What Do (and Don't) Forecasters Know About U.S. Inflation?*

Jane Ryngaert†
University of Notre Dame

July 2022

Abstract

This paper contributes to and extends our current understanding of information frictions in expectations. I propose a framework for estimating noisy information using individual forecasts. I extend this framework to incorporate misperceptions about the persistence of the underlying economic process. Applying this to the U.S. professional inflation forecasts suggests forecaster overestimation of inflation persistence in addition to the presence of noisy signals. Using a structural model that incorporates both noisy signals and misperceptions of persistence, I quantify the relative importance of each channel in expectations formation. Even among professional forecasters, multiple forces generate economically significant deviations from full information.

JEL Classification: E31, D83, D84

Keywords: inflation dynamics, inflation expectations, noisy information, parameter misperception, inflation persistence

* I thank Olivier Coibion, Saroj Bhattarai, and Tara Sinclair for their invaluable guidance and support. I also benefited from conversations with Andrew Glover, Christoph Boehm, Garrett Hagemann, Jeffrey Campbell, Spencer Krane, Leonardo Melosi, Allin Cottrell, Sandeep Mazumder, John Dalton, and Amanda Griffith as well as with many seminar participants from the University of Texas, The George Washington University, and the Federal Reserve Bank of Chicago. I further thank two anonymous reviewers.

† Assistant Professor, University of Notre Dame (e-mail: jryngaer@nd.edu)

1 Introduction

Expectations are a ubiquitous feature of macroeconomic models. Economic expectations and particularly expectations of the inflation rate, affect all manner of economic decisions. Firms must anticipate future costs and prices in setting their own prices and households must consider the path of future prices when planning the timing of purchases and borrowing. As the link between expectations and actions is so pervasive, expectations inevitably have consequences for economic dynamics. For this reason, there is a growing interest in understanding how economic agents form their expectations and the constraints they face in doing so.

Economists are increasingly using models relaxing the full information assumption of rational expectation models and limiting forecaster access to information. Agents in the sticky information model of Mankiw and Reis (2002) must pay fixed costs to obtain new information and therefore do so only periodically. Deviations from full information rational expectations in an agent's forecast come from the fact that, in any period that she does not update, she bases her entire expectation on outdated information. In a second class of models termed noisy information models, Sims (2003), Woodford (2002), and Mackowiak and Wiederholt (2009) restrict the agent's ability to observe the variable she is trying to forecast. Observing only signals about the fundamental rather than the fundamental itself, the forecaster engages in optimal signal processing and at least a portion of her new expectation is formed with dated information. While these models introduce constraints into the expectations formation process, they take for granted that agents understand the structure of the economy and therefore form expectations that, though partly out-of-date, are model-consistent. If forecasters face constraints in information collection and processing, it is reasonable to think that they also face difficulty making inferences about underlying economic structures.[1] Accordingly, this paper looks at these two issues jointly

[1]Thomas Sargent noted the implausibility of rational expectations with the following critque: "rational

with the ultimate goal of estimating the size and importance of each channel. I find that both noisy signals about inflation and misperceptions of the structural parameters governing inflation dynamics lead to economically significant deviations from full information rational expectations.

The first contribution of this paper is to develop a new framework for estimating noisy information using individual professional forecasts. Information frictions in the noisy information model derive from the inability of agents to observe inflation in real time and the resulting individual-specific error in observations of inflation. Recent approaches to estimating these frictions rely on mean forecasts, thus averaging across these individual signals and canceling out the variation that drives the need for signal processing. My approach instead utilizes these idiosyncratic signals and exploits both individual and time variation in forecast errors. Estimation using this approach results in a substantial efficiency gain over comparable estimation on aggregate forecasts. Applying this framework to professional forecasts of U.S. inflation since the early 1980s, I find estimates of noisy information implying that forecasters weigh new signals less than they weigh prior beliefs in forming their expectations. As they form more than half of their expectation with dated information, observation constraints constitute a relevant source of frictions that will affect macroeconomic dynamics. I then show that the individual approach to estimation produces economically different results than the more commonly used approach focusing on mean forecasts. The baseline noisy information model is unlikely to account for this difference across estimation at the individual and mean forecast levels.

In light of these results, a second contribution is to introduce an additional potential source of information frictions to the noisy information model. An extension of the noisy information model allows the agents to incorrectly perceive the structural parameters of the inflation process.[2] I derive the predicted path of forecast errors given both frictions:

expectations models impute much more knowledge to the agents within the model ... than is possessed by an econometrician, who faces estimation and inference problems that the agents in the model have somehow solved." Sargent (1993) noisy signals and mistaken parameters. This provides a simple framework that can simultaneously quantify the effect of noisy information as well as the magnitude and direction of forecaster misperception of parameters. My approach shows that since the early 1980s, forecasters have on average overestimated inflation persistence. This overestimation can account for the difference in the individual and aggregate estimates of noisy information. I provide additional evidence from the term structure of individual professional forecasts that forecasters misperceive the parameters of the inflation process. This prediction is in line with recent literature demonstrating that forecasters perception of persistence often differs from the value implied by time series estimates, e.g. Jain (2019). We can think about this additional friction as a departure from full information over and above the presence of noisy signals. This departure can create bias in the estimation of noisy information when the estimation fails to account for it. The direction of the bias depends on the sign of the misperception and can cause us to conclude that information is more or less noisy than it actually is.

The third contribution of this paper is to build a simple structural model that can be used to quantify the relative importance of each friction in explaining the predictability of forecast errors as well as the moments seen in the data. The model can be used to simultaneously estimate inflation persistence, forecaster misperception of persistence, and the strength of the noisy information friction. These estimates support the findings of the rest of the paper that forecasters both face real-time information constraints and overestimate inflation persistence. These estimates further show that information is much more noisy than previous estimates of the noisy information friction (that do not take the misperception friction into account) suggest.[3] The structural estimates imply that

[2]This can be interpreted in the context of the learning literature where agents form inferences about structural parameters. My paper looks primarily at the effects of misperception when it is present, not at the ways in which forecasters learn about structural parameters.

[3]See for example Coibion and Gorodnichenko (2015), Coibion and Gorodnichenko (2012), and Dovern et al. (2014).

forecasters base roughly 25 percent of their expectations on new information, leaving 75 percent of the expectation to be formed with prior information. While forecasters face economically relevant observation friction, the overestimation of persistence creates another relevant friction for expectations formation as forecasters will project the state forward using the wrong transition equation. Forecasters will make the wrong projections about the future path of inflation and overstate the longevity of shocks to inflation.

Jointly, my results suggest that, even for professional forecasters, we need a wider set of models to explain the formation of beliefs than is currently utilized. Professional forecasters display expectations consistent with multiple forms of frictions, both in observing the true value of inflation and in understanding the parameters governing inflation dynamics. To assess the effects of either of these types of frictions, one needs to consider them together.

This paper contributes to an empirical literature that attempts to assess the degree to which information is imperfect or rigid. A number of papers measure the noisiness or stickiness in information using predictability in forecast errors and revisions. These include Coibion and Gorodnichenko (2012, 2015), Bürgi (2017), Dovern et al. (2014), and Andrade and Le Bihan (2013). Notably, much of the previous work in estimating information frictions has focused on aggregate forecasts rather than individual forecaster data. While aggregate forecasts and forecast errors may show departures from full information rational expectations (FIRE), these findings are not necessarily representative of individual forecasters. Pesaran and Weale (2006) indicate that forecasters may diverge in ways that offset each other when aggregated. Crowe (2010), and Pesaran and Weale (2006) cite inefficiency in the consensus forecast even when individuals act rationally, implying that microdata is preferable to aggregate data whenever possible. While Bürgi (2017) and Dovern et al. (2014) consider individual settings for estimating noisy and sticky information, my paper interprets the difference between the individual and aggregate results and argues for a second type of information friction, parameter misperception. Other attempts to measure and characterize information frictions in the expectations formation process include Andrade and Le Bihan (2013) who study professional forecasters at the European Central Bank (ECB), and Sheng, Wallen, and An (2017) who perform nonparametric analyses on forecasters from both the SPF and Consensus Economics.

I also relate to two papers that discuss both individual and aggregate forecasts in terms of the over- and under-reaction to new information, Bordalo et al. (2020) and Fuhrer (2018). Bordalo et al. (2020) notes predictability of forecast errors with respect to forecast revisions at both the individual and aggregate levels. They argue that the sign of the predictability of aggregate forecast errors points to under-reaction to new information. This under-reaction and the predictability of forecast errors with respect to their own lags for both aggregate and individual forecast errors noted in this paper have the same source - the weighting of older beliefs in addition to new signals as a part of rational signal extraction. Bordalo et al. (2020) argue that the same analysis on individual forecast errors indicates the overreaction of expectations to new information and build a model of diagnostic expectations in which agents overestimate the probability of states that have become more likely. We could think about overestimated persistence as a form of overreaction to shocks. In a persistent series, a shock to the current level will be a part of the future level, but overestimated persistence will cause the forecaster to move expectations of the future series too far in the direction of this shock. An important difference between my model and diagnostic expectations is that diagnostic expectations allows forecasters to overreact to signals about the current state, whereas my model only allows for overreaction in forecasts. This paper cannot reject the presence of diagnostic expectations, but the results in my paper cannot be explained without noisy information or parameter misperception, even if expectations are also diagnostic. Furthermore, the overestimation of inflation has implications for the interpretation of Bordalo et al. 2020's analysis that are described in Appendix C. Fuhrer (2018) shows that expectations under-react to new information at both the individual and aggregate levels - as this paper does -

but credits this to forecaster inefficiency rather than information constraints. He examines the predictability of forecast revisions with respect to several measures of an individual's own forecast history as well as the past consensus values whereas the current paper looks at the persistence of forecast errors. Fuhrer (2018) argues against noisy information as a possible explanation, stating that individual forecast errors should not be predictable given the information in an individual's information set. In my model, forecasters do not observe their past forecast errors and persistence in forecast errors would be consistent with rational forecasters.

Andrade et al. (2016) is thematically similar to my paper as it combines multiple frictions in forecaster observation of the true state of the economy and the need to disentangle temporary and permanent disturbances to explain features of forecaster data. However, rather than looking at persistence in forecast errors as I do, Andrade et al. (2016) matches the term structure of forecaster disagreement in the Blue Chip Financial Forecasts. They explain the term structure of disagreement in CPI inflation as resulting from infrequent updating of information (consistent with my finding of a high degree of noise in expectations) and long term shocks that are relatively less volatile than short term shocks. My result could be interpreted as forecasters mistaking transitory components for long-lived components and placing too much weight on them in projections into the future. The noisy information model predicts that disagreement should decline with forecasting horizons as forecasts are formed by projecting the current state estimate forward according to the perceived transition equation. While the model in this paper cannot fully match the term structure of disagreement, the overestimation of persistence would flatten this term structure relative to the counterfactual case of no misperception. This moves the the pattern of near to mid-term disagreement towards the term structure observed in Andrade et al. (2016).

This paper also contributes to a literature arguing that lack of knowledge of structural parameters may constitute a relevant constraint for expectations formation and therefore economic dynamics. Orphanides and Williams (2004), for example, introduce limitations in forecaster understanding of parameters by having economic forecasters engage in perpetual learning about parameters with limited memory. They further argue that the assumption of full information about structural parameters can cause policy makers to choose the incorrect optimal policy with negative impacts on the economy. Other papers that limit forecaster knowledge of underlying structures governing dynamics and therefore requiring forecasters to learn include Milani (2007), Sargent (1993), and Cogley, Primiceri, and Sargent (2010).

Lastly, this paper contributes a new literature on forecasters' perceptions of inflation persistence. To the best of my knowledge, currently the only other paper that attempts to estimate perceived persistence using individual forecaster data is Jain (2019). Jain utilizes the term structure of forecaster beliefs in the Survey of Professional Forecasters CPI series and formulates reduced form estimates of perceived persistence. Notably, her estimates show that most forecasters perceive a level of persistence substantially different from that of a random walk and lower than estimates derived from time series inference. I find contrasting evidence in favor of the overestimation of persistence. A notable difference between my paper and hers is in the approach - she uses forecast revisions while I use the forecast and forecast errors. Assuming the same data-generating process for inflation, this difference alone would not explain the diverging results. More likely the difference in results is explained by the different handling of the constant term in the inflation transition equation. I choose to estimate this term directly, while Jain prefers to allow the constant term to cancel out by using revisions. This difference in approach highlights additional differences between my paper and hers. I assume an AR(1) specification of inflation with a constant term and assume my estimates recover the persistence parameter from the AR(1). This parameter and its perceived value are assumed to be constant across forecasters and estimated in a pooled regression. Jain estimates the perceived persistence of a state variable, x, which is a scalar index of all the variables a forecaster uses to forecast inflation. This variable is allowed to vary across forecasters and perceived persistence is therefore calculated separately for each forecaster.

The remainder of the paper proceeds as follows. In Section 2, I describe the simple noisy information model. Section 3 discusses the data and the initial results. Section 4 describes the model with an additional source of frictions and presents evidence for the underestimation of persistence. Section 5 describes the structural model used in simulations as well as its application to estimation. Section 6 discusses other potential explanations for the differences in the individual and aggregate estimates. Section 7 examines extensions to the model, including time variation in parameter misperception, modeling a different time series process for inflation, multivariable information, forecaster observation of realizations of inflation, and the noisy information model with public signals. Section 8 concludes.

2 Basic Noisy Information Model

Following Woodford (2002), Sims (2003) and Coibion and Gorodnichenko (2012), I present a noisy information model that generates predictability in individual forecast errors. In a noisy information context, a forecaster's difficulty in forming beliefs about the future stems from her inability to observe the present state clearly. Her signal extraction problem leads to additional persistence in her forecasts (above and beyond the persistence in the process being forecasted), causing the serial correlation of forecast errors over time.

Allow inflation to evolve according to an AR(1) with constant, μ.[4] Innovations to inflation arrive each quarter and are indicated by w_t, a Gaussian white noise term with variance σ_w^2. The long run mean of inflation is $\frac{\mu}{1-\rho}$.

[4] I use an AR(1) process for parsimony and ready comparison with related literature. This analysis may be extended to consider an AR(2) process with results that are compatible with the parameter misperception. I talk more in detail about this in Section 7.2

$$\pi_{t,t-1} = \mu + \rho\pi_{t-1,t-2} + w_t \tag{1}$$

Each period, agents receive a signal equivalent to the true value of inflation in the present period plus some individual-specific noise component, $v_t(i) \sim N(0, \sigma_v^2)$. They further know the structure of the economy and therefore know ρ and μ without error.

$$z_t(i) = \pi_{t,t-1} + v_t(i) \tag{2}$$

A forecaster combines her new signal, $z_t(i)$, with her beliefs about present inflation from the previous period, weighting the signal with the gain from the Kalman filter, k. The gain is derived optimally from the parameters of the process and measurement equations - ρ, σ_v^2, and σ_w^2 - as well as forecaster uncertainty about the state. The Kalman gain is defined as:

$$k = \frac{\Psi}{\Psi + \sigma_v^2}$$

where $\Psi = \rho^2 U^- + \sigma_w^2$ and U^- is the agent's uncertainty about the state before she receives her signal. The forecaster places the remaining weight, (1-k) , on her expectation of $\pi_{t,t-1}$ in time *t-1.* The result is the agent's optimal nowcast[5]:

$$\pi_{t,t-1|t}(i) = kz_t(i) + (1-k)\pi_{t,t-1|t-1}(i). \tag{3}$$

I adopt the following notation for agent forecasts: $\pi_{t+h,t|t-\tau}(i)$, with $\tau \geq 0$, is agent i's forecast of inflation from time t to t+h made with information available at time $t-\tau$. The corresponding average, or aggregate, forecast is denoted $\overline{\pi}_{t+h,t|t-\tau}$.

In Equation 3, (1-k) represents the percentage of the new expectation based on prior

[5]A nowcast is the agent's belief about current inflation formed in the current period.

information and we can interpret this as the degree of imperfection in information.[6] The Kalman gain is a measure of how much a forecaster can trust her signal. The more credible her signal is, the more weight she will assign to it in updating her expectations. The gain is increasing in process persistence, ρ, and the process noise, σ_w^2, and decreasing in the agents' noise variance σ_v^2. For a given ρ and σ_w^2, an increase in σ_v^2 will make the agent's signal noisier and less informative. Accordingly, the agent will give the signal a lower weight in her expectation. On the other hand, given a constant ρ and σ_v^2, a larger value of σ_w^2 means that a larger amount of the noise in the signal is attributable to the inflation innovation rather than signal noise, making the signal relatively more credible. Holding σ_v^2 and σ_w^2 constant, an increase in ρ means that the signal is relatively more dependent on lagged inflation (rather than noise) and deserves a greater weight.

Given a nowcast, $\pi_{t,t-1|t}(i)$, the agent will form forecasts by projecting the present belief forward according to the transition equation in Equation 1.

$$\begin{aligned}\pi_{t+1,t|t}(i) &= \mu + \rho\pi_{t,t-1|t}(i) \\ &= \mu + k\rho\pi_{t,t-1} + k\rho v_t(i) + \rho(1-k)\pi_{t,t-1|t-1}(i)\end{aligned} \tag{4}$$

To form the agent's forecast error, I subtract both sides of the above equation from $\pi_{t+1,t}$. The ex-post forecast error of agent i can then be written as:

$$FE_{t+1,t|t}(i) = \rho(1-k)FE_{t,t-1|t-1}(i) + w_{t+1} - k\rho v_t(i). \tag{5}$$

In this expression, the predictability of forecast errors relies on the existence of imperfect information and signals. If agents receive full information, $k = 1$, and forecast errors will not be serially correlated over time. Additionally, under full information, $\sigma_v^2 = 0$, and $v_t(i) = 0$, implying that the forecast error will collapse to $FE_{t+1,t|t}(i) = w_{t+1}$, or the full information rational expectations error. In this case, forecast errors arise only from

[6] In the perfect information model, the agent updates her expectation completely to the new information available to her, setting k=1.

the inability to observe future innovations or shocks to inflation, w_{t+1}, and not from constraints in observing past and current shocks. When agent signals are imperfect, $\sigma_v^2 > 0$ and $v_t(i) \neq 0$, meaning signals include idiosyncratic noise which obfuscates the true value of the current state. This introduces predictability in forecast errors as forecasters in this environment face constraints in observing past, present, and future innovations to the inflation process.

Taking the average of Equation 5 across agents gives the following relationship between consensus forecasts and their lags.

$$\overline{FE}_{t+1,t|t} = \rho(1-k)\overline{FE}_{t,t-1|t-1} + w_{t+1} \tag{6}$$

This is the standard approach used to estimate information rigidities, e.g. Coibion and Gorodnichenko (2012). The primary difference between this equation and Equation 5, aside from the use of aggregate forecast errors rather than individual forecast errors, is the construction of the error term. The signal noise term does not appear in the aggregate as it averages out across agents. Each of these equations can be estimated via the following reduced form equations.

$$FE_{t+1,t|t}(i) = \beta_0 + \beta_1 FE_{t,t-1|t-1}(i) + \varepsilon_t(i) \tag{7}$$

$$\overline{FE}_{t+1,t|t}(i) = \beta_0 + \beta_1 \overline{FE}_{t,t-1|t-1} + \varepsilon_t \tag{8}$$

The model predicts that, for both Equations 7 and 8, $\beta_0 = 0$ and, so long as the process persistence, ρ, is positive, that recovering a $\beta_1 > 0$ implies the presence of noisy information. Under full information, this equation will recover $\beta_1 = 0$ and we will not observe any persistence in forecast errors. A value of β_1 significantly greater than 0 leads us to reject the null hypothesis of full information.[7] Given the existence of imperfect signals [8] and positive process persistence, β_1 will not uniquely identify the degree of persistence or information rigidity, but a mixture of the two.[9] Because forecasters receive noisy signals of inflation and never observe the realization of inflation itself the correlation between forecast errors in t and $t+1$ is observable to the econometrician but not to the forecasters themselves.[10] This means that we could see correlation in individual forecast errors even among rational forecasters. This approach differs from Bordalo et al. (2020) in that it relates forecast errors to their own lags rather than to their revisions. As forecast revisions are observable to forecasters, the predictability of forecast errors as in Bordalo et al. (2020) constitutes a deviation from rational updating. Fuhrer (2018) considers forecast revisions as a dependent variable with a variety of independent variables.

2.1 Gains of the Individual Approach

To illustrate the gains from the individual approach over the aggregate approach, I simulate data to match the inflation process and forecasters' filtering problem in the basic noisy information model. I further calibrate the parameters of the model to standard values $\rho = 0.39$ and $k = 0.31$, though adjusting these parameters does not substantially change the gains from the individual approach. The value of σ_w^2 is set to 0.5 and σ_v^2 is set to 1.24 such that, given $\sigma_w^2 = 0.5$ and $\rho = 0.9$, $k = 0.5$ is optimal. We can adjust the value of σ_w^2 without significant change in the results as, in order to keep $k = 0.31$ as the optimal Kalman gain, we must adjust σ_v^2 such that the relative relationship between the

[7] Nordhaus (1987) proposed an equivalent specification and interpreted the null hypothesis $\beta_1 = 0$ as a test for forecaster rationality. In his model, only deviations from rational or optimal actions would create serial correlation in forecast errors over time. See also Keane and Runkle (1989) and Bonham and Cohen (1995). Coibion and Gorodnichenko (2015) show however, that for the aggregate forecast errors, we may reject the null $\beta_1 = 0$ with rational agents acting optimally under limited information.

[8] This is implied by the multiple tests finding positive values for these and related coefficients. See Coibion and Gorodnichenko (2012, 2015) and Dovern et al. (2014) for examples.

[9] I address this in Section 5.

[10] The assumption that forecasters do not observe ex-post realizations of inflation is relaxed in Section 7.4.

two variances is the same.

The simulated data matches the length of the sample period ($F = 154$) with a short 100-period burn-in period and contains $N = 184$ forecasters, as does the sample. I generate one-quarter ahead forecast errors for each forecaster in each time period. Following the burn-in period, the simulated data consists of an $N \times F$ matrix of forecast errors. As forecasters do not appear in sample for all time periods, I match each simulated forecaster to a forecaster from the sample data and populate only the periods in which the sample forecaster appeared. I then calculate mean forecasts for each period using only the forecasters that appear in that period. This allows me to match the number of observations in the data for both the pooled and time series approaches. I run the pooled and time series regressions estimating Equations 7 and 8 and record the coefficients for 1000 rounds of simulated data.

Table 1 shows the gains from my approach over the aggregate regression. Across the 1000 rounds of the simulation, both the mean and aggregate estimates recover approximately the correct coefficient on lagged forecast errors. The individual approach leads to a 29 percent reduction in the standard deviation of the 1000 estimates and a 27 percent reduction in the interquartile range of these estimates.

2.2 Forecast Errors at Horizons Greater than 1

We can also consider forecasts over longer horizons. Let h denote the forecasting horizon so that when h=1, a forecast covers inflation over a quarter. When h=2, however, a forecast is a projection over six months. Forecasts at h=3 provide a forecast of inflation over the next nine months and at h=4, over the next year. Longer horizon forecast errors must be regressed on lagged forecast errors at a lag length matching the horizon. If the lag of the forecast error is less than the horizon length, overlap in the shocks across forecast errors and lagged forecast errors will create serial correlation that does not come from information rigidity.[11]

I show the derivation of the predicted path of forecast errors for h=2 in Appendix A. Forecasts at horizons longer than two will share the same properties.

$$
\begin{aligned}
FE_{t+2,t|t}(i) &= \rho^2(1-k)^2 FE_{t,t-2|t-2}(i) - \rho(1-k)(1+\rho k)w_t + (1+\rho)w_{t+1} + w_{t+2} \\
&\quad - (1+\rho)\rho^2(1-k)kv_{t-1}(i) - \rho(1+\rho)kv_t(i)
\end{aligned}
$$

As forecasters access multiple signals and chances to update their expectations across longer-horizon forecasts, the coefficient on forecast errors shrinks exponentially with horizon length. For horizon h, the coefficient is $\rho^h(1-k)^h\ \forall h$. Forecast errors that are separated by a greater length of time should be less serially correlated than forecast errors in adjacent periods. In each period between these forecast errors, forecasters form new expectations with new information; this causes past beliefs to fade slowly from expectations.

The term w_t appears in both the error term and in $FE_{t,t-2|t-2}(i)$. [12] This endogeneity will be present as long as $k < 1$, that is when information imperfections are present. One can address this issue by adding a time fixed effect to control for w_t and remove the endogenous component of the error. This fixed effect will be horizon-specific as the effect of w_t on $FE_{t+h,t|t}(i)$ depends on the horizon.

For $h > 1$, the appropriate estimation equation is therefore,

$$FE_{t+h,t|t}(i) = \beta_0 + \beta_1 FE_{t,t-h|t-h}(i) + u_{t,h} + \varepsilon_t(i) \tag{9}$$

[11] For example, a forecast of year-ahead inflation in time t will only include shocks through time-t shocks. The realization of inflation, however, will contain shocks for four additional quarters. These shocks will enter into the year-ahead forecast error. The year-ahead forecast error, lagged by only a quarter, will include many of the same shocks in the forecast error, meaning that we would expect to see serial correlation across forecasts even under full information.

[12] See A-4. in Appendix A

The problematic term in the error is time-dependent rather than individual-dependent meaning that this term is still present and the endogeneity problem persists in the mean specification of higher-order forecast errors.[13] Moreover, as the aggregate specification is a time series, time period fixed effects cannot be used and the aggregate regression cannot be estimated by OLS for forecasting horizons greater than one. My individual framework is therefore the only approach that works for longer horizons and allows us to assess information rigidity for extended-period forecasts.

3 Forecaster Data and Noisy Information Predictions

3.1 Forecast Error Data

I use forecasts of headline CPI from the Survey of Professional Forecasters (SPF). Professional forecasters should be among the most well-informed agents in the economy. Accordingly, the presence of information frictions in their expectation signals that deviations from full information rational expectations are likely to be economically significant across firms and consumers as well. The survey is published quarterly by the Federal Reserve Bank of Philadelphia, though prior to 1990 it was operated by the American Statistical Organization and the National Bureau of Economic Research. Forecasts of the headline CPI inflation rate have been included in the survey since 1981Q3. This quarterly availability allows me to calculate forecast errors across subsequent periods. The survey includes forecaster predictions of the GDP deflator that can be used to generate an inflation rate since its inception in 1968Q4. I discuss the results of the analysis on this series and its implications for the conclusions of this paper in Appendix B

[13] The aggregate semi-annual forecast regressed on its two-quarter lag is $\overline{FE}_{t+2,t|t}(i) = \rho^2(1-k)^2\overline{FE}_{t,t-2|t-2}(i) - \rho(1-k)(1+\rho k)w_t + (1+\rho)w_{t+1} + w_{t+2}$.

The survey is sent out following the release of the advance report of the Bureau of Economic Analysis's advance report of the national income and product accounts, which occurs in the last week of the first month of the quarter. Recent values from this report and the reports of other government agencies are included in the questionnaire. Questionnaires are due back by the third week of the second month of the quarter. While there is an array of information available to the forecasters in the survey, we can think of the signals described in Section 2 as a weighting of information available to the public and to the forecaster as described in Bordalo et al. (2020) and Andrade et al. (2016).

Forecasts for the CPI are provided as annualized rates for quarter-over-quarter inflation. The variable CPI3 is the one-quarter ahead forecast used in this paper. Both Bordalo et al. (2020) and Fuhrer (2018) look at CPI data in their analyses. While Fuhrer (2018) uses one-quarter ahead forecasts, Bordalo et al. (2020) uses longer horizon foreacsts. Data on the realization of inflation comes from the St. Louis FRED CPI series for the corresponding sample. Realizations of inflation are formed from the final release measures and are calculated to match the format of the forecasts:

$$\pi_{t+h,t} = \left[\left(\frac{CPI_{t+h}}{CPI_t} \right)^{\frac{4}{h}} - 1 \right] \times 100$$

3.2 Results from Noisy Information Models

Using this data, I estimate Equations 7 and 8 for one-quarter ahead forecast errors. For forecasts at horizons greater than one, I estimate Equation 9. The results appear in Table 2. I include an indicator for 2008Q3, which included a large and atypical surprise in CPI inflation forecast errors.

For 1-quarter ahead forecasts, both the pooled regression implies some information rigidity in expectations of CPI inflation, while the aggregate regression cannot reject $\beta_1 = 0$. As we would expect, the individual forecaster results are much more precise due to a larger sample size with Newey-West standard errors roughly 75 percent lower than the standard errors from the mean regression. The coefficients on lagged forecast errors should map to $\rho(1-k)$ and, as both are significantly different than zero, we can reject the null hypothesis of full information rational expectations and assume that $k << 1$. Assuming that $\rho = 1$ gives an upper bound on the estimate of k. For the individual forecast, the estimate 0.27 implies a point estimate[14] of k as high as 0.73, indicating that forecasters base up to 73 percent of their new expectation on their most recent signal. The aggregate estimate, 0.22, implies an upper bound on k of 0.78. In the case of CPI inflation, it is important to consider the persistence portion of the regression coefficient as it is substantially lower than 1 and will reduce the predictability of forecast errors independently of the noisy information portion. Using $\rho = 0.39$ to be consistent with time series estimation of the transition equation gives estimates of $\hat{k}_{individual} = 0.31$ and $\hat{k}_{aggregate} = 0.44$. Both Dovern et al. (2014) and Coibion and Gorodnichenko (2015) find $k = 0.50$, approximately in line with the finding from the aggregate regression.

It is difficult to assess the difference in the two estimates statistically as the aggregate approach is a time series estimation in a small sample and therefore has large standard errors. This difference does, however, due to the low level of process persistence, imply an economically significant difference in the quantification of noise in forecasters' information. Comparing point estimates of k, the individual approach indicates that forecasters base roughly thirteen percent less of their new forecasts on new signals. This further suggests that signals are more noisy than the aggregate approach shows and forecasters can trust their information relatively less. Dovern et al. (2014) finds evidence for a greater amount of implied information rigidity in estimates obtained from mean forecasts rather than individual forecasts.[15] I will show that a difference in the individual and mean coefficients

[14]The point estimate of k is $1 - \frac{\beta_1}{\rho}$. To find the standard errors, it is necessary to use the delta method.

[15]Dovern et al. (2014) finds estimates of information rigidity at the individual level roughly half that of estimates found using consensus or aggregate forecasts.

can be a result of misperception and the direction of this difference is determined by the sign of the misperception, with overestimated persistence causing the aggregate coefficient to be lower than the individual coefficient and underestimated persistence causing the opposite.

The regressions for longer-horizon forecast errors ($h > 1$) include time fixed effects to control for the endogeneity that arises from the quarterly arrival of innovations to the inflation process. As the horizon and therefore the length of time between forecast errors in the regression increases, we expect to see an exponential decline over h in the coefficient on lagged forecast errors. Table 2 shows that we do not see this decline in the data. This may result from omitted variable bias coming from parameter misperception, further indicating the need for a broader set of models to describe inflation expectations.[16]

We may worry that forecaster entry and exit in the survey is non-random, that is that forecasters who provide better forecasters are more likely to stay. If some forecasters receive better signals than others and are therefore more likely to deliver better projections and stay in the survey, we might want to disregard forecasters who are in the survey for too short a period of time. Table 3 presents the same estimates from Table 2, but limiting the sample to forecasters who remain in the sample for at least 30 periods. There is no substantial difference in the findings between the two tables. The trimmed sample also does not cause the coefficients on higher order lags to decrease as predicted in the theory. Accordingly, for the rest of the analysis I use all the forecasters in the sample.

4 An Additional Source of Information Frictions

Noisy information first affects the way forecasters form expectations about current inflation, also called nowcasts. These nowcast errors are then perpetuated into forecasts errors as forecasters use the transition equation to form expectations of future inflation.

[16]See Section 4.

If forecasters use the correct model parameters to project the nowcast into the future, their forecast errors will consist only of future innovations to inflation (that are unobservable at time t) and their nowcast errors scaled by the persistence of the process and the degree of information rigidity they face. If, however, forecasters anticipate that shocks to inflation are less or more persistent than they actually are, they will project their forecasts forward at a rate different than the rate at which inflation actually moves. This misperception causes an additional predictability in forecast errors and serial correlation in forecast errors over time.

4.1 Forecast Errors with Incorrectly Perceived Persistence

When the forecaster's perceived inflation persistence is not consistent with the true value of persistence, the forecast error equation will take a slightly different form than in Equation 5.[17]

Define an agent's perceived persistence as $\rho_i = \rho + q_i$, so q_i measures the misperception of persistence. Assume all agents share the same misperception: ρ_i and q_i are identical $\forall i$. Using this similarity, denote ρ_i as $\tilde{\rho}$ and q_i as q. The forecast from Equation 4 takes the same form with one adjustment.

$$\begin{aligned} \pi_{t+1,t|t}(i) &= \mu + \tilde{\rho}\pi_{t,t-1|t}(i) \\ &= \mu + k\tilde{\rho}\pi_{t,t-1} + k\tilde{\rho}v_t(i) + \tilde{\rho}(1-k)\pi_{t,t-1|t-1}(i) \end{aligned} \tag{10}$$

where $\tilde{\rho}$ takes the place of ρ in front of $\pi_{t,t-1|t}(i)$ because we now anticipate that agents will forecast their beliefs about past inflation into beliefs about time t inflation using what they believe to be the correct persistence parameter. Subtracting both sides from π_t gives the following equation for forecast errors:

[17]As the misperception in the constant does not effect the coefficients on forecast errors that help determine the effects of the frictions, I focus my analysis on misperception in persistence. Appendix D shows this in detail.

$$FE_{t+1,1|t}(i) = \tilde{\rho}(1-k)FE_{t,t-1|t-1}(i) - q\pi_{t,t-1} + w_{t+1} - \rho k v_t(i) \tag{11}$$

Under noisy information ($k < 1$), forecast errors still exhibit serial correlation. However, even controlling for lagged forecast errors, the current forecast error shows additional predictability based on the current value of inflation. Fuhrer (2018) includes the lagged value of inflation in several of his regressions, but not the current value.

If $q > 0$ and forecasters overestimate persistence, we expect that the coefficient on $\pi_{t,t-1}$ will be negative. For underestimated persistence, we expect the opposite.

If $\tilde{\rho} = \rho$, this equation collapses to Equation 5 and there is no misperception effect. If, however, $\tilde{\rho} \neq \rho$ and forecasters do not form their forecasts with the true value of persistence, we should include inflation in our framework for estimating noisy information.

Using this equation, I estimate the following reduced form equation to uncover the misperception.[18] *** denotes significance at the 0.01 level.

$$\begin{aligned} FE_{t+1,t}(i) &= \beta_0 + \beta_1 FE_{t,t-1}(i) + \beta_2 \pi_{t,t-1}(i) + \epsilon_t(i) \\ &= 0.56^{***} + 0.43^{***} FE_{t,t-1|t-1}(i) - 0.21^{***} \pi_{t,t-1} \end{aligned} \tag{12}$$

Under the null that there is no misperception of persistence, $\beta_2 = 0$. This estimation implies that $q = 0.21$, meaning that agents overrestimate the persistence of the inflation process. The misperception of persistence is consistent with Jain (2019)'s findings, though she finds an underestimation as well as a much more dramatic deviation from time series estimates of ρ than my result suggests.[19] The interpretation on β_1 is now slightly different. Rather than $\rho(1-k)$, β_1 now maps to $\tilde{\rho}(1-k)$ and the perceived value of persistence takes the place of the true value of persistence. The finding that $\beta_1 > 0$ indicates that

[18] It is possible to estimate this equation using mean forecast errors, but the precision will be greatly reduced.

[19] Jain estimates that, across SPF forecasters, the 75th percentile of perceived coefficients on the persistent component of inflation is 0.40.

this specification also detects noisiness in information. [20] Note that modeling Equation 11 in the diagnostic framework of Bordalo et al. (2020) gives us:

$$FE_{t+1,1|t}(i) = \tilde{\rho}(1 - (1+\theta)k)FE_{t,t-1|t-1}(i) + w_{t+1} - \rho k v_t(i)$$

Bordalo et al. (2020)'s model of diagnostic expectations would mean that agents overweight the signal relative to the rational noisy information weight. We cannot disentangle the effect of increased k (better information) and increased θ (more diagnostic expectations). If the noisy information friction and misperception frictions were absent, however, we would recover a negative coefficient on lagged forecast errors and no predictability with respect to $\pi_{t,t-1}$. This means that, even if underlying expectations are diagnostic, information must be imperfect and persistence incorrectly estimated to see the documented results.

If we exclude inflation from the regression of forecast errors on their own lags, we create omitted variable bias in the coefficient on forecast errors. As inflation is positively correlated with forecast errors and the coefficient on inflation is negative, omitting inflation from the regression will lead to a negative bias on lagged forecast errors. In the context of this model, a downward biased coefficient will lead us to conclude that forecasters receive signals that are less noisy than they actually are. This can explain why β_1 is lower than it appears in Table 2, where $\pi_{t,t-1}$ is omitted.

The incorrect observation of ρ constitutes an additional deviation from full information rational expectations beyond the noisy information model of Coibion and Gorodnichenko (2015) and Coibion and Gorodnichenko (2012). The model in this paper permits two types of error. There is the error in observing the current state and then the error in moving the projection of the current state forward. The misperception of persistence is a

[20] The positive value of the β_0 implies that forecasters overestimate the constant by 0.56. See Appendix D.

deeper form of imperfect information that will cause consistent issues for forecasting.

4.2 Parameter Misperception at Longer Horizons

When I extend the model to consider forecast errors at longer horizons, forecaster misperception of inflation persistence has a more complicated effect on the predicted path of forecast errors than it does when $h = 1$. When forecasters use their perceived value of $\tilde{\rho} = \rho + q$, rather than ρ in forming expectations, the predicted path of semi-annual forecast errors is:[21]

$$\begin{aligned} FE_{t+2,t|t}(i) = {} & \tilde{\rho}^2(1-k)^2 FE_{t,t-2|t-2}(i) - \tilde{\rho}(1-k)(1+\tilde{\rho}k)\pi_{t-1,t-2} - (1+\tilde{\rho})q\pi_{t,t-1} - q\pi_{t+1,t} \\ & + \tilde{\rho}(1-k)(1+\tilde{\rho}k)w_t + (1+\tilde{\rho})w_{t+1} + w_{t+2} \\ & - (1+\tilde{\rho})\tilde{\rho}^2(1-k)kv_{t-1}(i) - (1+\tilde{\rho})\tilde{\rho}kv_t(i). \end{aligned}$$

Just as the shocks between $t-2$ and t all appear in the equation for $FE_{t+2,t|t}(i)$, the misperception of persistence causes the realizations of inflation from $\pi_{t-1,t-2}$ to $\pi_{t+1,t}$ to appear in the predicted path of forecaster errors. It is difficult to include these omitted variables to estimate the misperception as the time fixed effects terms will absorb the time-specific inflation terms.

4.3 Evidence from the Term Structure of Forecasts

Section 4 presented evidence that there are discrepancies between the true parameters of the inflation process and the forecaster's perception of these parameters with data on the forecast errors. Whether or not agents do in fact incorrectly perceive inflation parameters can also be tested given the term structure of forecast data available in the Survey of Professional Forecasters. Each period, respondents provide annualized quarter-

[21] This is derived in Appendix A-2.

over-quarter projections of the rate of CPI inflation over several successive quarters. To estimate the perceived inflation persistence for this series, I use the series for the nowcast through the quarterly projection for four quarters ahead.[22]

Applying the process implied by the transition equation in Equation 1, we can construct the relationships between each forecast the agent makes in period t.

$$\begin{aligned}\pi_{t+h,t+h-1|t}(i) &= \tilde{\mu} + \tilde{\rho}\pi_{t+h-1,t+h-2|t}(i) + E[w_{t+1}] \\ &= \tilde{\mu} + \tilde{\rho}\pi_{t+h-1,t+h-2|t}(i)\end{aligned} \tag{13}$$

where $\pi_{t+h,t+h-1|t}(i)$ represents agent i's forecast of inflation from period $t+h-1$ to $t+h$ given information available in period t. One can formulate reduced form regression equations matching the system of equations formed by the term structure of agent forecasts. [23] I exclude the relationship between the nowcast and the quarter ahead forecast, as this may be different from the relationship between other quarterly forecasts due to both information availability differences and forecaster perceptions of the nowcast. Accordingly, consider the following system of equations.

$$\pi_{t+s,t+s-1|t}(i) = \beta_0 + \beta_1\pi_{t+s-1,t+s-2|t}(i) + \epsilon_{st,i}, \ldots s = 1,2,3,4 \tag{14}$$

As the three reduced form equations described above represent the relationship between forecasts of quarter-ahead inflation at adjacent horizons, I combine the three and estimate the system of equations in one regression, including fixed effects for the distance, s, of each forecast from the forecasting period t.[24]

[22]These are the variables CPI2, CPI3, CPI4, CPI5, and CPI6 in the Survey of Professional Forecasters.

[23]Differing from Jain (2019), I formulate the relationship between forecasts themselves rather than the revisions in the forecasts. This holds an agent's information set constant within each observation and avoids canceling out perceptions of long run components of inflation that are important for understanding forecasters' perception of the inflation process.

[24]If each regression is performed separately, the results for each look very similar to the systems regression.

$$\pi_{t+s|t+s-1}(i) = \beta_0 + \beta_1 \pi_{t+s-1,t+s-2|t}(i) + \sum_{s=1}^{4} \gamma_s 1(s) + \epsilon_t(i) \tag{15}$$

The regression coefficient and constant provide estimates of the process parameters, $\beta_0 = \hat{\tilde{\mu}} = 0.73^{***}(0.02)$ and $\beta_1 = \hat{\tilde{\rho}} = 0.76^{***}(0.01)$. [25]

Running a corresponding AR(1) regression on the inflation process gives $\hat{\rho} = 0.39^{***}(0.04)$ and $\mu = 1.67^{**}(0.14)$, where *** and ** denote significance at the 0.01 and 0.05 percent levels, respectively. Agents therefore significantly overestimate the persistence of inflation and significantly underestimate the regression's constant term. Using the z-test from Paternoster et al. (1998) I reject the null that $\tilde{\rho} = \rho$ at the 99 percent level of confidence. I can also reject the null that $\tilde{\mu} = \mu$ at the 99 percent level. This supports the findings of parameter misperception.

I have presented evidence for both types of expectation formation friction. Both noisy information and forecaster mistakes about parameter values will lead to serial correlation in forecast errors and have implications for the magnitude of forecast errors. In the next section, I quantify the relative importance of each friction for the path of expectations.

5 A Structural Model of Forecasting with Two Frictions

In order to assess the relative importance of the noisy information and parameter misperception channels, I use the following structural model with variation in parameters, θ. I define:

[25] The coefficients γ_s are not statistically different from 0. This indicates that forecasters have stable beliefs about the constant despite the distance of the period in the future.

$$\theta = \begin{bmatrix} \rho \\ q \\ k \end{bmatrix}.$$

The parameters k and q represent the noisy information and parameter misperception frictions respectively. The gain, k, represents the weight forecasters give to new signals about inflation and quantifies the effect of noisy information on expectations. A lower value of k implies that forecasters face significant constraints in viewing the level of the target variable in real time. The size of the misperception, q, introduces forecaster errors in understanding the underlying structure governing inflation dynamics. I also include ρ in the parameter vector as it directly influences the predictability of forecast errors when $k < 1$.

I simulate forecasters forming predictions according to the data-generating process modeled in Section 2. The simulated forecasts can be used to estimate the moments from the data that depend on ρ, k, and q to be used in a simulated method of moments approach to the estimation of parameters. This model can also be used to assess other potential explanations for the differences in estimation at the individual and aggregate forecast levels. I discuss this in greater detail in Section 6.

Inflation evolves and signals arrive according to the following transition and measurement equations.[26]

$$\pi_{t+1,t} = \mu + \rho\pi_{t,t-1} + w_{t+1}$$

Forecasters again receive private signals:

[26]This follows from Equations 1 and 2

$$z_t(i) = \pi_{t,t-1} + v_t(i)$$

Agents process information through the Kalman filter, selecting the optimal gain for the given level of persistence (as they perceive it), process innovation variance, σ_w^2 and their own signal noise variance, σ_v^2. I allow forecasters to perfectly observe all parameters aside from persistence. Forecasters use $\tilde{\rho}$ rather as their estimate of ρ. As such - their optimal filtering process leads them to form optimal forecasts as in Section 4.1.

I generate data for the length of the sample period in the data ($F = 153$) with an extra 100-quarter burn in period. The total length of T is therefore 100+F. I further generate one-quarter ahead forecast errors for $N = 184$ forecasters. The simulated data following the burn-in period form an $N \times F$ matrix of individual forecast errors. Forecasters in the SPF enter and exit the survey throughout the sample period, generating a highly unbalanced panel structure. To mimic this structure in the simulated data, each forecaster in the data is matched to a row of the simulated forecast error matrix. Only the elements of each row corresponding to the time periods in which that forecaster participated in the survey are populated with forecast errors.[27] This allows me to match exactly the number of observations in the pooled and time series specifications. The pooled specification now consists of the non-missing observations in this matrix. I then calculate the aggregate forecast errors, $\overline{FE}_{t+t,t|t}$, by taking the mean of the individual forecast errors in each period.

[27]Note that it does not matter which rows of the simulated matrix we assign to the dates matching the forecasters in the data as, in the simulated data, the primary differences between the forecasters are the periods they appear in the survey and the draws of their individual specific noise, which are iid. Forecasters are matched to periods according to when they participated in the survey - such that a forecaster present only early in the sample period will appear early in the simulated data. The model does not include learning by doing. Accordingly, the forecasters still engage in the updating process in periods in which they are not forecasting. This is equivalent to forecasters making private forecasts every period, but reporting only on occasion.

Each simulation contains M iterations. For each iteration, m, I estimate the following three equations for 1-quarter ahead forecasts and save the coefficients of interest - that is those pertaining to ρ, q, and k - and their standard errors. I calibrate σ_w^2 to 0.5 and allow σ_v^2 to be determined such that given σ_w^2, ρ, and q, the chosen k is optimal. Changing this value does not substantially alter the results as σ_v^2 will always be chosen relative to the value of σ_w^2. Note that in the equations below, $\varepsilon_t^a(i)$ and ε_t^b are time and individual specific, whereas ε_t^c varies only with time.

$$FE_{t+1,t|t}(i) = \beta_0^a + \beta_1^a FE_{t,t-1|t-1}(i) + \varepsilon_t^a(i) \tag{16}$$

$$\overline{FE}_{t+1,t|t} = \beta_0^b + \beta_1^b \overline{FE}_{t,t-1|t-1} + \varepsilon_t^b \tag{17}$$

$$FE_{t+1,t|t}(i) = \beta_0^c + \beta_1^c FE_{t,t-1|t-1}(i) + \beta_2^c \pi_{t,t-1} + \varepsilon_t^c(i) \tag{18}$$

For a completed simulation, I calculate the average simulated values of the coefficient on lagged forecast errors in all three equations and the coefficient on lagged inflation in Equation 18. I define:

$$g(\theta) = \begin{bmatrix} \frac{1}{M}\sum_m \beta_1^a(m) & \frac{1}{M}\sum_m \beta_1^b(m) & \frac{1}{M}\sum_m \beta_1^c(m) & \frac{1}{M}\sum_m \beta_2^c(m) \end{bmatrix}$$

The simulated coefficients can be compared to the same set of coefficients from the data: $\hat{g} = \begin{bmatrix} 0.27, 0.22, 0.43, -0.21 \end{bmatrix}$.

5.1 Estimation by Simulated Method of Moments

The purpose of this estimation is to find the values of parameters most likely to generate the moments seen in the data. The parameters to be estimated are defined by θ. I define the objective function for minimization to be the difference between the coefficients from the data and the average coefficients from the simulation performed with a draw of θ. For the simulated method of moments, each simulation contains $M = 100$ iterations.

$$J(\theta) = [\hat{g} - g(\theta)]\Omega[\hat{g} - g(\theta)]' \tag{19}$$

Where $\hat{g}$ describes the relevant coefficients from the data, $g(\theta)$ is the corresponding coefficients simulated from a posited θ, and Ω is a weighting matrix scaling each component of $\hat{g} - g(\theta)$ in the objective function $J(\theta)$. I use the identity matrix for Ω such that each moment is weighted equally. The steps for the estimation algorithm proceed as follows:[28]

1. Draw a starting value of θ_0 from the acceptable range for each parameter. ρ_0 and k_0 are bounded between 0 and 1. The bounds for q_0 are determined by the draw for ρ such that $\tilde{\rho_0} \in (0, 1)$. Accordingly, $q_0 \in (-\rho_0, 1 - \rho_0)$.

2. Simulate data as described and calculate the value of the objective function: J_0. Save θ_0 and J_0.

3. Make a small and random perturbation to the input parameters: $\theta_1 = \theta_0 + \psi * x$. Here ψ is a 3×3 matrix that scales the value of the perturbation to each parameter. x is a vector of random variables. Simulate data again and calculate the value of the objective function J_1. Save θ_1 and J_1.

4. Compare J_0 and J_1. As the goal is minimization, accept the θ_1 over θ_0 if $J_1 < J_0$. If $J_1 > J_0$, accept θ_1 with some probability that is decreasing in the $J_1 - J_0$ such that values of J_1 that are close to undercutting J_0 are more likely to be accepted. For each step, save the accepted parameters as well as the value of the objective function associated with them.

5. Repeat steps 2-4 for the desired length of the simulation.

[28]This is a Metropolis Hastings algorithm form of Markov chain Monte Carlo method. I use the same set of random number draws for values of $v_t(i)$ and w_t, though the draws for $v_t(i)$ are scaled by different values of σ_v^2 as θ varies.

The estimated parameters $\hat{\theta}$ are constructed as a weighted average of the accepted parameters from the algorithm. The weights are determined by the relative value of the objective function. An estimation of parameters returns $\hat{\rho} = 0.32(0.005)$, $\hat{q} = 0.23(0.004)$, and $\hat{k} = 0.19(0.004)$. Interestingly, the estimate for $\hat{\rho}$ is similar to time series estimates, despite the time series estimate not appearing in the set of moments used to estimate parameters. The overestimation of persistence is surprising, but in line with the estimates derived from the term structure of forecasts in Section 4.3. The estimate $\hat{q}$ also exceeds the value implied by the coefficient on inflation in Equation 18. When the agent operates with an incorrect value of persistence, she uses a sub-optimal gain. This causes the incorrect weighting of the noise terms and biases the coefficient downwards, meaning it takes larger absolute values of q to match the moment corresponding to q. The estimate of k means that forecasters will form 19 percent of their new expectations with new information. This is lower than many recent findings of such as those of Afrouzi (2017), who derives a relatively low degree of information rigidity of $k \approx 0.7$ for firm managers in New Zealand. My estimate is also substantially different than estimates from Dovern et al. (2014) and Coibion and Gorodnichenko (2015) and implies that information is much more noisy than those estimates suggest. These papers find a Kalman gain of 0.5.

I simulate coefficients for the estimated values of the parameters and take the average (over 100 rounds of the simulation) point estimates of β_1^a and β_1^b from Equations 16 and 17. The individual coefficient takes an average value of 0.27, while the average coefficient on mean forecast errors takes a value of 0.20. The mean coefficient is higher than the individual coefficient, as we see in the data, but the simulation cannot match the magnitude of the difference in coefficients documented in Section 3.2. The true value of the coefficient under this specification is $\tilde{\rho}(1-k) = 0.45$.[29] When forecasters overestimate persistence, both the individual and panel approaches result in downward bias

[29] $\tilde{\rho}(1-k) = (0.32 + 0.23) \times (1 - 0.18) = 0.45$.

in the coefficient, making information appear less noisy than it actually is. The average simulated point estimates of β_1^c and β_2^c from Equation 18 are 0.45 and - 0.20, respectively. These match the corresponding moments in the data almost exactly. The value 0.45 also correctly recovers the true value of $\tilde{\rho}(1-k)$.

5.2 The Effect of Parameter Misperception

A simple counterfactual exercise demonstrates the importance of controlling for parameter misperception when using forecast errors to quantify noisy information. The coefficients on forecast errors in Equations 16, 17, and 18 should return $\tilde{\rho}(1-k)$, a combination of perceived persistence and the Kalman gain. When persistence is perceived correctly, all three regressions will result in unbiased coefficients. When persistence is misperceived, however, the coefficients β_1^a and β_1^b from Equations 16 and 17 will have omitted variable bias. I estimate the simulated coefficients for two cases: one with no parameter misperception and one where a substantial portion of $\tilde{\rho}$ is over-estimation. I allow k and $\tilde{\rho}$ to be equal across the two simulations such that the coefficients from each simulation should be equal; the composition of $\tilde{\rho}$ differs. The simulation shows that all three regression equations return the correct coefficient when persistence is correctly perceived. Introducing the overestimation of persistence causes us to underestimate the noise in information.

In the baseline case, the parameters are $k = 0.19$, $\rho = 0.32$, and $q = 0.23$. Persistence is over-estimated and the true value of the coefficients on the forecast errors is $\tilde{\rho}(1-k) = 0.45$. In the counterfactual case, I adjust the parameters to be $k = 0.19$, $\rho = 0.55$, and $q = 0$ such the misperception friction is no longer present and the degree of noisy information and true value of the coefficients are the same as in the baseline case. The estimates of β_1^a, β_1^b, and β_1^c from Equations 16, 17, and 18 under each case appear in Table 4. When the misperception is present, β_1^a and β_1^b are downwardly biased and also deviate from each other. When I control for the misperception using Equation 18, β_1^c returns the true value of the coefficient. In the counterfactual case with no misperception, all three coefficients return the true value.

In the counterfactual case, the individual and aggregate coefficients are 1.61 and 2.08 times higher than the coefficients that are estimated when the misperception is present but we fail to control for it as in Equation 18. This means that failing to account for the overestimation of persistence will result cause us to understate the noisiness in information. I conduct a back of the envelope calculation using these values and the coefficients from the data, 0.27 and 0.22. Scaling these coefficients for the effect of the misperception returns coefficients of 0.44 and 0.46 for the regressions of individual and aggregate forecast errors on their own lags, respectively. These estimates are more in line with the unbiased coefficient (0.43) that we recover when we control for parameter misperception by including the realization of inflation in the regression as in Equation 18.

Appendix B shows simulation results for PGDP inflation for both the full sample and the same period as CPI inflation. For the full sample, I find that the persistence of PGDP inflation is underestimated. This creates the opposite pattern of what we see in the CPI data. While the overestimation of persistence causes the coefficients from Equations 16 and 17 to be biased down, understating the amount of noise in expectations, underestimated persistence causes these noisy information coefficients to be biased upwards, causing us to conclude that information is noisier than it actually is. This means that considering parameter misperception does not always lead to the conclusion that information is noiser or less noisy than the method of Coibion and Gorodnichenko 2012. Rather, the quantitative estimates get closer or further from full information depending on the sign of the misperception. Regardless of the sign of q, Equation 18, recovers an unbiased coefficient and the correct estimate of $\tilde{\rho}(1-k)$. This result highlights a key takeaway from this paper, that it is important to assess noisy information allowing for simultaneous parameter misperception.

6 Other Potential Explanations

Up to this point, I examine forecaster underestimation of persistence in the context of noisy information as an explanation for the empirical moments this paper documents. In this section, I consider whether other possible deviations from the basic noisy information model can account for my results.

6.1 Changes in Parameter Values

This paper argues that persistence in forecaster errors can be explained by the interaction of noisy information and forecaster misperception of the parameters governing inflation dynamics. The sample period for this analysis, however, has seen notable changes in inflation dynamics, covering the time of the with the start of the data coinciding with the Volcker Disinflation and immediately following the Great Inflation of the 1970s and the Great Moderation beginning in the mid-1980s. Several papers argue that inflation persistence has changed over this period, and with it the volatility of inflation. Stock and Watson (2007) and Cogley, Primiceri, and Sargent (2010), for example, argue that inflation persistence has declined. Benati (2008) and Erceg and Levin (2003) argue that monetary policy regimes can influence inflation persistence, leading to changes over time.[30] In light of these concerns, I consider possible changes in the inflation process as well as changes in the parameters of the agent's signal processing problem as possible explanations for the empirical moments documented in this paper.

Using the simulated model, I can assess the effects on the regression coefficients of various changes in parameter values. For each σ_w^2, σ_v^2, ρ and μ, I simulate the data-generating process described in Section 5.1 with one adjustment - halfway through the sampling period the variable in question changes. As many of these variables affect the agent's choice of an optimal Kalman gain, I run each simulation in two ways. The first

[30] See Mishkin (2007) for a useful summary of changes in inflation dynamics since the early 1970s.

allows forecasters to observe the change in the variable and absorb it into their choice of Kalman gain. In the other, forecasters do not observe the change and are left with a suboptimal gain. The estimated coefficients from Equation 18 appear in Table 6. The implications of changes in parameter values for the estimates for Equations 16 and 17 for each simulation appear in Table 7.

I calibrate the model such that the starting parameter values match those estimated via simulated method of moments. As such, the starting values are: $k = 0.25$, $\rho = 0.35$, $\sigma_w^2 = 0.5$, and $\sigma_v^2 = 1.65$. The steady state value of U^- is set to 0.55 by the other parameters. As this value depends on other parameters, it will change with their values. I set the misperception, q, to zero to test alternative theories. The magnitudes of the changes in each parameter appear in Table 5. I use small changes in each parameter. For σ_w^2 and σ_v^2, I choose values that, given the calibration of the other parameters, will lead to a $\Delta k = \pm 0.10$.

- **Inflation Volatility**: The variance of innovations to inflation, σ_w^2, will factor into the optimal choice of Kalman gain. If the process is noisier, that is σ_w^2 is higher, an agents signal is more informative she will give it a relatively higher k. The tables show, however, that a change in σ_w^2 cannot explain the pattern of moments in the data.

- **Signal Noise Variance**: As the agent's signal noise variance, σ_v^2, changes, so does her optimal Kalman gain. We can think of a change in signal noise variance as a change in the signal quality, with a decline in the variance being an improvement and an increase in the variance as a decline in signal quality. Accordingly, if the agent observes this change in signal quality, she will either increase or decrease the weight that she gives the signal, k. Simulation shows a change in signal noise variance cannot explain the wedge between the aggregate and individual approach coefficients or a positive coefficient on inflation when that is included in the regression equation.

- **Inflation Constant**: A change in the constant will not change the agent's optimal gain parameter. If the change is observed, the constant term will drop out of the forecast error equation. If, however, the change is unobserved, there will be a structural break in the constant in the forecast error equation.[31] The simulated estimates show, however, that a change in the inflation constant cannot generate a wedge between coefficients or a significant coefficient on inflation in Equation 18.

- **Inflation Persistence**: Table 6 shows that, in the simulated model, a positive (negative) change in inflation persistence is the only change that can generate a significant coefficient on the inflation term in Equation 18. This feature only arises when forecasters are not permitted to observe the change in persistence and therefore underestimate (overestimate) persistence, meaning the misperception of persistence drives this finding rather than the change in persistence itself.[32] This is broadly consistent with the main result of this paper.

These show that a change in persistence is the most likely of these alternative candidate explanation. A change in persistence will replicate the moments from the data only when it is accompanied by an underestimation of this persistence on the part of forecasters. This points to the misperception of the parameters governing inflation dynamics as the primary explanation for the moments observed.

Table 7 shows that unobserved changes in persistence result in bias in the coefficients on lagged forecast errors in the opposite direction as q. This means that an overestimation of persistence makes information appear less noisy than it is while the underestimation of persistence makes information appear overly noisy. This is only true when we fail to control for inflation. Therefore, considering noisy information while considering the

[31] See Appendix D.

[32] The case where persistence decreases but the change is undetected is therefore in line with the literature that argues for a decline of inflation persistence over time.

misperception of inflation persistence can give a more accurate representation of both frictions.

7 Extensions

In this section I consider four possible extensions of the noisy information model with parameter misperception. First, I consider time variation in the parameters describing the two frictions forecasters face. This allows forecasters to misperceive inflation to a different degree and weight their signals differently across time periods. Second, I model the underlying inflation process with a richer AR(2) process. I then consider my analysis with multivariate information. Fourth, I relax the assumption that forecasters do not observe realizations of inflation by modeling the noisy information model where forecasters receive signals about the most recent past realization of inflation as well as current inflation. A final extension considers the case where forecasters receive both public and private signals about current inflation, meaning that their information noise has a component that is shared across forecasters.

7.1 Time Variation in Information Frictions

The inflation process over the sample period has seen many potential changes that could influence the Kalman gain that forecasters assign to their signals and the degree of their misperception of inflation persistence. Accordingly, I examine the possibility of time-variation in these parameters by performing rolling window regressions for the specifications in Equations 7, 8, and 12. The window width is 80 quarters, so each estimate is formed with 20 years of data. Figure 1 shows these coefficients plotted against the starting period of the window over which the coefficients are estimated.

This exercise reveals several features of time-variation in these parameters. There appears to be low frequency variation in the coefficient on forecast errors. This indicates changes in persistence, perceived persistence, the Kalman gain, or some combination of the three. This variation is more apparent in the point estimates at the aggregate level, though the standard errors for this approach are even larger than they are for estimates over the full sample. The standard errors for the aggregate approach are particularly large in the early part of the sample. In this period, the standard errors on the coefficient from the aggregate approach are so large that, given the time series estimate of ρ, nearly every possible value of k - 0 to 1 - is possible under the 95% confidence interval. The coefficient on forecast errors from the panel approach is very stable and has smaller standard errors across periods. We cannot reject the null of parameter stability over the period at the 95 percent level. This stability and precision is another advantage that my panel approach to estimation has over the mean forecast approach.

The evidence for time-variation in the misperception suggests that the finding that forecasters have, on average, underestimated the persistence of inflation since the late 1960s is largely driven by the forecasts of the 1970s. The bottom right panel of Figure 1 shows the coefficient on $\pi_{t,t-1}$, which can be interpreted as $-q$, or the negative misperception of persistence. As $\tilde{\rho} = \rho + q$, a positive value of this coefficient corresponds with underestimated persistence, while a negative value corresponds with an overestimation of persistence. The rolling window regression suggests that forecasters overestimate persistence in periods beginning 1980 and later.

We can make some sense of these results in light of Cogley and Sargent (2002), which argues that inflation was weakly persistent in the 1960s and 1990s and highly persistent in the 1970s with persistence increasing between 1965 and 1979 and declining between 1979 and 2000. If forecasters entered the 1980s with an expectation that inflation would follow the same high persistence as in the 1970s, they would overestimate its true persistence upon its decline. As forecasters then became accustomed to the lower level of inflation persistence, the misperception coefficient approaches 0.

Figure 2 shows that perceived persistence follows approximately this pattern as well as the estimated persistence of the inflation process. The figure plots an 80-quarter rolling window regression of $\tilde{\rho}$ as estimated from the term structure of forecasts (Equation 14). Perceived persistence is initially high, as it follows a period of highly persistent inflation according to Cogley and Sargent (2002), and decreases period marked by lower persistence. Forecasters enter the 1980s with a higher value of perceived persistence which then gradually adjusts downwards. As the window begins to include periods from the 1990s, the the persistence of inflation as well as the perceived persistence both decline. The size of the misperception is smaller, but forecasters continue to overestimate persistence. These results suggest that forecasters may learn about the persistence of the inflation process and that more consideration should be given to learning about parameters in the context of noisy information.

7.2 Alternative Time Series Process

The analysis in this paper is based on a simple underlying inflation process. While this makes the analysis interpretable and the results readily comparable to the related literature, we may worry that the underlying time series is more complicated and such a difference would change our results. In this section, I extend the analysis to a richer time series process.

Allow the underlying inflation process to follow an AR(2):

$$\pi_{t,t-1} = \mu + \phi_1 \pi_{t-1,t-2} + \phi_2 \pi_{t-2,t-3} + w_t \tag{20}$$

This transition equation can be given the following state space representation:

$$\begin{bmatrix} \pi_{t,t-1} \\ \pi_{t-1,t-2} \end{bmatrix} = \mu \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \Phi \begin{bmatrix} \pi_{t-1,t-2} \\ \pi_{t-2,t-3} \end{bmatrix} + \begin{bmatrix} w_t \\ 0 \end{bmatrix}.$$

where $\Phi = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}$. Forecasters still receive a private signal about the current level of inflation. As such, we can write the measurement equation as:

$$z_t(i) = H \begin{bmatrix} \pi_{t,t-1} \\ \pi_{t-1,t-2} \end{bmatrix} + v_t(i) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \pi_{t,t-1} \\ \pi_{t-1,t-2} \end{bmatrix} + v_t(i)$$

The optimal Kalman gain is now $\begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$, where k_1 and k_2 indicate how much weight the signal should have in forming the new expectations of $\pi_{t,t-1}$ and $\pi t-1, t-2$, respectively. The forecaster's posterior expectation of the state space is given by:

$$\begin{bmatrix} \pi_{t,t-1|t}(i) \\ \pi_{t-1,t-2|t}(i) \end{bmatrix} = (I - KH) \begin{bmatrix} \pi_{t,t-1|t-1}(i) \\ \pi_{t-1,t-2|t-1}(i) \end{bmatrix} + Kz_t(i)$$

This can be projected forward into a forecast using the transition matrix, Φ. Giving a one quarter ahead forecast error of:

$$\begin{aligned} FE_{t+1,t|t}(i) = {} & [\phi_1(1-k_1) - \phi_2 k_2]FE_{t,t-1|t-1} + \phi_2 FE_{t-1,t-2|t-1}(i) \\ & - (\phi_1 k_1 + \phi_2 k_2)v_t(i) + w_{t+1} \end{aligned}$$

Estimating this equation on the individual forecast error data gives the following estimates that suggest a ϕ_2 that is negative but not significant.

$$FE_{t+1,t|t}(i) = -0.01 + 0.30^{***} FE_{t,t-1|t-1}(i) - 0.09^{***} FE_{t-1,t-2|t-1}(i)$$

I add misperception such that $\tilde{\phi}_1 = \phi_1 + q_1$ and $\tilde{\phi}_2 = \phi_2 + q_2$. In this case, the forecast error follows:

$$FE_{t+1,t|t}(i) = [\tilde{\phi}_1(1-k_1) - \tilde{\phi}_2 k_2]FE_{t,t-1|t-1} + \tilde{\phi}_2 FE_{t-1,t-2|t-1}(i) - q_1\pi_{t,t-1} - q_2\pi_{t-1,t-2}$$
$$- (\phi_1 k_1 + \phi_2 k_2)v_t(i) + w_{t+1}$$

Estimating this gives:

$$FE_{t+1,t|t}(i) = 0.76^{***} + 0.42^{***}FE_{t,t-1|t-1} - 0.04FE_{t-1,t-2|t-1}(i) - 0.17^{***}\pi_{t,t-1} - 0.06^{**}\pi_{t-1,t-2}$$

This implies $q_1 > 0$ and $q_2 > 0$, or overestimation of both ϕ_1 and ϕ_2. It further implies $\tilde{\phi}_2 \approx 0$. This does not contradict the results of the AR(1) and, given the estimates of ϕ_2 and q_2, this equation will collapse to roughly Equation 18, or the model with misperception of an AR(1).

Using the CPI projections for various horizons as described in Section 4.3, I estimate the perceived AR(2) parameters implied by the term structure of forecasts. The equation is the same as Equation 14 with the exception that it includes an additional lag of inflation as a predictor. The dependent variables in the system of equations are CPI4, CPI5, and CPI6, which correspond to annualized quarter-over-quarter inflation 2 through 4 quarters ahead.

I estimate:

$$\pi_{t+s|t+s-1}(i) = \beta_0 + \beta_1\pi_{t+s-1,t+s-2|t}(i) + \beta_2\pi_{t+s-2,t+s-3|t}(i) + \sum_{s=2}^{4}\gamma_s 1(s) + \epsilon_t(i) \tag{21}$$

where β_1 and β_2 correspond to $\tilde{\phi}_1$ and $\tilde{\phi}_2$, respectively. This gives $\beta_1 = \hat{\tilde{\phi}}_1 = 0.80^{***}(0.01)$ and $\beta_2 = \hat{\tilde{\phi}}_2 = 0.08^{***}(0.01)$. Figure 3 shows the time-varying perceived AR(2) coefficients. The coefficient on the first lag follows the same pattern as the perceived persistence in an AR(1) as shown in Figure 2. The coefficient on the second lag hovers around or just above 0. This indicates that treating the perceived series as an AR(2) does not create substantially different results than treating it like an AR(1).

7.3 Multivariate Information

The baseline model assumes that inflation evolves as an AR(1). While this simple model allows for ready comparison to similar literature (Coibion and Gorodnichenko (2015), Bordalo et al. (2020)), we may be concerned that inflation depends more on other variables than on its own lag. In this extension, I consider the relationship between CPI inflation and its own lag, the current value of GDP growth, the unemployment rate, and the interest rate[33] and the CPI inflation forecast with the forecasts of these same variable.

Table 8 shows the results of a regression of inflation on its own lag and the other variables. This shows that for the period considered, inflation depends on both the interest rate and its own lag, with GDP growth and unemployment having small and insignificant coefficients. In the parallel regression of the forecast of inflation on the nowcast of inflation as well as the forecasts of GDP growth, unemployment, and the interest rate. The nowcast and the interest rate have significant coefficients, with the coefficient on the nowcast being larger than the coefficient of the realization on its own lag, consistent with the evidence presented in this paper that forecasters overestimate persistence.

While this table suggests that other variables may play into the forecasting process, I rerun the central regression controlling for the forecast errors for each variables as well as the time-t realization of each variable.

$$FE_{t+1,t|t}(i) = 0.46^{***}FE_{t,t-1|t-1}(i) - 0.25^{***}\pi_{t,t-1} + \ldots$$

The coefficients on the additional variables are all near zero and insignificant, while

[33] I use the rate on the 3-month Treasury bill as there is an equivalent forecast for this variable.

the coefficients considered in my analysis are nearly identical to the regression that does not control for them. Leaving out the additional variable therefore has relatively little impact on the analysis.

7.4 Forecasters Receive Signals about Past Inflation

In Section 2, I assumed that forecasters receive private signals about the state of inflation in time t and do not observe any information about past inflation other than the signals they received in previous periods. Forecasters do not observe true realizations of inflation and therefore do not observe their own lagged forecast errors. This is a substantial assumption, especially considering that agents receive some summary statistics and estimates about the previous inflation realization when they receive their forecasting questionnaires for the SPF. This assumption also creates a significant inequality in the information available to the forecasters and the information available to the econometrician. To lessen this inequality, I can introduce the following assumption about observations of inflation. Each period the statistical agencies in charge of releasing inflation estimates release the measure $\pi^M_{t,t-1}$. This value consists of both the true value of inflation and some measurement error e_t.

$$\pi^M_{t,t-1} = \pi_{t,t-1} + e_t$$

Each period, the agent receives information about $\pi^M_{t-1,t-2}$. The forecaster cannot disaggregate the lagged realization of true inflation from the error term, effectively prohibiting them from observing past realizations of inflation. This new information is effectively a public signal about the lagged state of inflation.

The state still evolves according to the transition equation in Equation 1. The agents now observe a vector of signals, one private and one public.

$$Z_t(i) = \begin{bmatrix} z_t(i) \\ \pi^M_{t-1,t-2} \end{bmatrix} = \begin{bmatrix} \pi_{t,t-1} \\ \pi_{t-1,t-2} \end{bmatrix} + \begin{bmatrix} v_t(i) \\ e_t \end{bmatrix}.$$

Following Nimark (2014), I reformulate the measurement vector as:

$$Z_t(i) = \begin{bmatrix} z_t(i) \\ \pi^M_{t-1,t-2} \end{bmatrix} = H_1\pi_t + H_2\pi_{t-1} + Ru_t.$$

In the above equation, $u_t \sim N(0,1)$, $H_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $H_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and $R = \begin{bmatrix} \sigma_v \\ \sigma_e \end{bmatrix}$. H_1 describes the signal components pertinent to the current state while H_2 allows signals to include information about the lagged state. R serves to scale the noise term u_t such that it matches the distributions of $v_t(i)$ and e_t.

Forecasters will update according to:

$$\pi_{t,t-1|t}(i) = \mu + \rho\pi_{t-1,t-2|t-1}(i) + K\left[Z_t(i) - H_1\mu - (H_1\rho + H_2)\pi_{t-1,t-2|t-1}(i)\right].$$

The forecasters's optimal Kalman gain $K = \begin{bmatrix} k_1 & k_2 \end{bmatrix}$ is now a 1×2 vector indicating how forecasters should weight each signal.[34]

The one quarter ahead forecast error in this example is:

$$FE_{t+1,t|t}(i) = \rho(1-k_1)FE_{t,t-1|t-1}(i) - k_2FE_{t-1,t-2|t-1}(i) - \rho k_1 v_t(i) - \rho k_2 e_t + w_{t+1}$$

Here the forecast error depends on both the lagged forecast error and the lagged nowcast error of inflation. The signal about the lagged state gives the forecasters an opportunity to revise their past beliefs. The full information model is nested in this

[34] In this specification, k_1 is the weight on the private signal and corresponds to k in Section 2 while k_2 is the weight the forecaster will optimally assign to the public signal. The gain matrix is given by: $K = [\rho U^-(H_1\rho + H_2)' + \sigma_w^2 H_1' + \sigma_w R'] \times [(H_1\rho + H_2)U^-(H_1\rho + H_2)' + (H_1\sigma_w + R)(H_1\sigma_w + R)']^{-1}$

equation the same way as in Section 2. When the signal about current inflation is perfect, that is $\sigma_v^2 = 0$ and $z_t(i) = \pi_{t,t-1}$, $k_1 = 1$ and $k_2 = 0$.[35] In this case, forecasters do not need to update their beliefs about the past state as they receive perfect information about the current state.

Estimating this equation on the individual forecast error data gives the following estimates that suggest a k_2 that is negative and slightly significant.

Considering this scenario with parameter misperception gives the following:

$$FE_{t+1,t|t}(i) = \tilde{\rho}(1 - k_1)FE_{t,t-1|t-1}(i) - k_2 FE_{t-1,t-2|t-1}(i) - q\pi_{t,t-1} - \tilde{\rho}k_1 v_t(i) - \tilde{\rho}k_2 e_t + w_{t+1}.$$

Estimating this equation

$$FE_{t+1,t|t}(i) = -0.56^{***} + 0.47^{***} FE_{t,t-1|t-1}(i) + 0.04 FE_{t-1,t-2|t-1}(i) - 0.21^{***}\pi_{t,t-1}$$

These estimates show a k_2 not statistically different from 0 and return coefficients on the lagged forecast error and the value of inflation very similar to those in Section 4.1. In other words, once we condition on the presence of noisy private signals and misperception about persistence, there is little additional statistical gain to modeling the release of inflation data.

7.5 Forecasters Receive Public Signals

We may also be concerned that forecasters have information that is related to that of other forecasters. In the previous subsection, forecasters shared information about past realizations of inflation. They may also receive public signals of the current state in

[35] The only other case where $k_2 = 0$ is when $\sigma_e = \sigma_w + \sigma_v$, meaning that the signal on past inflation is very noisy. A standard deviation of the noise term, e_t, is equal to a standard deviation of the process innovation, w_t, and private signal noise, $v_t(i)$ put together, making it relatively large and the signal effectively uninformative.

addition to their private signals. In this case, the measurement vector is defined as

$$Z_t(i) = \begin{bmatrix} z_t(i) \\ s_t \end{bmatrix} = \begin{bmatrix} \pi_{t,t-1} \\ \pi_{t,t-1} \end{bmatrix} + \begin{bmatrix} v_t(i) \\ \zeta_t \end{bmatrix} = H\pi_{t,t-1} + \begin{bmatrix} v_t(i) \\ \zeta_t \end{bmatrix}.$$

In the measurement equation, $H = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\zeta_t \sim N(0, \sigma^2_\zeta)$. The term ζ_t is a signal noise term like $v_t(i)$ but varies only with time as the signal s_t is the same for all forecasters. Forecasters update their expectations according to:

$$\pi_{t,t-1|t}(i) = \pi_{t,t-1|t-1}(i) + K\left[Z_t(i) - H\pi_{t,t-1|t-1}(i)\right].$$

Deriving the predicted path of quarter-ahead forecast errors produces something very similar to Equation 5.

$$FE_{t+1,t|t}(i) = \rho(1 - k_1 - k_2)FE_{t,t-1|t-1}(i) - k_1\rho v_t(i) - k_2\rho\zeta_t + w_{t+1}.$$

Under noisy information, forecast errors will still depend on lagged forecast errors, but the coefficient on these forecast errors will include both components of the Kalman gain, k_1 and k_2. With forecaster misperception, the forecast errors evolve according to the following equation.

$$FE_{t+1,t|t}(i) = \tilde{\rho}(1 - k_1 - k_2)FE_{t,t-1|t-1}(i) - q\pi_{t,t-1} - k_1\tilde{\rho}v_t(i) - k_2\tilde{\rho}\zeta_t + w_{t+1}.$$

This does not substantially change the predictions of the noisy information model - with or without parameter misperception.[36] It does, however, complicate the interpre-

[36] This equation is still estimable by pooled OLS. The errors are uncorrelated with the $FE_{t,t-|t-}(i)$ as the consist only of signal noise terms that arrive after time $t-1$ and inflation innovations that arrive after time t.

tation of the coefficient on forecast errors. The share of the agent's expectation that is formed with past beliefs is now $1-k_2 - k_2$, and so this is the degree of information rigidity.[37].

8 Concluding Remarks

Expectations influence the decisions of economic agents and therefore have a clear impact on economic dynamics. As such, it is important for economists and central bankers to consider the way that economic agents form their expectations. Recent work to this end has relaxed the assumptions of full information rational expectations and considered the limitations and frictions agents face when trying to form expectations of the future. This paper contributes to this discussion by modeling forecasters facing two simultaneous frictions.

This paper documents that forecasters face two different information frictions in forming their expectations. Forecasters receive imperfect information about inflation and misperceive the structural parameters influencing inflation dynamics. This second friction creates bias in existing approaches to estimating the first. In addition to observing noisy signals, forecasters do not know the true structure of the economy. In the presence of this friction, economic agents may make errors in forecasting beyond those implied by the noise in their information. This is a relevant point for monetary policy-makers, as they must consider not only the quality and credibility of the information they release but also the beliefs of economic agents that influence the expectations formation processes.

It is important to note that the misperception friction is unlikely to occur without the noisy information friction. If forecasters have perfect information, they should be able to estimate the parameters of the economy without error. The presence of noisy

[37] This is interpretation is the same as that in Section 2. In a model with one state variable, the information friction resulting from noisy information can be expressed as $1 - KH$ for both models, the number of columns in K and number of rows in H is equal to the number of signals the agent receives.

information may help us reconcile the idea that even professional forecasters with strong professional and financial incentives to make forecasts using the correct model of the economy incorrectly estimate underlying data-generating processes. Future research may focus on the link between imperfect information and forecaster estimation of the underlying parameters of the economy. This research may also consider alternative etiologies of the parameter misperception finding such as strategic concerns on the part of forecasters, rational inattention, or omitted variable bias on the part of forecasters or the econometrician.

Forecaster misperception of parameters closely relates to the learning literature, where forecasters must form inferences about the underlying structure of the economy or apply learning methods. The noisy information model I use does not provide a framework for thinking about the source of forecaster errors in estimating parameters or how forecaster awareness of these errors could change the signal processing problem. This paper focuses on assessing the effects of such errors on expectations and on existing approaches to quantifying information rigidity. I demonstrate that forecasters misperceive the values of inflation parameters *and* make forecasts as though they face constraints in observing the variable that they are attempting to forecast. This suggests that it is time to consider the noisy information and parameter learning approaches together in a more systematic way. The evidence on time-variation in parameter misperception and noisy information can further help to structure models combining the two frictions in future work. Future work may also consider the relationship between the overestimation of persistence and the diagnostic expectations framework of Bordalo et al. (2020).

References

Afrouzi, Hassan (2017). "Strategic Inattention, Inflation Dynamics and the Non-Neutrality of Money". Mimeo.

Andrade, Philippe and Hervé Le Bihan (2013). "Inattentive Professional Forecasters". In: *Journal of Monetary Economics* 60.8, pp. 967–982.

Andrade, Philippe et al. (2016). "Fundamental Disagreement". In: *Journal of Monetary Economics* 83, pp. 106–128.

Benati, Luca (2008). "Investigating Inflation Persistence Across Monetary Regimes". In: *Quarterly Journal of Economics* 123.3, pp. 1005–1060.

Bonham, Carl S. and Richard H. Cohen (1995). "Testing the rationality of price forecasts". In: *The American Economic Review* 85 (1), pp. 284–289.

Bordalo, Pedro et al. (2020). "Overreaction in Macroeconomic Expectations". In: *American Economic Review* 110.9, pp. 2748–2782.

Bürgi, Constantin (2017). "What Do We Lose When We Average Expectations?" Mimeo.

Cogley, Timothy, Giorgio E. Primiceri, and Thomas J. Sargent (2010). "Inflation-Gap Persistence in the US". In: *American Economic Journal: Macroeconomics* 2.1, pp. 43–69.

Cogley, Timothy and Thomas J Sargent (2002). "Evolving Post-World War II U.S. Inflation Dynamics". In: *NBER Macroeconomics Annual 2001, Volume 16*. MIT Press.

Coibion, Olivier and Yuriy Gorodnichenko (2012). "What Can Survey Forecasts Tell Us About Informational Rigidities?" In: *Journal of Political Economy* 120.1, pp. 116–159.

— (2015). "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts". In: *American Economic Review* 105.8, pp. 2644–2678.

Crowe, Christopher (2010). "Consensus Forecasts and Information Aggregation". IMF Working Paper 10/178.

Dovern, Jonas et al. (2014). "Information Rigidities: Comparing Average and Individual Forecasts for a Large International Panel". In: *International Journal of Forecasting* 31, pp. 144–154.

Erceg, Christopher and Andrew Levin (2003). "Imperfect credibility and inflation persistence". In: *Journal of Monetary Economics* 50, pp. 915–944.

Fuhrer, Jeffrey (2018). "Intrinsic Expectations Persistence: Evidence from Professional and Household Survey Expectations". Federal Reserve Bank of Boston WP 18-9.

Jain, Monica (2019). "Perceived Inflation Persistence". In: *Journal of Business and Economic Statistics* 37 (1), pp. 110 –120.

Keane, Michael P. and David E. Runkle (1989). "Are Economic Forecasts Rational". In: *Quarterly Review, Federal Reserve Bank of Minneapolis* Spring.

Mackowiak, Bartosz and Mirko Wiederholt (2009). "Optimal Sticky Prices under Rational Inattention". In: *American Economic Review* 99.3, pp. 769–803.

Mankiw, N. Gregory and Ricardo Reis (2002). "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve". In: *Quarterly Journal of Economis* 117.4, pp. 1295–1328.

Milani, Fabio (2007). "Expectations, Learning, and Macroeconomic Persistence". In: *Journal of Monetary Economics* 54, pp. 2065–2082.

Mishkin, Frederic (2007). "Inflation Dynamics". In: *International Finance* 10.3, pp. 317–334.

Nimark, Kristoffer P. (2014). "A Low Dimensional Kalman Filter for Systems with Lagged States in the Measurement Equation". Mimeo.

Nordhaus, William (1987). "Forecasting Efficiency: Concepts and Applications". In: *The Review of Economics and Statistics* 69, pp. 667–674.

Orphanides, Athanasios and John Williams (2004). "Imperfect Knowledge, Inflation Expectations, and Monetary Policy". In: *The Inflation-Targeting Debate*. University of Chicago Press.

Paternoster, Raymond et al. (Nov. 1998). "Using the Correct Statistical Test for the Equality of Regression Coefficients". In: *Criminology* 36.4, pp. 859–866.

Pesaran, M. Hashem and Martin Weale (2006). "Survey Expectations". In: *The Handbook of Economic Forecasting*.

Sargent, Thomas J. (1993). *The Conquest of American Inflation*. Princeton University Press.

Sheng, Xuguang (Simon), Jonathan Wallen, and Z. An (2017). "Inattention: Measurement, Determinants and Policy Implications".

Sims, Christopher (2003). "Implications of Rational Inattention". In: *Journal of Monetary Economics* 50.3, pp. 278–291.

Stock, James H. and Mark W. Watson (2007). "Why Has U.S. Inflation Become Harder to Forecast?" In: *Journal of Money, Credit, and Banking* 39.1, pp. 3–33.

Woodford, Michael (2002). "Imperfect Common Knowledge and the Effects of Monetary Policy". In: *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund Phelps*. Princeton Univ. Press.

Metric	Estimated coefficient on	
	$FE_{t,t-1\|t-1}(i)$	$\overline{FE}_{t,t-1\|t-1=1}$
Mean	0.26	0.26
SD	0.055	0.078
Range	0.36	0.53
IQR	0.08	0.11

Table 1: Gains from individual regression

Notes: This table presents statistics on regression coefficients generated in a simulation of the baseline noisy information model of Section 2. This simulation generates data and estimates the individual and mean equations, 7 and 8, 1000 times, storing the coefficients from both regressions for each simulation. The dependent variables for these regressions are $FE_{t+1,t|t}(i)$ and $\overline{FE}_{t+1,t|t}$, respectively. The simulation is calibrated with $\rho = 0.39$ and $k = 0.31$ such that the true coefficient on forecast errors should equal $\rho(1-k) = 0.27$. The table shows that the distribution of coefficients from both the mean and the aggregate approach center around the true value, but that the individual approach leads to reduced dispersion of estimates across simultations. The standard deviation, range, and interquartile range are all lower under the panel approach.

Horizon	Individual		Aggregate	
	$FE_{t,t-h\|t-h}(i)$	Constant	$\overline{FE}_{t,t-h\|t-h}$	Constant
$h=1$	0.27***	0.01	0.22***	-0.02
	(0.02)	(0.03)	(0.08)	(0.13)
$h=2$	0.30***	-	-	-
	(0.04)	-	-	-
$h=3$	0.42***	-	-	-
	(0.05)	-	-	-
$h=4$	0.38***	-	-	-
	(0.04)	-	-	-

Table 2: Coefficients from individual and aggregate regressions, full sample

Notes: The first two columns of the above table show the regression coefficients using individual-level forecaster errors from the Survey of Professional Forecasters for different horizons. Horizons greater than one include of a time dummy to control for time-specific endogeneity. For each horizon, the current forecast error is regressed on the lag from h quarters back to avoid overlap in the realizations of inflation included in the measures of inflation. The second set of columns shows the results for a time series regression on mean forecast errors. The standard errors are Newey-West with a HAC length of h-1. *** denotes significance at the 0.01 level. N = 3582 for the panel regression at h = 1 and N = 153 for the time series regression at h = 1. See Section 3.2 in text for details.

Horizon	Individual		Aggregate	
	$FE_{t,t-h\|t-h}(i)$	Constant	$\overline{FE}_{t,t-h\|t-h}$	Constant
$h = 1$	0.27***	0.04	0.21***	-0.02
	(0.02)	(0.03)	(0.08)	(0.13)
$h = 2$	0.33***	-	-	-
	(0.04)	-	-	-
$h = 3$	0.44***	-	-	-
	(0.05)	-	-	-
$h = 4$	0.43***	-	-	-
	(0.05)	-	-	-

Table 3: Coefficients from individual and aggregate regressions, trimmed sample

Notes: This table replicates Table 2, but rather than using the full sample of forecasters includes only those with 30 or more observations. Horizons greater than one include time fixed effects. The standard errors are Newey-West with a HAC length of h-1. *** denotea significance at the 0.01 levels. See Section 3.2 for details.

Panel A: simulation	$FE_{t,t-1\|t-1}(i)$ not controlled	$\overline{FE}_{t,t-1\|t-1}$	$FE_{t,t-1\|t-1}(i)$ controlled
Baseline: $k = 0.19$, $q = 0.23$	0.27	0.20	0.45
No misperception, $k = 0.19$, $q = 0$	0.43	0.42	0.45
Scaling factor,	1.61	2.08	1.00

Panel B: Data	$FE_{t,t-1\|t-1}(i)$ not controlled	$\overline{FE}_{t,t-1\|t-1}$	$FE_{t,t-1\|t-1}(i)$ controlled
Point estimate (from data)	0.27	0.22	0.43
Scaled estimates	0.44	0.46	0.43

Table 4: Simulation: the effect of parameter misperception on the estimation of noisy information

Notes: Panel A presents evidence from simulations in the case when forecasters underestimate inflation persistence and in the counterfactual case where they do not misperceive this parameter. The dependent variables are the individual and aggregate forecast errors, $FE_{t,t-1|t-1}(i)$ (from the regressions that do and do not control for $\pi_{t,t-1}$) and $\overline{FE}_{t,t-1|t-1}$. The estimates presented are the coefficients on lagged forecast errors. The scaling factor is calculated as the coefficient when we consider only noisy information divided by the coefficient when we consider both frictions. In the simulations, I set $k = 0.19$ and $\tilde{\rho} = 0.55$ to match the SMM estimates. Panel B shows the point estimates from the data as well as these estimates scaled by the scaling factor generated in the simulated model. The composition of $\tilde{\rho}$ changes between the baseline and counterfactual simulations. In the baseline case, $\rho = 0.32$ and $q = 0.23$. The counterfactual case holds $\tilde{\rho}$ constant, as its value influences the true value of the coefficient, $\tilde{\rho}(1-k)$, but changes the value of ρ to 0.55 such that $q = 0$. I calculate the scaled estimate by multiplying the point estimate by the noisy information share for each column.

	Δ	Observed $\overline{\tilde{\rho}(1-k)}$	Unobserved $\overline{\tilde{\rho}(1-k)}$	Observed $\overline{-q}$	Unobserved . $\overline{-q}$
$\Delta\mu$	$+0.2$	0.26	0.26	0	0
	-0.2	0.26	0.26	0	0
$\Delta\rho$	$+0.10$	0.30	0.26	0	0.05
	-0.10	0.23	0.26	0	-0.05
$\Delta\sigma_w^2$	$+0.31$	0.25	0.26	0	0
	-0.23	0.28	0.26	0	0
$\Delta\sigma_v^2$	$+1.50$	0.28	0.26	0	0
	-0.65	0.24	0.26	0	0

Table 5: True values for simulated coefficients from Section 6

Notes: This table shows the calibration of the changes in parameters from Section 6.1 as well as the true values of coefficients on lagged forecast errors and on inflation when these changes are unobserved. The coefficient on inflation is equal to the average $\tilde{\rho}(1-k)$ across the change while the coefficient on inflation is equal to $-q$. Tables 7 and 6 show the average simulated coefficients for each scenario. See Section 6.1 in the text for more details.

		Observed		Unobserved	
		$FE_{t,t-1\|t-1}(i)$	$\pi_{t,t-1}$	$FE_{t,t-1\|t-1}(i)$	$\pi_{t,t-1}$
$\Delta\mu$	+	0.27	-0.02	0.26	0
	-	0.27	-0.02	0.26	0.01
$\Delta\rho$	+	0.31	-0.02	0.27	**0.04**
	-	0.22	0.02	0.27	**-0.05**
$\Delta\sigma_w^2$	+	0.24	-0.02	0.27	-0.02
	-	0.28	-0.02	0.27	-0.02
$\Delta\sigma_v^2$	+	0.29	-0.02	0.27	-0.02
	-	0.25	-0.02	0.27	-0.02

Table 6: Coefficients from panel regression including lag of inflation with changing parameters

Notes: This table presents the estimates from the pooled regression on forecast errors and lagged inflation. The dependent variables are the individual and aggregate forecast errors, $FE_{t,t-1|t-1}(i)$ and $\pi_{t,t-1}$, respectively. The estimates presented are the mean coefficients on these variables. For each variable, σ_w^2, σ_v^2, ρ and μ, I simulate the model with both an increase and a decrease in that variable. I also perform a simulation in which forecasters observe the change and incorporate it into their optimal action and one in which they do not observe the change. The calibration of the changes as well as the true values of the parameters appear in Table 5. For more details, see Section 6.1.

		Observed		Unobserved	
		$FE_{t,t-1\|t-1}(i)$	$\overline{FE}_{t,t-1\|t-1}$	$FE_{t,t-1\|t-1}(i)$	$\overline{FE}_{t,t-1\|t-1}$
$\Delta\mu$	+	-0.01	-0.02	0	0.01
	-	-0.01	-0.02	0.01	0.01
$\Delta\rho$	+	0.00	-0.02	**0.04**	**0.05**
	-	-0.02	-0.02	**-0.05**	**-0.06**
$\Delta\sigma^2_w$	+	-0.02	-0.02	-0.01	-0.02
	-	-0.02	-0.02	-0.01	-0.02
$\Delta\sigma^2_v$	+	-0.02	-0.02	-0.01	-0.02
	-	-0.02	-0.02	-0.01	-0.02

Table 7: Deviations from true values for coefficients on individual and aggregate regressions with changing Parameters

Notes: This table presents results from the individual and aggregate regressions of forecast errors on their own lags with data simulated according to the noisy information model with changes in the variables. The dependent variables are the individual and aggregate forecast errors, $FE_{t,t-1|t-1}(i)$ and $\overline{FE}_{t,t-1|t-1}$, respectively. The estimates presented are the deviations of mean coefficients on lagged forecast errors from each approach from the true values of these coefficients implied by model parameters. For each variable, σ^2_w, σ^2_v ρ and μ, I simulate the model with both an increase and a decrease in that variable. I also perform a simulation in which forecasters observe the change and incorporate it into their optimal action and one in which they do not observe the change. The calibration of the changes as well as the true values of the parameters appear in Table 5. For more details, see Section 6.1.

	$\pi_{t+1,t}$		$\pi_{t+1,t\|t}(i)$
$\pi_{t,t-1}$	0.55^{***} (0.08)	$\pi_{t,t-1\|t}(i)$	0.36^{***} (0.02)
$g^y_{t+1,t}$	0.03^{*} (0.06)	$g^y_{t+1,t\|t}(i)$	-0 .01 (0.01)
$u_{t+1,t}$	-0.03^{*} (0.09)	$u_{t+1,t\|t}(i)$	0.09^{***} (0.01)
$i_{t+1,t}$	0.30^{***} (0.06)	$i_{t+1,t\|t}(i)$	0.20^{***} (0.01)

Table 8: Relationship of inflation and its forecast to other variables

Notes: This table shows the relationship of CPI inflation to other variables as well as its own lagged value. The second column shows the relationship of the forecast of inflation to the forecasts of other variables and the inflation nowcast.

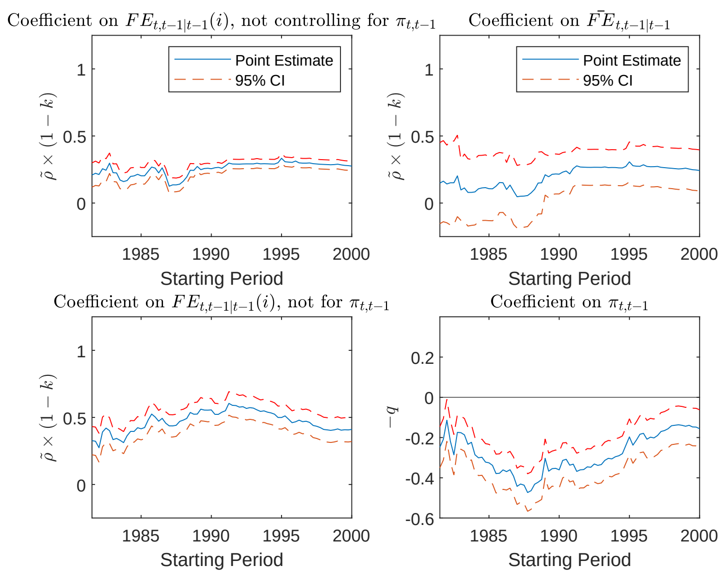

Figure 1: Time variation in information rigidity and parameter misperception

Notes: This figure shows results for rolling window regressions on Equations 7, 8 and 12. The top two plots show the coefficients from the individual and aggregate approaches described in Section 2. The bottom two plots show the coefficients when the value of inflation is included in the individual regression. I plot the point estimate for each coefficient along with its 95% confidence interval. See Section 7.1 for details.

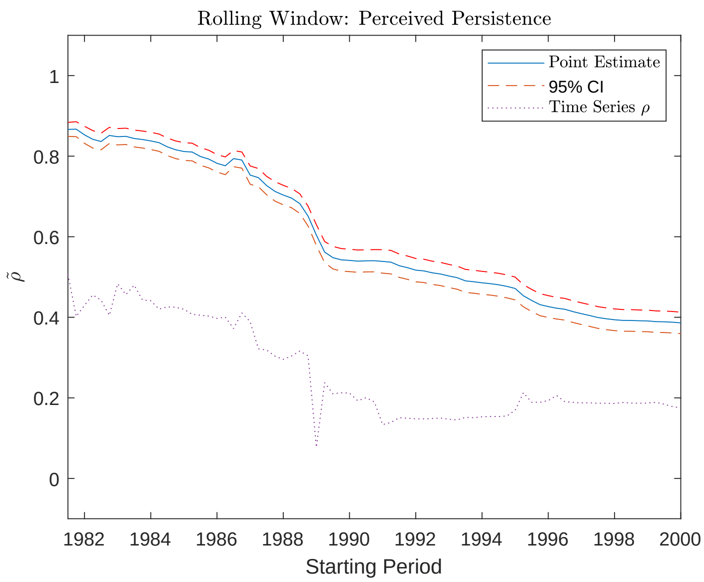

Figure 2: Time variation in perceived persistence

Notes: This figure shows an 80-window rolling regression of Equation 14, or time-variation in the forecasters' perceived persistence. The coefficient measuring perceived persistence as well as its 95% confidence interval is plotted against the first time period in each window. For reference, the persistence estimate from the time series regression on the process is also included. See Sections 4.3 and 7.1 for more details.

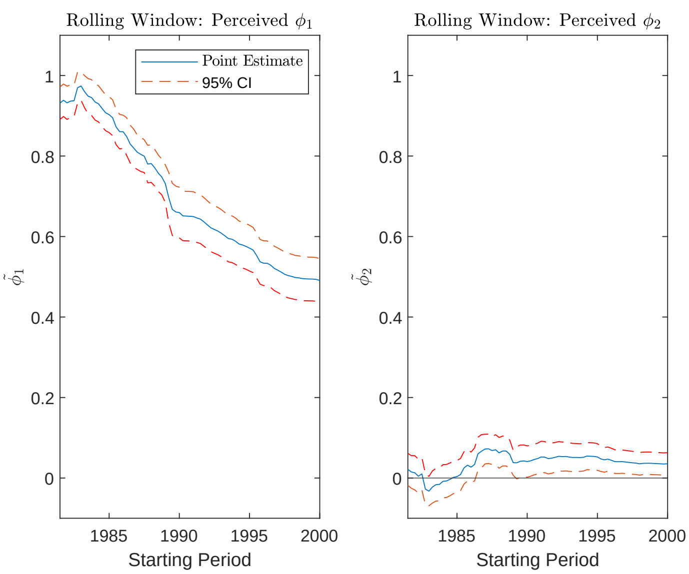

Figure 3: Time variation in perceived AR(2) parameters

Notes: This figure shows an 80-window rolling regression of Equation 21, or time-variation in the forecasters' perceived AR(2) parameters. The coefficients measuring perceived persistence as well as its 95% confidence interval is plotted against the first time period in each window. For reference, the persistence estimate from the time series regression on the process is also included. See Sections 7.1 and 7.2 for more details.