Misspecified Expectations Among Professional Forecasters*

Julio L. Ortiz †
Federal Reserve Board

July 2023

Abstract

Analyzing professional forecasts, I find that a theory of expectation formation in which respondents' perceived law of motion differs from the objective law of motion, tends to outperform alternative models in its ability to fit prediction errors and revisions. Misspecified expectations outperforms the alternatives for a variety of macroeconomic aggregates and subsamples. Misspecified expectations provides a successful description of the data in part because it is able to match patterns of overreaction, delayed overshooting, persistent disagreement, and updating behavior. I conclude that misspecified expectations can serve as a suitable theory to model expectation formation among professional forecasters.

Keywords: Non-rational expectations. Expectation formation. Noisy information. Overreactions. Misspecification.
JEL Codes: D83, D84, E70

*I would like to thank Francesco Bianchi, Carola Binder, Thiago Ferreira, Jaime Marquez, Stephen Terry, and Fabian Winkler for their valuable feedback and helpful conversations. The views expressed in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

†Email: julio.l.ortiz@frb.gov.
Address: 20th and C Street NW, Washington, DC 20551.

1 Introduction

There is ample survey-based evidence that forecast errors are predictable on the basis of information available to respondents in real time. This finding has animated a rich literature studying violations of full information rational expectations (FIRE) (Bordalo et al., 2020; Broer and Kohlhas, 2022; Coibion and Gorodnichenko, 2012, 2015; Farmer et al., 2021; Fuhrer, 2018; Kohlhas and Walther, 2021; Kucinskas and Peters, 2022; Ryngaert, 2017). While several theoretical departures from FIRE can explain forecast inefficiency, the literature has not yet settled on a benchmark non-FIRE model (Reis, 2020). This paper undertakes a formal analysis of competing theories of expectation formation in an effort to move toward establishing such a benchmark in the context of professional forecasting.

In this paper I study a noisy information rational expectations model along with three leading non-FIRE theories which are: diagnostic expectations (Bordalo et al., 2020), overconfidence (Daniel et al., 1998), and misspecified expectations (Fuster et al., 2010), the latter of which is a theory where forecasters mis-perceive the law of motion of the process that they are forecasting. Using panel data from the U.S. Survey of Professional Forecasters, I estimate these models and find that misspecified expectations best fits observed forecast errors and revisions across a range of macroeconomic variables. Misspecification is successful relative to the other theories in part because it is able to produce "overshooting" of aggregate forecast errors, persistent disagreement, and more empirically consistent updating rules among forecasters.

Misspecification, however, does not uniformly dominate the other non-FIRE theories. I find that diagnostic expectations outperforms the alternatives for some variables such as aggregate investment. Overall, however, misspecified expectations provides a superior fit to macroeconomic time series that exhibit partial reversals (i.e., the first order autocorrelation coefficient is positive while the second order autocorrelation coefficient is negative), in part because getting such dynamics wrong generates greater persistence in forecast errors and forecast disagreement which more closely matches the data. Taken together, my results suggest that a theory in which decision-makers adopt parsimonious models of richer underlying processes (e.g., specify an AR(1) instead of the true AR(2) process) can serve as a benchmark theory to describe expectation formation among professional forecasters for a variety macroeconomic aggregates.

I begin by illustrating a simple forecasting problem in a noisy information environment. The macroeconomic variable of interest follows an AR(2) process and is unobservable to forecasters. Forecasters issue predictions of this aggregate based on information gleaned from a contemporaneous private signal. In this linear Gaussian setting, forecasters employ the Kalman filter to obtain the optimal forecast which is consistent with the conditional expectation.

I next examine three models of expectation formation and assess their ability to fit the data relative to the rational baseline and to each other. In this paper, I focus on straightforward biases which can be flexibly embedded into more complex macroeconomic models. I estimate these models via maximum likelihood estimation (MLE). Inspecting the likelihood functions and information criteria, I find that misspecified expectations outperforms the alternatives both in and out of sample for real GDP growth, which I focus on for my baseline results, as well as a range of other macroeconomic variables.[1] While I abstract away from learning in my set up, I analyze the relative performance of the models over rolling windows of the data as well as by forecaster "age" in the survey, and argue that learning about the data generating process, which could favor misspecified expectations, does not drive my results.

Digging deeper, I show that misspecified expectations generally outperforms the alternatives because it is better able to match certain features of the data. First, I verify that all three estimated biased models generate forecaster-level overreaction based on the Bordalo et al. (2020) regression of errors-on-revisions. I also examine the consensus-level analog to this regression (Coibion and Gorodnichenko, 2015) and find that the estimated non-rational models are uniformly less successful in matching this moment. Second, I show that the estimated misspecified expectations and diagnostic expectations models are quantitatively able to produce overshooting (Angeletos et al., 2020b; Bianchi et al., 2022) in aggregate expectations, with misspecified expectations producing a more empirically consistent path of overshooting. Third, misspecified expectations is able to generate

[1] My results are also robust to assuming an AR(1) process rather than an AR(2) process, as detailed in Appendix C.

more persistent disagreement across horizons relative to the other models. Finally, misspecified expectations is better able to match updating behavior in the data based on the estimated weight placed on new information. The estimated overconfidence and diagnostic expectations models imply a weight placed on new information that is greater than or equal to one whereas the estimated misspecified expectations model delivers a weight of about 0.70 which is closer to what I estimate at the one-quarter ahead horizon.

Though I analyze three non-rational models, my findings indicate that unconditional overreaction, conditional overreaction (i.e., overshooting), persistent disagreement, and updating rules are all important features of the data to match. Hence, beyond the set of theories that I consider, the ideal alternative to FIRE should successfully match these moments. Furthermore, while misspecified expectations is best able to fit the data for a wider range of macroeconomic forecasts, it is important to note that my results are specific to the context of professional forecasting. The relative rankings of the candidate models that I examine may be different when analyzing household or firm expectations, or when studying micro-level expectations rather than forecasts about aggregate outcomes. Nonetheless, given that professional forecasters are arguably the most well-informed agents in the economy, the literature on survey expectations has found these predictions attractive in part to infer a lower bound on economy-wide information frictions and biases (Bianchi et al., 2022; Carroll, 2003; Coibion and Gorodnichenko, 2015; Cornand and Hubert, 2022).

My paper relates to a number of both longstanding and more recent contributions to the literature on survey expectations and violations of FIRE. One strand of the literature studies non-rational biases in survey forecasts. Bordalo et al. (2020) find that forecast errors and revisions are negatively correlated at the individual level. I use this correlation as the relevant measure of overreactions in Section 5. To explain this particular violation of FIRE, Bordalo et al. (2020) propose a theory of diagnostic expectations. In earlier contributions, Daniel et al. (1998) and Moore and Healy (2008) propose theories of overconfidence in which forecasters believe their private information to be more precise than it truly is. Other notable contributions can be found in the finance literature (Kent and Hirshleifer, 2015; Kent et al., 1998, 2001). Inspecting the term structure of forecast uncertainty, Binder et al. (2022) also documents evidence consistent with overconfidence. Other behavioral biases have also been proposed in the literature such as intrinsic expectations (Fuhrer, 2018), relative overconfidence (Broer and Kohlhas, 2022), and overpersistence bias (Rozsypal and Schlafmann, 2023). The misspecified expectations model that I estimate implies that forecasters tend to overextrapolate, in line with overpersistence bias.

Another strand of the literature studies rational deviations from FIRE. For instance, Azeredo da Silvera et al. (2020) proposes a theory of noisy memory in which forecasters optimize over their history of past signals.[2] Moreover, Farmer et al. (2021) argues that many of the forecasting anomalies in the data can be rationalized in a learning environment. More closely related to the notion of misspecification that I study, Branch and Evans (2006) and Pfajfar (2013) develop models featuring intrinsic heterogeneity in which agents optimally adopt different predictor functions. Therefore, while diagnostic expectations and overconfidence correspond to behavioral biases, misspecified expectations can arise due to a behavioral bias or because adopting parsimonious forecasting models is optimal.

The rest of the paper is organized as follows. Section 2 details the noisy information setting and the rational baseline model. Section 3 outlines the menu of biased models. Section 4 estimates these models via MLE and assesses their fit to the data. Section 5 explores why misspecified expectations outperforms the alternatives. Section 6 concludes.

2 A Baseline Rational Expectations Model

I begin by outlining a noisy information rational expectations model which will serve as the baseline non-FIRE theory against which I will compare the other non-FIRE models. The model outlined here is in the spirit of Coibion and Gorodnichenko (2015) and Bordalo et al. (2020), though under the assumption of a richer driving process which later allows me to model misspecified expectations

[2] Afrouzi et al. (2023) similarly propose a model of working memory to explain biases in expectations. The costly retrieval of past information in their model can be interpreted as emanating from memory constraints or availability biases.

as a world in which forecasters assume simpler dynamics.

Suppose that a forecaster wishes to predict some aggregate variable, x_t, which follows an AR(2) process:

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + w_t, \quad w_t \sim N(0, \sigma_w^2).$$

At a given point in time, t, the forecaster, indexed by i, has access to a noisy private signal,

$$y_t^i = x_t + v_t^i, \quad v_t^i \sim N(0, \sigma_v^2),$$

The forecaster's objective is to minimize her mean squared errors. Hence, the optimal forecast of x_t is the conditional expectation $\mathbb{E}(x_t|\mathcal{I}_t^i)$, where $\mathcal{I}_t^i$ denotes the forecaster's information set at time t. In this linear Gaussian setting, the forecaster can employ the Kalman filter to obtain an optimal forecast. From the Kalman filter, we obtain the following predict and update equations for x_t,

$$\text{Predict:} \quad x_{t|t-1}^i = \rho_1 x_{t-1|t-1}^i + \rho_2 x_{t-2|t-1}^i \tag{1}$$

$$\text{Update:} \quad x_{t|t}^i = x_{t|t-1}^i + \kappa_1 (y_t^i - x_{t|t-1}^i) \tag{2}$$

$$x_{t-1|t}^i = x_{t-1|t-1}^i + \kappa_2 (y_t^i - x_{t|t-1}^i), \tag{3}$$

where κ_1 and κ_2 denote the Kalman gains, which are the optimal weights placed on the private signal when updating the current prediction $x_{t|t}^i$ and the previous prediction $x_{t-1|t}^i$, respectively.

This baseline model yields a number testable predictions. For instance, the forecast error, $x_t - x_{t|t}^i$, should be uncorrelated with anything residing in the forecaster's information set at time t. In addition, the forecast revision, $x_{t|t}^i - x_{t|t-1}^i$, is uncorrelated with anything in the forecaster's information set through $t-1$. These orthogonality conditions, however, are violated in the data. I next consider a menu of alternative theories of expectation formation which deviate from this benchmark model.

3 Alternative Models of Beliefs

In this section, I describe three theories of expectation formation which can each, in principle, match well-known features of survey expectations.

3.1 Overconfidence (OC)

Overconfidence is a theory in which individuals believe their private signals to be more informative than they truly are (Daniel et al., 1998). As a result, forecasters trust their signals more than is warranted and consequently place excessive weight on new private information.

Rather than accurately processing their respective private signals, forecasters mis-perceive the amount of noise associated with private information:

$$y_t^i = x_t + v_t^i \qquad v_t^i \overset{\text{i.i.d.}}{\sim} N(0, \breve{\sigma}_v^2),$$

where $\breve{\sigma}_v = \alpha_v \sigma_v$ and $\alpha_v \in [0, 1]$. When $\alpha_v < 1$, then the forecaster is overconfident and when $\alpha_v > 1$, then the forecaster is underconfident. The predict and updating rules are as in the rational case outlined in the previous section, but with excessively large weights, $\widehat{\kappa}_1$ and $\widehat{\kappa}_2$, placed on new information.

3.2 Diagnostic Expectations (DE)

Diagnostic expectations is a theory that is rooted in the Kahneman and Tversky (1972) representativeness heuristic. According to diagnostic expectations, agents form their beliefs subject to a cognitive friction in which they conflate the objective likelihood of a type in a group with its representativeness (i.e., the frequency of the type within the group *relative* to a reference group). This is formalized in Gennaioli and Shleifer (2010) who define the representativeness of a type τ for group G as:

$$R(\tau, G) = \frac{h(T = \tau | G)}{h(T = \tau | -G)},$$

where $h(T = \tau|G)$ denotes the distribution of variable T in group G. According to diagnostic expectations, subjective probabilities are formed based on a distorted density,

$$h^{\varphi}(T = \tau|G) \propto h(T = \tau|G)R(\tau, G)^{\varphi}.$$

Under rational expectations, $\varphi = 0$, while under diagnostic beliefs, $\varphi > 0$.

In this context, forecasters form expectations about the future state of the world, and they judge future outcomes that are more representative to be more likely. Following Bordalo et al. (2019, 2020), since the latent state described in the previous section follows a Markov process with an objective conditional distribution of $f(x_t|x_{t-1})$ and a signal history, $Y_t^i = \{y_t^i\}_{t=0}$, we can characterize a diagnostic forecaster's perceived distribution as,

$$f^{\varphi}(x_t|Y_t^i) \propto f(x_t|Y_t^i)\left[\frac{f(x_t|Y_t^i)}{f\left(x_t|Y_{t-1}^i \cup x_{t|t-1}^i\right)}\right]^{\varphi}.$$

In this case, the reference group reflects a "no news" scenario in which today's signal is uninformative, $y_t^i = x_{t|t-1}^i$. Therefore, when $\varphi > 0$, the forecaster exhibits representativeness with respect to recent news. As a result, diagnostic expectations as devised here is a theory of overreaction to new information, and we can interpret this bias as reflecting imperfect memory and associativeness.

A diagnostic forecaster's current-period forecast of x_t is,

$$\widehat{x}_{t|t}^i = x_{t|t}^i + \varphi\left[x_{t|t}^i - x_{t|t-1}^i\right],$$

and based on the assumed AR(2) dynamics, her one-step ahead forecast is

$$\widehat{x}_{t+1|t}^i = \rho_1\widehat{x}_{t|t}^i + \rho_2\widehat{x}_{t-1|t}^i.$$

3.3 Misspecified Expectations (ME)

It is possible that forecasters misunderstand the data generating process more generally. I next consider a bias in which forecasters adopt simple models of the world in the spirit of Fuster et al. (2010) and Molavi (2022). Perhaps it is optimal for forecasters to use more parsimonious models rather than estimate the richer dynamics governing the true data generating process (Branch and Evans, 2006; Brock and Hommes, 1997; Pfajfar, 2013). Or perhaps forecasters are predisposed to use simpler models due to cognitive frictions. A number of theories fall under the umbrella of misspecification. For instance, Gabaix (2019) describes a bias in which agents anchor the autocorrelation of several processes to a reference persistence. A similar bias is documented in Rozsypal and Schlafmann (2023) for the case of overextrapolation. In a different setting, Barberis et al. (1998) presents a model in which an investor mis-perceives a random walk process.

To capture this form of expectation formation, I focus on a version of misspecification which is closest to natural expectations as modeled in Fuster et al. (2010). As with natural expectations, I assume that forecasters neglect longer lags in the data generating process. In its simplest form, the underlying state follows an AR(2) process:

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + w_t, \quad w_t \sim N(0, \sigma_w^2),$$

but forecasters treat x_t as an AR(1) process when devising their predictions

$$x_t = \widehat{\rho} x_{t-1} + u_t,$$

where $u_t = (\rho_1 - \widehat{\rho})x_{t-1} + \rho_2 x_{t-2} + w_t$.[3] Importantly, forecasters still understand the information structure. If the perceived persistence loads excessively onto the first lag, then forecasters will overextrapolate.

[3]This form of misspecification is technically different from Fuster et al. (2010). First, I do not explicitly model an AR(2) in levels and assume that agents forecast an AR(1) in growth rates. Second, here, the perceived persistence is estimated from the data rather than defined to be a function of the true autocorrelation parameters. Third, Fuster et al. (2010) define natural expectations as a weighted average of rational expectations and the naive AR(1) expectations. I do not make such an assumption under misspecified expectations.

3.4 Other Models

In general, existing alternatives to FIRE can be categorized into one of two groups: models that generate underreactions and models that generate overreactions. Theories of underreactions include revision smoothing (Scotese, 1994), sticky information (Mankiw and Reis, 2002), noisy information (Lucas, 1972), rational inattention (Woodford, 2001; Sims, 2003), and adaptive expectations (Cagan, 1956; Nerlove, 1958), among others. In this paper, I model different theories of overreaction within a noisy information environment. As a result, the models that I consider feature some scope for underreaction. I abstract away from theories that can only generate underreactions, however, mainly because they are unable to speak to the robust evidence of overreaction among individual forecasters.

Aside from the models considered in this paper, other theories that produce overreaction in expectations include imperfect memory (Afrouzi et al., 2023; Azeredo da Silvera et al., 2020), multi-frequency forecasting under aggregation constraints (Bürgi and Ortiz, 2023), asymmetric attention (Kohlhas and Walther, 2021), and learning (Farmer et al., 2021). I abstract away from these models because they cannot be flexibly nested into the current setting.[4] These rich theories require additional structure, different environments, and, crucially, additional parameters which could pose a challenge to identification.

While I limit the set of models considered here to a rational benchmark, diagnostic expectations, overconfidence, and misspecification, I supplement my results with various robustness checks. Overall, one might view these models as the most tractable among a broader set of theories of overreaction. Such tractability is desirable, particularly when embedding non-FIRE expectations into medium- and large-scale quantitative models. Moreover, extending beyond the set of models considered here, my results imply that a successful non-FIRE theory should be able to match particular features of the data, as I will show in Section 5.

[4] I revisit learning in the subsequent section and show that it does not drive my results.

4 Model Estimation

I next estimate the four models of expectation formation outlined above. My results indicate that the misspecified expectations model is best able to fit individual and aggregate survey expectations. I begin by reviewing the data and estimation approach, and then turn to the results.

4.1 Data: U.S. Survey of Professional Forecasters

To fit the parameters governing information frictions and biases in each model, I use professional forecasts from the U.S. Survey of Professional Forecasters (SPF). The SPF is a quarterly survey managed by the Federal Reserve Bank of Philadelphia. The survey began in the fourth quarter of 1968, and provides forecasts from several forecasters across a range of horizons. The SPF has been used in many studies in the survey expectations literature (Capistran and Timmermann, 2009; Clements, 2015; Coibion and Gorodnichenko, 2012).

For my baseline results, I estimate the models using data on real GDP growth, however, I also estimate and analyze the relative performance of the models for other macroeconomic variables. I consider forecasts made from 1992Q1 to 2019Q4. I choose this later sample in order to avoid estimating the model over a period that encompasses different regimes. In addition, the survey itself has undergone several changes since 1968 including a redefining of output in 1992 from GNP to GDP. With that said, in Appendix C I estimate the models over the full sample and show that my results remain qualitatively unchanged over this longer period. Moreover, in Section 4.5 I utilize the full sample to estimate the models over rolling windows and examine their relative performance over time.

The estimation procedure that I detail below requires unbroken sequences of observations, so I keep only forecasters' longest spell of reported forecasts. Furthermore, because entry and exit from the SPF may potentially be non-random (Engelberg et al., 2011), I keep only forecasters who have a minimum spell length of eight quarters. These choices yield a sample of 77 unique forecasters from 1992Q1 to 2019Q4. In Appendix C I re-estimate the models for different macroeconomic

Table 1: SPF Summary Statistics

	Mean	Median	Std. Dev.	25%	75%
One-quarter ahead forecast error	-0.272	-0.242	1.841	-1.349	0.773
One-quarter ahead forecast revision	-0.149	-0.053	0.822	-0.487	0.241
One-quarter ahead consensus error	-0.272	-0.155	1.692	-1.334	0.568
Real-time real GDP growth	2.550	2.585	1.853	1.580	3.541

Note: The table reports summary statistics for annualized real GDP growth forecast errors and revisions as well as real-time real GDP growth. The sample of forecast errors and revisions is constructed using the SPF and spans 1992Q1-2019Q4, with 77 unique forecasters and 1,520 forecaster-quarter observations.

variables by specifying minimum spell lengths of four and 12 quarters, and find that my results are qualitatively unchanged.

Table 1 reports summary statistics for real-time real GDP growth as well as the associated forecast errors and revisions in the SPF. I follow the literature in specifying real-time forecast errors rather than forecast errors based on revised figures (Bordalo et al., 2020; Coibion and Gorodnichenko, 2015; Kohlhas and Walther, 2021). In Appendix A, I provide additional details regarding the baseline sample.

4.2 Maximum Likelihood Estimation

I estimate the models following a three-step MLE procedure. First, I collect the parameters of the fundamental variable in a vector, $\theta = (\rho_1 \;\; \rho_2 \;\; \sigma_w)$, and estimate θ via MLE using x_t as the observation which corresponds to the real-time macroeconomic time series.[5] I fix the parameters obtained in the first step according to their point estimates so that they are the same across all models. Second, I estimate the signal noise dispersion parameter, σ_v, in the rational expectations model using the panel of forecast errors, forecast revisions, and consensus forecast errors from the SPF, and calibrate it to this point estimate across the remaining three models. Finally, for each of the remaining three models, I estimate the bias parameters, α_v, φ, and $\widehat{\rho}$ via MLE once again using the SPF data.

[5]I collect these real-time variables from the U.S. SPF Error Statistics dataset.

I take this approach for two reasons. First, by setting the fundamental and information parameters to be consistent across all three biased models, I am able to evaluate overconfidence, diagnostic expectations, and misspecified expectations solely on the basis of the biases that they generate. Second, this approach makes the identification of the parameters transparent and feasible. The fundamental parameters are identified from the macroeconomic time series while the information friction and the bias parameters are identified from survey expectations. More importantly, one would not be able to jointly identify σ_v and α_v in the overconfidence model. Appendix B provides additional details on the construction of the likelihood functions and the state space formulation of the different models.

Table 2 reports my baseline results. The estimates of the parameters governing the evolution of real GDP growth are reported in panel A and imply a relatively low persistence for real GDP growth. Panel B reports the rest of the parameters. The estimate of private noise dispersion, obtained from the estimated rational expectations model, implies a signal-to-noise ratio of $\frac{\sigma_w}{\sigma_v} \approx 1.32$.

Turning to the bias parameter estimates, the estimated degree of overconfidence is essentially zero. This is consistent with the discussion between diagnostic expectations and overconfidence in Bordalo et al. (2020). Although both models can produce overreaction to new information, the overconfidence model is restricted due to the fact that there is a limit to how precise decision-makers can believe their signals to be. Moreover, the estimated degree of diagnosticity is about 0.92, implying that forecasters overreact to new information. Finally, the perceived persistence in the misspecified expectations model is about 0.75. Given that the first order autocorrelation of real GDP growth is estimated to be 0.43, the misspecified expectations model implies that forecasters overestimate the first order autocorrelation of GDP growth and neglect the partial reversal which characterizes its second lag, thereby generating overreactions.

Finally, panel C reports model selection statistics. The maximized likelihoods for each model are listed in the first row. Comparing across columns, it is evident that the estimated misspecified expectations model produces the highest likelihood across all models. The second row of panel 2 reports the corresponding Akaike information criterion (AIC) for each model. The AIC provides

Table 2: Real GDP Growth Estimates

Panel A: Fundamental Parameters					
First order autocorrelation	ρ_1	0.434			
		(0.215)			
Second order autocorrelation	ρ_2	-0.006			
		(0.001)			
Persistent innovation dispersion	σ_w	1.663			
		(0.842)			
Panel B: Information/Bias Parameters					
		(1)	(2)	(3)	(4)
		RE	OC	DE	ME
Private noise dispersion	σ_v	1.257	1.257	1.257	1.257
		(0.021)	-	-	-
Overconfidence	α_v		0.0003		
			(0.0001)		
Diagnosticity	φ			0.919	
				(0.120)	
Perceived persistence	$\widehat{\rho}$				0.753
					(0.038)
Panel C: Model Selection					
Log likelihood		-8838.5	-6925.3	-6750.1	-6709.8
AIC		17682	13858	13508	13427
Δ_i		4255	431	81	0
w_i		0.00	0.00	0.00	1.00

Note: Panel A reports estimates for the first stage MLE which estimates the parameters governing the fundamental process. Block bootstrapped standard errors (Politis and Romano, 1994) are reported in parentheses. Panel B reports the parameter estimates based on the second and third steps of the estimation procedure. Column (1) reports the rational expectations (RE) model, column (2) reports the overconfidence (OC) model, column (3) reports the diagnostic expectations (DE) model, and column (4) reports the misspecified expectations (ME) model. Bootstrapped standard errors reported in parentheses. For each model, panel C reports the maximized log likelihood, AIC, the difference between a model's AIC and smallest AIC across the models (see equation (4)), and the Akaike weight, w_i (see equation (5)).

an estimate of the expected, relative distance between a given fitted model and the unknown true model that generated the data (Akaike, 1973, 1974). In this case, comparing the AIC across models, does not change the relative rankings that one would obtain by simply comparing the maximized likelihoods.[6]

While the level of a model's AIC on its own is not interpretable, its value relative to other AIC values in the set of candidate models is informative. The results indicate that the data favor misspecified expectations since it delivers the smallest AIC (i.e., the shortest distance from the unknown true model). To compare the AIC values, I follow Burnham and Anderson (2002) who define "AIC differences" as the difference between a given AIC and the minimum AIC in the set of models:

$$\Delta_i = AIC_i \ - \min_{k \in \mathcal{M}} AIC_k, \tag{4}$$

where $\mathcal{M}$ denotes the set of the four candidate models considered, and $i \in \mathcal{M}$.

For the model with the lowest AIC, $\Delta_i = 0$. For the other models, $\Delta_i \geqslant 0$. According to Burnham and Anderson (2002), $\Delta_i \in [0, 2]$ indicates substantial support for model i given the closeness of AIC_i to AIC_{min}. On the other hand, $\Delta_i \in [4, 7]$ suggests less support for model i, while $\Delta_i > 10$ indicates that there is effectively no support for model i. Based on these rules of thumb, I conclude that the misspecified expectations model provides a substantially better fit to the data relative to the other models.

Furthermore, the AIC differences can be used to construct Akaike weights (Akaike, 1981) which are defined as,

$$w_i = \frac{\exp(-\Delta_i/2)}{\sum_{k \in \mathcal{M}} \exp(-\Delta_k/2)}. \tag{5}$$

The Akaike weight, w_i, can be interpreted as the probability that model i is the best model given the data and the set of candidate models. The final row of Table 2 reports the Akaike weights for each model and implies that misspecified expectations is the best model with probability one.

To supplement the conclusions drawn from panel C of Table 2, I also conduct a likelihood ratio

[6]The results remain unchanged when selecting models on the basis of the Bayesian information criterion.

test for non-nested models (Vuong, 1989). I compare the misspecified expectations model to each of the other models separately. Under the null hypothesis, misspecified expectations is observationally equivalent to one of the other models. When the test statistic exceeds the critical value, we reject the null and conclude that there is evidence in favor of the misspecified expectations model relative to the model against which is being compared. The results of this exercise, reported in Appendix C, accord with those from the model comparisons based on Δ_i and w_i.

4.3 Other Macroeconomic Variables

I estimate the model for 13 other macroeconomic variables in the SPF. Table 3 reports w_i for each variable. For most of these aggregates, misspecified expectations provides the best fit to the data as indicated by the many probabilities equal to one in the ME column. Diagnostic expectations, however, outperforms the other models for inflation based on the GDP deflator, real nonresidential investment, real residential investment, and industrial production. I find that these variables are generally characterized by low but positive first and second order autocorrelation coefficients, ρ_1 and ρ_2, with relatively high fundamental shock volatilities, σ_w. On the other hand, misspecified expectations appears to outperform the other models among variables for which the fundamental process exhibits partial reversals (i.e., when the first autocorrelation coefficient is positive and second autocorrelation coefficient is negative).

Furthermore, the rational expectations model outperforms the other models for housing starts and the three-month Treasury bill, with an Akaike weight of 0.63 in both cases.[7] These two variables are characterized by exceptionally high signal-to-noise ratios which, in a noisy information environment, leads rational forecasters to almost fully update their forecasts according to their observed signals therefore leaving little scope for biases to drive expectation formation. The relatively large signal-to-noise ratios may partly have to do with the fact that these macroeconomic variables are observed at a higher frequency than the quarterly frequency at which the forecasts are reported.

[7] In both cases $\Delta_{OC} = 2$ and $\Delta_{DE} = 3$, while $\Delta_{ME} = 276$ for housing starts and $\Delta_{ME} = 1355$ for industrial production.

Table 3: Model Fit Across Macroeconomic Variables

	RE	OC	DE	ME
CPI	0.00	0.00	0.00	1.00
GDP Deflator	0.00	0.00	1.00	0.00
Housing starts	0.629	0.231	0.140	0.00
Industrial production	0.00	0.004	0.996	0.00
Payroll employment	0.00	0.00	0.00	1.00
Real consumption expenditures	0.00	0.00	0.00	1.00
Real federal government spending	0.00	0.00	0.00	1.00
Real GDP	0.00	0.00	0.00	1.00
Real nonresidential investment	0.00	0.00	1.00	0.00
Real residential investment	0.00	0.00	1.00	0.00
Real state and local government spending	0.00	0.00	0.00	1.00
Unemployment rate	0.00	0.00	0.00	1.00
3-month Treasury bill	0.629	0.231	0.140	0.00
10-year government bond	0.00	0.00	0.00	1.00

Note: Each row of the table reports the Akaike weight, w_i, for a given macroeconomic variable based on equation (5). 'RE' denotes rational expectations, 'OC' denotes overconfidence, 'DE' denotes diagnostic expectations, and 'ME' denotes misspecified expectations.

4.4 Out-of-Sample Performance Across Models

In addition to providing a superior in-sample fit for a wider range of macroeconomic variables, the misspecified expectations model also provides a better out-of-sample fit relative to the other models. I assess the out-of-sample fit by estimating each model based on data from the first half of the sample, 1992Q1-2005Q4, and then using these estimates along with data from the latter half of the sample to construct and compare the out-of-sample likelihood functions.[8]

Table 4 reports out-of-sample Akaike weights for each model across a range of macroeconomic variables. Here, again, misspecified expectations provide the best fit to more of the macroeconomic variables. However, diagnostic expectations outperforms all other models when considering the GDP deflator, real consumption expenditures, and real residential and non-residential investment. Moreover, the rational expectations model outperforms the others for CPI, payroll employment, and the 3-month Treasury bill.

[8]This approach is similar to Hansen et al. (2011).

Table 4: Out of Sample Model Fit

	RE	OC	DE	ME
CPI	0.576	0.212	0.212	0.00
GDP Deflator	0.00	0.00	1.00	0.00
Housing starts	0.00	0.00	0.00	1.00
Industrial production	0.00	0.00	0.00	1.00
Payroll employment	1.00	0.00	0.00	0.00
Real consumption expenditures	0.00	0.00	1.00	0.00
Real federal government spending	0.00	0.00	0.00	1.00
Real GDP	0.00	0.00	0.00	1.00
Real nonresidential investment	0.00	0.00	1.00	0.00
Real residential investment	0.00	0.00	1.00	0.00
Real state and local government spending	0.00	0.00	0.00	1.00
Unemployment rate	0.00	0.00	0.00	1.00
3-month Treasury bill	1.00	0.00	0.00	0.00
10-year government bond	0.00	0.00	0.00	1.00

Note: Each row of the table reports the out-of-sample Akaike weight, w_i, for a given macroeconomic variable based on equation (5). 'RE' denotes rational expectations, 'OC' denotes overconfidence, 'DE' denotes diagnostic expectations, and 'ME' denotes misspecified expectations.

These out-of-sample results might lead one to wonder whether forecasts could be improved in real time. There is a longstanding literature which finds that statistical models are often unable to outperform survey forecasts out of sample (Ang et al., 2007; Faust and Wright, 2013). Based on my results, one cannot necessarily conclude that forecasters could improve their forecasts out of sample because I include the contemporaneous individual and consensus forecast errors in my set of MLE observations, which are not known to forecasters at the time in which they issue their forecasts.

For a similar reason, my results are not inconsistent with Eva and Winkler (2023) which finds that error predictability regressions, such as the Coibion and Gorodnichenko (2015) regression, perform relatively poorly out of sample. The objective of this paper is to determine which non-FIRE model provides the best fit to the data. To answer this question, I estimate the different models using information available to the econometrician, which is not the same as the information available to forecasters in real time.

4.5 Learning About the Data Generating Process

Learning, which is a compelling and realistic feature of real-time forecasting, is another theory which can explain predictable forecast errors (Evans and Honkapohja, 2001; Farmer et al., 2021). Though Farmer et al. (2021) emphasize that initial beliefs rather than model misspecification are ultimately responsible for generating overreactions in their learning environment, in this section I explore whether learning may nonetheless bias my results in favor of misspecified expectations. Focusing on real GDP growth, I first examine the evolution of the Akaike weights for each model over rolling windows. Second, because the rolling windows approach does not account for composition, I split my sample by forecaster "age," estimate the models for each sub-sample, and then compare the Akaike weights across models.

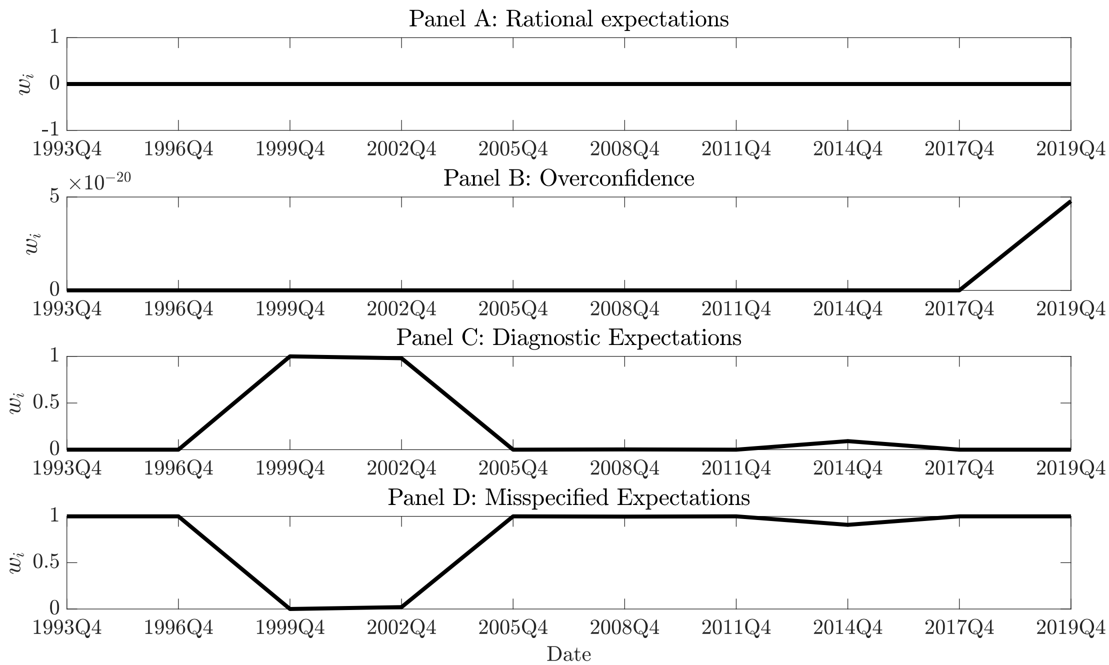
Rolling Windows Approach

If learning were to bias my results in favor of misspecified expectations, then it would likely to do so by providing a better fit to the data in the early part of the sample, during which time forecasters presumably employ simple forecasting models, and then providing a worse fit as forecasters learn about the underlying process.

To examine whether this is the case, I construct 25-year-long sub-samples with a three-year rolling window using the full pre-COVID-19 sample of the data (1968Q4-2019Q4). I repeat the MLE procedure and compute the AIC for each model from each sub-sample. Figure 1 plots the Akaike weight, w_i for each model over time. To reiterate, w_i can be interpreted as the probability that model i is the best model given the data and the set of models considered. Values closer to one indicate greater evidence for a given model while values closer to zero imply less evidence for the model.

Panel A plots w_i for the rational expectations model over time. Here, we see that there is little evidence for the rational model as its Akaike weight remains at zero since its AIC is substantially larger than the other models. Panel B plots w_i for the overconfidence model. In the earlier sub-samples, I find that the AIC for the overconfidence model was much higher than the other non-rational models.

Figure 1: Akaike Weights Across Time (Real GDP Growth)



Note: The figure plots the Akaike weight, w_i, for each model over rolling windows based on equation (5). The dates on the horizontal axis correspond to the end date of each 25-year window.

From the late 1990s, however, the relative AIC of the overconfidence model declined such that the Akaike weight registered a slight increase by the end of the sample. Still, based on the magnitudes reported on the vertical axis, we find that the Akaike weight for the overconfidence model remained at essentially zero over time.

Panel C plots w_i for the diagnostic expectations model. Diagnostic expectations outperforms misspecified expectations for two of the ten sub-periods. In the 1975Q1-1999Q4 sample period, $AIC_{ME} - AIC_{DE} \approx 16$ while in the 1978Q1-2002Q4 period, $AIC_{ME} - AIC_{DE} \approx 7$. Inspecting the estimated parameters for real GDP growth, I find that these two sub-periods are associated with significant declines in the estimated first order autocorrelation coefficient, declining from about 0.50 to 0.35. The second order autocorrelation coefficient, on the other hand, remained essentially unchanged.

Considering the evidence that overreactions are stronger for less persistent series (Afrouzi et al., 2023; Bordalo et al., 2020), it is possible that diagnostic expectations provides a better fit to these sub-samples since the degree of diagnosticity is unbounded while forecasters under misspecified expectations assume that the perceived AR(1) process is stationary thereby limiting the scope for overextrapolation. Consistent with this intuition, we observe a doubling of the diagnosticity parameter from about 1.2 to 2.4 from the 1972Q1-1996Q4 sample to the 1975Q1-1999Q4 sample, with little change in the misspecified expectation model's perceived persistence since it is already estimated to be roughly 0.95 in the 1972Q1-1996Q4 sample.

Though I find that diagnostic expectations offers a better fit relative to misspecified expectations for certain periods of my sample, the patterns observed in Figure 1 do not indicate that learning over time favors misspecified expectations.

Experience as a Proxy for Learning

The conclusions drawn from the exercise in the previous section rely on the assumption that we observe the same forecasters over time. In reality, this is not the case. Because entry and exit in the SPF can affect when a forecaster begins the learning process, I next explore whether learning affects the relative rankings of the candidate models by comparing "experienced" and "inexperienced" forecasters, where I proxy experience with forecaster age (tenure) in the sample.[9] I drop the eight-quarter spell length sample restriction and define inexperienced forecasters as those whose respective ages fall below the unconditional sample median while experienced forecasters are those whose ages reside above the median.[10]

Panel A of Table 5 reports Δ_i and panel B reports the Akaike weights for each model. The results imply that misspecified expectations outperforms the other models across the two sub-samples. This finding runs counter to the hypothesis that learning would favor misspecified expectations among inexperienced forecasters since they would be more likely to employ simple forecasting models. In

[9] Because the SPF is anonymized, I am unable to determine the "experience" of a given forecaster if she has a history of participating in other surveys as well. However, in this case my approach would tend to label a more seasoned forecaster as an "inexperienced" forecaster, which would work against misspecified expectations if more seasoned forecasters are less likely to use simple forecasting models.

[10] I have also completed this exercise by identifying experienced and inexperienced forecasters based on their *maximum* age attained in the sample rather than their current age in the sample. The results remain unchanged when applying this alternative definition.

Table 5: Model Fit by Forecaster Experience (Real GDP Growth)

	RE	OC	DE	ME
Panel A: Δ_i				
Below median age	4966	674	52	0
Above median age	5257	791	67	0
Panel B: w_i				
Below median age	0.0	0.0	0.0	1.0
Above median age	0.0	0.0	0.0	1.0

Note: Each row of panel A reports Δ_i based on equation (4). Each row of panel B reports the Akaike weight based on equation (5). 'RE' denotes rational expectations, 'OC' denotes overconfidence, 'DE' denotes diagnostic expectations, and 'ME' denotes misspecified expectations.

fact, based on panel A, rational expectations, overconfidence, and diagnostic expectations provide a slightly better fit to the data relative to misspecified expectations among inexperienced forecasters. Overall, while learning may indeed be happening in reality, it does not appear drive the relative model rankings to favor misspecified expectations.

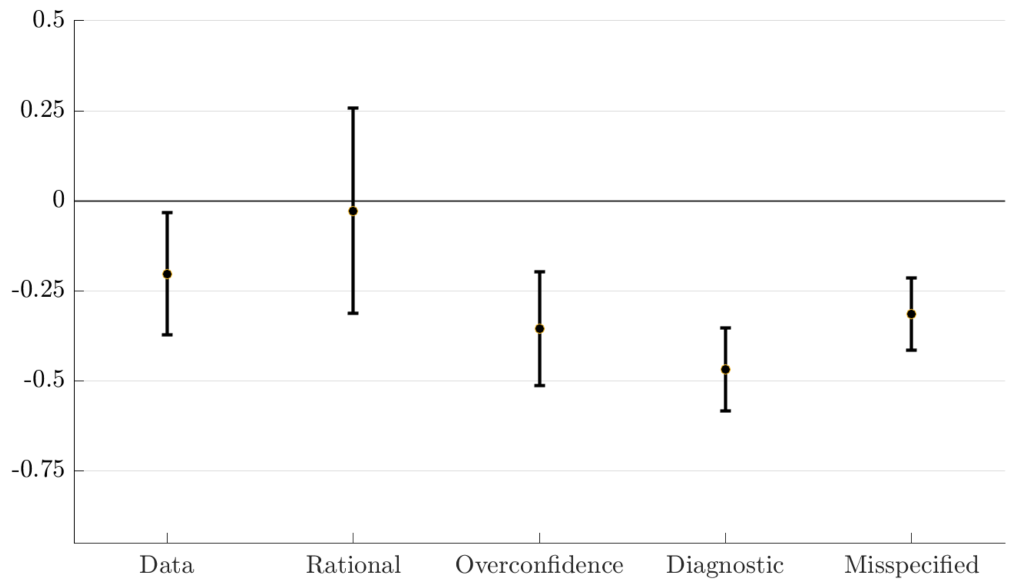
5 Misspecified Expectations Fits Important Moments

Of the different models considered, misspecified expectations most often outperforms the others. To understand why this might be the case, I focus on my baseline real GDP growth results and examine five features of the data that have been widely studied in the literature: unconditional overreaction, unconditional underreaction, overshooting (i.e., conditional over- and underreaction), persistent disagreement, and updating weights.

5.1 Overreaction

Bordalo et al. (2020) provide evidence of overreaction in macroeconomic expectations by running the Coibion and Gorodnichenko (2015) errors-on-revisions regression at the forecaster level. This

Figure 2: Overreaction in One Quarter Ahead Expectations



Note: The figure plots the 95% confidence intervals for the individual-level errors-on-revisions regression coefficient in the data as well as the four different models. The model-based coefficients are obtained by simulating 2,000 panels of data for each of the four models. Each model is simulated based on the estimates reported in Table 2.

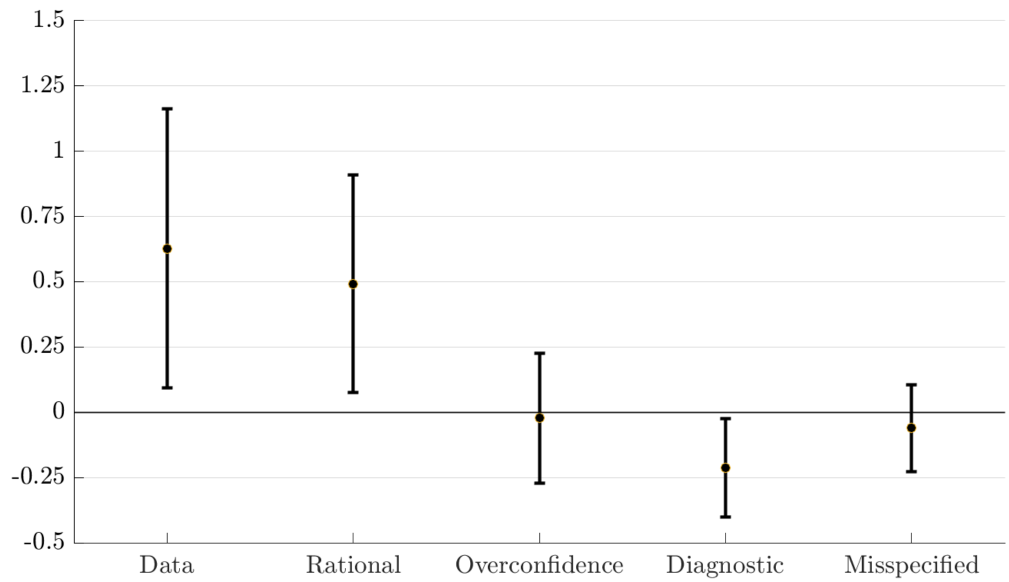
testable prediction has been studied extensively in the literature. The regression specification is:

$$x_{t+3} - \widehat{x}^i_{t+3|t} = \beta_0 + \beta_1[\widehat{x}^i_{t+3|t} - \widehat{x}^i_{t+3|t-1}] + \varepsilon^i_t. \tag{6}$$

Estimating $\widehat{\beta}_1 < 0$ implies that an upward ex-ante revision predicts a more negative ex-post error. This negative relation is interpreted as evidence of overreaction to new information.

Figure 2 displays 95% confidence intervals for the β_1 coefficient in (6) simulated across each of the models, along with the empirical estimate from the data. As expected, the rational expectations model cannot generate overreactions since forecast errors are orthogonal to anything residing in the forecaster's information set which includes the contemporaneous revision. On the other hand, the other three theories are able to generate negative β_1 coefficients. Furthermore, these simulated coefficients reside within the 95% confidence interval of the empirical estimate.

Figure 3: Consensus Errors-on-Revisions Regression



Note: The figure plots the 95% confidence intervals for the consensus-level errors-on-revisions regression coefficient in the data as well as the four different models. The model-based coefficients are obtained by simulating 2,000 panels of data for each of the four models. Each model is simulated based on the estimates reported in Table 2.

5.2 Underreaction

Although individual professional forecasts exhibit overreactions, it is well known that consensus expectations exhibit underreactions (Coibion and Gorodnichenko, 2015). In other words, running (6) at the aggregate level,

$$x_{t+3} - \widehat{x}_{t+3|t} = \alpha_0 + \alpha_1[\widehat{x}_{t+3|t} - \widehat{x}_{t+3|t-1}] + \epsilon_t, \tag{7}$$

generally delivers an estimate $\widehat{\alpha}_1 > 0$.

Figure 3 plots the consensus-level analogs to Figure 2. Based on the point estimates, we see that the rational expectations model best matches the data on consensus-level underreactions. This is because the rational model does not feature any scope for overreaction, so the simulated OLS coefficient reflects only the information friction arising from the noisy information environment.

Because the other models feature some overreaction at the individual level, and since this overreaction is quantitatively significant, none of the three estimated non-rational models successfully matches this moment. The 95% confidence intervals for the overconfidence and misspecified expectations models, however, include positive and negative values and overlap with the empirical 95% confidence interval. However, the point estimates for these models are negative.

The fact that the three biased models produce point estimates that are well below the empirical coefficient could suggest, based on the full information approach taken here, that such a moment is not particularly important to match. This would be consistent with Bianchi et al. (2022) which draws a similar conclusion in a data-rich context. Alternatively, the negative point estimates obtained across the different non-rational models could suggest that my setting abstracts away from other potential sources of inertia in expectations such as endogenous public signals (Broer and Kohlhas, 2022). A third potentially important explanation is that I estimate the biases in the non-rational models only after calibrating the information friction according to the estimates obtained in the rational model. This approach could deliver relatively less information rigidity than would be obtained by jointly estimating σ_v alongside the bias parameters for each model.[11]

I next turn to examining dynamics, where I find that misspecified expectations is able to generate a sign switch in the impulse response of the aggregate forecast error followed by a gradual convergence to zero, consistent with the data.

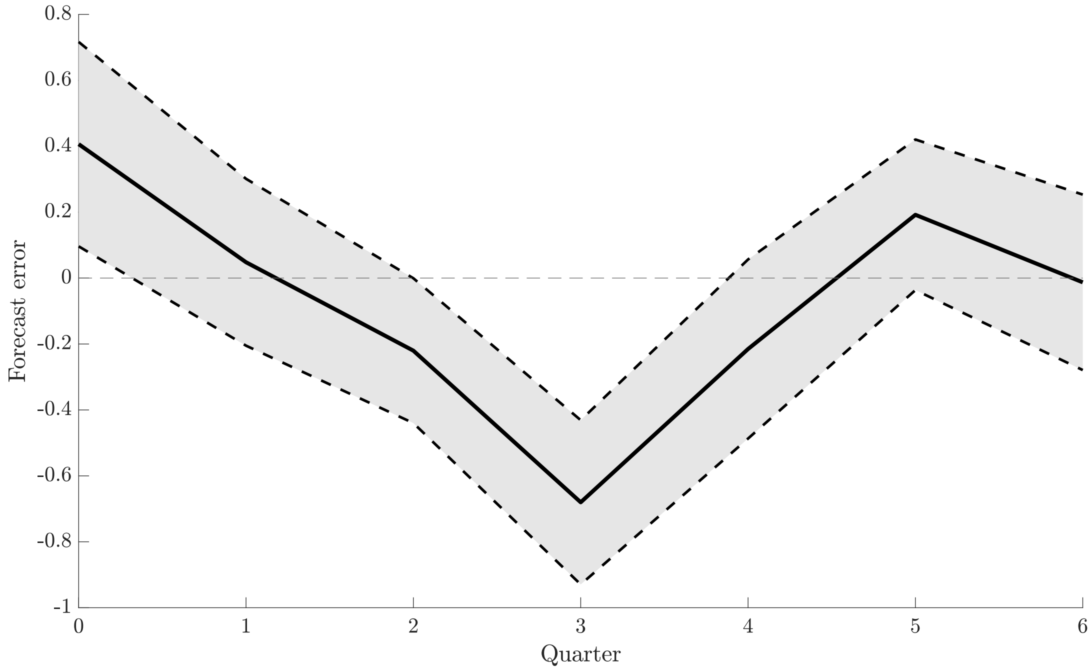
5.3 Overshooting

Angeletos et al. (2020b) documents evidence of delayed overshooting in the medium term, a form of conditional overreaction. This phenomenon can be observed by inspecting the impulse response of the consensus forecast error to a shock. If the impulse response function switches signs, then beliefs are said to exhibit overshooting.

Since Angeletos et al. (2020b) provides evidence of delayed overshooting for inflation and un-

[11] As discussed above, I choose to first estimate σ_v in the rational model and then calibrate this parameter in the other models to preserve uniformity in the estimation approach for the non-rational models because, otherwise, one would not be able to jointly identify σ_v and α_v in the overconfidence model.

Figure 4: Delayed Overshooting in Real GDP Growth Forecasts



Note: The figure plots the estimated impulse response of the one-quarter ahead consensus forecast error in my sample to an identified business cycle shock from Angeletos et al. (2020a) according to equation (8). Newey-West standard errors are specified and the shaded area reflects 68% confidence intervals.

employment rate forecast errors, I document evidence of delayed overshooting for one-quarter ahead real GDP growth forecast errors in my sample. To do so, I follow Angeletos et al. (2020b) by collecting the identified business cycle shocks from Angeletos et al. (2020a).[12] I estimate the impulse response of the one-quarter ahead consensus forecast error in my sample to a positive identified business cycle shock via local projections (Jordà, 2005). I run the following regression across various horizons, h,

$$\text{Error}_{t+h} = \beta_{0,h} + \beta_{1,h}\text{Shock}_t + \gamma_h' \mathbf{X}_{t-1} + \epsilon_{t+h}, \tag{8}$$

where I specify four lags of the consensus forecast, realized real-time GDP growth, and the identified shock as controls.

Figure 4 plots the response of the one-quarter ahead average forecast error to a positive demand

[12] I utilize the shock that they regard as the “main business cycle shock” which reflects a demand shock.

shock. We see that the forecast error is positive on impact, reflecting the upward surprise in real GDP growth. Thereafter, the one-quarter forecast error declines and begins turning negative two quarters following the shock before reverting back toward zero. The sign-switch in the impulse response of the forecast error observed here is evidence of delayed overshooting.

In principle, all three biased models can reproduce these dynamics given that each model features imperfect information, the source of initial underreaction, and overreactive biases, the source of delayed overreaction. Quantitatively, however, I find that only diagnostic expectations and misspecified expectations produce overshooting, without much initial underreaction, in response to an aggregate shock.

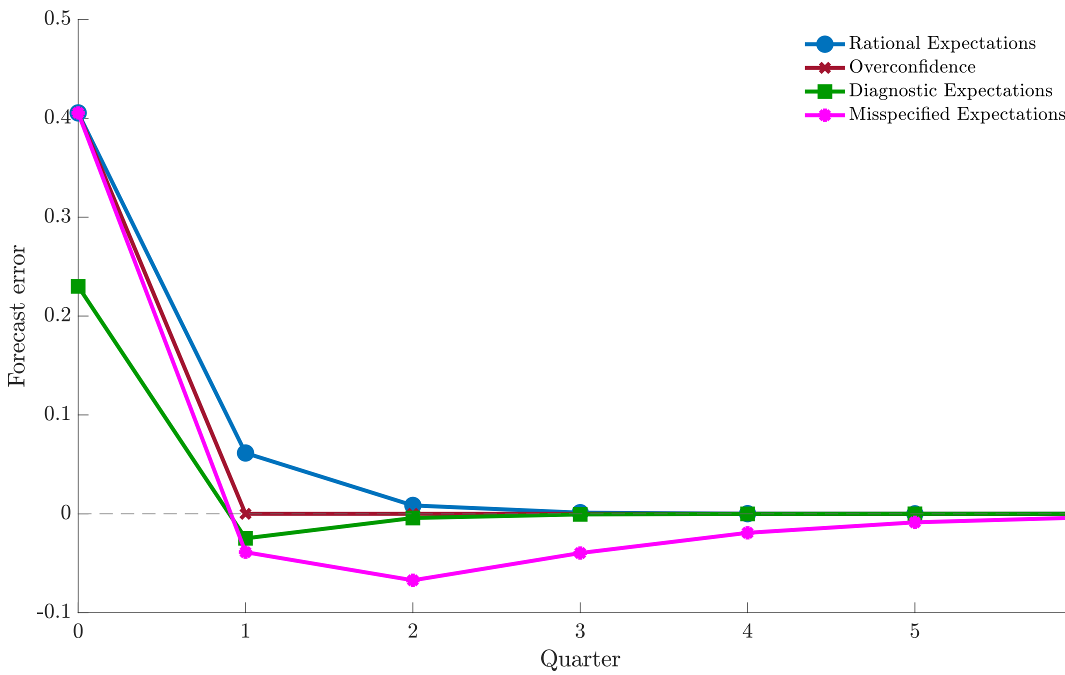
Figure 5 plots the response of the simulated consensus forecast error to a positive fundamental shock, $w_t > 0$, in each of the four models. I scale the shock to produce a 0.40 percentage point increase in the forecast error on impact, which coincides with the empirical estimate in Figure 4.

The consensus forecast error in the rational expectations model gradually converges to zero, with the evolution of the forecast error reflecting the rate of learning in the model. The consensus error in the overconfidence model features a stronger convergence to zero, but no sign switch. The diagnostic expectations model delivers a sign switch in period one, after which the consensus forecast error converges to zero from below relatively quickly. Finally, the misspecified expectations model also delivers a sign switch, with the extent of overshooting peaking two quarters following the shock, after which point the aggregate forecast error converges gradually to zero from below.

Angeletos et al. (2020b) note that overextrapolation is necessary to replicate delayed overshooting dynamics in the medium term. Misspecified expectations as modeled here is able to reproduce this pattern better than the other models. Intuitively, this has to do with the fact that forecasters operating under misspecified expectations exhibit overreaction because they mis-perceive the underlying data generating process. As a result, forecast errors are generally longer-lived under misspecified expectations, and, when coupled with a relatively large $\widehat{\rho}$, forecasters overextrapolate which generates delayed overshooting.

The relatively more persistent forecast errors exhibited by misspecified expectations can also

Figure 5: Simulated Overshooting



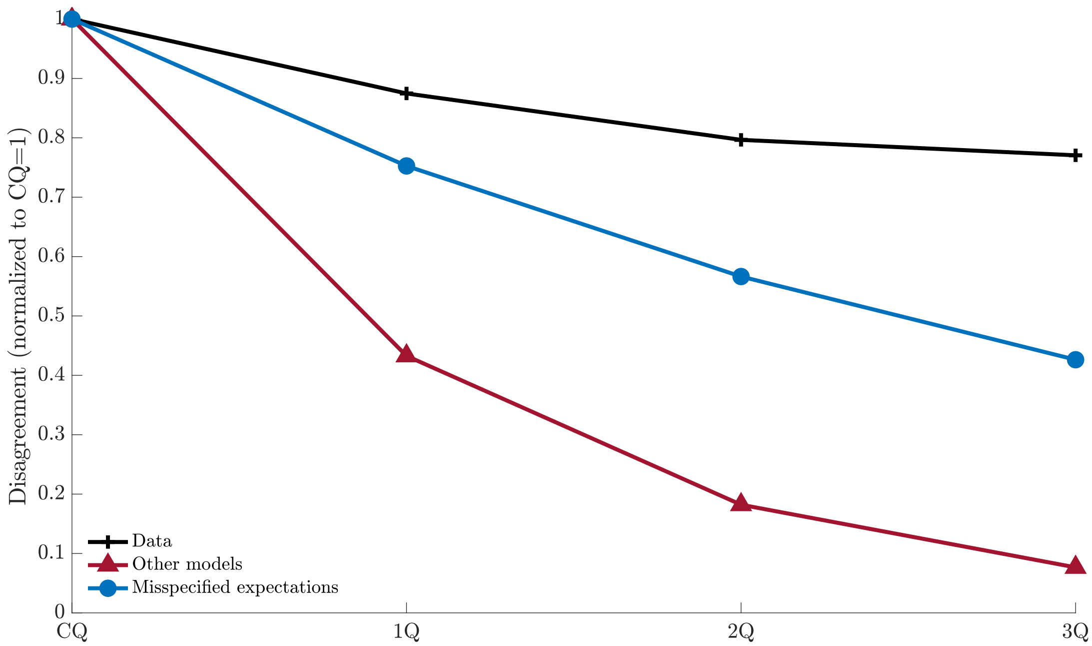
Note: The figure plots the simulated impulse response of the one-quarter ahead consensus forecast error to a positive shock to w_t scaled to generate a 0.40 percentage point increase in the one-quarter ahead consensus forecast error in the rational expectations model. The impulse responses are obtained by simulating 2,000 panels of forecasts for each of the four models. Each model is simulated based on the estimates reported in Table 2.

help explain why misspecified expectations outperforms the other models when examining forecaster disagreement, which I discuss in the next section.

5.4 Persistent Disagreement

Persistent disagreement is another well-documented puzzle in survey expectations (Andrade et al., 2016; Giacomini et al., 2020; Patton and Timmermann, 2010; Rich and Tracy, 2020). Misspecified expectations outperforms the other models in its ability to generate persistent disagreement across forecasters because it allows forecasters to misstate the degree of mean reversion of the data generating process. Figure 6 plots empirical and model-based forecaster disagreement, which is defined here as the dispersion of forecasts, at different horizons. Each line is normalized to equal one at the current-quarter horizon which facilitates the visual comparison of the persistence in disagree-

Figure 6: Persistence in Disagreement Across Horizons



Note: The figure plots a measure of forecaster disagreement at different horizons, where disagreement is defined as the unconditional standard deviation of forecasts. The values of disagreement are normalized so that disagreement in the current quarter is equal to one for all models. 'CQ' standard for 'current quarter', '1Q' stands for one-quarter ahead, '2Q' stands for two-quarters ahead, and '3Q' stands for three-quarters ahead. The model-based estimates are obtained by simulating 2,000 panels of data for each of the four models. Each model is simulated based on the estimates reported in Table 2.

ment across models. The black line reflects the data which confirms that disagreement is persistent across horizons. The blue line denotes the misspecified expectations model. The red line reflects the rational expectations, overconfidence, and diagnostic expectations models, all of which imply the same evolution in the dispersion of forecasts across horizons.

Because the rational expectations, overconfidence, and diagnostic expectations models assume that forecasters know that the data are generated by an AR(2) process, the dispersion of forecasts across horizons will evolve identically according to the AR(2) process. While each of these models implies a different *level* of disagreement, they all imply the same *persistence* in disagreement.[13]

[13]When comparing the levels of disagreement, the models have non-overlapping strengths. Diagnostic expectations provides the best fit to the level of disagreement in the current quarter while misspecified expectations provides the best fit to the level of disagreement from the one-quarter ahead horizon onward.

On the other hand, because forecasters in the misspecified expectations model mis-perceive the underlying process, disagreement based on news received today is longer lived. The misspecified expectations model will therefore exhibit a higher dispersion of forecasts over longer horizons relative to the other models.

5.5 Updating Rules

The forecasters populating the candidate models considered here are fundamentally Bayesian. In this section, I therefore compute the models' implied Kalman gains and compare the weights placed on the private signals across each model and to the data.

Based on the identical information structure assumed across all of the models, forecasters update their predictions as follows:[14]

$$\widehat{x}^i_{t|t} = \widehat{x}^i_{t|t-1} + \widehat{\kappa}_1(y^i_t - \widehat{x}^i_{t|t-1}).$$

This updating equation implies that forecasters place some weight on news, κ_1, and their prior, $(1-\kappa_1)$,

$$\widehat{x}^i_{t|t} = (1-\widehat{\kappa}_1)\widehat{x}^i_{t|t-1} + \widehat{\kappa}_1 x_t + \widehat{\kappa}_1 v^i_t, \tag{9}$$

Based on this updating rule, I run the following regression to estimate the updating weight,

$$\widehat{x}^i_{t|t} = \beta_0 + \beta_1\widehat{x}^i_{t|t-1} + \beta_2 x_t + \omega_{it}. \tag{10}$$

This specification projects the current-quarter forecast on the lagged one-quarter ahead forecast and the current realization of the macroeconomic variable. I report empirical and model-based estimates of regression (10) in panel A of Table 6. Based on column (1), there is strong evidence in the data for the updating rule implied by the models considered here. The sum of the estimated

[14]Under rational expectations, $\widehat{\kappa}_1$ coincides with the optimal Kalman gain, κ_1. Under overconfidence and misspecified expectations, $\widehat{\kappa}_1$ is the Kalman gain implied by the perceived signal precision, α_v, and perceived persistence, $\widehat{\rho}$, respectively. Under diagnostic expectations, $\widehat{\kappa}_1 = \kappa_1(1+\varphi)$.

Table 6: Updating Rule Regressions

	(1)	(2)	(3)	(4)	(5)
Panel A: No time fixed effects					
	Data	RE	OC	DE	ME
Lagged one-quarter ahead forecast	0.665***	0.349***	0.000	-0.136***	0.297***
	(0.056)	(0.011)	(0.012)	(0.012)	(0.006)
Current-quarter realization	0.348***	0.652***	1.000***	1.251***	0.704***
	(0.078)	(0.004)	(0.007)	(0.009)	(0.005)
Panel B: Time fixed effects					
	Data	RE	OC	DE	ME
Lagged one-quarter ahead forecast	0.421***	0.346***	-0.001	-0.130***	0.296***
	(0.036)	(0.022)	(0.022)	(0.022)	(0.012)

Note: Each column of the table reports regression results based on (10). Panel A reports the regression results without time fixed effects while panel B specifies time fixed effects. Column (1) reports the empirical regression results, for which I specify Driscoll and Kraay (1998) standard errors in parentheses. The remaining columns specify simulated regression coefficients and standard deviations (in parentheses) based on 2,000 simulated panels of forecasters. 'RE' denotes the rational expectations model, 'DE' denotes the diagnostic expectations model, 'OC' denotes the overconfidence model, and 'ME' denotes the misspecified expectations model. * denotes 10% significance, ** denotes 5% significance, and *** denotes 1% significance.

coefficients is close to one, and they imply that forecasters place a weight of about 0.35 on new information when updating their expectations.[15]

The other columns of panel A report the model-based regression results. Overall, the rational expectations model provides a better fit to the data, though among the non-rational models, misspecified expectations provides the relatively better fit. We see, however, that forecasters in the estimated models all place more weight on new information than their priors. The Kalman gain in the rational expectations model is 0.65. Among the non-rational models, the Kalman gain is greater than or equal to one under overconfidence and diagnostic expectations, whereas it is approximately 0.70 in the misspecified expectations model.

The set up that I consider assumes that forecasters only have access to a contemporaneous private signal. Notably, I abstract away from public information. In reality, however, forecasters also

[15] A test that the sum of the coefficients is different from one delivers a p-value of 0.76, leading to a failure to reject the null that they sum to one.

observe public signals. A better way to compare the empirical estimates to the model-based regression results would therefore require controlling for common signals by specifying time fixed effects in regression (10). I do so in panel B of Table 6.

By controlling for unobserved time variation, the coefficient in front of the current-quarter realization cannot be estimated since it is absorbed in the time fixed effects. I therefore only report estimates for β_1. Column (1) reports the empirical estimates under this alternative specification which reveals a meaningful decline in the magnitude of the estimate in front of the lagged one-quarter ahead forecast relative to panel A. These results now indicate that forecasters place more importance on new information rather than their priors when updating their expectations. While the empirical results in panel B are different from panel A, the simulated model-based results are nearly identical as expected. As a result, we once again conclude that among the non-rational models, misspecified expectations better matches updating rules in the data.

6 Conclusion

At present, a host of non-FIRE theories exist in the literature. As mentioned in Reis (2020), however, there is little agreement on a suitable non-FIRE benchmark. This paper offers a partial answer to this question by showing that misspecified expectations outperforms other non-FIRE models for a variety of macroeconomic aggregates for which professional forecasters issue predictions. Misspecified expectations can reproduce patterns of unconditional and conditional overreaction, persistent disagreement, and it delivers empirically similar updating rules. Misspecified expectations can be motivated from behavioral and non-behavioral foundations alike. Embedding the misspecification described here into a quantitative model merely requires introducing two parameters into an otherwise standard model: the second order autocorrelation coefficient for the underlying process, ρ_2, and a misspecification parameter, $\widehat{\rho}$. A promising avenue for future research could be to examine whether there is evidence favoring misspecified expectations in other settings.

References

Afrouzi, Hassan, Spencer Kwon, Augustin Landier, Yueran Ma, and David Thesmar (2023), "Overreaction in Expectations: Evidence and Theory." *The Quarterly Journal of Economics*, 138.

Akaike, Hirotugu (1973), "Information Theory and an Extension of the Maximum Likelihood Principle–In: Second International Symposium on Information Theory (Eds) BN Petrov, F." *Csaki. BNPBF Csaki Budapest: Academiai Kiado*, 267–281.

Akaike, Hirotugu (1974), "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control*, 716–723.

Akaike, Hirotugu (1981), "Likelihood of a Model and Information Criteria." *Journal of Econometrics*, 16, 3–14.

Andrade, Philippe, Richard K. Crump, Stefano Eusepi, and Emanuel Moench (2016), "Fundamental Disagreement." *Journal of Monetary Economics*, 83, 105–128.

Ang, Andrew, Geert Bekaert, and Min Wei (2007), "Do macro variables, asset markets, or surveys forecast inflation better?" *Journal of Monetary Economics*, 4.

Angeletos, George Marios, Fabrice Collard, and Harris Dellas (2020a), "Business-Cycle Anatomy." *American Economic Review*, 110, 3030–3070.

Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry (2020b), "Imperfect Macroeconomic Expectations; Evidence and Theory." *NBER Macroeconomics Annual 2020*, 35.

Azeredo da Silvera, Rava, Yeji Sung, and Michael Woodford (2020), "Optimally Imprecise Memory and Biased Forecasts." NBER Working Paper 28075.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny (1998), "A Model of Investor Sentiment." *Journal of Financial Economics*, 49.

Bianchi, Francesco, Sydney C. Ludvigson, and Sai Ma (2022), "Belief Distortions and Macroeconomic Fluctuations." *American Economic Review*, 112.

Binder, Carola, Tucker S. McElroy, and Xuguang S. Sheng (2022), "The Term Structure of Uncertainty: New Evidence From Survey Expectations." *Journal of Money, Credit, and Banking*, 54, 39–71.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (2020), "Overreaction in Macroeconomic Expectations." *American Economic Review*, 110, 2748–2782.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2019), "Diagnostic Expectations and Stock Returns." *Journal of Finance*, 74, 2839–2874.

Branch, William A. and George W. Evans (2006), "Intrinsic heterogeneity in expectation formation." *Journal of Economic Theory*, 127, 264–295.

Brock, William A. and Cars H. Hommes (1997), "A Rational Route to Randomness." *Econometrica*, 65, 1059–1095.

Broer, Tobias and Alexandre Kohlhas (2022), "Forecaster (Mis-)Behavior." *The Review of Economics and Statistics*, 1–45.

Bürgi, Constantin and Julio Ortiz (2023), "Overreaction Through Anchoring." Working Paper.

Burnham, Kenneth P. and David R. Anderson (2002), *Model Selection and Multimodel Inference*. Springer New York, NY.

Cagan, Phillip (1956), "The Monetary Dynamics of Hyperinflation." 25–117.

Capistran, Carlos and Allan Timmermann (2009), "Disagreement and Biases in Inflation Expectations." *Journal of Money, Credit and Banking*, 41, 365–396.

Carroll, Christopher D (2003), "Macroeconomic expectations of households and professional forecasters." *The Quarterly Journal of Economics*, 118, 269–298.

Clements, Michael P. (2015), “Are Professional Macroeconomic Forecasters Able to Do Better Than Forecasting Trends?” *Journal of Money, Credit and Banking*, 47, 349–381.

Coibion, Olivier and Yuriy Gorodnichenko (2012), “What Can Survey Forecasts Tell Us About Information Rigidities?” *Journal of Political Economy*, 120, 116–159.

Coibion, Olivier and Yuriy Gorodnichenko (2015), “Information rigidity and the expectations formation process: A simple framework and new facts.” *American Economic Review*, 109, 465–490.

Cornand, Camille and Paul Hubert (2022), “Information frictions across various types of inflation expectations.” *European Economic Review*, 146.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam (1998), “Investor Psychology and Security Market Under- and Overreactions.” *Journal of Finance*, 53, 1839–1885.

Driscoll, John C and Aart C Kraay (1998), “Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data.” *The Review of Economics and Statistics*, 80, 549–560.

Engelberg, Joseph, Charles F. Manski, and Jared Williams (2011), “Assessing the Temporal Varioation of Macroeconomic Forecasts by a Panel of Changing Composition.” *Journal of Applied Econometrics*, 26, 1059–1078.

Eva, Kenneth and Fabian Winkler (2023), “A Comprehensive Empirical Evaluation of Biases in Expectation Formation.” Finance and Economics Discussion Series 2023-042.

Evans, G.W. and S. Honkapohja (2001), *Learning and Expectations in Macroeconomics*. Princeton University Press.

Farmer, Leland, Emi Nakamura, and Jon Steinsson (2021), “Learning About the Long Run.” Working Paper.

Faust, Jon and Jonathan H. Wright (2013), “Chapter 1 Forecasting Inflation.” *Handbook of Economic Forecasting*, 2, 2–56.

Fuhrer, Jeff (2018), “Intrinsic Expectations Persistence: Evidence from Professional and Household Survey Expectations.” Working Paper.

Fuster, Andreas, Benjamin Hebert, and David Laibson (2012), “Investment Dynamics with Natural Expectations.” *International Journal of Central Banking*, 8, 243–265.

Fuster, Andreas, David Laibson, and Brock Mendel (2010), “Natural Expectations, Macroeocnomic Fluctuations.” *Journal of Economic Perspectives*, 24, 67–84.

Gabaix, Xavier (2019), “Behavioral Inattention.” *Handbook of Behavioral Economics: Applications and Foundations 1*, 2, 261–343.

Gennaioli, Nicola and Andrei Shleifer (2010), “What Comes to Mind?” *The Quarterly Journal of Economics*, 125, 1399–1433.

Giacomini, Raffaella, Vasiliki Skreta, and Javier Turen (2020), “Heterogeneity, Inattention, and Bayesian Updates.” *American Economic Journal: Macroeconomics*, 12, 282–309.

Hansen, Peter Reinhard, Zhuo Huang, and Howan Shek (2011), “Realized GARCH: A Joint Model For Returns and Realized Measures of Volatility.” *Journal of Applied Econometrics*, 27.

Jordà, Òscar (2005), “Estimation and Inference of Impulse Responses by Local Projections.” *American Economic Review*, 95, 161–182.

Kahneman, Daniel and Amos Tversky (1972), “Subjective probability: a judgment of representativeness.” *Cognitive Psychology*, 3, 430–454.

Kent, Daniel and David Hirshleifer (2015), “Overconfident Investors, Predictbable Returns, and Excessive trading.” *Journal of Economic Perspectives*, 29, 61–88.

Kent, Daniel, David Hirshleifer, and Avanidhar Subrahmanyam (1998), “Investor Psychology and Security Market Under- and Overreactions.” *Journal of Finance*, 53, 1839–1885.

Kent, Daniel, David Hirshleifer, and Avanidhar Subrahmanyam (2001), "Overconfidence, Arbitrage, and Equilibrium Asset Pricing." *Journal of Finance*, 56, 921–965.

Kohlhas, Alexandre N. and Answer Walther (2021), "Asymmetric Attention." *American Economic Review*, 111, 2879–2925.

Kucinskas, Simas and Florian Peters (2022), "Measuring Under- and Overreaction in Expectation Formation." *Review of Economics and Statistics*.

Lucas, Jr., Robert E. (1972), "Expectations and the Neutrality of Money." *Journal of Economic Theory*, 4, 103–124.

Mankiw, Gregory and Ricardo Reis (2002), "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics*, 117, 1295–1328.

Molavi, Pooya (2022), "Simple Models and Biased Forecasts." Working Paper.

Moore, Don and Paul J. Healy (2008), "The Trouble with Overconfidence." *Psychological Review*, 115, 502–517.

Nerlove, Marc (1958), "Adaptive Expectations and Cobweb Phenomena." *Quarterly Journal of Economics*, 72, 227–240.

Patton, Andrew J. and Allan Timmermann (2010), "Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion." *Journal of Monetary Economics*, 57, 803–820.

Pfajfar, Damjan (2013), "Formation of rationally heterogeneous expectations." *Journal of Economic Dynamics and Control*, 37, 1434–1452.

Politis, Dimitris N. and Joseph P. Romano (1994), "The Stationary Bootstrap." *Journal of the American Statistical Association*, 89, 1303–1313.

Reis, Ricardo (2020), “Comment on ‘Imperfect Expectations: Theory and Evidence’.” *NBER Macroeconomics Annual 2020*, 35, 99–111.

Rich, Robert and Joseph Tracy (2020), “A Closer Look at the Behavior of Uncertainty and Disagreement: Micro Evidence from the Euro Area.” *Journal of Money, Credit, and Banking*, 53, 233–253.

Rozsypal, Filip and Kathrin Schlafmann (2023), “Overpersistence Bias in Individual Income Expectations and its Aggregate Implications.” *American Economic Journal: Macroeconomics (Forthcoming)*.

Ryngaert, Jane (2017), “What Do (and Don't) Forecasters Know About U.S. Inflation?” Working Paper.

Scotese, Carol A. (1994), “Forecast Smoothing and the Optimal Under-utilization of Information at the Federal Reserve.” *Journal of Macroeconomics*, 16, 653–670.

Sims, Christopher (2003), “The Implications of Rational Inattention.” *Journal of Monetary Economics*, 50, 665–690.

Vuong, Quang H. (1989), “Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses.” *Econometrica*, 57, 307–333.

Woodford, Michael (2001), “Imperfect Common Knowledge and the Effects of Monetary Policy.” *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund Phelps*.

Appendix A SPF Sample

For my analysis, I use data from the Survey of Professional Forecasters website:
https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/survey-of-professional-forecasters.

I collect real GDP forecasts for my baseline results, but also collect other variables listed in the following section. Since forecasters issue predictions for all variables in levels, some forecasts in the dataset must be transformed into growth rates before utilizing them in the analysis. For these variables, such as real GDP, I construct the annualized one-quarter ahead forecast from the forecasted levels, $f^i_{t|t}$ as follows:

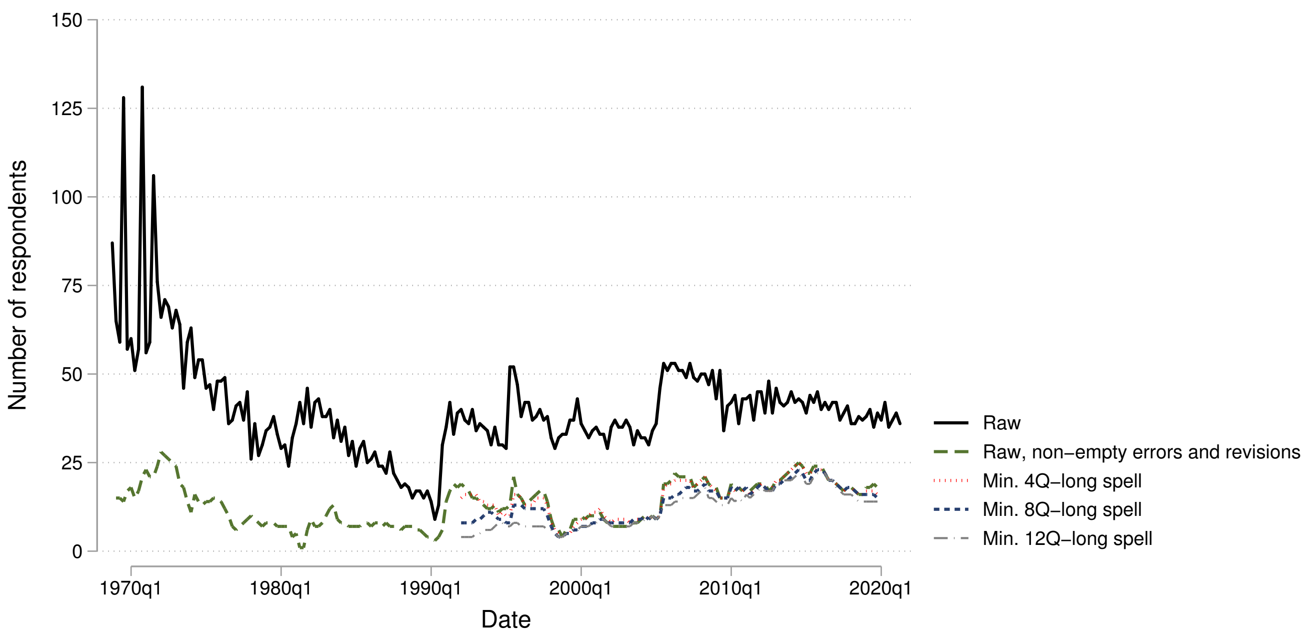
$$\widehat{x}^i_{t+1|t} = \left[\left(\frac{f^i_{t+1|t}}{f^i_{t|t}} \right)^4 - 1 \right] \times 100.$$

Other variables are already measured in growth rates such as CPI and do not require such a transformation.

The SPF is an unbalanced panel. Since forecasters can enter and exit the survey, the number of respondents varies over time. Figure A1 plots the number of respondents over time at different stages of data cleaning for my real GDP growth sample. The solid black line plots the number of respondents in each quarter in the raw data. The dashed green line plots the number of respondents in each quarter conditional on observing a one-quarter ahead forecast error and a one-quarter ahead forecast revision. This requires that a forecaster issue a one-quarter ahead forecast today and have issued a two-quarter ahead forecast in the previous quarter, the latter of which is necessary to construct the forecast revision, $\widehat{x}^i_{t+1|t} - \widehat{x}^i_{t+1|t-1}$. We see that this added restriction reduces the number of respondents in the survey, although the number of respondents is more stable over time.

The remaining lines in Figure A1 reflect increasingly more stringent sample restrictions. All of these other three time series start in 1992Q1 which is the selected date in which I begin my baseline sample. The dotted red line reflects the number of respondents when also requiring each respondent to have issued a minimum four-quarter-long string of forecasts. The short dashed navy blue line,

Figure A1: Respondents Over Time



Note: The figure plots the number of respondents in the SPF issuing real GDP growth forecasts across a range of different sample restrictions. My baseline specification imposes a minimum eight-quarter-long spell for each respondent, leaving me with 77 unique forecasters from 1992Q1 to 2019Q4. The four-quarter- and 12-quarter-long spell requirements are alternative specifications that I consider in Appendix C.

which reflects my baseline sample, imposes an eight-quarter-long spell requirement. Finally, the gray dashed-dotted line imposes a 12-quarter-long spell requirement for each respondent. Overall, we see that these requirements impose minor additional restrictions on the number of respondents observed in the sample.

Appendix B Estimation

In this section I detail the steps taken to estimate the models via MLE.

B.1 Maximum Likelihood Estimation

Across the different models, we have the following state space set up:

$$\begin{aligned}
\epsilon_{\mathbf{it}} &= \mathbf{A}\epsilon_{\mathbf{it-1}} + \mathbf{B}\eta_{\mathbf{it}} \\
\mathbf{z_{it}} &= \mathbf{C}\epsilon_{\mathbf{it}},
\end{aligned}$$

Where the latent state, ϵ_{it} is indexed by forecaster i and date t. This state vector includes the components of the macroeconomic variable which include the unobserved state, x_t and its innovations, w_t. In addition, this vector includes the unobserved Bayesian forecasts, $x^i_{t|t}$ and $x^i_{t|t-1}$ as well as their consensus analogs. The matrix $\mathbf{A}$ is the transition matrix. The vector η_{it} includes the state innovation, w_t, and signal noise, v^i_t.

The observation vector includes three measurements: individual one-quarter ahead forecast errors, individual one-quarter ahead forecast revisions, and one-quarter ahead consensus forecast errors. I keep only observations for which forecast errors and revisions are both populated, and fix a minimum spell length of eight quarters for which a forecaster must be observed in order to be included.

After stacking all of the forecasters, i, we can express the model in a form indexed only by date t. The state transition equation is

$$\epsilon_{\mathbf{t}} = \mathbf{T}\epsilon_{\mathbf{t-1}} + \mathbf{D}\mathbf{u_t}$$

where $\epsilon_{\mathbf{t}} = \begin{pmatrix} \epsilon_{\mathbf{1t}} \\ \epsilon_{\mathbf{2t}} \\ \vdots \\ \epsilon_{\mathbf{nt}} \end{pmatrix}$, $\mathbf{T} = \mathbf{I}_n \otimes \mathbf{A}$, and $\mathbf{D} = \mathbf{I}_n \otimes \mathbf{B}$, and $\mathbf{u_t} = \begin{pmatrix} \mathbf{u_{1t}} \\ \mathbf{u_{2t}} \\ \vdots \\ \mathbf{u}_{nt} \end{pmatrix}$.

The observation equation is:

$$\mathbf{z_t} = \mathbf{M_t W} \epsilon_{\mathbf{t}}$$

where: $\mathbf{z_t} = \begin{pmatrix} \mathbf{z_{1t}} \\ \mathbf{z_{2t}} \\ \vdots \\ \mathbf{z_{nt}} \end{pmatrix}$, and $\mathbf{W} = \mathbf{I}_n \otimes \mathbf{C}$.

The matrix $\mathbf{M_t}$ is a time-varying $3n_t \times 3n$ matrix, where n_t is the number of forecasters observed at time t. This matrix allows me to account for the unbalanced nature of the SPF panel data.

Defining $\mathbf{W_t} = \mathbf{M_t W}$, the Kalman filter equations are:

$$\begin{aligned}
\mathbf{F_t} &= \mathbf{W_t P_{t|t-1} W_t'} \\
\mathbf{K_t} &= \mathbf{P_{t|t-1} W_t' F_t^{-1}} \\
\mathbf{z_t^*} &= \mathbf{M_t z_t} - \mathbf{W_t z_{t|t-1}} \\
\mathbf{z_{t+1|t}} &= \mathbf{T}(\mathbf{z_{t|t-1}} + \mathbf{K_t z_t^*}) \\
\mathbf{P_{t+1|t}} &= \mathbf{T}((\mathbf{I_n} - \mathbf{K_t W_t})\mathbf{P_{t|t-1}})\mathbf{T}' + \mathbf{Q}
\end{aligned}$$

where $\mathbf{Q} = \mathbf{I_n} \otimes \begin{pmatrix} \sigma_w^2 & 0 & 0 \\ 0 & \sigma_e^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{pmatrix}$.

The log likelihood is therefore

$$LL = -\frac{1}{2}\Big(\sum_t n_t \log(2\pi) + \sum_t \log(\det \mathbf{F_t}) + \mathbf{S_{yy}} \Big)$$

where

$$\mathbf{S_{yy}} = \sum_t \mathbf{y_t^{*'}} \mathbf{F_t^{-1}} \mathbf{y_t^{*}}.$$

I estimate the model by constructing and maximizing the likelihood function numerically.

B.2 State Space Specifications for a Single Forecaster

Rational Expectations

State:

$$\begin{bmatrix} x_t \\ x_{t-1} \\ w_{t+1} \\ x^i_{t|t} \\ x^i_{t-1|t} \\ x^i_{t|t-1} \\ x^i_{t-1|t-1} \\ \overline{x}_{t|t} \\ \overline{x}_{t-1|t} \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \kappa_1\rho_1 & \kappa_1\rho_2 & 0 & (1-\kappa_1)\rho_1 & (1-\kappa_1)\rho_2 & 0 & 0 & 0 & 0 \\ \kappa_2\rho_1 & \kappa_2\rho_2 & 0 & 1-\kappa_2\rho_1 & -\kappa_2\rho_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_1 & \rho_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \kappa_1\rho_1 & \kappa_1\rho_2 & 0 & 0 & 0 & 0 & 0 & (1-\kappa_1)\rho_1 & (1-\kappa_1)\rho_2 \\ \kappa_2\rho_1 & \kappa_2\rho_2 & 0 & 0 & 0 & 0 & 0 & 1-\kappa_2\rho_1 & -\kappa_2\rho_2 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \\ w_t \\ x^i_{t-1|t-1} \\ x^i_{t-2|t-1} \\ x^i_{t-1|t-2} \\ x^i_{t-2|t-2} \\ \overline{x}_{t-1|t-1} \\ \overline{x}_{t-2|t-1} \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \kappa_1 & \kappa_1 \\ 0 & \kappa_2 & \kappa_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \kappa_1 & 0 \\ 0 & \kappa_2 & 0 \end{bmatrix} \begin{bmatrix} w_{t+1} \\ w_t \\ v^i_t \end{bmatrix}$$

Measurement:

$$\begin{bmatrix} x_{t+1} - \widehat{x}^i_{t+1|t} \\ \widehat{x}^i_{t+1|t} - \widehat{x}^i_{t+1|t-1} \\ x_{t+1} - x_{t+1|t} \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 & 1 & -\rho_1 & -\rho_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_1 & \rho_2 & -\rho_1 & -\rho_2 & 0 & 0 \\ \rho_1 & \rho_2 & 1 & 0 & 0 & 0 & 0 & -\rho_1 & -\rho_2 \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ w_{t+1} \\ x^i_{t|t} \\ x^i_{t-1|t} \\ x^i_{t|t-1} \\ x^i_{t-1|t-1} \\ \overline{x}_{t|t} \\ \overline{x}_{t-1|t} \end{bmatrix}$$

where $\begin{bmatrix} w_{t+1} \\ w_t \\ v^i_t \end{bmatrix} \sim N(\overline{\mu}, \Sigma)$, with $\overline{\mu} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sigma^2_w & 0 & 0 \\ 0 & \sigma^2_w & 0 \\ 0 & 0 & \sigma^2_v \end{bmatrix}$.

Overconfidence: The overconfidence model is similar, with the forecaster's filtering problem incorporating perceived signal noise, $\alpha_v\sigma_v$, which leads to distorted gains $\{\widehat{\kappa}_1, \widehat{\kappa}_2\}$.

Diagnostic Expectations

State:

$$
\begin{bmatrix} x_t \\ x_{t-1} \\ w_{t+1} \\ x^i_{t|t} \\ x^i_{t-1|t} \\ x^i_{t|t-1} \\ x^i_{t-1|t-1} \\ x^i_{t-2|t-1} \\ x^i_{t-1|t-2} \\ x^i_{t-2|t-2} \\ \overline{x}_{t|t} \\ \overline{x}_{t-1|t} \\ \overline{x}_{t|t-1} \\ \overline{x}_{t-1|t-1} \end{bmatrix}
=
\begin{bmatrix}
\rho_1 & \rho_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\kappa_1\rho_1 & \kappa_1\rho_2 & 0 & (1-\kappa_1)\rho_1 & (1-\kappa_1)\rho_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\kappa_2\rho_1 & \kappa_2\rho_2 & 0 & 1-\kappa_2\rho_1 & -\kappa_2\rho_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \rho_1 & \rho_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\kappa_1\rho_1 & \kappa_1\rho_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\kappa_1)\rho_1 & (1-\kappa_1)\rho_2 & 0 & 0 \\
\kappa_2\rho_1 & \kappa_2\rho_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-\kappa_2\rho_1 & -\kappa_2\rho_2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_1 & \rho_2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} x_{t-1} \\ x_{t-2} \\ w_t \\ x^i_{t-1|t-1} \\ x^i_{t-2|t-1} \\ x^i_{t-1|t-2} \\ x^i_{t-2|t-2} \\ x^i_{t-3|t-2} \\ x^i_{t-2|t-3} \\ x^i_{t-3|t-3} \\ \overline{x}_{t-1|t-1} \\ \overline{x}_{t-2|t-1} \\ \overline{x}_{t-1|t-2} \\ \overline{x}_{t-2|t-2} \end{bmatrix}
$$

$$
+
\begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 0 \\
1 & 0 & 0 \\
0 & \kappa_1 & \kappa_1 \\
0 & \kappa_2 & \kappa_2 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & \kappa_1 & 0 \\
0 & \kappa_2 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} w_{t+1} \\ w_t \\ v^i_t \end{bmatrix}
$$

$$\text{where} \begin{bmatrix} w_{t+1} \\ w_t \\ v_t^i \end{bmatrix} \sim N(\overline{\mu}, \Sigma), \text{ with } \overline{\mu} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \sigma_w^2 & 0 & 0 \\ 0 & \sigma_w^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{bmatrix}.$$

Measurement:

$$\begin{bmatrix} s_{t+1} - \widehat{x}^i_{t+1|t} \\ \widehat{x}^i_{t+1|t} - \widehat{x}^i_{t+1|t-1} \\ x_{t+1} - x_{t+1|t} \end{bmatrix} = \begin{bmatrix} c_{1,1} & c_{12} & c_{1,3} & c_{1,4} & c_{1,5} & c_{1,6} & c_{1,7} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{2,4} & c_{2,5} & c_{2,6} & c_{2,7} & c_{2,8} & c_{2,9} & c_{2,10} & 0 & 0 & 0 & 0 \\ c_{3,1} & c_{3,2} & c_{3,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{4,11} & c_{4,12} & c_{4,13} & c_{4,14} \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ w_{t+1} \\ x^i_{t|t} \\ x^i_{t-1|t} \\ x^i_{t|t-1} \\ x^i_{t-1|t-1} \\ x^i_{t-2|t-1} \\ x^i_{t-1|t-2} \\ x^i_{t-2|t-2} \\ \overline{x}_{t|t} \\ \overline{x}_{t-1|t} \\ \overline{x}_{t|t-1} \\ \overline{x}_{t-1|t-1} \end{bmatrix}$$

where

$$
\begin{aligned}
c_{1,1} &= \rho_1 \\
c_{1,2} &= \rho_2 \\
c_{1,3} &= 1 \\
c_{1,4} &= -(1+\varphi)\rho_1 \\
c_{1,5} &= -(1+\varphi)\rho_2 \\
c_{1,6} &= \varphi\rho_1 \\
c_{1,7} &= \varphi\rho_2 \\
c_{2,4} &= (1+\varphi)\rho_1 \\
c_{2,5} &= (1+\varphi)\rho_2 \\
c_{2,6} &= -\varphi\rho_1 \\
c_{2,7} &= -(1+\varphi)(\rho_1^2+\rho_2)+\varphi\rho_2 \\
c_{2,8} &= \varphi(\rho_1^2+\rho_2) \\
c_{2,9} &= -(1+\varphi)\rho_1\rho_2 \\
c_{2,10} &= \varphi\rho_1\rho_2 \\
c_{3,1} &= \rho_1 \\
c_{3,2} &= \rho_2 \\
c_{3,3} &= 1 \\
c_{3,11} &= -(1+\varphi)\rho_1 \\
c_{3,12} &= -(1+\varphi)\rho_2 \\
c_{3,13} &= \varphi\rho_1 \\
c_{3,14} &= \varphi\rho_2
\end{aligned}
$$

Misspecified Expectations

State:

$$\begin{bmatrix} x_t \\ x_{t-1} \\ w_{t+1} \\ x^i_{t|t} \\ x^i_{t|t-1} \\ \overline{x}_{t|t} \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \kappa_1\rho_1 & \kappa_1\rho_2 & 0 & (1-\kappa_1)\widehat{\rho} & 0 & 0 \\ 0 & 0 & 0 & \widehat{\rho} & 0 & 0 \\ \kappa_1\rho & \kappa_1\rho_2 & 0 & 0 & 0 & (1-\kappa_1)\widehat{\rho} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \\ w_t \\ x^i_{t-1|t-1} \\ x^i_{t-1|t-2} \\ \overline{x}_{t-1|t-1} \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \kappa_1 & \kappa_1 \\ 0 & 0 & 0 \\ 0 & \kappa_1 & 0 \end{bmatrix} \begin{bmatrix} w_{t+1} \\ w_t \\ v^i_t \end{bmatrix}$$

Measurement:

$$\begin{bmatrix} x_{t+1} - \widehat{x}^i_{t+1|t} \\ \widehat{x}^i_{t+1|t} - \widehat{x}^i_{t+1|t-1} \\ x_{t+1} - \widehat{x}^i_{t+1|t} \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 & 1 & -\widehat{\rho} & 0 & 0 \\ 0 & 0 & 0 & \widehat{\rho} & -\widehat{\rho} & 0 \\ \rho_1 & \rho_2 & 1 & 0 & 0 & -\widehat{\rho} \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ w_t \\ x^i_{t|t} \\ x^i_{t|t-1} \\ \overline{x}_{t|t} \end{bmatrix}$$

where $\begin{bmatrix} w_{t+1} \\ w_t \\ v^i_t \end{bmatrix} \sim N(\overline{\mu}, \Sigma)$, with $\overline{\mu} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sigma^2_w & 0 & 0 \\ 0 & \sigma^2_w & 0 \\ 0 & 0 & \sigma^2_v \end{bmatrix}$.

Appendix C Robustness

C.1 Likelihood Ratio Tests Across Variables

Table C1 reports the test statistics from the Vuong (1989) test across the different macroeconomic variables. For this one-sided test, a test statistic exceeding the critical value indicates a rejection of the null that misspecified expectations is observationally equivalent to a given model. For the majority of macroeconomic variables, misspecified expectations provides the superior fit to the data. Consistent with the results in Table 3, there is no evidence that misspecified expectations outperforms diagnostic expectations when modeling inflation based on the GDP deflator, real nonresidential investment, and industrial production. Furthermore, the evidence in favor of misspecified expectations for payroll employment and real residential investment is not statistically significant. For payroll employment, however, misspecified expectations is also not statistically different from overconfidence. Nonetheless, across most variables, there is evidence that the misspecified expectations model outperforms the other models.

Table C1: Likelihood Ratio Tests

	ME vs. RE	ME vs. OC	ME vs. DE
CPI	8.275***	6.768***	4.342***
GDP Deflator	7.848***	7.645***	-0.876
Housing starts	-2.561	-2.558	-2.541
Industrial production	2.566***	-1.299	-1.883
Payroll employment	2.019**	0.458	0.901
Real consumption expenditures	8.439***	7.809***	7.069***
Real federal government spending	5.669***	5.640***	4.579***
Real GDP	5.888***	2.574***	1.715**
Real nonresidential investment	4.429***	1.346*	-2.544
Real residential investment	3.843***	1.674**	-0.9561
Real state and local government spending	7.838***	7.047***	7.448***
Unemployment rate	3.200***	2.711***	1.640*
3-month Treasury bill	-5.477	-5.477	-5.473
10-year government bond	7.693***	7.538***	6.705***

Note: The table reports the test statistics from the Vuong (1989) test. Each row of the table refers to macroeconomic variable in the SPF, and each column denotes a pairwise comparison between the misspecified expectations model and another model. 'RE' denotes rational expectations, 'OC' denotes overconfidence, 'DE' denotes diagnostic expectations, and 'ME' denotes misspecified expectations. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance for these one-sided tests.

C.2 Estimates over Longer Sample Period

Table C2: Baseline Estimates (1968-2019)

Panel A: Stage 1 Parameter Estimates					
First order autocorrelation	ρ_1	0.494			
		(0.289)			
Second order autocorrelation	ρ_2	-0.007			
		(0.002)			
Persistent innovation dispersion	σ_w	2.582			
		(1.494)			
Panel B: Stage 2 Parameter Estimates					
		(1)	(2)	(3)	(4)
		RE	OC	DE	ME
Private noise dispersion	σ_v	1.980	1.980	1.980	1.980
		(0.065)	-	-	-
Overconfidence	α_v		0.00005		
			(0.000006)		
Diagnosticity	φ			1.084	
				(0.135)	
Perceived persistence	$\widehat{\rho}$				0.872
					(0.039)
Log likelihood		-18330	-14469	-13905	-13866
AIC		36669	28950	27822	27744
Δ_i		8925	1206	78	0
w_i		0.0	0.0	0.0	1.0

Note: Panel A reports estimates for the first stage MLE which estimates the parameters governing the fundamental process. Block bootstrapped standard errors (Politis and Romano, 1994) are reported in parentheses. Panel B reports the parameter estimates based on the second and third steps of the estimation procedure. Column (1) reports the rational expectations (RE) model, column (2) reports the overconfidence (OC) model, column (3) reports the diagnostic expectations (DE) model, and column (4) reports the misspecified expectations (ME) model. Bootstrapped standard errors reported in parentheses. For each model, panel C reports the maximized log likelihood, AIC, the difference between a model's AIC and smallest AIC across the models (see equation (4)), and the Akaike weight, w_i (see equation (5)).

Table C3: Other Macroeconomic Variables (Longer Sample)

	RE	OC	DE	ME
CPI	0.0	0.0	0.0	1.0
GDP Deflator	0.0	0.0	0.0	1.0
Housing starts	0.58	0.21	0.21	0.0
Industrial production	0.0	1.0	0.0	0.0
Payroll employment	0.0	0.0	1.0	0.0
Real consumption expenditures	0.0	0.0	0.0	1.0
Real federal government spending	0.0	0.0	0.0	1.0
Real GDP	0.0	0.0	0.0	1.0
Real nonresidential investment	0.0	0.0	1.0	0.0
Real residential investment	0.0	0.0	0.0	1.0
Real state and local government spending	0.0	0.0	0.0	1.0
Unemployment rate	0.0	0.0	0.0	1.0
3-month Treasury bill	0.0	0.0	0.0	1.0
10-year government bond	0.0	0.0	1.0	0.0

Note: Each row of the table reports the Akaike weight, w_i, for a given macroeconomic variable based on equation (5). 'RE' denotes rational expectations, 'OC' denotes overconfidence, 'DE' denotes diagnostic expectations, and 'ME' denotes misspecified expectations.

C.3 AR(1) Process

I re-estimate the baseline model for real GDP growth forecasts assuming that the data generating process follows an AR(1) rather than an AR(2). In this case, misspecified expectations imply that forecasters fully understand the AR process governing real GDP growth, $x_t = \rho x_{t-1} + w_t$, but they assign the wrong persistence to it, $\widehat{\rho}$ as in Fuster et al. (2012). Table C4 reports the results which show that misspecified expectations outperform the alternatives when considering simpler dynamics.

Table C4: AR(1) Estimates

Panel A: Stage 1 Parameter Estimates					
Autocorrelation	ρ	0.432			
		(0.001)			
Innovation dispersion	σ_w	1.663			
		(0.765)			
Panel B: Stage 2 Parameter Estimates					
		(1)	(2)	(3)	(4)
		RE	OC	DE	ME
Private noise dispersion	σ_v	1.262	1.262	1.262	1.262
		(0.021)	-	-	-
Overconfidence	α_v		0.00007		
			(0.0001)		
Diagnosticity	φ			0.926	
				(0.120)	
Perceived persistence	$\widehat{\rho}$				0.752
					(0.039)
Log likelihood		-8852.5	-6744.4	-6762.9	-6720.3
AIC		17713	13863	13504	13428
Δ_i		4284	425	76	0
w_i		0.0	0.0	0.0	1.0

Note: Panel A reports estimates for the first stage MLE which estimates the parameters governing the fundamental process. Block bootstrapped standard errors (Politis and Romano, 1994) are reported in parentheses. Panel B reports the parameter estimates based on the second and third steps of the estimation procedure. Column (1) reports the rational expectations (RE) model, column (2) reports the overconfidence (OC) model, column (3) reports the diagnostic expectations (DE) model, and column (4) reports the misspecified expectations (ME) model. Bootstrapped standard errors reported in parentheses. For each model, panel C reports the maximized log likelihood, AIC, the difference between a model's AIC and smallest AIC across the models (see equation (4)), and the Akaike weight, w_i (see equation (5)).

Table C5 reports the Akaike weights for each variable. For most variables, misspecified expectations outperforms the other models, consistent with the baseline AR(2) results.

Table C5: Other Macroeconomic Variables (AR1)

	RE	OC	DE	ME
CPI	0.0	0.0	1.0	0.0
GDP Deflator	0.0	0.0	1.0	0.0
Housing starts	0.0	0.0	0.0	1.0
Industrial production	0.0	0.0	1.0	0.0
Payroll employment	0.0	0.0	0.0	1.0
Real consumption expenditures	0.0	0.0	0.0	1.0
Real federal government spending	0.0	0.0	1.0	0.0
Real GDP	0.0	0.0	0.0	1.0
Real nonresidential investment	0.0	0.0	1.0	0.0
Real residential investment	0.0	0.0	0.0	1.0
Real state and local government spending	0.0	0.0	0.0	1.0
Unemployment rate	0.0	0.0	0.0	1.0
3-month Treasury bill	0.0	0.0	0.0	1.0
10-year government bond	0.0	0.0	0.0	1.0

Note: Each row of the table reports the Akaike weight, w_i, for a given macroeconomic variable based on equation (5). 'RE' denotes rational expectations, 'OC' denotes overconfidence, 'DE' denotes diagnostic expectations, and 'ME' denotes misspecified expectations.

C.4 Alternative Spell Lengths

Table C6: Real GDP Estimates (Minimum 4-Quarter Spells)

Panel A: Fundamental Parameters					
First order autocorrelation	ρ_1	0.434			
		(0.215)			
Second order autocorrelation	ρ_2	-0.006			
		(0.001)			
Persistent innovation dispersion	σ_w	1.663			
		(0.842)			
Panel B: Information/Bias Parameters					
		(1) RE	(2) OC	(3) DE	(4) ME
Private noise dispersion	σ_v	1.235	1.235	1.235	1.235
		(0.029)	-	-	-
Overconfidence	α_v		0.00006		
			(0.00001)		
Diagnosticity	φ			0.975	
				(0.117)	
Perceived persistence	$\widehat{\rho}$				0.768
					(0.036)
Panel C: Model Selection					
Log likelihood		-10200	-7902	-7611	-7565
AIC		20410	15816	15234	15142
Δ_i		5268	674	92	0
w_i		0.0	0.0	0.0	1.0

Note: Panel A reports estimates for the first stage MLE which estimates the parameters governing the fundamental process. Block bootstrapped standard errors (Politis and Romano, 1994) are reported in parentheses. Panel B reports the parameter estimates based on the second and third steps of the estimation procedure. Column (1) reports the rational expectations (RE) model, column (2) reports the overconfidence (OC) model, column (3) reports the diagnostic expectations (DE) model, and column (4) reports the misspecified expectations (ME) model. Bootstrapped standard errors reported in parentheses. For each model, panel C reports the maximized log likelihood, AIC, the difference between a model's AIC and smallest AIC across the models (see equation (4)), and the Akaike weight, w_i (see equation (5)).

Table C7: All Variables (Minimum 4-Quarter Spells)

	RE	OC	DE	ME
CPI	0.0	0.0	0.0	1.0
GDP Deflator	0.0	0.0	0.0	1.0
Housing starts	0.58	0.21	0.21	0.0
Industrial production	0.0	0.99	0.01	0.0
Payroll employment	0.0	0.0	0.0	1.0
Real consumption expenditures	0.0	0.0	0.0	1.0
Real federal government spending	0.0	0.0	0.0	1.0
Real GDP	0.0	0.0	0.0	1.0
Real nonresidential investment	0.0	0.0	0.0	1.0
Real residential investment	0.0	0.0	1.0	0.0
Real state and local government spending	0.0	0.0	0.0	1.0
Unemployment rate	0.0	0.0	0.0	1.0
3-month Treasury bill	0.05	0.02	0.94	0.0
10-year government bond	0.0	0.0	0.0	1.0

Note: Each row of the table reports the Akaike weight, w_i, for a given macroeconomic variable based on equation (5). ‘RE’ denotes rational expectations, ‘OC’ denotes overconfidence, ‘DE’ denotes diagnostic expectations, and ‘ME’ denotes misspecified expectations.

Table C8: Real GDP Estimates (Minimum 12-Quarter Spells)

Panel A: Fundamental Parameters					
First order autocorrelation	ρ_1	0.434			
		(0.215)			
Second order autocorrelation	ρ_2	-0.006			
		(0.001)			
Persistent innovation dispersion	σ_w	1.663			
		(0.842)			
Panel B: Information/Bias Parameters					
		(1)	(2)	(3)	(4)
		RE	OC	DE	ME
Private noise dispersion	σ_v	1.263	1.263	1.263	1.263
		(0.367)	-	-	-
Overconfidence	α_v		0.00008		
			(0.00001)		
Diagnosticity	φ			0.914	
				(0.197)	
Perceived persistence	$\widehat{\rho}$				0.752
					(0.076)
Panel C: Model Selection					
Log likelihood		-7715.5	-6056.7	-5911.99	-5887.2
AIC		15441	12125	11836	11736
Δ_i		3655	339	49	0
w_i		0.0	0.0	0.0	1.0

Note: Panel A reports estimates for the first stage MLE which estimates the parameters governing the fundamental process. Block bootstrapped standard errors (Politis and Romano, 1994) are reported in parentheses. Panel B reports the parameter estimates based on the second and third steps of the estimation procedure. Column (1) reports the rational expectations (RE) model, column (2) reports the overconfidence (OC) model, column (3) reports the diagnostic expectations (DE) model, and column (4) reports the misspecified expectations (ME) model. Bootstrapped standard errors reported in parentheses. For each model, panel C reports the maximized log likelihood, AIC, the difference between a model's AIC and smallest AIC across the models (see equation (4)), and the Akaike weight, w_i (see equation (5)).

Table C9: All Variables (Minimum 12-Quarter Spells)

	RE	OC	DE	ME
CPI	0.0	0.0	0.0	1.0
GDP Deflator	0.0	0.0	1.0	0.0
Housing starts	0.52	0.21	0.21	0.0
Industrial production	0.0	0.0	1.0	0.0
Payroll employment	0.0	1.0	0.0	0.0
Real consumption expenditures	0.0	0.0	0.0	1.0
Real federal government spending	0.0	0.0	0.0	1.0
Real GDP	0.0	0.0	0.0	1.0
Real nonresidential investment	0.0	0.0	1.0	0.0
Real residential investment	0.0	0.0	1.0	0.0
Real state and local government spending	0.0	0.0	0.0	1.0
Unemployment rate	0.0	0.0	0.0	1.0
3-month Treasury bill	0.01	0.0	0.99	0.0
10-year government bond	0.0	0.0	0.0	1.0

Note: Each row of the table reports the Akaike weight, w_i, for a given macroeconomic variable based on equation (5). 'RE' denotes rational expectations, 'OC' denotes overconfidence, 'DE' denotes diagnostic expectations, and 'ME' denotes misspecified expectations.