The Dissertation Committee for Jane Maria Ryngaert
certifies that this is the approved version of the following dissertation:

Essays on Inflation Expectations and Information Frictions

Committee:

Olivier Coibion, Supervisor

Saroj Bhattarai

Christoph Boehm

Tara Sinclair

Essays on Inflation Expectations and Information Frictions

by

Jane Maria Ryngaert

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2018

Dedicated to my parents,

To John,

And to Estella, my greatest expectation.

Acknowledgments

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, David Papell as well as with many seminar participants from the University of Texas, The George Washington University, and the Federal Reserve Bank of Chicago.

I also particularly thank Vivian Goldman-Leffler and Katie Foster. And of course, I thank my family and friends for the constant support and encouragement.

Essays on Inflation Expectations and Information Frictions

Publication No. __________________

Jane Maria Ryngaert, Ph.D.
The University of Texas at Austin, 2018

Supervisor: Olivier Coibion

This dissertation empirically investigates the expectations formation process and the constraints that economic agents face in forming beliefs about macroeconomic variables. Chapter 1 contributes to and extends our current understanding of information frictions in expectations. I first propose a new framework for estimating noisy information using individual forecasts, rather than mean forecasts as commonly done in previous work. This approach provides more power for identifying underlying information rigidities. I further extend this framework to incorporate misperceptions on the part of economic agents about the persistence of the underlying process being forecasted. Applying this framework to the U.S. inflation forecasts of professional forecasters points toward significantly less noisy information than previous estimates suggest but reveals a systematic underestimation on the part of forecasters of the persistence of inflation. Using a structural model that incorporates both

noisy signals and misperceptions of persistence, I quantify the relative importance of each channel in accounting for the expectations formation process of these agents. The results indicate that, even for professional forecasters, there are multiple forces that generate economically significant deviations from full information.

Chapter 2 is joint work with Olivier Coibion, Yuriy Gorodnichenko, and Saten Kumar. Using novel survey questions on the higher-order expectations of firm managers, we study the formation and evolution of these beliefs. A unique experimental approach allows us to characterize the degree of higher-order thinking of economic agents and how this degree of higher-order thinking affects managers? expectations as well as their economic decisions. We then relate these results to macroeconomic models in which higher-order thinking matters for dynamics.

Chapter 3 is develops a method for measuring the information flow of economic agents at a given point in time using survey data. I document a reduction in attention to several macroeconomic variables over time. I further document that in periods in which agents are paying more attention to a specific variable, there is also greater cross-sectional dispersion in attention across agents.

Table of Contents

List of Tables

List of Figures

Chapter 1

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

1.1 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 comes from the fact that, in any period that she does not update, she based 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 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

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

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 late 1960s, I find estimates of noisy information implying that forecasters weigh new signals slightly more than they weigh prior beliefs in forming their expectations. As they still form almost 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 cannot 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 pre-

[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.

dicted path of forecast errors given both frictions: 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 late 1960s, forecasters have on average underestimated inflation persistence. This underestimation 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 forecaster underestimation of persistence such as Jain (2017).

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 explaining 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 do face both real-time information constraints and underestimate inflation persistence. These estimates further show that information is much less noisy than previous estimates of the noisy information friction (that do not take the misperception friction into account) suggest.[3]

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

The structural estimates imply that forecasters base as much as 66 percent of their expectations on new information, leaving only 34 percent of the expectation to be formed with prior information. While forecasters still face an economically relevant observation friction, they respond more to new signals than previous literature reports. However, the underestimation of inflation persistence creates a another relevant friction for expectations formation as forecasters will project the state forward using the wrong transition equation. Even in the case where their signals are highly credible, forecasters will make the wrong projections about the future path of inflation and fail to recognize 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. Andrade et al. (2016) is thematically similar to my paper as it combines multiple frictions limitations 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 terms structure of forecaster disagreement in the Blue Chip Financial Forecasts. 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 et al. (2017) who perform nonparametric analyses on forecasters from both the SPF and Consensus Economics.

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 et al. (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 (2017). 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. This is consistent with the findings from the current paper, though my findings do not differ as radically from time series estimates as Jain's. My paper provides further insights by deriving not only perceived persistence, but also the perceived constant in inflation.[4]

[4]These two together generate perceived long-run beliefs about inflation. Forecasts of long run beliefs are of particular importance to the Federal Reserve in assessing the anchoring of

The remainder of the paper proceeds as follows. In Section 3.2, I describe the simple noisy information model. Section 3.4 discusses the data and the initial results. Section 1.4 describes the model with an additional source of frictions and presents evidence for the underestimation of persistence. Section 1.5 describe the structural model used in simulations as well as its application to estimation. Section 1.6 discusses other potential explanations for the differences in the individual and aggregate estimates. Section 1.7 examines extensions to the model, including time variation in parameter misperception, forecaster observation of realizations of inflation, and the noisy information model with public signals. Section 3.7 concludes.

1.2 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. In-

agent expectations and the stability of these anchors to macroeconomic events and shocks. The Survey of Professional forecasters recognized this importance and began tracking long run forecasts for CPI inflation in 1991. My approach can be used to track low frequency movements in these beliefs.

novations 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 $\frac{\mu}{1-\rho}$.

$$\pi_{t,t-1} = \mu + \rho\pi_{t-1,t-2} + w_t \tag{1.1}$$

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.

$$z_t(i) = \pi_{t,t-1} + v_t(i) \tag{1.2}$$

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]:

$$\pi_{t,t-1|t}(i) = kz_t(i) + (1-k)\pi_{t,t-1|t-1}(i). \tag{1.3}$$

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 1.3, (1-k) represents the percentage of the new expectation based on prior 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

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

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

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 1.1.

$$\begin{aligned}\pi_{t+1,t|t}(i) &= \mu + \rho\pi_{t,t-1|t}(i) \\ &= \mu + k\rho\pi_{t,t-1} + k\rho v_t(i) + \rho(1-k)\pi_{t,t-1|t-1}(i)\end{aligned} \tag{1.4}$$

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:

$$FE_{t+1,t|t}(i) = \rho(1-k)FE_{t,t-1|t-1}(i) + w_{t+1} - k\rho v_t(i). \tag{1.5}$$

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 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 now face constraints in observing past, present, and future innovations to the inflation process.

Taking the average of Equation 1.5 across agents gives the following relationship between consensus forecasts and their lags.

$$\overline{FE}_{t+1,t|t}(i) = \rho(1-k)\overline{FE}_{t,t-1|t-1}(i) + w_{t+1} \tag{1.6}$$

This is the standard approach used to estimate information rigidities, e.g. Coibion and Gorodnichenko (2012). The primary difference between this equation and Equation 1.5, 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.

$$FE_{t+1,t|t}(i) = \beta_0 + \beta_1 FE_{t,t-1|t-1}(i) + \varepsilon_t(i) \tag{1.7}$$

$$\overline{FE}_{t+1,t|t}(i) = \beta_0 + \beta_1 \overline{FE}_{t,t-1|t-1}(i) + \varepsilon_t \tag{1.8}$$

The model predicts that, for both Equations 1.7 and 1.8, $\beta_0 = 0$ 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]

1.2.1 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 $\rho = 0.9$ and $k = 0.5$, 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 0.84 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 $k = 0.5$ as the optimal Kalman gain, we must adjust σ_v^2 such that the relative relationship between the two variances is the same.

[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 1.5.

The simulated data matches the length of the sample period ($F = 195$) with a short 100-period burn-in period and contains $N = 249$ 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 1.7 and 1.8 and record the coefficients for 1000 rounds of simulated data.

Table 1.1 shows the gains from my approach over the aggregate regression. Across the 1000 rounds of the simulation, both the mean and aggregate estimate recover approximately the correct coefficient on lagged forecast errors. The individual approach leads to a 20 percent reduction in the standard deviation of the 1000 estimates and a 22 percent reduction in the interquartile range of these estimates.

1.2.2 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 error 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.[10]

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

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

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

[10]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.

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 $FE_{t,t-2|t-2}(i)$. [11] 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.

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

$$FE_{t+h,t|t}(i) = \beta_0 + \beta_1 FE_{t,t-h|t-h}(i) + \sum_t \gamma(t)\mathbf{1}(t) + \varepsilon_t(i) \tag{1.9}$$

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.[12] Moreover, as the aggregate specification is a time series, time period dummies 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.

[11] See 2.4. in Appendix 2

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

1.3 Forecaster Data and Noisy Information Predictions

1.3.1 Forecast Error Data

I use forecasts of the GDP deflator series 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. This quarterly availability allows me to calculate forecast errors across subsequent periods. I use GDP deflator inflation as opposed to CPI or PCE inflation because the survey contains forecaster responses predictions GDP deflator since its inception in 1968Q4. CPI and PCE inflation were added in 1980Q3 and 2007Q1, respectively.

I calculate annualized anticipated h-quarter GDP deflator inflation in time-t and time-t-1 using the following equations.[13]

$$\pi_{t+h,t|t}(i) = \left[\left(\frac{P_{t+h}(i)}{P_t(i)}\right)^{\frac{4}{h}} - 1\right] \times 100$$

Data on the realization of inflation comes from the St. Louis FRED GDP deflator series for the corresponding sample. Realizations of inflation are

[13]For a one-quarter ahead forecast, $P_{t+h}(i)$ is given by PGDP3 and $P_t(i)$ is given by PGDP2.

formed from the final release measures of the GDP deflator.

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

1.3.2 Results from Noisy Information Models

Using this data, I estimate Equations 1.7 and 1.8 for one-quarter ahead forecast errors. For forecasts at horizons greater than one, I estimate 1.9. The results appear in Table 1.2.

For 1-quarter ahead forecasts, both the pooled and time series regressions produce estimates that imply a non-trivial amount of information rigidity. As we would expect, the individual forecaster results are much more precise due to a larger sample size with Newey-West standard errors roughly 60 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 $k << 1$. For the individual forecast, the estimate 0.43 implies a point estimate[14] of k as high as 0.57, indicating that forecasters base up to 57 percent of their new expectation with their most recent signal. This estimate is formed assuming a $\rho = 1$. As ρ may be less than one, this estimate is an upper bound on k. The aggregate estimate,

[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.

0.53, implies an upper bound on k of 0.47.[15]. 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. The p-value of the test that the coefficients from the individual and aggregate regressions are equal is 0.24 when the estimation is performed allowing for Newey-West standard errors. Using the simulation in Section 1.2.1, however, none of the 1000 iterations generates a difference between the time series coefficient and the pooled coefficient as large as 0.10. This means that the basic noisy information model is unlikely to produce such a difference across estimation equations. Comparing point estimates, the individual approach indicates that forecasters base roughly ten percent more of their new forecasts on new signals. This further suggests that signals are not as noisy as the aggregate approach shows and forecasters can trust their information relatively more. Dovern et al. (2014) also finds evidence for a greater amount of implied information rigidity in estimates obtained from mean forecasts rather than individual forecasts.[16]

The regressions for longer-horizon forecast errors ($h > 1$) include time fixed effects to control for the heterogeneity that arises from the quarterly

[15]Using $\rho = 0.89$ to be consistent with time series estimation of the transition equation gives estimates of $\hat{k}_{individual} = 0.52$ and $\hat{k}_{aggregate} = 0.4$.

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

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 1.2 shows that we do not see this decline in the data. This may result from omitted variable bias coming from parameter misperception.[17]

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 1.3 presents the same estimates from Table 1.2, but limiting the sample to forecasters who remain in the sample for at least 30 periods. While the wedge between coefficients declines slightly, there is little 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 expected. Accordingly, for the rest of the analysis I use all the forecasters in the sample.

1.4 An Additional Source of Information Frictions

Noisy information affects first 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

[17] See Section 1.4.

to form expectations of future inflation. 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 between forecast errors over time.

1.4.1 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 1.5.[18]

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 1.4 takes the same form with one adjustment.

$$\begin{aligned} \pi_{t+1,t|t}(i) &= \mu + \tilde{\rho}\pi_{t,t-1|t}(i) \\ &= \mu + k\tilde{\rho}\pi_{t,t-1} + k\tilde{\rho}v_t(i) + \tilde{\rho}(1-k)\pi_{t,t-1|t-1}(i) \end{aligned} \tag{1.10}$$

[18]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 3 shows this in detail.

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:

$$FE_{t+1,1|t}(i) = \tilde{\rho}(1-k)FE_{t,t-1|t-1}(i) - q\pi_{t,t-1} + w_{t+1} - \rho k v_t(i) \qquad (1.11)$$

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.

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 1.5 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.[19] *** denotes significance at the 0.01 level.

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

$$\begin{aligned} FE_{t+1,t}(i) &= \beta_0 + \beta_1 FE_{t,t-1}(i) + \beta_2 \pi_{t,t-1}(i) + \epsilon_t(i) \\ &= -0.20^{***} + 0.38^{***} FE_{t,t-1|t-1}(i) + 0.08^{***} \pi_{t,t-1} \end{aligned} \tag{1.12}$$

Under the null that there is no misperception of persistence, $\beta_2 = 0$. This estimation implies that $q = -0.08$, meaning that agents underestimate the persistence of the inflation process. The underestimation of persistence is consistent with Jain (2017)'s findings, though she finds a much more dramatic deviation from time series estimates of ρ than my result suggests.[20] 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 this specification also detects noisiness in information. [21]

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 positive, omitting inflation from the regression will lead to a positive bias on lagged forecast errors. In the context of this model, an upwardly biased coefficient will lead us to conclude that forecasters receive signals that are noisier than they actually are. This can explain why β_1 is lower than it appears in Table 1.2, where $\pi_{t,t-1}$ is omitted.

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

[21] The negative value of the β_0 implies that forecasters underestimation the constant by -0.20. See Appendix 3.

1.4.2 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:[22]

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

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 as the time fixed effects terms will absorb the time-specific inflation terms.

1.4.3 Evidence from the Term Structure of Forecasts

Section 1.4 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

[22]This is derived in Appendix 2.2.

structure of forecast data available in the Survey of Professional Forecasters. Each period, respondents forecast the target variable several periods ahead, indicating how they believe the variable will develop over time. To generate an estimate of perceived inflation persistence for GDP deflator inflation, I form the following measures of annualized one-quarter ahead forecasts for each forecaster:

$$\pi_{t+h,t+h-1|t}(i) = \left[\left(\frac{P_{t+h}}{P_{t+h-1}} \right)^4 - 1 \right]$$

For this exercise, $\pi_{t+h,t+h-1|t}(i)$ represents agent i's forecast of inflation from period $t+h-1$ to $t+h$ given information available in period t.

Applying the process implied by the transition equation in Equation 1.1, we can construct the relationships between each forecast the agent makes in period t.

$$\begin{aligned} \pi_{t+h,t+h-1|t}(i) &= \tilde{\mu} + \tilde{\rho}\pi_{t+h-1,t+h-2|t}(i) + E[w_{t+1}] \\ &= \tilde{\mu} + \tilde{\rho}\pi_{t+h-1,t+h-2|t}(i) \end{aligned} \tag{1.13}$$

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

[23]Differing from Jain (2017), 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.

to both information availability differences and forecaster perceptions of the nowcast.[24] Accordingly, consider the following system of equations.

$$\pi_{t+s,t+s-1|t}(i) = \beta_0 + \beta_1 \pi_{t+s-1,t+s-2|t}(i) + \epsilon_{st,i}, \ldots s = 1,2,3 \tag{1.14}$$

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.[25]

$$\pi_{t+s|t+s-1}(i) = \beta_0 + \beta_1 \pi_{t+s-1,t+s-2|t}(i) + \sum_{s=2}^{4} \gamma_s 1(s) + \epsilon_t(i) \tag{1.15}$$

The regression coefficient and constant provide estimates of the process parameters, $\beta_0 = \hat{\hat{\mu}} = 0.78(0.05)$ and $\beta_1 = \hat{\hat{\rho}} = 0.76(0.01)$. [26]

Running a corresponding AR(1) regression on the inflation process gives $\hat{\rho} = 0.89^{***}(0.04)$ and $\mu = 0.36^{**}(0.14)$, where *** and ** denote significance at the 0.01 and 0.05 percent levels, respectively. Agents therefore significantly

[24] Forecasters may, for example, believe that their nowcast should represent a prediction of the revision the statistical agency will make to its preliminary estimate.

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

[26] 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.

underestimate the persistence of inflation and significantly overestimate 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.

There is 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.

1.5 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:

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

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 3.2. 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 1.6.

Inflation evolves and signals arrive according to the following transition and measurement equations.[27]

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

Forecasters again receive private signals:

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

Agents process information through the Kalman filter, selecting the optimal gain for the given level of persistence (as they perceive it), process

[27]This follows from Equations 1.1 and 1.2

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 1.4.1.

I generate data for the length of the sample period in the data ($F = 195$) 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 $N = 249$ forecasters. The simulated data following the burn-in period is 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.[28] 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.

[28]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 for each round. 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, ε_a and ε_c are time and individual specific, whereas ε_c varies only with time.

$$FE_{t+1,t|t}(i) = \beta_0^a + \beta_1^a FE_{t,t-1|t-1}(i) + \varepsilon_t^a(i) \tag{1.16}$$

$$\overline{FE}_{t+1,t|t} = \beta_0^b + \beta_1^b \overline{FE}_{t,t-1|t-1} + \varepsilon_t^b \tag{1.17}$$

$$FE_{t+1,t|t}(i) = \beta_0^c + \beta_1^c FE_{t,t-1|t-1}(i) + \beta_2^c \pi_{t,t-1} + \varepsilon_t^c(i) \tag{1.18}$$

For a completed simulation I calculate the average values of the simulated coefficients using the coefficient on lagged forecast errors in all three equations and the coefficient on lagged inflation in Equation 1.18. I define:

$$g(\theta) = \left[\tfrac{1}{M}\textstyle\sum_m \beta_1^a(m) \quad \tfrac{1}{M}\textstyle\sum_m \beta_1^b(m) \quad \tfrac{1}{M}\textstyle\sum_m \beta_1^c(m) \quad \tfrac{1}{M}\textstyle\sum_m \beta_2^c(m)\right]$$

The simulated coefficients can be compared to the same set of coefficients from the data: $\hat{g} = \left[0.43, 0.53, 0.38, 0.08\right]$.

1.5.1 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.

$$J(\theta) = [\hat{g} - g(\theta)]\Omega[\hat{g} - g(\theta)]' \tag{1.19}$$

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:[29]

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, $q_0 \in (-\rho_0, 1 - \rho_0)$.

[29]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.

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.[30]. 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.[31]

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 $\hat{\rho} = 0.88(0.005)$, $\hat{q} = -0.18(0.004)$, and $\hat{k} = 0.66(0.004)$. Interestingly,

[30]Occasionally accepting θ_1 that does not produce $J_1 < J_0$ helps the algorithm avoid getting stuck in a local minimum. Given infinite time to run, this algorithm should search the entire space of parameters, spending the most time at its global minimum.

[31]This should be thousands of iterations at least. Note also that the first part of the thread will be removed as a burn-in period and the parameters calculated from only the remaining part of the thread.

the estimate for $\hat{\rho}$ is similar to time series estimates, despite these estimates not appearing in the set of moments used to estimate parameters. The sizeable underestimation of persistence is surprising, but is largely driven by the sub-optimality of the agent's choice of gain. This causes the wrong weighting of the noise terms and biases the coefficient on inflation from Equation 1.18 downwards, meaning it takes larger absolute values of q to match that moment. The estimate of k means that forecasters will form 66 percent of their new expectations with new information. This is consistent with the recent findings of Afrouzi (2017), which also derives a relatively low degree of information rigidity of $k \approx 0.7$ for firm managers in New Zealand. My estimate is also substantially lower than estimates from Dovern et al. (2014) and Coibion and Gorodnichenko (2015) and implies that information is much less noisy than those estimates suggest. These papers find a Kalman gain of 0.5. Assuming $\rho = 0.88$, my estimate implies a half-life of forecast errors of 1.8 months, while the 0.5 estimate implies a half-life of 2.5 months. Accordingly, it takes nearly a month longer to reduce the forecast error by half under previously estimated values of the Kalman gain.

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 1.16 and 1.17. The individual coefficient takes an average value of 0.44, while the average coefficient on mean forecast errors takes a value of 0.47. The mean coefficient is higher than the individual coefficient, as we see in the data, but we cannot match the magnitude of the difference in coefficients documented in Section 3.5. The true value of the coefficient under this specification is $\tilde{\rho}(1-k) = 0.24$.[32] When forecasters underestimate persistence, both the individual and panel approaches result in upward bias in the coefficient, making information appear more noisy than it actually is. The average simulated point estimates of β_1^c and β_2^c from Equation 1.18 are 0.38 and 0.09, respectively. These match the corresponding moments in the data almost exactly.

1.5.2 The Effect of Parameter Misperception

To quantify the relative importance of parameter misperception and noisy information in generating the patterns of forecaster expectations, I shut down the underestimation of persistence as a friction in the simulation ($k = 0.66$, $q = 0$). Without the underestimation of persistence, the counterfactual coefficients from the aggregate and individual regressions come very close both to each other and to the true value of the coefficient on lagged forecast errors. When I add the misperception friction back to the simulation ($k = 0.66$, $q = -0.17$), both coefficients become biased upwards and begin to separate from each other. These estimates appear in Table 1.4. The aggregate point estimate is higher than the individual point estimate. Of the overall predictability of forecast errors, the noisy information friction can account for 64 percent at the individual level and and 60 percent at the aggregate level. This is calculated as the counterfactual estimate with no misperception divided by the estimate

[32] $\tilde{\rho}(1-k) = (0.88 - 0.17) \times (1 - 0.66) = 0.24$.

with both frictions present. The remaining 36 percent and 40 percent of the predictability in forecast errors comes from the underestimation of persistence. This indicates that both misperception and noisy information are relevant for generating the predictability of forecast errors.

A back-of-the-envelope calculation using these estimates of the relative importance of each friction on the estimates from the data suggests that, in the absence of parameter misperception, the individual and aggregate approaches would return coefficients on forecast errors of 0.28 and 0.32, respectively.[33] These scaled estimates are remarkably similar to the true value of $\rho(1-k)$ used in the simulation (0.30). This is an interesting result as the simulation has a hard time matching the magnitude of the differences in the aggregate and individual point estimates for any realistic value of q.[34]

1.6 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.

[33] To find the scaled estimates, I multiply the point estimate by the noisy information share as estimated in the simulation. For the individual forecasts $0.43 \times 0.64 = 0.28$ and for the aggregate, $0.53 \times 0.60 = 0.32$.

[34] Identification in the MCMC comes primarily from Equation 1.18.

1.6.1 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 Great Inflation and Volcker Disinflation 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 et al. (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.[35] In light of these concerns, I consider possible changes in the inflation process as well as changes in the agent's signal processing parameters 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, U^-, ρ and μ, I simulate the data-generating process described in Section 1.5.1 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 allows forecasters

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

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 results of the estimation for Equations 1.16 and 1.17 for each simulation appear in Table 1.6 and the estimated coefficients from Equation 1.18 appear in Table 1.7.

I calibrate the model such that the starting values match those estimated via simulated method of moments. As such, the starting values are: $k = 0.66$, $\rho = 0.88$, $\sigma_w^2 = 0.5$, and $\sigma_v^2 = 0.35$. The starting steady state value of U^- is set to 0.68 by the other parameters. I set the misperception, q, to zero to test alternative theories. The magnitudes of the changes in each parameter appear in Table 1.5. 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.

- **Agent Uncertainty**: U^- describes the forecaster's uncertainty about inflation before she receives her signal each period. If this uncertainty is higher, the signal is relatively more valuable and the forecaster will assign a higher Kalman gain to her new information. A change in uncertainty is also unable to replicate the moments from the data.

- **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.[36] The simulated estimates show, however, that a change in the inflation constant cannot generate a wedge between coefficients or a positive coefficient on inflation.

- **Inflation Persistence**: In the simulated model, an increase in inflation persistence is the only change that can generate the difference between the individual and aggregate approaches. It is also the only change in

[36] See Appendix 3.

parameters that leads to a significantly positive coefficient on lagged inflation when the pooled regression is performed using lagged forecast errors and the lagged value of inflation. However, these features only arise when forecasters are not permitted to observe the increase in persistence and therefore underestimate persistence, meaning the underestimation of persistence drives this finding.[37]

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.

1.7 Extensions

In this section I consider three 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. I then relax the assumption that forecasters do not observe realizations of inflation by modeling the noisy information

[37]It is possible to generate similar results when forecasters start the period underestimating persistence and then persistence declines and either reduces or erases the effect of their underestimation. In this sense, this finding is not out of line with the literature that argues for a decline of inflation persistence over time.

model where forecasters receive signals about the most recent past realization of inflation as well as current inflation. A third 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.

1.7.1 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 1.7 , 1.8 and 1.12. The window width is 80 quarters, so each estimate is formed with 20 years of data. Figure 1.1 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 become particularly large as the window covers longer portions of the Great Recession and recovery periods post-2007. The coefficient on forecast errors is bounded by the interval 0 to 1. In the windows covering the last portion of the sample, the standard errors on the coefficient from the aggregate approach are so large that every value in this interval is contained in 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 1.1 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 underestimate persistence early in the sample but that, once the 1970s are no longer included in the window, forecasters overestimate persistence. After the 1980s are no longer included in the window, the coefficient and therefore the magnitude of the misperception are not significantly different from 0.

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 1970s with an expectation that inflation would follow the same low persistence as in the 1960s, they would underestimate its true persistence. By the same logic, if by the 1980s, their perceived value of persistence had increased or adapted to the higher level, they would end up overestimating persistence upon its decline. As forecasters then became accustomed to the lower level of inflation persistence, the misperception coefficient would go to 0, as we see.

Figure 1.2 shows that perceived persistence follows approximately this pattern. The figure plots an 80-quarter rolling window regression of $\tilde{\rho}$ as estimated from the term structure of forecasts (Equation 1.14). Perceived persistence is low as the interval moves into the 1970s, a time of highly persistent inflation according to Cogley and Sargent (2002), and increases as the interval nears the 1980s, a period marked by lower persistence. Forecasters enter the 1980s with a higher value of perceived persistence which then gradually adjusts downwards. Following a brief increase for windows beginning in the late 1980s, the perception of persistence begins declining over the remainder of the sample. 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.

1.7.2 Forecasters Receive Signals about Past Inflation

In Section 3.2, 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 questionaires 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 1.1. 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:

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

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.[38]

$$K = [\rho U^- (H_1 \rho + H_2)' + \sigma_w^2 H_1' + \sigma_w R'] \times [(H_1 \rho + H_2) U^- (H_1 \rho + H_2)' + (H_1 \sigma_w + R)(H_1 \sigma_w + R)']$$

The one quarter ahead forecast error in this example is:

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

[38] In this specification, k_1 is the weight on the private signal and corresponds to k in Section 3.2 while k_2 is the weight the forecaster will optimally assign to the public signal.

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 equation the same way as in Section 3.2. 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 $k_2 = 0$.[39] 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.

$$FE_{t+1,t|t}(i) = 0.04 + 0.43^{***} FE_{t,t-1|t-1}(i) + 0.07^{**} FE_{t-1,t-2|t-1}(i)$$

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

$$FE_{t+1,t|t}(i) = -0.19^{***} + 0.41^{***} FE_{t,t-1|t-1}(i) + 0.04 FE_{t-1,t-2|t-1}(i) - 0.07^{***} \pi_{t,t-1}$$

[39] 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. Derivations of k_1 and k_2 appear in Appendix 1.3.

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 1.4.1. 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.

1.7.3 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 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 $\zeta_t \sim N(0, \sigma_\zeta^2)$. 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 error produces something very similar to Equation 1.5.

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

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.[40] It does, however, complicate the interpretation 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.[41].

1.8 Concluding Remarks

Expectations influence the decisions of economic agents and therefore have a clear impact on economic dynamics. As such, it is important for

[40]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.

[41]This is interpretation is the same as that in Section 3.2. 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.

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. Joint estimation of the two frictions shows that information is less noisy than estimates that do not control for parameter misperception. This is an economically important finding as it means that forecasters actually do utilize most of the information available to them. It suggests, however, a second friction that creates problems for expectations formation, that is that forecasters do not know the true structure of the economy. In the presence of this friction, economic agents may receive full information but may still make errors in forecasting beyond the full information rational expectations error. 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.

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.

Table 1.1: Gains from Individual Regression

Metric	Estimated Coefficient on	
	$FE_{t,t-1\|t-1}(i)$	$\overline{FE}_{t,t-1\|t-1=1}$
Mean	0.44	0.44
Standard Deviation	0.052	0.065
Range	0.30	0.39
IQR	0.07	0.09

Notes: This table presents statistics on regression coefficients generated in a simulation of the baseline noisy information model of Section 3.2. This simulation generates data and estimates the individual and mean equations, 1.7 and 1.8, 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 $\rho = 0.9$ and $k = 0.5$ such that the true coefficient on forecast errors should equal $\rho(1-k) = 0.45$. 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.

Table 1.2: Coefficients from Individual and Aggregate Regressions, Full Sample

Horizon	Individual		Aggregate	
	$FE_{t,t-h\|t-h}(i)$	Constant	$\overline{FE}_{t,t-h\|t-h}$	Constant
$h=1$	0.43***	0.08***	0.53***	0.01
	(0.03)	(0.02)	(0.08)	(0.08)
$h=2$	0.24***	-	-	-
	(0.05)	-	-	-
$h=3$	0.21***	-	-	-
	(0.05)	-	-	-
$h=4$	0.29***	-	-	-
	(0.05)	-	-	-

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. $N = 4548$ for the panel regression at $h = 1$ and $N = 195$ for the time series regression at $h = 1$. See Section 3.5 in text for details.

Table 1.3: Coefficients from Individual and Aggregate Regressions, Trimmed Sample

Horizon	Individual		Aggregate	
	$FE_{t,t-h\|t-h}(i)$	Constant	$\overline{FE}_{t,t-h\|t-h}$	Constant
$h = 1$	0.45***	0.08***	0.52***	0.01
	(0.03)	(0.03)	(0.08)	(0.08)
$h = 2$	0.33***	-	-	-
	(0.05)	-	-	-
$h = 3$	0.25***	-	-	-
	(0.06)	-	-	-
$h = 4$	0.35***	-	-	-
	(0.05)	-	-	-

Notes: This table replicates Table 1.2, 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. *** denotes significance at the 0.01 level. See Section 3.5 for details.

Table 1.4: Simulation: The Share of Predictability Coming from Noisy Signals

Panel A: Simulation		
	$FE_{t,t-1\|t-1}(i)$	$\overline{FE}_{t,t-1\|t-1}$
Baseline: $k = 0.66$, $q = -0.17$	0.44	0.46
No Misperception, $k = 0.66$, $q = 0$	0.28	0.28
Scaling Factor, Noisy Information Share	65%	60%

Panel B: Data		
	$FE_{t,t-1\|t-1}(i)$	$\overline{FE}_{t,t-1\|t-1}$
Point Estimate	0.43	0.53
Scaled Estimates	0.28	0.32

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, $FE_{t,t-1|t-1}(i)$ and $\overline{FE}_{t,t-1|t-1}$, respectively. The estimates presented are the coefficients on lagged forecast errors. The noisy information share is calculated as the percentage of the coefficient on lagged forecast errors when both frictions are present that is generated when only noisy information is present. This is the second row of the table divided by the first. In the simulations, I set $k = 0.66$ and $\rho = 0.88$ 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. I calculate the scaled estimate by multiplying the point estimate by the noisy information share for each column.

Table 1.5: Data Values for Simulated Coefficients

	Δ	Observed $\overline{\tilde{\rho}(1-k)}$	Unobserved $\overline{\tilde{\rho}(1-k)}$	Observed $\overline{-q}$	Unobserved . $\overline{-q}$
$\Delta\mu$	$+0.2$	0.30	0.30	0	0
	-0.2	0.30	0.30	0	0
$\Delta\rho$	$+0.08$	0.31	0.30	0	0.04
	-0.08	0.29	0.30	0	-0.04
ΔU^-	$+0.10$	0.30	-	0	-
	-0.10	0.30	-	0	-
$\Delta\sigma_w^2$	$+0.35$	0.26	0.30	0	0
	-0.20	0.34	0.30	0	0
$\Delta\sigma_v^2$	$+0.25$	0.34	0.30	0	0
	-0.16	0.25	0.30	0	0

Notes: This figure shows the calibration of the changes in parameters from Section 1.6.1 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 1.6 and 1.7 show the average simulated coefficients for each scenario. See Section 1.6.1 in the text for more details.

Table 1.6: Coefficients on Individual and Aggregate Regressions with Changing Parameters

		Observed		Unobserved	
		$FE_{t,t-1\|t-1}(i)$	$\overline{FE}_{t,t-1\|t-1}$	$FE_{t,t-1\|t-1}(i)$	$\overline{FE}_{t,t-1\|t-1}$
$\Delta\mu$	+	0.28	0.28	0.30	0.30
	-	0.28	0.28	0.30	0.31
$\Delta\rho$	+	0.29	0.29	**0.58**	**0.61**
	-	0.27	0.27	0.27	0.27
ΔU^-	+	0.27	0.27	-	-
	-	0.30	0.30	-	-
$\Delta\sigma_w^2$	+	0.24	0.24	0.28	0.28
	-	0.33	0.33	0.29	0.28
$\Delta\sigma_v^2$	+	0.34	0.33	0.29	0.28
	-	0.23	0.24	0.28	0.28

Notes: This table presents the estimates from the individual and aggregate regressions of forecast errors on their own lags for 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 mean coefficients on lagged forecast errors from each approach. For each variable, σ_w^2, σ_v^2, U^-, 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. I do not perform a simulation where forecasters do not observe the change in U^- as it does not make sense to think about the case where forecasters cannot see their own subjective uncertainty. The calibration of the changes as well as the true values of the parameters appear in Table 1.5. For more details, see Section 1.6.1.

Table 1.7: Coefficients from Pooled Regression Controlling for Misperception

		Observed		Unobserved	
		$FE_{t,t-1\|t-1}(i)$	$\pi_{t,t-1}$	$FE_{t,t-1\|t-1}(i)$	$\pi_{t,t-1}$
$\Delta\mu$	+	0.28	-0.01	0.29	0.02
	-	0.28	-0.02	0.29	0.01
$\Delta\rho$	+	0.29	0.00	**0.35**	**0.10**
	-	0.27	0	0.27	0.00
ΔU^{-}	+	0.27	-0.02	-	-
	-	0.30	-0.02	-	-
$\Delta\sigma_w^2$	+	0.24	-0.02	0.28	-0.02
	-	0.33	-0.02	0.28	-0.02
$\Delta\sigma_v^2$	+	0.34	-0.02	0.29	-0.02
	-	0.23	-0.02	0.28	-0.02

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, U^{-}, 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. I do not perform a simulation where forecasters do not observe the change in U^{-} as it does not make sense to think about the case where forecasters cannot see their own subjective uncertainty. The calibration of the changes as well as the true values of the parameters appear in Table 1.5. For more details, see Section 1.6.1.

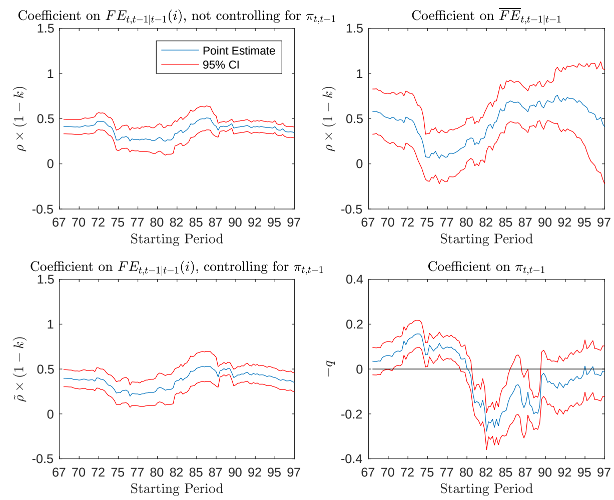

Figure 1.1: Time Variation in Information Rigidity and Parameter Misperception

Notes: This figure shows results for rolling window regressions on Equations 1.7, 1.8 and 1.12. The top two plots show the coefficients from the individual and aggregate approaches described in Section 3.2. 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 1.7.1 for details.

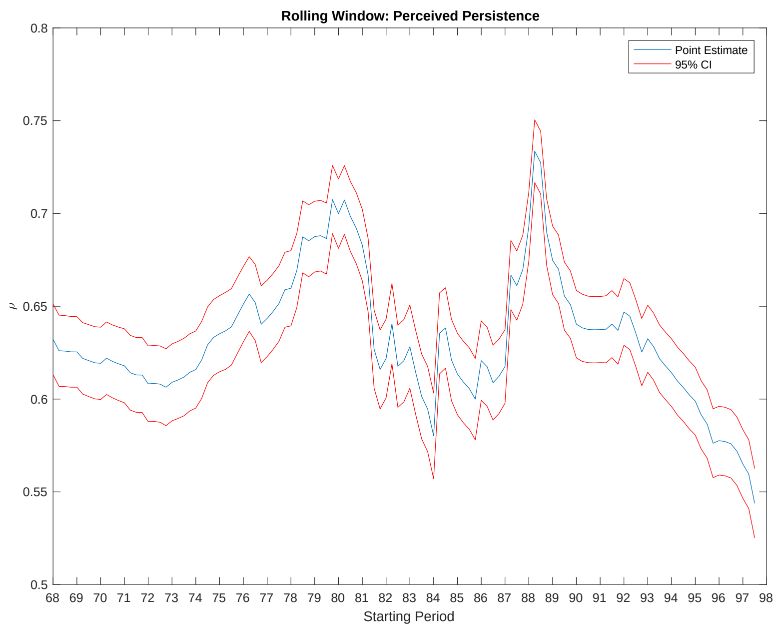

Figure 1.2: Time Variation in Perceived Persistence

Notes: This figure shows an 80-window rolling regression of Equation 1.14, 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. See Sections 1.4.3 and 1.7.1 for more details.

Chapter 2

Do You Know that I Know that You Know...? Higher Order Beliefs in Survey Data

2.1 Introduction

Firms anticipate the future and make pricing and employment decisions that have direct bearing on economic dynamics. For this reason, a firm's macroeconomic expectations are of particular interest to central banks and policymakers. When firms have strategic incentives to price their products similarly to their competitors, they must anticipate the expectations of other managers. Such expectations of the expectations of others are known as higher order expectations and play a key role alongside strategic complementarities in generating amplification and propagation in macroeconomic models.

Higher order expectations have implications for models above and beyond the implications of an agent's own expectation. Angeletos and La'O (2009) highlight the importance of considering higher order beliefs separately from an agents own beliefs. They argue that in a noisy or imperfect information context, the precision of information does not predict higher order beliefs the way it does own expectations. Bacchetta and Wincoop (2008) shows that the difference between higher order and own expectations is important for determining the pricing volatility of assets as well as the link between asset pricing and expectations of future asset payoffs. Therefore, we need data assessing both types of expectation directly in order to properly calibrate macroeconomic models and consider the implications of higher order expectations on economic dynamics. The current paper addresses this need.

As higher order expectations require that agents solve a problem where they iteratively consider the best responses of other managers to their own actions, this is a natural place to think about level-k thinking, or limitations in the strategic reasoning of agents. Most macroeconomic models consider the case in which economic agents perform infinite iterations to solve the problem. Such high level reasoning is, however, computationally expensive and difficult. Experiments (Nagel and Duffy (1997), Nagel (1995), Camerer et al. (2004), Stahl and Wilson (1995), Costa-Gomes and Crawford (2006)) suggest that agents often do not perform infinite iterations of the problem, or "degrees of reasoning," but often stop close to 2 or 3 iterations of a problem. Models of level-k thinking allow managers to make decisions with limited strategy. A level-0 player will respond to a game non-strategically, possibly even without regard to the rules of the game. A level-1 player will follow the rules of the game, but strategize as if other players are all level-0 players. A level-2 player will respond as if all other players are level-1, etc. Such models allow for imperfect reasoning and strategy on the part of agents and are also gaining popularity in macroeconomics. Fahri and Werning (2017) show that level-k thinking coupled with heterogeneity in market frictions creates mitigates the effect of monetary policy, addressing the forward guidance puzzle. Garcia-Schmidt and Woodford (2015) use level-k reasoning to show that monetary policy commitments to keep the nominal interest rate very low need not be deflationary. As level-k thinking can dramatically alter the predictions of a higher order expectation model, we consider these two features of expectations jointly and provide survey evidence on their interplay.

The first contribution of this paper is to introduce a novel set of questions on higher order expectations to a survey of firm managers. Strategic complementarities in pricing make it necessary for firm managers to anticipate the beliefs and actions of other managers, giving rise to higher order expectations. We introduce new questions to an existing survey of firm managers in New Zealand[1] that allow us to characterize the features of these expectations and compare them to the features of a managers' own expectations. We compare our data from these questions to the predictions of models of higher order expectations where agents form beliefs about the beliefs of other agents, as in Morris and Shin (2002), and find some facts inconsistent with managers reasoning to such a high degree.

Accordingly, our second contribution is to consider empirically whether managers engage such high orders of reasoning when forming these higher order expectations. Models of level-k thinking such as Nagel (1995) and Camerer et al. (2004) limit an agent's ability to anticipate the general equilibrium effect

[1] See Coibion et al. (2018) and Afrouzi et al. (2015).

of his own actions or expectations. In other words, the agent does not fully consider the response of other agents to his own actions, even if other agents would respond in the same way that he does. This limitation could arise from a cognitive constraint or the costs associated with higher level computation. We introduce questions to the survey that allow us to characterize the thinking type of each manager and compare the properties of level-k thinking to existing models of level-k thought and cognitive hierarchy. We can also assess whether thinking types predict differences in managers' higher order expectations. We find that managers do not exhibit behavior completely consistent with current models of level-k thinking. We also do not find any differences in higher order expectations that can be explained by differences in reasoning type.

Lastly, we assess the effects of information by conducting an experiment gauging the response of manager expectations to signals. Noisy information models such as Woodford (2002) and Sims (2003) argue that agents will partially update their beliefs given new information or signals. The degree to which they update indicates how much informative content they believe the signals contain. Our experiment introduces different types of signals about inflation itself and different orders of expectations about inflation. We find that managers respond strongly to information about past inflation and about the *higher order expectations* of other managers, but more weakly to the inflation expectations of other managers. This holds regardless of the thinking type of the manager.

The remainder of the paper is organized as follows. Section 2.2 describes a model of higher order expectations under strategic complementarities in price and compares the expectations of managers in New Zealand to the predictions of this model. Section 2.3 describes our results on level-k thinking and Section 2.4 describes the results of the experiment. Section 2.5 concludes.

2.2 A Model of Higher Order Expectations

Strategic complementarities in pricing behavior require that firms think not only of their own expectations of a fundamental, but also of other firms' expectations and actions. Firm A must think about the fundamental and what Firm B thinks of the fundamental. Firm B then anticipates the fundamental, what firm A thinks of the fundamental, and what Firm A thinks that Firm B thinks. Firm A's expectations must respond accordingly and this continues. As firms anticipate each others' actions, they form higher order beliefs that involve iterating a problem to progressively higher levels of reasoning. We use the static model of Morris and Shin (2002) to demonstrate how the expectations and higher order expectations of the firms in our survey compare to the predictions of a model of strategic complementarities where firms perform infinite regress in their expectations.

2.2.1 Strategic Complementarities in Pricing

A firm chooses to set its optimal price, p_i as a linear combination of its expectations of a fundamental, m[2], and its expectation of the average price

[2]We can think of the fundamental as monetary policy.

level in the economy, $\overline{p}$:

$$p_i = \alpha E_i[m] + (1-\alpha)E_i[\overline{p}] \tag{2.1}$$

In the above equation, $\alpha \in (0,1)$ describes the degree of complementarity in firm pricing. As all firms behave in the same way, manager i can iterate the optimal price equation forward by basing his beliefs about the aggregate price level on the aggregate pricing decision.

$$p_i = \alpha E_i[m] + (1-\alpha)E_i\left[\int p_j dj\right]$$

The aggregate price level becomes an average of progressively higher order expectations of m, weighted by the complementarities present at each step.

$$\overline{p} = \alpha\overline{E}[m] + \alpha(1-\alpha)\overline{E}^2[m] + \alpha^2(1-\alpha)\overline{E}^3[m] + \ldots \tag{2.2}$$

The optimal choice of p_i then depends on the manager's expectation of each event in Equation 2.2

$$p_i = \alpha E_i[\overline{E}[m]] + \alpha(1-\alpha)E_i[\overline{E}^2[m]] + \alpha^2(1-\alpha)E_i[\overline{E}^3[m]] + \ldots \tag{2.3}$$

2.2.2 Noisy Information

Firms operate under imperfect information. This means that, rather than observing m completely, they see noisy public and private signals that include the true value of m and some noise. Allow $m \sim N(y, \frac{1}{\kappa_y})$, where y is a public signal about the fundamental. Firms also receive a private signal about m, $x_i = m + v_i$, with $v_i \sim N(0, \frac{1}{\kappa_x})$. Firms weight their signals according to the relative noise in each into an individual expectation of m:

$$E_i[m] = \frac{\kappa_y}{\kappa} y + \frac{\kappa_x}{\kappa} x_i \tag{2.4}$$

where $\kappa = \kappa_x + \kappa_y$. Aggregating gives:

$$\overline{E}[m] = \frac{\kappa_y}{\kappa} y + \frac{\kappa_x}{\kappa} m$$

One can obtain progressively higher order expectations of m by continuing to substitute $E_i[m]$ for m. Higher orders of reasoning will depend more on the public signal as they rely more on average, rather than idiosyncratic, beliefs.

Referring to the firm's optimal price-setting equation, 2.3, we can substitute for the manager's expectations of m at various orders.

$$p_i = (1-\alpha) \sum_{k=0}^{\infty} \alpha^k \left[\left[1 - \left(\frac{\kappa_x}{\kappa} \right)^k \right] y + \left(\frac{\kappa_x}{\kappa} \right)^k x_i \right]$$

Given a that k goes to infinity, every agent sets the optimal price:

$$p_i = \phi_y y + \phi_x x_i \tag{2.5}$$

with $\phi_y = \frac{\kappa_y}{(1-\alpha)\kappa_x + \kappa_y}$ and $\phi_x = \frac{(1-\alpha)\kappa_x}{(1-\alpha)\kappa_x + \kappa_y}$. The realization of the aggregate price is the integral of Equation 2.5 across the support of all managers.

$$\overline{p} = \phi_y y + \phi_x m \tag{2.6}$$

2.2.3 Comparing Own Expectations and Higher Order Expectations

We compare the predictions of the above model of pricing and expectations to the moments of managers' expectations of inflation and higher order expectations of inflation in the New Zealand data. A summary of the moments for each type of expectation can be found in Table 2.1.

2.2.3.1 Means

Consistent with models of noisy information under strategic complementarities, the mean of the distribution of firms' own expectations of the aggregate price level is similar to that of the firms' higher order expectation of the aggregate price level - or their expectation of other managers' expectation.

$$E_i[\overline{p}] = \phi_y y + \phi_x (\frac{\kappa_y}{\kappa} y + \frac{\kappa_x}{\kappa} x_i) \tag{2.7}$$

Aggregating across agents gives the mean of agents' own expectations.

$$\overline{E}[\overline{p}] = \phi_y y + \phi_x(\frac{\kappa_y}{\kappa}y + \frac{\kappa_x}{\kappa}m) \tag{2.8}$$

The individual expectation of Equation 2.8 is an individual manager's higher order expectation:

$$E_i[\overline{E}[\overline{p}]] = \phi_y y + \phi_x \left[\left(1 - \left(\frac{\kappa_x}{\kappa}\right)^2\right) y + \left(\frac{\kappa_x}{\kappa}\right)^2 x_i\right] \tag{2.9}$$

Aggregating Equation 2.9 again gives the mean of the higher order expectation.

$$\overline{E}^2[\overline{p}] = \phi_y y + \phi_x \left[\left(1 - \left(\frac{\kappa_x}{\kappa}\right)^2\right) y + \left(\frac{\kappa_x}{\kappa}\right)^2 m\right] \tag{2.10}$$

Given $\frac{\kappa_x}{\kappa} < 1$, which is true as long as $\kappa_y > 0$, higher order expectations become slightly more weighted to the public signal. We do not expect, however, a substantial difference in managers' own expectations and higher order expectations. This holds in the data, with $\overline{E}[\overline{p}] = 3.41$ and $\overline{E}^2[\overline{p}] = 3.50$ for managers in our sample.

2.2.3.2 Variance

The noisy information model predicts that the cross sectional variance of higher order expectations will be smaller than the variance of the managers' own expectations. This happens as higher order expectations become more weighted toward the common signal, which is perfectly observed by all agents. The variance of own expectations of the aggregate price level is derived from the individual and aggregate own expectations, Equations 2.7 and 2.8.

$$\begin{aligned} V[E_i[\overline{p}]] &= E[(E_i[\overline{p}]] - \overline{E}[\overline{p}]])^2] \\ &= \phi_x^2 \left(\frac{\kappa_x}{\kappa}\right)^2 E[(x_i - m)^2] \\ &= \phi_x^2 \left(\frac{\kappa_x}{\kappa}\right)^2 \frac{1}{\kappa_x} \end{aligned} \tag{2.11}$$

The cross-sectional variance of higher order expectations derives from Equations 2.9 and 2.10:

$$\begin{aligned} V[E_i[\overline{E}[\overline{p}]] &= E[(E_i[\overline{E}[\overline{p}]] - \overline{E}^2[\overline{p}]])^2] \\ &= \phi_x^2 \left(\frac{\kappa_x}{\kappa}\right)^4 E[(x_i - m)^2] \\ &= \phi_x^2 \left(\frac{\kappa_x}{\kappa}\right)^4 \frac{1}{\kappa_x} \end{aligned} \tag{2.12}$$

As $\frac{\kappa_x}{\kappa} < 1$, the variance of the higher order expectation is lower than the variance of managers' own expectations. This feature is also present in our data, with the variance of higher order expectations at 9.36 and the variance of managers' own expectations at 5.90.

2.2.3.3 Uncertainty

As all distributional parameters are known to all managers, the subjective uncertainty about both the expectation of the price level and the higher order expectation of the price level should be the same as the cross sectional variances of these objects and given by Equations 2.11 and 2.12.. Therefore, we expect to see in the data that managers' subjective uncertainty about their own prediction of inflations should be higher than their uncertainty regarding the expectations of other managers. This holds true in the data with the uncertainty average uncertainty about own expectations and higher order expectations at 1.10 and 0.89, respectively. However, these numbers are much smaller than the respective cross sectional variances and therefore do not line up with the predictions of the model under infinite orders of reasoning.

2.2.3.4 Correlation

This model posits perfect correlation between the expectations and higher order expectations of managers.

$$Corr(E_i[\overline{p}], E_i[\overline{E}[\overline{p}]) = \frac{Cov(E_i[\overline{p}], E_i[\overline{E}[\overline{p}])}{SD(E_i[\overline{p}]) \times SD(E_i[\overline{E}[\overline{p}])}$$

As both own expectations and higher order expectations vary only with realizations of x_i and y, the correlation between the two will equal 1:

$$Corr(E_i[\overline{p}], E_i[\overline{E}[\overline{p}]) = \frac{\phi_x^3 \left(\frac{\kappa_x}{\kappa}\right)^3 E[(x_i - m)^2]}{(\phi_x \left(\frac{\kappa_x}{\kappa}\right) E[(x_i - m)]) \times (\phi_x^2 \left(\frac{\kappa_x}{\kappa}\right)^2 E[(x_i - m)])} = 1$$

In the data, we see a deviation from this prediction, with the correlation between the two types of expectations equal to 0.63.

2.2.3.5 Regression Coefficient

Figure 2.1 shows a scatter plot and regression line plotting higher order expectations against own expectations in the New Zealand data. The regression coefficient on managers' own expectations is $0.54^{***}(0.02)$. This finding is consistent with the predictions of the model so far. The regression coefficient is given by the following equation.

$$\beta = Corr(E_i[\overline{p}], E_i[\overline{E}[\overline{p}]) \frac{SD(E_i[\overline{E}[\overline{p}])}{SD(E_i[\overline{p}])}$$

Using the fact that the correlation between own and higher order expectations should equal 1 and the standard deviations derived from Equations 2.11 and 2.12, we find that the regression coefficient is equal to the following expression:

$$\begin{aligned}\beta &= \frac{\phi_x \left(\frac{\kappa_x}{\kappa}\right)^2 \frac{1}{\sqrt{\kappa_x}}}{\phi_x \left(\frac{\kappa_x}{\kappa}\right) \frac{1}{\sqrt{\kappa_x}}} \\ &= \frac{\kappa_x}{\kappa}\end{aligned}$$

Under the assumptions of the model, this is strictly less than one.

2.3 Level-k Thinking

Section 2.2 described a series of predictions that hold if agents undertake infinite degrees of reasoning about the pricing decisions of others. Reasoning of this sort is, however, difficult and computationally intensive. Managers are therefore likely, due to either cognitive constraints or recognizing the costs of such reasoning, limit their degrees of thinking to levels well below infinity.[3]. We introduce questions into our survey that allow us to categorize the thinking types of firm managers in New Zealand and compare the properties of level-k behavior in our survey with the predictions of existing models.

2.3.1 Categorizing Types

To characterize firms' degree of reasoning, we ask the following question. We also time the firm managers as they answer this question to provide another measure of the amount of thinking that managers do.

"Please choose a number from zero to 100. We will take your number as well as the numbers chosen by other managers to calculate the average pick. The winning number will be the number that is closest to two-thirds (2/3) of the average.

The individual(s) with the winning number will receive (or share with other winners in case of tie) \$500."

As in Nagel (1995) and Nagel and Duffy (1997), we can then define the degree of reasoning by the manager's answer to this question. A kth-level thinker provides the following guess:

[3]Experimental literature on this topic suggests that experiment participants rarely get beyond reasoning at level 3 (Nagel and Duffy (1997), Nagel (1995), Camerer et al. (2004), Stahl and Wilson (1995), Costa-Gomes and Crawford (2006)).

$$g(k) = \left(\frac{2}{3}\right)^k \times 50 \tag{2.13}$$

The distribution of guesses appears in Figure 2.2. In the full sample, firm managers provide guesses throughout the full interval. However, when we restrict the sample to those managers who spend at least 20 seconds on the question, the guesses pile on integers associated with reasoning types between $k = 1$ and $k = 5$, with the number of firms of each type declining with k. Accordingly, we classify these managers by their guess and assign $k = 0$ to those who do less than 20 seconds of thinking. The guesses associated with $k = 0$ therefore fall throughout the interval of allowable guesses, rather than at 50.[4]

Table 2.2 shows the breakdown of types in our survey and in two papers that use experiments to identify agents' depths of reasoning. Nagel and Duffy (1997) runs experiments with small groups of students to see what depths of reasoning appear in the first round of the beauty contest game. This model and definition of k-types requires that a level-k thinker believes that everyone else in the game performs at level-$k-1$. Camerer et al. (2004) develops a model of 'cognitive hierarchy' that allows agents to form beliefs about the distribution of other reasoning types in the sample. A level-k thinker is assumed to observe the correct frequency of thinkers at his type and types below, but to incorrectly

[4]In the Nagel (1995) model, $g(k = 0) = 50$ while guesses throughout the interval are associated with $k < 0$.

assume that there are no thinkers at types above his own. As a result, he posits inaccurate relative frequencies of thinkers. As a thinker's reasoning type, k, increases and he observes the true frequencies of a greater number of types, his expectation of the density over the sample becomes "increasingly rational" and closer to the true distribution of types. The thinking types of managers in our survey appear more dispersed than in other surveys. We see a greater density of thinkers at $k = 0$, partially due to the way we assign this rating (as anyone who does less than 20 seconds of thinking about the question). In our survey, 36.8 percent of managers are $k = 0$, as opposed to 20 to 27.3 percent in the other papers. We also see a higher densities of managers at higher levels, $k = 3$ and $k \geq 4$, with 12.9 and 13.5 percent of managers at these levels. In other surveys we see 3 to 4 percent of people at these levels.

2.3.2 Higher Order Beliefs

We assess the beliefs of managers about the distribution of other managers' types by asking them to provide a probability distribution over ranges of other managers' guesses. Specifically, we ask:

"Other managers are asked to guess a number from zero to 100, with the goal of making their guess as close as possible to two-thirds of the average guess of all those participating in the contest. *What percentage of other managers' guesses do you think will fall in each of the following ranges?*[5]

[5]The ranges include $0 - 9.99$, $10 - 19.99$, $20 - 20.99$ and so on through $90 - 100$.

The results of the distribution question are not fully consistent with either the Nagel (1995) or Camerer et al. (2004) models. Table 2.3 reports results on these beliefs by reasoning type. Roughly 80 percent of managers of all reasoning levels assign positive probability to multiple bins a fact that is not consistent with the Nagel model of reasoning. A level-k thinker as defined by the guess in Equation 2.13 should report positive probability on only one bin, the one associated with the level-$k-1$ guess. This level-k thinker will also not place positive probability on the bin associated with her own guess. For types $k = 1, 2$, and 3, managers place an average probability of 0.72 to 0.77 on this bin, meaning they think that between 72 and 77 percent of other managers are the same type as them. All levels, $k \geq 1$, assign positive probability to bins associated with thinkers beneath their own level, consistent with both Nagel (1995) and Camerer et al. (2004). However, only types $k \geq 4$ report believing that a majority of managers will fall into bins associated with lower level thinkers. Thinkers at types $k \leq 3$ also report positive probability on bins associated with levels of k *above* their own. This result proposes a puzzle not explained by current depths of reasoning models. In such models, a level-k thinker cannot fathom the existence of a level-$k+1$ thinker, as to do so would be to engage in a higher order of reasoning himself.

Figure 2.3 shows the average believed distribution of guesses for each thinking type, k, as well as the true density function across guesses.[6] This

[6] Using Equation 2.13, we can interpret a manager's beliefs about the guesses of other managers as his beliefs about their types. For example, a level-1 thinker places a guess

illustrates that all types believe that the majority of managers share their own type and all managers at levels $k \leq 3$ assign probability to guesses associated with players at levels both higher and lower than their own. All managers underestimate the true dispersion of guesses. Accordingly, none of the thinking types correctly observe the true density of types, nor do beliefs about the density become closer to the truth with increasing k, as in Camerer et al. (2004).

Managers' reported beliefs about the guesses of other managers mean that their own guesses are not consistent with the rules of game. Managers believe the average guess, $E_i[Guess_{HO}]$ to be close to their own. To win the prize, the manager should submit a guess of two-thirds of his believed average guess. It may be the case that when asked directly about higher order expectations, managers will engage an additional level of reasoning that was not present when they formed their own guess or expectation. In this case, part of being a lower level thinker is failing to realize when expectations and higher order expectations are inconsistent with each other.[7]. If this is true, a thorough model of level-k thinking will allow thinkers to perform at different levels when considering their own expectations or actions and when considering the expectations or actions of others.

To test the consistency of the agents' guesses in the beauty contest

---

in the bin $20 - 20.99$, so placing positive probability in this bin indicates that a manager believes some percentage of other managers are type $k = 1$.

[7] This is a reasonable proposition, as the concept of level-k thinking itself introduces failures or limitations in reasoning to the agent's problem.

game with their beliefs about the average guess of other managers, we can check the following relationship:

$$Guess_{Own}(i) = c + \beta E_i[Guess_{HO}] + \epsilon_t \tag{2.14}$$

The constant term, c, should be equal to 0 and $\beta = 0.67$. Estimating this equation for managers who spent more than 20 seconds on the guessing game question gives the following result:

$$Guess_{Own} = 0.15 + 0.91^{***} E_i[Guess_{HO}] + \epsilon_t \tag{2.15}$$

$$(0.26)\,(0.01) \tag{2.16}$$

where robust standard errors are presented in parentheses. We can reject the null that $\beta = 0.67$. If we restrict the sample to higher levels of k ($k > 2$) we find

$$Guess_{Own} = 5.04 + 0.50^{***} E_i[Guess_{HO}] + \epsilon_t \tag{2.17}$$

$$(1.10)\,(0.08) \tag{2.18}$$

Here, we can no longer reject $H_0 : \beta = 0.67$, but we find a constant term significantly different from 0. These results indicate that thinkers of all types guess too high relative to their reported beliefs of the guesses of other managers. Additionally, all managers at levels lower than $k = 3$ - or 58 percent of those that spent at least 20 seconds on the beauty contest question - show behavior consistent with different levels of reasoning for guess and their higher order expectation of the guesses of others.

2.4 Experiment

In a noisy information environment, firm managers will update their expectations of inflation as well as of other managers' expectations when they receive new information. However, they will update their expectation only partially, reflecting that they may not believe the information to be fully credible or noiseless. To assess empirically the degree of noise that managers perceive in signals, we introduce signals in an experimental context and gauge the response of managers' expectations to these signals.

We are interested to see how managers respond, not only to information, but to *different kinds of information*. Following the initial survey where we ask about inflation expectations and higher order expectations, we perform the following experiment. We divide managers into five groups. Group A is a control group and does not receive any information. Group B receives information about the average beliefs of survey participants about inflation: $\overline{E}[\pi]$. Group C receives information about the average *higher order* inflation expectations of survey participants: $\overline{E}^2[\pi]$. Group D's signal consists of both information about average expectations and average higher order expectations. We utilize Group E to compare the impact of information about other managers' beliefs to information about the target variable. Managers in Group E receive a signal about lagged inflation.

To assess the impact of the signal on agent believes, we run the following regression for each experimental group:

$$Posterior_{i,Group} = \alpha_{Group} + \beta_{Group} Prior_{i,Group} \tag{2.19}$$

The coefficient on the prior expectation has a different interpretation based on its value. If $\hat{\beta}_{Group} \approx 1$, managers see the signal as uninformative and do not update their prior beliefs at all. If $0 < \hat{\beta}_{Group} < 1$, the signal is partially informative and managers will update their posterior somewhat, but will still rely partially on the prior. A $\hat{\beta}_{Group} = 0$ indicates a completely informative signal that causes managers to discard their priors in favor of the signal.

Table 2.4 shows the coefficient on the prior expectation for the regression of posterior expectations on prior expectations for own inflation expectations and higher order inflation expectations. Each experimental group has its own coefficient for each variable. We find that all managers update their expectations slightly, as evidenced by $\beta_A < 1$ for the control group. However, we see that agents perceive information about past realization and the average *higher order* beliefs of other managers more informative than information about the other managers' average expectation of inflation. The coefficient on the prior inflation expectation is 0.090 and 0.096 for the groups receiving information about the average higher order expectation. It is 0.059 for the group receiving information about the past realization. This contrasts with a coefficient of 0.502 for the group receiving information about the average inflation expectation. We see a similar discrepancy for the coefficients on the prior higher order expectation. For the groups receiving information about the average higher order expectation in the survey, this coefficient is 0.118 and 0.071. For those receiving information about the past realization of inflation, it is 0.062, and for those receiving information about the average own expectation, it is 0.430. These results imply that managers view signals about higher levels of thought as more informative than signals about lower levels.[8]

We also ran these regressions separately for different levels of thinking, but did not find any significant differences in the way managers at different reasoning types processed signals.

2.5 Concluding Remarks

This paper presents novel survey evidence on higher order expectations as well as level-k thinking. We find evidence broadly in line with noisy information models with strategic complementarities in pricing that require firm managers to form higher order expectations. We also find, however, evidence that managers may not reason to an infinite degree. We back this up further by characterizing managers by their levels of thinking. We further show that levels of thinking conform to aspects of different models of level-k thinking, but

[8]It is worth further investigation to see if this fact is consistent with the finding in Section 2.3 that a manager's guess in the beauty contest game is not necessarily consistent with his beliefs about the guesses of others.

do not line up with others. This means there is work to be done in correctly modeling and defining the behavior of a level-k thinker.

This paper also measures the treatment effects of information on firm expectations. We find differential responses in responses to different types of information. According to our analysis, depths of reasoning do not significantly impact manager responses to treatments.

Jointly, these results contribute to a broader research agenda explaining the expectations formation of agents. Central banks may find this work particularly interesting as our results challenge certain model-based predictions of how expectations are formed and how decision-makers reason through information problems.

Future work[9] will focus on how a firm's higher order expectations and reasoning type effect its actions and its long term responses to information.

[9]Pending a follow-up wave of the survey

Table 2.1: Summary Statistics

	Own Expectation	**HO Expectation**
Mean	3.41	3.50
Standard Deviation	3.06	2.43
Uncertainty	1.10	0.89

Notes: This table gives summary statistics for managers' own expectations and higher order expectations of inflation, where the higher order expectation is defined as the expectation of the average of *other managers'* expectation. Consistent with the predictions of a model with strategic interaction in pricing, these expectations have similar means and the standard deviation and uncertainty of higher order expectations is less than those of own expectations. There is, however, a substantial difference in the standard deviation and uncertainty terms for each variable. This is inconsistent with the model of strategic interaction in which agents infinitely reason the behavior of their competitors. See Section 2.2 for more information.

Table 2.2: Breakdown of Reasoning Types

Level of Thinking, k	**NZ Managers**	**Duffy and Nagel 1997**	**Camerer et al 2004**
$k = 0$	36.8	~ 20.0	~ 27.3
$k = 1$	21.2	~ 50.0	~ 35.4
$k = 2$	15.6	~ 22.0	~ 23.0
$k = 3$	12.9	~ 4.5	~ 10.0
$k = 4+$	13.5	~ 3.5	~ 3.2

Notes: This table shows the distribution of level-k types from our survey, from the mean game of Nagel and Duffy (1997) and from the cognitive hierarchy characterization of Camerer et al. (2004). Our results are broadly similar to other experimental approaches for characterizing depths of reasoning. We see a slightly higher density of participants at $k = 0$. This results from our definition of a $k = 0$ manager as someone who spends less than 20 seconds on the beauty contest question. See Section 2.3 for more information.

Table 2.3: Beliefs about Other Managers' Guesses

Level of Thinking, k	Fraction Reporting: Positive Probability on More than One Bin	Average Probability on Bin: of Own Guess	Average Probability on Bin: of Lower k than Guess	Average Probability on Bin: of Higher k than Guess
$k = 0$	0.87	-	-	-
$k = 1$	0.82	0.72	0.14	0.14
$k = 2$	0.78	0.77	0.11	0.13
$k = 3$	0.82	0.74	0.16	0.10
$k = 4+$	0.83	0.36	0.64	-

Notes: This table describes managers' beliefs about other managers' guesses in the beauty contest game. The first column gives the fraction of each thinking type reporting that they believe other managers provide guesses in more than one bin. The remaining columns show the average probability that agents of each thinking type place on bins associated with levels of k above, below, and the same as their own. For more information on how these results compare with existing models of level-k thinking, see Section 2.3.

Table 2.4: Responses to Experiment

Experiment Group	Treated With:	Own Expectations	HO Expectations
A	-	0.727 (0.020)	0.700 (0.021)
B	$\overline{E}[\pi]$	0.502 (0.041)	0.430 (0.039)
C	$\overline{E}^2[\pi]$	0.090 (0.018)	0.118 (0.024)
D	$\overline{E}[\pi]$ & $\overline{E}^2[\pi]$	0.096 (0.022)	0.071 (0.020)
E	π_{t-1}	0.059 (0.015)	0.062 (0.021)

Notes: This table shows the results of the experiment. We provided each group with the signal described in the second column and asked for a posterior expectation. Group A did not receive a signal as the control group. The remaining columns of the table present the coefficient on the prior expectation from a regression of the posterior on the prior. A coefficient close to one indicates that managers view the signal as uninformative and stick closely to their priors. As the coefficient values get closer to zero, it indicates that managers find the signal more informative and more greatly update their priors. Here, the most informative signals were information about the lagged value of inflation itself and information about the average higher order expectation. Information on the average expectation of inflation had a lesser impact on posterior expectations. See Section 2.4 for more details.

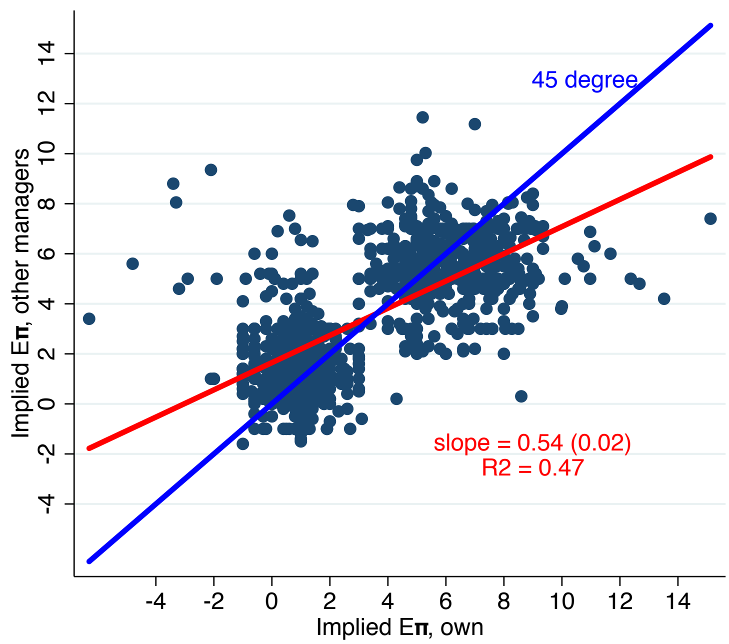

Figure 2.1: Own Expectations and Higher Order Expectations

Notes: This figure shows managers' higher order expectations plotted against their own expectations, along with the regression line relating these two. The See 2.2 for more information.

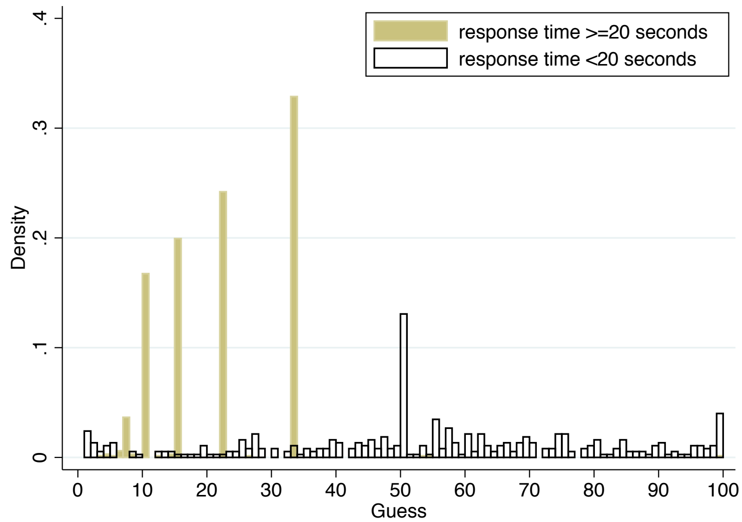

Figure 2.2: Distribution of Reasoning Types

Notes: This figure shows the distribution of guesses from the beauty contest game. We asked managers to provide a guess between zero and 100 with the guess closest to $\frac{2}{3}$ of the average guess receiving a prize. For managers who spent at least 20 seconds in considering their guess, we see clumping of guesses at those points which correspond neatly with level-k types as defined in Nagel (1995). Those managers who answered the question in less than 20 seconds made guesses dispersed across the full interval. See Section 2.3 for more information.

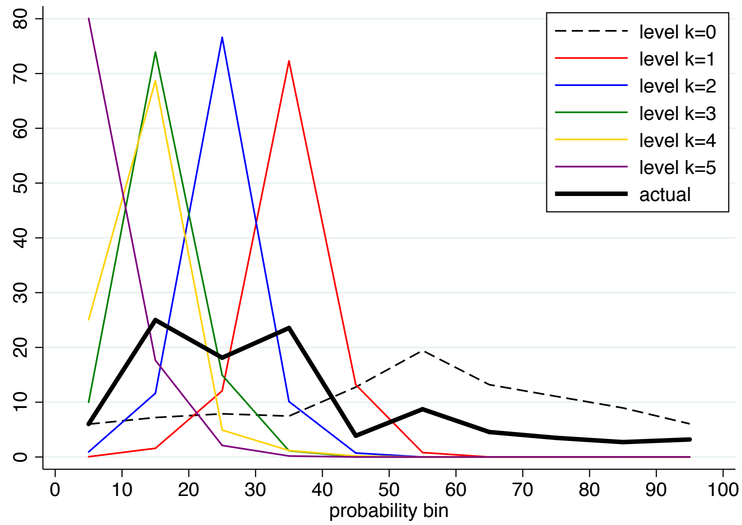

Figure 2.3: Beliefs About Distribution of Other Managers' Guesses

Notes: This figure shows the average beliefs about the distribution of other managers' guesses in the beauty contest game by reasoning type. Individuals of each type believe that the majority of other managers provide guesses similar to their own. All thinking types also underestimate the true dispersion of the guesses of other managers. See Section 2.3 for more information.

Chapter 3

Time Varying Attention: Evidence from Survey Data

3.1 Introduction

Central banks and policymakers care about the economic expectations of agents in the economy as the actions of these agents directly affect policy objectives. For this reason, there is a growing interest in how economic agents approach the problem of collecting information and using it to form beliefs. Amid this growing interest is an interest on rational inattention models, where rather than facing exogenous constraints on information, agents optimally choose some level of attention to assign to a target variable subject to a capacity constraint.

Rationally inattentive agents choose the degree of attention they pay to a variable in response to economic dynamics. Therefore, this degree of attention is likely be subject to changes given different monetary policy regimes and underlying economic conditions. This paper proposes a framework for estimating attention that tracks time variation. This allows us to view changes in the expectations formation process in light of changes to macroeconomic processes over time. I find that forecasters' degree of inattention has indeed changed over time, declining for most variables since the late 1960s. I further argue that this change is consistent with a rationally inattentive response to underlying economic conditions.

Interestingly, in periods associated with higher attention, we also see a greater dispersion in attention across agents. This implies that underlying macroeconomic conditions that generate a greater attention response are also more likely to generate a heterogeneous response in attention across agents. This is a new stylized fact that may prove useful in modeling rational inattention in future work.

This paper relates to a literature on rational inattention, using the seminal models of Sims (2003) and Mackowiak and Wiederholt (2009), to empirically assess the degree to which agents are inattentive. Afrouzi (2017) does something similar, calibrating a rational inattention model and backing out estimates of noisy information parameters. My approach begins with the empirical estimation of a noisy information model, making this paper similar to Coibion and Gorodnichenko (2015), Coibion and Gorodnichenko (2012), Dovern et al. (2014), Andrade and Le Bihan (2013), and Chapter 1 of this dissertation.

The paper is organized as follows. Sections 3.2 and 3.3 present the model and estimation strategy, respectively. Section 3.4 presents the data. I discuss the results in Section 3.5. Section 3.6 provides a brief discussion and Section 3.7 concludes.

3.2 Model of Inattention

I begin with a simple Kalman filter model in which agents predict a fundamental, x_t, according to a linear combination of the their past beliefs about the fundamental and a private signal that they receive. While signal processing via methods like the Kalman filter commonly appear in noisy information models, these same methods can be used in a rational inattention framework. In noisy information models, agents receive exogenously imperfect signals. As their information is imperfect, they cannot fully trust their signals and place a portion of the weight of their new expectation on prior beliefs. On the other hand, rationally inattentive agents face a capacity constraint on processing information and therefore view attention to any particular variable as costly. The precision of agent signals in rational inattention models is, therefore, endogenously generated and left up to agent choice.

Mackowiak et al. (2017) show that, given a fundamental that evolves according to an AR(1) process, the optimal signal chosen by a rationally inattentive agent has the same structure as the signals from a basic noisy information model. Accordingly, I allow the fundamental to follow an AR(1):

$$x_t = \mu + \rho x_{t-1} + w_t. \tag{3.1}$$

Agents optimally select signals of the form:

$$z_t(i) = x_t + v_t(i) \tag{3.2}$$

where $w_t \sim N(0, \sigma_w^2)$ and $v_{it} \sim N(0, \sigma_v^2(i))$. Each forecaster chooses their signal noise variance, $\sigma_v^2(i)$, consistent with rational inattention models over noisy information models. The constant in the transition equation, μ, allows the fundamental to converge to a non-zero long run mean. The agent forms her time-t expectation of x_t according to the following equation:

$$\begin{aligned}
\tilde{x}_{t|t}(i) &= k_i(z_t) + (1-k_i)\tilde{x}_{t|t-1} \\
&= k_i(\pi_t + v_{it}) + (1-k_i)\tilde{x}_{t|t-1} \\
&= k_i\pi_t + (1-k_i)\tilde{x}_{t|t-1} + k_i v_{it}
\end{aligned} \tag{3.3}$$

The steady state Kalman gain, k_i has the following representation:

$$k_i = \frac{P_{t|t-1}(i)}{\sigma_{v,x}(i)^2 + P_{t|t-1}(i)}$$

The a priori covariance of the estimate, $P_{t|t-1}(i)$, represents the forecaster's perceived variance, or uncertainty, of the time-t state conditioned on signals received up to period t-1. As she enters time t and receives signal z_t(i), she updates this uncertainty estimate to $P_{t|t}(i) = (1-k_i)P_{t|t-1}(I)$. A forecaster can reduce her uncertainty about a particular variable by allocating attention to it. Models of rational inattention rely on placing constraints on information flow, or the reduction in uncertainty that results from attention; a forecaster can generate information flow up to a specified capacity constraint and will therefore optimally choose to limit attention.

Using entropy as a measure of uncertainty, as in Mackowiak and Wiederholt (2009), it is possible to define information flow as the difference between the entropy of a random variable, $\mathbf{x}$, prior to receiving the signal z_{it} and the conditional entropy of $\mathbf{x}$ given the signal $z_t(i)$. The entropy of $\mathbf{x}$ prior to receiving $z_t(i)$, given that x_t is normally distributed with conditional variance $\sigma^2_{x|z_{t-1}(i)}$ is:

$$H(x|z_{t-1}(i)) = \frac{1}{2} log_2(2xe\sigma^2_{x|z_{t-1}(i)}). \tag{3.4}$$

The conditional entropy given signal $z_t(i)$ is:

$$H(x|z_{it}) = \frac{1}{2} log_2(2xe\sigma_{x|z_t}(t)^2). \tag{3.5}$$

Given these two terms, the information flow for a univariate process is equal to the mutual information between the two, defined as:

$$I(x; z_t(i)) = H(x|z_{t-1}(i)) - H(x|z_t(i)). \tag{3.6}$$

As the a priori estimate covariance establishes the agent's uncertainty about the state before realizing her signal, let $\sigma^2_{x|z_{t-1}(i)}$ take the value of the a priori variance of the state estimate, $P_{t|t-1}(i)$. Following the observation of the signal, the agent's estimate covariance updates to $P_{t|t}(i)$. Therefore, we can consider $P_{t|t}(i)$ an estimate of the conditional variance $\sigma_{x|z_t(i)}$. Each individual's information flow now takes the form:

$$
\begin{aligned}
I(x; z_t(i)) &= \frac{1}{2} log_2(2xeP_{t|t-1}(i)) - \frac{1}{2} log_2(2xeP_{t|t}(i)) \\
&= \frac{1}{2} \left[log_2(P_{t|t-1}) - log_2(P_{t|t}) \right]
\end{aligned}
$$

Using the relationship between the two uncertainty estimates, $P_{t|t}(i) = (1 - k_i)P_{t|t-1}(i)$, it is possible to rewrite this as:

$$
\begin{aligned}
I(x; z_t(i)) &= \frac{1}{2} \left[log_2(P_{t|t-1}(i)) - log_2((1 - k_i)P_{t|t-1}(i)) \right] \\
&= \frac{1}{2} \left[log_2(P_{t|t-1}(i)) - log_2(1 - k_i) - log_2(P_{t|t-1}(i)) \right] \\
&= -\frac{1}{2} log_2((1 - k_i)).
\end{aligned}
\tag{3.7}
$$

As k_i is bounded between 0 and 1, this term is guaranteed to be non-negative. Information flow is also greater than zero as long as $k_i > 0$ or as long as agent signals contain some informative content.

3.3 Estimation Strategy

Changing underlying economic conditions over time will induce a different choice of signal noise variance, $\sigma_v^2(i)$. This choice leads to a different Kalman gain, k_i, and a different information flow, $I(x|z_t(i))$. We should accordingly expect information flow to differ across agents based on the time periods that they participate in the sample. Accordingly, we should be able to track subtle changes in attention by tracking the average attention across agents present in each time period.

From Equation 3.7, it is clear that one only needs an estimate of the individual-specific Kalman gain to obtain an estimate of the individual's information flow. I estimate k_i for the individuals in the sample by running a constrained regression of the agent's forecast of x_t on her lagged forecast of the same event and on the realization of x_t, which is the observable part of her signal.

$$\tilde{x}_{t|t} = \beta_0(i) + \beta_1(i)x_t + \beta_2(i)\tilde{x}_{t|t-1} + \epsilon_t \tag{3.8}$$

I impose the constraint $\beta_1 + \beta_2 = 1$ as $\beta_1 = \hat{k}_i$ and $\beta_2 = (1 - \hat{k}_i)$. The errors in the above equation have a structural interpretation as signal noise terms, scaled by each agent's Kalman gain. I expect that the agents' forecasting equation will not include a constant term and therefore want to check that the regression constants do not differ from 0.

The information flow for each individual is then calculated as a monotonic transformation of $\hat{k}_i$, $\hat{I}(x; z_t(i)) = -\frac{1}{2}log_2((1-\hat{k}_i))$. For each time period, I then take the average of information flow across agents and estimate the cross sectional dispersion in information flow to see how the agents' attention problem has changed over time.

$$\overline{I(x|z_t)} = \frac{\sum_i \hat{I}(x; z_t(i))1_t(i)}{\sum_i 1_t(i)}$$

While changes in the time-average of attention occur due to changes in the composition of sample, it remains the case that, over time, the sample may transition to higher or lower attention individuals. As rationally inattentive forecasters will respond to changing underlying economic conditions with adjustments to attention, observing forecasters present in different periods will give a sense of forecasters' response to changing regimes.

3.4 Data

The data for this estimation comes from the Survey of Professional Forecasters, a quarterly survey conducted by the Federal Reserve Bank of Philadelphia.[1]. Several design features of this survey make it desireable for this estimation. First, the survey began in 1968 tracking several macroeconomic variables. This allows for the examination of attention across several influential periods in monetary policy including the high inflation of the 1970s followed by the Volcker Disinflation and Great Moderation. Second, the survey consists of a highly unbalanced panel, meaning that not all forecasters are observed for all periods. This allows me to estimate the information flow for individuals who are present in different periods and therefore differentially exposed to different

[1]The American Statistical Organization and the National Bureau of Economic Research conducted the survey prior to 1990.

policy regimes and underlying conditions.

The survey includes several macroeconomic variables since its inception in 1968. These variables include the GDP price deflator, unemployment, real GDP, industrial production, and housing starts. I form each of these variables except for unemployment into projected growth rates. I use the quarterly nowcast of each variable as $\tilde{x}_{t|t}(i)$ and use the final release measure of each as variable to represent the realization or the observable part of the signal. Each forecaster's lagged expectation of time-t x is given by the quarter lag of the quarter-ahead forecast.

3.5 Results

Figure 3.1 shows the distribution of Kalman gains across individuals for each of the five macroeconomic variables available since the survey's inception. Table 3.1 gives summary statistics about the individual Kalman gains for each variable. There is heterogeneity across variables in the average Kalman gain, though the average gain is similar for GDP price inflation, industrial production, and housing starts. Variables with relatively higher and lower degrees of attention also appear to have less dispersion in individual Kalman gains.

Figures 3.2 and 3.3 show the average information flow and the cross sectional standard deviation of information flow for each time period for price inflation, real GDP, industrial production, and housing starts. These graphs appear in a separate figure, 3.4, for unemployment.

- **GDP Price Inflation**. The average Kalman gain for inflation is 0.41, meaning that the average forecaster forms her nowcast giving 41 percent weight to the signal and 59 percent to her lagged expectation. We also see that attention to inflation, or information flow, has decreased over time. Dispersion has decreased with attention.

- **Unemployment** Interpreting the individual Kalman gain as the amount of weight the agent gives to her new signal, we see that agents pay the most amount of attention to new signals about unemployment, with 73 percent of their nowcast coming from the signal and only 27 percent on their past expectation. Figure 3.4 shows the time variation in information flow and the standard deviation in information flow for unemployment. I present these results differently as forecasters show low-frequency variation in attention to unemployment rather than declining attention over time. We can see, however, that aggregate attention and dispersion in attention covary for unemployment.

- **Real GDP**: While agents weight signals about unemployment relatively highly, they pay relatively little attention to signals about real GDP. Real GDP has the lowest average Kalman gain at 0.33. This implies that forecasters give more weight to their past expectations than to their signals when forming a nowcast about GDP. Both average information flow and the dispersion in information flow has decreased over time for real GDP.

- **Industrial Production** As measured by the Kalman gain, forecasters pay a moderate amount of attention to industrial production, with the average Kalman gain at 0.43. This implies the nowcast gives 43 percent weight to the new signal. Attention and dispersion in attention to industrial production have declined over time.

- **Housing Starts**. Similar to industrial production, forecasters pay a moderate amount of attention to housing starts with an average Kalman gain at 0.43. We see declining attention and dispersion in attention over time for housing starts.

Figure 3.5 shows the standard deviation of information flow plotted against the aggregate information flow. This pools across periods and variables. As the time trend in unemployment appears to be different than for the other variables, I show the relationship between information flow and dispersion separately for unemployment. The relationship between aggregate information flow and the cross sectional standard deviation is positive for unemployment and for the pooled remaining variables, though the magnitude of the relationship is different. Running the following regression:

$$Std.\,Dev._{Variable,t} = c + \beta Aggregate\ Attention_{Variable,t} + \epsilon_{Variable,t} \tag{3.9}$$

$\hat{\beta} = 0.09^{**}(0.04)$ for unemployment and $\hat{\beta} = 0.25^{***}(0.01)$for the remaining variables. ** and *** indicate the 5 and 1 percent significance levels, respectively.

3.6 Discussion

These results suggest that a forecaster's attention problem has changed over time. A full rational inattention model would include a capacity constraint on information flow. Seeing a reduction in information flow to so many variables suggests that either forecasters' capacity has declined, or their payoff from attention has declined to the point that they allocate their capacity elsewhere.

A reduction in attention to variables due to macroeconomic conditions is a possibility for the period over which we see the reduction. Forecasters pay more attention during the volatile period of the 1970s and show a reduction in attention moving into the Great Moderation as well as the more macroeconomically stable 1990s. A rationally inattentive agent is less likely to pay attention to a stable process as there will not be many unpredictable changes.

3.7 Concluding Remarks

This paper provides an estimation approach and initial results regarding time-variation in information flow in the survey of professional forecasters. I find that forecasters have reduced attention to inflation, real GDP, industrial production, and housing starts while holding attention to unemployment relatively constant. I also find that cross sectional dispersion in attention increases with the average degree of attention. Future work may focus on sources of this link between attention and increased heterogeneity.

Table 3.1: Individual Kalman Gains

	GDP Inflation	Unemployment	Real GDP	Industrial Production	Housing Starts
Mean	0.46	0.74	0.36	0.48	0.48
Median	0.41	0.74	0.32	0.45	0.42
Standard Deviation	0.26	0.16	0.23	0.25	0.24
Interquartile Range	0.31	0.19	0.24	0.34	0.25

Notes: This table provides descriptive statistics of the distributions of individual Kalman gains for each of the following variables: GDP deflator inflation, unemployment, real GDP, industrial production, and housing starts. Each variable is approximately normally distributed with a median value close to the mean. There is variation in the average Kalman gain across variables, with forecasters showing a high degree of attention to unemployment and a relatively low degree of attention to real GDP. Variables with average Kalman gains closer to the center of the acceptable interval $(0, 1)$ also show greater dispersion across individuals.

Figure 3.1: Distribution of Individual Kalman Gains

Notes: This figure shows the distribution of individual Kalman gains for each variable available since the survey's initial release in 1968. See Section 3.5 and Table 3.1 for more information.

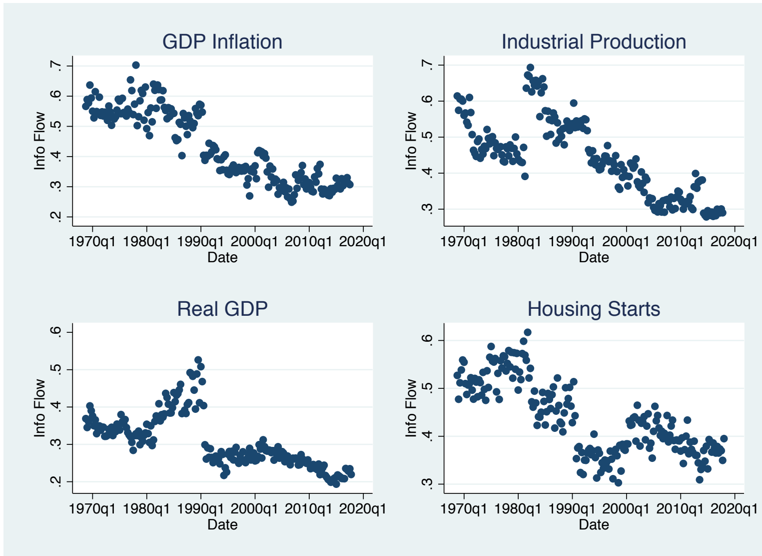

Figure 3.2: Time Variation in Information Flow

Notes: This figure shows the average information flow across individuals present in each time period. For each of the above variables, information flow appears to decline at later dates.

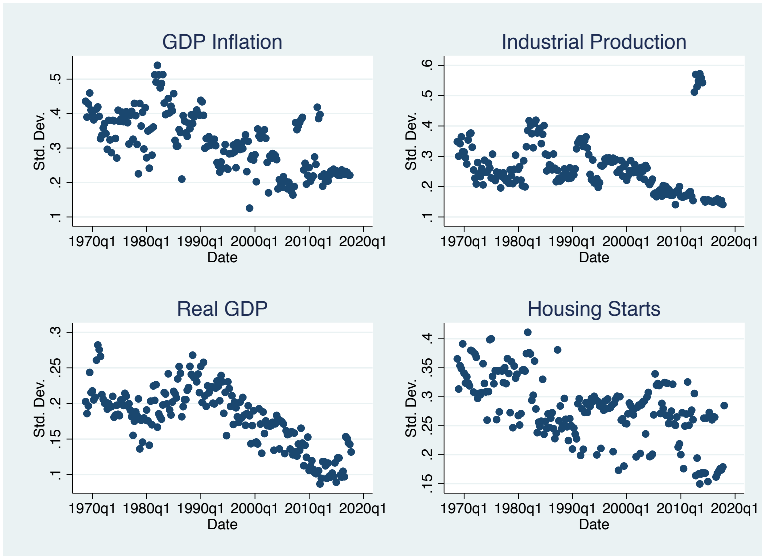

Figure 3.3: Time Variation in Dispersion

Notes: This figure shows the standard deviation in information flow in each time period. For each of the above variables, cross sectional dispersion appears to decline at later dates.

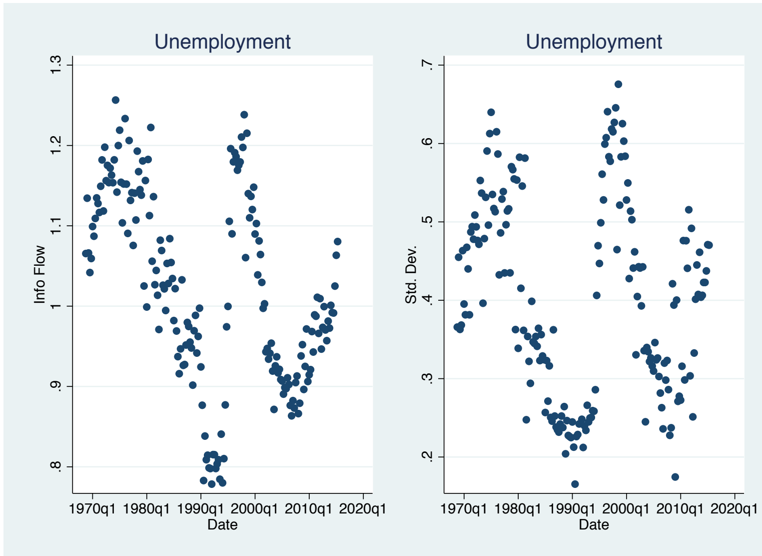

Figure 3.4: Time Variation in Information Flow and Dispersion - Unemployment

Notes: This figure shows time variation in the mean and standard deviation of information flow in unemployment. We see low frequency changes in both over time, but not a systematic time trend.

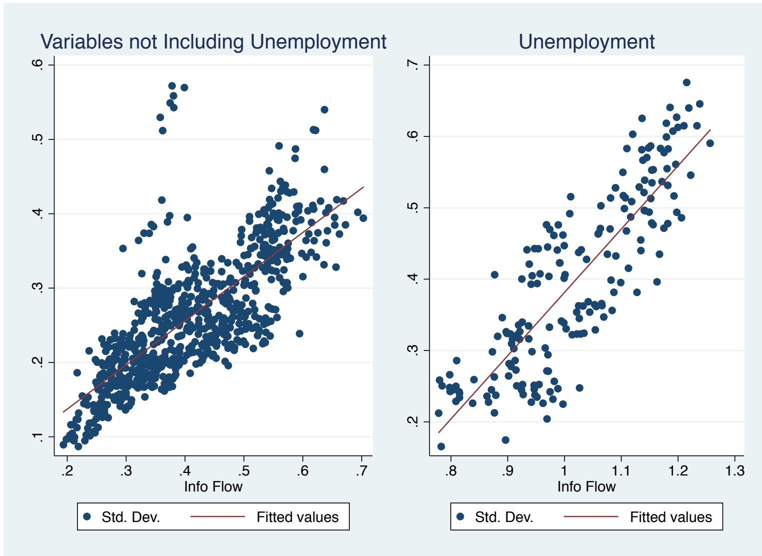

Figure 3.5: Information Flow and Dispersion

Notes: This figure shows the relationship between information flow and cross sectional dispersion for these variables. For the variables that exhibit declining information flow over time, standard deviation increases with information flow. The same is true for unemployment, even though we do not see a time trend in attention to unemployment.

Appendix

Appendix 1

Derivations

1.1 Basic Noisy Information Model

Derivation of Equation 1.4 :

$$
\begin{aligned}
\pi_{t+1,t|t}(i) &= \mu + \rho\pi_{t,t-1|t}(i) \\
&= \mu + \rho(kz_t(i) + (1-k)\pi_{t,t-1|t-1}(i)) \\
&= \mu + k\rho\pi_{t,t-1} + k\rho v_t(i) + \rho(1-k)\pi_{t,t-1|t-1}(i)
\end{aligned} \tag{1.1}
$$

Derivation of Equation 1.5:

To form the agent's forecast error, I subtract both sides of the above equation from $\pi_{t+1,t}$.

$$
\begin{aligned}
FE_{t+1,t|t}(i) &= \pi_{t+1,t} - (\mu + k\rho\pi_{t,t-1} + \rho(1-k)\pi_{t,t-1|t-1}(i)) \\
&= \mu + \rho\pi_{t,t-1} + w_{t+1} - (\mu + k\rho\pi_{t,t-1} + \rho(1-k)\pi_{t,t-1|t-1}(i)) \\
&= \rho(1-k)(\pi_{t,t-1} - \pi_{t,t-1|t-1}(i)) + w_{t+1} - k\rho v_t(i) \\
&= \rho(1-k)FE_{t,t-1|t-1}(i) + w_{t+1} - k\rho v_t(i)
\end{aligned} \tag{1.2}
$$

1.2 Model with Two Frictions

Derivation for Equation 1.10

$$\begin{aligned}
\pi_{t+1,t|t}(i) &= \mu + \tilde{\rho}\pi_{t,t-1|t}(i) \\
&= \mu + \tilde{\rho}(kz_t(i) + (1-k)\pi_{t,t-1|t-1}(i)) \qquad (1.3) \\
&= \mu + k\tilde{\rho}\pi_{t,t-1} + k\tilde{\rho}v_t(i) + \tilde{\rho}(1-k)\pi_{t,t-1|t-1}(i)
\end{aligned}$$

Derivation for Equation 1.11

$$\begin{aligned}
FE_{t+1,1|t}(i) &= \pi_{t+1,t} - (\mu + k\tilde{\rho}\pi_{t,t-1} + k\tilde{\rho}v_t(i) + \tilde{\rho}(1-k)\pi_{t,t-1|t-1}(i)) \\
&= \mu + \rho\pi_{t,t-1} + w_{t+1} - \mu - k\tilde{\rho}\pi_{t,t-1} - k\tilde{\rho}v_t(i) - \tilde{\rho}(1-k)\pi_{t,t-1|t-1}(i) \\
&= (\rho - \tilde{\rho}k)\pi_{t,t-1} - \tilde{\rho}(1-k)\pi_{t,t-1|t-1}(i) + w_{t+1} - \rho k v_t(i) \\
&= (\tilde{\rho} - q - \tilde{\rho}k)\pi_{t,t-1} - \tilde{\rho}(1-k)\pi_{t,t-1|t-1}(i) + w_{t+1} - \rho k v_t(i) \\
&= \tilde{\rho}(1-k)FE_{t,t-1|t-1}(i) - q\pi_{t,t-1} + w_{t+1} - \rho k v_t(i)
\end{aligned} \qquad (1.4)$$

1.3 Extensions

When forecaster i observes a private signal about current inflation and a public signal about lagged inflation, she forms her optimal nowcast according to:

$$\begin{aligned}
\pi_{t,t-1|t}(i) &= \mu + \rho\pi_{t-1,t-2|t-1}(i) + K\left[Z_t(i) - H_1\mu - (H_1\rho + H_2)\pi_{t-1,t-2|t-1}(i)\right] \\
&= \mu + \rho\pi_{t-1,t-2|t-1}(i) + k_1(\pi_{t,t-1} + v_t(i) - \mu - \rho\pi_{t-1,t-2|t-1}(i)) \\
&+ k_2(\pi_{t-1,t-2} + e_t - \pi_{t-1,t-2|t-1}(i))
\end{aligned} \qquad (1.5)$$

Regrouping terms gives:

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

Using the fact that $\pi_{t,t-1|t-1}(i) = \mu + \rho\pi_{t-1,t-2|t-1}(i)$:

$$\pi_{t,t-1|t}(i) = (1-k_1)\pi_{t,t-1|t-1}(i) - k_2\pi_{t-1,t-2|t-1}(i) + k_1\pi_{t,t-1} + k_2\pi_{t-1,t-2} + k_1 v_t(i) + k_2 e_t$$

The forecaster's projection for 1-quarter ahead inflation will take the following form:

$$\begin{aligned}
\pi_{t+1,t|t}(i) &= \mu + \rho((1-k_1)\pi_{t,t-1|t-1}(i) - k_2\pi_{t-1,t-2|t-1}(i) + k_1\pi_{t,t-1} + k_2\pi_{t-1,t-2} + k_1 v_t(i) + k_2 e_t) \\
&= \mu + \rho(1-k_1)\pi_{t,t-1|t-1}(i) - \rho k_2\pi_{t-1,t-2|t-1}(i) + \rho k_1\pi_{t,t-1} + \rho k_2\pi_{t-1,t-2} + \rho k_1 v_t(i) + \rho k_2 e_t
\end{aligned} \tag{1.6}$$

Subtracting both sides from $\pi_{t+1,t}$, where $\pi_{t+1,t} = \mu + \rho\pi_{t,t-1} + w_{t+1}$ gives the following equation for one-quarter ahead forecast errors:

$$\begin{aligned}
FE_{t+1,t|t}(i) &= \rho(1-k_1)(\pi_{t,t-1} - \pi_{t,t-1|t-1}(i)) + \rho k_2(\pi_{t-1,t-2|t-1}(i) - \pi_{t-1,t-2}) \\
&- \rho k_1 v_t(i) - \rho k_2 e_t + w_{t+1} \\
&= \rho(1-k_1)FE_{t,t-1|t-1}(i) - \rho k_2(\pi_{t-1,t-2} - \pi_{t-1,t-2|t-1}(i)) \\
&- \rho k_1 v_t(i) - \rho k_2 e_t + w_{t+1} \\
&= \rho(1-k_1)FE_{t,t-1|t-1}(i) - k_2(\rho\pi_{t-1,t-2} - \rho\pi_{t-1,t-2|t-1}(i)) \\
&- \rho k_1 v_t(i) - \rho k_2 e_t + w_{t+1}
\end{aligned} \tag{1.7}$$

We can then express the above equation as a relationship between quarter-ahead forecast errors and its lagged value as well as the lag of the nowcast error.

$$FE_{t+1,t|t}(i) = \rho(1-k_1)FE_{t,t-1|t-1}(i) - k_2(FE_{t-1,t-2|t-1}(i)) - \rho k_1 v_t(i) - \rho k_2 e_t + w_{t+1} \tag{1.8}$$

Derivation and interpretation of Kalman gain terms, k_1 and k_2

$$K = [\rho U^-(H_1\rho + H_2)' + \sigma_w^2 H_1' + \sigma_w R'] \times [(H_1\rho + H_2)U^-(H_1\rho + H_2)' + (H_1\sigma_w^2 + R)(H_1\sigma_w^2 + R)']$$

$$K = \begin{bmatrix} \rho U^-\rho + \sigma_w + \sigma_w\sigma_v & \rho U^- + \sigma_w\sigma_e \end{bmatrix} \times \begin{bmatrix} \rho U^-\rho + \sigma_w^2 + 2\sigma_w\sigma_v + \sigma_v^2 & \rho U^- + \sigma_w\sigma_e + \sigma_v\sigma_e \\ \rho U^- + \sigma_w\sigma_e + \sigma_v\sigma_e & U^- + \sigma_e^2 \end{bmatrix}^{-1}$$

Allow χ to symbolize the determinant of the 2×2 matrix in the above equation.

$$\chi = \frac{1}{U^-\sigma_w^2 + 2U^-\sigma_w\sigma_v + U^-\sigma_v^2 + \rho U^-\rho\sigma_e^2 - 2\rho U^-\sigma_w\sigma_e - 2\rho U^-\sigma_v\sigma_e}$$

$$K = \begin{bmatrix} \rho U^-\rho + \sigma_w^2 + \sigma_w\sigma_v & \rho U^- + \sigma_w\sigma_e \end{bmatrix} \times \begin{bmatrix} \frac{U^- + \sigma_e^2}{\chi} & \frac{-(\rho U^- + \sigma_w\sigma_e + \sigma_v\sigma_e)}{\chi} \\ \frac{-(\rho U^- + \sigma_w\sigma_e + \sigma_v\sigma_e)}{\chi} & \frac{\rho U^-\rho + \sigma_w^2 + 2\sigma_w\sigma_v + \sigma_v^2}{\chi} \end{bmatrix}$$

Multiplying through,

$$\begin{aligned} k_1 &= (\rho U^-\rho + \sigma_w^2 + \sigma_w\sigma_v) \times (U^- + \sigma_e^2) - (\rho U^- + \sigma_w\sigma_e) \times (\rho U^- + \sigma_w\sigma_e + \sigma_v\sigma_e) \times \chi \\ &= \frac{U^-\sigma_w^2 + U^-\sigma_w\sigma_v + \rho U^-\rho\sigma_e^2 - 2\rho U^-\sigma_w\sigma_e - \rho U^-\sigma_v\sigma_e}{U^-\sigma_w^2 + 2U^-\sigma_w\sigma_v + U^-\sigma_v^2 + \rho U^-\rho\sigma_e^2 - 2\rho U^-\sigma_w\sigma_e - 2\rho U^-\sigma_v\sigma_e} \end{aligned}$$

$$k_2 = -\left(\rho U^- \rho + \sigma_w^2 + \sigma_w \sigma_v\right) \times \left(\rho U^- + \sigma_w \sigma_e + \sigma_v \sigma_e\right) + \left(\rho U^- + \sigma_w \sigma_e\right) \times \left(\rho U^- \rho + \sigma_w^2 + 2\sigma_w \sigma_v + \sigma\right.$$

$$= \frac{\rho U^- \sigma_v^2 + \rho U^- \sigma_w \sigma_v - \rho U^- \rho \sigma_v \sigma_e}{U^- \sigma_w^2 + 2U^- \sigma_w \sigma_v + U^- \sigma_v^2 + \rho U^- \rho \sigma_e^2 - 2\rho U^- \sigma_w \sigma_e - 2\rho U^- \sigma_v \sigma_e}$$

Under full information, the forecaster receives a signal about today's inflation that is equal to the true value of inflation. As such, $v_t(i) = 0$ and $\sigma_v = 0$. Substituting into the equations for k_1 and k_2:

$$k_1 = \frac{U^- \sigma_w^2 + +\rho U^- \rho \sigma_e^2 - 2\rho U^- \sigma_w \sigma_e}{U^- \sigma_w^2 + \rho U^- \rho \sigma_e^2 - 2\rho U^- \sigma_w \sigma_e} = 1$$

$$k_2 = \frac{0}{U^- \sigma_w^2 + \rho U^- \rho \sigma_e^2 - 2\rho U^- \sigma_w \sigma_e} = 0$$

k_2 is only equal to zero under one other condition, when a standard deviation in the error noise is as large as the sum of a standard deviation in the inflation innovation and a standard deviation of the private signal noise. In this case the signal about past inflation is too noisy to be informative.

$$\rho U^- \sigma_v^2 + \rho U^- \sigma_w \sigma_v - \rho U^- \rho \sigma_v \sigma_e = 0 \Leftrightarrow \sigma_v + \sigma_w = \sigma_e$$

We can also consider the situation where the forecaster perfectly observes the lagged value of inflation, or $e_t = 0$ and $\sigma_e = 0$.

$$k_1 = \frac{U^- \sigma_w^2 + U^- \sigma_w \sigma_v}{U^- \sigma_w^2 + 2U^- \sigma_w \sigma_v + U^- \sigma_v^2}$$

$$k_2 = \frac{\rho U^- \sigma_v^2 + \rho U^- \sigma_w \sigma_v}{U^- \sigma_w^2 + 2U^- \sigma_w \sigma_v + U^- \sigma_v^2}$$

Under this circumstance, the agent weights the two signals according to the relative noise in the process innovation, w_t, and in her signal $v_t(i)$, with the signal about the past receiving more weight if the signal about the present is noisier and the signal about the present receiving more weight if inflation is more volatile. k_1 and k_2 should sum to 1.

Appendix 2

Forecasts at Longer Horizons

Deriving the relationship between forecast errors and lagged forecast errors for longer horizons requires transforming forecasts from higher-frequency observations to lower-frequency observations. As shocks compound quarterly, this transformation introduces endogeneity to the relationship between forecast errors and lagged forecast errors at longer horizons. This appendix presents derivations for longer horizon forecasts.

2.1 Two-Quarter Horizon

Inflation follows the same AR(1) process with shocks arriving each quarter. Agents further receive the same signals each period. [1] $\pi_{t+1,t}$ is inflation from period t to period $t+1$.

The forecast will follow:

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

and the forecast error:

[1] See Equations 1.1 and 1.2.

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

As shown in Section 3.2, we can estimate this equation by OLS under basic assumptions. The following derivation of semi-annual forecast errors

$$\begin{aligned} \pi_{t+2,t} &= \pi_{t+1,t} + \pi_{t+2,t+1} \\ &= \mu + (1+\rho)\pi_{t+1,t} + w_{t+2} \end{aligned} \tag{2.1}$$

The expectation of this event takes the same form.

$$\begin{aligned} \pi_{t+2,t|t}(i) &= \pi_{t+1,t|t}(i) + \pi_{t+2,t+1|t}(i) \\ &= \mu + (1+\rho)\pi_{t+1,t|t}(i) \\ &= \mu + (1+\rho)\left[\mu + \rho k \pi_{t,t-1} + \rho k v_t(i) + (1-k)\pi_{t,t-1|t-1}(i)\right] \end{aligned} \tag{2.2}$$

From Equations 2.1 and 2.2 the semi-annual forecast error is:

$$FE_{t+2,t|t}(i) = (1+\rho)FE_{t+1,t|t}(i) + w_{t+2} \tag{2.3}$$

This provides the semi-annual forecast error in terms of the lagged quarterly forecast error. The desired relationship is the semi-annual forecast error and the lagged semi-annual forecast error. As time is denominated in quarters, the desired lag of two-quarter inflation occurs at time $t-2$ rather than time $t-1$. Mirroring the structure of Equation 2.3, $FE_{t,t-2|t-2}(i)$ is given by

$$FE_{t,t-2|t-2}(i) = (1+\rho)FE_{t-1,t-2|t-2}(i) + w_t \tag{2.4}$$

Note that we can derive the relationship between the one-quarter ahead forecasts that appear in the Equations 2.3 and 2.4.

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

Plugging this into Equation 2.3:

$$\begin{aligned}
FE_{t+2,t|t}(i) &= (1+\rho)\big[\rho^2(1-k)^2 FE_{t-1,t-2|t-2}(i) + \rho(1-k)w_t + w_{t+1} - \rho^2(1-k)kv_{t-1}(i) - \rho k v_t(i \\
&+ w_{t+2} \\
&= (1+\rho)\rho^2(1-k)^2 FE_{t-1,t-2|t-2}(i) + (1+\rho)\rho(1-k)w_t + (1+\rho)w_{t+1} + w_{t+2} \\
&- (1+\rho)\rho^2(1-k)kv_{t-1}(i) - \rho(1+\rho)kv_t(i)
\end{aligned} \tag{2.5}$$

Rearranging Equation 2.4 gives:

$$FE_{t-1,t-2|t-2}(i) = \frac{1}{1+\rho}FE_{t,t-2|t-2}(i) - \frac{1}{1+\rho}w_t \tag{2.6}$$

Substituting this into Equation 2.5 gives the desired relationship between a semi-annual forecast and its appropriate lag.

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

The error term consists of signal noise terms for the periods between the two forecasting periods and shocks that occur in t, $t+1$, and $t+2$. The presence of w_t in the error term means that the error term is correlated with the dependent variable and this equation cannot be estimated by OLS.

2.2 Two-Quarter Horizon with Misperceived Persistence

The agent's quarter-ahead forecast will now follow:

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

and the forecast error:

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

The expectation of two-quarter ahead inflation takes the following form, while the realization is the same as in Equation 2.1.

$$
\begin{aligned}
\pi_{t+2,t|t}(i) &= \pi_{t+1,t|t}(i) + \pi_{t+2,t+1|t}(i) \\
&= \mu + (1+\tilde{\rho})\pi_{t+1,t|t}(i)
\end{aligned}
$$

The forecast error for two-quarter ahead inflation can then be written as:

$$
\begin{aligned}
FE_{t+2,t|t}(i) &= \mu + (1+\rho)\pi_{t+1,t|t} + w_{t+2} - (\mu + (1+\tilde{\rho})\pi_{t+1,t|t}(i)) \\
&= (1+\tilde{\rho}-q)\pi_{t+1,t|t} - (1+\tilde{\rho})\pi_{t+1,t|t}(i) + w_{t+2} \qquad (2.7) \\
&= (1+\tilde{\rho})FE_{t+1,t|t}(i) - q\pi_{t+1,t} + w_{t+2}
\end{aligned}
$$

Similarly, we can write the two-quarter ahead forecast error from two quarters ago as:

$$
FE_{t,t-2|t-2}(i) = (1+\tilde{\rho})FE_{t-1,t-2|t-2}(i) - q\pi_{t-1,t-2} + w_t \qquad (2.8)
$$

We can derive the relationship between the quarter ahead forecast errors in Equations 2.7 and 2.8:

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

Plugging this into 2.7 gives:

$$
\begin{aligned}
FE_{t+2,t|t}(i) = & (1+\tilde{\rho})\tilde{\rho}^2(1-k)^2 FE_{t-1,t-2|t-2}(i) - (1+\tilde{\rho})\tilde{\rho}(1-k)\pi_{t-1,t-2} - (1+\tilde{\rho})q\pi_{t,t-1} - q\pi_{t+1,t} \\
& + (1+\tilde{\rho})\tilde{\rho}(1-k)w_t + (1+\tilde{\rho})w_{t+1} + w_{t+2} \\
& - (1+\tilde{\rho})\tilde{\rho}^2(1-k)kv_{t-1}(i) - (1+\tilde{\rho})\tilde{\rho}kv_t(i)
\end{aligned}
\tag{2.9}
$$

Rearranging 2.8 gives us the following:

$$
FE_{t-1,t-2|t-2}(i) = \frac{1}{1+\tilde{\rho}} FE_{t,t-2|t-2}(i) + \frac{q}{1+\tilde{\rho}}\pi_{t-1,t-2} - \frac{1}{1+\tilde{\rho}} w_t. \tag{2.10}
$$

We can then substitute this into Equation 2.11 to obtain:

$$
\begin{aligned}
FE_{t+2,t|t}(i) = & \tilde{\rho}^2(1-k)^2 FE_{t,t-2|t-2}(i) - \tilde{\rho}(1-k)(1+\tilde{\rho}k)\pi_{t-1,t-2} - (1+\tilde{\rho})q\pi_{t,t-1} - q\pi_{t+1,t} \\
& + \tilde{\rho}(1-k)(1+\tilde{\rho}k)w_t + (1+\tilde{\rho})w_{t+1} + w_{t+2} \\
& - (1+\tilde{\rho})\tilde{\rho}^2(1-k)kv_{t-1}(i) - (1+\tilde{\rho})\tilde{\rho}kv_t(i)
\end{aligned}
\tag{2.11}
$$

Where I use that $[\tilde{\rho}^2(1-k)^2 - (1+\tilde{\rho})\tilde{\rho}(1-k)] = \tilde{\rho}(1-k)(1+\tilde{\rho}k)$. The predicted path of forecast errors now includes multiple realizations of inflation in addition to the endogeneity problem identified in the previous section. When we add time fixed effects to this regression to control for the w_t, it will absorb the effect of the realizations of the time-dependent realizations of inflation.

Appendix 3

Misperception of the Constant

3.1 Forecast Errors with Incorrectly Observed Constant

A misperception of the constant of the inflation process, μ, creates a different form of forecast errors. Define $\mu_i = \mu + d_i \ \forall i$. Further assume constant beliefs across forecasters and call $\mu_i = \tilde{\mu}$ and $d_i = d$ for all forecasters. Forecasters will form their a priori beliefs about future inflation using their perceived constant.

$$\begin{aligned}
\pi_{t+1,t|t}(i) &= \tilde{\mu} + \rho \pi_{t,t-1|t}(i) \\
&= \tilde{\mu} + \rho(kz_t(i) + (1-k)\pi_{t,t-1|t-1}(i) \\
&= \tilde{\mu} + \rho k \pi_{t,t-1} + \rho(1-k)\pi_{t-1|t-1}(i) + \rho k v_t(i)
\end{aligned} \tag{3.1}$$

Subtracting both sides from the realization of $\pi_{t+1,t}$ and substituting $\tilde{\mu} = \mu + d$ gives the following equation for forecast errors.

$$\begin{aligned}
FE_{t+1,t|t}(i) &= = \mu + \rho \pi_{t,t-1} + w_{t+1} - \tilde{\mu} + \rho k \pi_{t,t-1} + \rho(1-k)\pi_{t-1|t-1}(i) + \rho k v_t(i) \\
&= -d + \rho(1-k)FE_{t,t-1|t-1}(i) - q\pi_{t,t-1} + w_{t+1} - \rho k v_t(i)
\end{aligned} \tag{3.2}$$

Should $\tilde{\mu} = \mu$ and $d = 0$ for all time periods, as is the case when the constant is observed, the constant will drop from the forecast error equation.

If forecasters misperceive the constant, the estimation will simply produce a nonzero constant term.

3.2 Forecast Errors if Both Parameters are Incorrectly Observed

If forecasters mis-estimate both parameters, the quarter-ahead forecast errors will follow a pattern combining the effects of the last two sections.

$$FE_{t+1,t|t}(i) = -d + \rho(1-k)FE_{t,t-1|t-1}(i) - q\pi_{t,t-1} + w_{t+1} - \rho k v_t(i) \quad (3.3)$$

Using this equation, I estimate the following reduced form equation to uncover the parameters d and q .

$$\begin{aligned} FE_{t+1,t}(i) &= \beta_0 + \beta_1 FE_{t,t-1}(i) + \beta_2 \pi_{t,t-1}(i) + \epsilon_t(i) \\ &= -0.2003^{***} + 0.3764^{***} FE_{t,t-1|t-1}(i) + 0.0849^{***} \pi_{t,t-1} \end{aligned} \quad (3.4)$$

Under the null that there is no misperception of parameters, $\beta_0 = 0$ and $\beta_2 = 0$. This estimation implies that $d = 0.200$ and $q = -0.085$, meaning that agents overestimate the regression constant and underestimate persistence in for the inflation process.

Appendix 4

Derivations and Results for Nowcast Errors

4.1 Predicted Path of Nowcast Errors

The Kalman filter model consists of the following process and measurement equations.

$$\pi_{t,t-1} = \mu + \rho\pi_{t-1,t-2} + w_t \tag{4.1}$$

$$z_t(i) = \pi_{t,t-1} + v_t(i) \tag{4.2}$$

The optimal nowcast is a linear combination of the signal and the agent's prior expectation.

$$\pi_{t,t-1|t}(i) = kz_t(i) + (1-k)\pi_{t,t-1|t-1}(i) \tag{4.3}$$

Substituting the process and measurement equations, 4.1 and 4.2, into the optimal nowcast equation, 4.3 gives the following relationship:

$$
\begin{aligned}
\pi_{t,t-1|t}(i) &= k(\pi_{t,t-1} + v_t(i)) + (1-k)\pi_{t,t-1|t-1}(i) \\
&= k(\mu + \rho\pi_{t-1,t-2} + w_t + v_t(i)) + (1-k)(\mu + \rho\pi_{t-1|t-1}(i)) \\
&= \mu + k(\rho\pi_{t-1,t-2} + w_t + v_t(i)) + (1-k)(\rho\pi_{t-1|t-1}(i))
\end{aligned}
$$

Subtracting both sides from the true value of $\pi_{t,t-1}$:

$$
\begin{aligned}
\pi_{t,t-1} - \pi_{t,t-1|t}(i) &= \mu + \rho\pi_{t-1,t-2} + w_t - \mu - \rho k\pi_{t-1,t-2} - \rho(1-k)\pi_{t-1|t-1}(i) - kw_t - kv_t(i) \\
&= \rho(1-k)(\pi_{t-1} - \pi_{t-1|t-1}(i)) + (1-k)w_t - kv_t(i)
\end{aligned}
$$

Note that the constant from the transition equation drops out of the forecast error equation. For ease of notation, let $FE_{t,t-1|t}(i)$ take the place of $\pi_{t,t-1} - \pi_{t,t-1|t}(i)$.

$$
\begin{aligned}
FE_{t,t-1|t}(i) &= \rho(1-k)FE_{t-1,t-2|t-1}(i) + (1-k)w_t - kv_t(i) \\
FE_{t,t-1|t}(i) &= 0.06^{***} + 0.29^{***}FE_{t-1,t-2|t-1}(i)
\end{aligned}
$$

4.2 Predicted Path of Nowcast Errors with Mis-perceived Persistence

Forecasters form their a priori beliefs about inflation using their perceived persistence, $\tilde{\rho} = \rho + q$.

$$\begin{aligned}
\pi_{t,t-1|t}(i) &= k(\pi_{t,t-1} + v_t(i)) + (1-k)\pi_{t,t-1|t-1}(i) \\
&= k(\mu + \rho\pi_{t-1,t-2} + w_t + v_t(i)) + (1-k)(\mu + \tilde{\rho}\pi_{t-1,t-2|t-1}(i)) \\
&= \mu + k(\rho\pi_{t-1,t-2} + w_t + v_t(i)) + \tilde{\rho}(1-k)\pi_{t-1,t-2|t-1}(i)
\end{aligned}$$

The corresponding forecast error is therefore:

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

4.3 Predicted Path of Nowcast Errors with Mis-perceived Inflation Constant

In this case the forecaster applies the transition equation to $\pi_{t,t-1|t-1}(i)$ with the incorrect constant, $\tilde{\mu} = \mu + d$.

$$\begin{aligned}
\pi_{t,t-1|t}(i) &= k(\pi_{t,t-1} + v_t(i)) + (1-k)\pi_{t,t-1|t-1}(i) \\
&= k(\mu + \rho\pi_{t-1,t-2} + w_t + v_t(i)) + (1-k)(\tilde{\mu} + \rho\pi_{t-1,t-2|t-1}(i)) \\
&= k\mu + (1-k)\tilde{\mu} + k(\rho\pi_{t-1,t-2} + w_t + v_t(i)) + (1-k)\rho\pi_{t-1,t-2|t-1}(i) \\
&= k\mu + (1-k)(\mu + d) + k(\rho\pi_{t-1,t-2} + w_t + v_t(i)) + (1-k)\rho\pi_{t-1,t-2|t-1}(i) \\
&= \mu - (1-k)d + k(\rho\pi_{t-1,t-2} + w_t + v_t(i)) + (1-k)\rho\pi_{t-1,t-2|t-1}(i)
\end{aligned}$$

The nowcast error under these circumstances:

$$FE_{t,t-1|t}(i) = (1-k)d + \rho(1-k)FE_{t-1,t-2|t-1}(i) + (1-k)w_t - kv_t(i)$$

4.4 Predicted Path of Nowcast Errors with Misperceptions of Both Persistence and the Constant

With misperception of both parameters, the predicted path of nowcast errors is:

$$\begin{aligned} FE_{t,t-1|t}(i) &= (1-k)d + \tilde{\rho}(1-k)FE_{t-1,t-2|t-1}(i) - q(1-k)\pi_{t-1,t-2} + (1-k)w_t - kv_t(i) \\ &= -0.06 + 0.27^{***}FE_{t-1,t-2|t-1}(i) + 0.04^{***}\pi_{t-1,t-2} \end{aligned}$$

This again provides evidence for the underestimation of persistence as the interpretation of the coefficient on $\pi_{t-1,t-2}$ is $-q$. A positive coefficient implies that q is negative, or that forecasters underestimate inflation persistence.

process.

Bibliography

Afrouzi, H. (2017). Strategic inattention, inflation dynamics and the non-neutrality of money. Mimeo.

Afrouzi, H., Coibion, O., Gorodnichenko, Y. and Kumar, S. (2015). Inflation targeting does not anchor inflation expectations: Evidence from firms in new zealand, *Brookings Papers on Economic Activity* **Fall**: 151–225.

Andrade, P., Crump, R. K., Eusepi, S. and Moench, E. (2016). Fundamental disagreement, *Journal of Monetary Economics* **8**(C).

Andrade, P. and Le Bihan, H. (2013). Inattentive professional forecasters, *Journal of Monetary Economics* **60**(8): 967–982.

Angeletos, G.-M. and La'O, J. (2009). Incomplete information, higher-order beliefs and price inertia, *Journal of Monetary Economics* **56**.

Atkeson, A. and Ohanian, L. E. (n.d.). Are phillips curves useful for forecasting inflation?

Bacchetta, P. and Wincoop, E. V. (2008). Higher order expectations in asset pricing, *Journal of Money, Credit, and Banking* **40**: 837–866.

Benati, L. (2008). Investigating inflation persistence across monetary regimes, *Quarterly Journal of Economics* **123**(3).

Bonham, C. S. and Cohen, R. H. (1995). Testing the rationality of price forecasts, *The American Economic Review* **85**.

Bonham, C. S. and Cohen, R. H. (2017). To aggregate, pool, or neither: Testing the rational expectations hypothesis using survey data, *Journal of Business and Economic Statistics* **19**.

Bürgi, C. (2017). What do we lose when we average expectations? Mimeo.

Camerer, C. F., Ho, T.-H. and Chong, J.-K. (2004). A cognitive hierarchy model of games, *Quarterly Journal of Economics* **119**: 861–898.

Cogley, T., Primiceri, G. E. and Sargent, T. J. (2010). Inflation-gap persistence in the us, *American Economic Journal: Macroeconomics* **2**(1).

Cogley, T. and Sargent, T. J. (2002). Evolving post-world war ii u.s. inflation dynamics, *NBER Macroeconomics Annual 2001, Volume 16*, MIT Press.

Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about informational rigidities?, *Journal of Political Economy* **120**(1): 116–159.

Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts, *American Economic Review* **105**(8): 2644–2678.

Coibion, O., Gorodnichenko, Y. and Kumar, S. (2018). How do firms form their expectations? new survey evidence. Forthcoming: The American Economic Review,.

Costa-Gomes, M. A. and Crawford, V. P. (2006). Cognition and behavior in two-person guessing games: An experimental study, *American Economic Review* **96**: 1737–1768.

Crowe, C. (2010). Consensus forecasts and information aggregation. IMF Working Paper 10/178.

Dovern, J., Fritsche, U., Loungani, P. and Tamirisa, N. (2014). Information rigidities: Comparing average and individual forecasts for a large international panel, *International Journal of Forecasting* **31**: 144–154.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data, *The Review of Economics and Statistics* **80**: 549–560.

Erceg, C. and Levin, A. (2003). Imperfect credibility and inflation persistence, *Journal of Monetary Economics* **50**: 915–?944.

Fahri, I. and Werning, I. (2017). Monetary policy, bounded rationality, andincomplete markets. Mimeo,.

Garcia-Schmidt, M. and Woodford, M. (2015). Are low interest rates deflationary? a paradox of perfect foresight analysis. Mimeo,.

Jain, M. (2017). Perceived inflation persistence, *Journal of Business and Economic Statistics* **Online**.

Keane, M. P. and Runkle, D. E. (1989). Are economic forecasts rational, *Quarterly Review, Federal Reserve Bank of Minneapolis* **Spring**.

Mackowiak, B. and Wiederholt, M. (2009). Optimal sticky prices under rational inattention, *American Economic Review* **99**(3): 769–803.

Mackowiak, B., Wiederholt, M. and Matejka, F. (2017). Strategic inattention, inflation dynamics and the non-neutrality of money. Mimeo.

Mankiw, N. G. and Reis, R. (2002). Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve, *Quarterly Journal of Economis* **117**(4): 1295–1328.

Milani, F. (2007). Expectations, learning, and macroeconomic persistence, *Journal of Monetary Economics* **54**: 2065–2082.

Mishkin, F. (2007). Inflation dynamics, *International Finance* **10**(3).

Morris, S. and Shin, H. S. (2002). The social value of public information, *American Economic Review* **92**: 1521–1534.

Nagel, R. (1995). Unraveling in guessing games: An experimental study, *American Economic Review* **85**: 1313–1326.

Nagel, R. and Duffy, J. (1997). On the robustness of behavior in experimental beauty-contest, *Economic Journal* **107**: 1684–1700.

Nimark, K. P. (2014). A low dimensional kalman filter for systems with lagged states in the measurement equation. Mimeo.

Nordhaus, W. (1987). Forecasting efficiency: Concepts and applications, *The Review of Economics and Statistics* **69**: 667–674.

Orphanides, A. and Williams, J. (2004). Imperfect knowledge, inflation expectations, and monetary policy, *The Inflation-Targeting Debate*, University of Chicago Press.

Paternoster, R., Brame, R., Mazerolle, P. and Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients, *Criminology* **36**(4): 859–866.

Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics* **68**: 79–113.

Pesaran, M. H. and Weale, M. (2006). Survey expectations, *The Handbook of Economic Forecasting*.

Phelps, E. S. (n.d.). Phillips curves, expectations of inflation and optimal unemployment over time.

Pivetta, F. and Reis, R. (2007). The persistence of inflation in the united states, *Journal of Economic Dynamics and Control* **31**(4).

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

Sheng, X. S., Wallen, J. and An, Z. (2017). Inattention: Measurement, determinants and policy implications.

Sims, C. (2003). Implications of rational inattention, *Journal of Monetary Economics* **50**(3).

Stahl, D. O. and Wilson, P. W. (1995). Cognition and behavior in two-person guessing games: An experimental study, *Games and Economic Behavior* **10**: 218–254.

Stock, J. H. and Watson, M. W. (2007). Why has u.s. inflation become harder to forecast?, *Journal of Money, Credit, and Banking* **39**(1).

Woodford, M. (2002). Imperfect common knowledge and the effects of monetary policy, *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund Phelps*, Princeton Univ. Press.