb{E}(\hat{\pi}_t^2)} = \frac{b+\rho}{1+\rho b} \tag{C.1}$$

For a CE equilibrium to exist, we must have that $F(\gamma) = \gamma$ for at least one value of $\gamma \in (0,1)$. Moreover, $F(\gamma)$ is an increasing function of γ, with $F(0) = \rho > 0$ and $F(1) = \frac{\beta n(1-\alpha)+\rho(1-\alpha\beta n)}{1-\alpha\beta n+\rho\beta n(1-\alpha)}$, where $\rho \leq F(1) < 1$. Therefore, $F(\gamma)$ crosses the 45° line at least once; that is, a CE equilibrium is guaranteed to exist. Since $F(\gamma) \geq \rho$, it follows that $\gamma^* \in [\rho, 1)$.

To show that the CE equilibrium is unique, I show that $F(\gamma)$ is convex whenever it intersects with the 45° line, i.e., whenever (C.1) holds. Note that $F(\gamma)$ is an increasing function of γ, such that $F(0) = \rho$ and $F(1) < 1$. Therefore, if multiple fixed points existed for $\gamma \in [0,1)$, it must be that at least one fixed point, $F(\gamma)$, is concave.

$$F''(\gamma)\mid_{\gamma=\gamma^*} = (1-\rho\gamma^*)\frac{b''(1+\rho b)-\rho(b')^2}{(1+\rho b)^2} \tag{C.2}$$

where $b' = \partial b/\partial\gamma$ and b'' denotes the second-order partial derivative of b w.r.t. γ. Therefore, $F''(\gamma=\gamma^*)>0 \iff b'' > \frac{\rho(b')^2}{1+\rho b}$. One can show that

$$b'' = \frac{2\beta n(1-\alpha)}{(1-\alpha\beta n\gamma^*)^3} \tag{C.3}$$

Then,

$$\begin{aligned} b'' - \frac{\rho(b')^2}{1+\rho b} &= \frac{2\beta n(1-\alpha)}{(1-\alpha\beta n\gamma^*)^3} - \frac{\rho(\beta n\gamma^*(1-\alpha))^2(2-\alpha\beta n\gamma^*)^2}{(1-\alpha\beta n\gamma^*)^3(1-\alpha\beta n\gamma^*+\beta n\rho(1-\alpha)(\gamma^*)^2)} \\ &= \underbrace{\frac{\beta n\gamma^*(1-\alpha)}{(1-\alpha\beta n\gamma^*)^3(1-\alpha\beta n\gamma^*+\beta n\rho(1-\alpha)(\gamma^*)^2)}}_{(+)} \\ &\quad\times \underbrace{\big(2(1-\alpha\beta n\gamma^*)+2(\beta n\rho(1-\alpha)(\gamma^*)^2)-\rho\beta n(\gamma^*)^2(1-\alpha)(2-\alpha\beta n\gamma^*)^2\big)}_{G(\gamma)} \end{aligned} \tag{C.4}$$

Hence, the sign of $F''(\gamma=\gamma^*)$ is determined by the sign of $G(\gamma)$, which is always positive.

$$\begin{aligned} G(\gamma) &= 2(1-\alpha\beta n\gamma^*) + \beta n\rho(1-\alpha)(\gamma^*)^2(2-4+4\alpha\beta n\gamma^* - (\alpha\beta n\gamma^*)) \\ &= 2(1-\alpha\beta n\gamma^*)(1-\beta n\rho(1-\alpha)(\gamma^*)^2) + \alpha\beta^2n^2\rho(1-\alpha)(\gamma^*)^3)(2-\alpha\beta n\gamma^*) \geq 0 \end{aligned} \tag{C.5}$$

C.2 Corollary 1

Consider $F(\gamma)$, with $F(\gamma)$ as defined in (C.1). Since the CE equilibrium is unique, γ^*, following a change in price stickiness or myopia, will change in the same direction as $F(\gamma)$. Taking the first-order partial derivative with respect to α of $F(\gamma)$ yields

$$\frac{\partial F(\gamma)}{\partial\alpha} = \frac{1-\rho^2}{(1+\rho b)^2}\underbrace{\frac{\partial b}{\partial\alpha}}_{(-)} < 0 \tag{C.6}$$

Similarly, taking the first-order partial derivative with respect to n of $F(\gamma)$ yields

$$\frac{\partial F(\gamma)}{\partial n} = \frac{1-\rho^2}{(1+\rho b)^2} \underbrace{\frac{\partial b}{\partial n}}_{(+)} < 0 \tag{C.7}$$

C.3 Proposition 2

The actual law of motion for inflation along the CE equilibrium is $\hat{\pi}_t = a\hat{mc}_t + b\hat{\pi}_{t-1}$, and the forecast about next period's inflation along the equilibrium path is $\widetilde{\mathbb{E}}_t\hat{\pi}_{t+1} = n(\gamma^*)^2\hat{\pi}_{t-1}$. Hence, the h-period-ahead forecasting error about inflation in period $(t+k)$, following a one-time shock ε_t in period t, is

$$\begin{aligned}
\hat{\pi}_{t+k} - \widetilde{\mathbb{E}}_{t+k-h}\hat{\pi}_{t+k} &= a\hat{mc}_{t+k} + b\hat{\pi}_{t+k-1} - n(\gamma^*)^2\hat{\pi}_{t+k-h-1} \\
&= a\rho^k\varepsilon_t + ab(\rho^{k-1} + b\rho^{k-2} + ... + b^{k-1})\varepsilon_t - an^h(\gamma^*)^{h+1}(\rho^{k-h-1} + ... + b^{k-h-1})\varepsilon_t \\
&= a\rho^{k-h-1}\left(\rho^{h+1} + b\rho^h\left(1 + ... + \left(\frac{b}{\rho}\right)^h + ... + \left(\frac{b}{\rho}\right)^{k-1}\right)\right) \\
&- a\rho^{k-h-1}\left(n^h(\gamma^*)^{h+1}\left(1 + ... + \left(\frac{b}{\rho}\right)^{k-h-1}\right)\right)\varepsilon_t \\
&= a\rho^{k-h-1}\left(\left(b^{h+1} - n^h(\gamma^*)^{h+1}\right)\sum_{j=0}^{k-h-1}\left(\frac{b}{\rho}\right)^j + \rho\left(\rho^h + b\rho\frac{\rho^h - b^h}{\rho - b}\right)\right)\varepsilon_t
\end{aligned} \tag{C.8}$$

The effect of $\varepsilon_t > 0$ on the forecasting error for $k = 0$ is positive; hence, forecasters under-react on impact. Moreover, $\lim_{k\to\infty}\rho^{k-h-1} = 0$, and therefore the forecasting error will eventually dissipate at some point in the future. The question remains whether, as $k \to \infty$, we approach the 0 forecasting errors from below (delayed over-shooting) or above. Given that $a > 0$ and $\rho > 0$, delayed over-shooting is guaranteed to occur if

$$\lim_{k\to\infty}\left(\left(b^{h+1} - n^h(\gamma^*)^{h+1}\right)\sum_{j=0}^{k-h-1}\left(\frac{b}{\rho}\right)^j + \rho\left(\rho^h + b\rho\frac{\rho^h - b^h}{\rho - b}\right)\right) < 0 \tag{C.9}$$

One can easily show that $(b^{h+1} - n^h(\gamma^*)^{h+1}) < 0$. Then, if $b > \rho$, we have that

$$\lim_{k\to\infty}\left(b^{h+1} - n^h(\gamma^*)^{h+1}\right)\sum_{j=0}^{k-h-1}\left(\frac{b}{\rho}\right)^j = -\infty$$

so

$$\lim_{k\to\infty}\left(\left(b^{h+1}-n^h(\gamma^*)^{h+1}\right)\sum_{j=0}^{k-h-1}\left(\frac{b}{\rho}\right)^j+\rho\left(\rho^h+b\rho\frac{\rho^h-b^h}{\rho-b}\right)\right)=-\infty \tag{C.10}$$

On the other hand, if $b<\rho$, we have that $\lim_{k\to\infty}\left(b^{h+1}-n^h(\gamma^*)^{h+1}\right)\sum_{j=0}^{k-h-1}\left(\frac{b}{\rho}\right)^j=\frac{\rho(b^{h+1}-n^h(\gamma^*)^{h+1})}{\rho-b}$, so

$$\lim_{k\to\infty}\left(\left(b^{h+1}-n^h(\gamma^*)^{h+1}\right)\sum_{j=0}^{k-h-1}\left(\frac{b}{\rho}\right)^j+\rho\left(\rho^h+b\rho\frac{\rho^h-b^h}{\rho-b}\right)\right)=\frac{\rho(\rho^{h+1}-n^h(\gamma^*)^{h+1})}{\rho-b} \tag{C.11}$$

Hence, when $b<\rho$, delayed over-shooting is guaranteed to exist if $\rho^{h+1}<n^h(\gamma^*)^{h+1}$. Mis-specified, to show that the two conditions stated above are sufficient for late over-response, we have to show that if the forecast error response turns negative, it will never become positive. Showing this proves that if the forecast error impulse response approaches 0 from above in the limit as $k\to\infty$, it has never been negative before. Suppose there exists $k^*\geq 1$, such that for $k\geq k^*$,

$$\mathbb{I}_{k,h}=\frac{\partial(\hat{\pi}_{t+k}-\widetilde{\mathbb{E}}_{t+k-h}\hat{\pi}_{t+k})}{\partial\varepsilon_t}=a\rho^{k-2}\left(\left(b^2-n(\gamma^*)^2\right)\sum_{j=0}^{k-h-1}\left(\frac{b}{\rho}\right)^j+\rho(b+\rho)\right)<0 \tag{C.12}$$

Since $(b^{h+1}-n^h(\gamma^*)^{h+1})<0$, as k increases the impulse response of forecast errors becomes more negative, and the sign of $\mathbb{I}_{k,h}$ can never flip as k increases.

C.4 Proposition 3

I first derive a number of important moments. Consider first the covariance between $\hat{\pi}_{t+h}$ and $\hat{\pi}_t$ for any $h>0$:

$$Cov(h)=\mathbb{E}(\hat{\pi}_{t+h}\hat{\pi}_t)=a^2\left(\frac{\rho(\rho^h-b^h)}{(\rho-b)(1-\rho b)}+\frac{b^h(1+\rho b)}{(1-b^2)(1-\rho b)}\right)\mathbb{E}(\hat{mc}_t^2)$$

Next, I derive the covariance between the ex-post forecast errors, $FE_{t,t+h}$ and ex-ante forecast revisions, $FR_{t,t+h}$:

$$
\begin{aligned}
\mathbb{E}(FE_{t,t+h}FR_{t,t+h}) &= n^h(\gamma^*)^{h+1}\left[a\rho^{h+1}(1-n\gamma^*)\mathbb{E}(\hat{mc}_t\hat{\pi}_t) + b(Cov(h) - n\gamma^* Cov(h+1)) - n^h(\gamma^*)^{h+1}(\mathbb{E}(\hat{\pi}_t^2) - \right. \\
&= \frac{a^2 n^h(\gamma^*)^{h+1}\rho^{h+1}(1-n\rho\gamma^*)}{1-\rho b}\mathbb{E}(\hat{mc}_t^2) \\
&+ \frac{a^2 n^h(\gamma^*)^{h+1}}{1-\rho b}\frac{b^{h+1}(1+\rho b)(1-nb\gamma^*) - n^h(\gamma^*)^{h+1}(1+\rho b - n\gamma^*(\rho+b))}{1-b^2}\mathbb{E}(\hat{mc}_t^2) \\
&+ \frac{a^2 n^h(\gamma^*)^{h+1}}{1-\rho b}\left[b\rho^h(1-n\rho\gamma^*)\sum_{j=0}^{h-1}\left(\frac{b}{\rho}\right)^j - n\rho\gamma^* b^{h+1}\right]\mathbb{E}(\hat{mc}_t^2)
\end{aligned}
\tag{C.13}
$$

On the other hand, one can show that the variance of forecast errors is given by

$$
\mathbb{E}(FR_{t,t+h}^2) = a^2 n^{2h}(\gamma^*)^{2(h+1)}\frac{(1+n^2(\gamma^*)^2 - 2n(\gamma^*)^2)(1+\rho b)}{(1-\rho b)(1-b^2)}\mathbb{E}(\hat{mc}_t^2)
$$

Finally, K_h is given by the covariance between forecast errors and forecast revisions and divided by the variance of forecast revisions, that is,

$$
\begin{aligned}
K_h = &\underbrace{\frac{\rho^h(1-b^2)(1-n\rho\gamma^*)\left(\rho + b\sum_{j=0}^{h-1}\left(\frac{b}{\rho}\right)^j\right) + b^{h+1}(\rho(b-n\gamma^*) + 1 - nb\gamma^*)}{n^h(\gamma^*)^{h+1}(1+n^2(\gamma^*)^2 - 2n(\gamma^*)^2)(1+\rho b)}}_{(+)} \\
&\underbrace{-\frac{\rho(b-n\gamma^*) + 1 - nb\gamma^*}{(1+n^2(\gamma^*)^2 - 2n(\gamma^*)^2)(1+\rho b)}}_{(-)}
\end{aligned}
\tag{C.14}
$$

Now, I compute the covariance between forecast errors and inflation realized in period t,

$$
\begin{aligned}
\mathbb{E}(FE_{t,t+h}\hat{\pi}_t) &= \frac{a^2\rho^h}{1-\rho b}\mathbb{E}(\hat{mc}_t^2) + bCov(h-1) - n^h(\gamma^*)^{h+1}Cov(1) \\
&= a^2\left[\frac{\rho(\rho^h - b^h)}{(\rho-b)(1-\rho b)} + \frac{b^h(1+\rho b)}{(1-\rho b)(1-b^2)} - \frac{n^h(\gamma^*)^{h+1}(b+\rho)}{(1-\rho b)(1-b^2)}\right]\mathbb{E}(\hat{mc}_t^2)
\end{aligned}
\tag{C.15}
$$

Dividing the expression above by the variance of inflation, I derive M_h:

$$
M_h = \underbrace{\frac{\rho^h}{1+\rho b}\left[\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j - b^2\sum_{j=0}^{h-2}\left(\frac{b}{\rho}\right)^j\right]}_{(+)}\underbrace{-\frac{n^h(\gamma^*)^{h+1}(b+\rho)}{1+\rho b}}_{(-)}
\tag{C.16}
$$

C.5 Corollary 3

First, from Proposition 2, it is trivial to see that delayed over-shooting is guaranteed to occur for any parameterization of the model.

Second, I re-write the condition for which $K_h > 0$ as follows

$$K_h = \rho(1-b^2)(1-\rho\gamma^*)\frac{\rho^h - b^h}{\rho - b} + (b^{h+1} - (\gamma^*)^{h+1})(1+\rho b - \gamma^*(b+\rho)) > 0$$

Simple re-arrangement gives rise to the following inequality,

$$\begin{aligned}(\gamma^*)^{h+1} &< b^{h+1} + \rho(1-b^2)(1-\rho\gamma^*)\frac{\rho^{h+1}-b^{h+1}}{(\rho-b)(1+\rho b-\gamma^*(b+\rho))} \\ &= b^{h+1} + \rho(1-b^2)(1-\rho^2) + \frac{\rho^{h+1}-b^{h+1}}{(\rho-b)(1-\rho^2)(1-b^2)} \\ &= \rho^{h+1} + \rho^h b + \ldots + \rho b^h + b^{h+1} = \bar{\gamma}\end{aligned} \tag{C.17}$$

where for the second equality, I rely on the fact that along the CE equilibrium, $\gamma^* = \frac{b+\rho}{1+\rho b}$, as shown in Proposition 1.

Third, I re-write the condition for which $M_h < ol.0$ as follows

$$M_h = \frac{\rho^h}{1+\rho b}\left[\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j - b^2\sum_{j=0}^{h-2}\left(\frac{b}{\rho}\right)^j\right] - \frac{(\gamma^*)^{h+1}(b+\rho)}{1+\rho b} < 0$$

from which it follows that

$$(\gamma^*)^{h+1} > \frac{\rho^{h+1} - b^{h+1} - \rho^2 b^2(\rho^{h-1} - b^{h-1})}{\rho^2 - b^2} = \underline{\gamma}$$

C.6 Proposition 4

As shown in the main text, when myopia is combined with well-specified forecasting rules, the aggregated optimal pricing rule can be written as $\hat{\pi}_t = \kappa \hat{mc}_t + \beta n \mathbb{E}_t \hat{\pi}_{t+1}$. The solution for inflation then is given by $\hat{\pi}_t = a_0 \hat{mc}_t$, where $a_0 = \frac{\kappa}{1-\beta n \rho}$.

1. I show that $\mathbb{I}_{k,h} \geq 0$:

$$\mathbb{I}_{k,h} = a_0 \hat{mc}_{t+k} - a_0 \rho^h \hat{mc}_{t+k-h} = a_0(\rho^k - n\rho^k)\varepsilon_t \geq 0 \tag{C.18}$$

for any $k \geq 0$. From here, it follows that if $n = 1$, i.e., if we impose well-specified forecasts absent myopia (FIRE), $\mathbb{I}_{k,h} = 0$ for any $k \geq 0$.

2. Next, I compute the covariance between forecast errors and forecast revisions:

$$\mathbb{E}\left((\hat{\pi}_{t+h} - n^h \mathbb{E}_t \hat{\pi}_{t+h})(\mathbb{E}_t \hat{\pi}_{t+h} - \mathbb{E}_{t-1}\hat{\pi}_{t+h})\right) = n^h \rho^{2h}(1 - n\rho^2)(1 - n^h)\mathbb{E}(\hat{\pi}_t^2)$$

Dividing the expression above by the variance of forecast revisions delivers

$$K_h = \frac{(1 - n^h)(1 - n\rho^2)}{n^h(1 + n^2\rho^2 - 2n\rho^2)} \geq 0$$

3. Finally, I compute the covariance between forecast errors and current inflation:

$$\mathbb{E}\left((\hat{\pi}_{t+h} - n^h \mathbb{E}_t \hat{\pi}_{t+h})\hat{\pi}_t\right) = \rho^h(1 - n^h)\mathbb{E}(\hat{\pi}_t^2)$$

Dividing the expression above by the variance of current inflation gives rise to

$$M_h = \rho^h(1 - n^h) \geq 0$$

D Data and Additional Results

D.1 Data

I use quarterly data from 1966 to 2018. All data are extracted from FRED and described as follows

$$y_t = 100ln\left(\frac{GDPC1_t}{POP_{index,t}}\right)$$

$$y_t^{potential} = 100ln\left(\frac{GDPPOT_t}{POP_{index,t}}\right)$$

$$x_t^{obs} = y_t - y_t^{potential}$$

$$\pi_t^{obs} = 100ln\left(\frac{GDPDEF_t}{GDPDEF_{t-1}}\right)$$

$$R_t^{obs} = \frac{Funds_t}{4}$$

where

- $GDPC1$ – Real GDP, Billions of Chained 2012 Dollars, Seasonally Adjusted Annual Rate.

- $POP_{index} = \frac{CNP160V}{CNP160V_{1992Q3}}$.
- $CNP160V$ – Civilian non-institutional population, thousands, 16 years and above.
- $GDPPOT$ – Real potential GDP, Billions of Chained 2012 Dollars, as reported by the US Congressional Budget Office.
- $GDPDEF$ – GDP-Implicit Price Deflator, 2012 = 100, Seasonally Adjusted.
- $Funds$ – Federal funds rate, daily figure averages in percentages.

D.2 Additional Results on Bayesian Estimation

Figures 12 and 13 show the evolution of two crucial convergence statistics in the cases when myopia is combined with AR(1) forecasting rules and VAR(1) forecasts, respectively. Convergence is assessed based on the Brooks and Gelman (1998) methodology. To compute $\hat{B}$, I first estimate the evolution of the mean across draws for each parameter for each one of the two chains of the Metropolis-Hastings. Then, $\hat{B}$ equals the variance of the two means over time. On the other hand, to compute $\hat{W}$, I first estimate the evolution of the draws variance for each parameter for each one of the two chains of the Metropolis-Hastings. Then, $\hat{W}$ equals the mean of the two computed variance values. Convergence is achieved when the evolution of $\hat{W}$ and $\hat{W} + \hat{B}$ converge to one another and remain stable.

When the expectations formation process is characterized by a combination of AR(1) forecasts with myopia, convergence is achieved for all paramaters. In the case of VAR(1) forecasts combined with myopia, convergence is achieved for most of the parameters: $\hat{W}$ and $\hat{W} + \hat{B}$ exhibit some divergence for σ_e.

D.2.1 Robustness

Table 6 shows characteristics of the posterior distribution when the degree of myopia is assumed to have a uniform prior with mean 0.5 and standard deviation $1/\sqrt{12}$.

The estimated posterior distributions for all three specifications exhibit similar characteristics relative to the ones presented in the main text. In terms of model fit, I set the model with AR(1) forecasts and myopia to be the benchmark. The Bayes factor analysis shows i) positive evidence in favor of the benchmark compared to the well-specified forecasts case; and ii) strong evidence in favor of a combination of VAR(1) forecasts relative to the benchmark. Overall, one of the main

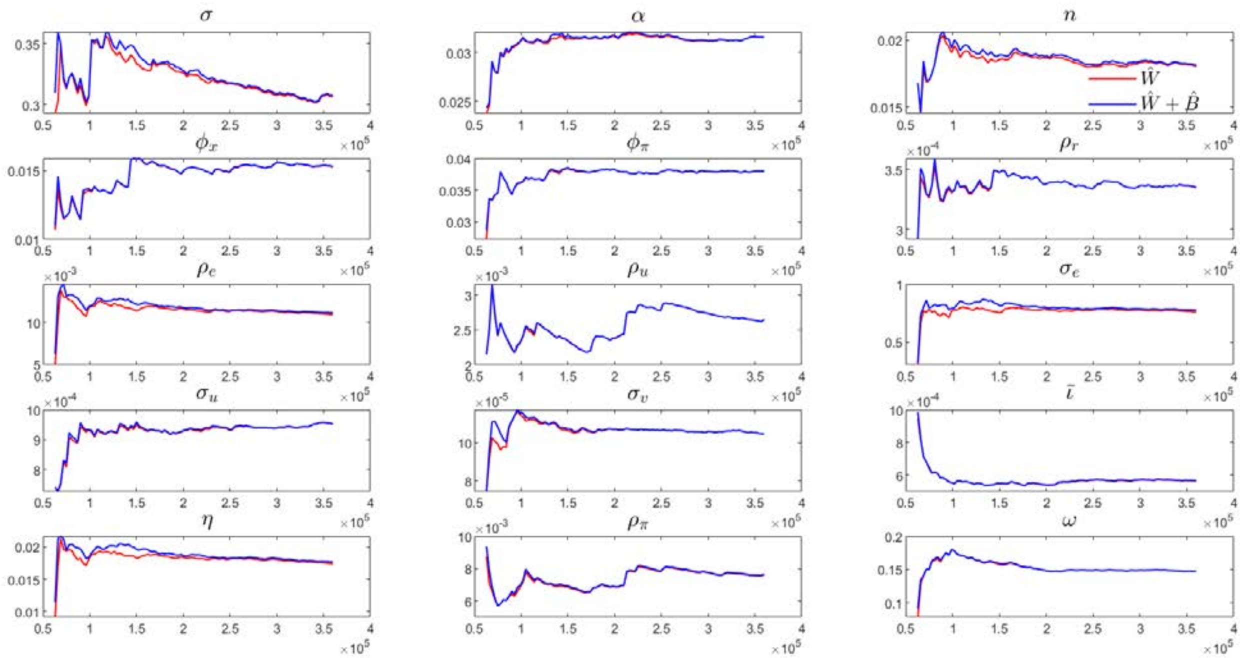


Figure 12: Convergence statistics when the expectations formation process is characterized by a combination of AR(1) mis-specified forecasts and myopia.

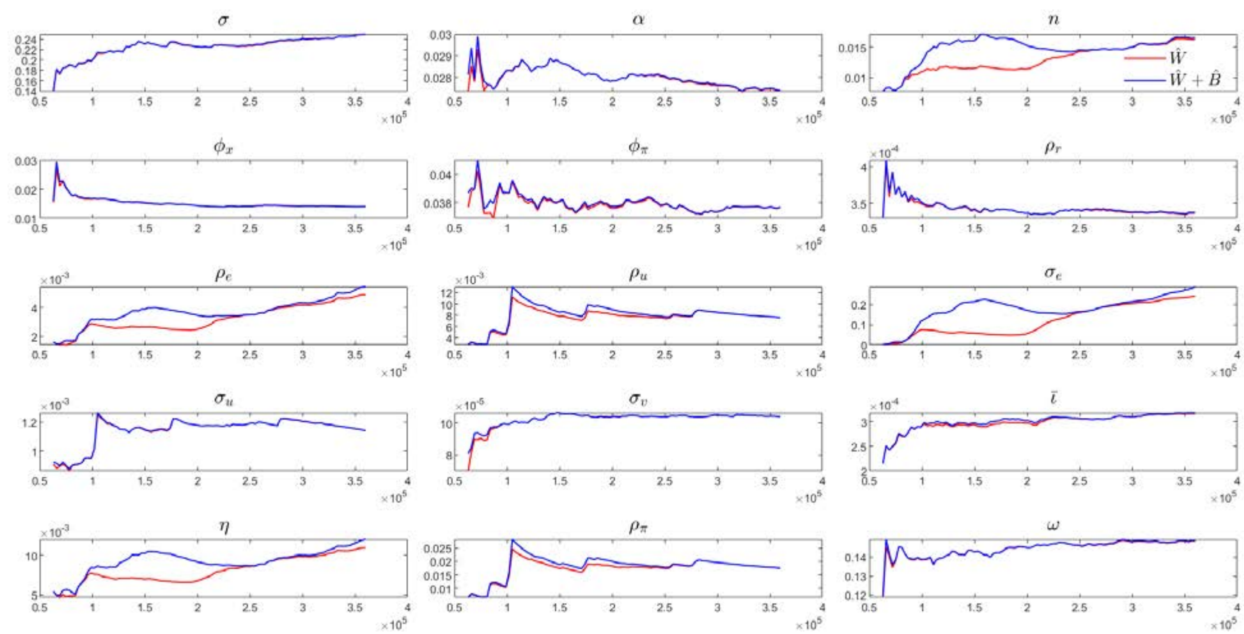


Figure 13: Convergence statistics when the expectations formation process is characterized by a combination of VAR(1) mis-specified forecasts and myopia.

results of the paper - that autoregressive mis-specified forecasts are preferred over well-specified forecasting rules - holds true.

	(1) Well-specified			(2) AR(1)			(3) VAR(1)		
Parameters	mean	5%	95%	mean	5%	95%	mean	5%	95%
Calvo parameter, α	-	-	-	0.44	0.15	0.74	0.37	0.12	0.67
Degree of myopia, n	0.26	0.00	0.49	0.72	0.51	0.87	0.73	0.51	0.88
Inverse IES coefficient, σ	1.99	1.15	2.82	2.41	1.57	3.35	2.23	1.47	3.08
Habit in consumption, η	0.73	0.43	0.92	0.43	0.30	0.56	0.40	0.27	0.54
Inflation indexation, ρ_π	0.87	0.81	0.92	0.76	0.54	0.91	0.77	0.55	0.92
Elasticity mc, ω	0.94	0.30	1.58	0.98	0.33	1.62	0.94	0.31	1.58
Feedback to output gap, ϕ_x	0.42	0.23	0.59	0.43	0.26	0.63	0.42	0.25	0.66
Feedback to inflation, ϕ_π	1.45	1.13	1.77	1.46	1.13	1.79	1.45	1.13	1.78
Interest rate smoothing, ρ_r	0.92	0.89	0.95	0.92	0.89	0.95	0.92	0.88	0.95
Demand shock autocorr., ρ_e	0.56	0.33	0.88	0.80	0.72	0.88	0.83	0.74	0.90
Supply shock autocorr., ρ_u	0.05	0.01	0.10	0.10	0.02	0.22	0.11	0.03	0.26
Demand shock std., σ_e	5.17	1.77	8.53	1.48	1.10	1.76	1.21	0.75	1.62
Supply shock std., σ_u	0.26	0.24	0.28	0.33	0.28	0.38	0.34	0.28	0.39
Monetary shock std., σ_v	0.21	0.19	0.22	0.21	0.19	0.23	0.21	0.19	0.23
Learning gain, $\bar{\iota}$	-	-	-	0.08	0.04	0.12	0.07	0.04	0.10
Log marg. data dens.									
Modified Harmonic Mean	-259.195			-256.751			-253.100*		
Bayes factor	($e^{-2.44}$)			(e^{0})			($e^{3.65}$)		

Table 6: Posterior distribution of the model for various assumptions on the expectations formation process with myopia. The prior for myopia is uniform with mean 0.5 and standard deviation $1/\sqrt{12}$. Values in parentheses denote the Bayes factor of the model relative to the benchmark specification with mis-specified forecasts and myopia. The asterisk denotes strong evidence in favor of the model relative to the benchmark one.

D.3 Additional Results on Forecast Error Behavior

In this subsection, I repeat the analysis of Section 5 when the expectations formation process is characterized by mis-specified forecasts of the AR(1) or VAR(1) type and *no* myopia.

In particular, Figure 14 plots the three-dimensional IRFs of annual inflation forecasting errors to a demand, a cost-push, and a monetary shock over the estimation period from 1966:Q1 through 2018:Q3. The top three panels exhibit the response of forecast errors in the case of AR(1) forecasts, and the bottom three panels show the responses in the case of VAR(1) forecasts. In all six subplots, inflation forecast errors exhibit delayed over-shooting.

Finally, I estimate the regressions in (15) and (16) with simulated inflation forecasting data,

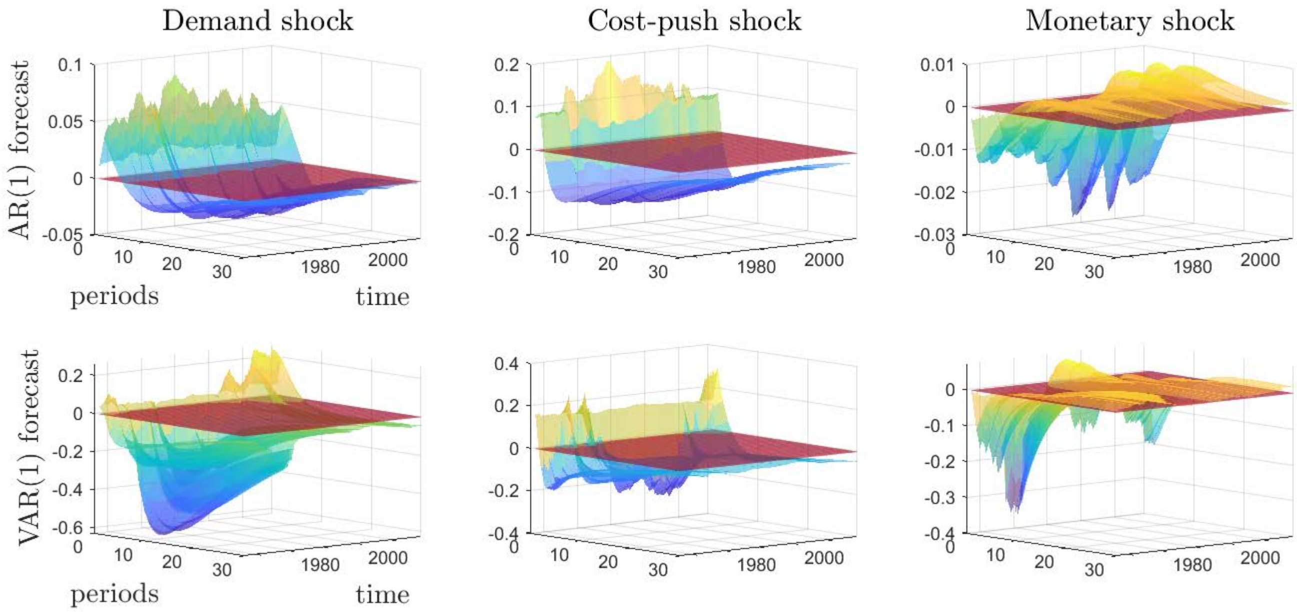


Figure 14: Three-dimensional impulse response functions of annual inflation forecasting errors to a one standard deviation positive demand, cost-push, and monetary shock in the model with mis-specified forecasts and no myopia. Top three panels: AR(1) mis-specified forecasts; bottom three panels: VAR(1) mis-specified forecasts. Model parameters are set equal to the posterior mean as shown in Table 3 and 4. In dark red: (response, time periods) plane.

when the expectations formation process is characterized by mis-specified forecasting rules (of the AR(1) or VAR(1) type) only. Table 7 presents the estimates of coefficients K_4 and M_4: panel A shows both estimates over the full sample from 1967:Q1 through 2018:Q3, whereas panel B exhibits estimates of M_4 starting from 1981:Q3 through 2018:Q3.

Estimates in Table 7 show that annual inflation forecast errors depend positively on ex-ante inflation forecast revisions and negatively on current inflation realizations, regardless of how the model is parameterized and of the structure of the mis-specified forecasting rule. On the other hand, when the sample starts from the early 1980s, the sign of the correlation between ex-post forecast errors and current inflation realizations is always negative for both AR(1) and VAR(1) forecasts, regardless of how the model is parameterized. This is different from the case when mis-specified forecasts are combined with some myopia, where the sign of the correlation between ex-post forecast errors and current realizations depends on the model parameterization.

	AR(1) and no myopia			*VAR(1) and no myopia*		
Panel A: 1967:Q1 - 2018:Q3						
	5%	mean	95%	5%	mean	95%
Revision	0.54 (0.21)	0.68 (0.22)	0.77 (0.23)	0.48 (0.20)	0.51 (0.19)	0.65 (0.20)
Current	0.19 (0.04)	0.11 (0.04)	0.07 (0.04)	0.37 (0.04)	0.25 (0.04)	0.19 (0.04)
Panel B: 1981:Q3 - 2018:Q3						
	5%	mean	95%	5%	mean	95%
Current	-0.36 (0.05)	-0.32 (0.06)	-0.44 (0.06)	-0.25 (0.06)	-0.35 (0.05)	-0.43 (0.06)

Table 7: Estimates of regressions in (15) and (16) on simulated forecasting data. Model parameters are set at the posterior 5^{th} percentile, the mean, and the 95^{th} percentile of the distribution for the model with mis-specified forecasts and no myopia as documented in column (3) of Tables 3 and 4. All regressions include a constant term, and standard errors are given in parenthesis. Panel A: estimates of K_4 (revision) and M_4 (current) over the full sample for AR(1) and VAR(1) mis-specified forecasts. Panel B: estimates of M_4 (current) since the early 1980s for AR(1) and VAR(1) mis-specified forecasts.

References

Adam, Klaus. 2007. "Experimental Evidence of the Persistence of Output and Inflation." *The Economic Journal* 117 (520):603–636. URL https://doi.org/10.1111/j.1468-0297.2007.02043.x.

Airaudo, Marco and Ina Hajdini. 2021. "Consistent Expectations Equilibria in Markov Regime Switching Models and Inflation Dynamics." *International Economic Review* 62 (4):1401–1430. URL https://doi.org/10.1111/iere.12529.

Angeletos, George-Marios and Zhen Huo. 2021. "Myopia and Anchoring." *American Economic Review* 111 (4):1166–1200. URL https://doi.org/10.1257/aer.20191436.

Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry. 2021. "Imperfect Macroeconomic Expectations: Evidence and Theory." *NBER Macroeconomics Annual* 35:1–86. URL https://doi.org/10.1086/712313.

Atkeson, Andrew and Lee E. Ohanian. 2001. "Are Phillips Curves Useful for Forecasting Inflation?" *Quarterly Review* 25 (Win):2–11. URL https://ideas.repec.org/a/fip/fedmqr/y2001iwinp2-11nv.25no.1.html.

Bianchi, Francesco. 2013. "Regime Switches, Agents' Beliefs, and Post-World War II US Macroeconomic Dynamics." *The Review of Economic Studies* 80 (2):463–490. URL https://doi.org/10.1093/restud/rds032.

Bianchi, Francesco, Cosmin Ilut, and Hikaru Saijo. 2022. "Diagnostic Business Cycles." Working paper 28604, National Bureau of Economic Research. URL https://doi.org/10.3386/w28604.

Bils, Mark and Peter Klenow. 2004. "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy* 112 (5):947–985. URL https://doi.org/10.3386/w28604.

Branch, William, Bruce McGough, and Mei Zhu. 2022. "Statistical Sunspots." *Theoretical Economics* 17 (1):291–329. URL https://doi.org/10.3982/TE3752.

Brooks, Stephen P. and Andrew Gelman. 1998. "General Methods for Monitoring Convergence of Iterative Simulations." *Journal of Computational and Graphical Statistics* 7 (4):434–455. URL https://doi.org/10.1080/10618600.1998.10474787.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1):1–45. URL https://doi.org/10.1086/426038.

Coibion, Olivier and Yuriy Gorodnichenko. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105 (8):2644–2678. URL https://doi.org/10.1257/aer.20110306.

Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar. 2018. "The Formation of Expectations, Inflation, and the Phillips Curve." *Journal of Economic Literature* 56 (4):1447–1491. URL https://doi.org/10.1257/jel.20171300.

Cole, Stephen J. and Fabio Milani. 2019. "The Misspecification of Expectations in New Keynesian Models: A DSGE-VAR Approach." *Macroeconomic Dynamics* 23 (3):974–1007. URL https://doi.org/10.1017/S1365100517000104.

Del Negro, Marco and Stefano Eusepi. 2011. "Fitting Observed Inflation Expectations." *Journal of Economic Dynamics and Control* 35 (12):2105–2131. URL https://doi.org/10.1016/j.jedc.2011.04.005.

Eusepi, Stefano and Bruce Preston. 2018. "The Science of Monetary Policy: An Imperfect Knowledge Perspective." *Journal of Economic Literature* 56 (1):3–59. URL https://doi.org/10.1257/jel.20160889.

Fuster, Andreas, Benjamin Hebert, and David Laibson. 2012. "Natural Expectations, Macroeconomic Dynamics and Asset Pricing." *NBER Macroeconomics Annual* 26:1–48. URL https://doi.org/10.1086/663989.

Fuster, Andreas, David Laibson, and Brock Mendel. 2010. "Natural Expectations and Macroeconomic Dynamics." *Journal of Economic Perspectives* 24 (4):67–84. URL https://doi.org/10.1257/jep.24.4.67.

Gabaix, Xavier. 2020. "A Behavioral New Keynesian Model." *American Economic Review* 110 (8):2271–2327. URL https://doi.org/10.1257/aer.20162005.

Galí, Jordi. 2008. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications*. Princeton University Press.

Ganong, Peter and Pascal Noel. 2019. "Consumer Spending during Unemployment: Positive and Normative Implications." *American Economic Review* 109 (7):2383–2424. URL https://doi.org/10.1257/aer.20170537.

Gaus, Eric and Christopher G. Gibbs. 2018. "Expectations and the Empirical Fit of DSGE Models." Tech. rep.

Geweke, John. 1999. "Simulation Methods for Model Criticism and Robustness Analysis." In *Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting*, vol. 6, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith. Oxford University Press, 275–299.

Giannoni, Marc and Michael Woodford. 2004. "Optimal Inflation-Targeting Rules." University of Chicago Press. URL https://www.nber.org/books-and-chapters/inflation-targeting-debate/optimal-inflation-targeting-rules.

Greenwood, Robin and Andrei Shleifer. 2014. "Expectations of Returns and Expected Returns." *The Review of Financial Studies* 27 (3):714–746. URL https://doi.org/10.1093/rfs/hht082.

Hajdini, Ina and André Kurmann. 2022. "Predictable Forecast Errors in Full-Information Rational Expectations Models with Regime Shifts." URL https://dx.doi.org/10.2139/ssrn.4124427.

Haubrich, Joseph G., George Pennacchi, and Peter Ritchken. 2008. "Estimating Real and Nominal Term Structures Using Treasury Yields, Inflation, Inflation Forecasts, and Inflation Swap Rates." Working Paper 08-10, Federal Reserve Bank of Cleveland. URL https://doi.org/10.26509/frbc-wp-200810.

Herbst, Edward P. and Frank Schorfheide. 2016. *Bayesian Estimation of DSGE Models*. Princeton University Press. URL https://doi.org/10.23943/princeton/9780691161082.001.0001.

Hommes, Cars H. 2013. *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*. Cambridge University Press.

Hommes, Cars H., Kostas Mavromatis, Tolga Ozden, and Mei Zhu. 2019. "Behavioral Learning Equilibria in the New Keynesian Model." Tech. Rep. 654, Netherlands Central Bank, Research Department.

Hommes, Cars H. and Gerhard Sorger. 1998. "Consistent Expectations Equilibria." *Macroeconomic Dynamics* 2 (3):287–321. URL https://doi.org/10.1017/S1365100598008013.

Hommes, Cars H. and Mei Zhu. 2014. "Behavioral Learning Equilibria." *Journal of Economic Theory* 150:778–814. URL https://doi.org/10.1016/j.jet.2013.09.002.

Ilabaca, Francesco, Greta Meggiorini, and Fabio Milani. 2020. "Bounded Rationality, Monetary Policy, and Macroeconomic Stability." *Economics Letters* 186 (108522). URL https://doi.org/10.1016/j.econlet.2019.07.007.

Ilut, Cosmin and Hikaru Saijo. 2021. "Learning, confidence, and business cycles." *Journal of Monetary Economics* 117:354–376. URL https://doi.org/10.1016/j.jmoneco.2020.01.010.

Kass, Robert E. and Adrian E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90 (430):773–795. URL https://doi.org/10.1080/01621459.1995.10476572.

Kohlhas, Alexandre N. and Ansgar Walther. 2021. "Asymetric Attention." *American Economic Review* 111 (9):2879–2925. URL https://doi.org/10.1257/aer.20191432.

Maćkowiak, Bartosz and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review* 99 (3):769–803. URL https://doi.org/10.1257/aer.99.3.769.

Malmendier, Ulrike and Stefan Nagel. 2016. "Learning from Inflation Experiences." *The Quarterly Journal of Economics* 131 (1):53–87. URL https://doi.org/10.1093/qje/qjv037.

Mankiw, Gregory N. and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *The Quarterly Journal of Economics* 117 (4):1295–1328. URL https://doi.org/10.1162/003355302320935034.

Milani, Fabio. 2006. "A Bayesian DSGE Model with Infinite-Horizon Learning: Do 'Mechanical' Sources of Persistence Become Superfluous?" *International Journal of Central Banking* 2 (3):87–106.

———. 2007. "Expectations, Learning and Macroeconomic Persistence." *Journal of Monetary Economics* 54 (7):2065–2082. URL https://doi.org/10.1016/j.jmoneco.2006.11.007.

Ormeño, Arturo and Krisztina Molnár. 2015. "Using Survey Data of Inflation Expectations in the Estimation of Learning and Rational Expectations Models." *Journal of Money, Credit and Banking* 47 (4):673–699. URL https://doi.org/10.1111/jmcb.12224.

Petersen, Luba. 2015. "Do Expectations and Decisions Respond to Monetary Policy?" *Journal of Economic Studies* 42 (6):972–1004. URL https://doi.org/10.1108/JES-08-2015-0149.

Preston, Bruce. 2005. "Learning about Monetary Policy Rules When Long-Horizon Expectations Matter." *International Journal of Central Banking* 1 (2):81–126.

Reis, Ricardo. 2006. "Inattentive Producers." *The Review of Economic Studies* 73 (3):793–821. URL https://doi.org/10.1111/j.1467-937X.2006.00396.x.

Rychalovska, Yuliya. 2016. "The Implications of Financial Frictions and Imperfect Knowledge in the Estimated DSGE Model of the U.S. Economy." *Journal of Economic Dynamics and Control* 73:259–282. URL https://doi.org/10.1016/j.jedc.2016.09.014.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3):665–690. URL https://doi.org/10.1016/S0304-3932(03)00029-1.

Slobodyan, Sergey and Raf Wouters. 2012a. "Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models." *American Economic Journal: Macroeconomics* 4 (2):65–101. URL https://doi.org/10.1257/mac.4.2.65.

———. 2012b. "Learning in an Estimated Medium-Scale DSGE Model." *Journal of Economic Dynamics and Control* 36 (1):26–46. URL https://doi.org/10.1016/j.jedc.2011.01.016.

Smets, Frank and Raf Wouters. 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* 1 (5):1123–1175. URL https://doi.org/10.1162/154247603770383415.

Smets, Frank and Rafael Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3):586–606. URL https://doi.org/10.1257/aer.97.3.586.

Stock, James H. and Mark W. Watson. 2007. "Why Has US Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking* 39 (s1):3–33. URL https://doi.org/10.1111/j.1538-4616.2007.00014.x.

Tversky, Amos and Daniel Kahneman. 1973. "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology* 5 (2):207–232. URL https://doi.org/10.1016/0010-0285(73)90033-9.

———. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* URL https://doi.org/10.1126/science.185.4157.1124.

Woodford, Michael. 2003a. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information, and Expectations in Modern Macroeconomics, in Honor of Edmund S. Phelps*, edited by Philippe Aghion, Roman Frydman, Joseph E. Stiglitz, and Michael Woodford. Princeton University Press, 25–58.

———. 2003b. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, N.J.: Princeton University Press.