Time-Varying Volatility, Underreaction, and Overreaction*

Julio Ortiz [†]

December 2020

Abstract

Two seemingly contradictory patterns coexist in data on professional forecasters. After positive news and upward forecast revisions, predictions made by the same person are sometimes systematically too optimistic, "overreacting," while they are also sometimes predictably too pessimistic, "underreacting." Making sense of both patterns within the same model proves difficult for a wide range of theories of belief dynamics. But I show that such patterns are to be expected in an environment with time-varying volatility about which agents are imperfectly informed. In states of the world where volatility exceeds agents' perceptions, forecasters appear to underreact, while states in which volatility is lower than agents perceive cause apparent overreaction. I provide empirical evidence consistent with this mechanism, emphasizing the importance of accounting for the impact of volatility shifts for belief dynamics.

Keywords: Rational expectations. Noisy information. Forecasting. Stochastic volatility. Nonlinear filtering.
JEL Codes: C11, C13, C15 D83, D84

---

*This paper was previously circulated under the title of "Time-Varying Volatility as a Source of Overreactions." I would like to thank Constantin Bürgi, Ryan Charhour, Adam Guren, Tara Sinclair, Stephen Terry, Rosen Valchev for their valuable feedback and helfpul conversations. I would also like to thank the participants of the GW Forecasting Seminar Series, the 2019 annual CEBRA meeting, and the 2019 BC-BU Green Line Macro meeting.

†Boston University, Department of Economics, 270 Bay State Road, Boston, MA 02215; Phone: 201-230-1960; Email: jlortiz@bu.edu.

"...the structure of the economy is constantly evolving in ways that are imperfectly understood by both the public and policymakers..." Ben Bernanke (2007)

1 Introduction

Professional forecasts exhibit error predictability. Specifically, the covariance between ex-post errors and ex-ante revisions is non-zero and can run in either direction. A negative covariance is interpreted as an overreaction whereas a positive covariance is interpreted as an underreaction. At the same time, macroeconomic and financial time series have been found to exhibit complex dynamics such as stochastic volatility, structural breaks, and regime switching. Whereas existing models of belief formation do not generally accommodate simultaneous over- and underreactions, I show that these patterns can arise in an otherwise standard noisy information setting that incorporates unobserved time-varying volatility and heterogeneous forecasting techniques.

Error predictability in the model arises due to updating mistakes committed by forecasters. With unobserved volatility, the optimal weight to place on new information is not exactly known. In addition, there are costs associated with devising quantitative predictions. For instance, producing a forecast requires (computing) time and cognitive effort. To the extent that macroeconomic dynamics vary in their complexity, it stands to reason that forecasters tailor their models to each time series. Furthermore, subject to these costs, forecasters may select simpler misspecified models. Therefore with time-varying volatility and heterogeneous forecasting models, revisions can hold predictive power over errors, and the nature of this relationship can be variable-specific. In spite of this, my framework is compatible with rationality in the sense that reported forecasts are optimal outcomes.

Survey data has traditionally been used to test theories of expectation formation. In this paper, I make use of the Survey of Professional Forecasters (SPF) which provides a panel of multi-horizon forecasts across several macroeconomic variables.[1] In the data, over- and underreactions arise along

[1] Examples other than the SPF include the Livingston survey, the Michigan Survey of Consumers, the NY Fed Survey of Consumer Expectations, Blue Chip forecasts, the ECB Survey of Professional Forecasters, and the daily Focus Survey from the Central Bank of Brazil, among others.

different dimensions. First, across levels of aggregation, consensus forecasts broadly exhibit underreactions while forecaster-level predictions tend to imply overreactions. Second, across variables and forecasters, overreactions appear for some variables and underreactions for others. Both of these facts have been previously documented in the literature. This paper offers a third empirical fact: across variable, within forecaster, the same respondent appears to over- and underreact to distinct macroeconomic variables.

The presence of simultaneous over- and underreactions prompts several fundamental questions about belief formation. Are professional forecasters, presumably the most informed private agents in the economy, rational? Alternatively, do behavioral biases govern the manner in which expectations are formed? As policymakers increasingly pursue expectations-based policies such as forward guidance, taking a step toward reconciling theories of expectations formation with the data is of first-order importance.

Against this backdrop, I develop a noisy information model with unobserved time-varying volatility. Rather than obtaining an exact solution to the optimal inference problem, forecasters must approximate the posterior distribution. They may choose from a finite set of approximation methods. The available methods vary in complexity, and adopting a given method is subject to a cost that is increasing in forecasting model sophistication. Forecasters generate a prediction that minimizes the sum of their mean squared errors and model adoption costs.

Importantly, some forecasters adopt suboptimal forecasting models to predict variables thereby generating error predictability. I consider a stylized version of this unobserved volatility model in which forecasters can select either a suboptimal Kalman filter or an asymptotically efficient particle filter, the former being less costly to adopt than the latter.[2] I find that the underlying signal-to-noise ratio governs the extent to which over- and underreactions arise. Intuitively, the optimal weight to place on new information is increasing in the time-varying signal-to-noise ratio. Predictions based on the suboptimal model, however, erroneously update new information in a constant fashion. As a result, forecasters will tend to underreact to variables for which the average signal-to-noise ratio is

[2]The particle filter is an asymptotically efficient approximation method. I provide further details on the filter in Appendix D.

high, and will overreact to variables for which the average signal-to-noise ratio is low. Put another way, there are certain features inherent to a given macroeconomic time series that explain why forecasters appear to either over- or underreact to that particular variable.

After providing simulation results that confirm the above intuition, I examine these implications in the data by exploiting the cross-section of macroeconomic variables for which forecasters report predictions in the SPF. I then parameterize the stylized model and show that it can match a quantitatively relevant share of simultaneous over- and underreactions. Taken together, my findings demonstrate that time-varying volatility, coupled with costly forecast model adoption, can rationalize important features of survey expectations data.

In their seminal paper, Coibion and Gorodnichenko (2015), henceforth CG, make sense of forecast error inefficiency while preserving the assumption of rationality. Using consensus-level data, CG show that projecting ex-post forecast errors on ex-ante forecast revisions delivers an estimate of information rigidity. More recently, many studies have used forecaster-level data to test for rationality.[3] In doing so, much of this literature preserves the linearity assumption made in CG, and ultimately rejects rational expectations even under imperfect information. To make sense error predictability at the forecaster-level while also matching the CG finding of underreactions at the aggregate level, several theories of non-rational expectations have been proposed.[4] My paper relates to this strand of the literature in many ways, and provides an alternate interpretation of the errors-on-revisions coefficient.

In a recent contribution, Kohlhas and Walther (2020) also examines simultaneous over- and underreactions. The authors are able to explain overreaction to news coupled with underreaction on average with a model of asymmetric attention. Although I ground over- and underreactions from a slightly different empirical perspective, I view my paper as complementary to theirs. Whereas my model is based on heterogeneity in the underlying volatility across state variables, Kohlhas and

[3]Examples include Bordalo et al. (2020), Fuhrer (2018), Dovern et al. (2015), Andrade and Bihan (2013), Broer and Kohlhas (2019), Bürgi (2016).

[4]For instance Bordalo et al. (2020) rule out rationality in favor of diagnostic expectations. Other studies such as Fuster et al. (2012) argue in favor of models featuring misperception at long horizons. Daniel et al. (1998) argues for a model of overconfidence while Broer and Kohlhas (2019) present a model of relative overconfidence.

Walther (2020) present a model of costly attention which delivers distinct signal precisions for different components of the state. In both cases, however, the underlying signal-to-noise ratio is the relevant object that varies across variables or components.[5] This paper is also related to Gabaix (2018) which proposes a model in which agents over- and underreact due to misperceived persistence of the data generating process. The focus of my model, however, is in how forecasters assess volatility. Nonetheless, my model can in general speak to sources of misperceived persistence, for example, unobserved structural breaks.

Moreover, the unobserved volatility noisy information model in this paper is in the spirit of Branch (2004), Evans and Ramey (1992) and Brock and Hommes (1997) who define adaptively rational equilibrium dynamics (ARED). Branch (2004) was the first to introduce this concept to expectations formation. My paper builds on his important insights in key ways. First, I present more complex dynamics for the state variable. Introducing nonlinearities, such as stochastic volatility, provides an even stronger justification for the use of different predictor functions. Second, I explicitly model heterogenous expectations through private information whereas in Branch (2004) predictions are assumed to be homogeneous among all who adopt a specific predictor function.[6] Taken together, my model is able to reproduce the empirical facts relating to simultaneous over- and underreactions across level of aggregation, by variable across forecasters, and by variable within forecaster.

While a discussion of nonlinearities has generally been absent in the survey expectations literature, the finance literature has previously tied nonlinear dynamics to error predictability. For instance, Lewis (1989) considers error predictability concerning dollar forecasts in the context of a structural break. Veronesi (2015) finds that over- and underreactions arise in a regime switching model of asset pricing. More recently, Lansing et al. (2020) attribute the predictability of excess returns to either volatility or deviations from rationality. To this end, my paper also relates to the literature on volatility in macroeconomics.[7]

[5] Relatedly, Broer and Kohlhas (2019) present a model of over- and underreactions. The focus in this paper is to match simultaneous over- and underreactions to endogenous public signals.

[6] Heterogeneity in this model comes from idiosyncratic "trembles" in the reported prediction.

[7] See for instance, Justiniano and Primiceri (2008), Kim and Nelson (1999), McConnell and Perez-Quiros (2000),

Finally, due to the assumptions imposed on the state dynamics, this paper relates to the literature on nonlinear filtering. Several approximation methods have been devised in order to deal with nonlinearities in the evolution of a state variable. These methods include generalizations to Kalman filtering as well as importance sampling algorithms, among others.[8] A strand of this literature has formalized some basic efficiency properties of particle filtering.[9]

The rest of the paper is organized as follows. Section 2 presents previously documented facts about error predictability at the forecaster and consensus levels, as well as a novel fact pertaining to simultaneous over- and underreactions. Section 3 presents the noisy information model subject to unobserved time-varying volatility. Section 4 introduces a stylized version of the model and provides simulation results. Section 5 documents empirical evidence consistent with the model. Section 6 parameterizes the stylized model to show that it can generate within forecaster over- and underreactions. Finally, Section 7 concludes.

2 Evidence from Survey Data

The SPF is a quarterly survey provided by the Federal Reserve Bank of Philadelphia. The survey began in 1968Q4 and provides forecasts from several forecasters across a number of macroeconomic variables over many horizons, h. The variables of interest in this paper are the forecast error and the forecast revision. To construct forecast errors from forecaster i about variable x,

$$FE^i_{t+h,t} = x_{t+h} - x^i_{t+h|t},$$

I take the difference between the realized real-time value for x at $t + h$ and the forecaster's h-step ahead prediction generated at time t. To compute forecast revisions, I exploit the term structure of

and Stock and Watson (2007)

[8]Julier and Uhlmann (2004) develop a Kalman filter for nonlinear settings while Doucet and Johansen (2009) discuss particle filtering methods.

[9]See Crisan and Doucet (2002) and Hu et al. (2011)

Table 1: Pooled OLS Forecast Error Predictability Regressions

	Nowcast		One-Quarter Ahead		Two-Quarters Ahead	
	β_1	α_1	β_1	α_1	β_1	α_1
Estimate	-0.317***	0.569***	-0.231**	1.011***	-0.344***	0.565**
	(0.050)	(0.128)	(0.067)	(0.201)	(0.058)	(0.272)
Observations	65,070	2,323	54,067	2,309	52,220	2,295

Note: The table reports the estimated coefficients of forecast error predictability at the current, one-, and two-quarter ahead horizons. Across all horizons, column (1), refers to the forecaster-level errors-on-revisions regression. Column (2) refers to the consensus-level errors-on-revisions regression. Standard errors for forecaster-level regressions are as in Driscoll and Kraay (1998), while Newey-West standard errors are used for aggregate-level specifications. Data used for estimation come from SPF. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

forecasts generated by the survey respondents

$$FR^i_{t,t-1} = x^i_{t+h|t} - x^i_{t+h|t-1}.$$

This requires making use of the h-step ahead forecasts formulated in periods t and $t-1$. In other words, I consider the fixed horizon, h and take the difference between two adjacent forecasts.

In the data, simultaneous over- and underreactions arise along different dimensions: across level of aggregation, across SPF variables pooled over forecasters, and across SPF variables within forecaster. CG present the following testable implication at the consensus-level which holds for an arbitrary horizon:

$$FE_{t+h,t} = \alpha_0 + \alpha_1 FR_{t+h,t} + \epsilon_t. \tag{1}$$

CG find that in the data, $\alpha_1 > 0$ for most variables which indicates that consensus forecasts underreact to new information. More recently, Bordalo et al. (2020) estimate the same regression at the forecaster-level:

$$FE^i_{t+h,t} = \beta_0 + \beta_1 FR^i_{t+h,t} + \varepsilon^i_t, \tag{2}$$

and find that $\beta_1 < 0$ for most macroeconomic series. The interpretation is that forecasters overreact to new information.

Table 1 reports estimates of β_1 and α_1, using data from the SPF. The estimates are obtained

Table 2: Pooled OLS Regressions at $h = 0$, by Variable

Variable	Mnemonic	β_1	α_1
Consumer price inflation	CPI	-0.085	0.868***
Employment	EMP	-0.123	0.564***
Housing starts	HOUSING	0.063	0.359***
Industrial production	IP	-0.147*	0.513***
Nominal GDP	NGDP	-0.310***	0.421**
GDP Deflator	PGDP	-0.363***	0.350**
Real consumption	RCONSUM	-0.401***	0.098
Real federal government spending	RFEDGOV	-0.483***	0.377
Real GDP	RGDP	-0.264***	0.350**
Real nonresidential investment	RNRESIN	-0.499**	0.362
Real residential investment	RRESINV	-0.234***	0.925***
Real state/local government spending	RSLGOV	-0.660***	-0.381
3-month Treasury bill	TBILL	0.010	0.178***
10-year Treasury bond	TBOND	0.020	0.154***
Unemployment rate	UNEMP	0.082**	0.247***

Note: The table reports the OLS coefficients from errors-on-revisions regressions across 15 macroeconomic variables reported in the Survey of Professional Forecaters. Column (3) reports the coefficient in front of the revision at the forecaster-level while column (4) reports the analogous coefficient using consensus-level data. The errors and revisions are for current period forecasts ($h = 0$). All forecasts refer to growth rates with the exception of consumer price inflation (CPI), 3-month treasury bill (TBILL), 10-year bond (TBOND), and unemployment rate (UNEMP). *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

via OLS regressions, pooling across both forecasters and macroeconomic variables. Estimates are reported for three different horizons. Across all horizons considered, it is clear that overreactions dominate at the individual-level, while underreactions arise at the aggregate-level.

However even at the forecaster-level there is evidence of simultaneous over- and underreactions across macroeconomic variables. Table 2 reports variable-by-variable results for nowcasts ($h = 0$) at the forecaster- and consensus-levels. The results point to individual overreactions for most variables but underreactions for some variables such as the unemployment rate.

These findings are neither driven by entry and exit among SPF forecasters, nor different forecasters systematically reporting predictions for select macroeconomic variables. Instead, the same respondent simultaneously over- and underreacts to distinct macroeconomic variables. To show this, I estimate regression (2) for each forecaster i forecasting a specific variable j. This delivers an $N_i \times N_j$ matrix of estimates $\widehat{\beta}_{1,ij}$. I keep only those estimates that are significant at the 5%

level. I then fix a pair of SPF variables j and k, and compute the number of forecasters such that $\widehat{\beta}_{1,ij} < 0$ and $\widehat{\beta}_{1,ik} > 0$, and normalize by the number of total forecasters reporting predictions about variables j and k. More formally, I estimate a matrix P whose elements are p_{jk} with

$$p_{jk} = \frac{\sum_i \mathbb{1}\left(\widehat{\beta}_{1,ij} < 0 \text{ and } \widehat{\beta}_{1,ik} > 0\right)}{\min\{N_j, N_k\}}$$

where N_x denotes the number of forecasters providing predictions of variable x and $\mathbb{1}(\cdot)$ is the indicator function. The elements of matrix P, therefore, denote the share of forecasters who simultaneously overreact to the row variable and underreact to the column variable. When p_{jk} is close to one, this means that nearly all forecasters overreact to variable j and underreact to variable k. On the other hand, when p_{jk} is close to zero, then almost no forecaster overreacts to variable j while also underreacting to variable k.

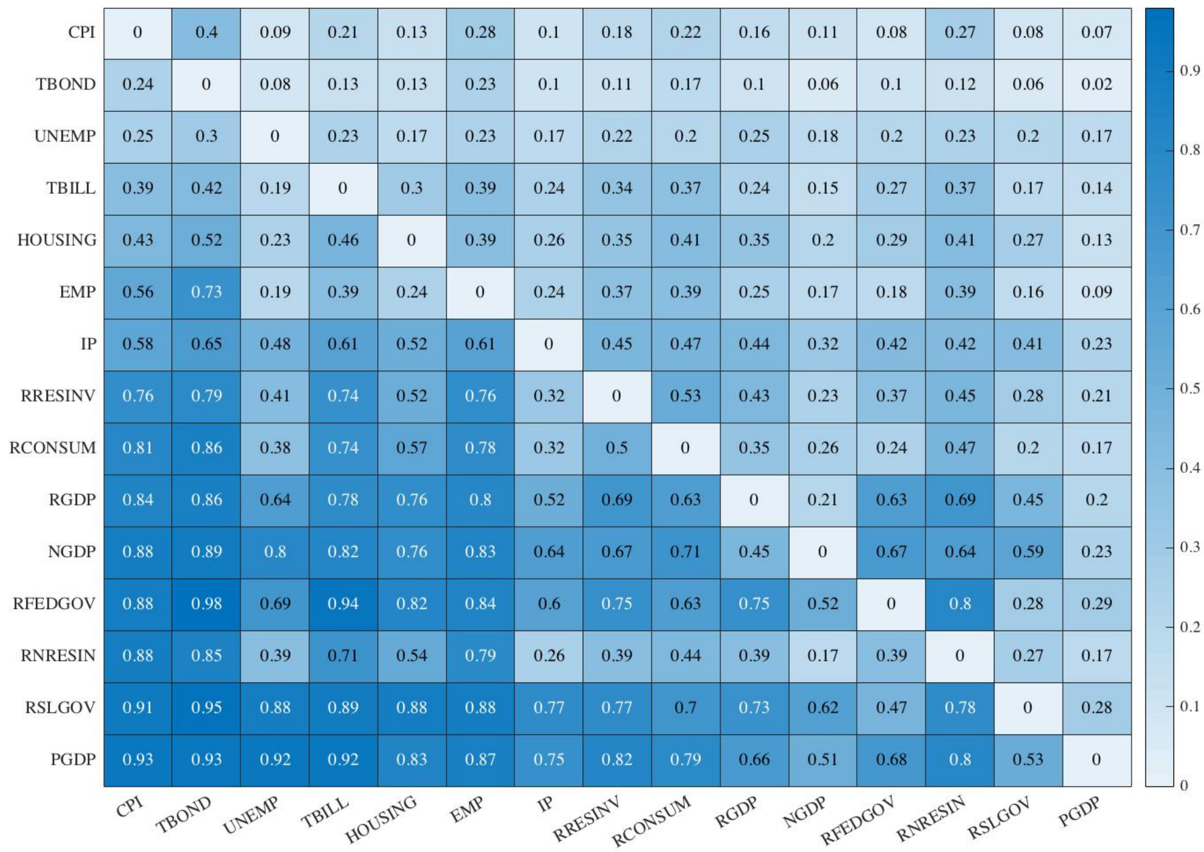
Figure 1 reports the results from this exercise. The heatmap verifies that a given forecaster tends to overreact to some variables and underreact to others. For instance, 84% of professional forecasters exhibit overreactions when forecasting growth in real gross domestic product (RGDP) while simultaneously underreacting to information regarding inflation based on the consumer price index (CPI).

In order to understand how individuals formulate these expectations, a theory of expectations formation must take into account that a single agent may overreact and underreact to different variables. I propose a model of time-varying volatility and heterogeneity in forecasting models in order to account for this fact.

3 Model

I begin with a simple regime switching example to highlight the intuition. I then proceed to develop a generalized noisy information rational expectations model with unobserved time-varying volatility and heterogeneous forecasting methods. In the next section, I narrow my focus to a stylized

Figure 1: Frequency of Over- and Underreaction



Note: The heatmap displays the share of forecasters who overreact to the row variable and simultaneously underreact to the column variable.

version of this model which I later parameterize.

3.1 A Simple Model

Suppose that the state is described as follows:

$$s_t = w_t \quad w_t \sim \mathcal{N}(0, \sigma_t^2),$$

where

$$\sigma_t^2 = \begin{cases} \sigma_L^2 & \text{with probability } q \\ \sigma_H^2 & \text{with probability } 1-q. \end{cases}$$

The forecaster cannot directly observe the latent state s_t or its volatility σ_t^2. More specifically, the probability q is unknown. Instead, the forecaster receives a signal each period that is contaminated with noise:

$$y_t = s_t + v_t \quad v_t \sim \mathcal{N}(0, \sigma_v^2).$$

Since the state is i.i.d., the optimal weight to place on new information (the Kalman gain) is:

$$\kappa_t = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_v^2}.$$

Without full knowledge of the probability q, the optimal weighting of new information is unknown. Suppose that the forecaster finds it too time consuming or otherwise prohibitive to ascertain q. Instead, the forecaster simply assess the state variance to be $\sigma^2 = \frac{1}{2}(\sigma_L^2 + \sigma_H^2)$. The associated Kalman gain is

$$\kappa = \frac{\sigma^2}{\sigma^2 + \sigma_v^2}.$$

The resulting weighting error can be expressed as:

$$\kappa_t - \kappa = \frac{[\sigma_t^2 - \frac{1}{2}(\sigma_L^2 + \sigma_H^2)]\sigma_v^2}{[\frac{1}{2}(\sigma_L^2 + \sigma_H^2) + \sigma_v^2](\sigma_t^2 + \sigma_v^2)}.$$

When $\sigma_t^2 = \sigma_L^2$, the weighting error is negative meaning that the forecaster puts undue weight on his signal thereby overreacting. On the other hand, when $\sigma_t^2 = \sigma_H^2$, the weighting error is positive and the forecaster underreacts to new information.

Based on this simple example, the magnitude of the overreaction depends importantly on the signal-to-noise ratio. Noisier environments deliver more negative β_1 coefficients. On the other hand, underreactions arise when q is closer to zero (i.e., the underlying signal-to-noise ratio is high). In this case, as the forecaster generates predictions, he believes the state to be less variable than it truly is thereby placing less weight on news. This mutes the effects of signal noise and creates more inertia in expectation formation than is optimal. The result is a more positive β_1 coefficient.

The weighting errors in my model stem from: (i) unobserved volatility and (ii) incentives to adopt parsimonious approximations of the volatility. With these two assumptions, over- and underreactions can arise depending on the underlying state dynamics. Assumption (ii) is crucial because if forecasters could easily observe q, then there would be no need to approximate the variance of the state. In this case, error orthogonality would hold despite the regime switching nature of the variance. However, if adopting different forecasting models comes at a cost (be it cognitive, timing, or otherwise), forecasters may find it optimal to make use of such approximations.

3.2 A Model of Unobserved Time-Varying Volatility

Having illustrated the basic intuition that delivers simultaneous over- and underreactions, I now turn to presenting the general model. Nonlinearities such as time-varying volatility in the underlying state complicates the forecaster's problem as he must now formulate expectations about levels and the volatilities. Supposing that there are n latent state variables and m exogenous signals, the state and observations equations are:

$$\begin{aligned} \overline{\mathbf{s}}_{\mathbf{t}} &= F(\overline{\mathbf{s}}_{\mathbf{t-1}}, \overline{\mathbf{w}}_{\mathbf{t}}) \\ \mathbf{z}^{\mathbf{i}}_{\mathbf{t}} &= \mathbf{C}\overline{\mathbf{s}}_{\mathbf{t}} + \mathbf{D}\mathbf{v}^{\mathbf{i}}_{\mathbf{t}}, \end{aligned} \tag{3}$$

where $\overline{\mathbf{s}}_{\mathbf{t}}$ is an $n \times 1$ vector, $\mathbf{z}_{\mathbf{t}}$ is $m \times 1$, $\mathbf{C}$ is $m \times n$, $\mathbf{D}$ is $m \times m$ and $\mathbf{v}^{\mathbf{i}}_{\mathbf{t}}$ is $m \times 1$. There are no other restrictions placed on the model. In particular, $\mathbf{s}_{\mathbf{t}}$ can be a vector of many different state variables, or lags of itself. Furthermore, $\mathbf{z}^{\mathbf{i}}_{\mathbf{t}}$ can include an arbitrary finite number of observed signals. The noise vector $\mathbf{v_t^i}$ can include private or public noise.[10]

In a linear model, $\mathbf{s_t} = \mathbf{As_{t-1}} + \mathbf{Bw_t}$. The crucial difference between a linear model and this one is the unobserved time-varying covariance matrix $\mathbf{B}_t$ which implies an expanded state space, $\mathbf{\bar{s}_t} = \begin{pmatrix}\mathbf{s_t} & diag(\mathbf{B_t})\end{pmatrix}'$. As a result, the error now enters *multiplicatively* into the state. This nonlinearity is modeled by the function $F(\cdot)$ which governs the evolution of the state. While the state now exhibits stochastic volatility, the shocks remain normal, and the signal structure is unchanged. Hence, the measurement equation remains linear.[11] [12]

Whereas Kalman filtering delivers an exact optimal solution in a linear Gaussian environment, it is no longer optimal in this context. The reason for this is that the Kalman filter requires one to evaluate the expected value of $\mathbf{\bar{s}_t}$ conditional on the history of signals $\mathcal{Z}_t^i = \{\mathbf{z_1^i}, \ldots, \mathbf{z_t^i}\}$. This is made intractable due to the lack of knowledge about the underlying conditional distribution. To see this more clearly, consider the scalar case where the state is s_t and there is only a private signal available to the forecaster, z_t^i. The observation equation can be expressed as a conditional likelihood, $p(z_t^i|s_t)$ and the state evolution as $p(s_{t+1}|s_t)$. The optimal filter computes $p(s_t|\mathcal{Z}_t^i)$ from a predict-update procedure

$$p(s_t|\mathcal{Z}_{t-1}^i) = \int p(s_t|s_{t-1})p(s_{t-1}|\mathcal{Z}_{t-1}^i)ds_{t-1} \quad \text{(Predict)}$$

$$p(s_t|\mathcal{Z}_t^i) = \frac{p(z_t^i|s_t)p(s_t|\mathcal{Z}_{t-1}^i)}{p(z_t^i|\mathcal{Z}_{t-1}^i)} \quad \text{(Update)}$$

where $p(z_t|Z_{t-1}^i) = \int p(z_t^i|s_t)p(s_t|Z_{t-1}^i)ds_t$.

In a linear Gaussian environment, this can be exactly computed via the Kalman recursions.[13] In a nonlinear setting, however, computing $p(s_t|\mathcal{Z}_t^i)$ is not feasible as the density cannot be obtained analytically.

In light of this, forecasters must approximate the nonlinear state. I assume that this is done

[10] I index this vector by i in general to allow for forecaster-specific signals.

[11] This could be generalized to a nonlinear measurement as well. I abstract away from this for simplicity.

[12] While I consider stochastic volatility, any nonlinearity can deliver the results presented in the paper. In particular, this model can also speak to unobserved changes in the persistence of macroeconomic time series.

[13] In particular, $p(s_t|\mathcal{Z}_t^i) \sim \mathcal{N}(s_{t|t}^i, \Psi_{t|t}^i)$, where $s_{t|t}^i$ is the expected value of the posterior density and $\Psi_{t|t}^i$ is the variance.

by selecting from a set of costly approximation functions, $A \in \mathcal{A}$. Forecasters first select an approximation function so as to obtain an estimate of the posterior density of the underlying state. Forecasters then report their predictions, the first moment of this approximated density. Hence, the forecaster's loss function can be defined as

$$\mathcal{L} = \min_{A \in \mathcal{A}} \left[\left(\mathbf{z^i_{t+h}} - \widehat{\mathbf{z}}^{\mathbf{i,A}}_{\mathbf{t+h|t}}\right)' \left(\mathbf{z^i_{t+h}} - \widehat{\mathbf{z}}^{\mathbf{i,A}}_{\mathbf{t+h|t}}\right) + c^i_A \right], \tag{4}$$

where the first term is the mean square error arising from individual i's forecast which makes use of approximation function A, and the second term denotes the cost associated with adopting approximation function A.[14] I assume that these forecaster-specific costs are drawn randomly $c^i_A \sim U(0, \bar{c}_A)$.[15] This cost embodies unobserved heterogeneity among forecasters that result in the use of different forecasting models.

After applying their approximations of the state, forecasters generate a prediction and an update according to the new information received. Since agents are formulating a forecast subject to an approximation of the state, I call these approximate predictions. An approximate prediction is defined as

$$\widehat{\bar{\mathbf{s}}}^{\mathbf{i}}_{\mathbf{t|t}} = \int \bar{\mathbf{s}}_{\mathbf{t}}\, \widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}} | \mathcal{Z}^i_t)\, d\bar{\mathbf{s}}_{\mathbf{t}}. \tag{5}$$

In essence, the forecaster predicts the current state according to the approximated density $\widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}} | \mathcal{Z}^i_t)$. In a linear Gaussian setting, the density would be obtained exactly so that $\widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}} | \mathcal{Z}^i_t) = p(\bar{\mathbf{s}}_{\mathbf{t}} | \mathcal{Z}^i_t)$ and errors would be orthogonal.

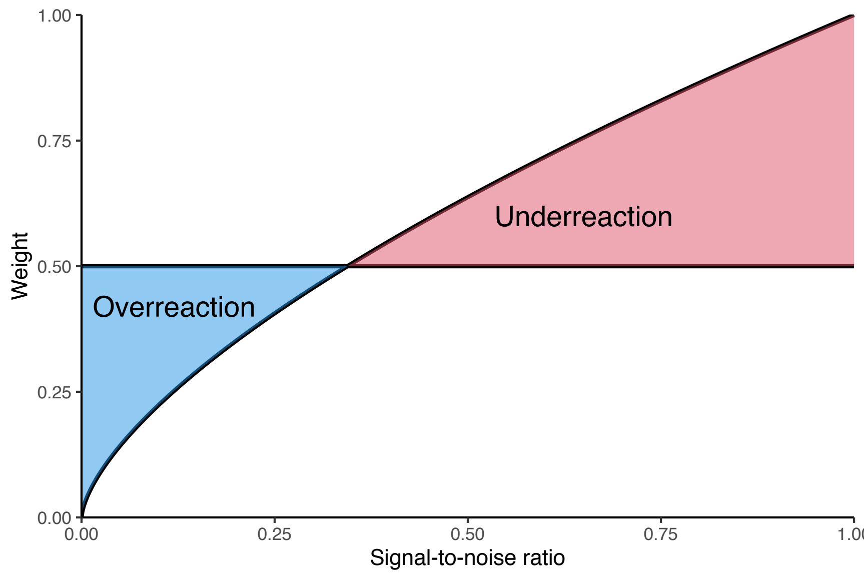
One can express the approximate prediction as a deviation from the optimal minimum mean square error forecast

$$\widehat{\bar{\mathbf{s}}}^{i}_{t|t} = \underbrace{\mathbb{E}(\bar{\mathbf{s}}_t | \mathcal{Z}^i_t)}_{\text{Optimal}} + \underbrace{\int \bar{\mathbf{s}}_t [\widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}} | \mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \mathbf{p}(\bar{\mathbf{s}}_{\mathbf{t}} | \mathcal{Z}^{\mathbf{i}}_{\mathbf{t}})] d\bar{\mathbf{s}}_{\mathbf{t}}}_{\text{Approximation error}}. \tag{6}$$

[14]Forecasters have knowledge of the mean square error associated with each A.

[15]One could alternatively assume heterogeneous signal precision. Models featuring heterogeneous signal-to-noise ratios have been proposed in the literature, particularly to explain forecast disagreement.

Figure 2: Scope for Over- and Underreaction



Note: The figure illustrates the optimal Kalman gain (upward sloping curve) and the suboptimal Kalman gain (horizontal line) as functions of the signal-to-noise ratio.

Whereas existing theories of expectation formation restrict the deviation from the optimal forecast to be either positive (overreactions) or negative (underreactions), the direction of the error here is unrestricted.

3.3 Scope for Over- and Underreaction

Error predictability is due to the presence of suboptimal models in the set $\mathcal{A}$. Importantly, some forecasters must select suboptimal approximations from this menu of models. A sufficient condition for the presence of over- and underreactions is that some approximation $A \in \mathcal{A}$ involves a time-invariant weight used to update new information which resides in the interval $(0, 1)$.

Intuitively, the optimal weight to place on new information is increasing in the signal-to-noise ratio. If a forecaster partially incorporates the most recent signal with a naive constant weight, then the forecaster will overweight new information when signals tend to be imprecise, and will underweight new information when signals tend to be more precise. Figure 2 illustrates this intuition.

The gaps between the curve and the horizontal line reflect weighting errors. These weighting errors are made each period among those who choose the suboptimal constant-weight forecasting model. The scope for underreaction rises with the signal-to-noise ratio.[16]

The constant weight assumption is made for the purpose of transparency. There are other forecasting techniques that could deliver misspecified time-varying weights which could deliver a similar intuition to the one displayed in Figure 2.

Aggregate error predictability arises because the consensus revision does not reside in any forecaster's information set. Appendix B proves that $\alpha_1 \geq \beta_1$ which implies that underreaction can arise at the aggregate level. Aggregate unnderreactions arise in part due to aggregation, but also because not all forecasters adopt suboptimal models in this noisy information environment.

3.4 Relation to Some Theories of Expectation Formation

Amid the mounting evidence against full information rational expectations, several theories of expectations formation have been proposed in the literature. Here, I consider a few prominent theories and assess their ability to generate the empirical facts presented in Section 2.[17]

According to diagnostic expectations, forecasters over-weight new information according to a parameter $\theta > 0$ which ultimately governs the extent of overreaction. This parameter comes from the representativeness heuristic of Tversky and Kahneman (1974). The diagnostic nowcast is defined as:

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

This theory makes use of a distorted Kalman filter called the diagnostic Kalman filter, which, as shown in Bordalo et al. (2020), is able to generate $\beta_1 < 0$ and $\alpha_1 > 0$. Because $\theta > 0$, however, this theory cannot accommodate underreactions. In other words, the sign restriction imposed on θ implies that diagnostic errors and revisions always covary negatively.

[16] It is important to note that the optimal weight will vary along the curve according to the underlying stochastic volatility. Ultimately, the optimal weight is a complicated function of the volatility and persistence of the state as well as the noisiness of the signals, so the black curve is specific to the macroeconomic variable in question.

[17] I do not include a formal discussion of sticky information or linear noisy information models as their shortcomings in this respect have already been documented in the literature. Appendix B provides some detail on why these popular models are at odds with the data. In particular, they imply error orthogonality at the forecaster-level.

Models of overconfidence also distort the Kalman gain. The distortion stemming from overconfidence, however, is different. According to this theory, forecasters misperceive the precision of their own signals. Suppose that forecasters observe only one private signal:

$$z_t^i = s_t + v_t^i, \qquad v_t^i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_v^2),$$

but they perceive $\widetilde{\sigma}_v < \sigma_v$. Put another way, individual forecasters believe their own private signals to be more precise than they truly are. This results in an erroneous assessment of the noise in the system. Overconfident beliefs are recursive so that the distorted gain injects a bias to the update in each period. These beliefs are then projected forward only to be further distorted by the overweighting of new information in the subsequent period. In other words, at an arbitrary point in time, forecasters exhibit both a non-zero ex-ante forecast error as well as a weighting error. Models of overconfidence can generate individual overreactions as well as aggregate underreactions, however, overconfidence is similarly unable to generate individual underreactions.

Strategic interaction models can also generate error predictability. For instance, strategic substitution can drive errors and revisions in opposite directions. This is because forecasters have a dual objective of not only minimizing their errors but also of distinguishing themselves from the average forecast. These models differ from the previous two in that strategic interaction models are rational. While this class of models can generate either overreaction or underreaction depending on the strategic motive assumed, it is unable to jointly deliver $\beta_1 > 0$ and $\beta_1 < 0$.

Models of noisy memory, first introduced in Azeredo da Silvera and Woodford (2019), can also generate overreactions. In a noisy memory model, forecasters do not have access to their full history of signals due to finite memory capacity. While models of rational inattention can explain individual underreactions, noisy memory may explain individual overreactions. Developing a hybrid rational inattention-noisy memory model could plausibly deliver simultaneous over- and underreactions.

Several other theories of expectations formation have been found to be inconsistent with the data. For instance, models of reputational concerns imply smoothing which can only generate underreaction. Moreover, asymmetric loss functions deliver counterfactually biased expectations, whereas the data show that professional forecasts are not unconditionally biased.

4 Stylized Model

To extract further insight as to how time-varying volatility can generate over- and underreactions, I consider next a stylized model of noisy information and unobserved volatility. I provide simulation results that describe the source of over- and underreactions. In the subsequent section I document cross-sectional evidence consistent with this mechanism.

4.1 Set Up

For simplicity, I suppose that forecasters can choose between two models: a Kalman filter (KF) and a particle filter (PF). Forecasters utilizing KF ignore the stochastic volatility and assess only the unconditional volatility of the state when formulating predictions. By ignoring time-varying volatility, forecasts based on the Kalman filter are suboptimal and generate error predictability among forecasters.

To reiterate, the Kalman gain is increasing in the signal-to-noise ratio, as illustrated in Figure 2. In a world with stochastic volatility, the optimal weight placed on new information varies over time. If the nature of the volatility were known in each period, then forecasters could update their predictions efficiently according to the weights traced out by the curve. If one were to ignore the time-varying volatility and filter with only the unconditional variance of the state, then he would update according to the constant Kalman gain akin to the horizontal line in the figure.

The underlying state in the stylized model is described as follows:

$$\begin{aligned} s_t &= \rho s_{t-1} + e^{h_t/2} w_t, \qquad && w_t \sim \mathcal{N}(0,1) \\ h_t &= \phi_0 + \phi_1 h_{t-1} + \sigma_\eta \eta_t, && \eta_t \sim \mathcal{N}(0,1). \end{aligned} \tag{7}$$

Furthermore, forecasters observe a contemporaneous private signal as well as a lagged public signal[18]

$$
\begin{aligned}
y_t^i &= s_t + \sigma_v v_t^i, \qquad & v_t^i &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1) \\
x_{t-1} &= s_{t-1} + \sigma_e e_{t-1}, \quad & e_{t-1} &\sim \mathcal{N}(0,1).
\end{aligned} \tag{8}
$$

In addition to being unable to observe s_t, forecasters are also unable to observe h_t. As such, the innovation w_t now enters into the state multiplicatively. Moreover, let $\mathbf{z}_t^i = \begin{pmatrix} y_t^i & x_{t-1} \end{pmatrix}'$ denote the vector of signals observed by forecasters.

Each period consists of two stages. In the first stage, a forecaster observes $\mathbf{z_t^i}$ and selects an approximation function. The use of these models comes at a random cost, $c^i \sim U(0, \overline{c})$ such that the PF cost distribution has a higher upper bound relative to KF. Then in the second stage, given the predictor function and the history of signals $\mathcal{Z}_t^i$, the forecaster reports a prediction of the public signal $\widehat{x}_{t|t}^i$ which is the macroeconomic variable in question.[19][20] For simplicity, I normalize $\overline{c}_{KF} = 0$. With this in mind, forecasters minimize the following loss function:

$$
\mathcal{L} = \min \left[MSE_{KF}^i, MSE_{PF}^i + c_{PF}^i \right], \quad c_{PF}^i \sim U(0, \overline{c}_{PF}).
$$

A forecaster will choose to make use of the more sophisticated PF if and only if

$$
MSE_{KF}^i - MSE_{PF}^i \geq c_{PF}^i. \tag{9}
$$

The lefthand side of the inequality reflects the benefit to adopting the PF, which manifests itself in a lower mean square error, whereas the righthand side denotes the relative cost to adopting the PF.

[18] One can alternatively envision that forecasters observe a macroeconomic variable with a transitory (e_t) and persistent (s_t) component. The persistent component is what is relevant for forecasting the target variable, though it is unobserved.

[19] The public signal, x, is the SPF variable to be forecasted as it is an observable. The latent state, s_t, is unobserved to the forecaster and the econometrician.

[20] At t, forecasters make use of their full history of signals in order to formulate a state estimate, $\widehat{\mathbf{s}}_{\mathbf{t|t}}^{\mathbf{i}}$. The first element of the forecasted state vector is $\widehat{s}_{t|t}^i$. Based on the assumption that $x_t = s_t + e_t$, it follows that $\widehat{x}_{t|t}^i = \widehat{s}_{t|t}^i$.

Table 3: Signal-to-Noise Ratio and Implied OLS Coefficients

Individual	Aggregate	SNR	β_1	α_1
Underreaction	Underreaction	1.40	0.12	0.46
Overreaction	Underreaction	0.40	-0.14	0.18

Note: The table simulates the errors-on-revision coefficient at the forecaster-level (β_1) and the consensus-level (α_1) for two different simulated signal-to-noise ratios (SNR).

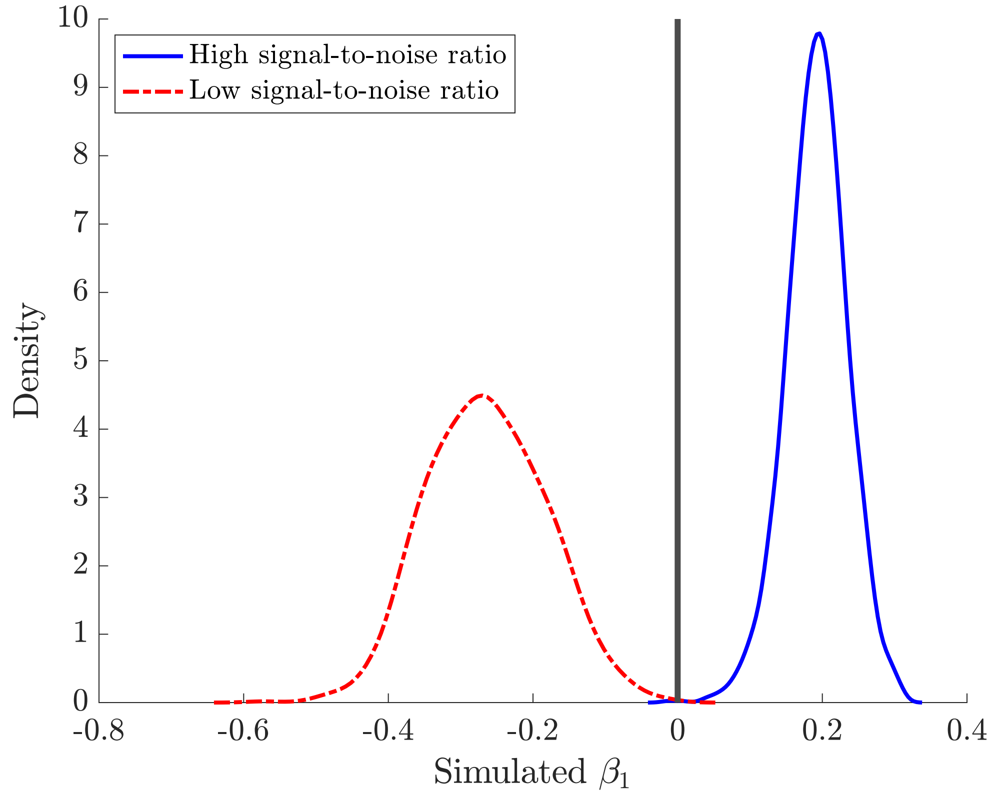
The adoption cost embodies unobserved heterogeneity in model adoption. As previously described, these reduced form costs can reflect heterogeneous time constraints among professional forecasters, different levels of training or experience in forecasting techniques, institution-specific frictions that make it more difficult to adopt a particular forecasting model, etc.

Fundamentally, this is a noisy information environment in which forecasters infer the state subject to private and public signals. Therefore, as in the simple model described in the previous section, the sign of the covariance between errors and revisions depends on the underlying signal-to-noise ratio. As the signal-to-noise ratio falls, forecast revisions are increasingly driven by the noise in the system. In this case, it is as if forecasters report their predictions with measurement error since an upward revision in the reported forecast will mechanically result in a more negative forecast error. On the other hand, when the signal-to-noise ratio is high, then fluctuations in the underlying state drive the forecast revisions. As a result, an upward revision delivers a more positive forecast error.

4.2 Simulation Results

The simulation results reported in Table 3 confirm that the model is able to qualitatively explain over- and underreactions across level of aggregation and variable. Differences across level of aggregation are seen through differences in simulated individual errors-on-revisions coefficient (β_1) and its consensus analog (α_1). Moreover, consider the difference in parameter values between models of individual underreactions and models of individual overreactions. As the signal-to-noise ratio

Figure 3: Overreactions, Underreactions, and Driving Process



Note: The figure plots two simulated densities of β_1 arising from a pooled individual-level errors-on-revisions regression from the stylized model. The red dashed line plots the simulation in which private signal noise variance is relatively high whereas the solid blue line plots the simulation in which private signal noise variance is relatively low.

(SNR) is reduced from 1.40 to 0.40, the sign of β_1 changes.[21] This result suggests that observed over- and underreactions at the individual level can be explained by different underlying data generating processes.

I also plot the simulated β_1 distribution across high signal-to-noise ratio and low signal-to-noise ratio parameterizations. For each of 2,000 simulations, I generate a panel of 250 forecasters over 200 periods. I then collect the errors and revisions for these forecasters and compute β_1. Figure 3 plots the density of β_1 across the simulations. The results confirm that the model can generate error predictability, and that over- and underreactions depend on the signal-to-noise ratio.[22] I next turn

[21] Although robust and reliable estimates of the SNR are currently sparse in this literature, CG provide some estimates using cross-country data which are in line with the simulated SNR values here. In addition, these simulated SNR values are similar to those that I quantify in Section 6.

[22] If all forecasters made use of the particle filter, however, then in large samples we would expect this distribution to be centered at zero.

to the SPF data to document facts consistent with this mechanism.

5 Evidence from the Survey of Professional Forecasters

This section exploits variation across the macroeconomic variables reported in the SPF in order to document evidence consistent with the idea that the signal-to-noise ratio is the key driver of over- and underreactions. The next section will parameterize the model in order to speak to simultaneous over- and underreactions within forecaster.

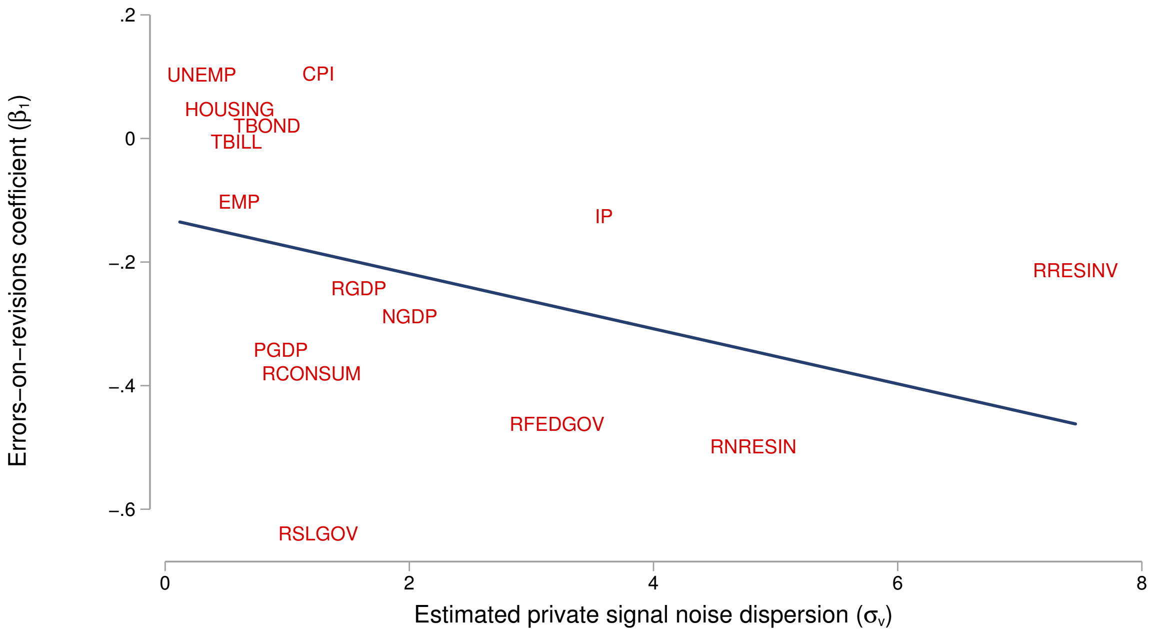
Each of the variables is presumed to follow a specific data generating process. As a result, $\beta_1 = \beta_1(\rho, \sigma_v, \sigma_e, \phi_0, \phi_1, \sigma_\eta)$ will, in general, vary in the cross-section of SPF variables. I document four facts by measuring proxies for signal and noise. I find that variables exhibiting greater unconditional volatility tend to be variables for which we observe underreactions. On the other hand, variables that feature elevated amounts of noise are associated with observed overreactions.

Testable Prediction 1: Error Predictability and Private Noise

In the stylized model, forecasters revise their predictions according to the realization of the lagged public signal as well as their contemporaneous private signal. The noise therefore feeds into the forecast revision. From the perspective of the model, the variance of private signal noise determines the amount of dispersion in revisions across forecasters. More dispersed signal noise admits more pronounced cross-sectional differences in revisions.

With this insight, I collect the pooled β_1 coefficients across SPF variables and plot these against the interquartile range of revisions across forecasters for each variable. Figure 4 displays the results. As the model suggests, variables exhibiting greater dispersion in revisions tend to be those for which forecasters overreact.

Figure 4: Error Predictability and Revision Dispersion

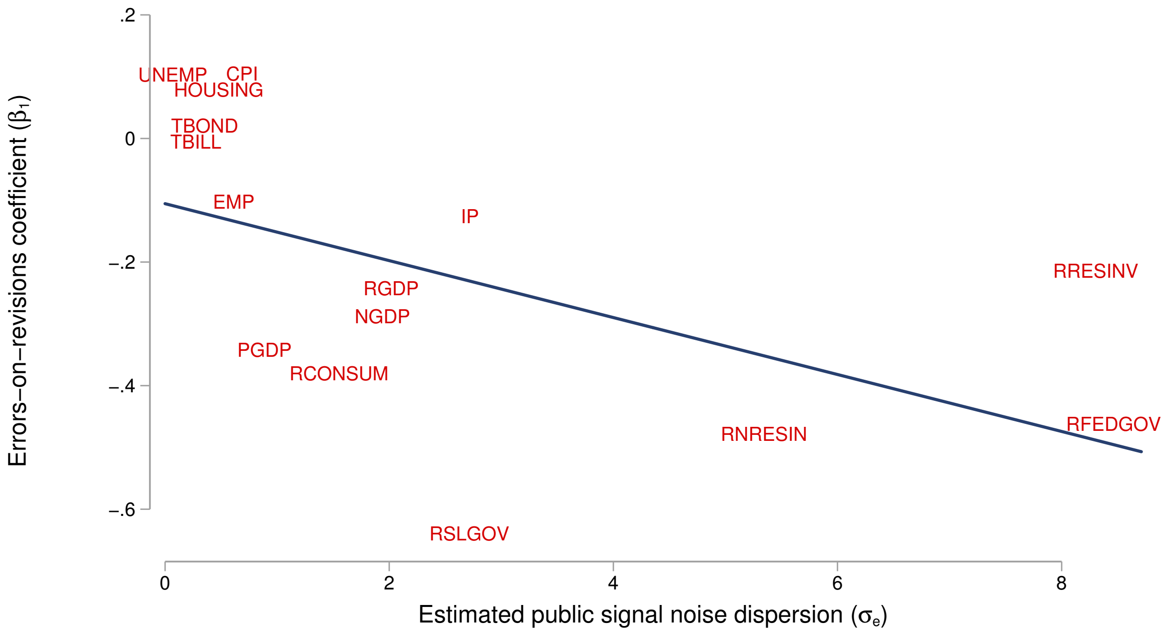


Note: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against estimated private noise dispersion, proxied by the interquartile range of forecast revisions. Slope of fitted line is -0.045.

Testable Implication 2: Error Predictability and Public Noise

While Figure 4 relates β_1 to private signal noise, there is also common noise present in the model. I next turn to measure the noisiness of the public signal. Whereas the SPF variable of interest has sometimes been modeled as the latent state in the literature, it is best thought of as a lagged public signal. This is because the SPF variables are observed by all forecasters with a lag. With this in mind, the official government revisions made to these variables across different vintages can provide a partial measure of public signal noise. Assuming that the vintages following the initial real-time release of the variable eliminate some of the common noise, one can quantify these revisions over time. As a matter of notation, define x_t^I as the real-time initial data release for a given variable, and x_t^L as the last release of the variable. Then, we can define $\text{noise}_t^{\text{public}} = \text{Var}(x_t^I - x_t^L)$. I construct this variable from the first and last data vintage for all SPF variables in my sample, and then measure the dispersion of this public noise over time. Figure 5 relates β_1 with this measure of

Figure 5: Error Predictability and Public Noise



Note: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against estimated public noise dispersion, proxied by the standard deviation of government revisions to real-time data. Slope of fitted line is -0.046.

public signal noise. The results are consistent with the intuition of the model: variables exhibiting higher measured noise dispersion tend to deliver observed overreactions.

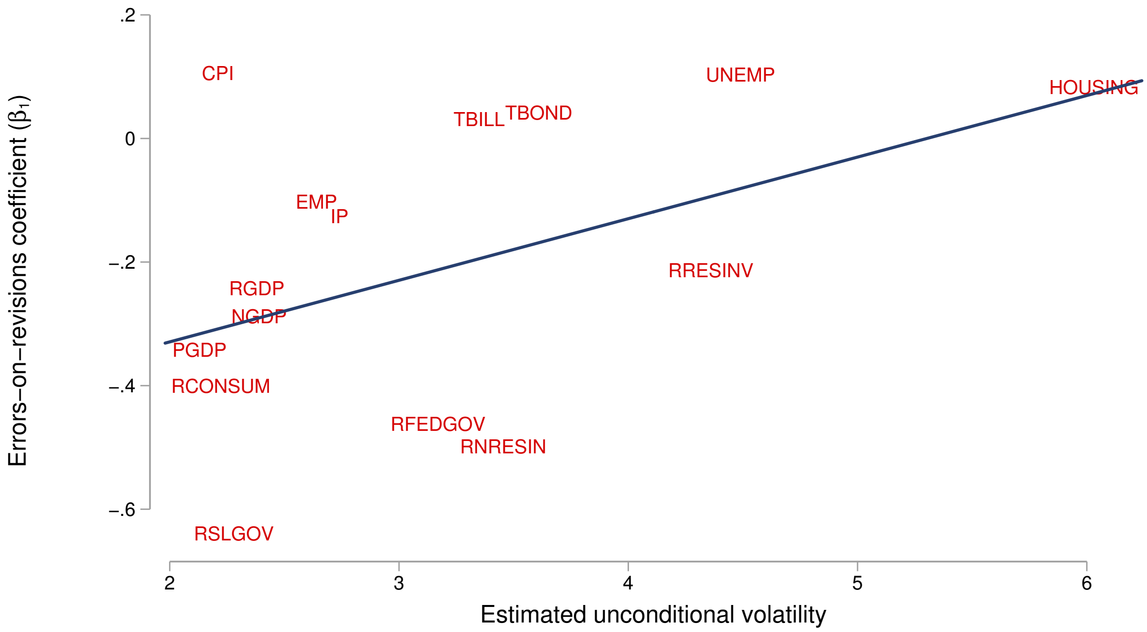
Testable Prediction 3: Error Predictability and Unconditional Volatility

Moreover, the model predicts that with more unconditional variability in the state, there is less scope for overreaction. To test this, I proceed to estimate $\{\phi_{j,0}, \phi_{j,1}\}$ for each SPF variable, j. I then construct an estimate of unconditional volatility:[23]

$$\text{vol}_j = \exp\left(\frac{\widehat{\phi_{j,0}}}{2(1-\widehat{\phi}_{j,1}^2)}\right).$$

[23] I estimate the parameters of the stochastic volatiliity model, $\{\phi_0, \phi_1, \sigma_\eta\}$ using MCMC techniques (Kastner and Frühwirth-Schnatter, 2014).

Figure 6: Error Predictability and Unconditional Volatility



Note: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against estimated unconditional volatility of the state, $\exp\left(\frac{\widehat{\phi_{j,0}}}{2(1-\widehat{\phi}_{j,1}^2)}\right)$. Slope of fitted line is 0.100.

Figure 6 relates β_1 to vol_j. The figure supports the hypothesis that variables exhibiting more variability in the state tend to provide greater scope for underreactions. Furthermore, note that the variance of the state is increasing in ρ. Hence, the model predicts that more persistent variables will reduce the scope for overreactions. This is consistent with Bordalo et al. (2020) who verify this empirically.

Testable Prediction 4: Error Predictability and Release Frequency

As an additional way to measure signal precision, I consider the frequency with which these different variables are made available to the public. While professional forecasters report predictions in each quarter, some variables are made available at higher frequencies. Specifically, the SPF conducts its survey at roughly the middle of each quarter. However, some of the SPF variables are released at a monthly frequency. For instance employment statistics are released on the Friday of each month. The survey asks forecasters to provide a quarterly average of these series. Furthermore, the financial time series are available at a daily frequency. As a result, forecasters have arguably more information pertaining to the eventual value of some variables in a given quarter than others. This reduces the effective noise in the lagged public signal.[24] Hence, variables available at higher frequencies should raise the scope for underreaction. Note that this does not preclude overreactive behavior in financial markets as has been readily documented. Here, I simply argue that quarterly (average of daily observations) predictions of a financial variable are better informed by the presence of daily observations through the middle of the quarter when the reported forecast is requested. On the other hand, the latest information that forecasters have for quarterly variables, such as GDP, is the previous quarter's release and an advance estimate. Since there is additional information available for some variables and not others, and the existence of this additional information depends on the variable frequency, then it follows that there is more scope for underreaction among variables that are available at higher frequencies. Figure 7 confirms this.

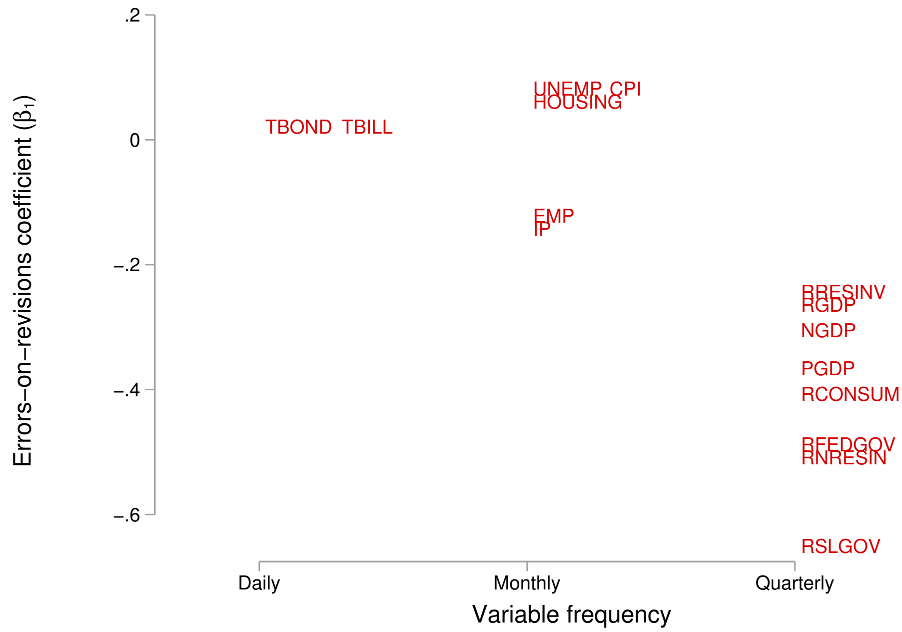
Jointly Testing for Overreaction and Underreaction Channels

As an additional check, I formally test for these channels jointly. For the data to accord with this theory of expectations, it should be the case that an interaction of the forecast revision with each of these variables either raises or reduces the extent to which β_1 is negative in the pooled specification (column 1 of Table 1). To complete this exercise, I incorporate two new regressors (and all possible interactions), each capturing a source of either noise or state volatility. As a measure of noise, I select the release frequency explained above. For my measure of fundamental volatility, I take a factor analysis approach. Since the latent state and its volatility are unobservable, it is natural to consider an index of the shared variation among all SPF variables. From this exercise, I obtain a time-varying index of what I call fundamental volatility.[25]

[24] Alternatively, one could suppose that forecasters receive an additional informative public signal for daily or monthly SPF variables.

[25] For this analysis, I drop nominal GDP since its components reside in my data set. Furthermore, I exclude CPI due to its shorter available history. For the remaining macroeconomic variables, I compute five-year rolling standard

Figure 7: Error Predictability and Release Frequency



Note: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against variable's release frequency {Daily, Monthly, Quarterly}.

Given these two new regressors, I modify the baseline errors-on-revisions regression (pooling across all SPF variables as in Table 1). Specifically, in addition to projecting errors on revisions, I specify a quarterly release indicator and the constructed fundamental volatility index. I also include interactions of each of these with the forecast revision as well as all interactions with each other. The regression results are reported in Table 4.

The first column of the table reproduces the first column of Table 1 for the relevant observations. The second column reports the fully specified regression. The relationship of interest remains the extent to which the forecast error and revision are related. The only interaction terms to enter statistically significantly are those crossed with the forecast revision. Furthermore, the signs of

deviations and then estimate underlying principal factors. The results deliver two factors that explain roughly equal amounts of the common variance of the final vintage of SPF variables. Based on the factor loadings, I call the first factor a real residential factor, and the second a real non-residential factor (the residential factor loads highly on housing and real residential investment whereas the second factor does not). While both factors deliver the correct sign in my regression specification, I report the regression that specifies the real non-residential factor as it delivers statistically detectable results. Appendix A confirms that the results are robust to window length.

Table 4: Modified Forecast Error Predictability Regressions

	Forecast Error	
	(1)	(2)
Revision	-0.314***	-0.165**
	(0.0414)	(0.073)
Revision × Quarterly		-0.194**
		(0.089)
Revision × Fundamental Volatility		0.106**
		(0.041)
Observations	58,740	58,740

Note: The table reports estimated coefficients of forecast error predictability across two specifications. The Quarterly indicator is equal to 1 if the SPF variable is released at a quarterly frequency and 0 otherwise. The Fundamental Volatility variable is a time-varying index constructed as described in the text. In addition to the interactions reported in the table, the column (2) specification includes the individual variables and their interactions as controls. Standard errors are as in Driscoll and Kraay (1998). Data used for estimation come from SPF (1964Q4-2018Q3). *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

these two interactions are consistent with the expected signs according to my theory of expectations under unobserved time-varying volatility. In particular, noise raises the scope for overreactions as evidenced by the negative cross term between the quarterly frequency indicator and the revision. On the other hand, fundamental volatility reduces the scope for overreaction as seen by the positive coefficients in front the volatility index and the revision.

The cross-sectional correlations and regression results reported above confirm that the model mechanism is consistent with the SPF. In the next section, I parameterize the model in order to show that it is also consistent with the extent to which a forecaster simultaneously over- and underreacts to different variables as in the data.

6 Parameterization

The model is able to generate over- and underreactions across levels of aggregation and across variables, pooling over forecasters. In this section, I parameterize the stylized model for real GDP and unemployment in order to demonstrate that it can also generate simultaneous over- and underreactions *within* forecaster.

Specifically, I calibrate the public noise dispersion, the persistence of the latent state, and the stochastic volatility parameters $\{\sigma_e, \rho, \phi_0, \phi_1, \sigma_\eta\}$. I then find the values of $\bar{c}_{PF}$ and σ_v that minimize the distance between the model-simulated and empirical estimates of the errors-on-revisions coefficients at the forecaster and consensus-levels (β_1 and α_1).

For real GDP and unemployment respectively, I set σ_e equal to the standard deviation of the data revisions made to each variable over the sample period. The data revision is taken to be the difference between the first and final release of the data series. For the remaining parameters, I consider the revised data rather than the real-time data. Intuitively, these series should be more highly correlated with the unobserved latent state. I then estimate an AR(1) on the revised series and set ρ equal to the estimated AR(1) coefficient. Finally, I collect the squared residuals from this autoregression and estimate $\{\phi_0, \phi_1, \sigma_\eta\}$. Panel A of Table 5 reports the calibration for each variable.

I then parameterize σ_v and $\bar{c}_{PF}$ by minimizing the distance between the model-implied $\{\beta_1, \alpha_1\}$ from its empirical counterpart. Since I am calibrating these parameters for GDP and unemployment, the procedure amounts to searching a four-dimensional parameter space and matching four moments. For each simulation, I generate two state variables according to the dynamics described in Section 4. I then simulate the lagged public signal as well as the contemporaneous private signal for each variable. In every period, forecasters report a forecast for each variable according to the state dynamics, signals received, and loss function described in Section 5. From this simulated panel of forecasters, I construct the errors-on-revisions coefficients. I minimize the distance between the simulated and empirical OLS coefficients by making use of simulated annealing, a standard global stochastic optimization routine.

The results, reported in Panel B of Table 5, indicate that real GDP is characterized by more private signal noise than the unemployment rate. Furthermore, as shown in Panel C, the implied signal-to-noise ratio for real GDP is about 0.50 whereas it is 1.36 for unemployment. This is consistent with the intuition of the model as well as the cross-sectional evidence in the previous section:

Table 5: Parameterization

Panel A: External			
Parameter	Description	Unemployment	Real GDP
ρ	State persistence	0.98	0.30
σ_e	Standard deviation of public noise	0.07	2.02
ϕ_0	Level of log variance	-0.76	0.14
ϕ_1	Persistence of log variance	0.73	0.92
σ_η	Volatility of log variance	0.69	0.39
Panel B: Internal			
Parameter	Description	Unemployment	Real GDP
σ_v	Standard deviation of private noise	0.12	2.81
$\bar{c}_{PF}$	PF cost upper bound	7.84	3.79
Panel C: Implied values			
	Description	Unemployment	Real GDP
	Signal-to-noise ratio	1.36	0.50
	Share using PF	0.53	0.82

Note: The table reports parameterization for unemployment and real GDP. Panel A reports the external parameterization. The stochastic volatility parameters $\{\phi_0, \phi_1, \sigma_\eta\}$ are estimated according to the algorithm presented in Kastner and Frühwirth-Schnatter (2014). Panel B reports internal parameterization obtain through the minimum distance procedure described in the text. Based on this calibration, Panel C reports the implied signal-to-noise ratio and share of forecasters that utilize the PF model.

variables that exhibit higher signal-to-noise ratios tend to be the variables for which underreactions are observed.

In addition, the cost distribution parameters indicate that costs to implementing the particle filter for real GDP are lower on average relative to the unemployment rate. The discrepancy between these two cost parameters can be attributed to the fact that mean square errors in the data are much larger in magnitude for real GDP. These costs govern in part the incentives to adopt the particle filter and, as reported in Panel C, imply that roughly 82% of forecasters optimally select to forecast with the particle filter for real GDP. On the other hand, 53% of forecasters choose the particle filter as the forecasting model of choice for unemployment.

Table 6 reports the model fit. The minimum distance procedure was able to successfully match patterns of over- and underreactions observed in the pooled forecaster-level and consensus regressions for real GDP. Though the fit for unemployment is not as tight, the model can fairly closely

Table 6: Model Fit

	Model	Data
Unemployment		
Errors-on-revisions, forecaster-level (β_1)	0.149	0.082
Errors-on-revisions, consensus (α_1)	0.212	0.247
Real GDP		
Errors-on-revisions, forecaster-level (β_1)	-0.263	-0.264
Errors-on-revisions, consensus (α_1)	0.352	0.350

Note: The table reports empirical and model-implied moments. The calibration directly targets the errors-on-revisions moments for unemployment and real GDP at the forecaster and consensus-levels. The final row reports the share of forecasters that overreact to real GDP and simultaneously underreact to unemployment.

match the consensus-level coefficient and can qualitatively match the forecaster-level coefficient.

Lastly, I assess the calibrated model's ability to match untargeted moments. Table 7 reports the mean square error, standard deviation of errors, and the share of over- and underreaction. Although the model generally overstates the mean square error, it successfully matches the relative magnitudes across both variables. Furthermore, the model is broadly successful in matching the dispersion of forecast errors. Lastly, the stylized model is able to successfully match the share of simultaneous over- and underreactions. Figure 1 reports that about 64% of forecasters in the sample overreact to real GDP while simultaneously underreacting to unemployment. Based on the parameterization devised here, the stylized model 63% of simulated forecasters overreact real GDP and underreact to unemployment.

6.1 Implications for Information Rigidities

What does time-varying volatility coupled with noisy information imply about interpreting the coefficient α_1 as an information rigidity? Based on my model, it is apparent that α_1 does not cleanly map to the Kalman gain as it does in the scalar linear context.[26] The key intuition of Bayesian filtering, however, still holds, and the optimal weight placed on innovation errors remains a sufficient

[26] See Appendix B.

Table 7: Untargeted Moments

	Unemployment	Real GDP
Mean square error		
Model	0.107	7.150
Data	0.045	6.005
Standard deviation of forecast error		
Model	0.320	2.670
Data	0.209	2.451
Share that overreact to real GDP and underreact to unemployment		
	0.632	0.641

Note: The table reports empirical and model-implied moments. The calibration directly targets the errors-on-revisions moments for unemployment and real GDP at the forecaster and consensus-levels. The final row reports the share of forecasters that overreact to real GDP and simultaneously underreact to unemployment.

statistic for capturing the rate of learning. This weight depends on the covariances of the state estimation error and the measurement error. Quantifying the rate of learning, however, is not readily feasible from a projection of mean errors on mean revisions as has traditionally been suggested in the literature (Coibion and Gorodnichenko, 2015; Dovern et al., 2015; Larsen et al., 2020).

In fact, from the perspective of the stylized model, the coefficient coming from errors on revisions regressions at the consensus-level may reveal misleading insights on the extent of information rigidity. Whereas a large α_1 coefficient would typically imply more information rigidities, here, α_1 is larger when the signal-to-noise ratio is high. On the other hand, a negative α_1 arises when signals are less informative. This suggests that the reduced form coefficient α_1 may be limited in what it reveals about genuine information frictions.

6.2 Implications for State Dependence

What does this mean for state dependence? If β_1 and α_1 were to rise in recessions, then the model would imply that the signal to noise ratio is countercyclical and information rigidities fall during economic downturns. If, on the other hand, these coefficients fall, then the signal-to-noise ratio is procyclical and information rigidities actually rise in recessions. CG document evidence indicating that α_1 falls in recessions. They interpret this as a reduction in information rigidities, however, my model would suggest that this implies a *rise* in information rigidities since it implies that the system experiences elevated amounts of noise. This is an important distinction between my model and the extant literature as it delivers an opposite answer to the question of whether individuals trust their signal more or less in recessions.

However, after performing a similar exercise to that in Table 4 by interacting a quarterly recession indicator with forecast revisions, I find no evidence that β_1 or α_1 changes with the business cycle. I also run this exercise by replacing the recession indicator with revised real GDP growth. It is possible, however, that the signal-to-noise ratio is insensitive to business cycle fluctuations because both the state and the signal experience stochastic volatility. I abstract away from volatility in signal precision, so there is a limit as to what one could glean from this model as it pertains to state dependence of information rigidities.

Nonetheless, one could distinguish between two types of uncertainty: fundamental uncertainty and information uncertainty. The first maps to time-varying volatility in the state while the latter arises when signal noise experiences stochastic volatility. There is a literature that stresses the importance of uncertainty shocks. These are often modeled as fundamental uncertainty shocks. Shocks to information precision are also studied in the literature (Dun Jia, 2016). According to my model, the state dependence of β_1 and α_1 depend on the signal-to-noise ratio which in turn depends on how these two types of uncertainty evolve relative to one another over the business cycle. If both rise in recessions, then is possible that the signal-to-noise ratio is acyclical thereby rendering β_1 and α_1 roughly constant over the cycle as well.

7 Conclusion

This paper documents that individual forecasters appear to simultaneous over- and underreact to new information. Existing models of belief formation are unable to flexibly accommodate these empirical patterns. This paper shows that a noisy information model incorporating unobserved time-varying volatility can make sense of these facts. Forecasters optimally select different models based on the complexity of the state dynamics. Heterogeneity in predictor functions can jointly deliver coincident over- and underreactions among forecasters. In particular, forecasters overreact to variables that exhibit more noise whereas they underreact to variables that are characterized by less noise. I uncover evidence in favor of this mechanism, demonstrating that fluctuations in volatility matter for belief dynamics.

References

Andrade, Philippe and Herve Le Bihan (2013), "Inattentive Professional 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.

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.

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

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

Bürgi, 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.

Crisan, Dan and Arnaud Doucet (2002), "A Survey of Convergence Results on Particle Filtering Methods for Practitioners." *IEEE Transactions on Signal Processing*, 50, 736–746.

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

Doucet, Arnaud and Adam M. Johansen (2009), "A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later." *Handbook of Nonlinear Filtering*.

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.

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.

Dun Jia, Calvin (2016), "Disagreement vs. Uncertainty: Investment Dynamics and Business Cycles." Working Paper.

Evans, George W. and Garey Ramey (1992), "Expectation Calculation and Macroeconomic Dynamics." *American Economic Review*, 82, 207–224.

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, Macroeocnomic Dynamics, and Asset Pricing." *NBER Macroeconomics Annual*, 26, 1–48.

Gabaix, Xavier (2018), "Behavioral Inattention." *Handbook of Behavioral Economics*.

Hu, Xiao-Li, Thomas B. Schon, and Lennart Ljung (2011), "A General Convergence Result for Particle Filtering." *IEEE Transactions on Signal Processing*, 59, 3424–3429.

Julier, Simon J. and Jeffrey K. Uhlmann (2004), "Unscented filtering and nonlinear estimation." *IEEE Transactions on Signal Processing*, 92, 401–422.

Justiniano, Alejandro and Giorgio E. Primiceri (2008), "The Time Varying Volatility of Macroeconomic Fluctuations." *American Economic Review*, 98, 604–641.

Kastner, Gregor and Sylvia Frühwirth-Schnatter (2014), "Ancillarity-Sufficiency Interweaving Strategy (ASIS) for Boosting MCMC Estimation of Stochastic Volatility Models of the Business Cycle." *Computational Statistics and Data Analysis*, 76, 408–423.

Kim, Chang-Jin and Charles Nelson (1999), “Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle.” *Review of Economics and Statistics*, 81, 608–616.

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

Lansing, Kevin J., Stephen F. LeRoy, and Jun Ma (2020), “Examining the Sources of Excess Return Predictability: Stochastic Volatility or Market Inefficiency?” Federal Reserve Bank of San Francisco Working Paper 2018-14.

Larsen, Vegard H., Leif Anders Thorsrud, and Julia Zhulanova (2020), “News-driven inflation expectations and information rigidities.” *Journal of Monetary Economics (Forthcoming)*.

Lewis, Karen (1989), “Changing beliefs and systematic rational forecast errors with evidence from foreign exchange.” *American Economic Review*, 99, 769–803.

McConnell, Margaret M. and Gabriel Perez-Quiros (2000), “Output Fluctuations in the United States: What Has Changed Since the Early 1980’s?” *American Economic Review*, 90, 1464–1476.

Stock, James and Mark Watson (2007), “Why has U.S. Inflation Become Harder to Forecast?” *Journal of Money, Credit and Banking*, 39, 4–33.

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

Veronesi, Pietro (2015), “Stock Market Overreactions to Bad News in Good Times: A Rational Expectations Equilibrium Model.” *The Review of Financial Studies*, 12, 975–1007.

Appendix A Empirics

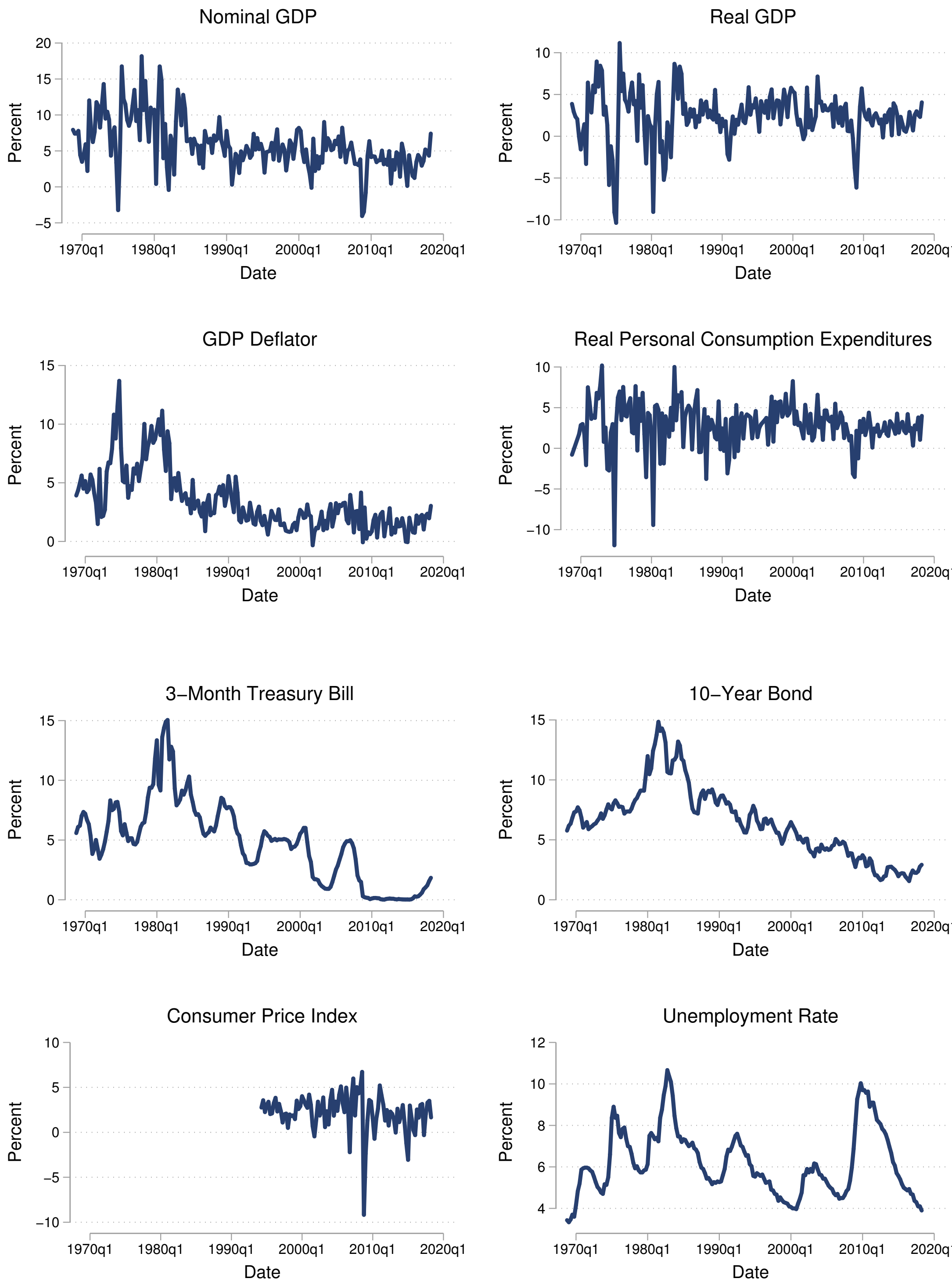
A.1 SPF: Variable Descriptions

While the paper focuses on inflation forecasts based on the GDP deflator, in this subsection I report additional results that make use of several other variables. Before presenting these results, I provide the variable descriptions below:

- NGDP–Quarterly nominal GDP growth forecast (seasonally adjusted, annual rate). Prior to 1992, these are forecasts for nominal GNP.
- RGDP–Quarterly real GDP growth forecast (seasonally adjusted, annual rate).
- PGDP–Quarterly GDP price index growth forecast (seasonally adjusted, annual rate). From 1992 - 1995, GDP implicit deflator is used, and prior to 1992, GNP implicit deflator.
- UNEMP–Forecasts for the quarterly average unemployment rate (seasonally adjusted, average of underlying monthly levels).
- EMP–Quarterly average growth of nonfarm payroll employment (seasonally adjusted, average of underlying monthly levels).
- RNRESIN–Quarterly growth forecast of real nonresidential fixed investment. Also known as business fixed investment (seasonally adjusted, annual rate).
- RRESINV–Quarterly growth forecast of real residential fixed investment (seasonally adjusted, annual rate).
- TBILL–Quarterly forecast of average three-month Treasury bill rate (percentage points, average of underlying daily levels).
- HOUSING– Quarterly growth forecast of average housing starts (seasonally adjusted, annual rate, average of underlying monthly levels).

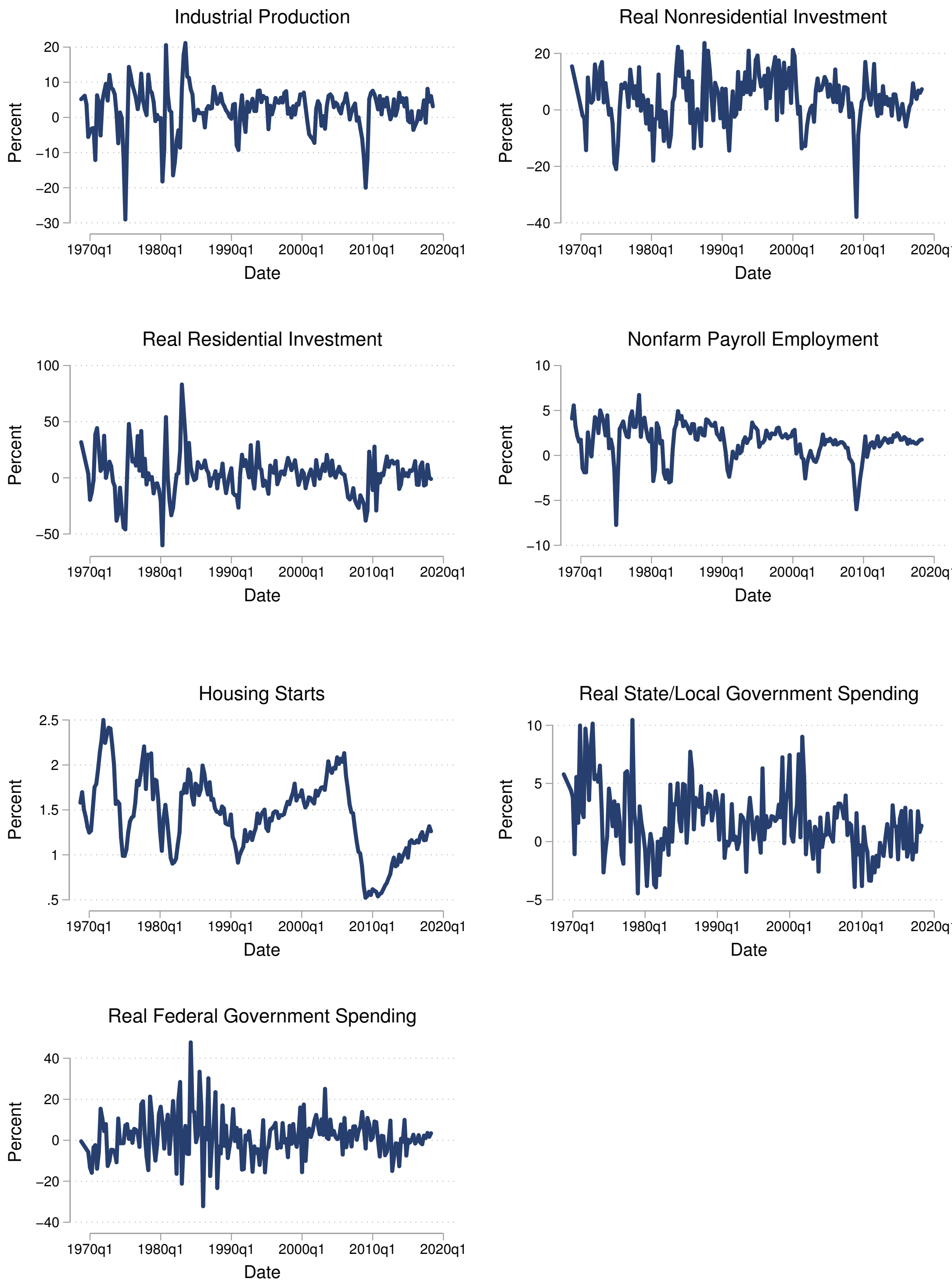
- CPI–Quarterly forecasts of the headline CPI inflation rate (percentage points, seasonally adjusted, annual rate). Quarterly forecasts are annualized q/q percent changes of quarterly average price index level (average of underlying monthly levels).
- RCONSUM – Quarterly growth forecast of real personal consumption expenditures (seasonally adjusted, annual rate).
- RFEDGOV –Quarterly growth forecast of real federal government consumption and gross investment (seasonally adjusted, annual rate).
- INDPROD – Quarterly forecasters of level of the index of industrial production, seasonally adjusted (quarterly forecasts are for quarterly average of underlying monthly levels).
- TBOND–Quarterly average 10-year Treasury bond rate (percentage points, average of the underlying daily levels). the underlying daily levels
- RSLGOV–Quarterly growth forecast of real state and local government consumption and gross investment (seasonally adjusted, annual rate).

Figure A1: Real-Time Macroeconomic Time Series



Source: Survey of Professional Forecasters

Figure A2: Real-Time Macroeconomic Time Series



Source: Survey of Professional Forecasters

A.2 Modified Error Predictability Regressions (Robustness)

This subsection reports the robustness results for the modified regressions in Section 5 of the main text. Table A1 reports the results by defining 7-year year windows for the rolling standard deviations. Table A2 reports the results from a 10-year rolling window specification.

Table A1: Modified Forecast Error Predictability Regressions (7 Year Window)

	Forecast Error	
	(1)	(2)
Revision	-0.308***	-0.136*
	(0.052)	(0.071)
Revision × Quarterly		-0.194**
		(0.090)
Revision × Fundamental Volatility		0.093**
		(0.040)
Observations	58,739	58,739

Note: The table reports estimated coefficients of forecast error predictability across two specifications. The variable *Quarterly* is equal to 0 if the SPF variable is released at a non-quarterly frequency and 1 otherwise. The *Fundamental Volatility* variable is a time-varying index of fundamental volatility constructed as described in the text. Column (1) reports a simple regression of errors-on-revisions while columns (2) includes the two proxies described. In addition to the variables reported in the table, column (2) includes the proxies individually as well as all of their interactions . Standard errors for forecaster-level regressions are as in Driscoll and Kraay (1998). Data used for estimation come from SPF (1964Q4-2018Q3). *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

Table A2: Modified Forecast Error Predictability Regressions (10-Year Window)

	Forecast Error	
	(1)	(2)
Revision	-0.278***	-0.143**
	(0.053)	(0.071)
Revision × Quarterly		-0.209**
		(0.091)
Revision × Fundamental Volatility		0.118**
		(0.036)
Observations	48,827	48,827

Note: The table reports estimated coefficients of forecast error predictability across two specifications. The variable *Quarterly* is equal to 0 if the SPF variable is released at a non-quarterly frequency and 1 otherwise. The *Fundamental Volatility* variable is a time-varying index of fundamental volatility constructed as described in the text. Column (1) reports a simple regression of errors-on-revisions while columns (2) includes the two proxies described. In addition to the variables reported in the table, column (2) includes the proxies individually as well as all of their interactions . Standard errors for forecaster-level regressions are as in Driscoll and Kraay (1998). Data used for estimation come from SPF (1964Q4-2018Q3). *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

Appendix B Model

B.1 General Linear Noisy Information RE Model

Theories of linear rational expectations are unable to account for over- and underreactions. Full information rational expectations counterfactually imply that errors are unpredictable. In addition, models of sticky information imply that forecast errors and forecast revisions are unrelated at the forecaster level.[27] In this sub-section, I focus on a general linear rational expectations model and provide analytical results about error predictability.

Consider a linear Gaussian state space model. Suppose there are n latent state variables and m

[27] If a forecaster updates, he does so with full information rational expectations so that the subsequent error is unrelated to the revision. On the other hand, if the forecaster does not update, then there is no revision.

exogenous signals.

$$\mathbf{s_t} = \mathbf{A s_{t-1}} + \mathbf{B w_t} \tag{10}$$
$$\mathbf{z_t^i} = \mathbf{C s_t} + \mathbf{D v_t^i}$$

Note that $\mathbf{s_t}$ is an $n \times 1$ vector, $\mathbf{A}$ is $n \times n$, $\mathbf{B}$ is $n \times n$ and $\mathbf{w_t}$ is $n \times 1$. Furthermore, $\mathbf{z_t}$ is $m \times 1$, $\mathbf{C}$ is $m \times n$, $\mathbf{D}$ is $m \times m$ and $\mathbf{v_t^i}$ is $m \times 1$. There are no other restrictions placed on the model. In particular, $\mathbf{s_t}$ can be a vector of many different state variables, or lags of itself. $\mathbf{B}$ need not be a diagonal matrix. Furthermore, $\mathbf{z_t^i}$ can include an arbitrary finite number of observed signals. The noise vector $\mathbf{v_t^i}$ can include private or public noise.[28]

From the Kalman filter, the optimal state estimate is defined as

$$\mathbf{s^i_{t|t}} = \mathbf{s^i_{t|t-1}} + \kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}}) \tag{11}$$

where κ is the (constant) Kalman gain. Since the state is unobservable, forecasters can only formulate predictions of the signals and assess the mistakes made with regard to these observables. The optimal forecast of the signal vector $\mathbf{z_t^i}$ is

$$\mathbf{z^i_{t+1|t}} = \mathbf{z^i_{t+1|t-1}} + \mathbf{CA}\kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}}) \tag{12}$$

Forecast errors for the generalized linear model can be expressed as follows

$$\mathbf{z^i_{t+1}} - \mathbf{z^i_{t+1|t}} = (\mathbf{z^i_{t+1}} - \mathbf{z^i_{t+1|t-1}}) - \mathbf{CA}\kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}}) \tag{13}$$

Furthermore, the forecast revision is

$$\mathbf{z^i_{t+1|t}} - \mathbf{z^i_{t+1|t-1}} = \mathbf{CA}\kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}}) \tag{14}$$

[28]I index this vector by i in general to allow for forecaster-specific signals.

Using these expressions, one can derive the two testable implications presented in the main text.

Proposition 1. *The generalized linear model implies the following:*

(i) $\beta_1 = 0$

(ii) $\alpha_1 = \mathbf{CA}(\mathbf{I} - \mathbf{C}\kappa)(\kappa\mathbf{C})^{-\mathbf{1}}(\mathbf{CA})^{-\mathbf{1}} > 0$

Proof. We have the following expressions for forecast errors and revisions, respectively:

$$\mathbf{FE^i} = \mathbf{CA}(\mathbf{I} - \kappa\mathbf{C})(\mathbf{s_t} - \mathbf{s^i_{t|t-1}}) + \mathbf{CBw_{t+1}} + \mathbf{Dv^i_{t+1}} - \mathbf{CA}\kappa\mathbf{Dv^i_t}$$

$$\mathbf{FR^i} = \mathbf{CA}\kappa\mathbf{Dv^i_t} + \mathbf{CA}\kappa\mathbf{C}(\mathbf{s_t} - \mathbf{s^i_{t|t-1}})$$

Then,

(a) $\beta_1 \propto \text{Cov}(FE^i, FR^i) = \mathbf{CA}(\mathbf{I} - \kappa\mathbf{C})\mathbf{\Psi}(\mathbf{CAKC})' - \mathbf{CAK}(\mathbf{Dv^i_t v^i_t D})(\mathbf{CA}\kappa)'$

where $\mathbf{\Psi}$ denotes the state estimation error variance. This becomes

$$\begin{aligned}
\beta_1 &\propto CA(I - \kappa C)\Psi(CA\kappa C)' - CAK(Dv_t^i v_t^i D)(CA\kappa)' \\
&= CA\Big\{(I - \kappa C)\Psi C - \kappa D v_t^i v_t^i D\Big\}(CA\kappa)' \\
\beta_1 &= 0
\end{aligned}$$

because the term in brackets is zero by the definition of the Kalman gain.

(b) Denoting $\overline{FE}$ and $\overline{FR}$ as the cross-sectional mean of the forecast error and revision, respectively, we have

$$\alpha_1 \propto \text{Cov}(\overline{\mathbf{FE}}, \overline{\mathbf{FR}}) = CA(I - \kappa C)\overline{\Psi}(CA\kappa C)'$$

The variance of the average revision is $\text{Var}(\overline{FR}) = CA\kappa C\Psi(CA\kappa C)'$ Thus, we have

$$\alpha_1 = CA(I - \kappa C)(CA\kappa C)^{-1}$$

$\square$

The proofs are straightforward: (a) holds given the orthogonality condition that must be satisfied at the individual-level under rational expectations. Forecast error orthogonality implies that $\mathbb{E}\big[(\mathbf{z_t^i} - \mathbf{z^i_{t|t}})\mu\big] = \mathbf{0}$ for any μ residing in the forecaster's information set.[29] Put another way, rationality implies the optimal use of information so that no variable residing in one's information set may predict the forecast error. This very general model precludes the predictability of forecast errors at the individual-level. As a result, any such linear Gaussian model with mean square loss cannot generate error predictability, regardless of the signal structure.

Moreover, (b) is a generalization of the CG result. The extent to which the mean revision predicts mean errors is determined by the Kalman gain matrix and the matrix $\mathbf{C}$ which maps the underlying state to the observed signal vector. The generalized linear model nests the CG result. Letting $\mathbf{C} = 1$, $\mathbf{D} = \sigma_v$, $\mathbf{A} = \rho$ and $\mathbf{B} = \sigma_w$, it follows that $\alpha_1 = \frac{1-\kappa}{\kappa}$. In this limiting case, one can recover an estimate of information rigidity by projecting consensus errors on consensus revisions. Importantly, the signal structure must be such that $\mathbf{C} = 1$. If, instead, the elements of $\mathbf{C}$ include additional parameters, or there is common noise in the signal vector, then it is no longer possible to cleanly extract an the Kalman gain from a standard OLS regression.[30]

As a result, a highly generalized linear rational expectations model is unable to explain the patterns in the data.

[29] Similarly, there is a revision orthogonality condition implied by rationality which states that $\mathbb{E}(\mathbf{z^i_{t|t}} - \mathbf{z^i_{t|t-1}})|\mathcal{I}_t^i) = 0$. See Pesaran and Weale (2006).

[30] See CG for a discussion of the bias in estimated information rigidities induced by public noise.

B.2 Error Predictability Under Time-Varying Volatility

From the general nonlinear model described in the main text, the covariance of errors and revision can be signed as follows:

$$
\begin{aligned}
\beta_1 \propto \mathbb{C}\Bigg(\bar{\mathbf{s}}_{\mathbf{t}}, \int \bar{\mathbf{s}}_{\mathbf{t}}[\widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})]\mathbf{d}\bar{\mathbf{s}}_{\mathbf{t}}\Bigg) \\
- \mathbb{C}\Bigg(\int \bar{\mathbf{s}}_{\mathbf{t}}\widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}})d\bar{\mathbf{s}}_{\mathbf{t}}, \int \bar{\mathbf{s}}_{\mathbf{t}}[\widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})]\mathbf{d}\bar{\mathbf{s}}_{\mathbf{t}}\Bigg)
\end{aligned}
$$

When there are no approximation errors, error orthogonality holds and $\beta_1 = 0$. In the case of non-zero approximation errors, however, the first term is the source of observed underreaction while the second term governs the extent of overreaction. When forecast revisions are more closely related to the underlying state, then underreactions arise as the first term dominates the second. If instead, forecast revisions covary more with the current prediction than the underlying state, then overreactions result. In essence, when the approximate revision incorporates more noise than is optimally called for, then forecasters will overreact.

Similar to the approximate prediction defined above, the consensus forecast arising from approximate predictions is defined as follows

$$
\begin{aligned}
\alpha_1 \propto \mathbb{C}\Bigg(\bar{\mathbf{s}}_{\mathbf{t}}, \int\int \bar{\mathbf{s}}_{\mathbf{t}}\Big[\widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})\Big]d\bar{\mathbf{s}}_{\mathbf{t}}di\Bigg) \\
- \mathbb{C}\Bigg(\int\int \bar{\mathbf{s}}_{\mathbf{t}}\widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}})d\bar{\mathbf{s}}_{\mathbf{t}}di, \int\int \bar{\mathbf{s}}_{\mathbf{t}}\Big[\widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \widehat{\mathbf{p}}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})\Big]d\bar{\mathbf{s}}_{\mathbf{t}}di\Bigg)
\end{aligned}
$$

More volatile revisions increase the scope for overreaction. Upon aggregating (symmetrically) across several individual forecasts, the consensus revision will exhibit more persistence than the individual revisions. This motivates the following result

Proposition 2. *In the nonlinear noisy information model, $\alpha_1 \geq \beta_1$.*

Proof. Recall that

$$\alpha_1 = [(\widehat{\mathbf{z}}_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1})'(\widehat{\mathbf{z}}_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1})]^{-1}(\widehat{\mathbf{z}}_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1})'(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})$$

and

$$\beta_1 = [(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})'(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})]^{-1}(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})'(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})$$

To prove the proposition, I will show that the covariance between consensus errors and revisions is weakly greater than that for pooled errors and revisions. I will then show that the variance of the consensus revision is weakly smaller than the variance of the pooled variance.

We can express the covariance between errors and revisions as

$$\begin{aligned}\mathbb{C}(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t}, \widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}) &= \int\int(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})didt \\ &- \int\int(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})didt - \int\int(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})didt\end{aligned}$$

and at the consensus level

$$\begin{aligned}\mathbb{C}(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}_{t+h|t}, \widehat{\mathbf{z}}_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1}) &= \int\left(\mathbf{z}_{t+h} - \int\widehat{\mathbf{z}}^i_{t+h|t}di\right)\left(\int[\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]di\right)dt \\ &- \int\left(\mathbf{z}_{t+h} - \int\widehat{\mathbf{z}}^i_{t+h|t}di\right)dt - \int\int(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1})didt\end{aligned}$$

We wish to show that the second equation is weakly greater than the first. One can note immediately that the second and third terms of both equations are equal (given the linearity of the expectations operator), and so they cancel out. The resulting inequality that we wish to verify is

$$\int\left(\mathbf{z}_{t+h} - \int\widehat{\mathbf{z}}^i_{t+h|t}di\right)\left(\int[\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]di\right)dt \geq \int\int(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})didt$$

By distributing the revision into the error on either side of the inequality, we can express each side as the sum of two terms. The first of these will drop out as we will have

$$\int \mathbf{z}_{t+h}\bigg(\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]di\bigg)dt$$

on the LHS and

$$\int\int \mathbf{z}_{t+h}(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})didt$$

on the RHS. Again, due to the linearity of the expectations operator, these terms cancel out. The remaining inequality is therefore

$$\begin{aligned}
-\int\bigg(\int \widehat{\mathbf{z}}^i_{t+h|t}di\bigg)\bigg(\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]di\bigg)dt &\geq -\int\int \widehat{\mathbf{z}}^i_{t+h|t}(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})didt \\
\int\bigg(\int \widehat{\mathbf{z}}^i_{t+h|t}di\bigg)\bigg(\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]di\bigg)dt &\leq \int\int \widehat{\mathbf{z}}^i_{t+h|t}(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})didt \\
\int\bigg(\int \widehat{\mathbf{z}}^i_{t+h|t}di\bigg)^2 dt - \int\bigg(\int \widehat{\mathbf{z}}^i_{t+h|t}di\bigg)\bigg(\int \widehat{\mathbf{z}}^i_{t+h|t-1}di\bigg)dt &\leq \int\int \widehat{\mathbf{z}}^{i2}_{t+h|t}didt - \int\int \widehat{\mathbf{z}}^i_{t+h|t}\widehat{\mathbf{z}}^i_{t+h|t-1}didt \\
\int\int \widehat{\mathbf{z}}^i_{t+h|t}\widehat{\mathbf{z}}^i_{t+h|t-1}didt - \int\bigg(\int \widehat{\mathbf{z}}^i_{t+h|t}di\bigg)\bigg(\int \widehat{\mathbf{z}}^i_{t+h|t-1}di\bigg)dt &\leq \int\int \widehat{\mathbf{z}}^{i2}_{t+h|t}didt - \int\bigg(\int \widehat{\mathbf{z}}^i_{t+h|t}di\bigg)^2 dt \\
\int\bigg[\int \widehat{\mathbf{z}}^i_{t+h|t}\widehat{\mathbf{z}}^i_{t+h|t-1}di - \bigg(\int \widehat{\mathbf{z}}^i_{t+h|t}di\bigg)\bigg(\int \widehat{\mathbf{z}}^i_{t+h|t-1}di\bigg)\bigg]dt &\leq \int\bigg[\int \widehat{\mathbf{z}}^{i2}_{t+h|t}di - \bigg(\int \widehat{\mathbf{z}}^i_{t+h|t}di\bigg)^2\bigg]dt
\end{aligned}$$

which is true since the terms in hard brackets on the RHS is the cross-sectional variance of the forecast whereas the term in hard brackets on the LHS is a cross-sectional covariance. Hence, the covariance of the consensus errors with consensus revisions is weakly greater than the covariance of individual-level pooled errors and revisions.

Finally, I show that the variance of the consensus revision is weakly smaller than the variance of the pooled revision. This is simpler to verify. Note that

$$\mathbb{V}(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}) = \int\int (\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})^2 didt - \bigg(\int\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]didt\bigg)^2$$

and

$$\mathbb{V}(\widehat{\mathbf{z}}_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1}) = \int \left(\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]di \right)^2 dt - \left(\int \left[\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]di \right] dt \right)^2$$

Once again, the second term in each of the above revision variance equations will cancel out. The resulting condition that we wish to verify is

$$\int \left(\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]di \right)^2 dt \leq \int \int (\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})^2 di dt$$

which holds by Jensen's inequality. □

The model implies that the OLS coefficient estimated from an errors on revisions regression will be weakly greater than the analogous coefficient obtained from a pooled regression of individual forecasters. This result does depend on the presence of nonlinear dynamics. In fact, this holds in the linear setting as well (see Appendix B).[31]

[31] This need not be the case in other economic settings in which agents take actions, and their decisions are aggregated in a manner other than by taking a simple mean. Depending on the context, it is possible for the aggregate decision to also exhibit overreaction or excess volatility. Bordalo et al. (2020) provide a discussion of this.

Appendix C Calibration

I internally calibrate four parameters: $\{\sigma_{v,1}, \sigma_{v,2}, \overline{c}_{PF,1}, \overline{c}_{PF,2}\}$ where the subscript one denotes the first variable (real GDP) and two denotes the second variable (unemployment). These parameters are calibrated to match four moments: $\{\beta_{1,1}, \beta_{1,2}, \alpha_{1,1}, \alpha_{1,2}\}$.

Based on the calibration, the two state variables must only be simulated once. Following this, I simulate a panel of forecasters who select KF or PF depending on the mean square errors and their model adoption cost draw. Following the endogenous model selection decision, I simulate a panel of errors and revisions from which I then compute model-implied errors-on-revisions coefficients. The simulated panel of forecasters is roughly 7 times the size of the size of the panel of SPF forecasters.[32].

I then collect the targeted empirical moments in a stacked vector $m(X)$ which comes from the SPF sample. I next stack the model-based moments, which depend on $\theta = (\beta_{1,1} \;\; \alpha_{1,1} \;\; \beta_{1,2} \;\; \alpha_{1,2})'$, in the vector $m(\theta)$. Finally I search the parameter space to find the $\widehat{\theta}$ that minimizes the following objective

$$\min_{\theta} \big(m(\theta) - m(X)\big)' W \big(m(\theta) - m(X)\big)$$

where the weighting matrix is set to be the identity matrix, $W = I$.

[32] I also discard the first 1,000 observations of the simulated state variables

Appendix D Details on Particle Filtering

In this section I briefly summarize the particle filter which is a popular nonlinear filter that have been devised to handle state dynamics such as unobserved stochastic volatility.

In their seminal paper, Gordon et. al. (1993) propose the bootstrap filter which is a popular variant to the particle filter. In principle, this approach makes use to mass points (particles) to approximate the underlying filtering density, $p(s_t|Z_t^i)$. This is done by defining the set of particles and associated weights: $\chi = \{s^{(n)}, \omega^{(n)}\}_{n=1}^N$.

Importantly, the filter still follows a general predict-update algorithm. For each particle n, the forecaster propagates the estimate through the nonlinear system

$$s_t^{i,(n)} = F\big(s_{t-1}^{i(n)}, w_t\big)$$

and then updates the weight,[33]

$$\widetilde{\omega}_t^{i,(n)} = \omega_{t-1}^{i,(n)} \cdot p\big(z_t^i | s_t^{i,(n)}\big)$$

The forecaster then normalizes the weights

$$\omega_t^{i,(n)} = \frac{\widetilde{w}_t^{i,(n)}}{\sum_{n=1}^N \widetilde{\omega}_t^{i,(n)}}$$

so that they sum to one. Lastly, the nowcast of the state is computed as a weighted average of the particles

$$\widehat{s}_{t|t}^i = \sum_{n=1}^N s_t^{i,(n)} \cdot \omega_t^{i,(n)}.$$

One common issue with sequential importance sampling is that the sample of particles tends to degenerate as few particles are given most of the weight. As a result, I make use of the common sequential importance resampling scheme in which I resample the particles, each with a probability equal to its weight.

[33] The precise manner in which the weights are updated depends on choices for the importance distribution.

Forecast errors and revisions are analogous to the formulation with the Kalman filter generalizations. The only difference is that the particle filtered estimates are not formulated by making use of the Kalman filtering equations. Nonetheless, these estimates approximate the optimal forecast.