**Federal Reserve Bank of Cleveland Working Paper Series**

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

Ina Hajdini

Working Paper No. 22-03

February 2022

**Suggested citation:** Hajdini, Ina. 2022. "Mis-specified Forecasts and Myopia in an Estimated New Keynesian Model." Working Paper No. 22-03. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202203.

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

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

Ina Hajdini†

February 15, 2022

**Abstract**

The paper considers a New Keynesian framework in which agents form expectations based on a combination of mis-specified forecasts and myopia. The proposed expectations formation process is found to be consistent with *all three empirical facts* on consensus inflation forecasts, namely, that forecasters under-react to ex-ante forecast revisions, that forecasters over-react to recent events, and that the response of forecast errors to a shock initially under-shoots but then over-shoots. 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) The data strongly prefer the combination of autoregressive mis-specified forecasting rules and myopia over other alternatives; (ii) The best fitting expectations formation process for both households and firms is characterized by high degrees of myopia and simple AR(1) forecasting rules; (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; E13; E30; E50; E52; 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 with 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 forecasting data evidence from the US Survey of Professional Forecasters (SPF). 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 of the paper is that both the analysis of the forecasting data and the likelihood-based Bayesian estimation of the full New Keynesian model prefer the specification in which economic agents rely on a combination of autoregressive mis-specified forecasting rules and myopia over other alternatives. Importantly, I prove that such a combination 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 et al. (2021)).[3]

[1] See for instance, Tversky and Kahneman (1973, 1974), Adam (2007), Hommes (2013); Hommes et al. (2019), Petersen (2015), Malmendier and Nagel (2016), 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 et al. (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] The three papers use various sources of forecasting data to validate the three facts. However, the US SPF data

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 Rational Expectations. 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 toward the steady-state of the economy. Consecutively, expectations about aggregate variables many periods ahead shrink toward steady-state values. 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 the myopic adjustment is made to *well-specified* forecasting rules, in the present paper such an adjustment is made 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, i.e. 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 - which 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 rules,

are their common denominator. In particular, Angeletos et al. (2021) show delayed over-shooting using 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.

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

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 (respectively, higher) degree of myopia or, equivalently, more (respectively, less) forward-lookingness in firms' decisions induces higher (respectively, lower) 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 et al. (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 *endogenous* over-extrapolation, i.e., if the equilibrium persistence of inflation sufficiently exceeds that of the marginal cost. Similarly, Angeletos et al. (2021) show that there should be over-extrapolation and sufficiently noisy information for delayed over-shooting to be replicated theoretically. In contrast to their paper where over- or under-extrapolation is an *exogenous* feature, in the present paper over-extrapolation is an *endogenous outcome* of the CE equilibrium solution.[7] Second, I prove that the proposed expectations formation process in the present paper reflects the empirical fact that ex-post forecasting errors are positively predicted by ex-ante forecast revisions, as found in Coibion and Gorodnichenko (2015). The presence of myopia generally facilitates matching such evidence by slowing down the update of forecasts as new information becomes available to forecasters. Moreover, consistent with the findings in Coibion and Gorodnichenko (2015), myopia also ensures that inflation lags have no predictive power over inflation forecasting errors, once one controls for ex-ante forecast revisions.[8] Third, I prove that the expectations formation process reflects the empirical fact brought forward by Kohlhas and Walther (2021) that forecasters over-react to information at the time of forecast. Mis-specified forecasting rules give rise to over-extrapolation, regardless of the degree of myopia, and thus guarantee that inflation at the time of forecast negatively predicts inflation forecasting

[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 myopia as in Gabaix (2020) or FIRE do not match late over-shooting.

[8] Further, I show that mis-specified forecasts *absent* myopia can mirror such a fact *if and only if* the equilibrium persistence of inflation coincides with the inertia of the marginal cost, which in turn can only happen if the marginal cost has no inertia. Importantly, even if the equilibrium inflation persistence equals that of the marginal cost, the condition for delayed over-shooting will not hold. On the contrary, the presence of myopia relaxes the constraint and allows for ex-ante revisions to be the only significant predictor of ex-post forecasting errors without violating the condition for delayed over-shooting.

errors.

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, consistent with forecasting data evidence, macroeconomic data strongly prefer the model whose expectations formation process is a combination of autoregressive forecasts and myopia over the other aforementioned alternatives.[10] Second, the best-fitting expectations formation process is characterized by significant degrees of myopia and AR(1) forecast rules. More elaborate VAR-based forecasts do not enhance the model's fit and do not provide additional information in terms of forecasting. Third, compared to FIRE, frictions such as habit in consumption are significantly less important in the presence of myopia combined with AR(1) forecasts, because such a combination strengthens the internal propagative features of the model: autoregressive forecasts can induce excess persistence, whereas myopia can induce excess volatility. Therefore, the proposed expectations formation process will often deliver impulse response functions to demand, cost-push, and monetary shocks that are more persistent and volatile (especially for inflation), relative to a case of no myopia or FIRE.

Finally, setting the model's parameters to their estimated posterior mean, I simulate inflation annual forecasting data according to 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 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*

[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 when there is no myopia.

[10]The second best-fitting model is the one with well-specified forecasts and myopia.

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).[11] 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 et al. (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. Nevertheless, as mentioned in Angeletos et al. (2021), asymmetric attention is not reflective of forecast errors' delayed over-shooting. Angeletos et al. (2021), on the other hand, show that noisy information combined with *exogenous* over-extrapolation can match the list of aforementioned empirical facts on forecasting data. The present paper shows that myopia combined with *endogenous* over-extrapolation matches all three empirical facts on consensus forecasting data.[12]

The paper contributes with additional evidence to a rich body of literature that validates usage of simple forecasting processes by the private sector (e.g., Tversky and Kahneman (1973, 1974), Adam (2007), Hommes (2013); Hommes et al. (2019), Greenwood and Shleifer (2014), Petersen (2015), and Malmendier and Nagel (2016)).[13] The paper also relates to a series of papers that discuss the analytical implications of mis-specified forecasting rules, as in Hommes and Sorger (1998), Fuster et al. (2010, 2012), and Hommes and Zhu (2014), among others. In particular, the paper relies on the solution concept of a first-order Consistent Expectations equilibrium, developed by Hommes and Sorger (1998) and Hommes and Zhu (2014).

The paper shares a common idea with Gabaix (2020) about myopia being excess discounting of future deviations from the steady-state. However, differently from the present paper, in Gabaix (2020) forecasts are based on structurally well-specified rules, and as mentioned earlier, the forecasting data evidence presented in this paper stands in favor of a combination of mis-

[11]See Coibion et al. (2018) as well for a review.

[12]While the current work focuses on representative agents model and, thus, aggregate/consensus forecasting data, Gabaix (2020) shows that cognitive discounting (myopia) can be microfounded through noisy signals. As argued by Gabaix (2020), a more explicit modelling of individual noisy signals would give rise to individual forecasts' over-reaction to new individual information as documented by Bordalo et al. (2020).

[13]Experimental evidence in Adam (2007), Hommes (2013); Hommes et al. (2019) and Petersen (2015), 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, evidence in Malmendier and Nagel (2016) shows that expectations are history dependent not model-based rational. 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)).

specified forecasting rules and myopia. Evidence of myopia 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).

This work also shares common insights with the literature that posits that imperfect common knowledge can explain observed persistence better than its FIRE counterpart (see, e.g., Milani (2006, 2007), Slobodyan and Wouters (2012a), Hommes et al. (2019)). The novelty of the present paper, however, is the finding that myopia is an important ingredient in improving the fit of the model.[14] Relatedly, 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. Even though this 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.[15]

The paper is also 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), Cole and Milani (2019), and Gaus and Gibbs (2018). Differently, the present paper estimates the model conditional on the novel combination of autoregressive mis-specified forecasts and myopia. Mis-specified, due to imperfect common knowledge, the partial equilibrium and full New Keynesian models share the same infinite horizon structure as in Preston (2005).[16]

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 US SPF data. Section 4 nests the

[14] In fact, the economy constrained to no myopia resembles to a certain extent the one under FIRE - consistent with recent laboratory experimental findings in Evans et al. (2019) showing that short-horizon forecasts are characterized by more substantial deviations from FIRE than long-horizon forecasts.

[15] 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.

[16] See Eusepi and Preston (2018) as well for a review.

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 expectations 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[17]

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

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, 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 $\hat{p}^*_t$. 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 (3) - once aggregated - will not produce the standard Phillips curve. For simplicity purposes only, I assume that firms understand that marginal costs evolve according to (4).[18] 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 RE, 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 RE. To model myopia, I build on the idea of cognitive

[17] See Appendix A for more details.

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

discounting of 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 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,

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

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

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

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,

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

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 $\widetilde{\mathbb{E}}^{\star}_t \hat{\pi}_{t+h}$. 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,[19]

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

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 $\epsilon_t$ is perceived to follow a white noise process. The value of $\epsilon_t$ 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 $(\delta, \gamma)$ will be pinned down using the solution concept of a Consistent Expectations 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 RE. Therefore, firms can use their own optimal condition in (6) to form expectations about inflation, and the infinite-horizon New

[19] 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.[20]

<u>*Remark 2:*</u> 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 (8) is a valid perceived law of motion for inflation if and only if its parameters, which represent the perceived unconditional mean ($\delta$) and first-order autocorrelation ($\gamma$), are consistent with the same moments from the data-generating process for inflation in (11). Coefficients $\delta$ and $\gamma$ in equilibrium are pinned down through the solution concept of a CE equilibrium, as defined by Hommes and Zhu (2014):

**Definition 1** *A pair $(\delta^*, \gamma^*)$, where $\delta^*$ and $\gamma^*$ 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 $\delta^*$* and *the stationary stochastic process defined by (11) has unconditional first-order autocorrelation coefficient $\gamma^*$.*

Along the RE 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*.

[20] 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,

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

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

**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 $(\delta^*, \gamma^*)$ 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* $\alpha$*) leads to lower* $\gamma^*$ *in equilibrium.*

*ii) A higher degree of myopia (lower* $n$*) leads to lower* $\gamma^*$ *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. Furthermore, 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. 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). Mis-specified, due to the nature of the expectations formation process, the actual law of motion for inflation in equilibrium resembles the one that would be derived under RE in a setting with inflation indexation/backward-looking pricing (even though these features are missing in this Section's setup).

# 3 Forecasting Data Evidence

To assess the empirical relevance of the proposed expectations formation process in Section 2, I derive three testable implications for forecasting errors and show that US SPF inflation data are consistent with the combination of mis-specified forecasts and myopia.

## 3.1 Delayed Over-shooting

Angeletos et al. (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. The authors define this phenomenon as *delayed over-shooting.*[21] The present paper shows that a combination of myopia with autoregressive mis-specified forecasts delivers delayed over-shooting if there is sufficient *endogenous* over-extrapolation. This result differs from Angeletos et al. (2021), who show that a model with *exogenous* over-extrapolation and noisy information can replicate this fact.

**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 $\varepsilon_t$, i.e.,*

$$\mathbb{I}_{k,h} = \frac{\partial(\hat{\pi}_{t+k} - \widetilde{\mathbb{E}}_{t+k-h}\hat{\pi}_{t+k})}{\partial\varepsilon_t} \tag{15}$$

*Let the expectations formation process be a combination of autoregressive mis-specified forecasting rules and myopia. Then,*

*i) There will be delayed over-shooting for any parameterization that yields $b \geq \rho$.*

*ii) For $b < \rho$, delayed over-shooting occurs if $\rho^{h+1} < n^h(\gamma^*)^{h+1}$.*

[21] I refer the reader to Section 5 in Angeletos et al. (2021) for a description of their IRF estimation methodologies, as well as Figures 3 and 4 for a visualization of their results.

**Proof.** See Appendix C.3. ■

Proposition 1 showed that $\gamma^* \geq \rho$; therefore, the condition for delayed over-shooting is satisfied if there is sufficient endogenous over-extrapolation, i.e., if $\gamma^* >> \rho$. Note that for $n = 1$, delayed over-shooting is always satisfied, since the condition from Proposition 2 translates into $\gamma^* > \rho$, which, from the analysis in the previous section, is always the case. On the contrary, if $n = 0$, the condition for late over-reaction fails to hold.

Figure 1 visualizes the results of Proposition 2 for the annual inflation forecast errors ($h = 3$) 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, i.e., 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.

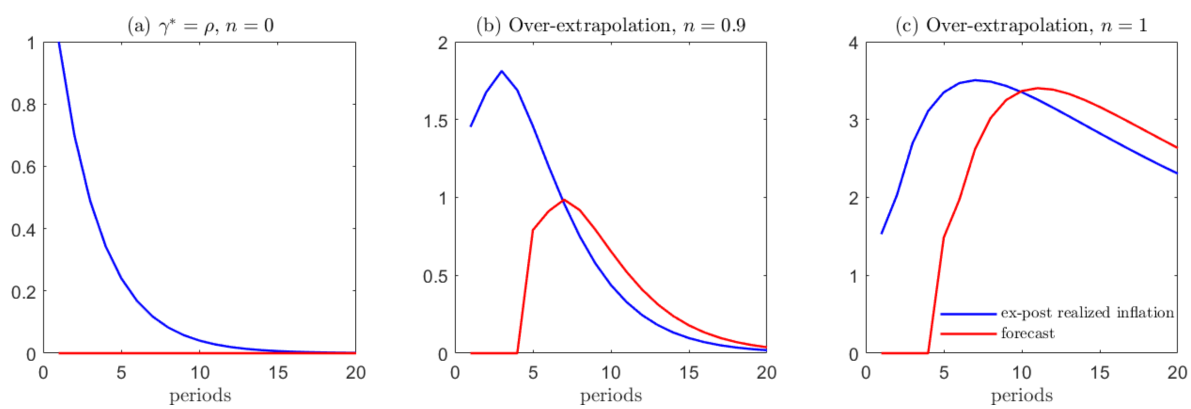



Figure 1: **Evolution of the IRFs of the annual forecasts and ex-post realized inflation for various values of $n$.** Parameterization: $\kappa = 1$, $\rho = 0.7$, $\alpha = 0.5$, $\sigma_\varepsilon = \sqrt{5}$. The implied equilibrium first-order autocorrelation coefficients of inflation are, respectively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

I now turn to the other two expectations alternatives, namely, well-specified forecasting rules absent and with myopia. Proposition 3 proves that neither of these two cases are consistent with a flip in the sign of $\mathbb{I}_k$ for some $k > 0$.

**Proposition 3** *Let $\mathbb{I}_{k,h}$ be defined as in (15). If $\widetilde{\mathbb{E}}_t\hat{\pi}_{t+h} = n^h\mathbb{E}_t\hat{\pi}_{t+h}$, where $\mathbb{E}_t$ is the RE operator, then $\mathbb{I}_{k,h} \geq 0$ for any $k \geq 0$. If $n = 1$, i.e., $\widetilde{\mathbb{E}}_t\hat{\pi}_{t+h} = \mathbb{E}_t\hat{\pi}_{t+h}$, then $\mathbb{I}_{k,h} = 0$ for any $k \geq 0$.*

**Proof.** See Appendix C.4. ■

The analysis in this subsection showed that, so far, the admissible expectational assumptions would be mis-specified forecasts with or absent myopia, i.e., combined with $n \in (0,1)$ or $n = 1$, respectively.

## 3.2 Forecast Errors as a Function of Ex-ante Forecast Revisions

This subsection studies the structure of ex-post forecasting errors implied by mis-specified forecasts combined with and absent myopia in more details. As derived in Section 2.4, the data-generating process for inflation along the CE equilibrium path is given by equation (13), $\hat{\pi}_t = a\hat{mc}_t + b\hat{\pi}_{t-1}$, where $a = \frac{\kappa}{1-\alpha\beta\rho n}$ and $b = \frac{\beta n(1-\alpha)}{1-\alpha\beta\gamma^*}(\gamma^*)^2 < 1$. Proposition 4 shows that ex-post inflation forecast errors can be generally written as a function of ex-ante forecast revisions and the second lag of inflation.

**Proposition 4** *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 ex-post forecast error can be written as*

$$\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t\hat{\pi}_{t+h} = K_h FR_{t,t+h} + L_h\hat{\pi}_{t-2} + \boldsymbol{\varepsilon_{t:t+h}} \tag{16}$$

*where $FR_{t,t+h}$ denotes ex-ante forecast revisions; $K_h = -1 + \frac{1}{n^h}\left(\frac{\rho}{\gamma^*}\right)^{h+1}\sum_{j=0}^{h+1}\left(\frac{b}{\rho}\right)^j$ and $L_h = n^{h+1}(\gamma^*)^{h+2}\left(K_h - \left(\frac{b}{\gamma^*}\right)\left(\frac{\rho}{n\gamma^*}\right)^{h+1}\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j\right)$. The term $\boldsymbol{\varepsilon_{t:t+h}}$ is accumulated noise, uncorrelated with $FR_{t,t+h}$ or $\hat{\pi}_{t-2}$.*

**Proof.** See Appendix C.5. ■

The regression in (16) can be estimated through OLS since the error term, $\boldsymbol{\varepsilon_{t:t+h}}$, is uncorrelated with any of the regressors. The insightful work of Coibion and Gorodnichenko (2015) has shown that the estimate of $K_h$ for the annual inflation forecasting data ($h = 4$) from the US SPF is positive. Relatedly, they have shown that once one controls for ex-ante forecast revisions, inflation lags have no predictive power over ex-post forecast errors. The present paper re-estimates equation (16) with annual ($h = 3$) US SPF forecasting data on the GNP/GDP deflator (1968:Q4 -

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $\hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+3}$ | | | | |
| *Panel A: GNP/GDP Deflator* | | | *Panel B: CPI Inflation* | |
| $\left(\widetilde{\mathbb{E}}_t \hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_{t-1} \hat{\pi}_{t+3}\right)$ | 1.189*** | 1.195*** | 0.455** | 0.389*** |
| | (0.377) | (0.389) | (0.097) | (0.123) |
| $\hat{\pi}_{t-2}$ | | -0.010 | | -0.068 |
| | | (0.047) | | (0.066) |
| Constant | -0.112 | -0.084 | -0.108 | 0.091 |
| | (0.127) | (0.207) | (0.111) | (0.205) |
| R-squared | 0.148 | 0.148 | 0.050 | 0.069 |
| Observations | 197 | 196 | 151 | 150 |

Newey-West standard errors in parentheses based on Barlett kernel with 6 lags.
*** Robust at 1%, ** Robust at 5%, * Robust at 10%.

Table 1: Forecasters' under-reaction to ex-ante forecast revisions. Panel A estimates regression in (16) with GNP/GDP inflation forecast data from 1968:Q4 to 2020:Q1. Panel B estimates regression in (16) with CPI inflation forecast data from 1981:Q3 to 2020:Q1. In columns (2) and (4) I control for the second lag of inflation.

2020:Q1) and CPI inflation (1981:Q3 - 2020:Q1). Results are reported in Table 1. As anticipated, the OLS estimate of $K_4$ is significantly positive, and there is no significant predictability of annual inflation forecast errors by inflation realized in period $(t-2)$.

In the context of equation (16) in Proposition 4, this implies that for a realistic expectations formation process $K_h > 0$ and $L_h \approx 0$ should be feasible. Leveraging on the evidence presented in Table 1, Proposition 5 provides conditions under which $K_h > 0$. Importantly, whether there is under-reaction to ex-ante forecast revisions is tightly connected with the presence or lack thereof of delayed over-shooting as defined in Proposition 2. If there is *no* delayed over-shooting, i.e., there is no sufficiently high over-extrapolation, forecasts are guaranteed to update slowly as new information becomes available, leading to under-reaction of forecast errors to ex-ante forecast revisions. In the presence of delayed over-shooting, over-extrapolation is already implied. High levels of over-extrapolation, i.e., large $(n^h(\gamma^*)^{h+1} - \rho^{h+1})$, impede the slow update of forecasts, thus leading to forecasters' over-reaction to ex-ante forecast revisions. On the contrary, sufficiently low levels of over-extrapolation, as defined in Proposition 5, allow for slow forecast update even in the presence of delayed over-shooting.

**Proposition 5** *Consider the expression for forecasting errors in (16), where* $n \in (0,1]$*, and let delayed over-shooting be defined as in Proposition 2. Let* $d_h = b\rho^{h+1}\sum_{j=0}^{h+1}\left(\frac{b}{\rho}\right)^j > 0$*. Then, the following statements are true:*

*i) For any parameterization that leads to* <u>no</u> *delayed over-shooting in equilibrium* $K_h$ *is guaranteed to be positive.*

*ii) For any parameterization that leads to delayed over-shooting,* $K_h > 0$ *if* $0 < n^h(\gamma^*)^{h+1} - \rho^{h+1} < d_h$*.*

**Proof.** See Appendix C.6. ■

As noted in Corollary 1, a lower degree of myopia lands the economy in an equilibrium characterized by higher inflation persistence; so a lack of myopia can increase the degree of over-extrapolation. Therefore, as myopia fades away, the likelihood of forecasters' under-reaction to ex-ante forecast revisions diminishes.

**Proposition 6** *Consider the expression for forecasting errors in (16), where* $n \in (0,1]$ *. Then,* $L_h = 0$ *is guaranteed to occur for some* $n \in \left(\left(\frac{\rho}{\gamma^*}\right)^2, 1\right)$*. If* $n = 1$*,* $L_h = 0$ *can happen only if* $\rho = 0$*.*

**Proof.** See Appendix C.7. ■

Proposition 6 shows that the combination of mis-specified forecasts and myopia is consistent with the evidence in Table 1 that the second lag of inflation is insignificant in explaining forecast errors once forecast revisions are being controlled for. On the contrary, mis-specified forecasts absent myopia can mirror such evidence only if the marginal cost has no persistence. Importantly, Corollary 2 shows that $L_h = 1$ in the case of mis-specified forecasting rules absent myopia violates the condition for delayed over-shooting.

**Corollary 2** *Consider the expression for forecasting errors in (16), with* $n = 1$*. If* $L_h = 0$*, the condition for delayed over-shooting in Proposition 2 is violated.*

**Proof.** For $n = 1$, $L_h = 0$ only if $\rho = 0$. The lack of marginal cost persistence implies that inflation will have no inertia either in equilibrium and that $\widetilde{\mathbb{E}}_t\hat{\pi}_{t+k} = 0$. Hence, the impulse response function of forecast errors to a marginal cost innovation will never exhibit delayed over-shooting. ■

### 3.3 Forecast Errors as a Function of Current Realizations

Proposition 7 shows that ex-post forecasting errors generated under a combination of mis-specified forecasts and myopia can be negatively predictable by inflation realized at the time of forecast.

**Proposition 7** *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 ex-post forecast error is*

$$\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+h} = M_h \hat{\pi}_t + error_{t,t+h} \tag{17}$$

*with* $M_h = \frac{b^{h+1} - n^h (\gamma^*)^{h+1}}{b} \leq 0$. *The term* $error_{t,t+h}$ *is positively correlated with* $\hat{\pi}_t$*, and the OLS estimate of* $M_h$ *is biased upward.*

**Proof.** See Appendix C.8. ■

Note that the sign of $M_h$ would be negative in the case of mis-specified forecasts absent myopia as well ($b^{h+1} - n^h(\gamma^*)^{h+1} \leq 0$ for $n = 1$). Mis-specified forecasts induce over-reaction of forecasts to inflation realized at the time of forecast.[22] Why? The correlation between inflation in period $(t + h)$ and inflation in period $t$, conditional on the path of innovations $\boldsymbol{\varepsilon}_{t:t+h}$, is $b^{h+1}$. On the other hand, the correlation between anticipated inflation in period $(t + h)$ and inflation in period $t$, conditional on innovation $\varepsilon_t$, is $bn^h(\gamma^*)^{h+1} \geq b^{h+1}$.

Kohlhas and Walther (2021) provide evidence that on average forecasters over-react to information realized at the time of forecast, i.e., that the sign of $M_h$ is negative. In particular, the authors show that the consensus annual forecast errors for CPI inflation depend negatively on inflation at the time of forecast, and in that case, the OLS estimate of $M_3$ is $-0.19$ with a robust standard error equal to 0.07 (estimation period 1982:Q3 to 2021:Q1). Regarding inflation measured by the GNP/GDP deflator, the authors report a positive dependence of ex-post annual forecast errors ($h = 3$) on inflation at the time of forecast for the period 1970:Q4 to 2020:Q1 (see Table C.7 in their online appendix). However, as shown in Table 2, for the period from 1981:Q1 to 2020:Q1 the OLS estimate of $M_3$ is instead *negative* and *statistically significant* so.

[22] In the case of well-specified forecasting rules and myopia, it is easy to show that

$$\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+h} = \underbrace{\rho^h(1 - n^h)}_{(+)} \hat{\pi}_t + error_{t,t+h}$$

where $error_{t+1,t+h}$ is uncorrelated with $\hat{\pi}_t$. If myopia is shut down, i.e., FIRE is imposed, $\rho^h(1 - n^h) = 0$.

| | (1) | (2) |
|---|---|---|
| $\hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+3}$ | | |
| $\hat{\pi}_t$ | 0.104 | -0.077*** |
| | (0.067) | (0.027) |
| Constant | -0.458*** | -0.214* |
| | (0.167) | (0.117) |
| R-squared | 0.039 | 0.024 |
| Time period | 1968:Q4 - 2020:Q1 | 1981:Q1 - 2020:Q1 |

Newey-West standard errors in parentheses based on Barlett kernel with 5 lags.
*** Robust at 1%, ** Robust at 5%, * Robust at 10%.

Table 2: Estimates of regression (17) on US SPF GNP/GDP deflator annual forecasting data.

Therefore, an expectations formation process characterized by mis-specified forecasts and myopia (or absent it) would be consistent with the empirical findings of Kohlhas and Walther (2021), which are further complemented by the results in Table 2.

### 3.4 Taking Stock

I now turn to the parametric region of myopia, $n$, and shock inertia, $\rho$, parameters that ensure all three aforementioned empirical facts are matched. Figure 2 exhibits the regions where, in equilibrium, there is under-reaction to ex-ante forecast revisions (areas in white and dark gray); non-significant second lag of inflation once ex-ante forecast revisions are controlled for in (16) (areas in light gray and white); and delayed over-shooting (area to the right of the dashed red curve). Note that Proposition 2 shows that mis-specified forecasts generate over-reaction of forecast errors to inflation realized at the time of forecast ($M_h \leq 0$), regardless of the degree of myopia, forecasting horizon, or model's parameterization. Therefore, any parameter combination between the degree of myopia and persistence of the marginal cost in Figure 2 guarantees that $M_h \leq 0$ as defined in (17).

Note that the condition for over-extrapolation necessitates sufficiently high endogenous over-extrapolation, which, as shown in Figure 2, is guaranteed to occur for relatively low degrees of myopia. Furthermore, as predicted by Proposition 5, there exists a region within the parametric space of delayed over-shooting where forecasters over-react not only to recent inflation but also to

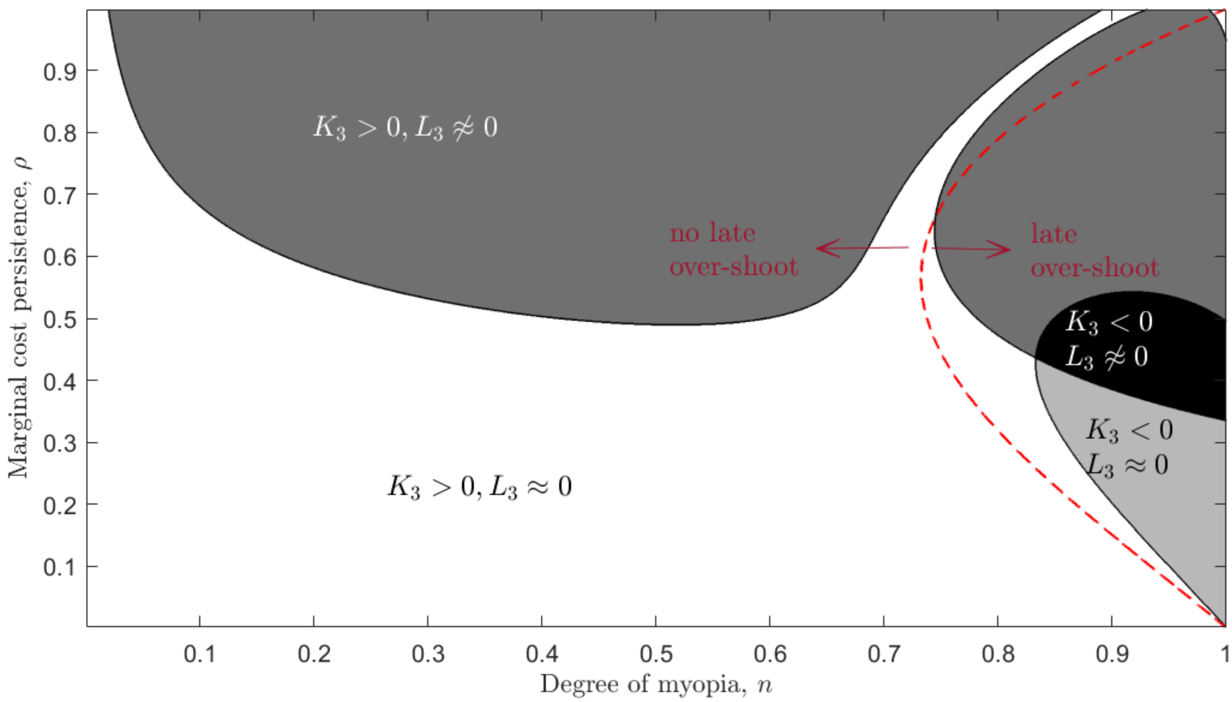



Figure 2: Regions of delayed over-shooting, under-reaction to ex-ante forecast revisions, and non-significant second lag of inflation in (16). Delayed over-shooting: region to the right of the dashed red curve. Under-reaction to ex-ante forecast revisions: regions in white and dark gray. Non-significant second lag of inflation in (16), $L_3 \approx 0$: white and light gray regions. $L_3 \approx 0$ ($L_3 \not\approx 0$) is defined as $|L_3| < 0.01$ ($|L_3| \geq 0.01$). Over-reaction to current inflation occurs for all combinations of $n$ and $\rho$. Forecasting horizon: $h = 3$. Parameterization: $\alpha = 0.5$, $\kappa = 1$, $\sigma = \sqrt{5}$.

ex-ante forecast revisions, i.e., $K_3 < 0$. The region where *all* empirical facts are matched is the area in white to the right of the red dashed curve.

Figure 3 shows how the different regions change as prices become stickier; that is, as $\alpha$ increases. The threshold red curve shifts to the right, thus shrinking the parametric space that guarantees delayed over-shooting. Corollary 1 proved that as prices become stickier, the equilibrium persistence decreases. Under such conditions, to generate sufficiently high over-extrapolation a lower degree of myopia is necessary. The parametric region where all empirical facts are matched allows for higher marginal cost persistence as prices become stickier. Figure 4 provides a visual sensitivity analysis of how the different regions change as the forecasting horizon, $h$, increases. The implications are very similar to the ones pertaining to Figure 3.

To summarize, the analytical analysis has revealed strong evidence in favor of a combination of mis-specified autoregressive forecasting rules and myopia as it is consistent with *all three facts*

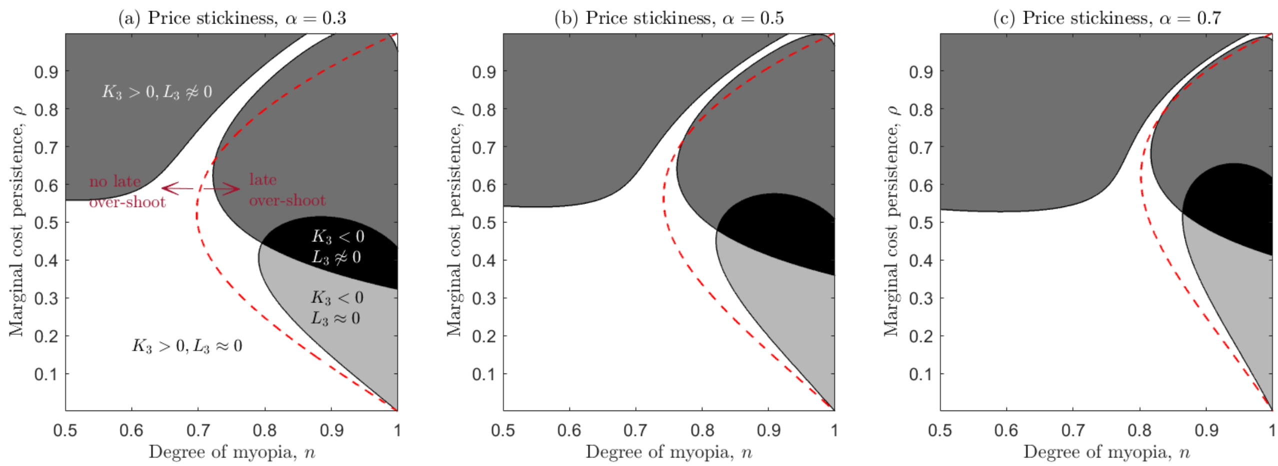



Figure 3: Sensitivity analysis with respect to price stickiness. Interpretation of regions, the forecasting horizon and parameterization of the model is the same as in Figure 2.

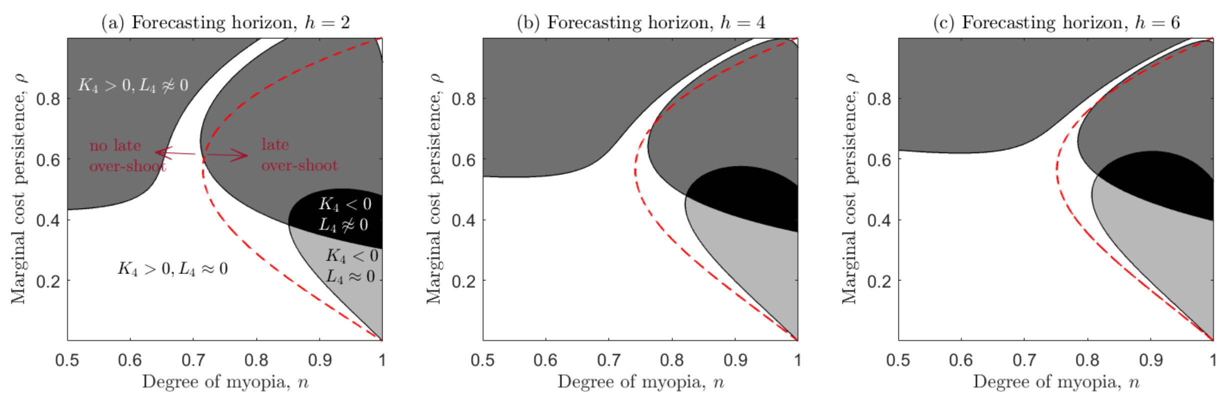



Figure 4: Sensitivity analysis with respect to the forecasting horizon. Interpretation of regions, the forecasting horizon and parameterization of the model is the same as in Figure 2.

concerning inflation consensus forecast data identified in the literature: forecasters' i) delayed over-shooting; ii) under-reaction to ex-ante forecast revisions; and iii) over-reaction to inflation realized at the time of forecast. On the other hand, mis-specified autoregressive forecasting rules absent myopia are consistent with delayed over-shooting and over-reaction to inflation at the time of forecast, but stand at odds with ex-post forecast errors depending on ex-ante forecast revisions *only*.

**Relation to other expectations formation processes.** Coibion and Gorodnichenko (2015) show that equation (16) 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 et al. (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 et al. (2021). While in Angeletos et al. (2021) over-extrapolation is an exogenous feature, in the present paper over-extrapolation is an organic implication of forecasters relying on mis-specified forecasting rules about inflation that are consistent with the inflation data-generating process up to the first-order autocorrelation. Furthermore, in the insightful paper of 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 procyclical variables. Nevertheless, as mentioned in Angeletos et al. (2021), asymmetric attention cannot match delayed over-shooting.

# 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 macroeconomic fluctuations. mis-specified, setting the model parameters to their estimated posterior mean, I simulate forecasting data and re-evaluate the three main facts.

## 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.[23] 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

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

demand of the $i^{th}$ household for the $j^{th}$ good is given by

$$c_{it}(j) = \left(\frac{P_{jt}}{P_t}\right)^{-\zeta} c_{it} \tag{18}$$

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 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 its budget constraint. The problem each household faces is

$$\max_{c_{it},H_{it},B_{it}} \widetilde{\mathbb{E}}_{it} \sum_{h=0}^{\infty} \beta^h \xi_{t+h} \left( \frac{(c_{it} - \eta c_{i,t-1})^{1-\sigma}}{1-\sigma} - \psi \frac{H_{it}^{1+\varphi}}{1+\varphi} \right) \tag{19}$$

s.t.

$$\frac{R_{t-1}}{\pi_t} b_{i,t-1} = b_{it} - w_t H_{it} - d_{it} + c_{it} \tag{20}$$

where $\beta \in (0,1)$ is the discount factor; $0 \leq \eta < 1$ measures the degree of habit in consumption; $\sigma$ is the inverse intertemporal elasticity of substitution; $\widetilde{\mathbb{E}}_{it}$ is the expectations operator described in the previous section; $\xi_t$ 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

$$\hat{x}_t = \frac{\eta}{1+n\eta\upsilon}\hat{x}_{t-1} + n\frac{\upsilon - n\beta\eta(1-\beta)(1-\eta)}{1+n\eta\upsilon}\widetilde{\mathbb{E}}_t^{\star}\hat{x}_{t+1} + \frac{\beta n^2(1-\beta)(1-\eta)(1-n\beta\eta)}{1+n\eta\upsilon}\widetilde{\mathbb{E}}_t^{\star}\sum_{h=0}^{\infty}(\beta n)^h \hat{x}_{t+h+2}$$
$$- \widetilde{\mathbb{E}}_t^{\star}\sum_{h=0}^{\infty}(\beta n)^h \frac{1-\beta\eta}{\sigma(1+n\eta\upsilon)}\left(\hat{R}_{t+h} - \hat{\pi}_{t+h+1} - \hat{e}_{t+h}\right) \tag{21}$$

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

$$\hat{e}_t = \rho_e \hat{e}_{t-1} + \sigma_e \varepsilon_t^e,\ \varepsilon_t^e \sim \mathcal{N}(0,1) \tag{22}$$

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 exogenously given technology, $z_t$, with labor input, $h_{jt}$, to produce output $y_{jt}$ as follows

$$y_{jt} = z_t h_{jt}^{a_h} \tag{23}$$

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 it according to the following indexation rule (see Christiano et al. (2005)),

$$P_{jt} = P_{j,t-1}\pi_{t-1}^{\rho_\pi} \tag{24}$$

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.,

$$\max_{P_{jt}^*} \widetilde{\mathbb{E}}_{jt} \sum_{h=0}^{\infty} (\alpha\beta)^h Q_{t+h} \left( \frac{P_{jt}^*}{P_{t+h}} \left( \frac{P_{t+h-1}}{P_{t-1}} \right)^{\rho_\pi} y_{j,t+h} - w_{t+h} h_{j,t+h} \right) \tag{25}$$

s.t.

$$y_{jt} = \left( \frac{P_{jt}^*}{P_t} \right)^{-\zeta} y_t \tag{26}$$

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

$$P_t = \left[ \alpha \left( P_{t-1} \left( \frac{P_{t-1}}{P_{t-2}}^{\rho_\pi} \right) \right)^{1-\zeta} + (1-\alpha)(P_t^*)^{1-\zeta} \right]^{\frac{1}{1-\zeta}} \tag{27}$$

Solving the firm's optimization problem, log-linearizing around the steady-state equilibrium, and applying the myopic adjustment process, one derives the aggregate supply as follows[24]

$$\hat{\pi}_t = \frac{1}{1 + n\beta\rho_\pi(1-\alpha)} (\rho_\pi \hat{\pi}_{t-1} - \kappa\eta\tau\hat{x}_{t-1}) + \frac{\kappa(\omega + \tau(1 - n\eta\beta(\alpha - \eta))}{1 + n\beta\rho_\pi(1-\alpha)} \hat{x}_t + \frac{1}{1 - \alpha\beta n\rho_u} \hat{u}_t$$
$$+ \frac{n\beta}{1 + n\beta\rho_\pi(1-\alpha)} \widetilde{\mathbb{E}}_t^\star \sum_{h=0}^{\infty} (\alpha\beta n)^h \left( (1-\alpha)(1 - \alpha\beta n\rho_\pi)\hat{\pi}_{t+h+1} + \kappa(\alpha\omega + \tau(\alpha - \eta)(1 - \alpha\beta n\eta)\hat{x}_{t+h+1} \right) \tag{28}$$

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

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

follow an AR(1) process,

$$\hat{u}_t = \rho_u \hat{u}_{t-1} + \sigma_u \varepsilon_t^u,\ \varepsilon_t^u \sim \mathcal{N}(0,1) \tag{29}$$

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 log-linearized policy rule is

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1-\rho_r)\phi_\pi \hat{\pi}_t + (1-\rho_r)\phi_x \hat{x}_t + \sigma_v \varepsilon_t^v,\ \varepsilon_t^v \sim \mathcal{N}(0,1) \tag{30}$$

**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

$$A_0(\Theta)S_t = A_1(\Theta)S_{t-1} + \widetilde{\mathbb{E}}_t^\star \sum_{h=0}^{\infty} (F(\Theta))^h A_2(\Theta) S_{t+h+1} + B(\Theta)\mathcal{E}_t \tag{31}$$

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 (31) 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 take into account an infinite horizon of autoregressive forecasts about deviations from the steady-state to make optimal decisions, exhibiting no myopia at all.[25]

## 4.2 SAC Learning

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

$$S_{1:3,t} = \boldsymbol{\delta}_{t-1} + \boldsymbol{\gamma}_{t-1}(S_{1:3,t-1} - \boldsymbol{\delta}_{t-1}) + \boldsymbol{\epsilon}_t \tag{32}$$

[25] Preston (2005) and Milani (2007) have used a model similar to (31) 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).

where $\boldsymbol{\delta}_{t-1}$ is the mean of $S_{1:3,0:t-1}$ series; $\boldsymbol{\gamma}_{t-1}$ represents the first-order correlation matrix between $S_{1:3,0:t-2}$ and $S_{1:3,1:t-1}$ series; and $\boldsymbol{\epsilon}_t$ is a white noise process. The formulation in (32) 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

$$\widetilde{\mathbb{E}}_t^{\star} S_{1:3,t+h} = \boldsymbol{\delta}_{t-1} + (\boldsymbol{\gamma}_{t-1})^{h+1}(S_{1:3,t-1} - \boldsymbol{\delta}_{t-1}) \tag{33}$$

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

$$\begin{aligned}
\boldsymbol{\delta}_t &= \boldsymbol{\delta}_{t-1} + \iota(S_{1:3,t} - \boldsymbol{\delta}_{t-1}) \\
\boldsymbol{\gamma}_t &= \boldsymbol{\gamma}_{t-1} + \iota\left((S_{1:3,t} - \boldsymbol{\delta}_{t-1})(S_{1:3,t-1} - \boldsymbol{\delta}_{t-1})' - \boldsymbol{\gamma}_{t-1}(S_{1:3,t} - \boldsymbol{\delta}_{t-1})(S_{1:3,t} - \boldsymbol{\delta}_{t-1})'\right)\boldsymbol{\eta}_t^{-1} \\
\boldsymbol{\eta}_t &= \boldsymbol{\eta}_{t-1} + \iota\left((S_{1:3,t} - \boldsymbol{\delta}_{t-1})(S_{1:3,t} - \boldsymbol{\delta}_{t-1})' - \boldsymbol{\eta}_{t-1}\right)
\end{aligned} \tag{34}$$

where $\boldsymbol{\eta}_t$ is the second moment matrix and $\iota$ 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.[26]

[26] 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 et al. (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 data, whereas constant-gain learning is robust to it.

## 4.3 Bayesian Estimation

Incorporating (32) into (31) yields the state-space representation of the model, described by

$$S_t = C_0(\Theta, \boldsymbol{\gamma}_{t-1})\boldsymbol{\Delta}_{t-1} + C_1(\Theta, \boldsymbol{\gamma}_{t-1})S_{t-1} + C_2(\Theta)\mathcal{E}_t \tag{35}$$

$$Y_t - \bar{Y} = PS_t \tag{36}$$

together with the dynamic system in (34), 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 is measured as the log difference between real GDP and potential GDP.[27] I refer the reader to Appendix D for more details on data preparation. The state-space form of the model in (34)-(36) is a Gaussian system; hence, I evaluate the likelihood function using the Kalman filter. The posterior distribution then is computed as

$$p(\Theta \mid Y_{1:T}) \propto p(Y_{1:T} \mid \Theta)p(\Theta) \tag{37}$$

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.[28]

I fix the discount factor $\beta = 0.99$. 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

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

[28] 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.

| Parameters | | pdf | mean | standard deviation |
|---|---|---|---|---|
| Calvo parameter | $\alpha$ | $\mathcal{B}$ | 0.5 | 0.2 |
| Degree of myopia | $n$ | $\mathcal{U}$ | 0.5 | $1/\sqrt{12}$ |
| Inverse IES coefficient | $\sigma$ | $\mathcal{G}$ | 2 | 0.5 |
| Phillips curve elasticity | $\kappa$ | $\mathcal{B}$ | 0.3 | 0.15 |
| Habit in consumption | $\eta$ | $\mathcal{U}$ | 0.5 | $1/\sqrt{12}$ |
| Inflation indexation | $\rho_\pi$ | $\mathcal{U}$ | 0.5 | $1/\sqrt{12}$ |
| Elasticity mc | $\omega$ | $\mathcal{N}$ | 0.8975 | 0.4 |
| Feedback to x | $\phi_x$ | $\mathcal{N}$ | 0.5 | 0.25 |
| Feedback to $\pi$ | $\phi_\pi$ | $\mathcal{N}$ | 1.5 | 0.25 |
| Interest rate smooth | $\rho_r$ | $\mathcal{B}$ | 0.5 | 0.2 |
| Autocorr. $e$ | $\rho_e$ | $\mathcal{U}$ | 0.5 | $1/\sqrt{12}$ |
| Autocorr. $u$ | $\rho_u$ | $\mathcal{U}$ | 0.5 | $1/\sqrt{12}$ |
| Std. $\varepsilon^e$ | $\sigma_e$ | $\mathcal{IG}$ | 0.1 | 2 |
| Std. $\varepsilon^u$ | $\sigma_u$ | $\mathcal{IG}$ | 0.1 | 2 |
| Std. $\varepsilon^v$ | $\sigma_v$ | $\mathcal{IG}$ | 0.1 | 2 |
| Gain parameter | $\bar{\iota}$ | $\mathcal{G}$ | 0.035 | 0.015 |

Table 3: Priors

Schorfheide (2016). Priors are given in Table 3. The Calvo parameter, $\alpha$, is following a beta prior with mean 0.5 and standard deviation 0.2. The prior for $n$ follows a non-informative uniform distribution with mean 0.5 and standard deviation of $1/\sqrt{12}$. The inverse intertemporal elasticity of substitution (IES) coefficient, $\sigma$, follows a gamma distribution with mean 2. The Phillips curve elasticity with respect to current output follows a beta prior with mean 0.3 and standard deviation 0.15. Habit in consumption and inflation indexation parameters follow a uniform distribution with mean 0.5. 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 output from their steady-state values are normally distributed with mean 1.5 and 0.5, respectively, and 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 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. Misspecified, the learning gain parameter follows a gamma distribution prior with mean 0.035 and standard deviation 0.015.

#### 4.3.1 Posterior Distribution

Table 4 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 or absent it.[29] Data fit is judged based on the evaluation of the log marginal data likelihood, computed through the modified harmonic mean method in Geweke (1999).[30] Generally, the presence of myopia ensures a better fit of the US macroeconomic data for both specifications with well- and mis-specified forecasting rules. I will then compare the fit of the other specifications relative to the one where well-specified forecasts are combined with myopia; that is, that specification will be the benchmark for model fit comparison purposes. The model where households and firms combine learning of AR(1) mis-specified forecasting rules with myopia (column 4 in Table 4) fits the data best. Specifically, values in parentheses in the last row of Table 4 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. 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 Bayes factor for the model with FIRE as well as the one with AR(1) forecasting rules absent myopia is much less than unity. Thus, both of these model specifications fit the data worse than the benchmark and, as a result, worse than mis-specified forecasts combined with myopia.

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 0.186 for the model with well-specified forecasting rules and 0.693 for the model with AR(1) mis-specified forecasts. In the model specifications with mis-specified forecasting rules, one can separately identify the Calvo parameter, $\alpha$, and price elasticity with respect to marginal costs, $\kappa$. In particular, the posterior mean of the Calvo parameter is estimated to be 0.35 for the model with myopia and 0.85 for the model absent

[29] Posterior distributions are generally well-behaved. I rely on the method proposed by Brooks and Gelman (1998) to analyze convergence statistics.

[30] 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.

myopia. Note that the specification with no myopia necessitates much higher price stickiness relative to the case with myopia. The implied expected price duration for the benchmark and no myopia specifications is, respectively, 4.6 and 20 months on average. The former is in accordance with findings in Bils and Klenow (2004) that, for most goods, prices change on average once every 4 months, while the latter generates too much price stickiness. The elasticity with respect to marginal costs is generally higher in the presence of myopia. At the posterior mean, $\kappa \approx 0$ under FIRE and $\kappa = 0.002$ for well-specified forecasts with myopia, whereas for mis-specified forecasts with and absent myopia, respectively, $\kappa = 0.012$ and $\kappa = 0.004$.

The model estimated under FIRE requires significantly higher degrees of habit in consumption (with a posterior mean of $\eta = 0.939$), relative to all of the other three assumptions about expectations, to fit US macro data. On the other hand, the posterior mean of $\eta$ for the benchmark model is 0.577, and for the model specifications with mis-specified forecasts $\eta \approx 0.4$. The degree of inflation indexation, $\rho_\pi$, is estimated to be around 0.3 - 0.4 for the model specifications absent myopia. In the presence of myopia, quite some degree of inflation indexation is necessary at the posterior mean ($\rho_\pi = 0.881$ for mis-specified forecasts and $\rho_\pi = 0.862$ for well-specified forecasts), but there is no need for a persistent cost-push shock ($\rho_u \approx 0.03$). The specification with no myopia, on the contrary, requires much less inflation indexation ($\rho_\pi \approx 0.3$ at the posterior mean), but it needs a persistent cost-push shock ($\rho_u \approx 0.5$).

The inverse elasticity of intertemporal substitution, $\sigma$, at the posterior mean is estimated to be around 2 for all specifications except the one with mis-specified forecasts and no myopia, where the estimate increases to around 3.9. Given the separable utility preferences for households, $\sigma$ coincides with the relative risk aversion coefficient; hence, forward-lookingness is associated with a higher degree of risk aversion.

Policy parameters are generally robust across specifications, with the posterior mode estimates of policy's reaction to the output gap being between 0.3 and 0.4, reaction to inflation around 1.5, and the interest rate smoothing parameter being close to 0.9. The posterior mean of $\omega$ is around 0.8 and 1 across the different specifications.

The posterior mean values for the learning gain parameter $\bar{\iota}$ are 0.076 when $n \in (0, 1)$ and 0.033 for $n = 1$, which further implies that a rolling window of about 13 and 33 quarters is used to update the forecasting process in the presence and absence of myopia, respectively.

Myopia plays a central role in determining how much of the observed inertia is due to autoregressive forecasting rules. The posterior mean estimate of the demand shock autocorrelation in the model with mis-specified forecasts and myopia is about 0.83 and it is significantly lower than its estimate of 0.98 in the model with mis-specified forecasts and no myopia. Furthermore, at the estimated posterior mean the demand shock is less persistent under FIRE ($\rho_e \approx 0.61$) relative to the specification in column (4) with $\rho_e \approx 0.83$. On the other hand, demand shock innovations are significantly less volatile under mis-specified forecasts than under FIRE: the posterior mode estimate for the innovation to the demand shock is 1.35 for the former and about 6.4 for the latter. 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% |
| $\alpha$ | - | - | - | - | - | - | 0.853 | 0.576 | 0.99 | 0.346 | 0.121 | 0.619 |
| $n$ | - | - | - | 0.186 | 0 | 0.401 | - | - | - | 0.693 | 0.464 | 0.862 |
| $\sigma$ | 2.356 | 1.356 | 3.448 | 2.045 | 1.166 | 2.843 | 3.855 | 2.898 | 4.944 | 2.185 | 1.463 | 3.028 |
| $\kappa$ | 0.000 | 0.000 | 0.000 | 0.004 | 0.001 | 0.008 | 0.004 | 0.002 | 0.008 | 0.012 | 0.006 | 0.021 |
| $\eta$ | 0.939 | 0.906 | 0.973 | 0.577 | 0.336 | 0.899 | 0.421 | 0.264 | 0.657 | 0.400 | 0.243 | 0.589 |
| $\rho_\pi$ | 0.387 | 0.077 | 0.680 | 0.887 | 0.834 | 0.940 | 0.331 | 0.003 | 0.94 | 0.861 | 0.695 | 0.980 |
| $\omega$ | 1.022 | 0.461 | 1.562 | 0.862 | 0.206 | 1.470 | 0.877 | 0.199 | 1.573 | 0.824 | 0.189 | 1.463 |
| $\phi_x$ | 0.351 | 0.224 | 0.484 | 0.420 | 0.234 | 0.600 | 0.399 | 0.233 | 0.61 | 0.413 | 0.251 | 0.622 |
| $\phi_\pi$ | 1.485 | 1.184 | 1.785 | 1.449 | 1.134 | 1.759 | 1.42 | 1.148 | 1.695 | 1.455 | 1.146 | 1.783 |
| $\rho_r$ | 0.888 | 0.858 | 0.919 | 0.915 | 0.886 | 0.945 | 0.914 | 0.882 | 0.943 | 0.916 | 0.885 | 0.944 |
| $\rho_e$ | 0.609 | 0.544 | 0.671 | 0.723 | 0.391 | 0.940 | 0.981 | 0.949 | 0.997 | 0.826 | 0.680 | 0.929 |
| $\rho_u$ | 0.495 | 0.412 | 0.574 | 0.020 | 0.000 | 0.045 | 0.449 | 0.006 | 0.862 | 0.036 | 0.002 | 0.107 |
| $\sigma_e$ | 6.399 | 5.715 | 7.114 | 3.568 | 1.357 | 6.904 | 0.189 | 0.075 | 0.399 | 1.350 | 0.614 | 2.320 |
| $\sigma_u$ | 0.012 | 0.012 | 0.012 | 0.214 | 0.150 | 0.272 | 0.174 | 0.047 | 0.347 | 0.350 | 0.300 | 0.405 |
| $\sigma_v$ | 0.209 | 0.192 | 0.226 | 0.208 | 0.191 | 0.224 | 0.208 | 0.192 | 0.225 | 0.209 | 0.193 | 0.226 |
| $\bar{\iota}$ | - | - | - | - | - | - | 0.033 | 0.004 | 0.065 | 0.076 | 0.043 | 0.117 |
| **Log marg. data dens.** | | | | | | | | | | | | |
| Modified Harmonic Mean | -520.27 | | | -264.20 | | | -284.35* | | | -258.16* | | |
| Bayes factor | ($e^{-262.11}$) | | | (1) | | | ($e^{26.19}$) | | | ($e^{6.04}$) | | |

Table 4: Posterior distribution of the model for various assumptions on expectations. Values in parentheses denote the Bayes factor of the model relative to the benchmark model with well-specified forecasts and myopia. The asterisk denotes strong evidence in favor of the model relative to the benchmark one.

Table 5 re-estimates the model under FIRE, while fixing the standard deviations of the innovations to the estimated posterior mean under mis-specified forecasts and myopia. FIRE then delivers much higher estimates of inflation indexation ($\rho_\pi \approx 0.7$) and the cost-push shock is estimated to be almost a random walk process. Habit in consumption is estimated to be about 0.3, close to the estimate under mis-specified forecasts and myopia in Table 4, but the estimate of the demand shock persistence is lower than the one found under mis-specified forecasts and myopia.

Overall, myopia and autoregressive mis-specified forecasts substitute for other frictions, such as habit in consumption and cost-push shock persistence, while being consistent with the evidence from forecasting data.

| | **FIRE** | | |
|---|---|---|---|
| Parameters | mean | 5% | 95% |
| $\sigma$ | 2.966 | 1.560 | 4.301 |
| $\kappa$ | 0.127 | 0.073 | 0.1781 |
| $\eta$ | 0.302 | 0.185 | 0.421 |
| $\rho_\pi$ | 0.699 | 0.314 | 1.000 |
| $\omega$ | 0.835 | 0.259 | 1.395 |
| $\phi_x$ | 0.058 | 0.014 | 0.098 |
| $\phi_\pi$ | 2.188 | 1.989 | 2.389 |
| $\rho_r$ | 0.790 | 0.764 | 0.816 |
| $\rho_e$ | 0.478 | 0.431 | 0.525 |
| $\rho_u$ | 0.994 | 0.987 | 1.000 |

Table 5: Posterior distribution of the model under FIRE. The standard deviations of the shock innovations to the estimated posterior mean under mis-specified forecasts and myopia, i.e., $\sigma_e = 1.35$, $\sigma_u = 0.209$, and $\sigma_v = 0.076$.

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 mode 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, in the early 1980s. Before the early 1980s, recessions were associated with increasing beliefs about the mean and

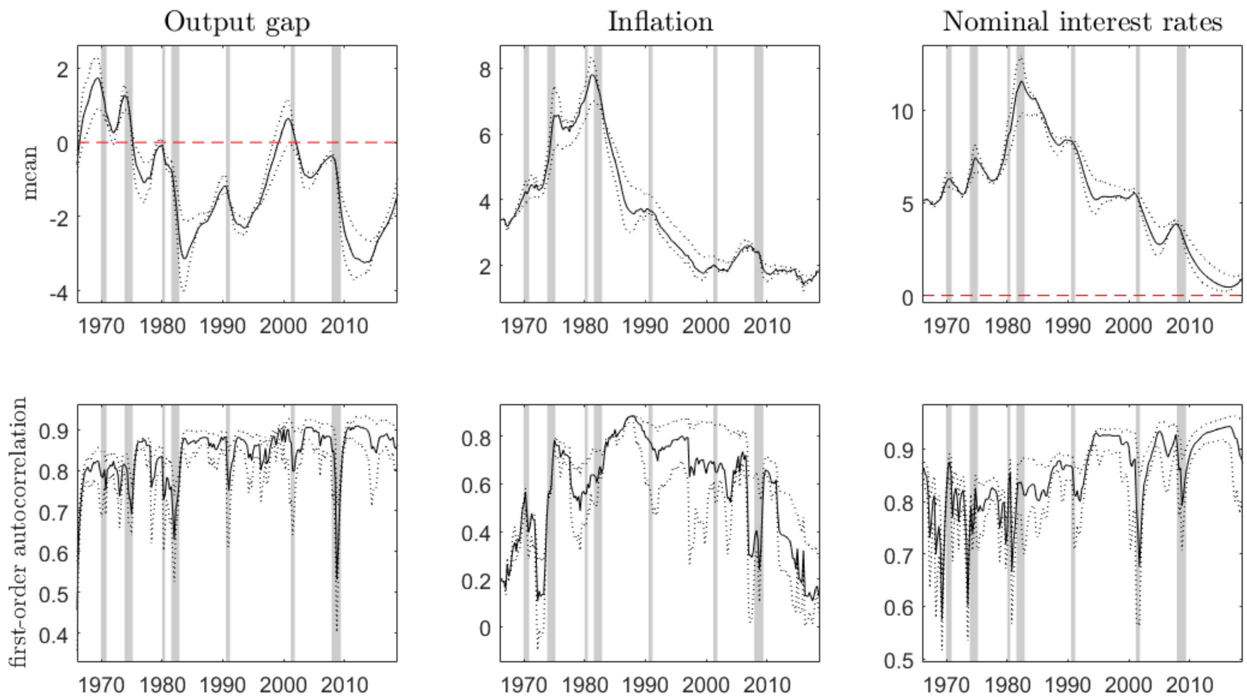



Figure 5: Evolution of the AR(1) forecast coefficients in the model with myopia. The black and dotted curves plot perceived moments for structural parameters set at their estimated posterior mode and 90 percent highest posterior density, respectively. Gray areas indicate recessionary periods as reported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

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.[31] Another interesting observation from Figure 5 is that the implied beliefs about the annualized mean of inflation over the last decade 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 and myopia under constant-gain learning.

The characteristics of the posterior distribution of parameters are exhibited in Table 6. Esti-

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

| Parameters | mean | 5% | 95% |
|---|---|---|---|
| $\alpha$ | 0.344 | 0.116 | 0.608 |
| $n$ | 0.704 | 0.407 | 0.898 |
| $\sigma$ | 2.015 | 1.384 | 2.732 |
| $\kappa$ | 0.010 | 0.004 | 0.018 |
| $\eta$ | 0.372 | 0.230 | 0.612 |
| $\rho_\pi$ | 0.878 | 0.740 | 0.986 |
| $\omega$ | 0.824 | 0.217 | 1.436 |
| $\phi_x$ | 0.415 | 0.251 | 0.638 |
| $\phi_\pi$ | 1.449 | 1.133 | 1.789 |
| $\rho_r$ | 0.915 | 0.883 | 0.945 |
| $\rho_e$ | 0.854 | 0.685 | 0.945 |
| $\rho_u$ | 0.047 | 0.002 | 0.127 |
| $\sigma_e$ | 1.109 | 0.409 | 2.374 |
| $\sigma_u$ | 0.352 | 0.295 | 0.418 |
| $\sigma_v$ | 0.208 | 0.192 | 0.225 |
| $\bar{\iota}$ | 0.072 | 0.044 | 0.102 |
| **Log marginal data density** | | | |
| Modified Harmonic Mean | -255.41 | | |

Table 6: Posterior distribution for the model with VAR(1) forecasting rules combined with myopia.

mates of parameters are very much similar to the ones found in Table 4 for AR(1) forecasting rules, and the log marginal data density is 2.75 units below the measure for AR(1) rules. According to Kass and Raftery (1995), this would be positive, not strong, evidence in favor of VAR(1) forecasts. Moreover, as Figure 6 shows, when agents engage in constant-gain learning of a VAR(1) forecasting process, 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 any significant information to households and firms in terms of forecasting on average, and it will not enhance the model's fit of the data. Moving forward then, I remain focused on AR(1) forecasting rules.

#### 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, the response of the perceived mean and first-order autocorrelation to

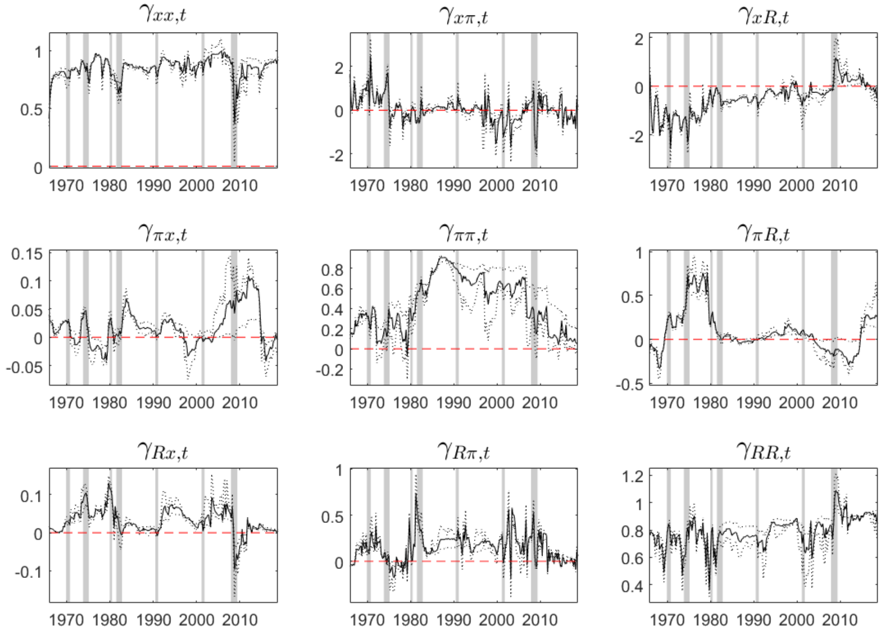



Figure 6: Evolution of the VAR(1) forecast coefficients in the model with myopia. The black and dotted curves plot implied beliefs for structural parameters set at their estimated posterior mode and 90 percent highest posterior density, respectively. Grey areas indicate recessionary periods as reported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

shocks, which - differently from, e.g., FIRE - enters the model's solution non-linearly, evolves jointly with the response of aggregates. To make the IRFs comparable across time periods and expectations' assumptions, I assume that the economy prior to the shock is at its steady-state.[32] 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, calibrated at the estimated posterior means of the model with AR(1) mis-specified forecasting rules combined with myopia. The three-dimensional IRFs show that the response of aggregates to various fundamental shocks depends on the perceived persistence of the underlying variable prior to the shock.

Figure 8 plots the IRFs projected on the [response - periods of response] plane to get an idea of the change in the magnitude of aggregates' response to shocks over the years. An interesting observation is that since the late 1980s, the response of the output gap to monetary shocks has increased, while the reaction of inflation to the same shocks has subsided.[33]

[32] Consecutively, the perceived mean of the aggregates is set to 0, whereas the prior perceived first-order auto-correlations each period are set to their values as implied by the model.

[33] This outcome would be consistent with the idea that the slope of the Phillips curve has flattened over recent decades.

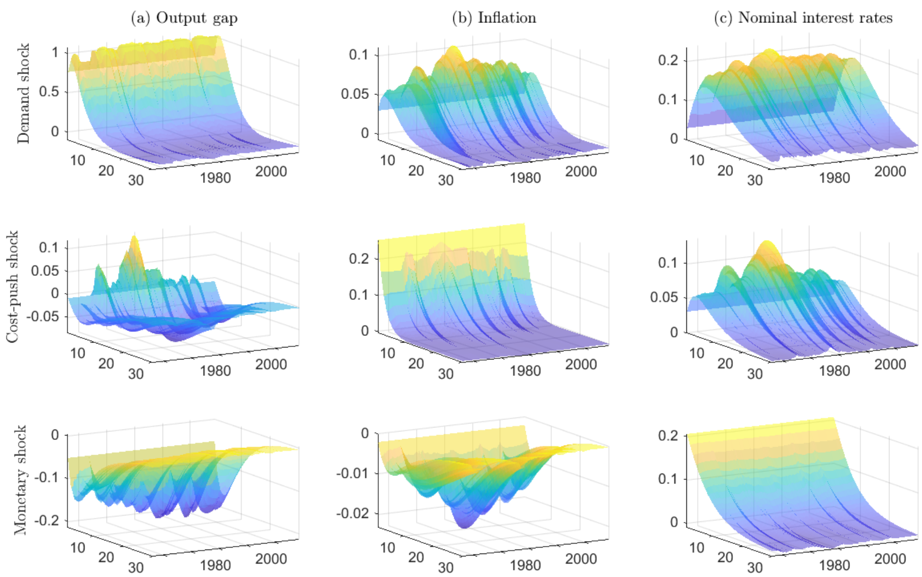



Figure 7: Three-dimensional impulse response functions to a one standard deviation positive demand, cost-push, and monetary shock for the model with mis-specified forecasts and myopia. Parameters are set at their estimated posterior mean as shown in Table 6. X axis: periods of response; y axis: estimation periods; z axis: impulse response function.

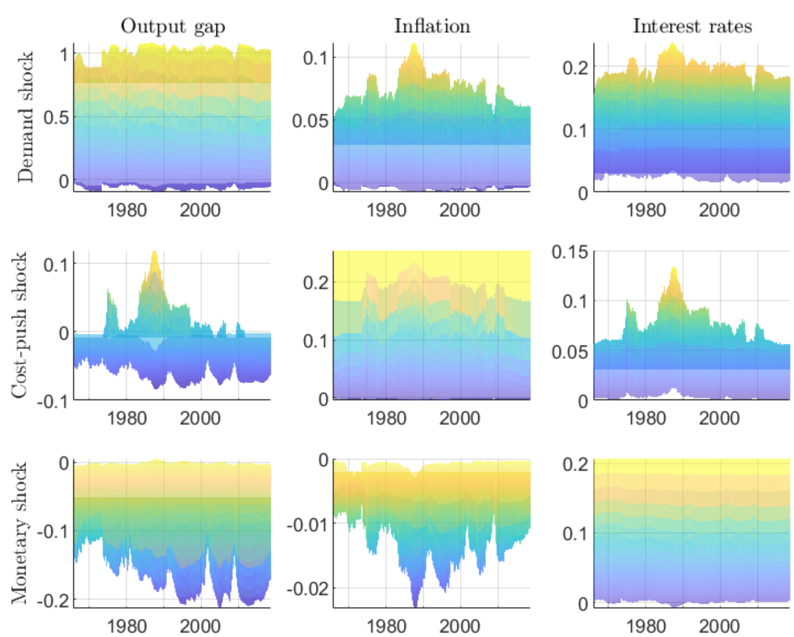
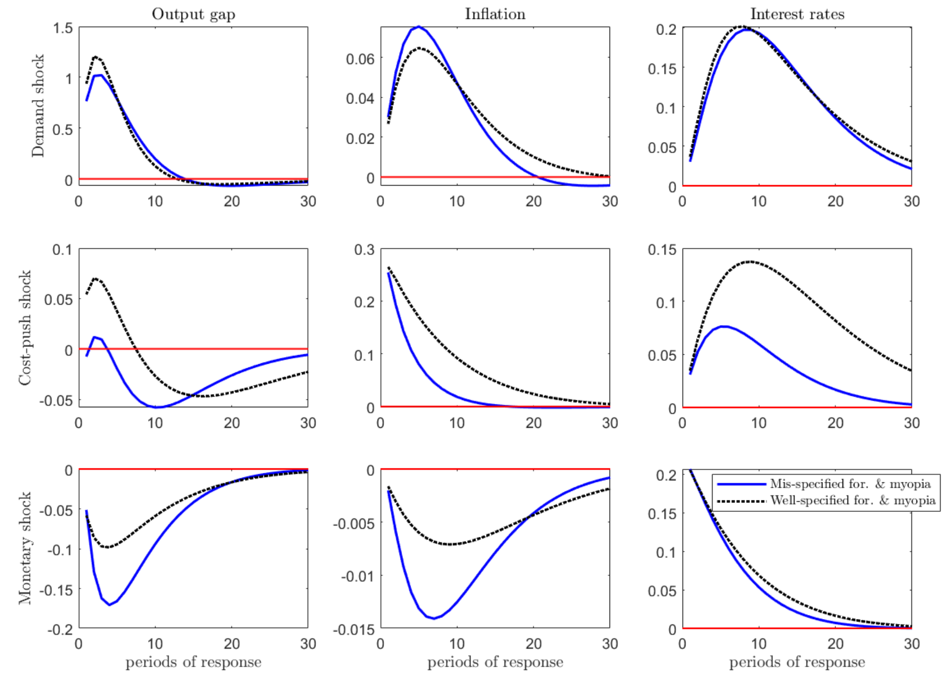
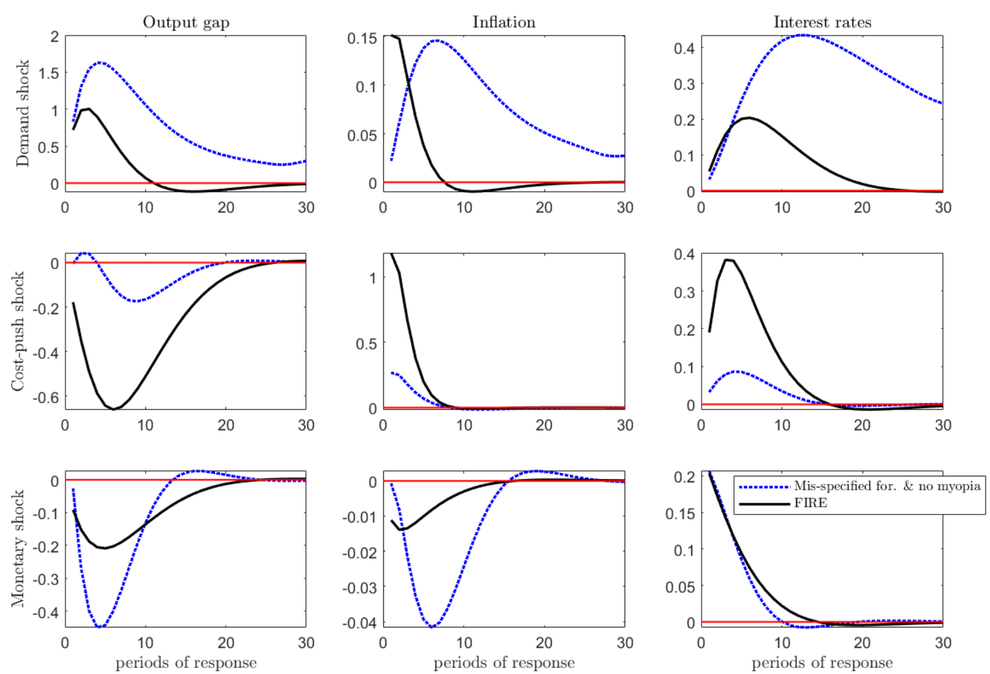



Figure 8: Impulse response functions to a one standard deviation demand, cost-push, and monetary shock in the model with mis-specified forecasts and myopia, projected on the [response - periods of response] plane.


Output gap
Inflation
Interest rates
Demand shock
Cost-push shock
Monetary shock
periods of response
Mis-specified for. & myopia
Well-specified for. & myopia

(a)


Output gap
Inflation
Interest rates
Demand shock
Cost-push shock
Monetary shock
periods of response
Mis-specified for. & no myopia
FIRE

To compare the IRFs across different assumptions about expectations, Figure 9 plots the average impulse responses of the output gap, inflation, and nominal interest rates for FIRE, well-specified forecasts with myopia, and mis-specified forecasts with and absent myopia to a one standard deviation demand, cost-push, and monetary shock. Parameters are set at their posterior mean as reported in Table 4. Due to the presence of habit in consumption and inflation indexation, all four specifications about expectations deliver hump-shaped responses. The left panel exhibits IRFs for mis-specified and well-specified forecasts combined with myopia, which are at the same time the two best performing assumptions on expectations in terms of model fit. Mis-specified forecasting rules deliver a more amplified response of the output gap and inflation to monetary shocks, as well as a more persistent response of both variables to demand shocks. Similarly, relative to FIRE the model under SAC learning with myopia generates significantly more persistent and volatile response functions of the output gap and inflation, especially to demand and monetary shocks.

# 5 Implications for Forecasting Errors

In this final section, I revisit the three empirical facts about inflation consensus forecasting errors analyzed in Section 3. Setting the model parameters to the posterior mean as found in Table 4, I generate the implied annual inflation forecast errors as well as their impulse responses to a demand, cost-push, and monetary shock over the estimation period of 1966:Q1 - 2018:Q3.

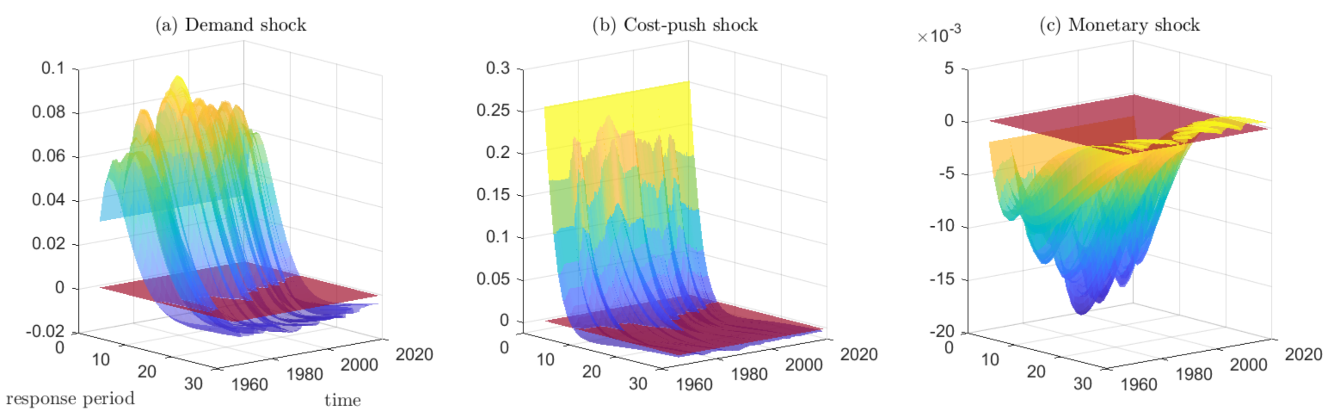



Figure 10: Three-dimensional impulse response functions of annual inflation forecasting errors to a one standard deviation positive demand, cost-push, and monetary shock in the model with mis-specified forecasts and myopia. In red: (response periods,time) plane.

**Delayed over-shooting.** Figure 10 plots the three-dimensional IRFs of annual inflation forecasting errors to demand, cost-push, and monetary shocks over the estimation period. Following a positive demand or cost-push shock, annual inflation forecast errors start positive but they turn negative after a few periods. The evolution of inflation and its annual forecast averaged across the estimation period in Figure 11 confirms this. On the other hand, following a monetary shock, the inflation forecast errors are initially negative and they turn positive at some later period. This is still consistent with delayed over-shooting: when the monetary shock hits the economy, the response of inflation is higher in absolute value relative to its forecast, but as shown in Figure 11, after a number of periods the response of annual forecasts becomes more negative (larger in absolute value) than ex-post realized inflation.

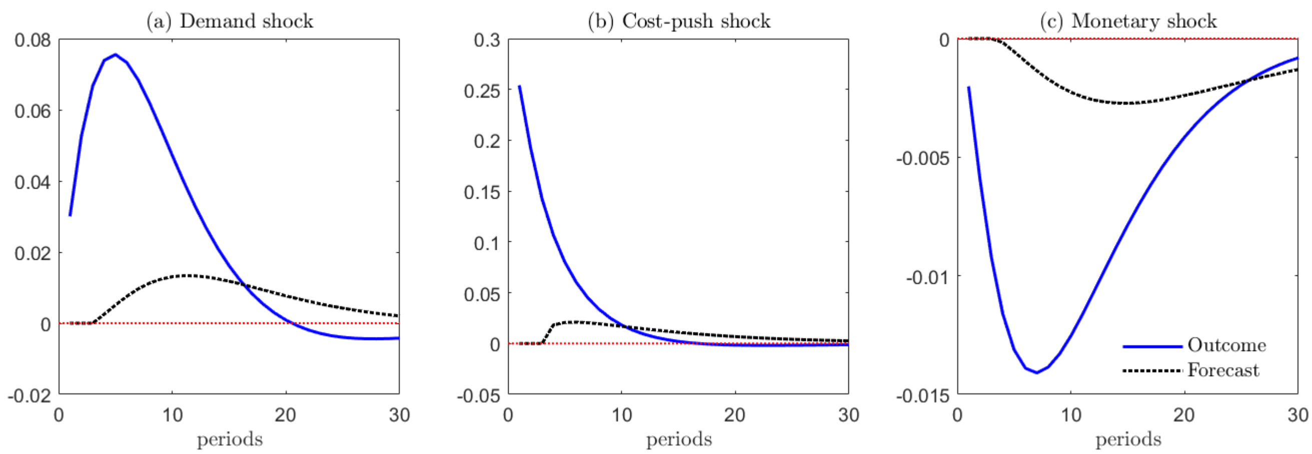



Figure 11: Impulse response functions of annual inflation forecasting errors to a one standard deviation positive demand, cost-push, and monetary shock in the model with mis-specified forecasts and myopia, averaged across the estimation period. In blue: Ex-post realized inflation. In dashed black: annual inflation forecast.

**Under-reaction to ex-ante forecast revisions, over-reaction to current inflation.** I estimate the following two regressions with simulated annual forecasting data over the Bayesian estimation period.

$$\hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+3} = c + K_4(\widetilde{\mathbb{E}}_t \hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_{t-1} \hat{\pi}_{t+3}) + O_z z_{t-2,t} + error_{t,t+4} \tag{38}$$

$$\hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+3} = c + M_3 \hat{\pi}_t + O_z z_{t-1,t} + error_{t+1,t+4} \tag{39}$$

where $z_{t-2,t}$ refers to a set of control variables in periods $(t-2)$, $(t-1)$, and $t$. The error terms in (38) and (39) are uncorrelated with any of the regressors. Since the data-generating process for inflation is much more complex than the one studied in the first part of the paper, the structure of inflation forecast errors is also more complicated. Specifically, inflation forecast errors may (theoretically) depend on other variables apart from aggregate inflation at the time of forecast or current and past forecasts. Therefore, the two estimates of interest, namely, $\hat{K}_3$ and $\hat{M}_3$, might suffer from omitted variable bias if one does not control for other variables that affect annual inflation forecast error data simulated through the full New Keynesian model. In Appendix E, I derive the set of relevant controls for both regressions. In other words, the first equation is an extended version of the main regression in Coibion and Gorodnichenko (2015), whereas the second one is an extended version of the main regression in Kohlhas and Walther (2021).

| | **Ex-ante forecast revisions** | **Current inflation** | |
|---|---|---|---|
| | (1) | (2) | (3) |
| $\hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_t\hat{\pi}_{t+3}$ | | | |
| $\widetilde{\mathbb{E}}_t\hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_{t-1}\hat{\pi}_{t+3}$ | 2.442*** | | |
| | (0.261) | | |
| $\hat{\pi}_t$ | | 0.267 | -1.002** |
| | | (0.441) | (0.456) |
| Constant | 0.523*** | 0.362* | 0.749*** |
| | (0.164) | (0.192) | (0.163) |
| Controls | ✓ | ✓ | ✓ |
| R-squared | 0.765 | 0.757 | 0.612 |
| Time period | 1966:Q1 - 2018:Q3 | 1966:Q1 - 2018:Q3 | 1981:Q1 - 2018:Q3 |

Newey-West standard errors in parentheses.
*** Robust at 1%, ** Robust at 5%, * Robust at 10%.

Table 7: Estimates of regressions in (38) and (39) on simulated forecasting data. Model parameters are set at the posterior mean for the model with mis-specified forecasts and myopia as documented in Table 4.

As reported in Table 7, annual inflation forecast errors depend positively on ex-ante inflation forecast revisions, speaking to forecasters' under-reaction to revisions. With respect to the fact that forecasters over-react to information at the time of forecast, I find results that are consistent with the US SPF data on the GNP/GDP deflator (note that inflation in the DSGE model is measured through the GDP deflator). Specifically, US SPF forecast errors of the GNP/GDP deflator reflect under-reaction to inflation at the time of forecast over the full sample from 1968:Q4 to 2018:Q3, *but* there is over-reaction for the data from 1981:Q1 to 2018:Q3 (see Table 2). The estimates of $M$ in columns (2) and (3) in Table 7 reflects that.

# 6 Concluding Remarks

The present paper simultaneously incorporates two of the most prominent deviations from the FIRE assumption, 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 the US SPF 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 three novel results. First, both estimation approaches speak strongly in favor of the model in which agents use mis-specified forecasts and myopia over a number of other alternatives. Importantly, a combination of mis-specified forecasting rules and myopia is consistent with consensus inflation forecasts' i) delayed over-shooting; ii) under-reaction to ex-ante forecast revisions *and* no reaction to inflation lags; and iii) over-reaction to recent events. Second, the best fitting expectations formation process for both households and firms is characterized by high degrees of myopia and simple AR(1) forecasts, and more elaborate VAR(1) rules do not add any useful information to private agents. Third, the estimated high degree of myopia - in the presence of mis-specified forecasts - generates substantial internal persistence and amplification to exogenous shocks.

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 would be to allow for heterogeneity in the degree of myopia, as well as in the structure of forecasting rules. Another potential avenue would be to re-examine ex-post forecast errors, while allowing forecasters' perpetual learning of mis-specified forecasting rules.

# Appendix

## 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 $\zeta$ 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 $\chi_{it}$ 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 $\xi_t$ 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

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

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

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

So,

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

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

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

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

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

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 $\lim_{h\to\infty}\widetilde{\mathbb{E}}_{it}\hat{cc}_{i,t+h} = 0$. 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 $\hat{cc}_{it} = \hat{c}_{it} - \eta\hat{c}_{i,t-1} - \beta\eta\widetilde{\mathbb{E}}_{it}(\hat{c}_{i,t+1} - \eta\hat{c}_{it})$, from which it follows that

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

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

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

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

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

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

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

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:

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

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

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

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,

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

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 $\beta n$ and $\alpha\beta n$, respectively. Using results from the previous subsection, the perceived law of motion (PLM) in matrix form can be written as

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

where $\boldsymbol{\delta}_t = \begin{bmatrix}\boldsymbol{\delta}_t' & \mathbf{0}_{1x2}\end{bmatrix}'$; $\boldsymbol{\Gamma}_t = \begin{bmatrix}\boldsymbol{\gamma}_t & \mathbf{0}_{3x2} \\ \mathbf{0}_{2x3} & \mathbf{0}_{2x2}\end{bmatrix}$; $H$ 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

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

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

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

where

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

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

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

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,

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

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,

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

Hence,

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

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

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

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, $\gamma^*$, 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 $\alpha$ 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 $\varepsilon_t$ in period $t$, is

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

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 will first show that $\mathbb{I}_{k,h}\geq 0$ when myopia is combined with well-specified forecasting rules. As shown in the main text, the aggregated optimal pricing rule in this case can be written as $\hat{\pi}_t=\kappa\hat{mc}_t+\beta n\mathbb{E}_t\hat{\pi}_{t+1}$, and the solution for inflation is $\hat{\pi}_t=a_0\hat{mc}_t$, where $a_0=\frac{\kappa}{1-\beta n\rho}$. Therefore,

$$\mathbb{I}_{k,h}=a_0\hat{mc}_{t+k}-a_0\rho^h\hat{mc}_{t+k-h}=a_0(\rho^k-n\rho^k)\varepsilon_t\geq 0 \tag{C.13}$$

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 $k\geq 0$.

## C.5 Proposition 4

Again, 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 $b = \frac{\beta n(1-\alpha)}{1-\alpha\beta n\gamma^*}(\gamma^*)^2$. The forecast error at horizon $h \leq 1$ is

$$\begin{aligned}
\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t\hat{\pi}_{t+h} &= a\hat{mc}_{t+h} + b\hat{\pi}_{t+h-1} - n^h(\gamma^*)^{h+1}\hat{\pi}_{t-1} \\
&= a\rho^{h+1}\hat{mc}_{t-1} + ab\hat{mc}_{t+h-1} + b^2\hat{\pi}_{t+h-2} - n^h(\gamma^*)^{h+1}(a\hat{mc}_{t-1} + b\hat{\pi}_{t-2}) + f(\varepsilon_t, ..., \varepsilon_{t+h}) \\
&= a(\rho^{h+1} + b\rho^h + ... + b^h\rho - n^h(\gamma^*)^{h+1})\hat{mc}_{t-1} + b^{h+1}\hat{\pi}_{t-1} - bn^h(\gamma^*)^{h+1}\hat{\pi}_{t-2} + \boldsymbol{\varepsilon_{t:t+h}} \\
&= \left(\rho^{h+1}\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j - n^h(\gamma^*)^{h+1}\right)(\hat{\pi}_{t-1} - b\hat{\pi}_{t-2}) + b^{h+1}\hat{\pi}_{t-1} - bn^h(\gamma^*)^{h+1}\hat{\pi}_{t-2} + \boldsymbol{\varepsilon_{t:t+h}} \\
&= \left(\rho^{h+1}\sum_{j=0}^{h+1}\left(\frac{b}{\rho}\right)^j - n^h(\gamma^*)^{h+1}\right)\hat{\pi}_{t-1} - b\rho^{h+1}\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j\hat{\pi}_{t-2}) + \boldsymbol{\varepsilon_{t:t+h}} \\
&= \underbrace{\left(-1 + n\left(\frac{\rho}{n\gamma^*}\right)^{h+1}\sum_{j=0}^{h+1}\left(\frac{b}{\rho}\right)^j\right)}_{K_h}\widetilde{\mathbb{E}}_t\hat{\pi}_{t+h} - \frac{b}{\gamma^*}\left(\frac{\rho}{n\gamma^*}\right)^{h+1}\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j\widetilde{\mathbb{E}}_{t-1}\hat{\pi}_{t+h} + \boldsymbol{\varepsilon_{t:t+h}} \\
&= K_h(\widetilde{\mathbb{E}}_t\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_{t-1}\hat{\pi}_{t+h}) + n^h(\gamma^*)^{h+1}\left(K_h - \frac{b}{\gamma^*}\left(\frac{\rho}{n\gamma^*}\right)^{h+1}\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j\right)\hat{\pi}_{t-2}
\end{aligned} \tag{C.14}$$

where $\boldsymbol{\varepsilon_{t:t+h}}$ is a linear function of $\varepsilon_t, ..., \varepsilon_{t+h}$.

## C.6 Proposition 5

i) Suppose the conditions for lack of delayed over-shooting are satisfied, i.e., $b < \rho$ *and* $n^h(\gamma^*)^{h+1} < \rho^{h+1}$. Then,

$$K_h = \frac{(-n^h(\gamma^*)^{h+1} + \rho^{h+1}) + \rho^h b + ... + b^{h+1}}{n^h(\gamma^*)^{h+1}} > 0 \tag{C.15}$$

ii) Now suppose the conditions for delayed over-shooting are satisfied, i.e., either $b > \rho$, or $n^h(\gamma^*)^{h+1} > \rho^{h+1}$ if $b < \rho$.

$$K_h = \frac{-n^h(\gamma^*)^{h+1}(\rho - b) + \rho^{h+2} - b^{h+2}}{n^h(\gamma^*)^{h+1}(\rho - b)} \tag{C.16}$$

Consider $b > \rho$ first. Then, $K_h > 0$ if $n^h(\gamma^*)^{h+1} < \frac{b^{h+2}-\rho^{h+2}}{b-\rho}$, or put differently if

$$n^h(\gamma^*)^{h+1} - \rho^{h+1} < \frac{b^{h+2} - \rho^{h+2}}{b - \rho} - \rho^{h+1} = b\rho^{h+1}\sum_{j=0}^{h+1}\left(\frac{b}{\rho}\right)^j \tag{C.17}$$

Note that $b > \rho$ implies that $n^h(\gamma^*)^{h+1} - \rho^{h+1} > 0$. Now consider $b < \rho$ and $n^h(\gamma^*)^{h+1} > \rho^{h+1}$. $K_h > 0$ if $0 < n^h(\gamma^*)^{h+1} < \frac{b^{h+2}-\rho^{h+2}}{b-\rho}$, which is the exact same condition as in (C.17).

## C.7 Proposition 6

Define $L_h(n) = n^{h+1}(\gamma^*)^{h+2}\left(K_h - \frac{b}{\gamma^*}\left(\frac{\rho}{n\gamma^*}\right)^{h+1}\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j\right)$. Then,

$$L_h(n) = n^{h+1}(\gamma^*)^{h+2}\left(-1 + \left(n - \frac{b}{\gamma^*}\right)\left(\frac{\rho}{n\gamma^*}\right)^{h+1}\sum_{j=0}^{h}\left(\frac{b}{\rho}\right)^j + n\left(\frac{b}{n\gamma^*}\right)^{h+1}\right) \tag{C.18}$$

$$\begin{aligned} L_h(1) &= (\gamma^* - b)(\rho(\rho^h + \rho^{h-1}b + ... + b^h) - \gamma^*((\gamma^*)^h + (\gamma^*)^{h-1}b + ... + b^h) \\ &= (\gamma^* - b)\left(\rho(\rho^h + \rho^{h-1}b + ... + b^h\right) - \gamma^*\left((\gamma^*)^h + (\gamma^*)^{h-1}b + ... + b^h\right) \\ &= (\gamma^* - b)\sum_{j=1}^{h+1} b^{h+1-j}(\rho^j - (\gamma^*)^j) \leq 0 \end{aligned} \tag{C.19}$$

where the last inequality follows from the fact that $\gamma^* \geq \rho$ and $\gamma^* > b$. Note that $L_h(1) = 0$ if and only if $\gamma^* = \rho$. Consider $n = \left(\frac{\rho}{\gamma^*}\right)^2$, which is less than unity for any $\rho \in (0, 1)$. I show that $L_h\left(\left(\frac{\rho}{\gamma^*}\right)^2\right) > 0$. Define $b_0 = \frac{\beta(1-\alpha)}{1-\alpha\beta n\gamma^*}(\gamma^*)^2$, such that $b_0 n = b$. Then, we have

$$\begin{aligned} L_h\left(\left(\frac{\rho}{\gamma^*}\right)^2\right) &= \rho^{h+1}\frac{\rho^2(\gamma^*)^{h+1}\left((\gamma^*)^h(\gamma^* - b_0) - \rho^{h-1}((\gamma^*)^2 - b_0\rho)\right) + (b_0\rho)^{h+1}b(\gamma^* - \rho)}{(\gamma^*)^{2*h+1}((\gamma^*)^2 - b_0\rho)} \\ &= \rho^{h+1}\frac{\rho^2(\gamma^*)^{h+1}\left(\left((\gamma^*)^{h-1} - \rho^{h-1}\right)\left((\gamma^*)^2 - b_0\right)\right) + b_0(\gamma^*)^{h-1}(1 - \gamma^*)\right)}{(\gamma^*)^{2*h+1}((\gamma^*)^2 - b_0\rho)} \\ &+ \rho^{h+1}\frac{(b_0\rho)^{h+1}b(\gamma^* - \rho)}{(\gamma^*)^{2*h+1}((\gamma^*)^2 - b_0\rho)} \geq 0 \end{aligned} \tag{C.20}$$

where the last inequality follows from the fact that $(\gamma^*)^2 > b_0$ and $(\gamma^*)^2 > b_0\rho$. Note that $L_h\left(\left(\frac{\rho}{\gamma^*}\right)^2\right) = 0$ if and only if $\gamma^* = \rho$.

In the case of mis-specified forecasting rules absent myopia, it is easy to see from the expression for $L_h(1)$ above that $L_h(1) = 0$ if and only if $\gamma^* = \rho$, which can be achieved only if $\rho = 0$.

## C.8 Proposition 7

As before,

$$\begin{aligned}
\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+h} &= a\hat{mc}_{t+h} + b\hat{\pi}_{t+h-1} - n^h(\gamma^*)^{h+1}\hat{\pi}_{t-1} \\
&= a\rho^h \sum_{l=0}^{h-1} \left(\frac{b}{\rho}\right)^l \hat{mc}_t + b^h \hat{\pi}_t - n^h(\gamma^*)^{h+1}\hat{\pi}_{t-1} + \boldsymbol{\varepsilon_{t+1:t+h}} \\
&= a\rho^h \sum_{l=0}^{h-1} \left(\frac{b}{\rho}\right)^l \hat{mc}_t + b^h \hat{\pi}_t - \frac{n^h(\gamma^*)^{h+1}}{b}(\hat{\pi}_t - a\hat{mc}_t) + \boldsymbol{\varepsilon_{t+1:t+h}} \\
&= \underbrace{\frac{b^{h+1} - n^h(\gamma^*)^{h+1}}{b}}_{M_h} \hat{\pi}_t + error_{t,t+h}
\end{aligned} \tag{C.21}$$

where $\boldsymbol{\varepsilon_{t+1:t+h}}$ is a function of $\varepsilon_{t+1}, ..., \varepsilon_{t+h}$, and $error_{t,t+h} = a\left[\rho^h \sum_{l=0}^{h-1}\left(\frac{b}{\rho}\right)^l + \frac{n^h(\gamma^*)^{h+1}}{b}\right]\hat{mc}_t + \boldsymbol{\varepsilon_{t+1:t+h}}$. It is straightforward to see that $b^{h+1} < n^h(\gamma^*)^{h+1}$ for any $h \geq 1$. Suppose we try to estimate

$$\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+h} = c + m_h \hat{\pi}_t + error_{t,t+h} \tag{C.22}$$

using OLS. We then have that $\hat{m}_h^{OLS} = M_h + \underbrace{a\left[\rho^h \sum_{l=0}^{h-1}\left(\frac{b}{\rho}\right)^l + \frac{n^h(\gamma^*)^{h+1}}{b}\right]\frac{\mathbb{E}(\hat{\pi}_t \hat{mc}_t)}{\mathbb{E}(\hat{\pi}_t^2)}}_{(+)}$. Hence, $M_h = m_h^{OLS} - a\left[\rho^h \sum_{l=0}^{h-1}\left(\frac{b}{\rho}\right)^l + \frac{n^h(\gamma^*)^{h+1}}{b}\right]\frac{\mathbb{E}(\hat{\pi}_t \hat{mc}_t)}{\mathbb{E}(\hat{\pi}_t^2)} < m_h^{OLS}$. Kohlhas and Walther (2021) show that the OLS estimate of $m_h$ for the annual CPI inflation forecast in the US SPF data is significantly negative. Therefore, the true estimate of $M_h$ implied by a combination of mis-specified forecasts and myopia is even more negative than the OLS estimate.

# D 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.

# E Forecast Errors for More Complicated Data-Generating Processes

The purpose of this section is to get a sense of the variables one should control for when regressing ex-post forecast errors on current inflation or current and past inflation forecasts, such that the omitted variable bias is limited as much as possible. Hence, I think of a simpler, non-time-varying process describing the dynamics of the state vector, relative to the one in (35), i.e., let

$$S_t = D_1 S_{t-1} + D_2 \mathcal{E}_t \tag{E.1}$$

where $S_t$ and $\mathcal{E}_t$ are as defined in the main text. Hence, the $h$-period-ahead forecast error is given by

$$S_{t+h} - \widetilde{\mathbb{E}}_t S_{t+h} = D_1^h S_t - n^h \Gamma^{h+1} S_{t-1} + \mathcal{E}_{t+1:t+h} \tag{E.2}$$

where $\mathcal{E}_{t+1:t+h}$ is a function of the i.i.d. demand, cost-push, and monetary innovations from period $(t+1)$ to $(t+h)$, and $\Gamma$ is a diagonal matrix. Isolating the inflation forecast error yields

$$\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t \hat{\pi}_{t+h} = (D_1^h)_\pi S_t - n^h \gamma_\pi^{h+1} \hat{\pi}_{t-1} + \mathcal{E}_{t+1:t+h} \tag{E.3}$$

where $(D_1^h)_\pi$ is the second row of matrix $D_1^h$ (since inflation is positioned second in the state vector $S_t$). We have that

$$(D_1^h)_\pi S_t = f(\hat{\pi}_t, \hat{x}_{t-1}, \hat{\pi}_{t-1}, \hat{R}_{t-1}, e_{t-1:t}, u_{t-1:t}, \varepsilon^v_{t-1:t}) \tag{E.4}$$

and

$$\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t\hat{\pi}_{t+h} = g(\hat{\pi}_t, \hat{x}_{t-1}, \hat{\pi}_{t-1}, \hat{R}_{t-1}, e_{t-1:t}, u_{t-1:t}, \varepsilon^v_{t-1:t}) \tag{E.5}$$

where $f(.)$ and $g(.)$ are linear functions. Therefore, to minimize biases coming from omitted variables from regressing equation (17), I control for the first lag of the output gap, inflation, and nominal rates, and the current and first lag of the demand, cost-push, and monetary shock.

To find out the relevant controls for the regression in (38) the inflation ex-ante forecast revisions can be written as

$$\widetilde{\mathbb{E}}_t\pi_{t+h} - \widetilde{\mathbb{E}}_{t-1}\pi_{t+h} = n^h\gamma_\pi^{h+1}(\hat{\pi}_{t-1} - n\gamma\hat{\pi}_{t-2}) = ff(\hat{x}_{t-2}, \hat{\pi}_{t-2}, \hat{R}_{t-2}, e_{t-2:t-1}, u_{t-2:t-1}, \varepsilon^v_{t-2:t-1}) \tag{E.6}$$

where $ff(.)$ is a linear function. Further, I rewrite the ex-post inflation forecast error as

$$\begin{aligned}\hat{\pi}_{t+h} - \widetilde{\mathbb{E}}_t\hat{\pi}_{t+h} &= gg(\hat{x}_{t-2}, \hat{x}_{t-2}, \hat{R}_{t-2}, e_{t-2:t}, u_{t-2:t}, \varepsilon^v_{t-2:t}) \\ &= gg(\widetilde{\mathbb{E}}_t\pi_{t+h} - \widetilde{\mathbb{E}}_{t-1}\pi_{t+h}, \hat{x}_{t-2}, \hat{R}_{t-2}, e_{t-2:t}, u_{t-2:t}, \varepsilon^v_{t-2:t})\end{aligned} \tag{E.7}$$

where $gg(.)$ is a linear function. Therefore, to minimize biases coming from omitted variables in regression (38), I control for the second lag of the output gap and nominal rates, as well as for the current, first, and second lags of the demand, cost-push and monetary shock.

Table 8 exhibits a full report on the estimates of (38) and (39) after the added control variables.

| | Ex-ante forecast revisions | Current inflation | |
|---|---|---|---|
| $\hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_t\hat{\pi}_{t+3}$ | (1) | (2) | (3) |
| $\widetilde{\mathbb{E}}_t\hat{\pi}_{t+3} - \widetilde{\mathbb{E}}_{t-1}\hat{\pi}_{t+3}$ | 2.442*** | | |
| | (0.261) | | |
| $\hat{\pi}_t$ | | 0.267 | -1.002** |
| | | (0.441) | (0.456) |
| $\hat{x}_{t-1}$ | | 0.220*** | -0.021 |
| | | (0.073) | (0.057) |
| $\hat{\pi}_{t-1}$ | | 0.208 | 1.102*** |
| | | (0.415) | (0.369) |
| $\hat{R}_{t-1}$ | | -0.089*** | -0.044 |
| | | (0.032) | (0.038) |
| $\hat{x}_{t-2}$ | 0.212*** | | |
| | (0.079) | | |
| $\hat{R}_{t-2}$ | -0.114*** | | |
| | (0.024) | | |
| $\hat{e}_t$ | 0.125*** | 0.062 | 0.202*** |
| | (0.039) | (0.057) | (0.056) |
| $\hat{u}_t$ | 1.740*** | 0.967 | 4.294*** |
| | (0.175) | (1.358) | (1.328) |
| $\varepsilon_t^v$ | 0.708 | 0.545 | 0.622** |
| | (0.454) | (0.371) | (0.254) |
| $\hat{e}_{t-1}$ | 0.021 | -0.075 | 0.015 |
| | (0.037) | (0.056) | (0.051) |
| $\hat{u}_{t-1}$ | -0.056 | -0.127 | -0.343* |
| | (0.199) | (0.250) | (0.199) |
| $\varepsilon_{t-1}^v$ | -0.212 | 0.308 | -0.085 |
| | (0.224) | (0.305) | (0.188) |
| $\hat{e}_{t-2}$ | -0.084 | | |
| | (0.062) | | |
| $\hat{u}_{t-2}$ | -0.003 | | |
| | (0.162) | | |
| $\varepsilon_{t-2}^v$ | 0.647* | | |
| | (0.358) | | |
| Constant | 0.523*** | 0.362* | 0.749*** |
| | (0.164) | (0.192) | (0.163) |
| R-squared | 0.765 | 0.757 | 0.612 |
| Time period | 1966:Q1 - 2018:Q3 | 1961:Q3 - 2018:Q3 | 1981:Q1 - 2018:Q3 |

Newey-West standard errors in parentheses.
*** Robust at 1%, ** Robust at 5%, * Robust at 10%.

Table 8: **Estimates of regressions in (38) and (39) on simulated forecasting data w/ added controls.**

# References


Klaus Adam. Experimental evidence of the persistence of output and inflation. *The Economic Journal*, 117(520):603–636, April 2007. ISSN 0013-0133. URL https://doi.org/10.1111/j.1468-0297.2007.02043.x.

George-Marios Angeletos and Zhen Huo. Myopia and anchoring. *American Economic Review*, 111 (4):1166–1200, April 2021. ISSN 0002-8282. URL https://doi.org/10.1257/aer.20191436.

George-Marios Angeletos, Zhen Huo, and Karthik A. Sastry. Imperfect macroeconomic expectations: Evidence and theory. *NBER Macroeconomics Annual*, 35:1–86, May 2021. ISSN 0889-3365. URL https://doi.org/10.1086/712313.

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

Francesco Bianchi. Regime switches, agents' beliefs, and post-World War II US macroeconomic dynamics. *The Review of Economic Studies*, 80(2):463–490, April 2013. ISSN 0034-6527. URL https://doi.org/10.1093/restud/rds032.

Mark Bils and Peter J. Klenow. Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112(5):947–985, October 2004. ISSN 0022-3808. URL https://doi.org/10.1086/422559.

Pedro Bordalo, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. Overreaction in macroeconomic expectations. *American Economic Review*, 110(9):2748–2782, September 2020. ISSN 0002-8282. URL https://doi.org/10.1257/aer.20181219.

Stephen P. Brooks and Andrew Gelman. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455, December 1998. ISSN 1061-8600. URL https://doi.org/10.1080/10618600.1998.10474787.

Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45, February 2005. ISSN 0022-3808. URL https://doi.org/10.1086/426038.

Olivier Coibion and Yuriy Gorodnichenko. Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–2678, August 2015. ISSN 0002-8282. URL https://doi.org/10.1257/aer.20110306.

Olivier Coibion, Yuriy Gorodnichenko, and Rupal Kamdar. The formation of expectations, inflation, and the Phillips curve. *Journal of Economic Literature*, 56(4):1447–1491, December 2018. ISSN 0022-0515. URL https://doi.org/10.1257/jel.20171300.

Stephen J. Cole and Fabio Milani. The misspecification of expectations in New Keynesian models: A DSGE-VAR approach. *Macroeconomic Dynamics*, 23(3):974–1007, April 2019. ISSN 1365-1005, 1469-8056. URL https://doi.org/10.1017/S1365100517000104.

Marco Del Negro and Stefano Eusepi. Fitting observed inflation expectations. *Journal of Economic Dynamics and Control*, 35(12):2105–2131, December 2011. ISSN 0165-1889. URL https://doi.org/10.1016/j.jedc.2011.04.005.

Stefano Eusepi and Bruce Preston. The science of monetary policy: An imperfect knowledge perspective. *Journal of Economic Literature*, 56(1):3–59, March 2018. ISSN 0022-0515. URL https://doi.org/10.1257/jel.20160889.

George Evans, Cars Hommes, Bruce McGough, and Isabelle Salle. Are long-horizon expectations (de-)stabilizing? Theory and experiments. Staff working paper 19-27, Bank of Canada, August 2019. URL https://ideas.repec.org/p/bca/bocawp/19-27.html.

Andreas Fuster, David Laibson, and Brock Mendel. Natural expectations and macroeconomic dynamics. *Journal of Economic Perspectives*, 24(4):67–84, December 2010. ISSN 0895-3309. URL https://doi.org/10.1257/jep.24.4.67.

Andreas Fuster, Benjamin Hebert, and David Laibson. Natural expectations, macroeconomic dynamics and asset pricing. *NBER Macroeconomics Annual*, 26:1–48, 2012. URL https://doi.org/10.1086/663989.

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

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

Peter Ganong and Pascal Noel. Consumer spending during unemployment: Positive and normative implications. *American Economic Review*, 109(7):2383–2424, July 2019. URL https://doi.org/10.1257/aer.20170537.

Eric Gaus and Christopher G. Gibbs. Expectations and the empirical fit of DSGE models. Technical report, 2018.

John Geweke. Simulation methods for model criticism and robustness analysis. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, editors, *Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting*, volume 6, pages 275–299. Oxford University Press, October 1999. ISBN 978-0-19-850485-6.

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

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

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

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

Cars H. Hommes and Mei Zhu. Behavioral learning equilibria. *Journal of Economic Theory*, 150: 778–814, March 2014. ISSN 0022-0531. URL https://doi.org/10.1016/j.jet.2013.09.002.

Cars H. Hommes, Kostas Mavromatis, Tolga Ozden, and Mei Zhu. Behavioral learning equilibria in the New Keynesian model. Technical Report 654, Netherlands Central Bank, Research Department, September 2019.

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

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

Bartosz Maćkowiak and Mirko Wiederholt. Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803, June 2009. URL https://doi.org/10.1257/aer.99.3.769.

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

N. Gregory Mankiw and Ricardo Reis. Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve. *The Quarterly Journal of Economics*, 117(4):1295–1328, November 2002. URL https://doi.org/10.1162/003355302320935034.

Fabio Milani. A Bayesian DSGE model with infinite-horizon learning: Do 'mechanical' sources of persistence become superfluous? *International Journal of Central Banking*, 2(3):87–106, September 2006.

Fabio Milani. Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics*, 54(7):2065–2082, October 2007. URL https://doi.org/10.1016/j.jmoneco.2006.11.007.

Arturo Ormeño and Krisztina Molnár. 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, 2015. URL https://doi.org/10.1111/jmcb.12224.

Luba Petersen. Do expectations and decisions respond to monetary policy? *Journal of Economic Studies*, 42(6):972–1004, January 2015. URL https://doi.org/10.1108/JES-08-2015-0149.

Bruce Preston. Learning about monetary policy rules when long-horizon expectations matter. *International Journal of Central Banking*, 1(2):81–126, September 2005.

Ricardo Reis. Inattentive producers. *The Review of Economic Studies*, 73(3):793–821, July 2006. ISSN 0034-6527. URL https://doi.org/10.1111/j.1467-937X.2006.00396.x.

Yuliya Rychalovska. 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, December 2016. URL https://doi.org/10.1016/j.jedc.2016.09.014.

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

Sergey Slobodyan and Raf Wouters. Learning in a medium-scale DSGE model with expectations based on small forecasting models. *American Economic Journal: Macroeconomics*, 4(2):65–101, April 2012a. URL https://doi.org/10.1257/mac.4.2.65.

Sergey Slobodyan and Raf Wouters. Learning in an estimated medium-scale DSGE model. *Journal of Economic Dynamics and Control*, 36(1):26–46, January 2012b. URL https://doi.org/10.1016/j.jedc.2011.01.016.

Frank Smets and Raf Wouters. An estimated dynamic stochastic general equilibrium model of the Euro area. *Journal of the European Economic Association*, 1(5):1123–1175, September 2003. URL https://doi.org/10.1162/154247603770383415.

Frank Smets and Rafael Wouters. Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review*, 97(3):586–606, June 2007. URL https://doi.org/10.1257/aer.97.3.586.

James H. Stock and Mark W. Watson. Why has US inflation become harder to forecast? *Journal of Money, Credit and Banking*, 39(s1):3–33, 2007. URL https://doi.org/10.1111/j.1538-4616.2007.00014.x.

Amos Tversky and Daniel Kahneman. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2):207–232, September 1973. URL https://doi.org/10.1016/0010-0285(73)90033-9.

Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, September 1974. URL https://doi.org/10.1126/science.185.4157.1124.

Michael Woodford. Imperfect common knowledge and the effects of monetary policy. In Philippe Aghion, Roman Frydman, Joseph E. Stiglitz, and Michael Woodford, editors, *Knowledge, Information, and Expectations in Modern Macroeconomics, in Honor of Edmund S. Phelps*, pages 25–58. Princeton University Press, 2003a. ISBN 978-0-691-22393-3.

Michael Woodford. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton, N.J., 2003b. ISBN 978-0-691-01049-6.