A New Fact to Discipline Models of Beliefs[*]

Julio Ortiz [†]

December 2020

Abstract

Updates made to existing forecasts that are systematically correlated with subsequent forecast errors yield "error predictability," a salient feature of survey data. Full information rational expectations (FIRE) models, by contrast, produce unpredictable errors, a finding that motivates many non-FIRE theories. Within this non-FIRE group, I demonstrate that error predictability is not typically enough to distinguish across alternative belief systems. Instead, I highlight an additional empirical fact that, paired with error predictability, can serve to further disentangle competing models of beliefs. In particular, I emphasize that linear rational models featuring imperfect information tend to require that forecast revisions be serially uncorrelated. Applying this logic to two widely studied theories and taking it to data from the Survey of Professional Forecasters favors a model of diagnostic expectations over a model of beliefs driven by strategic interaction.

Keywords: Rational expectations. Noisy information. Overreactions.
JEL Codes: D83, D84

[*]I would like to thank Stephen Terry and Constantin Bürgi for their valuable feeback and helpful conversations
[†]Boston University, Department of Economics, 270 Bay State Road, Boston, MA 02215; Phone: 201-230-1960; Email: jlortiz@bu.edu.

1 Introduction

Expectations are ubiquitous in economics. Nonetheless, the manner in which expectations are formed remains an open question. The benchmark theory of full information rational expectations (FIRE) posits that agents form beliefs in an optimal manner such that ex-post forecast errors are unpredictable. Error orthogonality, however, has been found to be inconsistent with survey data on individual expectations. In particular, ex-ante updates made to forecasts are systematically correlated with ex-post errors. Such error predictability has motivated departures from FIRE in the literature, particularly in favor of non-rational theories (Bordalo et al., 2020; Broer and Kohlhas, 2019; Kohlhas and Walther, 2020; D'Acunto et al., 2019).

This paper shows that error predictability, while sufficient to reject FIRE, is not particularly informative about individual rationality. I show this by way of example. Specifically, I narrow my focus to models that explain overreaction, a salient feature of survey data. I consider three prominent non-FIRE theories, embedded in a noisy information environment: overconfidence (Daniel et al., 1998), diagnostic expectations (Bordalo et al., 2020), and strategic interaction (Woodford, 2001). I then prove that all three of these models can deliver identical patterns of error predictability. Whereas the first two models feature non-rational expectations, the model of strategic interaction is rational. Intuitively, each of these theories features a key parameter that governs the extent of overreactive behavior. I show that the keys parameters can be mapped to each other, thereby delivering identical relationships between ex-post expectation errors and ex-ante revisions.

The three candidate models, however, are not observationally equivalent in general. I offer a simple testable implication that is able to distinguish across all three of these models, and that can provide a way forward more broadly as the literature assesses which non-FIRE theories are consistent with the data. The relevant testable implication is the persistence of revisions about a fixed event. Models of overconfidence imply a negative autocorrelation coefficient for revisions, distinct from the aforementioned errors-on-revisions coefficient. Whereas diagnostic expectations also deliver a negative autocorrelation coefficient for revisions, this theory implies that the autocorrelation of revisions is *identical* to the errors-on-revisions coefficient. On the other hand, the model of strategic interactions requires that revisions made to a fixed event forecast be unpredictable.

When taking this testable implication to data from the Survey of Professional Forecasters (SPF), I document evidence in against models of strategic interaction. For most variables in the SPF, I find that estimates of error predictability and revision persistence are not statistically distinct from one another. As a result, the data are broadly consistent with diagnostic expectations. Nonetheless, there are some variables for which these estimates diverge significantly, indicating that a theory of overconfidence provides a better fit for those variables.

My findings support non-rational theories of overreaction under standard linear dynamics.[1] Whereas models featuring Bayesian updating can deliver error predictability conditional on modifications made to the forecaster's objective, these models are generally unable to deliver revision predictability.[2] Bayesian updating requires that information be used efficiently, thereby preventing past revisions from holding predictive power over current revisions. Hence, analyzing the time series properties of revisions serves as a way to narrow the set of non-FIRE models that are consistent with the data.

Several studies have used survey data to test for FIRE.[3] In essence, FIRE implies that expectation errors, e_t, are orthogonal to the individual forecaster's information set, $\mathcal{I}_t$:

$$\mathbb{E}(e_t|\mathcal{I}_t) = 0.$$

This implication, however, is at odds with the data. In particular, ex-ante forecast revisions predict ex-post errors. Since a rejection of FIRE can be due to either a rejection of full information or a rejection of rationality, existing non-FIRE theories can be classified as fundamentally rational or non-rational. Rational non-FIRE theories typically incorporate a particular motive for deviating from the minimum mean square forecast. On the other hand, non-rational theories introduce a

[1] Models featuring nonlinear dynamics have also been found to rationalize the data Ortiz (2020).

[2] There are two documented exceptions in the extant literature: models featuring smoothing motives and noisy memory. Section 5 discusses these theories in more detail.

[3] Examples include Coibion and Gorodnichenko (2015), Bordalo et al. (2020), Fuhrer (2018), Dovern et al. (2015), Crowe (2010), Paloviita and Viren (2013), Burgi (2016), Andrade and Bihan (2013)

behavioral bias such that. In either case, expectation errors can be characterized as

$$\widehat{\mathbb{E}}(e_t|\mathcal{I}_t) \neq 0.$$

Nordhaus (1987) first introduced the testable implication proposed in this paper. For the same reason that errors must be unpredictable under FIRE, so must revisions, r_t:

$$\mathbb{E}(r_t|\mathcal{I}_t) = 0.$$

The key contribution of this paper is to study non-FIRE theories in which:

$$\widehat{\mathbb{E}}_t(r_t|\mathcal{I}_t) = 0 \quad \text{despite} \quad \widehat{\mathbb{E}}_t(e_t|\mathcal{I}_t) \neq 0,$$

and to show that such theories are inconsistent with the data. In doing so, this paper offers a new fact to discipline models of beliefs.

The focus of this paper is on overreaction. It should be noted that a recent literature has documented the existence of both over- and underreaction in survey expectations (Broer and Kohlhas, 2019; Kohlhas and Walther, 2020; Ortiz, 2020). There is at present no unified definition of over- and underreaction. For instance Broer and Kohlhas (2019) document overreaction in the form of a negative relation between errors and revisions, and underreaction to some endogenous public signals. On the other hand, Kohlhas and Walther (2020) define overreactions as a negative relation between errors and current output growth, while underreactions are defined as a the relation between errors and revisions at the consensus-level. Finally, (Ortiz, 2020) documents that the same forecaster simultaneously exhibits positive and negative coefficients of error predictability for distinct macroeconomic variables. Though I abstract from these empirical regularities here, my findings are relevant for this strand of the literature as well. In particular, theories of simultaneous over- and underreaction that are to be consistent with professional forecasts must be able to deliver serially correlated revisions.

The rest of the paper is organized as follows. In Section 2, I document empirical facts related to overreaction in survey expectations, based on both error and revision predictability. In Section 3, I introduce the three candidate models of overreaction. Section 4 maps the models' key parameters to each other and shows that all models can deliver the same relation between errors and revisions. I then characterize revision predictability from the lens of each model in Section 5. Section 6 concludes.

2 Evidence of Overreaction in Survey Expectations

Many non-FIRE theories have been devised to explain overreactive behavior(Bordalo et al., 2020; Afrouzi et al., 2020; Rozsypal and Schlafmann, 2019; Fuster et al., 2012; Benigno and Karantounias, 2019). The literature often utilizes survey data from the Survey of Professional Forecasters in order to assess the empirical performance of different theories. The SPF is a quarterly survey provided by the Federal Reserve Bank of Philadelphia. The survey provides forecasts from several forecasters across a number of macroeconomic variables over many horizons.

A popular test often implemented for empirical motivation projects expectation errors on revisions. Suppose that x_t is the target variable and $x^i_{t+h|t}$ is forecaster i's forecast devised at time t for horizon h. Then the empirical test is defined as:

$$\underbrace{x_{t+h} - x^i_{t+h|t}}_{\text{Error}} = \beta_0 + \beta_1[\underbrace{x^i_{t+h|t} - x^i_{t+h|t-1}}_{\text{Revision}}] + \epsilon^i_{t+h}. \tag{1}$$

At the forecaster-level, the coefficient in front of the revision is found to be negative for a range of macroeconomic variables, β_1.[4] Table 1 confirms this finding. Pooling across horizons, the results indicate that errors covary negatively with revisions. For instance, a one percentage point upward revision for real GDP is associated with a 0.23 percentage point more negative forecast

[4]Though there is also evidence of underreaction for some variables, I focus on overreaction due to its pervasiveness at the forecaster-level. Moreover, the model comparison exercise that I later undertake is without much loss of generality. In particular, a theory of underreaction that features strategic complementarity would imply a positive covariance between errors and revisions, but it would nonetheless require that revisions be serially uncorrelated.

Table 1: Forecast Error Predictability

Variable	Estimate	Standard Error	Observations
CPI	-0.103	0.122	7,052
Industrial production	-0.177	0.055	11,491
Nominal GDP	-0.291	0.044	11,994
GDP deflator	-0.297	0.057	11,951
Real consumption	-0.373	0.072	8,865
Real federal government spending	-0.499	0.061	8,273
Real GDP	-0.229	0.058	12,110
Real nonresidential investment	-0.228	0.088	8,604
Real residential investment	-0.192	0.059	8,587
Real state and local government spending	-0.542	0.027	8,282
Ten-year government bond	-0.079	0.059	7,246

Note: The table reports the estimated β_1 coefficient of error predictability from (1). Samples for reach regression are pooled across horizons. Standard errors are clustered by forecaster and quarter.

error. Following positive news, forecasters tend to revise upward excessively such that the overly-optimistic update results in a systematically more negative subsequent error. Hence, professional forecasters tend to over-revise their forecasts.

Overreaction can similarly be observed by analyzing the persistence of fixed-event updates. Empirically, revision predictability estimates the following relationship:

$$\underbrace{x^i_{t+h|t} - x^i_{t+h|t+h-1}}_{\text{Revision}} = \gamma_0 + \gamma_1[\underbrace{x^i_{t+h|t+h-1} - x^i_{t+h|t+h-2}}_{\text{Previous revision}}] + \nu^i_{t+h}, \tag{2}$$

If $\gamma_1 \neq 0$, then the fixed event revisions are serially correlated (predictable). Whereas testing for error predictability is more common in the literature, revision predictability was previously analyzed in (Dovern et al., 2015).

Table 2 reports estimates of (2). Across the same set of variables for which $\beta_1 < 0$, there is also ample evidence of $\gamma_1 < 0$. Focusing on real GDP, a one percentage point upward revision in the forecast for $t + h$ devised today is associated with a 0.23 percentage point downward revision in the same forecast made tomorrow. Jointly, the results could be interpreted as follows: amid the realization of positive news today, forecasters appear to over-revise their predictions upward only

Table 2: Forecast Revision Predictability

Variable	Estimate	Standard Error	Observations
CPI	-0.250	0.064	7,052
Industrial production	-0.202	0.030	11,491
Nominal GDP	-0.223	0.032	11,994
GDP deflator	-0.255	0.031	11,951
Real consumption	-0.240	0.042	8,865
Real federal government spending	-0.283	0.029	8,273
Real GDP	-0.211	0.033	12,110
Real nonresidential investment	-0.165	0.042	8,604
Real residential investment	-0.165	0.039	8,587
Real state and local government spending	-0.317	0.025	8,282
Ten-year government bond	0.003	0.052	7,246

Note: The table reports the estimated β_1 coefficient of error predictability from (2). Samples for reach regression are pooled across horizons. Standard errors are clusted by forecaster and quarter.

to observe a systematically more negative forecast error tomorrow, which subsequently contributes to a more pessimistic revision tomorrow.

Overreaction is pervasive among professional forecasters. In the subsequent section, I review three prominent models of overreaction that have been proposed in the literature.

3 Three Models of Overreactive Behavior

All three models of overreaction are grounded in a noisy information setting. Optimal forecasts in this context are consistent with the mathematical expectations operator, $\mathbb{E}$. This will be the benchmark from which all other expectations will deviate. In keeping with Coibion and Gorodnichenko (2015) and Bordalo et al. (2020) consider the following set up:

$$\text{Exogenous State:} \quad x_t = \rho x_{t-1} + w_t, \quad w_t \sim \mathcal{N}(0, \sigma_w^2)$$

$$\text{Private Signal:} \quad y_t^i = x_t + v_t^i, \quad v_t^i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_v^2).$$

The exogenous fundamental follows a simple AR(1) process.[5] Moreover, agents have access only to private information in the form of a noisy private signal y_t^i observed each period. I abstract away from more complex signal structures for simplicity, however, the assumptions on how agents receive information are unimportant for the results presented in subsequent sections.

From the Kalman filter (Anderson and Moore, 1979), the optimal nowcast of x_t is:

$$\mathbb{E}(x_t|\mathcal{I}_t^i) = x_{t|t}^i = (1-\kappa)x_{t|t-1}^i + \kappa y_t^i,$$

where $\mathcal{I}_t^i$ denotes individual i's information set at time t, and κ refers to the steady-state Kalman gain, $\kappa = \frac{\text{Var}(x_t - x_{t|t-1}^i)}{\text{Var}(x_t - x_{t|t-1}^i)+\sigma_v^2}$ which is the optimal (in the mean-square sense) weight placed on new information.[6]

3.1 Overconfidence

First, I consider a non-rational model of overconfidence. Daniel et al. (1998) presents a theory of overconfidence in which individuals perceive their private signals to be more precise than they truly are. Intuitively, because forecasters believe their information to be more precise, they will tend to trust their signals more than is optimal. As a result, overconfidence forecasters over-weight new information, thereby generating a over-revisions and delivering a negative covariance between updates and errors.

Specifically, forecasters perceive

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

where $\breve{\sigma}_v = \alpha\sigma_v$ such that $\alpha \in [0,1]$.

As a matter of notation, let the forecaster's current period forecast be denoted by $\breve{x}_{t|t}^i$, and his one step ahead forecast by $\breve{x}_{t|t-1}^i$. Forecasters invoke the Kalman filter in order to formulate their

[5]The dynamics of the state are innocuous for the results of the paper.

[6]Note that the imperfect information enviornment is a more general formulation than a full-information environment. The model collapses to full-information rational expectations when $\sigma_v = 0$ so that $\kappa = 1$.

expectations. As a result, expectations are determined according to the following predict-update procedure:

$$\begin{aligned} \breve{x}^i_{t|t-1} &= \rho \breve{x}^i_{t-1|t-1} && \text{(Predict)} \\ \breve{x}^i_{t|t} &= \breve{x}^i_{t|t-1} + \breve{\kappa}(y^i_t - \breve{x}^i_{t|t-1}) && \text{(Update).} \end{aligned}$$

Having defined the expectations formation process under overconfidence, we obtain the following result.

Proposition 1. *The OLS coefficient arising from an errors-on-revisions regression in the overconfidence model is:*

$$\beta_1^{OC} = \frac{\mathbb{C}(x_t - \breve{x}^i_{t|t}, \breve{x}^i_{t|t} - \breve{x}^i_{t|t-1})}{\mathbb{V}(\breve{x}^i_{t|t} - \breve{x}^i_{t|t-1})} = \frac{(\alpha^2 - 1)\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2} < 0.$$

Proof. See Appendix A. □

The coefficient β_1^{OC} is always negative since $\alpha < 1$. In the absence of overconfidence, $\alpha = 1$ and $\beta_1 = 0$. Hence, the underlying overconfidence on the part of forecasters generates observed overreaction.

3.2 Diagnostic Expectations

Second, I consider a non-rational theory of diagnostic expectations (Bordalo et al., 2020). Under diagnostic expectations, overreactions arise due to the representativeness heuristic of Tversky and Kahneman (1974). Intuitively, more recent information is more easily recalled and therefore overweighted when formulating beliefs. The predict-update procedure for diagnostic forecasters can be described as follows:

$$x^{i,\theta}_{t|t-1} = \rho x^{i,\theta}_{t-1|t-1} \qquad \text{(Predict)}$$

$$x^{i,\theta}_{t|t} = x^{i}_{t|t} + \theta(x^{i}_{t|t} - x^{i}_{t|t-1}) \qquad \text{(Update)},$$

where θ is the belief distortion parameter that governs the extent of overreaction. Under $\theta = 0$, there is no belief distortion, and expectations collapse to the standard rational expectations benchmark.

Bordalo et al. (2020) provide additional details including a derivation of the coefficient of error predictability:

$$\beta_1^{DE} = \frac{\mathbb{C}(x_t - x^{i,\theta}_{t|t}, x^{i,\theta}_{t|t} - x^{i,\theta}_{t|t-1})}{\mathbb{V}(x^{i,\theta}_{t|t} - x^{i,\theta}_{t|t-1})} = \frac{-\theta(1+\theta)}{(1+\theta)^2 + \theta^2\rho^2} < 0.$$

3.3 Strategic Interaction

Overreactions can also arise rationally, that is, as an optimal outcome. I next consider a model of strategic interaction. The model presented in this section draws from Morris and Shin (2002) and Woodford (2001). To obtain overreactions, I assume that forecasters are characterized by strategic substitutability. Intuitively, forecasters have the dual objective to minimize their squared errors and also to distinguish themselves from the average forecast. More specifically, each forecaster wishes the minimize the following loss function:

$$\min_{\{\widetilde{x}^i_{t|t}\}} \mathbb{E}\left[(x_t - \widetilde{x}^i_{t|t})^2 + R(\widetilde{x}^i_{t|t} - F_t)^2 | \mathcal{I}^i_t\right]. \tag{3}$$

where x_t is the realized fundamental, $\widetilde{x}^i_{t|t}$ is forecaster i's *reported* current-period forecast, $\mathcal{I}^i_t$ is forecaster i's information set at time t, F_t is the consensus forecast at time t, and $R < 0$ is the degree of strategic substitutability.[7] There are a number of possible microfoundations for strategic substitutability. Most prominently, see Ottaviani and Sorensen (2006). When $R = 0$, the loss function collapses to the familiar mean-squared loss.

[7]For $R > 0$, forecasters exhibit strategic complementarities. That is, forecasters have an incentive to stay close to the consensus forecast.

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

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

where $\mathbb{E}^{(k)}$ is the k^{th}-order expectation of x_t.

Taking the first order conditions of (3), it follows that the optimal reported prediction is:

$$\widetilde{x}^i_{t|t} = \frac{1}{1+R}x^i_{t|t} + \frac{R}{1+R}F^i_{t|t}.$$

where $x^i_{t|t}$ is the optimal current-period forecast for the state and $F^i_{t|t}$ is forecaster i's prediction about what the consensus nowcast is at time t.

From this, it follows that the forecast error in this model is: [8]

$$x_t - \widetilde{x}^i_{t|t} = (1-\lambda)\left[x_t - \frac{1}{1+R}x^i_{t|t-1} - \frac{R}{1+R}\rho F^i_{t-1|t-1}\right] - \lambda v^i_t.$$

and the forecast revision is defined as:

$$\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} = \lambda(x_t - x^i_{t|t-1} + v^i_t).$$

where $\lambda = \frac{\kappa_1 + R\kappa_2}{1+R}$ and κ_1 and κ_2 are the elements of the $2\times$ 1 Kalman gain vector, κ.[9]

Having defined the errors and revisions according to this model, we can derive the following result.

Proposition 2. *The errors-on-revisions regression coefficient in the strategic substitution model is:*

$$\beta_1^{SI} = \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2} < 0.$$

[8]See Appendix A for details.

[9]The Kalman gain vector is two-dimensional because forecasters generate predictions for the unobserved state, x_t and the unobserved consensus forecast, F_t.

Proof. See Appendix A. □

As expected, when $R = 0$, the coefficient $\beta_1^{SI} = 0$, consistent with rational expectations under standard mean squared loss. Given the assumption placed on R, this model can generate overreactions.[10]

All three of these models are capable of explaining the over-revisions observed in the survey data and reported in Table 1. Collectively, these three models encompass both rational and non-rational theories of expectation formation. I next demonstrate how these three models can deliver the same patterns of error predicability, implying that the forecast error predictability, while evidence against FIRE, is not particularly informative about rationality. Following this, I show how revision predictability is a more desirable test because it can speak to deviations from FIRE and, more specifically, deviations from rationality.

4 Matching Error Predictability

All three models discussed in the previous section are able to generate forecaster-level overreactions. Table 3 summarizes the updating rules for each of the three models. Note that these models can be expressed as a (positive) deviation from the conditional expectation, $\mathbb{E}(x_t|\mathcal{I}_t^i) = x_{t|t}^i$. Hence, forecast updates exceed what is called for by the optimal minimum mean square estimate. Importantly, all three theories can deliver an identical β_1 coefficient.

Proposition 3. *Fix the state and signal parameters, ρ, σ_w, and σ_v. Given one of the parameters in $\{\alpha, \theta, R\}$, there exist values for the other two such that :*

$$\beta_1^{OC} = \beta_1^{DE} = \beta_1^{SI}.$$

Proof. See Appendix A. □

[10] One might wonder whether forecasters can exhibit overreactions under strategic complementarities ($R > 0$). This can only occur if the weight placed private information when predicting the consensus forecast exceeds the weight placed on private information when predicting the state ($\kappa_2 > \kappa_1$). Given that the signal is more informative about x_t than F_t, it is never optimal for the forecaster to set $\kappa_2 > \kappa_1$.

Table 3: Update Rules Across Models

Model	Update rule
Overconfidence	$\breve{x}^i_{t\|t} = x^i_{t\|t} + (1-\breve{\kappa})\breve{x}^i_{t\|t-1} - (1-\kappa)x^i_{t\|t-1} + (\breve{\kappa}-\kappa)y^i_t$
Diagnostic Expectations	$x^{i,\theta}_{t\|t} = x^i_{t\|t} + \theta(x^i_{t\|t} - x^i_{t\|t-1})$
Strategic Interaction	$\widetilde{x}^i_{t\|t} = x^i_{t\|t} - \frac{R}{1+R}(x^i_{t\|t} - F^i_{t\|t})$

Note: The table reports the updating rules for models of overconfidence, diagnostic expectations, and strategic interaction.

This is fundamentally an exercise in matching moments. In light of the recent use of β_1 to motivate non-FIRE models, this result demonstrates that a key assumption made in non-FIRE models (rationality or non-rationality) are equally capable of delivering the same patterns of error predictability. I next summarize precisely how these coefficients are equalized across theories of overreaction with details provided in Appendix A.

Consider first the mapping to the strategic interaction model. Given $\beta_1^{DE} = \beta_1(\rho, \sigma_v, \sigma_w, \theta)$ or $\beta_1^{OC} = \beta_1(\rho, \sigma_v, \sigma_w, \alpha)$, we can match β_1 in the strategic interaction model by setting the degree of strategic substitution to be:

$$R = \frac{[\kappa_1(1+\theta) - \kappa_2](1+\theta) + (\kappa_1 - \kappa_2)\rho^2\theta^2}{(1+\theta)^2 + \rho^2\theta^2}, \text{ and}$$
$$R = \frac{(\alpha^2-1)\sigma_v^2\kappa_1}{(\kappa_1-\kappa_2)[\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2] - (\alpha^2-1)\sigma_v^2\kappa_2},$$

respectively.

Similarly, there is a mapping to diagnostic expectations. Given $\beta_1^{SI} = \beta_1(\rho, \sigma_v^2, \sigma_w^2, R)$ or $\beta_1^{OC} = \beta_1(\rho, \sigma_v^2, \sigma_w^2, \alpha)$, one could construct a model diagnostic expectations that delivers an equivalent β_1 by setting the degree of diagnosticity, θ, to be equal to the largest root of:

$$0 = \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2} + \left(\frac{3R\kappa_1 - R\kappa_2}{\kappa_1 + R\kappa_2}\right)\theta + \left[\frac{(1+R)\kappa_1 + \rho^2 R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2}\right]\theta^2, \text{ and}$$
$$0 = \frac{(\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2} + \left(\frac{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + (2\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2}\right)\theta + \left(\frac{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + [\alpha^2 + \rho^2(\alpha^2-1)]\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2}\right)\theta^2,$$

respectively.

Finally, consider the mapping to overconfidence. Given $\beta_1^{DE} = \beta_1(\rho, \sigma_v^2, \sigma_w^2, \theta)$ or $\beta_1^{SI} = \beta_1(\rho, \sigma_v^2, \sigma_w^2, R)$ one could construct a model of overconfidence by finding the α parameter such that:

$$\alpha = \sqrt{\frac{(1+\theta)[\sigma_v^2 - \theta\mathbb{V}(x_t - \breve{x}^i_{t|t-1})]}{[(1+\theta)^2 + \rho^2\theta^2]\sigma_v^2}}, \text{ and}$$
$$\alpha = \sqrt{\frac{(1+R)\kappa_1\sigma_v^2 + R(\kappa_1 - \kappa_2)\mathbb{V}(x_t - \breve{x}^i_{t|t-1})}{(\kappa_1 + R\kappa_2)\sigma_v^2}},$$

respectively. Note that $\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) \equiv \breve{\Psi}^i_{t|t-1}(\alpha)$ is the forecast error variance in the overconfidence model which is itself a function of α.

Hence, by simply assessing the regression coefficient in an errors-on-revisions regression, one cannot necessarily distinguish across noisy information models of rational and non-rational expectations. The panels in Figure 1 plot the relationship between the θ, α, and R that are key in delivering identical β_1 coefficients.

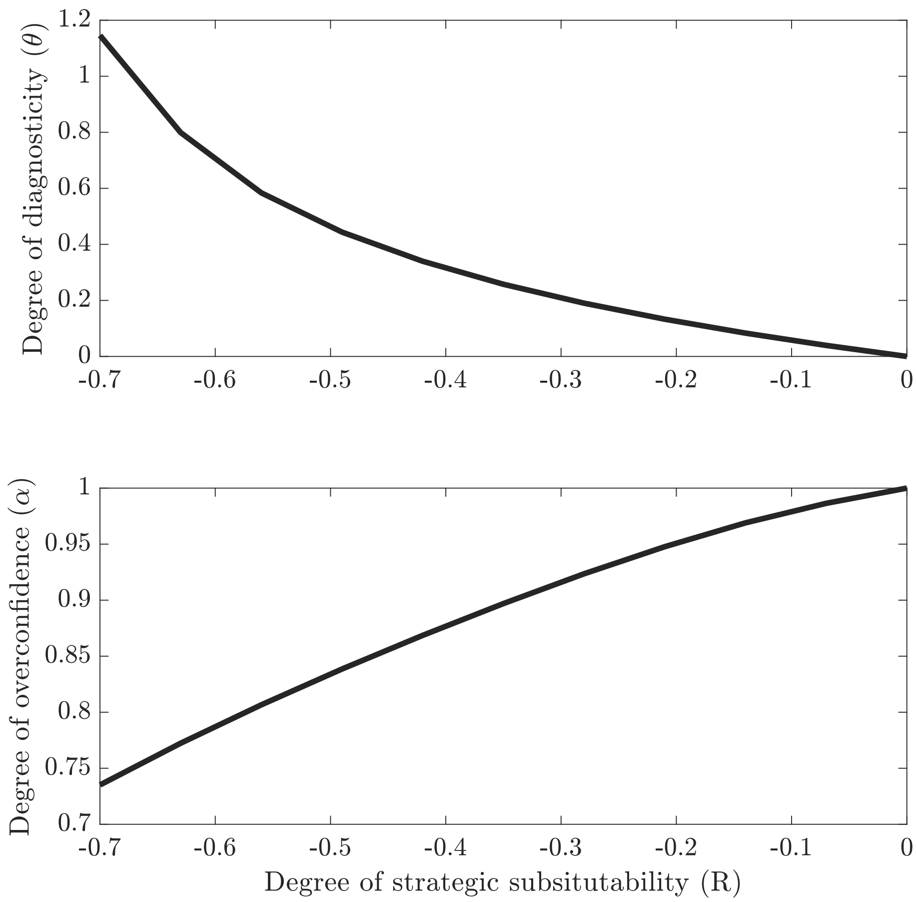
5 Revision Predictability

While both models can deliver identical β_1 coefficients, they are not observationally equivalent, in general. Hence, with enough data, one can discern between the two. More broadly, with enough data, we can distinguish between two subsets of non-FIRE models. I show that we can make progress on this front by merely focusing on an additional fact: the persistence of fixed-even revisions.

Beyond forecast error orthogonality, Nordhaus (1987) notes that revisions must be "informationally efficient." This requires the following condition to hold:

$$\mathbb{E}(x^i_{t|t} - x^i_{t|t-1} | \mathcal{I}^i_t) = 0.$$

Figure 1: Mapping β_1 Across Models of Overreactions



Note: The figure plots the degree of strategic substitutability, R, that generates the same β_1 that is obtained by non-rational models of overreactions denoted by the degree of diagnosticity, θ (for diagnostic expectations) and the parameter governing perceived information precision, α (for overconfidence).

In words, forecast revisions must be orthogonal to any variable residing in the forecasters information set,

$$\mathbb{E}[(x^i_{t|t} - x^i_{t|t-1})\mu] = 0 \quad \text{for } \mu \in \mathcal{I}^i_t.$$

This is akin to the error orthogonality condition which has been the focus of conventional efficiency tests. However, whereas error orthogonality can be violated for some linear noisy information rational expectations models, revision orthogonality cannot. This is an artifact of Bayesian updating in a linear setting. In such models, the forecast revision is equal to the innovation error observed when the signal is received, scaled by the optimal Kalman gain. These innovation errors are unpredictable by definition. Motivated by this insight, I present the next result.

Proposition 4. *The three models deliver distinct implications about the serial correlation of revisions. In particular:*

(i) The overconfidence model implies: $\gamma_1 = \frac{(\alpha^2-1)\sigma_v^2\breve{\kappa}}{\mathbb{V}(x_t-\breve{x}^i_{t|t-1})+\sigma_v^2} \leqslant \beta_1$.

(ii) The diagnostic expectations model implies: $\gamma_1 = -\frac{\theta(1+\theta)}{(1+\theta)^2+\theta^2\rho^2} = \beta_1$.

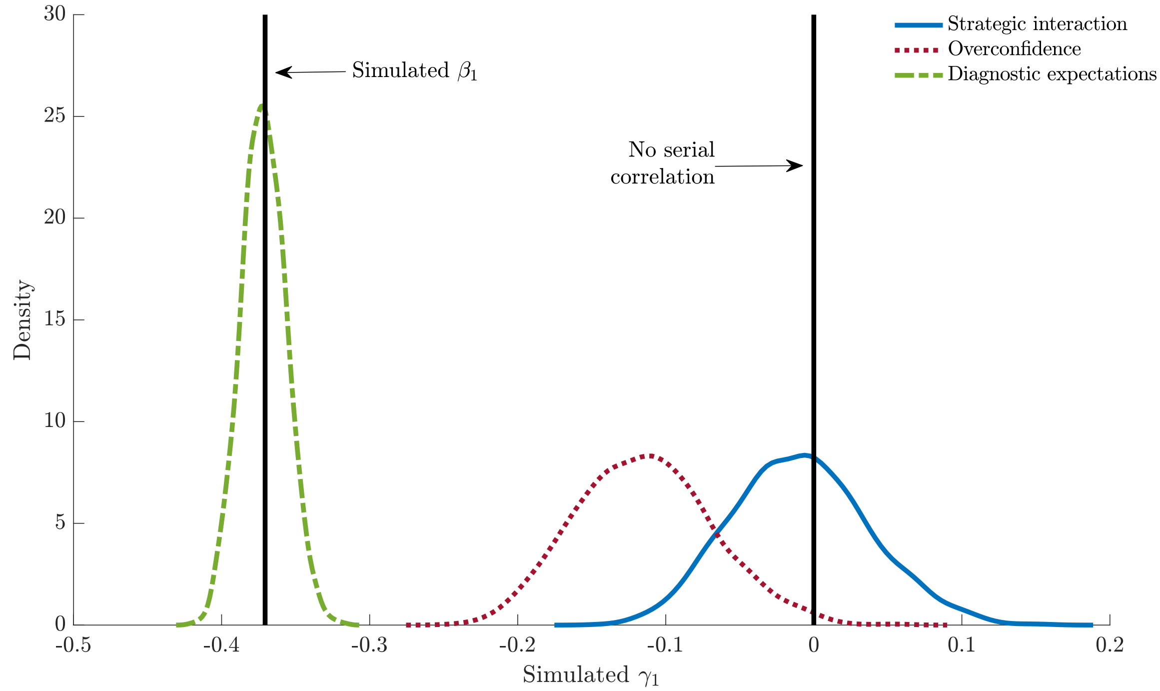
(iii) The strategic interaction model implies: $\gamma_1 = 0$.

Proof. See Appendix A. □

The autocorrelation of revisions is able to distinguish between all three models. In particular, the overconfidence model implies that revision persistence is negative and distinct from the errors-on-revisions coefficient. On the other hand, diagnostic expectations models imply that the persistence of revisions is identical to the errors-on-revisions coefficient. Finally, the strategic interaction model requires that revisions be serially uncorrelated.

Figure 2 plots a set of simulations results from all three models. I first fix the parameters of the strategic interaction model and I find the $\{\alpha, \theta\}$ that replicate β_1^{SI} in the other two models. I then compute the simulated revision persistence coefficient, γ_1 for the three models. The figure verifies that while all models can share identical β_1 coefficients, they have distinct implications for revision persistence. Consistent with Proposition 4, the overconfidence model implies that the first-order autocorrelation of revisions is smaller than the error predictability coefficient. Furthermore, diagnostic expectations requires that these two coefficients be equal. Finally, strategic interaction model requires lagged revisions to have no predictive power over current revisions. Observed overreaction on the basis of errors-on-revisions regressions is consistent with several theories, rational and non-rational alike. However, revision persistence is capable of discerning across non-FIRE models, and in the process, is more information about rationality. Based on negatively serially correlated revisions in the data, a rational model of strategic diversification can be rejected.

Figure 2: Simulated First-Order Autocorrelation of Revisions



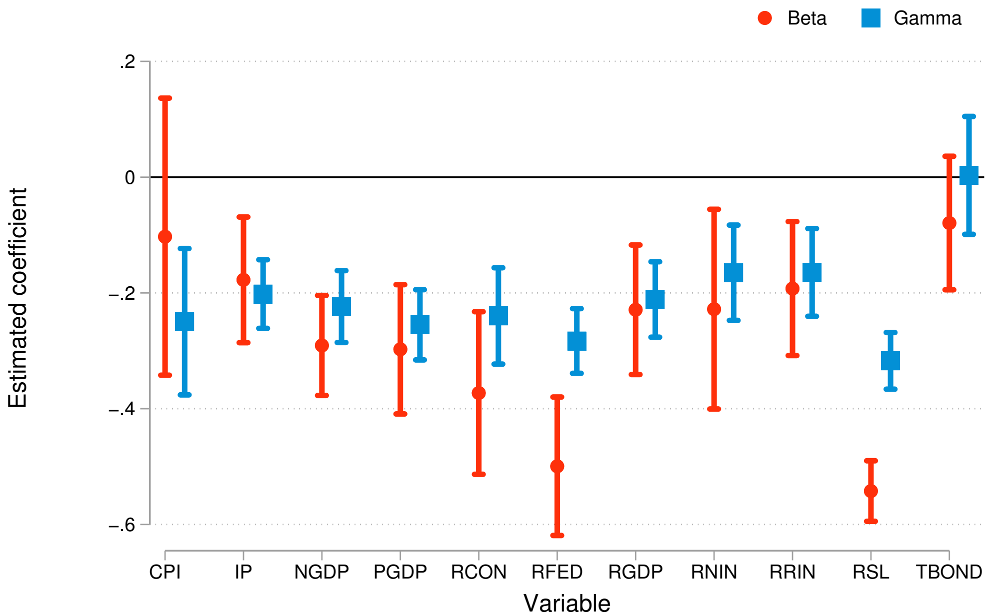
Note: The figure plots the simulated γ_1 coefficients from the three models, each of which delivers the same simulated β_1 coefficient of -0.37. The mean first-order autocorrelations of revisions are: -0.009, -0.11, and -0.37 for the strategic interaction, overconfidence, and diagnostic expectations models, respectively.

5.1 Empirical Evidence Against Strategic Interaction

Based on Table 2, it is clear that the data are inconsistent with serially uncorrelated revisions. Furthermore, Figure B1 visualizes the 95% confidence intervals of the error predictability estimates alongside revision persistence.[11] The estimates of β_1 and γ_1 are mostly statistically indistinguishable from one another. This lends support to the of overreaction based on diagnostic expectations. Nonetheless, for some variables there is evidence that the error predictability coefficient is different from the autocorrelation of revisions, namely, for macroeconomic variables related to government spending. In these cases, the data support a model of overconfidence.

[11] Estimates are pooled across horizon. Appendix B details horizon-by-horizon results.

Figure 3: Error Predictability and Revision Persistence



Note: Figure displays 95% confidence intervals for estimated coefficients of error predictability and revision persistence, (β_1 and γ_1 in the main text, respectively). Double clustered standard errors are specified. EMP- Employment, IP - Industrial Production, NGDP - Nominal GDP, PGDP - GDP Deflator, RCON - Real Consumption, RFED - Real Federal Government Spending, RGDP - Real GDP, RNIN - Real Nonresidential Investment, RRIN - Real Residential Investment, and RSL - Real State and Local Government Spending.

5.2 Broader Implications of Revision Predictability

As demonstrated in Table 3, non-FIRE forecasts can be generally expressed as some deviation from the conditional expectation:

$$x^i_{t|t} = \mathbb{E}(x_t|\mathcal{I}_t) + \tau^i_{t|t}$$
$$x^i_{t|t-1} = \mathbb{E}(x_t|\mathcal{I}_{t-1}) + \tau^i_{t|t-1},$$

where $\tau^i_{t|t}$ and $\tau^i_{t|t-1}$ describe the nature of the deviation from the minimum mean square forecast.[12] These deviations, which are either due to a strategic motive or a behavioral bias, can generate error predictability. However, models featuring strategic motives (i.e. alterations to the standard mean

[12]The benchmark noisy information model is a special case where $\tau^i_{t|t} = \tau^i_{t|t-1} = 0$.

square error objective) tend to imply zero autocovariance of revisions, implying:

$$\text{Cov}(\tau^i_{t|t} - \tau^i_{t|t-1}, \tau^i_{t|t-1} - \tau^i_{t|t-2}) = 0.$$

Examples of such models include those incorporating strategic interaction, heterogeneous priors, and asymmetric attention.[13] Notably, models featuring smoothing motives and noisy memory (Azeredo da Silvera and Woodford, 2019; Afrouzi et al., 2020) are an exception to this condition. In the case of smoothing motives, the forecast revision is explicitly incorporated in the forecaster's objective. As a result, the optimal reported forecast is a linear combination of the conditional expectation and the forecast revision. Nonetheless, as shown in Appendix D, such models imply a counterfactually positive autocorrelation of revisions.

Moreover, models featuring noisy memory can successfully deliver serially correlated revisions because forecasters optimize over the set of past signal realizations.[14] In particular, forecasters face a cost to retrieve past information. Intuitively presence of such a cost allows for information sets across adjacent periods to differ, thereby admitting for serially correlated revisions. These two theories serve as exceptions since they are both rational models that are able to generate non-zero revision persistence.

In general, theories that invoke Bayesian updating require informational efficiency despite the modified incentives to optimally deviate from the conditional expectation. Besides models in which forecasters strategically track their path of revisions or optimally choose their information sets each period, these models preclude revisions from being serially correlated. The class of models featuring behavioral biases, on the other hand, are generally better able to produce persistence in fixed-event revisions. Besides the overconfidence and diagnostic expectations models considered here, other theories that can make sense of this empirical fact include, for instance, models of relative overconfidence (Broer and Kohlhas, 2019) and natural expectations (Fuster et al., 2010).

[13] See Appendix D for a discussion of alternative models.

[14] Such models imply that the scope for serially correlated revisions is decreasing in the underlying persistence of the variable in question. In particular, when $\rho = 1$, these models imply zero autocorrelation of revisions.

6 Conclusion

In this paper, I argue that the popular errors-on-revisions coefficient used in the expectations formation literature is insufficient to motivate departures from rationality. By way of example, I demonstrate that two prominent models of non-rational expectations can deliver the same errors-on-revisions coefficient as in a rational strategic interaction model. Motivated by this finding, I offer a new fact for further discerning across non-FIRE models which requires projecting revisions on their past values. Using survey from the Survey of Professional Forecasters, I find evidence favoring diagnostic expectations and overconfidence over strategic interaction. To make progress on understanding how different economic actors formulate their expectations, it is necessary to have a set of rich theories. Conditional on a set of existing theories, however, it is important to assess the ways in which the models the fit the data. Studying the time series properties of updates made to fixed-event forecasts serves as a powerful way to discerning across existing non-FIRE models.

References

Afrouzi, Hassan, Spencer Kwon, Augustin Landier, Yueran Ma, and David Thesmar (2020), "Overreaction and Working Memory." Working Paper.

Anderson, Brian D. O. and John B. Moore (1979), "Optimal Filtering." *Englewood Cliffs*, 21, 22–95.

Andrade, Philippe and Herve Le Bihan (2013), "Inattentive Pprofessional Forecasters." *Journal of Monetary Economics*, 60, 967–982.

Azeredo da Silvera, Rava and Michael Woodford (2019), "Noisy Memory and Overreaction to News." NBER Working Paper 25478.

Benigno, Pierpaolo and Anastasios G. Karantounias (2019), "Overconfidence, Subjective Perception and Pricing Behavior." *Journal of Economic Behavior and Organization*, 164, 107–132.

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

Branch, William (2004), "The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations." *The Economic Journal*, 114, 592–621.

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

Burgi, Constantin (2016), "What Do We Lose When We Average Expectations?" Working Paper.

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.

Crowe, Christopher (2010), "Consensus Forecasts and Inefficient Information Aggregation." Working Paper.

D'Acunto, Francesco, Daniel Hoang, Maritta Paloviita, and Michael Weber (2019), "IQ, Expectations, and Choice." Becker Friedman Institute Working Paper.

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

Dovern, Jonas, Ulrich Fritsche, Prakash Loungani, and Natalia Tamirisa (2015), "Information Rigidities: Comparing Average and Individual Forecasts for a Large International Panel." *International Journal of Forecasting*, 31, 144–154.

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

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

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

Kohlhas, Alexandre N. and Answer Walther (2020), "Asymmetric Attention." Working Paper.

Morris, Stephen and Hyun Song Shin (2002), "Social Value of Public Information." *American Economic Review*, 92, 1521–1534.

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

Ortiz, Julio (2020), "Time-varying volatility as a source of error predictability." Working Paper.

Ottaviani, Marco and Peter Sorensen (2006), "The strategy of professional forecasting." *Journal of Financial Economics*, 2, 441–466.

Paloviita, Maritta and Matti Viren (2013), "Are Individual Survey Expectations Internally Consistent?" Working Paper.

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.

Peters, Florian and Simas Kucinskas (2018), "Measuring Biases in Expectation Formation." Working Paper.

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

Tversky, Amos and Daniel Kahneman (1974), "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185, 1124–1131.

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 Derivations

A.1 Proof of Proposition 1

Recall that $\breve{x}^i_{t|t} = \breve{x}^i_{t|t-1} + \breve{\kappa}(y^i_t - \breve{x}^i_{t|t-1})$. So the forecast error is $x_t - \breve{x}^i_{t|t} = (1-\breve{\kappa})(x_t - \breve{x}^i_{t|t-1}) - \breve{\kappa} v^i_t$, and the revision is $\breve{x}^i_{t|t} - \breve{x}^i_{t|t-1} = \breve{\kappa}[(x_t - \breve{x}^i_{t|t-1}) + v^i_t]$. Then, evaluating the population coefficient:

$$\begin{aligned}
\beta_1^{OC} &= \frac{\mathbb{C}(x_t - \breve{x}^i_{t|t}, \breve{x}^i_{t|t} - \breve{x}^i_{t|t-1})}{\mathbb{V}(\breve{x}^i_{t|t} - \breve{x}^i_{t|t-1})} \\
&= \frac{(1-\breve{\kappa})\breve{\kappa}\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) - \breve{\kappa}^2\sigma_v^2}{\breve{\kappa}^2[\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2]} \\
&= \frac{(1-\breve{\kappa})\mathbb{V}(x_t - x^i_{t|t-1}) - \breve{\kappa}\sigma_v^2}{\breve{\kappa}[\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2]}
\end{aligned}$$

Noting that the biased Kalman gain is, $\breve{\kappa} = \frac{\mathbb{V}(x_t - \breve{x}^i_{t|t-1})}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \breve{\sigma}_v^2}$, we can use this definition in the numerator to obtain:

$$\begin{aligned}
\beta_1^{OC} &= \frac{\frac{\breve{\sigma}_v^2\mathbb{V}(x_t - \breve{x}^i_{t|t-1})}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \breve{\sigma}_v^2} - \frac{\mathbb{V}(x_t - \breve{x}^i_{t|t-1})\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \breve{\sigma}_v^2}}{\breve{\kappa}[\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2]} \\
&= \frac{(\breve{\sigma}_v^2 - \sigma_v^2)\breve{\kappa}}{\breve{\kappa}[\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2]}
\end{aligned}$$

Finally, since $\breve{\sigma}_v = \alpha\sigma_v$, and $\alpha \in (0,1)$, we have:

$$\beta_1^{OC} = \frac{(\alpha^2 - 1)\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2} < 0.$$

A.2 Deriving Errors and Revisions for SI Model

In state space form, we have

$$\begin{bmatrix} x_t \\ F_t \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ \lambda\rho & (1-\lambda)\rho \end{bmatrix} \begin{bmatrix} x_{t-1} \\ F_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ \lambda \end{bmatrix} w_t$$

$$y_t^i = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ F_t \end{bmatrix} + v_t^i$$

Or more compactly,

$$\mathbf{s_t} = \mathbf{A}\mathbf{s_{t-1}} + \mathbf{B}w_t$$

$$y_t^i = \mathbf{C}\mathbf{s_t} + v_t^i$$

Invoking the Kalman filter

$$\mathbf{s^i_{t|t}} = (\mathbf{I} - \kappa\mathbf{C})\mathbf{A}\mathbf{s^i_{t-1|t-1}} + \kappa\mathbf{C}\mathbf{A}\mathbf{s_{t-1}} + \kappa\mathbf{C}\mathbf{B}w_t + \kappa v_t^i$$

$$\mathbf{s^i_{t|t-1}} = \mathbf{A}\mathbf{s^i_{t-1|t-1}}$$

which implies that

$$x^i_{t|t} = (1-\kappa_1)\rho x^i_{t-1|t-1} + \kappa_1 y_t^i$$

$$F^i_{t|t} = (\lambda - \kappa_2)\rho x^i_{t-1|t-1} + (1-\lambda)\rho F^i_{t-1|t-1} + \kappa_2 y_t^i$$

and

$$x^i_{t|t-1} = \rho x^i_{t-1|t-1}$$

$$F^i_{t|t-1} = \lambda\rho x^i_{t-1|t-1} + (1-\lambda)\rho F^i_{t-1|t-1}$$

where $\lambda = \frac{\kappa_1 + R\kappa_2}{1+R}$ and κ_1, κ_2 are the Kalman gains belonging to the two-dimensional column vector κ. Letting $\xi = \begin{bmatrix} \frac{1}{1+R} & \frac{R}{1+R} \end{bmatrix}$, we have

$$\begin{aligned}
\widetilde{x}^i_{t|t} &= \xi \mathbf{s^i_{t|t}} \\
&= \frac{1}{1+R} x^i_{t|t} + \frac{R}{1+R} F^i_{t|t} \\
&= \frac{1}{1+R}\Big[(1-\kappa_1)\rho x^i_{t-1|t-1} + \kappa_1 y_{it}\Big] + \frac{R}{1+R}\Big[(\lambda - \kappa_2)\rho x^i_{t-1|t-1} + (1-\lambda)\rho F^i_{t-1|t-1} + \kappa_2 y_{it}\Big] \\
&= \frac{1-\kappa_1 + R(\lambda-\kappa_2)}{1+R} x^i_{t|t-1} + \lambda y_{it} + \frac{(1-\lambda)R\rho}{1+R} F^i_{t-1|t-1} \\
&= \frac{1-\lambda}{1+R} x^i_{t|t-1} + \frac{(1-\lambda)R\rho}{1+R} F^i_{t-1|t-1} + \lambda y_{it}
\end{aligned}$$

and

$$\begin{aligned}
\widetilde{x}^i_{t|t-1} &= \xi \mathbf{s^i_{t|t-1}} \\
&= \frac{1}{1+R}\rho x^i_{t-1|t-1} + \frac{R}{1+R}\Big[\lambda \rho x^i_{t-1|t-1} + (1-\lambda)\rho F^i_{t-1|t-1}\Big] \\
&= \frac{1+R\lambda}{1+R} x^i_{t|t-1} + \frac{(1-\lambda)\rho R}{1+R} F^i_{t-1|t-1}
\end{aligned}$$

Hence,

$$\begin{aligned}
\widetilde{x}^i_{t|t} &= \frac{1-\lambda}{1+R} x^i_{t|t-1} + \frac{(1-\lambda)R\rho}{1+R} F^i_{t-1|t-1} + \lambda y^i_t \\
\widetilde{x}^i_{t|t-1} &= \frac{1+R\lambda}{1+R} x^i_{t|t-1} + \frac{(1-\lambda)\rho R}{1+R} F^i_{t-1|t-1}
\end{aligned}$$

Furthermore,

$$F_t = \widetilde{x}_{t|t} = (1-\lambda)\rho F_{t-1} + \lambda x_t$$

$$\widetilde{x}_{t|t-1} = \lambda \rho x_{t-1|t-1} + (1-\lambda)\rho F_{t-1}$$

From here, it follows that

$$
\begin{aligned}
x_t - \widetilde{x}^i_{t|t} &= x_t - \frac{1-\kappa_1+(\lambda-\kappa_2)R}{1+R}x^i_{t|t-1} - \lambda y^i_t - \frac{(1-\lambda)R\rho}{1+R}F^i_{t-1|t-1} \\
&= (1-\lambda)x_t - \frac{1-\kappa_1+(\lambda-\kappa_2)R}{1+R}x^i_{t|t-1} - \lambda v^i_t - \frac{(1-\lambda)R\rho}{1+R}F^i_{t-1|t-1} \\
&= (1-\lambda)x_t - \frac{1-\lambda}{1+R}x^i_{t|t-1} - \frac{(1-\lambda)R\rho}{1+R}F^i_{t-1|t-1} - \lambda v^i_t \\
&= (1-\lambda)\left[x_t - \frac{1}{1+R}x^i_{t|t-1} - \frac{R}{1+R}\rho F^i_{t-1|t-1}\right] - \lambda v^i_t
\end{aligned}
$$

and the forecast revision is

$$
\begin{aligned}
\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} &= \frac{(1-\kappa_1)+(\lambda-\kappa_2)R-(1+R\lambda)}{1+R}x^i_{t|t-1} + \lambda x_t + \lambda v^i_t \\
&= -\frac{\kappa_1-\kappa_2 R}{1+R}x^i_{t|t-1} + \lambda x_t + \lambda v^i_t \\
&= \lambda(x_t - x^i_{t|t-1} + v^i_t)
\end{aligned}
$$

A.3 Proof of Proposition 2

$$
\begin{aligned}
\beta_1^{SI} &= \frac{\text{Cov}(x_t - \widetilde{x}^i_{t|t}, \widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1})}{\text{Var}(\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1})} \\
&= \frac{(1-\lambda)\lambda \text{Cov}(x_t, x_t - x^i_{t|t-1}) - \lambda^2 \sigma_v^2}{\lambda^2 \left[\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2 \right]} \\
&= \frac{1-\lambda}{\lambda} \frac{\text{Var}(x_t) - \text{Cov}(x_t, x^i_{t|t-1})}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} - \frac{\sigma_v^2}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} \\
&= \frac{1-\lambda}{\lambda} \frac{\frac{\sigma_w^2}{1-\rho^2} - \frac{\rho^2 \kappa_1 \cdot \frac{\sigma_w^2}{1-\rho^2}}{1-(1-\kappa_1)\rho^2}}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} - (1-\kappa_1) \\
&= \frac{1-\lambda}{\lambda} \frac{\frac{\sigma_w^2}{1-\rho^2} \frac{1-\rho^2}{1-(1-G)\rho^2}}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} - (1-\kappa_1) \\
&= \frac{1-\lambda}{\lambda} \frac{\text{Var}(x_t - x^i_{t|t-1})}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} - (1-\kappa_1) \\
&= \frac{1-\lambda}{\lambda} \kappa_1 - (1-\kappa_1) \\
&= \frac{1-\kappa_1 + R(1-\kappa_2)}{\kappa_1 + R\kappa_2} \kappa_1 - (1-\kappa_1) \\
&= \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2}
\end{aligned}
$$

A.4 Proof of Proposition 3

Each model delivers a coefficient of error predictability that is a function of parameters θ, R, or α.

$$
\beta_1^{DE} = \frac{-\theta(1+\theta)}{(1+\theta)^2 + \rho^2\theta^2}, \quad \beta_1^{SI} = \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2}, \quad \beta_1^{OC} = \frac{(\alpha^2 - 1)\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2}.
$$

To show how the three models can match the identical error predictability coefficient, β_1, I consider the three pairs of models and draw the implications out in either direction to establish equality.

DE → SI

Given $\{\rho, \sigma_w, \sigma_v, \theta\}$ we set $\beta_1^{DE} = \beta_1^{SI}$ and solve for R:

$$\begin{aligned}
\frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2} &= \beta_1^{DE} \\
R(\kappa_1 - \kappa_2) &= \beta_1^{DE}(\kappa_1 + R\kappa_2) \\
R(\kappa_1 - \kappa_2) &= \beta_1^{DE}\kappa_1 + \beta_1^{DE}R\kappa_2 \\
R(\kappa_1 - \kappa_2) - \beta_1^{DE}R\kappa_2 &= \beta_1^{DE}\kappa_1 \\
R &= \frac{\beta_1^{DE}\kappa_1}{\kappa_1 - (1 + \beta_1^{DE})\kappa_2}
\end{aligned}$$

which, from the definition of β_1^{DE}, is equal to

$$R = \frac{[\kappa_1(1+\theta) - \kappa_2](1+\theta) + (\kappa_1 - \kappa_2)\rho^2\theta^2}{(1+\theta)^2 + \rho^2\theta^2}.$$

SI→DE

Given $\{\rho, \sigma_w, \sigma_v, R\}$ we solve for θ by setting $\beta_1^{DE} = \beta_1^{SI}$

$$\begin{aligned}
\frac{-\theta(1+\theta)}{(1+\theta)^2 + \rho^2\theta^2} &= \beta_1^{SI} \\
-\theta(1+\theta) &= \beta_1^{SI}(1+\theta)^2 + \beta_1^{SI}\rho^2\theta^2 \\
-\theta - \theta^2 &= \beta_1^{SI}(1 + 2\theta + \theta^2) + \beta_1^{SI}\rho^2\theta^2 \\
0 &= \beta_1^{SI} + 2\beta_1^{SI}\theta + \theta + \beta_1^{SI}\theta^2 + \theta^2 + \beta_1^{SI}\rho^2\theta^2 \\
0 &= \beta_1^{SI} + (2\beta_1^{SI} + 1)\theta + [1 + (1+\rho^2)\beta_1^{SI}]\theta^2.
\end{aligned}$$

From the definition of β_1^{SI}, we can express the above quadratic as:

$$\left[\frac{(1+R)\kappa_1 + \rho^2 R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2}\right]\theta^2 + \left(\frac{3R\kappa_1 - R\kappa_2}{\kappa_1 + R\kappa_2}\right)\theta + \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2} = 0$$

DE→OC

Given $\{\rho, \sigma_w, \sigma_v, \theta\}$, we solve for α by setting $\beta_1^{DE} = \beta_1^{OC}$

$$
\begin{aligned}
\frac{(\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2} &= \beta_1^{DE} \\
\alpha^2 &= \frac{\beta_1^{DE}}{\sigma_v^2}[\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2] + 1 \\
\alpha^2 - \frac{\beta_1^{DE}}{\sigma_v^2}\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) &= \beta_1^{DE} + 1.
\end{aligned}
$$

from the definition of β_1^{DE}, the above equality can be expressed as:

$$
\alpha = \sqrt{\frac{(1+\theta)[\sigma_v^2 - \theta\mathbb{V}(x_t - \breve{x}^i_{t|t-1})]}{[(1+\theta)^2 + \rho^2\theta^2]\sigma_v^2}}
$$

Note that due to the recursive nature of the overconfidence model, $\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) \equiv \breve{\Psi}^i_{t|t-1}(\alpha)$ is itself a function of α. As a result, there is no closed form expression that characterizes the mapping from θ to α.

OC→DE

Similar to the result relating SI→DE, we begin with:

$$
0 = \beta_1^{OC} + (2\beta_1^{OC} + 1)\theta + [1 + (1+\rho^2)\beta_1^{OC}]\theta^2.
$$

from the definition of β_1^{OC}, we obtain

$$
0 = \frac{(\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2} + \left(\frac{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + (2\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2}\right)\theta + \left(\frac{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + [\alpha^2 + \rho^2(\alpha^2-1)]\sigma_v^2}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2}\right)\theta^2
$$

SI→OC

Similar to the result relating DE→OC, we have

$$\alpha^2 - \frac{\beta_1^{SI}}{\sigma_v^2}\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) = \beta_1^{SI} + 1,$$

which implies:

$$\alpha = \sqrt{\frac{(1+R)\kappa_1\sigma_v^2 + R(\kappa_1 - \kappa_2)\mathbb{V}(x_t - \breve{x}^i_{t|t-1})}{(\kappa_1 + R\kappa_2)\sigma_v^2}}.$$

Again, there is no closed form expression that maps α to R since the state estimation error, $\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) \equiv \breve{\Psi}(\alpha)$, is a nonlinear function of α.

OC→SI

Finally, given α in the overconfidence model, we can equate β_1^{OC} to β_1^{SI} by setting R equal to

$$R = \frac{(\alpha^2 - 1)\sigma_v^2\kappa_1}{(\kappa_1 - \kappa_2)[\mathbb{V}(x_t - \breve{x}^i_{t|t-1}) + \sigma_v^2] - (\alpha^2 - 1)\sigma_v^2\kappa_2}.$$

A.5 Proof of Proposition 4

Recall that $\gamma_1 = \frac{\mathbb{C}(x^i_{t|t} - x^i_{t|t-1})}{\mathbb{V}(x^i_{t|t-1} - x^i_{t|t-2})}$.

(i) For the overconfidence model, the constant Kalman gain is:

$$\breve{\kappa} = \frac{\mathbb{V}(x_t - \breve{x}^i_{t|t-1}}{\mathbb{V}(x_t - \breve{x}^i_{t|t-1} + \sigma_v^2} \in [0, 1].$$

Given this, we can evaluate γ_1:

$$\begin{aligned}
\gamma_1 &= \frac{\mathbb{C}\left(\breve{\kappa}(x_t - \breve{x}^i_{t|t-1}) + \breve{\kappa}v^i_t, \rho\breve{\kappa}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \rho\breve{\kappa}v^i_{t-1}\right)}{\mathbb{V}\left(\rho\breve{\kappa}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \rho\breve{\kappa}v^i_{t-1}\right)} \\
&= \frac{\mathbb{C}[\breve{\kappa}(x_t - \breve{x}^i_{t|t-1}), \rho\breve{\kappa}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \rho\breve{\kappa}v^i_{t-1}]}{\rho^2\breve{\kappa}^2[\mathbb{V}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \sigma_v^2]} \\
&= \frac{\mathbb{C}[\breve{\kappa}\{\rho(1-\breve{\kappa})(x_{t-1} - \breve{x}^i_{t-1|t-2}) + w_t - \breve{\kappa}\rho v^i_{t-1}\}, \rho\breve{\kappa}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \rho\breve{\kappa}v^i_{t-1}]}{\rho^2\breve{\kappa}^2[\mathbb{V}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \sigma_v^2]} \\
&= \frac{\rho^2\breve{\kappa}^2(1-\breve{\kappa})\mathbb{V}(x_{t-1} - \breve{x}^i_{t-1|t-2}) - \breve{\kappa}^3\rho^2\sigma_v^2}{\rho^2\breve{\kappa}^2[\mathbb{V}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \sigma_v^2]} \\
&= \frac{(1-\breve{\kappa})\mathbb{V}(x_{t-1} - \breve{x}^i_{t-1|t-2}) - \breve{\kappa}\sigma_v^2}{\mathbb{V}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \sigma_v^2} \\
&= \frac{\widehat{\sigma}_v^2\widehat{\kappa} - \sigma_v^2\widehat{\kappa}}{\mathbb{V}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \sigma_v^2} \\
&= \frac{(\alpha^2 - 1)\sigma^2\breve{\kappa}}{\mathbb{V}(x_{t-1} - \breve{x}^i_{t-1|t-2}) + \sigma_v^2} = \breve{\kappa}\beta_1^{OC} \leqslant \beta_1^{OC}.
\end{aligned}$$

(ii) In the diagnostic expectations model, the current-period prediction is:

$$x^{i,\theta}_{t|t} = \mathbb{E}(x_t|\mathcal{I}_t) + \theta\big[\mathbb{E}(x_t|\mathcal{I}^i_t - \mathbb{E}(x_t|\mathcal{I}^i_{t-1})\big],$$

where $\mathbb{E}(x_t|\mathcal{I}^i_t) \equiv x^i_{t|t}$ and $\mathbb{E}(x_t|\mathcal{I}^i_{t-1}) \equiv x^i_{t|t-1}$. The forecast revision is therefore:

$$x^{i,\theta}_{t|t} - x^{i,\theta}_{t|t-1} = (1+\theta)(x^i_{t|t} - x^i_{t|t-1}) - \theta\rho(x^i_{t-1|t-1} - x^i_{t-1|t-2})$$

As a result, the first-order autocovariance of revisions is:

$$\mathbb{C}(x^{i,\theta}_{t|t} - x^{i,\theta}_{t|t-1}, x^{i,\theta}_{t|t-1} - x^{i,\theta}_{t-1|t-2}) = -(1+\theta)\theta\rho^2\mathbb{V}(x^i_{t-1|t-1} - x^i_{t-1|t-2})$$

and the variance of the lagged revision is:

$$\mathbb{V}(x^{i,\theta}_{t|t-1} - x^{i,\theta}_{t-1|t-2}) = \rho^2(1+\theta)^2\mathbb{V}(x^i_{t-1|t-1} - x^i_{t-1|t-2}) + \theta^2\rho^4\mathbb{V}(x^i_{t-2|t-2} - x^i_{t-2|t-3})$$

Noting that the error variance is time invariant, and evaluating the covariance divided by the variance, we obtain:

$$\gamma_1 = \frac{-\theta(1+\theta)}{(1+\theta)^2 + \theta^2\rho^2} = \beta_1^{DE}$$

(iii) It suffices to show that Cov$(\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1}, \widetilde{x}^i_{t|t-1} - \widetilde{x}^i_{t|t-2}) = 0.$ The current period revision is:

$$\begin{aligned}\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} &= \frac{1-\lambda-1-R\lambda}{1+R}x^i_{t|t-1} + \lambda y_{it} \\ &= \lambda(y_{it} - x^i_{t|t-1})\end{aligned}$$

and the previous period revision can be expressed as:

$$
\begin{aligned}
\widetilde{x}^i_{t|t-1} - \widetilde{x}^i_{t|t-2} &= \frac{1}{1+R}(x^i_{t|t-1} - x^i_{t|t-2}) + \frac{R}{1+R}(F^i_{t|t-1} - F^i_{t|t-2}) \\
&= \frac{\rho}{1+R}(x^i_{t-1|t-1} - x^i_{t-1|t-2}) + \frac{R}{1+R}\big[\lambda\rho x^i_{t-1|t-1} + (1-\lambda)\rho F^i_{t-1|t-1} - \lambda\rho x^i_{t-1|t-2} \\
&- (1-\lambda)\rho F^i_{t-1|t-2}\big] \\
&= \frac{\rho}{1+R}\big[\kappa_1(y_{it-1} - x^i_{t-1|t-2})\big] + \frac{R}{1+R}\big[\lambda\rho(x^i_{t-1|t-1} - x^i_{t-1|t-2}) \\
&+ (1-\lambda)\rho(F^i_{t-1|t-1} - F^i_{t-1|t-2})\big] \\
&= \frac{\rho\kappa_1}{1+R}\big[x_{t-1} - x^i_{t-1|t-2} + v_{it-1}\big] + \frac{R}{1+R}\big[\lambda\rho\kappa_1(x_{t-1} - x^i_{t-1|t-2} + v_{it-1}) \\
&+ (1-\lambda)\rho\kappa_2(x_{t-1} - x^i_{t-1|t-2} + v_{it-1})\big] \\
&= \left[\frac{\rho\kappa_1 + R\lambda\rho\kappa_1 + R(1-\lambda)\rho\kappa_2}{1+R}\right](x_{t-1} - x^i_{t-1|t-2} + v_{it-1}) \\
&= \frac{\rho(\kappa_1 + R\kappa_2) + \rho R\lambda(\kappa_1 - \kappa_2)}{1+R}(x_{t-1} - x^i_{t-1|t-2} + v_{it-1}) \\
&= \rho\lambda\left(1 + \frac{R(\kappa_1 - \kappa_2)}{1+R}\right)(x_{t-1} - x^i_{t-1|t-2} + v_{it-1})
\end{aligned}
$$

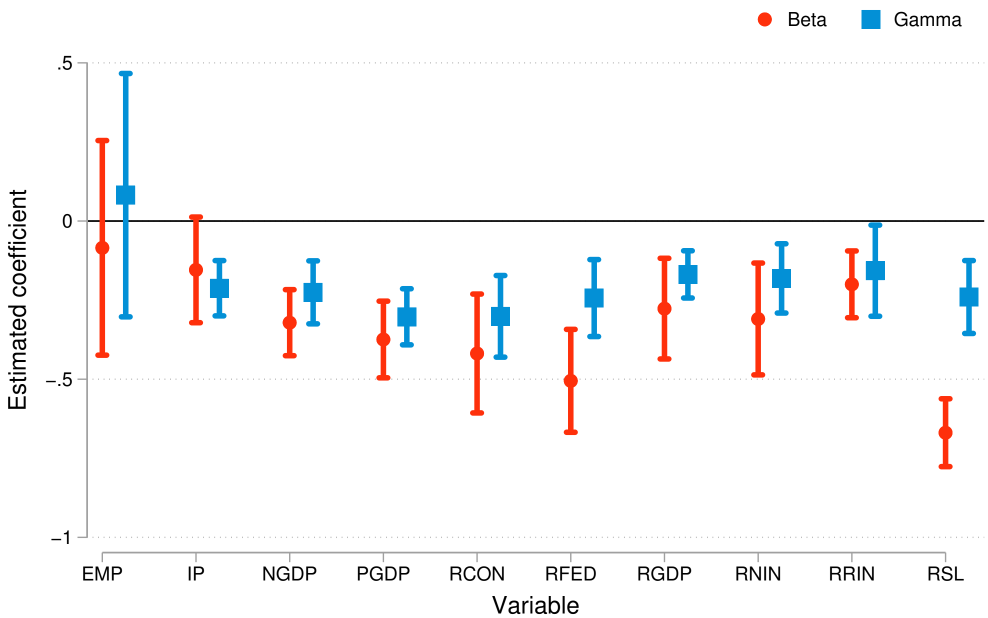
Then,

$$
\begin{aligned}
\mathrm{Cov}(\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1}, \widetilde{x}^i_{t|t-1} - \widetilde{x}^i_{t|t-2}) &= \mathrm{Cov}\left[\lambda(y_{it} - x^i_{t|t-1}), \rho\lambda\left(1 + \frac{R(\kappa_1 - \kappa_2)}{1+R}\right)(x_{t-1} - x^i_{t-1|t-2} + v_{it-1})\right] \\
&= \mathrm{Cov}\Big[\lambda(1-\kappa_1)\rho(x_{t-1} - x^i_{t-1|t-2}) - \lambda\kappa_1\rho v_{it-1} + \lambda(v_{it} + w_t), \\
&\rho\lambda\left(1 + \frac{R(\kappa_1 - \kappa_2)}{1+R}\right)(x_{t-1} - x^i_{t-1|t-2} + v_{it-1})\Big] \\
&= (\lambda\rho)^2\left[1 + \frac{R(\kappa_1 - \kappa_2)}{1+R}\right]\left[(1-\kappa_1)\Psi - \kappa_1\sigma_v^2\right] \\
&= 0
\end{aligned}
$$

By definition, $\kappa_1 = \frac{\Psi}{\Psi + \sigma_v^2}$ where Ψ is the steady state forecast error variance (i.e. the variance that solves the Ricatti equation: $\Psi = (1-\kappa_1)\rho^2\Psi + \sigma_w^2$). From this it follows that the last term in hard brackets is zero.

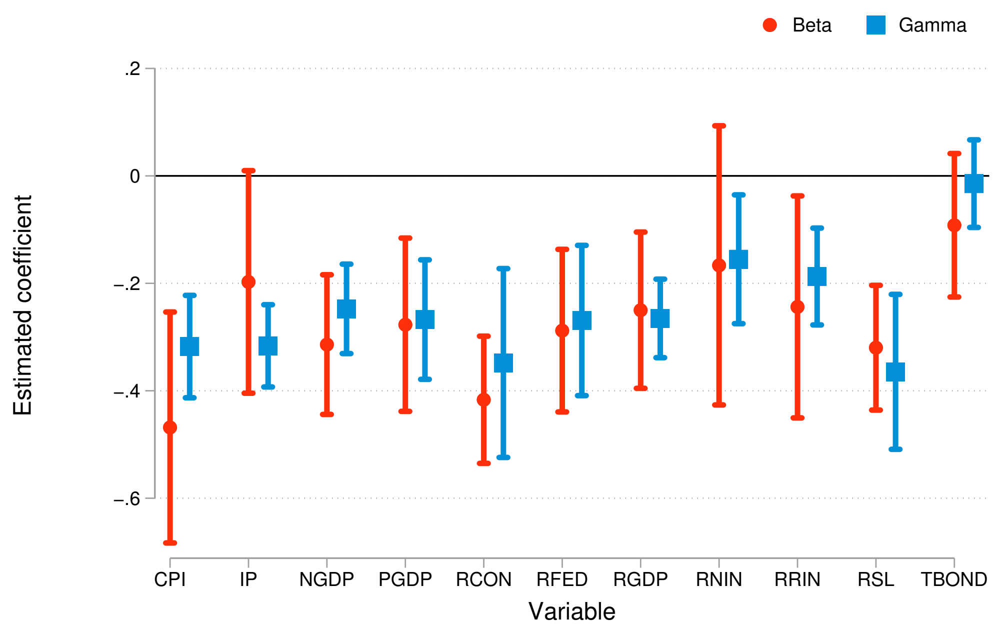
Appendix B Empirics

Figure B1: Error Predictability and Revision Persistence, $H = 0$



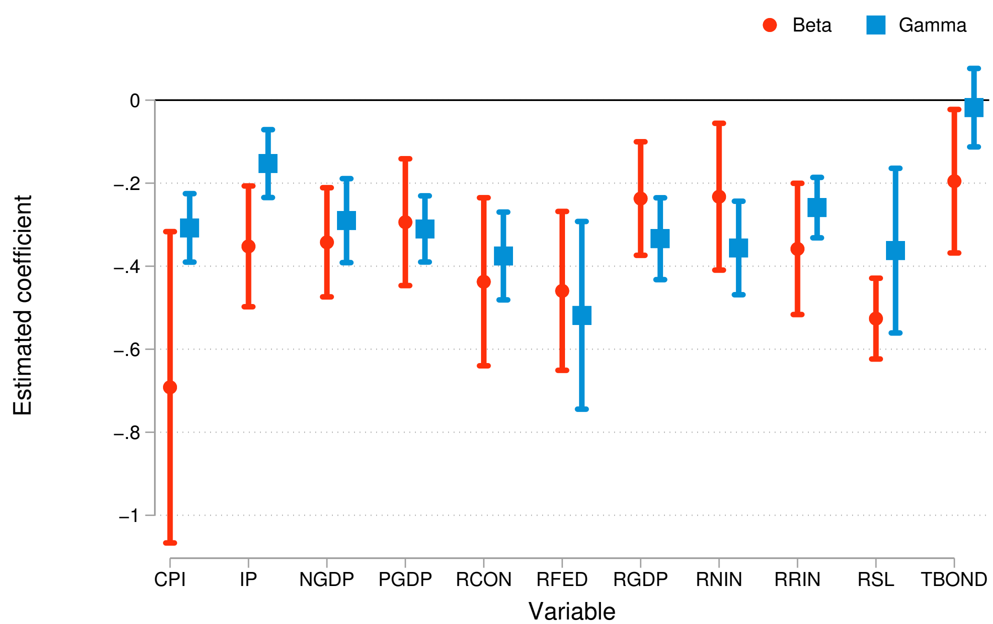
Note: Figure displays 95% confidence intervals for estimated coefficients of error predictability and revision persistence, (β_1 and γ_1 in the main text, respectively). Double clustered standard errors are specified. EMP- Employment, IP - Industrial Production, NGDP - Nominal GDP, PGDP - GDP Deflator, RCON - Real Consumption, RFED - Real Federal Government Spending, RGDP - Real GDP, RNIN - Real Nonresidential Investment, RRIN - Real Residential Investment, and RSL - Real State and Local Government Spending.

Figure B2: Error Predictability and Revision Persistence, $H = 1$



Note: Figure displays 95% confidence intervals for estimated coefficients of error predictability and revision persistence, (β_1 and γ_1 in the main text, respectively). Double clustered standard errors are specified. CPI - Consumer Price Index, IP - Industrial Production, NGDP - Nominal GDP, PGDP - GDP Deflator, RCON - Real Consumption, RFED - Real Federal Government Spending, RGDP - Real GDP, RNIN - Real Nonresidential Investment, RRIN - Real Residential Investment, RSL - Real State and Local Government Spending, and TBOND - 10 Year Government Bond.

Figure B3: Error Predictability and Revision Persistence, $H = 2$



Note: Figure displays 95% confidence intervals for estimated coefficients of error predictability and revision persistence, (β_1 and γ_1 in the main text, respectively). Double clustered standard errors are specified. CPI - Consumer Price Index, IP - Industrial Production, NGDP - Nominal GDP, PGDP - GDP Deflator, RCON - Real Consumption, RFED - Real Federal Government Spending, RGDP - Real GDP, RNIN - Real Nonresidential Investment, RRIN - Real Residential Investment, RSL - Real State and Local Government Spending, and TBOND - 10 Year Government Bond.

Appendix C Alternative Models

C.1 Heterogeneous Priors

In a model with heterogeneous priors (Patton and Timmermann, 2010), forecasters wish to minimize their mean squared forecast errors while tracking their long-run beliefs. Let the prior be μ^i. Then the forecaster's objective is:

$$\min_{\widetilde{x}^i_{t|t}} \left((x_t - \widetilde{x}^i_{t|t})^2 + \omega(\widetilde{x}^i_{t|t} - \mu^i)^2 \right)$$

The optimal reported forecast is:

$$\widetilde{x}^i_{t|t} = \frac{1}{1+\omega}\mathbb{E}(x_t|\mathcal{I}^i_t) + \frac{\omega}{1+\omega}\mu^i.$$

In this case, the deviation from the conditional expectation is $\tau^i_{t|t} = \frac{\omega}{1+\omega}(\mu^i - \mathbb{E}(x_t|\mathcal{I}^i_t)$. Furthermore, this model implies that revisions are unpredictable:

$$\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} = \frac{1}{1+\omega}\left[\mathbb{E}(x_t|\mathcal{I}^i_t) - \mathbb{E}(x_t|\mathcal{I}^i_{t-1})\right]$$

C.2 Measurement Error

In a model of measurement error, or trembling-hand noise as in Branch (2004), implies that the reported forecast is $\widetilde{x}^i_{t|t} = \mathbb{E}(x_t|\mathcal{I}^i_t) + \xi^i_t$. The measurement error, ξ^i_t is distributed $\xi^i_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2_\xi)$. While such models can deliver a negative coefficient of error predictability, the forecast revision is simply $\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} = \xi^i_t - \xi^i_{t-1}$, and hence, unpredictable.

C.3 Smoothing Motives

If forecasters wish to strategically temper their revisions, as in the smoothing motive models of Coibion and Gorodnichenko (2015) and Peters and Kucinskas (2018), then the optimal forecast is

$$\widetilde{x}^i_{t|t}(1-\phi)\mathbb{E}(x_t|\mathcal{I}^i_t) + \phi\widetilde{x}^i_{t|t-1}$$

where $\phi = \frac{\alpha(1-\delta)}{1+\alpha(1-\delta)}$. The parameter α governs the smoothing motive. As $\alpha \to \infty$, the smoothing motive (or quadratic revision cost) increases, and forecasters are less willing to update their forecasts. On the other hand, if $\alpha = 0$, then the smoothing motive disappears. Furthermore, δ is the discount factor.

From this model, the autocorrelation of revisions is simply $\phi > 0$. In other words, revisions are positively autocorrelated. Moreover, the only rational models that can deliver serially correlated revisions are those that explicitly incorporate the forecast revision in the forecaster's loss function. In this case:

$$\mathscr{L} = (x_t - \widetilde{x}^i_{t|t})^2 + \alpha(\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1})^2, \quad \alpha > 0.$$

C.4 Asymmetric Attention

Kohlhas and Walther (2020) present an elegant theory of asymmetric attention which is able to generate overreaction to current output realizations. This theory, however, is unable to replicate overreaction on the basis of forecast revisions. Furthermore, I demonstrate here that it is also unable to generate serially correlated revisions.[15]

I consider a stylized version of the baseline asymmetric attention model in Kohlhas and Walther (2020). In particular, suppose the target forecast variable is a function of two structural components:

$$y_t = x_{1t} + x_{2t},$$

[15]The extended version of this model that incorporates irrational overconfidence, however, would be able to deliver these facts.

where

$$x_{1t} = a_1\theta_t + b_1 u_{1t}$$

$$x_{2t} = a_2\theta_t + b_2 u_{2t}$$

$$\theta_t = \rho\theta_{t-1} + \eta_t,$$

where $u_{jt} \sim \mathcal{N}(0,1)$ and $\eta_t \sim \mathcal{N}(0,\sigma_\eta^2)$. Forecasters do not observe the structural components or the latent state, θ. Instead, they observe two private signals each period,

$$y^i_{1t} = x_{1t} + q_1\epsilon^i_{jt}$$

$$y^i_{2t} = x_{2t} + q_2\epsilon^i_{2t},$$

with $\epsilon^i_{jt} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. The precision of signals is component-specific so that weight placed on new information (i.e. attention) is asymmetric across components. This weight can be defined as:

$$m_j = \frac{b_j^2}{b_j^2 + q_j^2}.$$

As a result, the conditional expectation for a given component is:

$$\mathbb{E}(|\mathcal{I}^i_t) = m_j z^i_{jt} + (1-m_j)\mathbb{E}(x_{jt}|\mathcal{I}^i_{t-1}),$$

and the conditional expectation of the latent state is:

$$\mathbb{E}(\theta_t|\mathcal{I}^i_t) = \mathbb{E}(\theta_t|\mathcal{I}^i_{t-1}) + g_1[z^i_{1t} - \mathbb{E}(z^i_{1t}|\mathcal{I}^i_{t-1})] + g_2[z^i_{2t} - \mathbb{E}(z^i_{2t}|\mathcal{I}^i_{t-1})],$$

where g_j refers to the component-specific Kalman gain (a function of m_j).

From the above set up, one can define the forecast revision for y_{t+1} to be equal to the sum of the revisions made to the unobserved components x_{1t+1} and x_{2t+1}:

$$y^i_{t+1|t} - y^i_{t+1|t-1} = (x^i_{1,t+1|t} - x^i_{1,t+1|t-1}) + (x^i_{2,t+1|t} - x^i_{2,t+1|t-1}).$$

The righthand side of this equation, however, is simply equal to the sum of the innovation errors scaled by the attention coefficients,

$$y^i_{t+1|t} - y^i_{t+1|t-1} = m_1[z^i_{1,t} - \mathbb{E}(x_{1,t}|\mathcal{I}^i_{t-1})] + m_2[z^i_{2,t} - \mathbb{E}(x_{2,t}|\mathcal{I}^i_{t-1})],$$

which are themselves serially uncorrelated as a result of Bayesian updating. Hence, this theory is unable to accommodate autocorrelated revisions.