Misspecified Expectations *

Julio L. Ortiz †
Federal Reserve Board

June 2022

Abstract

Using data from professional forecasters, I find that a theory of misspecified expectations, in which forecasters' perceived law of motion differs from the objective law of motion, outperforms alternative models in its ability to fit prediction errors and revisions. Misspecification is successful in part because it matches updating behavior in the data. My framework delivers a novel testable implication through which I provide robust evidence of misspecification-related overextrapolation across a range of macroeconomic variables. I conclude that misspecified expectations can serve as a suitable benchmark alternative to full information rational expectations.

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

*I would like to thank Francesco Bianchi, Stephen Terry, and Fabian Winkler for their valuable feeback and helpful conversations. This paper subsumes a previous draft titled "A New Fact to Discipline Models of Beliefs." 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 Constitution Avenue NW, Washington, DC 20551.

1 Introduction

Survey forecasts appear to violate the Law of Iterated Expectations. The predictability of forecast errors on the basis of information available to respondents in real time has animated a rich literature studying deviations from full information rational expectations (FIRE) (Mankiw and Reis, 2002; Coibion and Gorodnichenko, 2015; Bordalo et al., 2020; Kohlhas and Walther, 2021). 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 establish such a benchmark in the literature.

I study a noisy information rational expectations model along with three leading non-FIRE theories of: diagnostic expectations (Bordalo et al., 2020), overconfidence (Daniel et al., 1998), and misspecified expectations (Fuster et al., 2010). Using panel data from the Survey of Professional Forecasters (SPF), I estimate these models and find that misspecified expectations best fits observed forecast errors and revisions. Misspecification is successful relative to the other theories because, in addition to generating overreactions, it is able to produce realistic updating behavior among forecasters. Finally, I derive a novel testable implication through which I document reduced form evidence of overextrapolation stemming from misspecified expectations. Taken together, my results suggest that a theory in which forecasters adopt parsimonious models of richer underlying processes (e.g., specify an AR(1) instead of the true AR(2) process) can serve as a successful benchmark departure from FIRE.

I begin by illustrating a simple forecasting problem in a noisy information environment. The macroeconomic variable of interest has a persistent and transitory component which cannot be separately observed. Forecasters issue predictions of this aggregate based on information gleaned from its lagged realization and a contemporaneous private signal. The inclusion of the former as an observable is directly motivated by the timing of the survey. In this linear Gaussian setting, forecasters employ the Kalman filter in order 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. In this paper, I focus on straightforward biases which can be flexibly embedded into more complex macroeconomic models. In particular, I study theories of (i) diagnostic expectations, in which forecasters overreact to all information, (ii) overconfidence, in which forecasters overreact to private information, and (iii) misspecified expectations, in which forecasters overextrapolate. I estimate these models via maximum likelihood estimation (MLE). Inspecting the likelihood functions and information criteria, I find that misspecified expectations outperforms the alternatives.

I then extend each of these models to feature strategic interactions. In this case, forecasters form expectations about the macroeconomic variable as well as the consensus forecast, making higher order beliefs important for expectation formation. Upon estimating these models, I find robust evidence of strategic substitutability among forecasters, consistent with Ottaviani and Sørensen (2006) and Gemmi and Valchev (2021). The results from this extension once again reveal that misspecified expectations outperforms the other theories.

My results are robust to a variety of alternative considerations. First, I show that misspecified expectations provide the best fit to forecasts of other aggregates such as inflation. Second, I show that my results remain unchanged when estimating the models over a different sample period. Finally, I find that my results hold under alternative assumptions about the objective data generating process.

Digging deeper, I show that misspecified expectations provides a better fit to the data because it can jointly generate individual overreaction and empirically realistic updating behavior. All three non-FIRE models can produce a negative linear relationship between errors and revisions, the moment of interest for measuring overreactions. The candidate models, however, are not observationally equivalent and imply distinct updating behavior. Forecasters update, in part, based on news about the past and the present. The misspecified expectations model delivers an updating rule in which forecasters place the most weight on their priors, followed by current news, and then past news, as in the data. Overconfident expectations, on the other hand, place too much weight on current news. Moreover, diagnostic forecasters counterfactually update in the opposite direction of their priors while also placing too much weight on past news. Thus, despite being able to produce overreactions, only misspecified expectations can accurately capture forecasters' updating behavior.

Lastly, I provide a new testable prediction through which I estimate misspecification-related overextrapolation in the data. This testable implication amounts to a regression of the forecast error on the current forecast, controlling for past forecast errors. After estimating this regression via OLS for a range of macroeconomic variables, I find that forecasters robustly overextrapolate the objective persistence of the driving process by 0.10 to 0.20 for most macroeconomic variables. These reduced form estimates are consistent with the maximum likelihood estimates.

Though I do not consider all models of expectation formation, my findings suggest that updating dynamics are an important feature of the data to match. Whereas many theories can explain overreaction in expectations, these overreactions emanate from different sources depending on the specifics of the model. Matching observed updating behavior, therefore, provides further discipline to these models and ultimately allows one to better understand why forecasts violate FIRE. Hence, beyond the set of theories that I consider, the ideal alternative to FIRE should generate empirically-consistent updating rules.

My paper relates to a number of both longstanding and more recent contributions to the literature. First, my paper relates to the literature studying 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 these apparent overreactions, 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 behavioral biases have also been proposed in the literature (Broer and Kohlhas, 2019; Rozsypal and Schlafmann, 2019).

At the same time, my paper relates to the literature studying rational deviations from FIRE. For instance, Azeredo da Silvera et al. (2020) propose a theory of noisy memory in which forecasters optimize over their history of past signals. In addition, Burgi and Ortiz (2022) offer a theory of low frequency smoothing, which, when linked to high frequency forecasts, can generate high frequency overreactions. Moreover, Farmer et al. (2021) propose an explanation relating to learning under nonlinear dynamics. This paper studies parsimonious deviations from FIRE and argues that misspecification, similar to Fuster et al. (2010), can serve as a benchmark non-FIRE theory to approximate empirically relevant belief dynamics in macroeconomic models. The underlying driver of misspecification can be rational or non-rational.

The rest of the paper is organized as follows. Section 2 details the noisy information setting and the rational baseline model. Section 3 estimates the menu of biased models. Section 4 extends these models to feature strategic interaction. Section 5 explores why misspecified expectations outperforms the alternatives. Section 6 estimates misspecification-related overextrapolation in the data. Section 7 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. Suppose that a forecaster wishes to predict some aggregate variable, x_t, which is the sum of a persistent and transitory component:

$$x_t = s_t + e_t, \quad e_t \sim N(0, \sigma_e^2),$$

where s_t evolves according to an AR(1) process,

$$s_t = \rho s_{t-1} + w_t, \quad w_t \sim N(0, \sigma_w^2).$$

At a given point in time, t, the forecaster, indexed by i, has access to a noisy private signal,[1]

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

[1] When $\sigma_v = \sigma_e = 0$, the model collapses to a FIRE model.

and also observes the previous period's realization of the aggregate variable, x_{t-1}.

The forecaster's objective is to minimize her mean squared errors. Hence, the optimal forecast of x_t is the conditional expectation $\mathbb{E}(x_t|\Omega_t^i)$, where Ω_t^i denotes the forecaster's information set at time t. In this linear Gaussian setting, forecaster can employ the Kalman filter to obtain the optimal forecast. Casting the model in state space form, we define the state and measurement equations, respectively, as:[2]

$$\begin{bmatrix} s_t \\ s_{t-1} \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} s_{t-1} \\ s_{t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} w_t$$
$$\begin{bmatrix} y_t^i \\ x_{t-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s_t \\ s_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_t^i \\ e_{t-1} \end{bmatrix}.$$

From the Kalman filter, we obtain the following predict and update equations for s_t,

$$\text{Predict:} \quad s_{t|t-1}^i = \rho s_{t-1|t-1}^i \tag{1}$$

$$\text{Update:} \quad s_{t|t}^i = s_{t|t-1}^i + \kappa_1(y_t^i - s_{t|t-1}^i) + \kappa_2(x_{t-1} - s_{t-1|t-1}^i), \tag{2}$$

where κ_1 and κ_2 denote the Kalman gains, which are the optimal weights placed on either signal when updating the current prediction $s_{t|t}^i$.

Note that in this setting, the optimal forecast of x_t coincides with the optimal forecast of the latent state, $x_{t|t}^i = s_{t|t}^i$. Furthermore, the forecast error, $x_t - x_{t|t}^i$, is uncorrelated with anything residing in the forecaster's information set in period t. At the same time, the forecast revision, $x_{t|t}^i - x_{t|t-1}^i$, is uncorrelated with anything in the forecaster's information set at time t (Nordhaus and Durlauf, 1984; Nordhaus, 1987). In the next section, I will consider a menu of alternative theories of expectation formation which deviate from this baseline model.

[2]Note that y_t^i, though observable to the forecaster, it is unobservable to the econometrician. This motivates the use of the statistical model in the next section.

3 Assessing Alternative Models of Beliefs

In this section, I estimate leading theories of expectation formation, and show that model misspecification is best able to explain the biases present in the data. Specifically, I consider models of diagnostic expectations, overconfidence, and misspecification.

3.1 Diagnostic Expectations (DE)

Diagnostic expectations is a theory of overreaction to new information. This theory involves a psychological distortion such that, when confronted with good news, forecasters place a higher probability on good future realizations since these developments are "top of mind." Note that the latent state of interest follows a Markov process with an objective conditional distribution of $f(s_{t+1}|s_t)$. A diagnostic forecaster believes the distribution to be

$$f^{\varphi}(s_{t+1}|s_t) \propto f(s_{t+1}|s_t)\left[\frac{f(s_{t+1}|s_t)}{f\big(s_{t+1}|\mathbb{E}_{t-1}(s_t)\big)}\right]^{\varphi}.$$

Under rational expectations, $\varphi = 0$. With diagnostic beliefs, $\varphi > 0$, which implies that a forecaster's current-period forecast of x_t is

$$\widehat{x}^i_{t|t} = \mathbb{E}_{it}(x_t) + \varphi\big[\mathbb{E}_{it}(x_t) - \mathbb{E}_{it-1}(x_t)\big],$$

and, based on the assumed dynamics, their h-step ahead forecast is

$$\widehat{x}^i_{t+h|t} = \rho^h \widehat{x}^i_{t|t}.$$

Diagnostic expectations is interpreted as a distortion relating to memory retrieval. Recent information is more readily accessible and therefore overemphasized.

3.2 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 optimal thereby placing excessive weight on incoming private information. In this case, forecasters perceive the following private signal:

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

where $\breve{\sigma}_v = \alpha_v \sigma_v$ and $\alpha_v \in (0,1)$. When $\alpha_v < 1$, then the forecaster is overconfident.[3] The predict and updating rules are as in the rational case, but with an excessively large weight placed on new information, $\widehat{\kappa}_1 > \kappa_1$, due to $\alpha_v < 1$. Importantly, whereas the diagnostic forecasters described in the previous section overreact to all news, overconfident forecasters overreact to *private* news.

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 a more parsimonious model rather than estimate the richer dynamics governing the true data generating process. Or perhaps forecasters are predisposed to use simpler models due to cognitive frictions. A number of theories fall within 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 (2019) for the case of overextrapolation. In a different setting, Barberis et al. (1998) present a model in which an investor misperceives a random walk process.

To capture this form of expectation formation, I focus on a version of misspecification which is closest to natural expectations (Fuster et al., 2010, 2012). As with natural expectations, I assume that

[3] In theory, it is also possible to have $\alpha_v > 1$, in which case the forecaster is underconfident.

forecasters neglect longer lags in the data generating process. In its simplest form, the underlying state follows an AR(2) process:

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

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

$$s_t = \widehat{\rho} s_{t-1} + u_t,$$

where $u_t = (\rho_1 - \widehat{\rho})s_{t-1} + \rho_2 s_{t-2} + w_t$.[4] Importantly, forecasters still understand the information structure. If the perceived persistence loads excessively onto the first lag, then forecasters will exhibit overreactions.

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/rational inattention (Woodford, 2001; Sims, 2003), and adaptive expectations (Cagan, 1956; Nerlove, 1958), among others. In the present setting, I introduce the aforementioned biases into a noisy information environment. As a result, the models that I consider feature some scope for underreaction. I abstract away from theories of pure underreaction, 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 of overreaction include imperfect memory (Azeredo da Silvera et al., 2020; Afrouzi et al., 2021), temporal consistency (Burgi and Ortiz, 2022), and asymmetric attention (Kohlhas and Walther, 2021). I abstract away from these models because they are not able to be flexibly nested into the current setting. The aforementioned

[4] This form of misspecification is technically different from natural expectations since I do not explictly model an AR(2) in levels and assume that agents forecast an AR(1) in growth rates. In addition, here, the perceived persistence is estimated from the data rather than defined to be a function of the true autocorrelation parameters.

theories require additional information such as the inclusion of annual forecasts, the introduction of countercyclical components, and additional parameters which would pose a further challenge to identification.[5]

While I limit the set of models considered to the rational baseline, diagnostic expectations, overconfidence, and misspecification, I supplement my results with various robustness checks. Overall, one might view these models as the most tractable among the broader set of theories of overreaction. This parsimony is desirable, particularly when embedding non-FIRE expectations into a macroeconomic model. Moreover, extending beyond the set of models considered here, my results imply that any successful non-FIRE theory should be able to generate realistic updating dynamics, as I will show in Section 5.

3.5 Maximum Likelihood Estimation

I estimate the rational model along with the three biased models via MLE. The parameters of the rational model are collected in the vector $\theta = (\rho \;\; \sigma_w \;\; \sigma_v \;\; \sigma_e)$. Each of the three biased models includes some additional parameter(s): $\{\alpha_v\}$, $\{\varphi\}$, and $\{\rho_1, \rho_2, \widehat{\rho}\}$. Using the expressions from the forecaster's signal extraction problem detailed in the previous section, (1) and (2), I define a separate filtering problem for the econometrician. Below I detail the state and measurement equations for a single forecaster i in the rational baseline model. Appendix C provides additional details on the construction of the likelihood function and the state space formulation of the biased models.
State:

$$\begin{bmatrix} s_t \\ w_{t+1} \\ e_{t+1} \\ e_t \\ s^i_{t|t} \\ s^i_{t|t-1} \\ s_{t|t} \end{bmatrix} = \begin{bmatrix} \rho & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \kappa_1\rho + \kappa_2 & \kappa_1 & 0 & \kappa_2 & (1-\kappa_1)\rho - \kappa_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho & 0 & 0 \\ \kappa_1\rho + \kappa_2 & \kappa_1 & 0 & \kappa_2 & 0 & 0 & (1-\kappa_1)\rho - \kappa_2 \end{bmatrix} \begin{bmatrix} s_{t-1} \\ w_t \\ e_t \\ e_{t-1} \\ s^i_{t-1|t-1} \\ s^i_{t-1|t-2} \\ s_{t-1|t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \kappa_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_{t+1} \\ e_{t+1} \\ v^i_t \end{bmatrix}$$

[5] Moreover, I abstract away from traditional theories of overextrapolation (Goodwin, 1947) because these expectations are backward looking.

Measurement:

$$
\begin{bmatrix} x_{t+1} - \widehat{x}^i_{t+1|t} \\ \widehat{x}^i_{t+1|t} - \widehat{x}^i_{t+1|t-1} \\ x_{t+1} - \widehat{x}_{t+1|t} \end{bmatrix} = \begin{bmatrix} \rho & 1 & 1 & 0 & -\rho & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho & -\rho & 0 \\ \rho & 1 & 1 & 0 & 0 & 0 & -\rho \end{bmatrix} \begin{bmatrix} s_t \\ w_{t+1} \\ e_{t+1} \\ e_t \\ s^i_{t|t} \\ s^i_{t|t-1} \\ s_{t|t} \end{bmatrix}
$$

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

I construct the likelihood function based on this state space model using the panel of SPF real GDP forecasts.[6] 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 horizons. Based on the timing of the survey, forecasters have access to the prior value of realized GDP, which motivates the use of the lagged outcome as a common signal.

In my baseline estimates, I consider forecasts made from 1992Q1 to 2019Q4. I choose this more recent 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 D I estimate the models over the full sample and show that my results remain qualitatively unchanged over this longer period.

The estimation is done as follows. From the SPF, I obtain a panel of one-quarter ahead forecast errors, forecast revisions, and consensus errors. These three observables serve as measurements, as detailed in the state space model above. For a given parameter guess, θ, I simulate the Kalman filter in order to retrieve the steady state Kalman gains, κ_1 and κ_2, which are nonlinear functions of the

[6] Summary statistics are reported in Appendix A.

model parameters. I then use θ along with the Kalman gains to construct the likelihood function, $\mathcal{L}(x;\theta)$. I search the parameter space to maximize the log likelihood, $\ell(x;\theta)$.

Table 1: Baseline Estimates

Panel A: Parameter Estimates					
Description	Parameter	(1) RE	(2) DE	(3) OC	(4) ME
Persistence	ρ	0.992	0.616	0.756	
		(0.022)	(0.065)	(0.064)	
Persistent innovation dispersion	σ_w	1.240	0.819	1.056	0.905
		(0.084)	(0.071)	(0.067)	(0.125)
Transitory innovation dispersion	σ_e	1.783	1.720	1.670	1.494
		(0.080)	(0.055)	(0.057)	(0.191)
Private noise dispersion	σ_v	0.975	0.748	1.005	0.989
		(0.192)	(0.154)	(0.165)	(0.463)
Diagnosticity	φ		1.780		
			(0.532)		
Overconfidence	α_v			0.279	
				(0.137)	
Perceived persistence	$\widehat{\rho}$				0.811
					(0.081)
First order autocorrelation	ρ_1				0.446
					(0.126)
Second order autocorrelation	ρ_2				0.552
					(0.157)
Panel B: Model Selection					
Log likelihood		-7658.4	-7444.0	-7518.2	-7293.5
AIC		15325	14898	15046	14602
BIC		15346	14925	15073	14635

Note: Panel A reports parameters estimates. Column (1) denotes the rational expectations model, column (2) reports the diagnostic expectations model, column (3) reports the overconfidence model, and column (4) reports the misspecified expectations model. Bootstrapped standard errors reported in parenthesis. Panel B reports the maximized log likelihood as well as AIC and BIC for each model.

The parameter estimates are reported in Panel A of Table 1. The estimated rational expectations model in column (1) implies a highly persistent latent process, with a signal-to-noise ratio of about 0.45. In the diagnostic expectations model, reported in column (2), the estimated persistence of the latent process declines notably, along with the variance of its innovations. The estimated degree of diagnosticity is larger than estimates in the literature (Bordalo et al., 2020, 2021). Moreover, in column (3), the estimated degree of overconfidence is 0.28 which means that forecasters believe the dispersion of their private signal noise is less than one third of what it truly is. Similar to the diagnostic expectations results in column (2), these estimates imply a strong amount of overreaction among forecasters in the model. Column (4) reports the misspecified expectations model. Forecasters believe the underlying process to be an AR(1) with a persistence of 0.81. In reality, the process follows an AR(2) with an estimated first order autocorrelation coefficient of 0.45. These estimates suggest that forecasters overextrapolate, consistent with the overreaction estimated in the other two non-FIRE models.

Panel B of Table 1 reports the maximized likelihood functions as well as the corresponding Akaike and Bayesian information criteria (AIC and BIC, respectively). After accounting for the additional parameter involved in fitting the misspecified expectations model, the data still favor misspecified expectations since it delivers the smallest information criteria.

I consider a range of robustness checks in Appendix D. I first re-estimate the models to fit CPI-based inflation forecasts. I then estimate the model across the full sample period in the SPF. Finally, I fit all of the models with the richer AR(2) process assumed in the misspecified expectations model.

4 Incorporating Strategic Interactions

In this section, I extend the four models to include strategic interactions. Strategic considerations have been shown to matter for expectation formation (Ottaviani and Sørensen, 2006; Broer and Kohlhas, 2019; Gemmi and Valchev, 2021). I introduce strategic incentives into the models which lead forecasters to form expectations about each others' forecasts. The results from this extension confirm that misspecified expectations best fit the survey data relative to the other models.

In models of strategic interaction, forecasters have the motive to either mimic or to deviate from the consensus forecast. Therefore, rather than traditional mean squared loss, the forecaster's loss function is:

$$\min_{\widehat{x}^i_{t|t}} (x_t - \widehat{x}^i_{t|t})^2 + \gamma(\widehat{x}^i_{t|t} - \widehat{x}_{t|t})^2,$$

where γ denotes the degree of strategic interaction. When $\gamma > 0$ then there is strategic complementarity and forecasters will tend to track the consensus belief. On the other hand, when $\gamma < 0$, there is strategic substitutability and forecasters will wish to distinguish themselves from the consensus belief.

The first order condition implies the following:

$$\widehat{x}^i_{t|t} = \frac{1}{1+\gamma}\mathbb{E}_{it}(x_t) + \frac{\gamma}{1+\gamma}\mathbb{E}_{it}(x_{t|t}),$$

which means the forecaster sets her reported forecast to be a combination of the the conditional expectation as well as the expectation over what the consensus forecast will be.

The presence of strategic incentives makes higher order beliefs crucial to this model. In particular, the consensus forecast at time t is denoted by F_t and it is defined as

$$F_t = \frac{1}{1+\gamma}\sum_{k=0}^{\infty}\left(\frac{\gamma}{1+\gamma}\right)^k \mathbb{E}^{(k)}(x_t) = \frac{1}{1+\gamma}x_{t|t} + \frac{\gamma}{1+\gamma}F_{t|t}.$$

where $\mathbb{E}^{(k)}$ is the k^{th}-order expectation of x_t. I solve this model by guessing and verifying a law of motion for the consensus forecast, as in Woodford (2001) and Coibion and Gorodnichenko (2015). Given my setting, I guess the following evolution of the consensus forecast: F_t:

$$F_t = c_1 s_{t-1} + c_2 F_{t-1} + c_3 w_t + c_4 e_{t-1},$$

where $c_1 = c_3\rho + c_4$, $c_2 = (1-\kappa_{11})\rho - \kappa_{12} + \gamma(c_1 - \kappa_{31}\rho)$, $c_3 = \frac{\kappa_{11}+\gamma\kappa_{31}}{1+\gamma}$, and $c_4 = \frac{\kappa_{12}+\gamma\kappa_{32}}{1+\gamma}$, and κ_{ij} are elements of the Kalman gain matrix, κ. Appendix B provides further details on the analytical derivation of these coefficients.

Table 2 report the results from the strategic interactions models. In addition to the usual parameters, this model also estimates the degree of strategic interaction, γ, and finds that it is robustly

Table 2: Alternative Models with Strategic Interaction

Panel A: Parameter Estimates					
Description	Parameter	(1) RE	(2) DE	(3) OC	(4) ME
Persistence	ρ	0.853	0.870	0.811	
		(0.096)	(0.044)	(0.133)	
Persistent innovation	σ_w	1.241	1.182	1.130	0.614
		(0.186)	(0.078)	(0.208)	(0.231)
Private noise dispersion	σ_e	1.78	1.750	1.707	1.590
		(0.087)	(0.133)	(0.061)	(0.331)
Transitory innovation	σ_v	1.234	1.234	1.208	1.179
		(0.169)	(0.118)	(0.301)	(0.434)
Degree of strategic interaction	γ	-0.684	-0.684	-0.615	-0.537
		(0.219)	(0.015)	(0.062)	(0.206)
Diagnosticity	φ		0.114		
			(0.047)		
Overconfidence	α_v			0.668	
				(0.249)	
Perceived persistence	$\widehat{\rho}$				0.594
					(0.133)
First order autocorrelation	ρ_1				0.429
					(0.180)
Second order autocorrelation	ρ_2				0.566
					(0.197)
Panel B: Model Selection					
Log likelihood		-7486.7	-7421.4	-7453.6	-7105.5
AIC		14983	14855	14919	14225
BIC		15010	14887	14952	14263

Note: Panel A reports parameters estimates. Column (1) denotes the rational expectations model, coumn (2) corresponds to the diagnostic expectations model, column (3) reports results for the overconfidence model, and column (4) reports results for the misspecified expectations model. Bootstrapped standard errors are reported in parentheses. Panel B reports the maximized log likelihood as well as AIC and BIC for each model.

negative across all models, consistent with Gemmi and Valchev (2021). Strategic substitutability is able to generate error and revision predictability. As a result, these estimates suggest a more moderate role for the biases as indicated by the smaller estimated magnitudes of diagnosticity, overconfidence, and overextrapolation. Nonetheless, my findings are qualitatively unchanged from the baseline estimates.

5 Why Does Misspecification Outperform the Alternatives?

Of the different models considered, misspecified expectations is the one that is best able to match the updating dynamics in the data while also reproducing unconditional overreactions. I begin by demonstrating that all three models are able to generate observed overreactions, which the baseline rational expectations model cannot do. I then examine the updating rules across the models and compare them with observed updating behavior in the data.

5.1 Overreaction

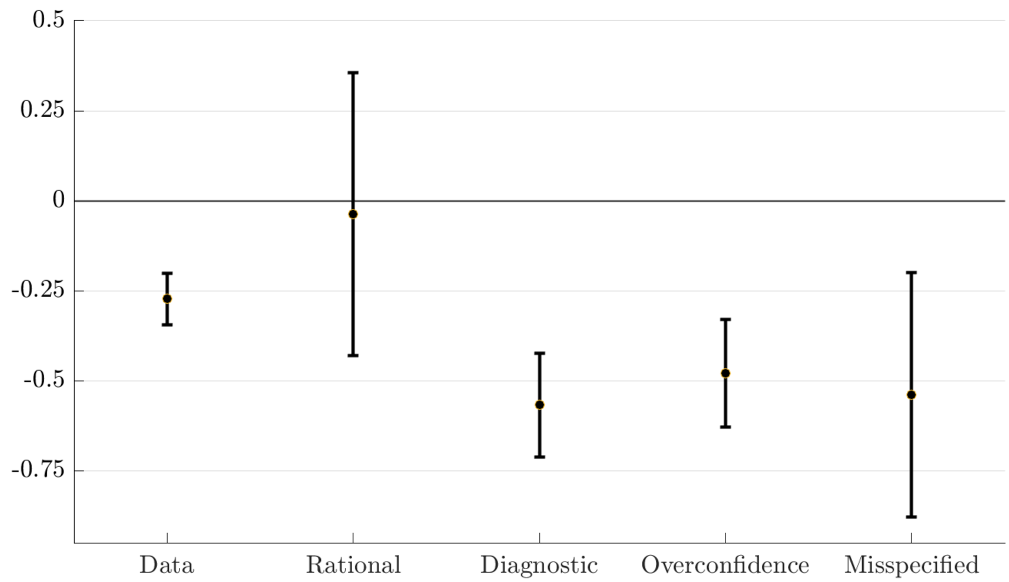
Bordalo et al. (2020) provides evidence of overreactions in macroeconomic expectations by running the Coibion and Gorodnichenko (2015) errors-on-revisions regression at the forecaster level. This testable prediction has been studied extensively in the literature (Kohlhas and Walther, 2021; Broer and Kohlhas, 2019; Burgi and Ortiz, 2022). I run the same regression using simulated panel data from each of the models. The regression specification is:

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

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 an overreaction to new information.

Figure 1 displays 95% confidence intervals for the β_1 coefficient in (3) simulated across each of the models, along with the empirical estimate from the data. As expected, the rational expectations

Figure 1: Overreaction in One Quarter Ahead Expectations



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

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 robust negative β_1 coefficients. Furthermore, these simulated coefficients reside within the 95% confidence interval of the empirical estimate for the overconfidence and misspecification models, while the estimated diagnostic expectations model delivers stronger-than-observed overreactions.

Beyond the unconditional correlations examined here, delayed overshooting, a form of conditional overreaction, has been documented in Angeletos et al. (2020) and Bianchi et al. (2021). Delayed overshooting can be observed by inspecting the consensus forecast error impulse response to a shock. If the impulse response function switches signs, then beliefs are said to exhibit 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 overshooting). As a result, I do not focus on these dynamics since they are less informative for understanding why misspecified expectations outperforms the alternatives. Instead, I proceed to examine updating rules across the different models.

5.2 Updating Rules

Based on the common information structure assumed across all of the models, forecasters update their predictions as follows:

$$\widehat{x}^i_{t|t} = \widehat{x}^i_{t|t-1} + \kappa_1(y^i_t - \widehat{x}^i_{t|t-1}) + \kappa_2(x_{t-1} - \widehat{x}^i_{t-1|t-1}).$$

This updating equation implies that forecasters place some weight on contemporaneous news, κ_1, past news, κ_2, and their prior, $(1-\kappa_1)\rho - \kappa_2$. We can express the updated forecast as a function of observables and shocks:[7]

$$\widehat{x}^i_{t|t} = \big[(1-\kappa_1)\rho - \kappa_2\big]\widehat{x}^i_{t-1|t-1} + \kappa_1 x_t + \kappa_2 x_{t-1} - \kappa_1 e_t + \kappa_1 v^i_t,$$

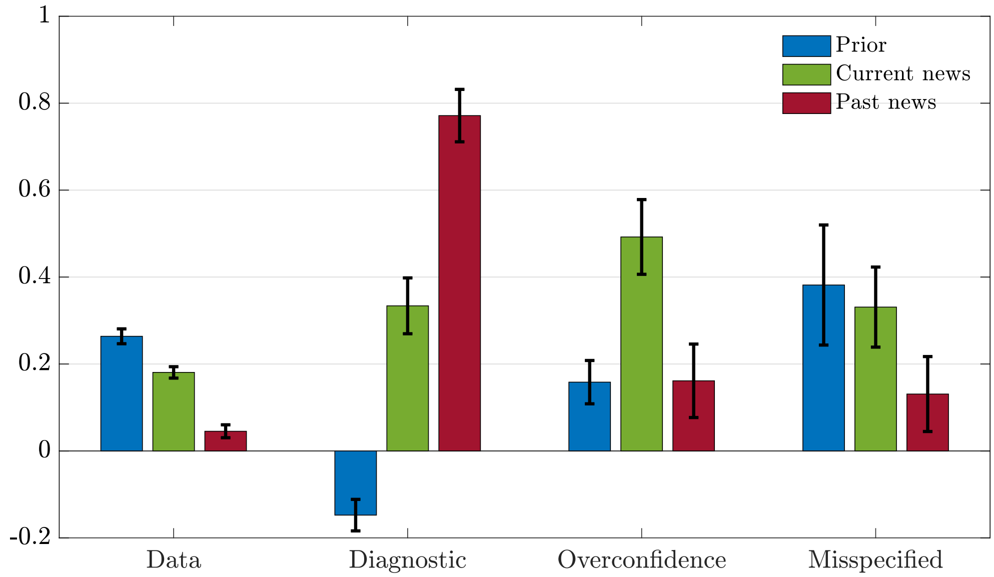
Each model considered generates overreactions in a different way. Overconfidence emanates from an erroneous belief that one's private information is more precise than it truly is. As a result, the bias enters in through κ_1. Under diagnostic expectations, agents overreact to all news so that κ_1 and κ_2 are both distorted. Finally, under misspecified expectations, the misperceived persistence enters into the weights κ_1 and κ_2, and also maps directly into the the importance placed on the prior through the persistence parameter. As a result, a regression of the current-quarter forecast on its lag, the contemporaneous outcome, and the lag of the contemporaneous outcome will deliver distinct reduced form coefficients across the three models.

I run the following regression in the data and also simulate it in the three models.[8] The results

[7] Under misspecified expectations the persistence parameter becomes the perceived persistence, $\widehat{\rho}$, rather than the objective persistence, ρ.

[8] Note that the coefficient in front of x_t will be biased downward from the perspective of the model due to the unobserved shock e_t. Addressing this bias, however, is not of first-order importance since I run the same biased regression in the data and across the three models.

Figure 2: Estimated Updating Coefficients



Note: The figure plots 95% confidence intervals of empirical and simulated sensitivites of updated forecasts to the prior forecast, current news, and past news based on the regression (4). Each model is simulated based on the estimates reported in Table 1.

are plotted in Figure 2.

$$\widehat{x}^i_{t|t} = \beta_0 + \beta_1 \widehat{x}^i_{t-1|t-1} + \beta_2 x_t + \beta_3 x_{t-1} + \omega_{it}. \tag{4}$$

Empirically, forecasters place relatively more importance on current news relative to past news as indicated by the second and third bars reflecting the estimated coefficients in the data in Figure 2. In addition, forecasts are positively autocorrelated with an estimated persistence of about 0.26.

Comparing these empirical estimates with the model-based estimates, we see that misspecification can best fit the updating patterns in the data. Under diagnostic expectations, forecasts are negatively autocorrelated due to the high amount of estimated diagnosticity. In addition, diagnostic forecasters place a counterfactually large amount of importance of past news when updating their expectations. On the other hand, overconfident forecasters place a counterfactually large amount of importance of current news. Since the only channel through which overconfident forecasters can overreact is through private news, the overconfident Kalman gain, $\widehat{\kappa}_1$, is relatively high which is reflected in the estimate for β_2. In addition, overconfident forecasters place equal importance to past news and their priors, which is not supported in the data. Finally, under misspecified expectations, forecasters place less importance on past news relative to current news or to their prior. Furthermore, based on the point estimates, forecasters place slightly more weight on their prior, thereby qualitatively delivering the patterns observed in the data.

While the set of models considered here is restricted to diagnostic expectations, overconfidence, and misspecification, the implications of my finding extend beyond this menu of models. In particular, a successful non-FIRE model should be able to deliver the patterns revealed in the data: persistent forecasts with relatively more importance placed on priors followed by current news and then past news ($\widehat{\beta}_1 > \widehat{\beta}_2 > \widehat{\beta}_3$).

6 Estimating Misspecification-Related Extrapolation

The misspecified expectations model delivers a testable implication which one can use to estimate overextrapolation via a simple regression. Assuming that the underlying data generating process follows an AR(2). The one period ahead forecast error is:[9]

$$\begin{aligned} x_{t+1} - \widehat{x}^i_{t+1|t} &= (1-\kappa_1)\rho_1(x_t - \widehat{x}^i_{t|t-1}) - \kappa_2\rho_1(x_{t-1} - \widehat{x}^i_{t-1|t-1}) \\ &\quad + (\rho_1 - \widehat{\rho})\widehat{x}^i_{t|t} + \rho_1(\kappa_1 - 1)e_t + \rho_2 s_{t-2} + w_{t+1} + e_{t+1} - \kappa_1\rho_1 v^i_t. \end{aligned}$$

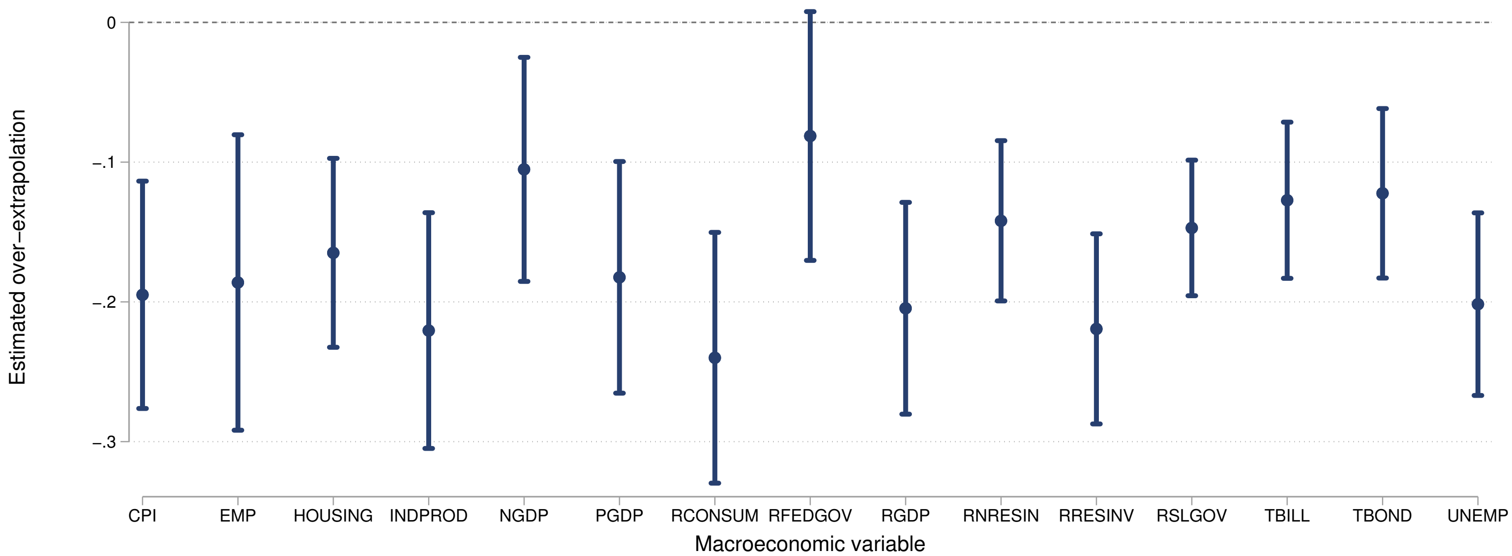
The gap between the objective autocorrelation coefficient, ρ_1, and the perceived autocorrelation, $\widehat{\rho}$, summarizes the extent of the extrapolation under misspecified expectations. As a result, we can estimate the following equation via OLS:

$$x_{t+1} - \widehat{x}^i_{t+1|t} = \beta_1(x_t - \widehat{x}^i_{t|t-1}) + \beta_2(x_{t-1} - \widehat{x}^i_{t-1|t-1}) + \beta_3\widehat{x}^i_{t|t} + \lambda_t + \varepsilon_{it}. \tag{5}$$

In words, we can regress the one period ahead forecast error on its lag, the lagged current period

[9] One could extend the data generating process to an arbitrary AR(p) and derive a similar testable prediction which would feature additional lags of the forecast errors.

Figure 3: Estimated Overextrapolation



Note: The figure plots 95% confidence intervals of $\widehat{\beta}_3$ from regression (5), which specify Driscoll and Kraay (1998) standard errors. The variable abreviations are defined in Appendix A as well as the SPF documentation.

error, and the current quarter forecast. The coefficient β_3 maps to $(\rho_1 - \widehat{\rho})$, meaning that under overextrapolation we would expect to find $\beta_3 < 0$.

Figure 3 plots the 95% confidence intervals of β_3 for 15 macroeconomic variables in the SPF. The β_3 estimate for real GDP is roughly -0.20. The estimated baseline model results reported in Table 1 imply a slightly larger extent of overextrapolation, though the 95% confidence intervals for the estimates overlap. Similarly, for inflation as measured by CPI, the results in Table D3 indicate that overextrapolation lies within the 95% interval of the reduced form estimate in Figure 3. The estimated β_3 coefficient is fairly stable, hovering between -0.10 and -0.20 for most variables.

7 Conclusion

At present, a host of non-FIRE theories exist in the literature. As mentioned in Reis (2020), however, there is little agreement on which theory could serve as a suitable non-FIRE benchmark. This paper offers an answer to this question: misspecified expectations can serve as a benchmark theory of non-FIRE expectations. Survey data from the SPF provide strong support for this finding. Misspecified expectations, which can generate overreaction and realistic updating rules, can be motivated from behavioral and non-behavioral foundations alike. Researchers interested in approximating belief dynamics in macroeconomic models can adopt misspecified expectations to do so. Embedding the misspecification described here into a macroeconomic model merely requires introducing two parameters into an otherwise standard model: the second order autocorrelation coefficient for the shock process, ρ_2, and a misspecification parameter, $\widehat{\rho}$. A promising avenue for future research could be to more closely examine the microfoundations of misspecification and whether they are due to inherent biases or optimizing behavior.

References

Afrouzi, Hassan, Spencer Kwon, Augustin Landier, Yueran Ma, and David Thesmar (2021), "Over-reaction in Expectations: Evidence and Theory." Working Paper.

Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry (2020), "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, Cosmin Ilut, and Hikaru Saijo (2021), "Diagnostic Business Cycles." NBER Working Paper 28604.

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

Bordalo, Pedro, Nicola Gennaioli, Andrei Shleifer, and Stephen Terry (2021), "Real Credit Cycles." NBER Working Paper 28416.

Broer, Tobias and Alexandre Kohlhas (2019), "Forecaster (Mis-)Behavior." Working Paper.

Burgi, Constantin and Julio Ortiz (2022), "Overreaction Through Expectation Smoothing."

Cagan, Phillip (1956), "The Monetary Dynamics of Hyperinflation." In *Studies in the Quantity Theory of Money*, 25–117, University of Chicago Press.

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.

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.

Farmer, Leland, Emi Nakamura, and Jon Steinsson (2021), "Learning About the Long Run."

Fuster, Andreas, Benjamin Hebert, and David Laibson (2012), "Natural Expectations, Macroeconomic Dynamics, and Asset Pricing." *NBER Macroeconomics Annual*, 26, 1–48.

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.

Gemmi, Luca and Rosen Valchev (2021), "Biased Surveys." Working Paper.

Goodwin, Richard M (1947), "Dynamical Coupling with Especial Reference to Markets Having Production Lags." *Econometrica*, 15, 181–204.

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

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), "SimpleModels 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.

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

Nordhaus, William D. and Steven N. Durlauf (1984), "Empirical Tests of the Rationality of Economic Forecasters: A Fixed Horizons Approach." *Cowles Foundation Discussion Papers. 955.*

Ottaviani, Marco and Peter Sørensen (2006), "The Strategy of Professional Forecasting." *Journal of Financial Economics*, 2, 441–466.

Reis, Ricardo (2020), "Comment on 'Imperfect Expectations: Theory and Evidence'." *NBER Macroeconomics Annual 2020*, 35, 99–111.

Rozsypal, Filip and Kathrin Schlafmann (2019), "Overpersistence Bias in Individual Income Expectations and its Aggregate Implications." 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.

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 Empirics

A.1 Sample

I obtain my 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. To construct the sample, I first transform the raw data from levels to annualized growth rates:

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

Because the underlying driving process has changed since the late 1960s, I narrow my focus to quarterly forecasts from 1992 to 2019. Moreover, the estimation procedure requires that forecasters contain unbroken sequences of observations. For this reason, I only keep each forecaster's longest spell of reported forecasts. Lastly, I require that forecasters be observed in the data for at least five quarters. Summary statistics of the cleaned sample are reported in Table A1.

Table A1: SPF Summary Statistics

	Mean	Median	Standard Deviation	25%	75%
Current-quarter forecast error	0.043	0.050	1.551	-1.006	1.000
Current-quarter forecast revision	-0.203	-0.102	1.178	-0.704	0.400
Current-quarter forecast	2.343	2.478	1.478	1.800	3.101
Real time real GDP growth	2.386	2.502	1.761	1.537	3.473
One-quarter ahead forecast error	-0.281	-0.188	1.979	-1.312	0.800
One-quarter ahead forecast revision	-0.110	-0.030	0.866	-0.412	0.272
One-quarter ahead consensus error	-0.105	-0.116	1.902	-1.086	0.778

Note: The table reports summary statistics for the relevant variables in estimation in the main text. The sample is constructed from the SPF. The current-quarter sample spans 1992-2019, with 118 unique forecasters and 1843 forecaster-date observations.

A.2 Macroeconomic Variable Definitions

Figure 3 estimates the degree of overextrapolation for a range of SPF variables. Below I detail the abbreviations used in the figure:

- CPI – Consumer price index
- EMP – Nonfarm payroll employment
- HOUSING – Housing starts
- INDPROD – Industrial production
- NGDP – Nominal GDP
- PGDP – GDP deflator
- RCONSUM – Real personal consumption expenditures
- RFEDGOV – Real federal government consumption and gross investment
- RGDP – Real GDP
- RNRESIN – Real business fixed investment
- RRESINV – Real residential fixed investment
- RSLGOV – Real state & local government consumption and gross investment
- TBILL – 3-month Treasury bill
- TBOND – 10-year Treasury bond
- UNEMP – Unemployment rate

Appendix B Model

In this section, I provide further detail on the extension of the baseline set of models to feature strategic interactions. I then derive the testable prediction estimated in Section 6 of the main text.

B.1 Strategic Interaction

In models of strategic interaction, forecasters have the motive to either mimic or to deviate from the consensus forecast. The forecaster's loss function is:

$$\min_{\widehat{x}^i_{t|t}} (x_t - \widehat{x}^i_{t|t})^2 + \gamma(\widehat{x}^i_{t|t} - \widehat{x}_{t|t})^2.$$

The first order condition implies

$$\widehat{x}^i_{t|t} = \frac{1}{1+\gamma}\mathbb{E}_{it}(x_t) + \frac{\gamma}{1+\gamma}\mathbb{E}_{it}(x_{t|t}),$$

which means the forecaster sets her reported forecast to be a combination of the the conditional expectation as well as the expectation over what the consensus forecast will be.

The presence of strategic incentives makes higher order beliefs crucial to this model. In particular, the consensus forecast at time t is denoted by F_t and it is defined as

$$F_t = \frac{1}{1+\gamma}\sum_{k=0}^{\infty}\left(\frac{\gamma}{1+\gamma}\right)^k \mathbb{E}^{(k)}(x_t) = \frac{1}{1+\gamma}x_{t|t} + \frac{\gamma}{1+\gamma}F_{t|t}.$$

where $\mathbb{E}^{(k)}$ is the k^{th}-order expectation of x_t. I guess and verify the following law of motion for F_t:

$$F_t = c_1 s_{t-1} + c_2 F_{t-1} + c_3 w_t + c_4 e_{t-1},$$

where $c_1 = c_3\rho + c_4$, $c_2 = (1-\kappa_{11})\rho - \kappa_{12} + \gamma(c_1 - \kappa_{31}\rho)$, $c_3 = \frac{\kappa_{11}+\gamma\kappa_{31}}{1+\gamma}$, and $c_4 = \frac{\kappa_{12}+\gamma\kappa_{32}}{1+\gamma}$.

From the Kalman filter (shown in the next section), this implies the following:

$$
\begin{aligned}
x^i_{t|t} &= x^i_{t|t-1} + \kappa_{11}(y^i_t - x^i_{t|t-1}) + \kappa_{12}(x_{t-1} - x^i_{t-1|t-1}) \\
x^i_{t+1|t} &= \rho x^i_{t|t} \\
F^i_{t|t} &= (\kappa_{31}\rho + \kappa_{32})s_{t-1} + \kappa_{31}w_t + \kappa_{32}e_{t-1} + (c_1 - \kappa_{31}\rho)s^i_{t-1|t-1} + c_2 F^i_{t-1|t-1} \\
F^i_{t+1|t} &= \frac{\rho + \gamma c_1}{1+\gamma} s^i_{t|t} + \frac{\gamma c_2}{1+\gamma} F^i_{t|t}
\end{aligned}
$$

Given these forecasting and updating rules, I cast this model into state space form and estimate the five parameters: $\rho, \sigma_w, \sigma_e, \sigma_v, \gamma$.

Deriving Coefficients c_1, c_2, c_3, and c_4

The reported forecast is:

$$\widehat{x}^i_{t|t} = \frac{1}{1+\gamma} x^i_{t|t} + \frac{\gamma}{1+\gamma} F^i_{t|t}.$$

Averaging across reported forecasts, we obtain

$$F_t = \frac{1}{1+\gamma} x_{t|t} + \frac{\gamma}{1+\gamma} F_{t|t}.$$

I next guess and verify a law of motion for F_t:

$$F_t = c_1 s_{t-1} + c_2 F_{t-1} + c_3 w_t + c_4 e_{t-1}.$$

This implies that the forecaster solves the following filtering problem:

$$\begin{bmatrix} s_t \\ s_{t-1} \\ F_t \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 \\ 1 & 0 & 0 \\ c_1 & 0 & c_2 \end{bmatrix} \begin{bmatrix} s_{t-1} \\ s_{t-2} \\ F_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ c_3 & c_4 \end{bmatrix} \begin{bmatrix} w_t \\ e_{t-1} \end{bmatrix}$$

$$\begin{bmatrix} y_t^i \\ x_{t-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} s_t \\ s_{t-1} \\ F_t \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_t^i \\ e_{t-1} \end{bmatrix}$$

Through the Kalman filtering equations, we can express the nowcast for the state vector as

$$\mathbf{x}_{t|t}^i = \mathbf{s}_{t|t}^i = (\mathbf{I} - \kappa\mathbf{C})\mathbf{s}_{t|t-1}^i + \kappa\mathbf{z_t^i},$$

where κ is the Kalman gain matrix, $\mathbf{C}$ is the matrix that maps the state into the observation vector, and $\mathbf{z}$ is the vector of observations. Aggregating this expression to the consensus-level, and using $x_{t|t}^i = s_{t|t}^i$, we have

$$x_{t|t} = \big[(1-\kappa_{11})\rho - \kappa_{12}\big]x_{t-1|t-1} + \kappa_{11}s_t + \kappa_{12}x_{t-1}$$
$$F_{t|t} = (c_1 - \kappa_{31}\rho - \kappa_{32})x_{t-1|t-1} + c_2F_{t-1|t-1} + \kappa_{31}s_t + \kappa_{32}x_{t-1}$$

Recall that the consensus forecast is $F_t = \frac{1}{1+\gamma}x_{t|t} + \frac{\gamma}{1+\gamma}F_{t|t}$. Substituting in the above expressions for $x_{t|t}$ and $F_{t|t}$:

$$\begin{aligned} F_t &= \frac{1}{1+\gamma}\Big\{\big[(1-\kappa_{11})\rho - \kappa_{12}\big]x_{t-1|t-1} + \kappa_{11}s_t + \kappa_{12}x_{t-1}\Big\} \\ &+ \frac{\gamma}{1+\gamma}\Big\{(c_1 - \kappa_{31}\rho - \kappa_{32})x_{t-1|t-1} + c_2F_{t-1|t-1} + \kappa_{31}s_t + \kappa_{32}x_{t-1}\Big\} \\ &= \frac{\big[(1-\kappa_{11})\rho - \kappa_{12}\big] + \gamma(c_1 - \kappa_{31}\rho - \kappa_{32})}{1+\gamma}x_{t-1|t-1} + \frac{\kappa_{11} + \gamma\kappa_{31}}{1+\gamma}s_t + \frac{\kappa_{12} + \gamma\kappa_{32}}{1+\gamma}x_{t-1} + \frac{c_2\gamma}{1+\gamma}F_{t-1|t-1} \end{aligned}$$

Note that $F_{t-1} = \frac{1}{1+\gamma}x_{t-1|t-1} + \frac{\gamma}{1+\gamma}F_{t-1|t-1}$. Rearranging, we have $F_{t-1|t-1} = \frac{1+\gamma}{\gamma}F_t - \frac{1}{\gamma}x_{t-1|t-1}$.

Substituting this expression for $F_{t-1|t-1}$ into the longer equation above, we can express F_t as a function of $x_{t-1|t-1}$, s_t, x_{t-1}, and F_{t-1}.

$$\begin{aligned}
F_t &= \frac{[(1-\kappa_{11})\rho - \kappa_{12}] + \gamma(c_1 - \kappa_{31}\rho - \kappa_{32})}{1+\gamma} x_{t-1|t-1} + \frac{\kappa_{11} + \gamma\kappa_{31}}{1+\gamma} s_t + \frac{\kappa_{12} + \gamma\kappa_{32}}{1+\gamma} x_{t-1} \\
&+ \frac{c_2\gamma}{1+\gamma}\left[\frac{1+\gamma}{\gamma} F_{t-1} - \frac{1}{\gamma} x_{t-1|t-1}\right] \\
&= \frac{[(1-\kappa_{11})\rho - \kappa_{12}] + \gamma(c_1 - \kappa_{31}\rho - \kappa_{32}) - c_2}{1+\gamma} x_{t-1|t-1} + \frac{\kappa_{11} + \gamma\kappa_{31}}{1+\gamma} s_t + \frac{\kappa_{12} + \gamma\kappa_{32}}{1+\gamma} x_{t-1} + c_2 F_{t-1}
\end{aligned}$$

Let $c_2 = (1-\kappa_{11})\rho + \gamma(c_1 - \kappa_{31}\rho - \kappa_{32})$. Expanding and simplifying the expressions for s_t and x_{t-1}, we obtain the consensus forecast as a function of the previous consensus forecast, the previous value of the latent state, the current innovation to the latent state, and the previous value of the transitory component of x_t:

$$F_t = \frac{(\kappa_{11} + \gamma\kappa_{31})\rho + (\kappa_{12} + \gamma\kappa_{32})}{1+\gamma} s_{t-1} + \frac{\kappa_{11} + \gamma\kappa_{31}}{1+\gamma} w_t + \frac{\kappa_{12} + \gamma\kappa_{32}}{1+\gamma} e_{t-1} + c_2 F_{t-1}.$$

Defining $c_3 = \frac{\kappa_{11}+\gamma\kappa_{31}}{1+\gamma}$, $c_4 = \frac{\kappa_{12}+\gamma\kappa_{32}}{1+\gamma}$, and $c_1 = (c_3\rho + c_4)$, we recover

$$F_t = c_1 s_{t-1} + c_2 F_{t-1} + c_3 w_t + c_4 e_{t-1},$$

as desired.

Note that the forecaster's filtering problem includes the coefficients c_1, c_2, c_3 and c_4 in the transition matrix, and these coefficients are, themselves, functions of the Kalman gains. As a result, the estimation of the models under strategic incentives requires solving a fixed point problem in each iteration. For a given parameter guess, I initiate a guess for the coefficients and run the filter for 1,000 periods. Based on the implied steady state Kalman gain matrix, κ, I update my guess for the coefficients and continue until the old and updated coefficients converge (up to a tolerance of 10^{-6}). With these coefficients and Kalman gain weights in hand, I proceed to construct the likelihood function.

Strategic Interaction with Diagnostic Expectations

Under diagnostic expectations, the coefficients are slightly different. Note that the diagnostic Kalman filter implies the following updating equation:

$$\mathbf{s}^i_{t|t} = \mathbf{s}^i_{t|t} + \varphi\kappa\big(\mathbf{z}^i_t - \mathbf{C}\mathbf{s}^i_{t|t-1}\big),$$

Following the steps outlined above, we can derive the law of motion for the diagnostic consensus forecast, $F^{\varphi}_t = c_1 s_{t-1} + c_2 F^{\varphi}_{t-1} + c_3 w_t + c_4 e_{t-1}$, where

$$\begin{aligned}
c_1 &= c_3\rho + c_4 \\
c_2 &= (\rho - (1+\varphi)(\kappa_{11}\rho + \kappa_{12})) + \gamma(c_1 - (1+\varphi)(\kappa_{31}\rho + \kappa_{32}) \\
c_3 &= (1+\varphi)\frac{\kappa_{11} + \gamma\kappa_{31}}{1+\gamma} \\
c_4 &= (1+\varphi)\frac{\kappa_{12} + \gamma\kappa_{32}}{1+\gamma}
\end{aligned}$$

Strategic Interaction with Misspecified Expectations

Under misspecified expectations, the coefficients are again different. In this case, it becomes important to distinguish the objective driving process form the perceived driving process. Given the assumed AR(2) dynamics of the objective driving process, I guess and verify:

$$F_t = c_1 s_{t-1} + c_2 F_{t-1} + c_3 w_t + c_4 e_{t-1} + c_5 s_{t-2},$$

where I introduce a new coefficient, c_5 because consensus beliefs will now also depend on the second lag of the latent state. The forecaster's filtering problem, however, remains unchanged. In other words, forecasters not only misspecify the underlying data generating process, but they also misspecify the law of motion of the aggregate belief. In both cases, the forecaster neglects the second lag s_{t-2}. As a result, the persistence of the consensus forecast, which is summarized by c_2, will depend on $\widehat{\rho}$. In addition, all of the other coefficients will be functions of the elements of a misspecified Kalman gain matrix.

I verify this new guess using the very same steps outlined above. In the final step, we obtain:

$$F_t = \frac{(\kappa_{11} + \gamma\kappa_{31})\rho_1 + (\kappa_{12} + \gamma\kappa_{32})}{1+\gamma} s_{t-1} + \frac{\kappa_{11} + \gamma\kappa_{31}\rho_2}{1+\gamma} s_{t-2} + \frac{\kappa_{11} + \gamma\kappa_{31}}{1+\gamma} w_t + \frac{\kappa_{12} + \gamma\kappa_{32}}{1+\gamma} e_{t-1} + c_2 F_{t-1},$$

which implies:

$$\begin{aligned}
c_1 &= \frac{(\kappa_{11} + \gamma\kappa_{31})\rho_1 + (\kappa_{12} + \gamma\kappa_{32})}{1+\gamma} \\
c_2 &= (1-\kappa_{11})\widehat{\rho} + \gamma(c_1 - \kappa_{31}\widehat{\rho}) \\
c_3 &= \frac{\kappa_{11} + \gamma\kappa_{31}}{1+\gamma} \\
c_4 &= \frac{\kappa_{12} + \gamma\kappa_{32}}{1+\gamma} \\
c_5 &= \frac{\kappa_{11} + \gamma\kappa_{31}\rho_2}{1+\gamma}
\end{aligned}$$

B.2 Deriving the Testable Implication in Section 6

The one-quarter ahead forecast error is

$$x_{t+1} - \widehat{x}^i_{t+1|t} = s_{t+1} + e_{t+1} - \widehat{\rho}\widehat{x}^i_{t|t}.$$

With some algebra, we re-write the one-quarter ahead error as a function of the current quarter error, the current quarter forecast, as well as a set of time-varying unobservables:

$$\begin{aligned}
x_{t+1} - \widehat{x}^i_{t+1|t} &= \rho_1 s_t + \rho_2 s_{t-1} + w_{t+1} + e_{t+1} - \widehat{\rho}\widehat{x}^i_{t|t} \pm \rho_1 e_t \\
&= \rho_1 x_t - \rho_1 e_t + \rho_2 s_{t-2} + w_{t+1} + e_{t+1} - \widehat{\rho}\widehat{x}^i_{t|t} \pm \rho_1 \widehat{x}^i_{t|t} \\
&= \rho_1 (x_t - \widehat{x}^i_{t|t}) - \rho_1 e_t + \rho_2 s_{t-2} + w_{t+1} + e_{t+1} + (\rho_1 - \widehat{\rho})\widehat{x}^i_{t|t}.
\end{aligned}$$

Expanding the first term and simplifying, we obtain:

$$
\begin{aligned}
x_{t+1} - \widehat{x}^i_{t+1|t} &= \rho_1\big(x_t - \widehat{x}^i_{t|t-1} - \kappa_1(y^i_t - \widehat{x}^i_{t|t-1}) - \kappa_2(x_{t-1} - \widehat{x}^i_{t-1|t-1})\big) - \rho_1 e_t + \rho_2 s_{t-2} \\
&+ w_{t+1} + e_{t+1} + (\rho_1 - \widehat{\rho})\widehat{x}^i_{t|t} \\
&= \rho_1\big[(x_t - \widehat{x}^i_{t|t-1}) - \kappa_1(x_t - \widehat{x}^i_{t|t-1}) - \kappa_1 v^i_t + \kappa_2 e_t - \kappa_2(x_{t-1} - \widehat{x}^i_{t-1|t-1})\big] \\
&- \rho_1 e_t + \rho_2 s_{t-2} + w_{t+1} + e_{t+1} + (\rho_1 - \widehat{\rho})\widehat{x}^i_{t|t} \\
&= (1-\kappa_1)\rho_1(x_t - \widehat{x}^i_{t|t-1}) - \kappa_1\rho_1 v^i_t + \rho_1(\kappa_1 - 1)e_t \\
&- \kappa_2\rho_1(x_{t-1} - \widehat{x}^i_{t-1|t-1}) + \rho_2 s_{t-2} + w_{t+1} + e_{t+1} + (\rho_1 - \widehat{\rho})\widehat{x}^i_{t|t},
\end{aligned}
$$

as desired.

Appendix C Estimation

In this section I detail the steps taken to estimate the models via MLE. I then report comparative statics to provide some evidence of identification.

C.1 Maximum Likelihood Estimation

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

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

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, s_t and its innovations, w_t, as well as the transitory component e_t. In addition, this vector includes the unobserved Bayesian forecasts, $s^i_{t|t}$ and $s^i_{t|t-1}$ 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 public and private noise: e_t and 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 five quarters for which a forecaster must be observed in order to be included. My findings are robust to shorter and longer spell lengths.

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

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

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:

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

where: $\mathbf{z_t} = \begin{pmatrix} \mathbf{z_{1t}} \\ \mathbf{z_{2t}} \\ \vdots \\ \mathbf{z_{nt}} \end{pmatrix}$, $\mathbf{z} = \begin{pmatrix} z_{11} \\ z_{12} \\ z_{13} \\ \vdots \\ z_{n1} \\ z_{n2} \\ z_{n3} \end{pmatrix}$, 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

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

where

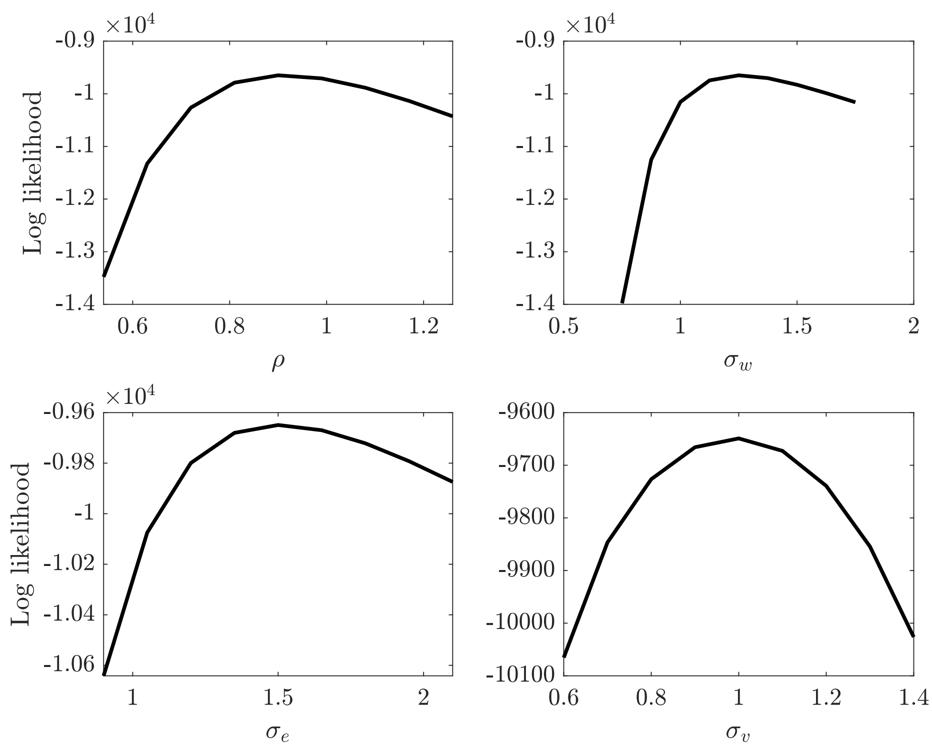
$$\mathbf{S_{yy}} = \sum_t \mathbf{y_t^{*'}} \mathbf{F_t^{-1}} \mathbf{y_t^{*}}.$$

I estimate the model by constructing and maximizing the likelihood function numerically, with judicious use of sparse matrices.

C.2 Comparative Statics

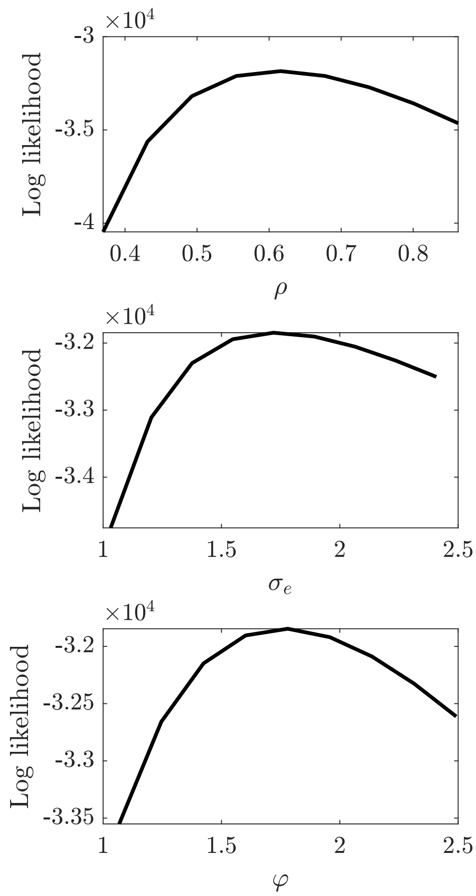
In this section, I complete several comparative statics exercises. For each model (rational expectation, diagnostic expectations, overconfidence, and misspecified expectations) I plot the likelihood functions across a range of parameter values, varying one parameter at a time.

Figure C1: Log Likelihood Across Parameter Values (RE)



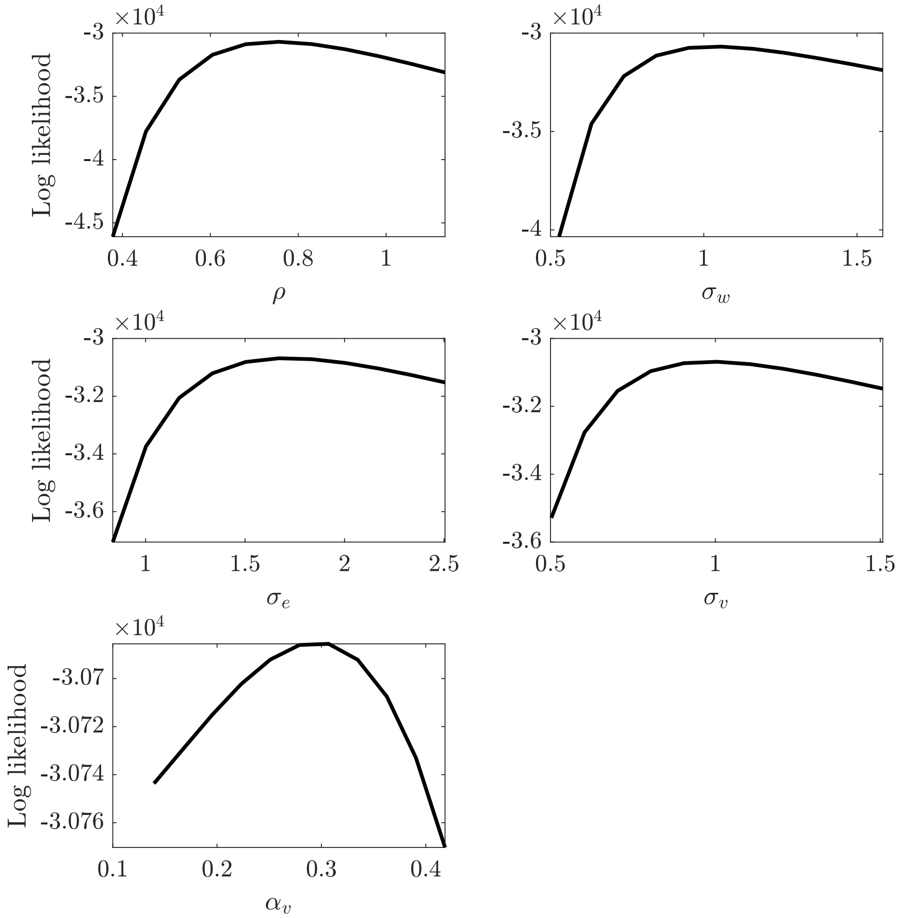
Note: Each panel plots the log likelihood function for a range of parameter values, holding the other parameters fixed.

Figure C2: Log Likelihood Across Parameter Values (DE)



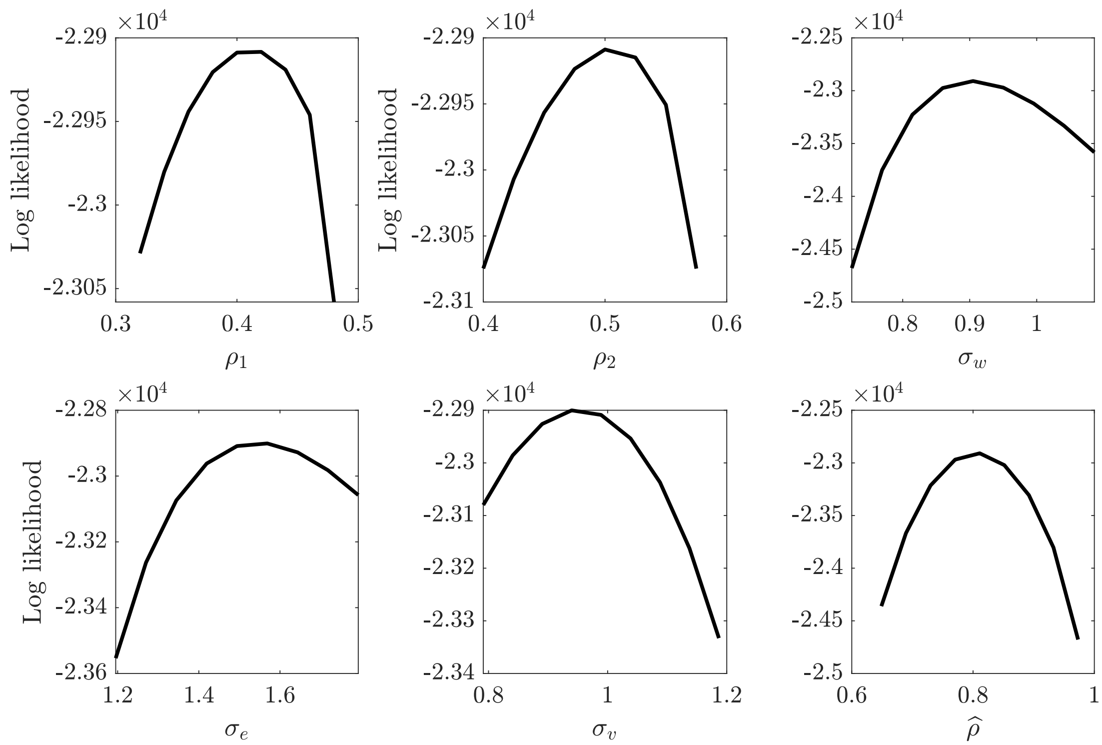
Note: Each panel plots the log likelihood function for a range of parameter values, holding the other parameters fixed.

Figure C3: Log Likelihood Across Parameter Values (OC)



Note: Each panel plots the log likelihood function for a range of parameter values, holding the other parameters fixed.

Figure C4: Log Likelihood Across Parameter Values (ME)



Note: Each panel plots the log likelihood function for a range of parameter values, holding the other parameters fixed.

C.3 State Space Specifications for a Single Forecaster

Rational Expectations

State:

$$\begin{bmatrix} s_t \\ w_{t+1} \\ e_{t+1} \\ e_t \\ s^i_{t|t} \\ s^i_{t|t-1} \\ s_{t|t} \end{bmatrix} = \begin{bmatrix} \rho & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \kappa_1\rho+\kappa_2 & \kappa_1 & 0 & \kappa_2 & (1-\kappa_1)\rho-\kappa_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho & 0 & 0 \\ \kappa_1\rho+\kappa_2 & \kappa_1 & 0 & \kappa_2 & 0 & 0 & (1-\kappa_1)\rho-\kappa_2 \end{bmatrix} \begin{bmatrix} s_{t-1} \\ w_t \\ e_t \\ e_{t-1} \\ s^i_{t-1|t-1} \\ s^i_{t-1|t-2} \\ s_{t-1|t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \kappa_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_{t+1} \\ e_{t+1} \\ v^i_t \end{bmatrix}$$

Measurement:

$$\begin{bmatrix} x_{t+1}-\widehat{x}^i_{t+1|t} \\ \widehat{x}^i_{t+1|t}-\widehat{x}^i_{t+1|t-1} \\ x_{t+1}-x_{t+1|t} \end{bmatrix} = \begin{bmatrix} \rho & 1 & 1 & 0 & -\rho & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho & -\rho & 0 \\ \rho & 1 & 1 & 0 & 0 & 0 & -\rho \end{bmatrix} \begin{bmatrix} s_t \\ w_{t+1} \\ e_{t+1} \\ e_t \\ s^i_{t|t} \\ s^i_{t|t-1} \\ s_{t|t} \end{bmatrix}$$

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

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

Diagnostic Expectations

State:

$$\begin{bmatrix} s_t \\ w_{t+1} \\ e_{t+1} \\ e_t \\ s^i_{t|t} \\ s^i_{t|t-1} \\ s^i_{t-1|t-2} \\ s_{t|t} \\ s_{t|t-1} \\ s_{t-1|t-2} \end{bmatrix} = \begin{bmatrix} \rho & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \kappa_1\rho+\kappa_2 & \kappa_1 & 0 & \kappa_2 & (1-\kappa_1)\rho-\kappa_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \kappa_1\rho+\kappa_2 & \kappa_1 & 0 & \kappa_2 & 0 & 0 & 0 & (1-\kappa_1)\rho-\kappa_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} s_{t-1} \\ w_t \\ e_t \\ e_{t-1} \\ s^i_{t-1|t-1} \\ s^i_{t-1|t-2} \\ s^i_{t-2|t-3} \\ s_{t-1|t-1} \\ s_{t-1|t-2} \\ s_{t-2|t-3} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \kappa_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_{t+1} \\ e_{t+1} \\ v^i_t \end{bmatrix}$$

Measurement:

$$\begin{bmatrix} x_{t+1} - \widehat{x}^i_{t+1|t} \\ \widehat{x}^i_{t+1|t} - \widehat{x}_{t+1|t-1} \\ x_{t+1} - \widehat{x}_{t+1|t} \end{bmatrix} = \begin{bmatrix} \rho & 1 & 1 & 0 & -\rho(1+\varphi) & \rho\varphi & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho(1+\varphi) & -\rho^2(1+2\varphi) & \rho^2\varphi & 0 & 0 & 0 \\ \rho & 1 & 1 & 0 & 0 & 0 & 0 & -\rho(1+\varphi) & \rho\varphi & 0 \end{bmatrix} \begin{bmatrix} s_t \\ w_{t+1} \\ e_{t+1} \\ e_t \\ s^i_{t|t} \\ s^i_{t|t-1} \\ s^i_{t-1|t-2} \\ s_{t|t} \\ s_{t|t-1} \\ s_{t-1|t-2} \end{bmatrix}$$

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

Misspecified Expectations

State:

$$\begin{bmatrix} s_t \\ s_{t-1} \\ w_{t+1} \\ e_{t+1} \\ e_t \\ s^i_{t|t} \\ s^i_{t|t-1} \\ s_{t|t} \\ s_{t|t-1} \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \kappa_1\rho_1+\kappa_2 & \kappa_1\rho_2 & \kappa_1 & 0 & \kappa_2 & (1-\kappa_1)\widehat{\rho}-\kappa_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \widehat{\rho} & 0 & 0 & 0 \\ \kappa_1\rho+\kappa_2 & \kappa_1\rho_2 & \kappa_1 & 0 & \kappa_2 & 0 & 0 & (1-\kappa_1)\widehat{\rho}-\kappa_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \widehat{\rho} & 0 \end{bmatrix} \begin{bmatrix} s_{t-1} \\ s_{t-2} \\ w_t \\ e_t \\ e_{t-1} \\ s^i_{t-1|t-1} \\ s^i_{t-1|t-2} \\ s_{t-1|t-1} \\ s_{t-1|t-2} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \kappa_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_{t+1} \\ e_{t+1} \\ v^i_t \end{bmatrix}$$

Measurement:

$$\begin{bmatrix} x_{t+1}-\widehat{x}^i_{t+1|t} \\ \widehat{x}^i_{t+1|t}-\widehat{x}^i_{t+1|t-1} \\ x_{t+1}-\widehat{x}^i_{t+1|t} \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 & 1 & 1 & 0 & -\widehat{\rho} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \widehat{\rho} & -\widehat{\rho} & 0 & 0 \\ \rho_1 & \rho_2 & 1 & 1 & 0 & 0 & 0 & -\widehat{\rho} & 0 \end{bmatrix} \begin{bmatrix} s_t \\ s_{t-1} \\ w_{t+1} \\ e_{t+1} \\ e_t \\ s^i_{t|t} \\ s^i_{t|t-1} \\ s_{t|t} \\ s_{t|t-1} \end{bmatrix}$$

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

Appendix D Robustness

AR(2) Dynamics for all Models

Since the ME models is the only one to assume AR(2) dynamics, in this subsection I fit all of the other models with fundamental AR(2) dynamics and show that the ME model remains the best fitting candidate model.

Table D2: Alternative Models under AR(2) Dynamics

Panel A: Parameter Estimates					
Description	Parameter	(1) RE	(2) DE	(3) OC	(4) ME
First order autocorrelation	ρ_1	0.710	0.632	0.623	0.446
		(0.043)	(0.070)	(0.132)	(0.126)
Second order autocorrelation	ρ_2	-0.999	-0.922	-0.877	0.552
		(0.101)	(0.0147)	(0.486)	(0.157)
Persistent innovation dispersion	σ_w	1.312	1.129	1.139	0.905
		(0.102)	(0.061)	(0.206)	(0.125)
Transitory innovation dispersion	σ_e	1.646	1.567	1.550	1.494
		(0.180)	(0.073)	(0.105)	(0.191)
Private noise dispersion	σ_v	0.873	0.800	0.841	0.989
		(0.363)	(0.105)	(0.230)	(0.463)
Diagnosticity	φ		0.307		
			(0.078)		
Overconfidence	α_v			0.314	
				(0.164)	
Perceived persistence	$\widehat{\rho}$				0.811
					(0.081)
Panel B: Model Selection					
Log likelihood		-7625.1	-7359.1	-7353.0	-7295.2
AIC		15089	14728	14716	14602
BIC		15111	14755	14743	14635

Note: Panel A reports parameters estimates. Column (1) denotes the rational expectations model, column (2) reports the diagnostic expectations model, column (3) reports the overconfidence model, and column (4) reports the misspecified expectations model. Bootstrapped standard errors reported in parenthesis. Panel B reports the maximized log likelihood as well as AIC and BIC for each model.

Inflation Forecasts

Table D3: Alternative Models for SPF Inflation (CPI)

Panel A: Parameter Estimates					
		(1)	(2)	(3)	(4)
Description	Parameter	RE	DE	OC	ME
Persistence	ρ	0.959	0.874	0.810	
		(0.029)	(0.032)	(0.067)	
Persistent innovation dispersion	σ_w	1.243	1.065	1.112	0.966
		(0.132)	(0.057)	(0.092)	(0.102)
Transitory innovation dispersion	σ_e	2.060	2.129	1.930	1.729
		(0.138)	(0.102)	(0.122)	(0.185)
Private noise dispersion	σ_v	1.399	1.396	1.427	1.350
		(0.177)	(0.218)	(0.179)	(0.617)
Diagnosticity	φ		0.382		
			(0.119)		
Overconfidence	α_v			0.528	
				(0.134)	
Perceived persistence	$\widehat{\rho}$				0.813
					(0.100)
First order autocorrelation	ρ_1				0.431
					(0.107)
Second order autocorrelation	ρ_2				0.567
					(0.115)
Panel B: Model Selection					
Log likelihood		-7033	-6993	-6961	-6843
AIC		14074	13996	13932	13699
BIC		14095	14023	13958	13731

Note: Panel A reports parameters estimates. Column (1) denotes the rational expectations model, column (2) reports the diagnostic expectations model, column (3) reports the overconfidence model, and column (4) reports the misspecified expectations model. Bootstrapped standard errors reported in parenthesis. Panel B reports the maximized log likelihood as well as AIC and BIC for each model.

Sample Period

Table D4: Alternative Models for 1969-2019

Panel A: Parameter Estimates					
Description	Parameter	(1) RE	(2) DE	(3) OC	(4) ME
Persistence	ρ	0.999	0.745	0.970	
		(0.002)	(0.078)	(0.046)	
Persistent innovation dispersion	σ_w	2.159	1.304	1.318	1.772
		(0.207)	(0.095)	(0.099)	(0.336)
Transitory innovation dispersion	σ_e	2.154	2.134	2.064	1.848
		(0.199)	(0.118)	(0.104)	(0.407)
Private noise dispersion dispersion	σ_v	1.969	1.365	1.498	2.079
		(0.455)	(0.419)	(0.182)	(1.023)
Diagnosticity	φ		1.600		
			(0.351)		
Overconfidence	α_v			0.139	
				(0.088)	
Perceived persistence	$\widehat{\rho}$				0.907
					(0.077)
First order autocorrelation	ρ_1				0.540
					(0.210)
Second order autocorrelation	ρ_2				0.456
					(0.238)
Panel B: Model Selection					
Log likelihood		-15712	-15300	-15366	-15224
AIC		31432	30611	30742	30460
BIC		31456	30640	30771	30495

Note: Panel A reports parameters estimates. Column (1) denotes the rational expectations model, column (2) reports the diagnostic expectations model, column (3) reports the overconfidence model, and column (4) reports the misspecified expectations model. Bootstrapped standard errors reported in parenthesis. Panel B reports the maximized log likelihood as well as AIC and BIC for each model.