Overreaction in Macroeconomic Expectations

Pedro Bordalo, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer[1]

March 2018

Abstract

We examine the rationality of individual and consensus professional forecasts of macroeconomic and financial variables using the methodology of Coibion and Gorodnichenko (2015), which focuses on the predictability of forecast errors from earlier forecast revisions. We document two principal findings: forecasters typically over-react to information individual level, while consensus forecasts exhibit under-reaction. To reconcile these findings, we combine the diagnostic expectations model of belief formation from Bordalo, Gennaioli, and Shleifer (2018) with Woodford's (2003) noisy information model of belief aggregation. The model accounts for the findings, but also yields a number of new implications related to the forward looking nature of diagnostic expectations, which we also test and confirm. Finally, we compare our model to mechanical extrapolation, rational inattention, and natural expectations.

[1] The authors are from Oxford Said Business School, Università Bocconi, Harvard University, and Harvard University, respectively. We thank Yuriy Gorodnichenko and participants at the 2018 AEA meeting for helpful comments, and we acknowledge the financial support of the Behavioral Finance and Finance Stability Initiative at Harvard Business School and the Pershing Square Venture Fund for Research on the Foundations of Human Behavior. Gennaioli thanks the European Research Council for Financial Support under the ERC Consolidator Grant. Johan Cassell, Spencer Kwon, Francesca Miserocchi, Johnny Tang, and Weijie Zhang provided outstanding research assistance.

I. Introduction

Since the advent of the Rational Expectations Hypothesis, the dominant approach in economics is to assume that market participants form their beliefs about the future, and make decisions, on the basis of statistically optimal forecasts. Recent research challenges this approach. Empirically, a growing body of work tests the Rational Expectations Hypothesis using survey data on the anticipations of households and professional forecasters. The evidence uniformly points to systematic departures from statistical optimality, which take the form of predictable forecast errors. Such departures have been documented, for example, in the context of forecasting inflation and other macro variables (Coibion and Gorodnichenko 2012, 2015, CG henceforth, Fuhrer 2017), the aggregate stock market (Bacchetta, Mertens, and Wincoop 2009, Amromin and Sharpe 2014, Greenwood and Shleifer 2014, Adam, Marcet, and Buetel 2017), the cross section of stock returns (e.g., La Porta 1996, Bordalo, Gennaioli, La Porta and Shleifer 2017, BGLS henceforth), credit spreads (Greenwood and Hanson 2013, Bordalo, Gennaioli, and Shleifer 2018), and corporate earnings (DeBondt and Thaler 1990, Ben-David et al. 2013, Gennaioli, Ma, and Shleifer 2015, Bouchaud, Kruger, Landier, and Thesmar 2017). Departures from optimal forecasts also obtain in controlled experiments (e.g., Hommes et al. 2004, Beshears et al. 2013, Frydman and Nave 2017, Landier, Ma, and Thesmar 2017).

On the theoretical side, various relaxations of the Rational Expectations Hypothesis have been proposed to account for the data. In macroeconomics, the main approach builds on rational inattention and information rigidities (Sims 2003, Woodford 2003, Carroll 2003, Mankiw and Reis 2005, Gabaix 2014). This view maintains the rationality of individual inferences, but relaxes the assumption of common information or full information processing. This is often justified by arguing that acquiring, absorbing, and processing information entails sizable material and cognitive costs. To economize on these costs, agents optimally revise their expectations only sporadically, or on the basis of selected news. As a consequence, expectations (and decisions) under-react to news relative to the world of unlimited information capacity. In an important empirical test of these theories, CG (2015) study predictability of errors in consensus macroeconomic forecasts of inflation and other variables, and find evidence consistent with under-reaction.

In finance, in contrast, although there is some evidence of momentum and under-reaction (Cutler, Poterba, and Summers 1990, Jegadeesh and Titman 1993), the dominant puzzle is over-reaction to news. This puzzle has been motivated by the evidence that stock prices move too much relative to the movements in fundamentals both in the aggregate (Shiller 1981) and in the cross section (De Bondt and Thaler 1985). The leading psychological mechanism for understanding over-reaction is Kahneman and Tversky's (1972) finding that, in reacting to news, people tend to overweight "representative" events (Barberis, Shleifer and Vishny 1998, Gennaioli and Shleifer 2010). For instance, exceptional past performance of a firm may cause overweighting of the probability that this firm is "the next google" because googles are representative of the group of well performing firms, even though they are rare in absolute terms. This approach is not inconsistent with limited information processing, but stresses that people infer too much from the information they attend to, however limited. Thus, beliefs and decisions move too much with news (Augenblick and Rabin 2017, Augenblick and Lazarus 2017). BGLS (2017) look at the cross section of stock returns and at analyst expectations about earnings growth and find support for over-reaction driven by representativeness.

This state of research motivates two questions. First, which departure from rational expectations is predominant, under- or over-reaction to news? At the least, can we identify circumstances in which either of them is more likely to prevail? Second, which mechanisms create these departures? Put differently, can one account for the main features in the data using a parsimonious model capturing precise cognitive mechanisms for under- and over-reaction?

This paper addresses these questions by studying the predictions of professional forecasters about 16 macroeconomic variables, which include but are not limited to those considered by CG (2015). We use both the Survey of Professional Forecasters (SPF) and the Blue Chip Survey, which gives us 20 expectations time series in total. These include forecasts of real economic activity, consumption, investment, unemployment, housing starts, and government expenditures, as well as multiple interest rate variables. We examine both consensus and individual level forecasts. SPF data are publicly available; Blue Chip data were purchased.

Section 3 addresses the first question above, namely what are the patterns of over- and under-reaction in different series. We follow CG's methodology of measuring a forecaster's news by his forecast revision, and of using this forecast revision to predict the forecast error, computed as the difference between the realization and the corresponding forecast. In this setting, under-reaction to news implies a positive correlation between forecast errors and forecast revisions, while over-reaction to news implies the opposite. Unlike CG, we examine not only consensus forecasts, defined as the average forecast across all analysts, but individual ones. We then explore the consequences of aggregating forecasts, which turns out to be crucial for understanding their properties.

For the case of consensus forecasts, our analysis confirms the CG findings of under-reaction: the average forecast revision positively predicts the average future forecast errors for most series. At the individual level, however, the opposite pattern emerges: for most series, the forecast revision of the average forecaster negatively predicts the same forecaster's future error. In stark contrast with the consensus results, at the level of the individual forecaster over-reaction is the norm, under-reaction the exception. These results are robust to a variety of potential sources of predictability, including forecaster heterogeneity, small sample bias, measurement error, non-standard loss functions, and non-normality of shocks.

In Section 4 we propose a model that reconciles these seemingly contradictory findings from the viewpoint of leading theories of under- and over-reaction to news. In our setup, agents must predict the future value of a state that follows an AR(1) process. Each agent observes a different noisy signal of the current value of this state. To exploit such noisy information optimally, forecasters should use the Kalman filter. This setup captures Woodford's (2003) "Noisy Information Model", which also describes CG's principal approach to rational inattention: noise stems from the cognitive costs of processing full information, but noisy signals are optimally evaluated using the Kalman filter. This setup can also capture a setting in which different forecasters, rather than being inattentive, simply observe different news (stemming for instance from their use of different models or different information sources, CG 2012).

To allow for over-reaction, we assume that – in processing the noisy signal – agents are swayed by representativeness. To formalize this heuristic we use the Gennaioli and Shleifer (2010) model, which was originally proposed to describe lab experiments on probabilistic judgments and later applied to social stereotypes (Bordalo, Coffman, Gennaioli, and Shleifer 2016), forecasts of credit spreads (BGS, 2018) and forecasts of firm performance (BGLS 2017). In this approach, the representativeness of a given future state is measured by the proportional increase in its probability in light of recent news. Agents exaggerate the probability of more representative states (states that have become relatively more likely) and underestimate the probability of others. Representativeness causes expectations to follow a modified Kalman filter that exaggerates the signal to noise ratio of news. As in earlier work, we call expectations distorted by representativeness "diagnostic."

In this model, under-reaction in the consensus can be reconciled with over-reaction at the individual level, *but only* when each forecaster over-reacts to the news he receives. When each forecaster over-reacts to his own information, the econometrician detects negative predictability of his forecast error at the individual level. At the consensus level, however, predictability may still be positive, provided the distortion caused by representativeness is not too strong. The reason is that, while over-reacting to his own signal, each individual forecaster does not react to the signals observed by the other forecasters. Because all signals are informative and on average correct about the state, the average forecast under-reacts to the average information.

Our analysis demonstrates that judging whether individuals under- or over-react to information on the basis of consensus forecasts may be misleading. Even if all forecasters over-react, as under diagnostic expectations, looking at consensus forecasts may point to under-reaction simply because different analysts over-react in different directions to partial information. In Section 5 we assess whether the data are consistent with further distinctive predictions of diagnostic expectations. These predictions allow us to distinguish the model from the mechanical updating rule of adaptive expectations. They also allow us to better compare our model to Rational Inattention. The general logic of these tests relies on the "kernel of truth" property of diagnostic expectations, which holds that belief updating exaggerates true patterns in the data. This property yields testable predictions both across different series and in the time series of individual variables.

We present cross sectional tests in Section 5.1. We show first that, upon receiving news, individuals' forecast revisions are stronger for variables whose time series exhibit more persistence. This is consistent with diagnostic expectations and with rational inattention, but not with adaptive expectations in which the updating rule is fixed. We then show that the individual-level CG coefficient of overreaction documented in Section 3 is closer to zero for series that are very persistent. This is in line with diagnostic expectations: as persistence increases, rational forecast revisions are more volatile (and in fact the signal to noise ratio increases) which reduces the scope for overreaction.

In Section 5.2 we develop a time-series test of the kernel of truth. We model individual series as AR(2) processes to account for long term reversals of actuals, consistent with the importance of hump shaped dynamics stressed by Fuster, Laibson, and Mendel (2010). We find that 12 out of 16 variables exhibit hump-shaped dynamics. We solve a diagnostic expectations model under AR(2) and show such dynamics have far-reaching implications for expectations under the kernel of truth property. In particular, they imply that: i) an upward forecast revision about the short term should predict excess pessimism about the long term, while ii) an upward forecast revision about the medium term should predict excess optimism about the long term. Put differently, diagnostic expectations exaggerate both short-term momentum and long-term reversals. We find that these predictions are borne out in the data. Besides strengthening the support for widespread over-reaction entailed by representativeness, these results also show the risks of using the CG method for AR(2) series. In fact, we show that overreaction to different time lags may contribute to finding apparent underreaction under an AR(1) specification. Taken together, the evidence is broadly consistent with the kernel of truth property of beliefs that is central to the diagnostic expectation mechanism.

In Section 6 we present a calibration exercise of our baseline model [to complete]. We also consider the realistic case where shocks are not normally distributed, where we replace the closed form Kalman filter with a numerical particle filter. We show that the model's predictions carry through.

The main contributions in this paper are to document empirically the prevalence of over-reaction to information in individual forecasts of macroeconomic variables, and to unify this finding with under-reaction in the consensus using diagnostic expectations. There have been several other approaches to similar phenomena. One of them is fixed-rule extrapolation or adaptive expectations; and we show throughout that the diagnostic expectations model both has better psychological foundations and yields predictions more consistent with the data. Another approach is the model of Natural Expectations of Fuster, Laibson, and Mendel (2010). Diagnostic expectations yield some patterns that are similar to natural expectations, but also make distinctive predictions – such as over-reaction to long-term reversals – that more closely reflect the data. Finally, diagnostic expectations are related to the idea of over-confidence. In particular, they imply an exaggeration of the perceived signal to noise ratio, which is a conventional formalization of overconfidence. We focus on diagnostic expectations rather than overconfidence, because – as we show in other work – they help explain beliefs even in settings where overconfidence can be ruled out (such as in cases where information is common and public). We return to these alternatives throughout our analysis.

2. The Data

Data on Forecasts

We collect forecast data from two sources: Survey of Professional Forecasters (SPF) and Blue Chip Financial Forecasts (Blue Chip).[2] SPF is a survey of professional forecasters currently run by the Federal Reserve Bank of Philadelphia. According to the enrollment form on Philadelphia Fed's website, "most of the survey's participants have formal and advanced training in economic theory and forecasting and use econometric models to generate their forecasts." Participation is also limited to "those who are currently generating forecasts for their employers or clients or those who have done so in the past." At a given point in time, around 40 forecasters contribute to the SPF anonymously. SPF is conducted on a quarterly basis, around the end of the second month in the quarter. It provides both consensus forecast data and forecaster-level data (identified by forecaster ID). Forecasters report forecasts for outcomes in the current and next four quarters, typically about the level of the variable in each quarter.

Blue Chip is a survey of panelists from around forty major financial institutions. The names of institutions and forecasters are disclosed. The survey is conducted around the beginning of each month. To match with the SPF timing most closely, we use Blue Chip forecasts from the end-of-quarter month survey

[2] Blue Chip provides two sets of forecast data: Blue Chip Economic Indicators (BCEI) and Blue Chip Financial Forecasts (BCFF). We do not use BCEI since historical forecaster-level data are only available for BCFF.

(i.e. March, June, September, and December). Blue Chip has consensus forecasts available electronically, and we digitize individual-level forecasts from PDF publications. Panelists forecast outcomes in the current and next four to five quarters. For variables such as GDP, they report (annualized) quarterly growth rates. For variables such as interest rates, they report the quarterly average level. For both SPF and Blue Chip, the median (mean) duration of a panelist contributing forecasts is about 16 (23) quarters.

Given the timing of the SPF and Blue Chip forecasts we use, by the time the forecasts are made in quarter t (i.e. around the end of the second month in quarter t), forecasters know the actual values of variables with quarterly releases (e.g. GDP) up to quarter $t-1$, and the actual values of variables with monthly releases (e.g. unemployment rate) up to the previous month.

Table 1 presents the list of variables we study, as well as the time range for which forecast data are available from SPF and/or Blue Chip. These variables cover both macroeconomic outcomes, such as GDP, price indices, consumption, investment, unemployment, government consumption, and financial variables, primarily yields on government bonds and corporate bonds. SPF covers most of the macro variables and selected interest rates (three month Treasuries, ten year Treasuries, and AAA corporate bonds). Blue Chip includes real GDP and a larger set of interest rates (Fed Funds, three month, five year, and ten year Treasuries, AAA as well as BAA corporate bonds). Relative to CG (2015), we add two SPF variables (nominal GDP and the 10Y Treasury rate) as well as the Blue Chip forecasts.[3]

Table 1. List of Variables

This table lists our outcome variables, the forecast source, and the period for which forecasts are available.

Variable	SPF	Blue Chip	Abbreviation
Nominal GDP	1968Q4--2014Q4	N/A	NGDP
Real GDP	1968Q4--2014Q4	1999Q1--2014Q4	RGDP
GDP Price Deflator	1968Q4--2014Q4	N/A	PGDP
Real Consumption	1981Q3--2014Q4	N/A	RCONSUM
Real Non-Residential Investment	1981Q3--2014Q4	N/A	RNRESIN
Real Residential Investment	1981Q3--2014Q4	N/A	RRESIN
Federal Government Consumption	1981Q3--2014Q4	N/A	RGF
State & Local Government Consumption	1981Q3--2014Q4	N/A	RGSL
Housing Starts	1968Q4--2014Q4	N/A	HOUSING
Unemployment Rate	1968Q4--2014Q4	N/A	UNEMP

[3] Relative to CG, we do not use SPF forecasts on CPI inflation and industrial production index, as real time macro data are missing for these two variables for a period of time.

Fed Funds Rate	N/A	1983Q1--2014Q4	FF
3M Treasury Rate	1981Q3--2014Q4	1983Q1--2014Q4	TB3M
5Y Treasury Rate	N/A	1988Q1--2014Q4	TN5Y
10Y Treasury Rate	1992Q1--2014Q4	1993Q1--2014Q4	TN10Y
AAA Bond Rate	1981Q3--2014Q4	1984Q1--2014Q4	AAA
BAA Bond Rate	N/A	2000Q1--2014Q4	BAA

The main forecast horizon we analyze is annual. For variables like GDP and inflation, we look at the annual growth rate from quarter $t-1$ to quarter $t+3$. In SPF, the forecasts for these variables are in levels (e.g. level of GDP), so we transform them into implied growth rates (actual GDP of quarter $t-1$ is known at the time of the forecast, so this transformation complies with the forecasters' information sets). In Blue Chip, the forecasts for these variables are in the form of quarterly growth rates, so we add up forecasts for growth rates in quarters t to $t+3$. For variables such as the unemployment rate and interest rates, we look at the level in quarter $t+3$. Both SPF and Blue Chip have direct forecasts of the quarterly average level in quarter $t+3$. Appendix A provides a description of variable construction.

Consensus forecasts are computed as means from individual-level forecasts available at a point in time. We calculate forecasts, forecast errors, and forecast revisions at the individual level, and then average them across forecasters to compute the consensus.[4]

Data on Actual Outcomes

The actual outcomes of macroeconomic variables are released quarterly but are often subsequently revised. To match as closely as possible the forecasters' information set, we focus on initial releases from Philadelphia Fed's Real-Time Data Set for Macroeconomists. For a given quarter, we proxy the forecasters' information set as the latest estimates available by the time of the forecast. Conversely, we measure the actual outcome that was forecasted using the initial release of the actuals in the corresponding time period. For example, for actual GDP growth from quarter $t-1$ to quarter $t+3$, we use the initial release of GDP_{t+3} (available in quarter $t+4$) divided by the initial release of GDP_{t-1} (available in quarter t, prior to

[4] There could be small differences in the set of forecasters who issue a forecast in quarter t, and the set of forecasters who revise their forecast at t (these forecasters need to be present at $t-1$ as well). Thus, simple averages of forecasts and forecast revisions may cover different sets of individuals. This issue does not affect the results much. We can restrict our calculation to forecasters that have both forecasts and forecast revisions and results are the same.

when the forecasts are made). For financial variables, the actual outcomes are available daily and are permanent (not revised). We use historical data from the Federal Reserve Bank of St. Louis.

Summary Statistics

Table 2 below presents the summary statistics of the variables, including the mean and standard deviation for the actuals being forecasted, as well as the consensus forecasts, forecast errors, and forecast revisions at a horizon of quarter t+3. The table also shows statistics for the quarterly share of forecasters with no meaningful revisions,[5] and the quarterly share of forecasters with positive revisions.

Table 2. Summary Statistics

Summary statistics of main variables; means and standard deviations are presented. All values are in percentages. Panel A shows the statistics of actuals, consensus forecasts, consensus errors and consensus revisions. Actuals are realized outcomes corresponding to the forecasts, and errors are actuals minus forecasts. Revisions are forecasts of the outcome made in quarter *t* minus forecasts of the same outcome made in quarter *t-1*. Panel B shows additional individual level statistics. The forecast dispersion column shows the mean of quarterly standard deviations of individual level forecasts. The revision dispersion column shows the mean of quarterly standard deviations of individual level forecast revisions. Non-revisions are instances where forecasts are available in both quarter *t* and quarter *t-1* and the change in the value is less than 0.01. The non-revision and up-revision columns show the mean of quarterly non-revision shares and up-revision shares. The final column of Panel B shows the fraction of quarters where less than 80% of the forecasters revise in the same direction.

Panel A. Consensus Statistics

		Actuals		Forecasts		Errors		Revisions	
Variable	Format	mean	sd	mean	sd	mean	sd	mean	sd
Nominal GDP (SPF)	Growth rate from end of quarter t-1 to end of quarter t+3	6.19	2.90	6.43	2.30	-0.24	1.75	-0.14	0.71
Real GDP (SPF)		2.56	2.31	2.73	1.38	-0.17	1.74	-0.18	0.64
Real GDP (BC)		2.66	1.55	2.62	0.86	0.03	1.30	-0.12	0.48
GDP Price Index (SPF)		3.56	2.49	3.63	2.03	-0.07	1.14	0.02	0.48
Real Consumption (SPF)		2.85	1.46	2.53	0.76	0.32	1.15	-0.05	0.51
Real Non-Residential Investment (SPF)		4.90	7.35	4.41	3.68	0.49	5.86	-0.26	1.78
Real Residential Investment (SPF)		2.77	11.68	2.67	6.19	0.11	8.71	-0.64	2.48
Real Federal Government Consumption (SPF)		1.36	4.59	1.34	2.61	0.02	3.22	0.13	1.24
Real State&Local Govt Consumption (SPF)		1.62	1.68	1.62	1.09	0.00	1.12	0.00	0.59
Housing Start (SPF)		1.67	22.16	4.75	15.33	-3.08	18.81	-2.41	5.97
Unemployment (SPF)	Average level in quarter t+3	6.38	1.55	6.38	1.43	0.00	0.76	0.06	0.33
Fed Funds Rate (BC)		4.10	2.99	4.53	2.94	-0.42	1.04	-0.18	0.54
3M Treasury Rate (SPF)		3.98	2.86	4.54	2.93	-0.56	1.15	-0.21	0.52
3M Treasury Rate (BC)		3.76	2.73	4.28	2.72	-0.52	1.02	-0.18	0.51

[5] We categorize a forecaster as making no revision if the forecaster provides non-missing forecasts in both quarters *t*-1 and *t*, and the forecasts change by less than 0.01. For variables in rates, the data is often rounded to the first decimal point, and this rounding may lead to a higher incidence of none-revision. For national accounts variables in SPF, which are provided in levels, we define no-revision as less than 0.01% change in the implied growth rate forecasts.

5Y Treasury Rate (BC)	4.45	2.24	4.86	2.05	-0.41	0.89	-0.15	0.45
10Y Treasury Rate (SPF)	4.49	1.56	4.99	1.40	-0.50	0.76	-0.12	0.37
10Y Treasury Rate (BC)	4.42	1.56	4.86	1.38	-0.44	0.75	-0.13	0.39
AAA Corporate Bond Rate (SPF)	7.26	2.4	7.74	2.52	-0.47	0.85	-0.11	0.39
AAA Corporate Bond Rate (BC)	6.84	1.94	7.26	2.01	-0.42	0.7	-0.12	0.37
BAA Corporate Bond Rate (BC)	6.30	1.08	6.75	0.95	-0.45	0.68	-0.14	0.31

Panel B. Additional Individual Level Statistics

		Forecasts	Revisions			
Variable	Format	Dispersion	Dispersion	non-rev share	up-rev share	Pr(<80% revise same direction)
Nominal GDP (SPF)	Growth rate from end of quarter *t*-1 to end of quarter t+3	0.59	1.13	0.02	0.45	0.79
Real GDP (SPF)		0.63	0.94	0.02	0.43	0.74
Real GDP (BC)		0.17	0.40	0.05	0.43	0.66
GDP Price Index (SPF)		0.52	0.75	0.05	0.49	0.79
Real Consumption (SPF)		0.68	0.76	0.03	0.48	0.76
Real Non-Residential Investment (SPF)		1.03	2.47	0.02	0.49	0.71
Real Residential Investment (SPF)		2.09	4.24	0.03	0.45	0.83
Real Federal Government Consumption (SPF)		1.38	2.25	0.06	0.52	0.87
Real State&Local Govt Consumption (SPF)		1.45	1.28	0.10	0.48	0.93
Housing Start (SPF)		5.46	8.61	0.00	0.39	0.68
Unemployment (SPF)	Average level in quarter *t*+3	0.13	0.30	0.18	0.42	0.77
Fed Funds Rate (BC)		0.33	0.48	0.22	0.30	0.68
3M Treasury Rate (SPF)		0.29	0.48	0.15	0.34	0.68
3M Treasury Rate (BC)		0.29	0.46	0.19	0.32	0.63
5Y Treasury Rate (BC)		0.15	0.42	0.12	0.35	0.61
10Y Treasury Rate (SPF)		0.09	0.38	0.10	0.35	0.65
10Y Treasury Rate (BC)		0.08	0.35	0.13	0.33	0.57
AAA Corporate Bond Rate (SPF)		0.25	0.51	0.09	0.38	0.73
AAA Corporate Bond Rate (BC)		0.22	0.47	0.12	0.34	0.71
BAA Corporate Bond Rate (BC)		0.12	0.41	0.13	0.32	0.81

Several patterns emerge from Table 2. First, the average forecast error is about zero. It does not appear that macro analysts have asymmetric loss functions that persistently bias their forecasts in a given direction. Second, for each variable, there is significant dispersion of forecasts and revisions at each point in time, as shown in Table 2 Panel B. Third, analysts frequently revise their forecasts, but they do so in different directions. For example, as shown by the final column of Panel B, it is uncommon to have quarters where more than 80% forecasters revise in the same direction. This suggests that the appropriate model is one in which different forecasters observe or attend to different news, either because they are exposed to different information or because they use different models, or both. Berger, Erhmann, and Fratzscher

(2011) show, for example, that the geographical location of forecasters influences their ability to predict monetary policy decisions. Different forecasters may have personal contacts with the industry, policymakers, etc., which offers one explanation for the disagreement we see in the data. In this sense, forecasting is in part psychological: it involves subjective weighting of model output with private signals.[6]

3. Over-reaction vs. Under-reaction: Basic Tests

Studies of the rational expectations hypothesis often test whether forecast errors can be predicted using information available at the time the forecast was made. Understanding whether departures from rational expectations are due to over- or under-reaction to information is more challenging, since the forecaster's full information set cannot be directly observed by the econometrician.

To confront this problem, CG (2015) measure the news observed by a forecaster by his forecast revision. Denote by $x^i_{t+h|t}$ the h-periods ahead forecast made at time t by forecaster i about the value x_{t+h} of a certain variable. Denote by $x^i_{t+h|t-1}$ his forecast in the previous period. The h-periods ahead forecast revision at t is given by $FR^i_{t,h} = \left(x^i_{t+h|t} - x^i_{t+h|t-1}\right)$, or the one period change in the forecast about x_{t+h}. This revision captures the information that the forecaster has observed and used to update his forecast.

CG analyze consensus forecasts, defined as the average of individual forecasters' predictions $x_{t+h|t} = \frac{1}{I}\sum_i x^i_{t+h|t}$, where $I > 1$ is the number of forecasters. In this setting, the h-periods ahead "consensus information" or forecast revision is given by the change in the consensus forecast, $FR_{t,h} = \left(x_{t+h|t} - x_{t+h|t-1}\right)$. The extent to which the consensus forecast under-reacts or over-reacts to information can then be assessed by estimating the regression:

$$x_{t+h} - x_{t+h|t} = \beta_0 + \beta_1 FR_{t,h} + \epsilon_{t,t+h}. \tag{1}$$

Under the Rational Expectations Hypothesis, the forecast error should be unpredictable using any current information, including the forecast revision itself, so $\beta_1 = 0$. When instead forecasters under-react

[6] This is well illustrated by a quote from Cleveland Fed President Pianalto, as cited by Coibion and Gorodnichenko (2012): "To paraphrase one of my colleagues, we are looking at flawed data through the lens of imperfect models. To try to clarify my perspective on the economy, I also spend a lot of time talking with businesspeople."

to information, we expect $\beta_1 > 0$. As shown by CG, this includes the case in which analysts, perhaps because they are inattentive, observe different noisy signals of x_{t+h} and update rationally based on those signals (see Section 4). At the same time, $\beta_1 > 0$ is also consistent with non-rational under-reaction, such as that arising under Adaptive Expectations. To see why $\beta_1 > 0$ captures under-reaction, suppose that positive information is received, leading to a positive forecast revision $FR_{t,h} > 0$. If the forecast under-reacts, the upward revision is insufficient, predicting a positive forecast error $\mathbb{E}_t\left(x_{t+h} - x_{t+h|t}\right) > 0$. The converse holds if negative information is received: the downward revision is insufficient, predicting a negative error. This is why, when forecasters under-react, forecast errors are positively correlated with forecast revisions.

By the same logic, when forecasters over-react to information we should expect $\beta_1 < 0$. Indeed, over-reaction means that after positive information $FR_{t,h} > 0$ forecasters are too optimistic, so the forecast error is negative $\mathbb{E}_t\left(x_{t+h} - x_{t+h|t}\right) < 0$. On the other hand, after negative information $FR_{t,h} < 0$ they are too pessimistic, so the error is positive $\mathbb{E}_t\left(x_{t+h} - x_{t+h|t}\right) > 0$. That is, over-reaction implies that the forecast error should be negatively correlated with the forecast revision.

To test for Rational Inattention, CG's baseline estimate of Equation (1) uses *consensus* SPF forecasts for the GDP price deflator (PGDP_SPF) at a horizon $h = 3$. This yields $\beta_1 = 1.2$, which is robust to a number of controls. They also run Equation (1) for 13 SPF variables by pooling forecast horizons from $h = 0$ to $h = 3$,[7] and find qualitatively similar results, with 8 out of 13 variables exhibiting significantly positive β_1's, and the average coefficient being close to 0.7. The general message is that consensus forecasts of macroeconomic variables display under-reaction.

We estimate Equation (1) for our 20 series for the same baseline horizon $h = 3$, using consensus forecasts. The results are reported in columns (1) through (3) of Table 3, and confirm the findings of CG. The estimated β_1 is positive for 14 out of 20 series, statistically significant for 8 of them at the 5% confidence level, and for a further two series at the 10% level. Our point estimate for inflation forecasts is exactly in line with CG's. While results for the other SPF series are not directly comparable (since CG

[7] These results are presented in Figure 1 Panel B of CG (2015).

pool across forecast horizons), the estimates lie in a similar range. The one exception is RGF_SPF (federal government spending) for which the estimated β_1 is negative and significant at the 5% level. Results from the Blue Chip survey align well with SPF where they overlap, but do not exhibit significant consensus overreaction for the remaining (exclusively financial variables) series.

We stress that the various forecast series are not independent. For instance, nominal and real GDP growth are naturally highly correlated; the different interest rate series are also closely connected. Nonetheless, the general message holds: for macro variables and short rates, under-reaction is common in the consensus forecast regressions, while such patterns are largely absent in long-term rates.

As mentioned above, insufficient updating of consensus beliefs may be due to aggregation issues, rather than to under-reaction to information by individual forecasters. As we saw in Table 2, individual forecasters often revise in different directions, perhaps because they look at different data or use different models. In this case, even if individual forecasters over-react to their own information, such over-reaction is attenuated by averaging individual revisions going in opposite directions.

Table 3. Error-on-Revision Regression Results

This table shows coefficients from the CG (error on revisions) regression (1). Coefficients are displayed for both consensus time-series regressions, and forecaster-level pooled panel regressions, together with standard errors and p-values. Standard errors are Newey-West for consensus time-series regressions, and clustered by both forecaster and time for individual level regressions.

	Consensus			Individual					
				No fixed effects			With fixed effects		
	β_1	s.e.	p-val	β_1^p	s.e.	p-val	β_1^p	s.e.	p-val
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nominal GDP (SPF)	0.48	0.22	0.03	-0.26	0.07	0.00	-0.30	0.06	0.00
Real GDP (SPF)	0.45	0.25	0.07	-0.23	0.08	0.00	-0.21	0.06	0.00
Real GDP (BC)	0.59	0.34	0.09	0.12	0.19	0.26	-0.02	0.17	0.93
GDP Price Index Inflation (SPF)	1.21	0.21	0.00	-0.07	0.10	0.46	-0.16	0.07	0.03
Real Consumption (SPF)	0.18	0.22	0.41	-0.34	0.11	0.00	-0.39	0.10	0.00
Real Non-Residential Investment (SPF)	0.93	0.38	0.02	0.01	0.13	0.93	-0.03	0.12	0.82
Real Residential Investment (SPF)	1.26	0.38	0.00	-0.02	0.10	0.82	-0.12	0.08	0.14
Real Federal Government Consumption (SPF)	-0.44	0.23	0.05	-0.62	0.07	0.00	-0.63	0.06	0.00
Real State&Local Govt Consumption (SPF)	-0.16	0.20	0.42	-0.71	0.14	0.00	-0.73	0.13	0.00
Housing Start (SPF)	0.45	0.31	0.14	-0.25	0.09	0.01	-0.28	0.08	0.00
Unemployment (SPF)	0.82	0.21	0.00	0.33	0.11	0.00	0.26	0.11	0.02
Fed Funds Rate (BC)	0.61	0.23	0.01	0.15	0.09	0.11	0.12	0.09	0.19
3M Treasury Rate (SPF)	0.71	0.26	0.01	0.24	0.09	0.01	0.19	0.09	0.04
3M Treasury Rate (BC)	0.67	0.25	0.01	0.20	0.09	0.02	0.16	0.08	0.06

5Y Treasury Rate (BC)	0.05	0.22	0.84	-0.12	0.10	0.23	-0.19	0.10	0.05
10Y Treasury Rate (SPF)	-0.01	0.28	0.97	-0.18	0.10	0.06	-0.23	0.09	0.01
10Y Treasury Rate (BC)	-0.06	0.25	0.81	-0.17	0.12	0.14	-0.25	0.11	0.02
AAA Corporate Bond Rate (SPF)	-0.01	0.24	0.97	-0.21	0.08	0.00	-0.26	0.07	0.00
AAA Corporate Bond Rate (BC)	0.21	0.21	0.31	-0.17	0.07	0.00	-0.22	0.06	0.00
BAA Corporate Bond Rate (BC)	-0.14	0.28	0.62	-0.28	0.10	0.00	-0.34	0.10	0.00

To assess whether individual forecasters over- or under-react to their own information, we continue to follow the CG methodology, but perform the analysis at the individual analyst level. Here $FR^i_{t,h} = \left(x^i_{t+h|t} - x^i_{t+h|t-1}\right)$ is the analyst-level revision, and the h-periods ahead individual forecast error is $x_{t+h} - x^i_{t+h|t}$. For each variable, we then pool all analysts and estimate the regression:

$$x_{t+h} - x^i_{t+h|t} = \beta^p_0 + \beta^p_1 FR^i_{t,h} + \epsilon^i_{t,t+h}. \tag{2}$$

Superscript p in the coefficients recognizes that these are estimated by pooling individual level data. The logic, however, does not change: $\beta^p_1 > 0$ indicates that the average analyst insufficiently adjusts his forecast on the basis of his own information, while $\beta^p_1 < 0$ indicates that the average analyst over-reacts.[8]

Columns (4) through (6) of Table 3 report the results of estimating Equation (2). Surprisingly, the picture is essentially reversed from the consensus analysis: at the individual level, the average analyst appears to over-react to information, as measured by a negative β^p_1 coefficient. The estimated β^p_1 is negative for 14 out of the 20 series, and significantly negative for 9 series at the 5% confidence level, and for one other series at the 10% level. Except for short rates (Fed Funds and 3-months T-bill rate), all financial variables display over-reaction, consistent with Shiller's evidence of excess volatility. But many macro variables also display over-reaction, including nominal GDP, real GDP (in SPF, not in Blue Chip), real consumption, real federal government expenditures, real state and local government expenditures. GDP price deflator inflation, real GDP in Blue Chip, and non-residential investment display neither over-

[8] The individual level coefficient β^p_1 can in principle be different from the consensus coefficient β_1: to the extent that some information is forecaster specific, and that individuals do not react to information they do not possess, errors $\epsilon^i_{t,t+h}$ may be correlated across individuals over time. In Section 4 we formalize this intuition.

nor under-reaction (β_1^p is close to zero). Only the 3-months T-bill rate and unemployment rate display individual level under-reaction with positive and statistically significant β_1^p.

In columns (7) to (9), we also analyze regressions with forecaster-level dummies to account for possible time-invariant differences among analysts. For example, some analysts may be consistently overly-optimistic or overly-pessimistic, perhaps due to differences in their prior beliefs. These tendencies could contribute to positive correlations between forecast errors and revisions. Specifically, the overly optimistic analysts systematically receive bad news, leading to negative revisions and negative forecast errors. Similarly, the overly pessimistic analysts systematically receive good news, leading to positive revisions and positive forecast errors. In the data, the results with and without forecaster fixed effects are similar. With forecaster fixed effects, the estimated β_1^p is negative for 17 series, and significantly negative for 13 series at the 5% confidence level. Overall, the broad message from Table 3 is clear: at the level of the individual forecaster, over-reaction is the norm.

Taken together, a seemingly contradictory picture emerges from these CG tests. At the consensus level, expectations typically under-react to information. At the individual level, in contrast, they typically over-react. We conclude this section with a number of robustness checks for these results. In Section 4, we present a model capable of reconciling these patterns.

3.1 Robustness Checks

Predictability of forecast errors might in principle arise from features of the data unrelated to individuals' under- or overreaction to information. In this section we show our results are robust to a variety of such confounds, including data limitations (forecaster heterogeneity, small sample bias, measurement error) as well as biases in reported forecasts (asymmetric loss functions, forecast smoothing). In Section 6, we use our model to make a simulation-based assessment of robustness with respect to non-normal shocks.

Heterogeneity among Forecasters. Coibion and Gorodnichenko (2015) point out that heterogeneity across forecasters, either in updating (e.g. heterogeneous signal to noise ratios) or in beliefs about long term means, may affect the predictability of forecast errors. They show that, under the appropriate econometric specifications, the predictability of errors from revisions of consensus forecasts continues to hold. To assess whether heterogeneity drives our results for individual level forecasts, we perform forecaster level regressions, focusing on forecasters with at least 10 observations. Table A1 in Appendix B compares the median coefficient from forecaster level regressions to the coefficients from pooled individual level regressions from Table 3. These coefficients are very similar, suggesting that the observed over-reaction represents the median forecaster. On average across series, we estimate a negative β_1^p for two thirds of forecasters. In some series, nearly every forecaster overreacts while in other series the distribution of β_1^ps is more balanced. It would be interesting to account for these patterns, but we do not try to do so here.

Small Samples. We also check whether the finite sample for individual forecasters can bias our results. Finite-sample Stambaugh bias exists in time series regressions (Kendall 1954, Stambaugh 1999) and panel regressions with fixed effects (Nickell, 1981), but not in pooled panel regressions as they have a common intercept (Hjalmarsson 2008). The baseline individual-level regressions in Table 3 do not have fixed effects; adding fixed effects does not change the results much, indicating the bias is not severe. Moreover, the finite sample biases are stronger when the predictor variables are persistent. The predictor variable in the CG regressions, forecast revision, has low persistence in the data (about zero for most variables at the individual level, and less than 0.5 at the consensus level). Finally, the simulation analysis of Section 6 suggests that, for the relevant parameters in our data, the coefficients are not biased.

Measurement Error. Forecasts measured with noise can mechanically lead to negative predictability of forecast errors in Equation (2): a positive shock increases the measured forecast revision and decreases the forecast error. In our case, since professional forecasters directly report their forecasts, it is hard to think of literal “measurement error.” Moreover, motivated by the fact that some series display an AR(2) structure, in Section 5 we regress the forecast error at $t + h$ on revisions of forecasts for *previous* periods $t + h - 1$ and $t + h - 2$ (Equation 13). In line with the predictions of the model (Proposition 3), but not with measurement error, we find strong predictability in these regressions as well (Table 6).

Asymmetric Loss Functions. An important issue for expectations data is that respondents might optimally report biased forecasts. One source of such bias is captured by asymmetric loss functions. Specifically, under-predicting may be more costly than over-predicting, or vice versa. In these cases, the average forecast would be biased upward (if under-predicting is more costly) or downward (if over-predicting is more costly). Such a bias might generate the predictability we document if combined with time varying volatility (Pesaran and Weale, 2006), but it would in any case generate an average forecast error. However, in the data we do not find that forecasts are systematically upward or downward biased on average. At the consensus level, the consensus forecast errors are both very small and insignificant (Table 2, panel A). At the individual level, we fail to reject that the average forecast error is different from zero for about 60% of forecasters for the macroeconomic variables.[9]

Another source of bias in reported expectations is that individuals may follow consensus forecasts (Morris and Shin 2002, Fuhrer 2017). To assess its implications, let $\tilde{x}^i_{t+h|t} = \alpha x^i_{t+h|t} + (1-\alpha)\tilde{x}_{t+h|t}$, where $x^i_{t+h|t}$ is the individual optimal forecast and $\tilde{x}_{t+h|t}$ is the average contemporaneous forecast under this bias (which coincides with the unbiased average). Our benchmark model has $\alpha = 1$ but for $\alpha < 1$ forecasters put weight on others' signals at the expense of their own. In this model, and according to intuition, following consensus forecasts leads to underreaction to own signals, contrary to our findings.[10]

A related source of biased forecasts is forecast smoothing for reputational reasons. In response to news at time t, forecasters may wish to minimize the path revisions in their forecasts about period $t+h$, taking into account the previous forecast $x^i_{t+h|t-1}$ as well as the future path of forecasts $x^i_{t+h|t+j}$. We assess the relevance of this mechanism in the data in two ways. First, it is easy to show that forecast smoothing implies a positive autocorrelation in forecast revisions; in contrast, this autocorrelation is close to zero for the vast majority of analyst-series pairs. Second, when revising forecasts for the current quarter

[9] No clear pattern emerges among the forecasters whose average error is significantly different from zero, consistently with the fact that such errors average out in the population for most series. For interest rates, average forecast errors tend to be negative, but this reflects not an asymmetric loss function but rather a secular decline over the time period we examine. In other cases, individual average errors may be due to relatively short samples.

[10] Formally, denote $\widetilde{FE}^i_{t+h,t} = x_{t+h} - \tilde{x}^i_{t+h|t}$ the forecast error and $\widetilde{FR}^i_{t+h,t} = \tilde{x}^i_{t+h|t} - \tilde{x}^i_{t+h|t-1}$ the forecast revision. It follows that $\widetilde{FE}^i_{t+h,t} = \alpha FE^i_{t+h,t} + (1-\alpha)FE_{t+h|t}$ and similarly $\widetilde{FR}^i_{t+h,t} = \alpha FR^i_{t+h,t} + (1-\alpha)FR_{t+h|t}$. Then $cov(\widetilde{FE}^i_{t+h,t}, \widetilde{FR}^i_{t+h,t}) > 0$ follows from $cov(FE^i_{t+h,t}, FR^i_{t+h,t}) = 0$ and $cov(FE_{t+h|t}, FR_{t+h|t}) > 0$ under noisy rational expectations, together with $cov(FE^i_{t+h,t}, FR_{t+h|t}), cov(FE_{t+h|t}, FR^i_{t+h,t}) > 0$.

($h = 0$), forecast smoothing reduces to minimizing the current revision alone. This leads to positive predictability of errors; in contrast, we find negative predictability even at this horizon (Table A2).

4. Diagnostic Expectations

We present a model that reconciles underreaction of consensus expectations with overreaction of individual level expectations. At each time t, forecasters forecast a certain variable x_{t+h}, whose current value x_t is not directly observed. What is observed instead is a noisy signal s_t^i:

$$s_t^i = x_t + \epsilon_t^i, \tag{3}$$

where ϵ_t^i is analyst specific noise, which is i.i.d. normally distributed across forecasters and over time, with mean zero and variance σ_ϵ^2. The hidden state x_t evolves according to an AR(1) process with persistence ρ:

$$x_t = \rho x_{t-1} + u_t, \tag{4}$$

where u_t is a normal shock with mean zero and variance σ_u^2. We consider the case in which fundamentals follow an AR(1) process (here and in Section 5) for two reasons. First, AR(1) is a benchmark that was also considered by CG (2015), so it allows us to compare our model to theirs. Second, AR(1) yields a closed form characterization of the model's predictions. In particular, it allows us to study how expectations depend on the persistence parameter ρ. In Section 6 we allow series to follow an AR(2) process and show that this has additional implications for our analysis.

This setup accommodates several interpretations. In CG (2015), the fact that x_t is unobservable stems from rational inattention (Sims 2003, Woodford 2003). Forecasters could in principle perfectly observe the true current value of x_t, say GDP, but doing so is too costly. As a consequence, they observe a noisy proxy for it. This version of rational inattention is called "Noisy Rational Expectations", to reflect the fact that individuals rationally update on the basis of noisy signals. It is a different formulation of rational inattention than, for example, sticky information models (Mankiw and Reis 2002), in which all forecasters observe the same information but only sporadically revise their predictions. As shown by CG, these two versions of rational inattention have similar predictions on the relationship between consensus forecast errors and consensus forecast revisions. For this reason, our model only considers Noisy Rational Expectations. When we discuss predictions, however, we refer more broadly to Rational Inattention.

Another interpretation of Equations (3) and (4) is that the current realization of a variable, say GDP, is influenced by a persistent component x_t and a transitory component. In predicting the future, forecasters must estimate the persistent component on the basis of the noisy signal s_t^i. In this interpretation, the forecaster specific shock ϵ_t^i captures the fact that different forecasters extract information using different models or pieces of evidence, and the variance σ_ϵ^2 of this shock captures the difficulty of the information extraction problem (which is shaped by the availability of reliable models and/or evidence). For professional forecasters, the latter interpretation is perhaps more compelling, since their job is to look at, and predict, the variables in question, so they are very attentive to these variables. The problem they face, though, when looking at, say, GDP statistics, is to assess whether shocks are transitory or persistent. Under both interpretations, the predictions of the model depart from full information rational expectations because forecasters form their revisions through rational updating on the basis of noisy signals. In this sense, it is not misleading to place both interpretations under the rubric of Noisy Rational Expectations.

A Bayesian, or rational, forecaster enters period t carrying from the previous period beliefs about the current persistent state x_t summarized by a probability density $f(x_t|S_{t-1}^i)$, where S_{t-1}^i denotes the full history of signals observed by this forecaster. In period t, the forecaster observes a new signal s_t^i. In light of this evidence, he updates his estimate of the current state using Bayes' rule:

$$f(x_t|S_t^i) = \frac{f(s_t^i|x_t)f(x_t|S_{t-1}^i)}{\int f(s_t^i|x)f(x|S_{t-1}^i)dx}. \tag{5}$$

Equation (5) iteratively defines the forecaster's beliefs. In the current setting with normal shocks, the distribution $f(x_t|S_t^i)$ is described by the Kalman filter. A rational forecaster should then estimate the current state to be $x_{t|t}^i = \int xf(x|S_t^i)dx$ and should forecast the economic series of interest by using the AR(1) structure of the state x_t, namely $x_{t+h|t}^i = \rho^h x_{t|t}^i$.

We allow beliefs to be distorted by Kahneman and Tverky's representativeness heuristic, as in our model of Diagnostic Expectations. In line with BGLS (2017), which applied Diagnostic Expectations to a (diagnostic) Kalman Filter, we define the representativeness of a state x_t at time t as the likelihood ratio:

$$R_t(x_t) = \frac{f\left(x_t|S_t^i\right)}{f\left(x_t|S_{t-1}^i \cup \left\{x_{t|t-1}^i\right\}\right)}. \tag{6}$$

State x_t is more representative at t if the signal s_t^i received in this period increases the probability of that state, relative to not receiving any news. Receiving no news means observing a signal equal to the ex-ante forecast, $s_t^i = x_{t|t-1}^i$, as described in the denominator of equation (6).

Intuitively, the most representative states are those whose likelihood has increased the most in light of recent data. The forecaster then overweighs representative states by using the distorted posterior:

$$f^{\theta}\left(x_t|S_t^i\right) = f\left(x_t|S_t^i\right)R_t(x_t)^{\theta}\frac{1}{Z_t}, \tag{7}$$

where Z_t is a normalization factor ensuring that $f^{\theta}\left(x_t|S_t^i\right)$ integrates to one. Parameter $\theta \geq 0$ denotes the extent to which beliefs are distorted by representativeness. For $\theta = 0$ beliefs are rational, described by the Bayesian conditional distribution $f\left(x_t|S_t^i\right)$. For $\theta > 0$ the diagnostic density $f^{\theta}\left(x_t|S_t^i\right)$ inflates the probability of highly representative states and deflates the probability of unrepresentative states. Mistakes occur because states that have become relatively more likely may still be unlikely in absolute terms.

This formalization of representativeness as relative likelihood, and its distortive effect on probability assessments, has been shown to unify well-known biases in probability assessments such as base rate neglect, the conjunction fallacy, and the disjunction fallacy (Gennaioli and Shleifer 2010). It has also been used to explain phenomena such as stereotyping (BCGS 2016), self-confidence (BCGS 2017), and expectation formation in financial markets (BGS 2018, BGLS 2017).

Equation (7) yields a very intuitive formalization of beliefs.

Proposition 1 *The distorted density $f^{\theta}\left(x_t|S_t^i\right)$ is normal. In the steady state it is characterized by a constant variance Σ and by a time varying mean $x_{t|t}^{i,\theta}$ which are given by:*

$$x_{t|t}^{i,\theta} = x_{t|t-1}^{i} + (1+\theta)\frac{\Sigma}{\Sigma + \sigma_{\epsilon}^{2}}\left(s_{t}^{i} - x_{t|t-1}^{i}\right), \tag{8}$$

$$\Sigma = \frac{-(1-\rho^{2})\sigma_{\epsilon}^{2} + \sigma_{u}^{2} + \sqrt{[(1-\rho^{2})\sigma_{\epsilon}^{2} - \sigma_{u}^{2}]^{2} + 4\sigma_{\epsilon}^{2}\sigma_{u}^{2}}}{2}. \tag{9}$$

In equations (8) and (9), $x_{t|t-1}^{i}$ refers to the rational forecast of the hidden state implied by the Kalman Filter. Diagnostic beliefs resemble rational beliefs in several respects. They have the same variance Σ, and their mean $x_{t|t}^{i,\theta}$ updates past rational beliefs $x_{t|t-1}^{i}$ with "rational news" $s_{t}^{i} - x_{t|t-1}^{i}$, to an extent that increases in the signal to noise ratio $\Sigma/\sigma_{\epsilon}^{2}$.

The difference between diagnostic and rational expectations is that the former overweigh the impact of news by the multiplicative factor θ in Equation (8). That is, representativeness induces forecasters to behave as if news is more informative than it actually is, exaggerating updating. As a consequence, the Diagnostic Kalman Filter generates over-reaction to information. This stands in contrast to models of information rigidity or inattention, which generate under-reaction to information.[11] Equation (8) highlights another important feature of diagnostic expectations: they display excess volatility. In particular, the discrepancy between rational and diagnostic expectations arises only in the presence of rational news, namely when $\left(s_{t}^{i} - x_{t|t-1}^{i}\right)$ is non-zero. Since rational news are zero on average, diagnostic expectations over-react on impact but then systematically revert to rationality, creating excess volatility.

In contrast to traditional departures from rationality such as adaptive expectations, diagnostic expectations are forward-looking in that they depend on parameters of the true data generating process. They are characterized by the "kernel of truth" property: they exaggerate true patterns in the data. Positive news are objectively associated with improvement, but representativeness causes excess focus on the right

[11] Equation (8) is reminiscent of overconfidence, which is in fact often modeled as inflating the signal to noise ratio of private information. In our current setup, where s_{t}^{i} is private to each forecaster, the Diagnostic Kalman Filter is equivalent to this formulation of overconfidence. In other settings it is possible to distinguish these mechanisms, which are psychologically very different. First, over-confidence refers only to private information, while representativeness causes over-reaction to all information, including the public one, as shown in BGLS (2017). Second, depending on the data generating process, representativeness and diagnostic expectations can cause distortions also in the variance and other moments. Finally, representativeness unifies distortions in expectations with a variety of errors in probabilistic judgments, including conjunction and disjunction fallacies, and also social stereotypes, which overconfidence cannot account for.

tail, generating excessive optimism. Likewise, expectation revisions should exaggerate the true properties of the underlying data generating process. If fundamentals are more persistent, kernel of truth implies that expectations should react more strongly to news. As we show in Section 5.2, the kernel of truth further implies that expectations exaggerate autocorrelation features of time series, so that the impact of longer lags may be overstated. In sum, the kernel of truth yields distinctive predictions that can be tested against conventional mechanical models of extrapolation such as adaptive expectations. We revisit such cross sectional and time-series implications in Sections 5 and 6.

Consider the implications of Diagnostic Expectations for forecasts and forecast errors. Section 3 presented two seemingly contradictory findings: predominance of under-reaction in consensus forecasts, and of over-reaction in individual forecasts. Define the consensus diagnostic forecast of x_{t+h} at time t as

$$x_{t+h|t}^{\theta} = \int x_{t+h|t}^{i,\theta} di = \rho^h \int x_{t|t}^{i,\theta} di,$$

so that the Diagnostic forecast error and revision are respectively given by $x_{t+h} - x_{t+h|t}^{\theta}$ and $x_{t+h|t}^{\theta} - x_{t+h|t-1}^{\theta}$. In the appendix, we prove the following result.

Proposition 2 *Under the Diagnostic Kalman Filter, the estimated coefficients of regression (2) at the consensus and individual level, β_1 and β_1^p, are given by:*

$$\frac{cov\left(x_{t+h} - x_{t+h|t}^{\theta}, x_{t+h|t}^{\theta} - x_{t+h|t-1}^{\theta}\right)}{var\left(x_{t+h|t}^{\theta} - x_{t+h|t-1}^{\theta}\right)} = (\sigma_\epsilon^2 - \theta\Sigma) g(\sigma_\epsilon^2, \Sigma, \rho, \theta) \tag{10}$$

$$\frac{cov\left(x_{t+h} - x_{t+h|t}^{i,\theta}, x_{t+h|t}^{i,\theta} - x_{t+h|t-1}^{i,\theta}\right)}{var\left(x_{t+h|t}^{i,\theta} - x_{t+h|t-1}^{i,\theta}\right)} = -\frac{\theta(1+\theta)}{(1+\theta)^2 + \theta^2\rho^2} \tag{11}$$

where $g(\sigma_\epsilon^2, \Sigma, \rho, \theta) > 0$ is a function of parameters. As a result, for $\theta \in (0, \sigma_\epsilon^2/\Sigma)$ the Diagnostic Kalman Filter entails a positive consensus coefficient $\beta_1 > 0$, and a negative individual coefficient $\beta_1^p < 0$.

When representative types are not too overweighed, $\theta < \sigma_\epsilon^2/\Sigma$, the Diagnostic Kalman Filter can reconcile positive consensus coefficients with negative individual level coefficients, consistent with the broad patterns of Section 3. Intuitively, when individual analysts over-react to their own information, a positive forecast revision by a given analyst is associated with excess optimism, while a negative revision is associated with his excess pessimism, which both imply $\beta_1^p < 0$. At the consensus level, however, matters are different. Individual analysts over-react to their own information but they don't react at all to the information received by the other analysts (which they do not observe). This is a force toward under-reaction to average information, which is particularly strong if individual analysts receive very noisy information. In fact, when σ_ϵ^2/Σ is high, a forecaster's neglect of the signals observed by the other forecasters entails a large loss of information. As a result, when noise is large enough, each analyst severely under-reacts to the information held by all other forecasters, so the average analyst consequently under-reacts to the average analyst information, even if each analyst over reacts to his own news.

The condition on θ that reconciles individual level overreaction and aggregate under-reaction has an intuitive interpretation. In fact, $\theta < \sigma_\epsilon^2/\Sigma$ is equivalent to the diagnostic Kalman gain $(1+\theta)\frac{\Sigma}{\Sigma+\sigma_\epsilon^2}$ being smaller than 1. This means that, as long as individual forecasters filter news to some extent, consensus forecasts exhibit underreaction, even if they discount information too little.

Compared to Diagnostic Expectations, Noisy Rational Expectations ($\theta = 0$) can generate under-reaction of consensus forecasts, $\beta_1 > 0$, but not over-reaction of individual analysts, $\beta_1^p < 0$. The reason is that in that model forecasters use the limited information at their disposal in an optimal, rational, way. As a result, their forecast error is uncorrelated with their own forecast revision. As evident from Equations (9) and (10), when $\theta = 0$ there is under-reaction at the consensus level but no individual-level predictability. This prediction is inconsistent with the evidence of Section 3.

Finally, Proposition 2 also illustrates the cross-sectional implications of the kernel of truth mentioned above: the predictability of forecast errors depends on the true features of the data generating process in the sense that the coefficients estimated at the pooled and individual analyst levels depend on the parameters characterizing the data generating process $(\sigma_\epsilon^2, \Sigma, \rho, \theta)$. In particular, higher noise to signal ratio σ_ϵ^2/Σ implies stronger consensus under-reaction (i.e. higher β_1).

Table 4 summarizes the predictions of three departures from rational expectations for the findings of Section 3. These include: Rational Inattention (which shares the broad predictions of Noisy Rational

Expectations), Diagnostic Expectations, and Mechanical Extrapolation (adaptive expectations). We evaluate these models according to three predictions: 1) consensus level predictability, 2) individual level predictability, and 3) dependence of forecast revisions on the true features of the data generating process.

Table 4.

Model	Consensus	Individual	Updating
Noisy Rational	Underreaction	no predictability	depends on fundamentals
Diagnostic	consistent with underreaction	overreaction	depends on fundamentals
Mechanical / Adaptive	Undetermined	underreaction for persistent series	does not depend on fundamentals

Let us compare Rational Inattention to Diagnostic Expectations. The broad pattern of Section 3 – the positive predictability of consensus forecast errors and the negative predictability of individual forecast errors for 9 or 10 out of 20 series -- is consistent with diagnostic expectations but not with rational inattention. The evidence for 4 series out of 20 – the GDP price deflator, the investment variables, and the Federal Funds rate – is consistent with rational inattention, featuring $\beta_1 > 0$ and β_1^p statistically indistinguishable from zero. Finally, the results for the 3-month T-bill rate (in SPF and Blue Chip) and the unemployment rate are consistent with neither Rational Inattention nor Diagnostic Expectations because they exhibit under-reaction at both the consensus and individual level, $\beta_1, \beta_1^p > 0$. Perhaps the behavior of these latter series, as well as of other series, may be accounted for by individual under-reaction due to mechanical adaptive expectations.

Overall, most of the evidence is consistent with Diagnostic Expectations, but Rational Inattention or Adaptive Expectations may play a role for some series. We further assess these models next.

5. Kernel of Truth

We run two sets of tests. The first is cross sectional, based on the persistence of the different series. By looking at how forecasts depend on persistence, we can check whether they are backward looking (as in Adaptive Expectations) or forward looking (as in Rational Inattention or Diagnostic Expectations) and, if the latter, which model best accounts for the evidence. The second test is a time series test, assessing whether expectations display over-reaction to longer lags of the series in question.

5.1 Persistence Tests

Under both Noisy Rational Expectations and Diagnostic Expectations, the forecast revision made at t about x_{t+h} is given by:

$$x_{t+h|t}^{i} - x_{t+h|t-1}^{i} = \rho\left(x_{t+h-1|t}^{i} - x_{t+h-1|t-1}^{i}\right).$$

The revision h periods ahead reflects the forecast revision about the same variable $h-1$ periods ahead, adjusted by the persistence ρ of the series. The idea is simple: when forecasts are forward looking, more persistent series should exhibit stronger news-based updating.

Under adaptive expectations, in contrast, updating is mechanical and should not depend on the true persistence of the forecasted process. Formally, in this case:

$$x_{t+h|t}^{i} - x_{t+h|t-1}^{i} = \mu\left(x_{t+h-1|t}^{i} - x_{t+h-1|t-1}^{i}\right),$$

where μ is a positive constant independent of ρ (we formally show this in the Appendix).

To assess this prediction, we fit an AR(1) for the actuals of each series and estimate ρ. The actuals have the same format as the forecast variables,[12] and we use the exact time period for which the forecasts are available.[13] We estimate the following individual level regression:

$$x_{t+3|t}^{i} - x_{t+3|t-1}^{i} = \gamma_o^p + \gamma_1^p\left(x_{t+2|t}^{i} - x_{t+2|t-1}^{i}\right) + \epsilon_{t+3}^{i}$$

[12] Here we follow CG and estimate persistence directly using autoregressions. Some of the series (e.g. interest rates) have time trends and are not stationary; in these cases we estimate persistence by fitting an ARIMA(1,1,0) process.
[13] Thus the properties of the actuals can be slightly different for the same variable from SPF and BlueChip (e.g. real GDP growth in SPF and Blue Chip), as these two datasets generally span different time periods.

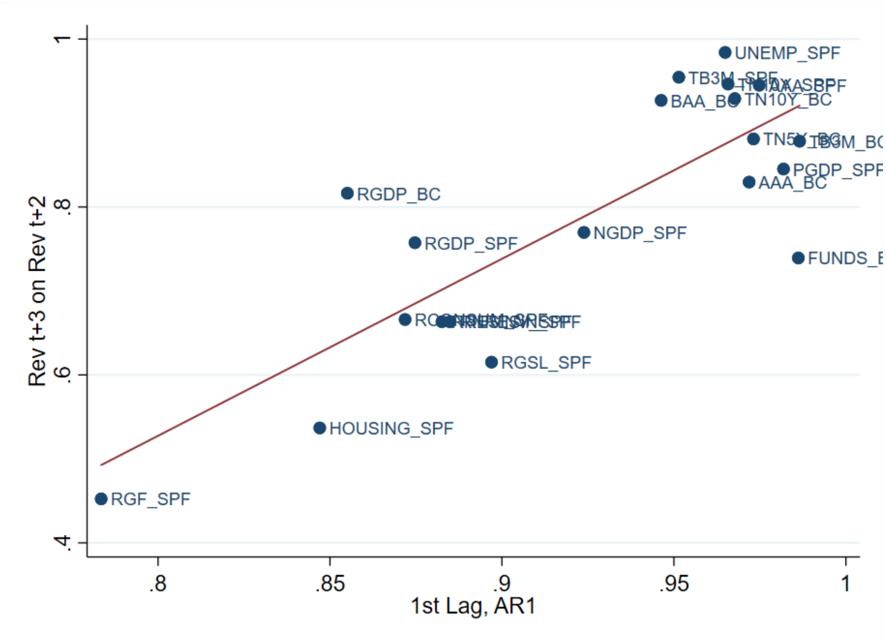
We estimate the same regression at the consensus level, which yields coefficients estimates γ_0 and γ_1. We then regress the slope coefficients γ_1^p and γ_1 on the estimated persistence $\hat{\rho}$ of each series. By integrating this equation, it is easy to see that consensus forecasts should satisfy the same condition.

The results of the exercise are reported in Figure 1 Panel A. At both the individual and the consensus level, the evidence shows that the more persistent series display larger forecast revisions. While we only have 20 series, the correlation is statistically different from zero with a p-value less than 0.001.[14] In line with forward-looking models, analysts take persistence into account when forming their forecasts. This evidence is inconsistent with adaptive expectations, where forecasters update mechanically, without taking into account the true properties of the data generating process, including persistence. This result is also robust to a series having richer dynamics, as it depends only on the first autocorrelation lag. The pattern is similar for consensus forecasts, as shown in Figure 1 Panel B.

Figure 1. Properties of Forecast Revisions and Actuals

In Panel A, the y-axis is the regression coefficient γ_1^p from regression $x_{t+3|t}^i - x_{t+3|t-1}^i = \gamma_o^p + \gamma_1^p(x_{t+2|t}^i - x_{t+2|t-1}^i) + \epsilon_{t+3}^i$. The x-axis is the persistence measured from an AR(1) regression of the actuals corresponding to the forecasts. For each variable, the AR(1) regression uses the same time period as when the forecast data is available. In Panel B, the y-axis is the regression coefficient from the parallel specification using consensus forecasts.

Panel A. Individual Level Coefficients



[14] The results in Figure 1 and 2 obtain also if we exclude the Blue Chip series that are also available in SPF (e.g. real GDP, 3-month Treasuries, 10-year Treasuries, AAA corporate bond rate).

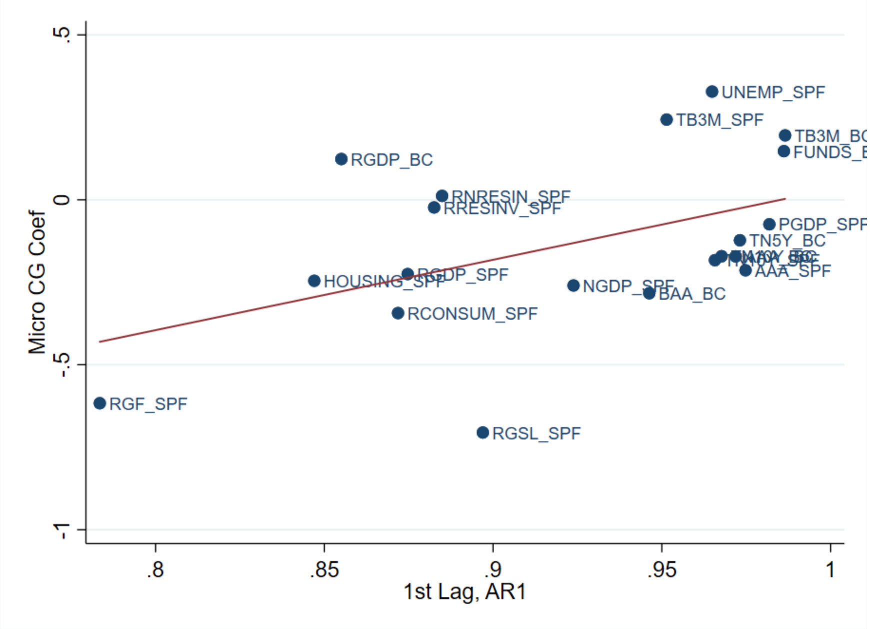
Panel B. Consensus Coefficients

Another strategy is to assess the correlation between the persistence of a series and the CG coefficient of reaction to news. For this test, Diagnostic Expectations do not have clear predictions at the consensus level. Indeed, the coefficient $(\sigma_\epsilon^2 - \theta\Sigma)g(\sigma_\epsilon^2, \Sigma, \rho, \theta)$ in Equation (10) can be either decreasing or increasing in persistence ρ, depending on parameter values. A more direct test is to check the correlation between the CG coefficient estimated at the individual level and the persistence of the series in question. In fact, Equation (11) predicts that the coefficient should increase, i.e. get closer to zero, as persistence ρ increases. The intuition is that when the series is more persistent, forecast revisions become more volatile, even if due to noise, which reduces their correlation with forecast errors. Under Noisy Rational Expectations, on the other hand, individual coefficients should be zero, so they should be uncorrelated with the persistence of the series that forecasters are trying to predict.

Figure 2 shows the correlation for the CG coefficient estimated from individual-level regressions. We find that the CG coefficient rises with persistence, which lends additional support for Diagnostic Expectations. The correlation is statistically different from zero with a *p*-value of 0.035.

Figure 2. CG Regression Coefficients and Persistence of Actual

Plots of individual level CG regression (forecast error on forecast revision) coefficients against the persistence of the actual variable (x-axis).



5.2. Kernel of Truth in the Time Series

We now analyze the possibility that some of the forecasted series may be influenced by longer lags, and in particular that may be better described by an AR(2) process. As discussed by Fuster, Laibson and Mendel (2010), several macroeconomic variables follow hump-shaped dynamics with short-term momentum and longer-term reversals. Considering this possibility is relevant for two reasons.

First, under the kernel of truth, forecasters should exaggerate true features of the data generating process, including the presence of long-term reversals. Checking whether longer lags are exaggerated in expectations thus allows us to further distinguish Diagnostic Expectations from Rational Inattention. This also allows us to compare these approaches to Natural Expectations, a model proposed by Fuster, Laibson and Mendel (2010) in which agents forecast and AR(2) process “as if” it was AR(1) in changes. In a stationary setting, this means that agents exaggerate the short run persistence of the process while dampening long-run reversals. Second, long-term reversals may help us to better understand our basic results in Section 3. In particular, AR(2) dynamics may contaminate the evidence of over and under-reaction to news documented in Section 3. We clarify this point below.

5.2.1 Diagnostic Expectations with AR(2) Processes

Suppose that the state which agents seek to forecast follows an AR(2) process:

$$x_{t+3} = \rho_2 x_{t+2} + \rho_1 x_{t+1} + u_t. \tag{12}$$

If $\rho_2 > 0$ and $\rho_1 < 0$, the variable displays short-term momentum and long-term reversal. Each agent now observes two signals, one about the current state $s_t^i = x_t + \epsilon_t^i$ and another about the past state $s_{t-1}^i = x_{t-1} + \epsilon_{t-1}^i$. The optimal solution to this inference problem is again provided by the Kalman filter.

Consider diagnostic expectations. Denote by $x_{t+s|t}^{i,\theta}$ the diagnostic expectation held by forecaster i about x_{t+s} at time t. The diagnostic expectation of x_{t+3} is then a linear combination of the diagnostic forecasts of x_{t+2} and x_{t+1} according to the autoregressive parameters ρ_1 and ρ_2:

$$x_{t+3|t}^{i,\theta} = \rho_2 x_{t+2|t}^{i,\theta} + \rho_1 x_{t+1|t}^{i,\theta}.$$

In the appendix we show that the diagnostic forecasts about these intermediate outcomes take the form of a distorted Kalman filter in which the signal to noise ratio of each signal is exaggerated. The Diagnostic forecast revision for $t+3$ at time t is a linear combination of the current Diagnostic Forecast revisions about the intermediate states:

$$x_{t+3|t}^{i,\theta} - x_{t+3|t-1}^{i,\theta} = \rho_2\left(x_{t+2|t}^{i,\theta} - x_{t+2|t-1}^{i,\theta}\right) + \rho_1\left(x_{t+1|t}^{i,\theta} - x_{t+1|t-1}^{i,\theta}\right). \tag{13}$$

This decomposition suggests a way to assess forecasters' reaction to information in an AR(2) setting, generalizing Equation (2). Denote by $FR_{t,t+1}^i$ the forecast revision at t about next period $t+1$. Likewise, denote by $FR_{t,t+2}^i$ the forecast revision at t about $t+2$. These forecast revisions are not indexed by θ because they represent data, not predictions of a diagnostic model. The forecasters' reaction to information can then be assessed by running the regression:

$$x_{t+3} - x_{t+3|t}^i = \delta_0^p + \delta_2^p FR_{t,t+2}^i + \delta_1^p FR_{t,t+1}^i + \epsilon_{t,t+h}. \tag{13}$$

In other words the forecast error at $t+3$ is predicted from current forecast revisions about the short and medium term. The Diagnostic Expectations model has the following implication.

Proposition 3. *Under the Diagnostic Kalman filter, the estimated coefficients* $\hat{\delta}_1^p$ *and* $\hat{\delta}_2^p$ *in Equation* (13) *are proportional to the negative of the AR(2) coefficients:*

$$\hat{\delta}_1^p \propto -\rho_1\theta, \tag{14}$$

$$\hat{\delta}_2^p \propto -\rho_2\theta. \tag{15}$$

As in the case of AR(1), rational expectations ($\theta = 0$) imply that individual forecast errors cannot be predicted by any forecast revisions, including those on the right hand side of (13). In contrast, diagnostic expectations imply that the coefficients should be non-zero, with flipped signs relative to the data generating process. In line with the kernel of truth, this predictability reflects over-reaction to the effect of lags in the true process. Suppose for instance that the process has short-term momentum, namely $\rho_2 > 0$. Then, over-reaction means that upward forecast revisions about x_{t+2} lead to exaggerated optimism about x_{t+3} and thus negative forecast errors. This yields $\hat{\delta}_2^p < 0$ in Equation (15), reproducing the basic insight of Equation (2). Suppose instead that the process has long-term reversal, namely $\rho_1 < 0$. Then, over-reaction to long-term reversal means that upward forecast revisions about x_{t+1} lead to exaggerated pessimism about x_{t+3} and thus positive forecast errors. This yields in $\hat{\delta}_1^p > 0$ in Equation (14).

Proposition 3 shows why assessing the AR(2) structure of our series is important to test the model. First, kernel of truth makes precise predictions for how forecast revisions should predict forecast errors on the basis of the data generating process. Diagnostic Expectations imply that forecast revisions about short and medium term conditions should predict forecast errors with a sign opposite to their true effects on x_{t+3}.

Second, Proposition 3 implies that the tests of Section 3 may not reliably distinguish over- or under-reaction when lags have different signs. Indeed, suppose that the AR(2) process features short term momentum, $\rho_2 > 0$, and long term reversals, $\rho_1 < 0$. Positive news at t may then trigger an upward revision of both the forecast of the short-term x_{t+1} and of the medium-term x_{t+2}. The former creates excess pessimism, the latter excess optimism. If the first effect is strong, it can reduce excess optimism after good news, making it harder to detect over-reaction using the specification of Section 3.

We can compare Diagnostic to Natural Expectations in this setting of the AR(2) process in Equation (12), which exhibits short term momentum and long term reversal. Under Natural Expectations, agents form their forecasts by fitting an AR(1) process in changes $(x_{t+1} - x_t) = \varphi(x_t - x_{t-1}) + v_{t+1}$. The resulting estimate for the autoregressive term is $\varphi = (\rho_1 - \rho_2 - 1)/2$. For a stationary process, this implies that Natural Expectations exaggerate the short run persistence of the series while dampening long-term reversals.[15] The overall comparison with Diagnostic Expectation is presented in Table 5.

Table 5.

Model	**Individual $\hat{\delta}_2^p$**	**Individual $\hat{\delta}_1^p$**
Noisy Rational	Zero	Zero
Diagnostic	Negative, due to overreaction to short term momentum	Positive, due to over-reaction to long term reversal
Natural Expectations	Negative, due to overreaction to short term momentum	Negative, due to under-reaction to long term reversal

Diagnostic and Natural Expectations share the same prediction concerning the coefficient on forecast revision at $t + 2$. In both models forecasters exaggerate short run persistence, predicting a negative $\hat{\delta}_2^p$. On the other hand, Diagnostic and Natural Expectations make opposite predictions about the coefficient on forecast revision at $t + 1$. Diagnostic Expectations over-react to long term reversals, predicting $\hat{\delta}_1^p > 0$, while Natural Expectations under-react to long term reversals, predicting $\hat{\delta}_1^p < 0$.

In the remainder of the section, we test the predictions of Proposition 3 about the term structure of expectations. Section 5.2.2 performs some tests to assess which of our 16 variables can be more accurately described by an AR(2) rather than an AR(1). We do not aim to find the unconstrained optimal ARMA(k, q) specification, which is a notoriously difficult task. We only wish to capture the simplest longer lags and see whether expectations react to them as predicted by the model. Section 5.2.3 then estimates Equation (13) for the variables that are found to be better approximated by an AR(2) process.

[15] Indeed, the "intuitive" process under this model is $x_{t+1} = (1 + \varphi)x_t - \varphi x_{t-1} + v_{t+1}$. The original AR(2) process is stationary if $\rho_1 - \rho_2 < 1$, $\rho_1 + \rho_2 < 1$ and $|\rho_2| < 1$. This implies that $1 + \varphi > \rho_1$ and that $0 < \varphi < |\rho_2|$.

5.2.2 AR(1) vs AR(2) Dynamics

To assess whether some series are better described by an AR(2) model, we first fit a quarterly AR(2) process to our 20 series.[16] Figure 4 below plots the estimates for the autocorrelation parameters ρ_1 and ρ_2 for the relevant series. As before, the actuals have the same format as the forecast variables, and for each series the regression covers the time period when the forecast data is available.

For all series, the sign of coefficients is indicative of positive momentum at short horizons ($\rho_2 > 0$) and long-run reversals ($\rho_1 < 0$).[17] To assess which dynamics better describe the series, we compare the AR(2) estimates to the AR(1) estimates from Section 5.1. Table 5 below shows the Bayesian Information Criterion (BIC) score associated with each fit.

For the majority of series, AR(2) dynamics are favored over AR(1). The tests favor AR(1) dynamics only for real consumption (SPF) and the BAA bond rate (BC), while for the 10-year Treasury rate series the tests are inconclusive.[18] Indeed, the estimates for the long-run reversals are weakest for these series. Taken together, Table 6 and Figure 4 indicate that hump shaped dynamics are a key feature of several series, which should shape expectations under the hypothesis of kernel of truth.

---

[16] Just like for the case of AR(1), for growth variables we run quarterly AR(2) regressions of growth from $t-1$ to $t+3$. For variables in levels, we run quarterly regressions in levels. We run separate regressions for the variables that occur both in SPF and BC, because they cover slightly different time periods.

[17] We check whether multicollinearity may affect our results in this Section, given that forecasts revisions at different horizons are often highly correlated. The standard issue with multicollinearity is the coefficients are imprecisely estimated, which we do not find to be the case. We also perform simulations to verify that the correlation among the right hand side variables by itself does not mechanically lead to the patterns we observe.

[18] The Akaike Information Criterion (AIC) yields similar results, except that it positively identifies the TN10Y series as AR(2). To interpret the IC scores, recall that lower scores represent a better fit. The likelihood ratio $\frac{\Pr(AR2)}{\Pr(AR1)}$ is estimated as $exp\left[-\frac{BIC_{AR2}-BIC_{AR1}}{2}\right]$, so that $\Delta BIC_{2-1} = -2$ means the AR(2) model is 2.7 times more likely than the AR(1) model. We follow the standard usage of considering scores below -2 as evidence for AR(2) over AR(1).

Figure 4. AR(2) Coefficients of Actuals

For each variable, the AR(2) regression uses the same time period as when the forecast data is available. The blue circles show the first lag and the red diamonds show the second lag. Standard errors are Newey-West, and the vertical bars show the 95% confidence intervals.

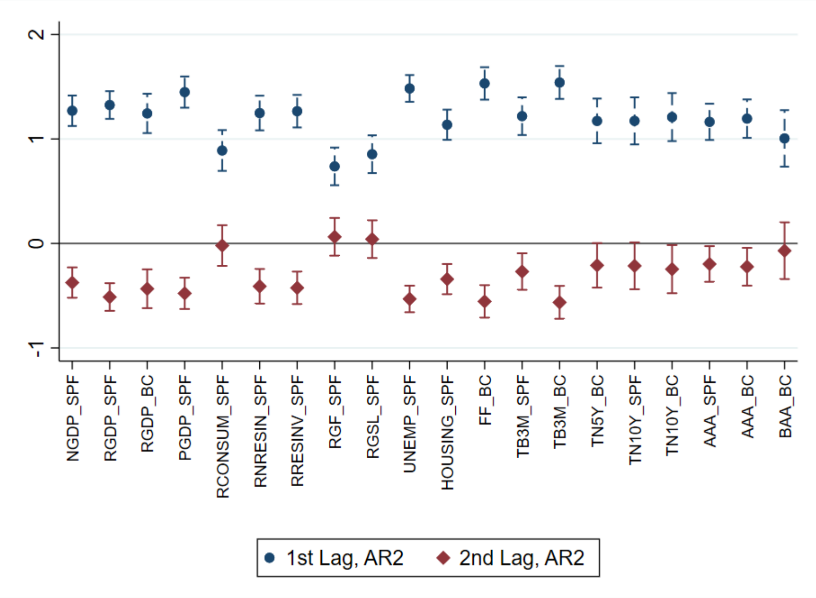


Table 6. BIC of AR(1) and AR(2) Regressions of Actuals

This table shows the BIC statistic corresponding to the AR(1) and AR(2) regressions of the actuals. The final column shows the specification that has a lower BIC (preferred).

Variable	BIC_{AR1}	BIC_{AR2}	$\Delta BIC_{2\text{-}1}$	model
Nominal GDP (SPF)	-1133.74	-1149.13	-15.39	AR(2)
Real GDP (SPF)	-1120.33	-1164.52	-44.19	AR(2)
Real GDP (BC)	-618.50	-626.83	-8.33	AR(2)
GDP Price Index Inflation (SPF)	-1423.70	-1456.90	-33.20	AR(2)
Real Consumption (SPF)	-924.47	-911.66	12.82	AR(1)
Real Non-Residential Investment (SPF)	-509.72	-524.37	-14.65	AR(2)
Real Residential Investment (SPF)	-375.81	-401.05	-25.25	AR(2)
Real Federal Government Consumption (SPF)	-560.97	-553.12	7.85	AR(1)
Real State&Local Govt Consumption (SPF)	-905.91	-896.23	9.68	AR(1)
Housing Start (SPF)	-250.88	-265.89	-15.01	AR(2)
Unemployment (SPF)	168.69	111.57	-57.12	AR(2)
Fed Funds Rate (BC)	191.89	149.87	-42.02	AR(2)
3M Treasury Rate (SPF)	240.87	232.25	-8.62	AR(2)
3M Treasury Rate (BC)	163.27	118.76	-44.51	AR(2)
5Y Treasury Rate (BC)	126.30	123.51	-2.79	AR(2)
10Y Treasury Rate (SPF)	89.66	89.91	0.25	AR(1)
10Y Treasury Rate (BC)	86.54	84.80	-1.74	AR(2)
AAA Corporate Bond Rate (SPF)	129.84	118.64	-11.20	AR(2)
AAA Corporate Bond Rate (BC)	86.05	84.72	-1.32	AR(2)
BAA Corporate Bond Rate (BC)	58.33	61.79	3.46	AR(1)

5.2.3 Empirical Tests of Over-Reaction with AR(2) dynamics

We next restrict the analysis to the series for which AR(2) is favored, and test the prediction of Proposition 3 by estimating Equation (13). Given that all AR(2) series we consider exhibit short term momentum $\rho_2 > 0$ and long-term reversals $\rho_1 < 0$, under Diagnostic Expectations the estimated coefficient on medium term forecast revision should be negative, $\hat{\delta}_2^p < 0$, while the estimated coefficient on short term forecast revision should be positive, $\hat{\delta}_1^p > 0$. To test this prediction, we run regressions on the pooled individual level data, as in Section 3.

Figure 5 shows, for each relevant series, the forecast error regression coefficients $\hat{\delta}_2^p$ and $\hat{\delta}_1^p$ estimated from Equation (13). Table 7 also displays these two coefficients, together with their corresponding standard errors and *p*-values. In line with the predictions of the model, the signs of the coefficients indicate that the short-term revision positively predicts forecast errors ($\hat{\delta}_1^p > 0$ for all 15 series, 10 of which are statistically significant at the 5% level) while the medium-term revision negatively predicts them ($\hat{\delta}_2^p < 0$ for 12 out of 15 series, 8 of which are statistically significant at the 5% level). To further assess the results, we perform a test of joint significance for $\hat{\delta}_2^p < 0$, $\hat{\delta}_1^p > 0$. We resample the data using block bootstrap, and calculate the fraction of times when $\hat{\delta}_2^p < 0$, $\hat{\delta}_1^p > 0$ holds, as shown in the last column of Table 7. The probability is greater than 95% for 8 out of the 15 series.

Figure 5. Coefficients in CG Regression AR(2) Version

This plot shows the coefficients δ_2^p (blue circles) and δ_1^p (red diamonds) from the regression in Equation (13). Standard errors are clustered by both forecaster and time, and the vertical bars shown the 95% confidence intervals.

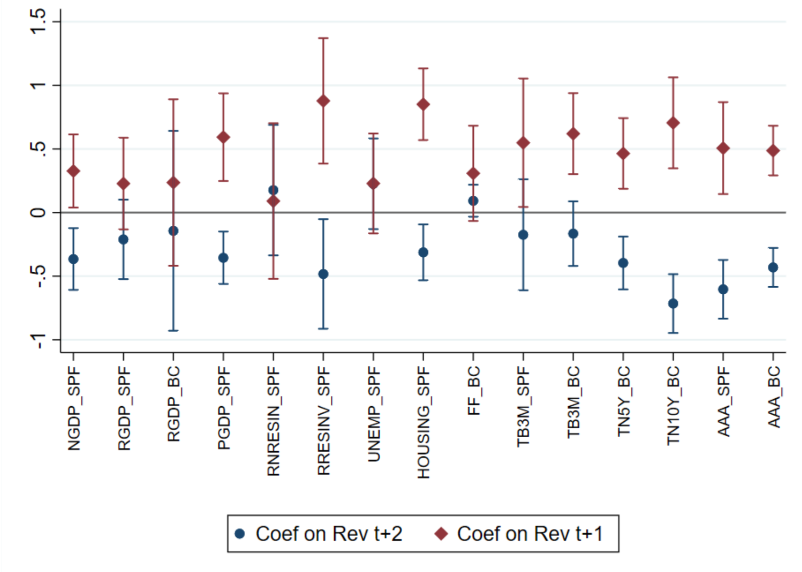


Table 7. Coefficients in CG Regression AR(2) Version

Coefficients δ_2^p and δ_1^p from the regression in Equation (13), together with the corresponding standard errors and *p*-values. The final column resamples the data using block bootstrap and shows the probability of $\delta_2^p < 0$ and $\delta_1^p > 0$.

Variable	δ_2^p	s.e.	*p*-val	δ_1^p	s.e.	*p*-val	Prob $\delta_2^p < 0$ & $\delta_1^p > 0$
Nominal GDP (SPF)	-0.37	0.12	0.00	0.33	0.15	0.03	0.99
Real GDP (SPF)	-0.21	0.16	0.19	0.23	0.18	0.22	0.86
Real GDP (BC)	-0.14	0.40	0.72	0.24	0.33	0.48	0.78
GDP Price Index Inflation (SPF)	-0.36	0.11	0.00	0.59	0.18	0.00	0.99
Real Non-Residential Investment (SPF)	0.18	0.26	0.50	0.09	0.31	0.77	0.11
Real Residential Investment (SPF)	-0.48	0.22	0.03	0.88	0.25	0.00	1.00
Housing Start (SPF)	-0.31	0.11	0.01	0.85	0.14	0.00	1.00
Unemployment (SPF)	0.23	0.18	0.22	0.23	0.20	0.26	0.03
Fed Funds Rate (BC)	0.09	0.06	0.15	0.31	0.19	0.11	0.40
3M Treasury Rate (SPF)	-0.17	0.22	0.43	0.55	0.26	0.03	0.85
3M Treasury Rate (BC)	-0.17	0.13	0.20	0.62	0.16	0.00	0.92
5Y Treasury Rate (BC)	-0.40	0.11	0.00	0.46	0.14	0.00	1.00
10Y Treasury Rate (BC)	-0.72	0.12	0.00	0.71	0.18	0.00	1.00
AAA Corporate Bond Rate (SPF)	-0.60	0.12	0.00	0.51	0.18	0.01	1.00
AAA Corporate Bond Rate (BC)	-0.43	0.08	0.00	0.49	0.10	0.00	1.00

The fact that short-term revisions negatively predict forecast errors conditional on longer-term revisions, and that they do so in different directions, is consistent with the idea that forecasters exaggerate true patterns in the data, including for longer-term dynamics. In contrast, this finding is harder to reconcile with Natural Expectations, where forecasters neglect longer lags (in the current setting, this means fitting an AR(1) model even for AR(2) series).

Overall, the AR(2) analysis confirms and in fact strengthens the evidence for the prevalence of over-reaction in the data. Indeed, four of the seven series (PGDP_SPF, RRESINV_SPF, TN5Y_BC and TN10Y_BC) for which individual level forecast errors seemed unpredictable (Table 3), and thus consistent with Noisy Rational Expectations, show evidence of over-reaction in the AR(2) setting. In addition, the two series that seemed to display under-reaction at the individual level, unemployment and the 3-months T Bill rate, now show evidence of over-reaction to long-term reversals ($\hat{\delta}_1^p > 0$), albeit not significantly. In all these cases, it is possible that over-reaction to long term reversals moved the individual level coefficient in Table 4 close to zero or above, giving the false impression of rationality or under-reaction. Only for the variable RGDP_SPF, which displayed significant over-reaction under the AR(1) specification loses its significance at conventional level under the more appropriate AR(2) case. Overall, the AR(2) specification strengthens the case for over-reaction.

6. Calibration and Additional Robustness Checks

In this section, we simulate the model to assess the statistical properties of the Diagnostic Kalman filter and compare them to our regression results. This allows us to calibrate the diagnostic parameter θ and to perform some additional robustness checks of our basic findings.

[CALIBRATION]

6.1 Robustness

Non-Normal Shocks. Our model in Section 4 assumes that shocks are normally distributed, so we can apply Kalman filtering techniques. In the data, macro shocks are not necessarily normally distributed, but commonly have fat tails. One question is whether fat-tailed distributions of shocks generate biases in CG regression coefficients.

When shocks are non-normal, the closed form Kalman filter is no longer an appropriate inference algorithm, and a numerical particle filter must be used instead. We find that the qualitative predictions of diagnostic expectations are unchanged when fundamental shocks and measurement noise are distributed according to t-distributions: forecasters over-react to news, and the CG regression coefficient is negative at the individual level. The coefficient can be positive in consensus regressions if the noise to signal ratio $\sigma_{\epsilon}^{2}/\Sigma$ is sufficiently large or the diagnostic parameter θ is close to zero. Appendix C describes particle filtering in detail and presents the results.

7. Conclusion

Using data from both the Blue Chip Survey and the Survey of Professional Forecasters, we have investigated how professional forecasters react to information using the methodology of Coibion and Gorodnichenko (2015). We have found that, for individual forecasters, over-reaction to information is the norm, whereas for the consensus forecast the norm is under-reaction. We then applied a psychologically founded model of belief formation, diagnostic expectations, to these data. Diagnostic expectations generate over-reaction in individual data, but also yields many additional predictions, and showed that this model is consistent with individual forecast data for many series. We further showed that because different forecasters see different information and use different models, the consensus forecast may theoretically exhibit under-reaction, as previously shown by Coibion-Gorodnichenko and confirmed in our data.

The ubiquity of over-reaction in individual macroeconomic forecasts helps reconcile distinctive evidence in finance and macroeconomics. Financial economics has put together a lot of evidence of over-reaction in individual markets, such as housing, credit, and equities. It would be puzzling if macroeconomic forecasts were completely the opposite, but as we show this may be largely a consequence of aggregation.

We also find that individual forecasts are better described by diagnostic expectations than by mechanical models of extrapolation. Adaptive expectations have been criticized by Lucas and others for assuming that people are entirely backward looking. Because with diagnostic expectations, forecasters are forward looking but their judgment is distorted by representativeness, this model is not vulnerable to the Lucas critique. Of course, diagnostic expectations can serve as a micro-foundation of adaptive expectations and extrapolation at a crude level. At the same time, the kernel of truth property of diagnostic expectations produces exact predictions on when we can see overreaction in forecasts, which becomes extremely important in some contexts, such as credit cycles (Bordalo, Gennaioli, and Shleifer 2018).

A final benefit of our approach is that it enables to reconcile diverse evidence that identified distinctive features of expectations. At the most basic level, we sought to reconcile the evidence on individual and consensus forecasts. Perhaps more subtly, diagnostic expectations when extended to the AR(2) context enable us to model expectations for hump shaped series. In this setting, diagnostic expectations capture some features of Natural Expectations (Fuster et al. 2010), such as exaggeration of short term persistence, but also yield over-reaction to long term reversal, which seems to be an important feature of the data.

Our paper leaves at least two important problems to future work. We have stressed over-reaction in individual time series, which seems to be the norm in our data, but other studies have also found some rigidity in expectations even in individual data. For example, in their experimental study, Landier, Ma, and Thesmar (2017) find that beliefs of experimental subjects are best characterized by a mixture of anchoring to old forecasts and over-reaction to news. In this paper we have combined over-reaction with rigidity by incorporating representativeness in a noisy information setting. The reconciliation of anchoring with over-reaction to information based on psychological foundations remains an open problem.

We have not addressed the basic theoretical question: do individual or consensus beliefs matter for macroeconomic outcomes? For aggregate outcomes, what may matter is consensus expectations, so all one needs to know is that consensus expectations under-react. There are two problems in this view. First, over-reaction by individual forecasters can influence aggregate outcomes through dispersion in beliefs. Certain firms or sectors will over-invest, others will under-invest, creating aggregate misallocations. Second, in some circumstances news may be correlated across different agents, for instance when major innovations are introduced, or when repeated news in the same direction are highly informative of persistent changes. In these cases, individual over-reaction will entail aggregate over-reaction. There is of course evidence of aggregate over-reaction in the stock market going back to Shiller (1981), and in credit markets as well (Greenwood and Hanson 2013, Lopez-Salido et al. 2017). Whether over-reaction can account for macroeconomic phenomena such as investment booms or business cycles is a key question in this research area.

References

Adam, Klaus, Albert Marcet, and Johannes Beutel. “Stock Price Booms and Expected Capital Gains.” *American Economic Review* 107, no. 8 (2017): 2352-2408.

Amromin, Gene, and Steven A. Sharpe. “From the Horse's Mouth: Economic Conditions and Investor Expectations of Risk and Return.” *Management Science* 60, no. 4 (2013): 845-866.

Augenblick, Ned, and Eben Lazarus. “Restrictions on Asset-Price Movements Under Rational Expectations: Theory and Evidence.” Working Paper (2017).

Augenblick, Ned, and Matthew Rabin. “Belief Movement, Uncertainty Reduction, and Rational Updating.” Working Paper (2017).

Bacchetta, Philippe, Elmar Mertens, and Eric Van Wincoop. “Predictability in Financial Markets: What Do Survey Expectations Tell Us?” *Journal of International Money and Finance* 28, no. 3 (2009): 406-426.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. “A Model of Investor Sentiment.” *Journal of Financial Economics* 49, no. 3 (1998): 307-343.

Ben-David, Itzhak, John R. Graham, and Campbell R. Harvey. “Managerial Miscalibration.” Quarterly Journal of Economics 128, no. 4 (2013): 1547-1584.

Berger, Helge, Michael Ehrmann, and Marcel Fratzscher. “Geography, Skills or Both: What Explains Fed Watchers’ Forecast Accuracy of US Monetary Policy?” *Journal of Macroeconomics* 33, no. 3 (2011): 420-437.

Beshears, John, James J. Choi, Andreas Fuster, David Laibson, and Brigitte C. Madrian. “What Goes Up Must Come Down? Experimental Evidence on Intuitive Forecasting.” *American Economic Review* 103, no. 3 (2013): 570-574.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. “Stereotypes.” *Quarterly Journal of Economics* 131, no. 4 (2016): 1753-1794.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. “Diagnostic Expectations and Stock Returns.” Working Paper (2017).

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. “Diagnostic Expectations and Credit Cycles.” *Journal of Finance* (forthcoming).

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar. “Sticky Expectations and the Profitability Anomaly.” Working Paper (2017).

Carroll, Christopher D. “Macroeconomic Expectations of Households and Professional Forecasters.” *Quarterly Journal of Economics* 118, no. 1 (2003): 269-298.

Coibion, Olivier, and Yuriy Gorodnichenko. “Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts.” *American Economic Review* 105, no. 8 (2015): 2644-2678.

Coibion, Olivier, and Yuriy Gorodnichenko. “What Can Survey Forecasts Tell Us about Information Rigidities?” *Journal of Political Economy* 120, no. 1 (2012): 116-159.

Cutler, David M., James M. Poterba, and Lawrence H. Summers. “Speculative Dynamics and the Role of Feedback Traders.” *American Economic Review* (1990): 63-68.

De Bondt, Werner FM, and Richard H. Thaler. “Do Security Analysts Overreact?” *American Economic Review* 80, no.2 (1990): 52-57.

Frydman, Cary, and Gideon Nave. “Extrapolative Beliefs in Perceptual and Economic Decisions: Evidence of a Common Mechanism.” *Management Science* 63, no. 7 (2016): 2340 - 2352.

Fuhrer, Jeff. “Intrinsic Expectations Persistence: Evidence from Professional and Household Survey Expectations.” Working Paper (2017).

Fuster, Andreas, David Laibson, and Brock Mendel. “Natural Expectations and Macroeconomic Fluctuations.” *Journal of Economic Perspectives* 24, no. 4 (2010): 67-84.

Gabaix, Xavier. “A Sparsity-Based Model of Bounded Rationality.” *Quarterly Journal of Economics* 129, no. 4 (2014): 1661-1710.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer. “Expectations and Investment.” *NBER Macroeconomics Annual* 30, no. 1 (2016): 379-431.

Gennaioli, Nicola, and Andrei Shleifer. “What Comes to Mind.” *Quarterly Journal of Economics* 125, no. 4 (2010): 1399-1433.

Greenwood, Robin, and Samuel G. Hanson. “Waves in Ship Prices and Investment.” *Quarterly Journal of Economics* 130, no. 1 (2014): 55-109.

Greenwood, Robin, and Andrei Shleifer. “Expectations of Returns and Expected Returns.” *Review of Financial Studies* 27, no. 3 (2014): 714-746.

Hjalmarsson, Erik. "The Stambaugh Bias in Panel Predictive Regressions." *Finance Research Letters* 5, no. 1 (2008): 47-58.

Hommes, Cars, Joep Sonnemans, Jan Tuinstra, and Henk Van de Velden. “Coordination of Expectations in Asset Pricing Experiments.” *Review of Financial Studies* 18, no. 3 (2004): 955-980.

Jegadeesh, Narasimhan, and Sheridan Titman. “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency.” *Journal of Finance* 48, no. 1 (1993): 65-91.

Kendall, Maurice G. "Note on Bias in the Estimation of Autocorrelation." *Biometrika* 41, no. 3-4 (1954): 403-404.

La Porta, Rafael. “Expectations and the Cross-section of Stock Returns.” *Journal of Finance* 51, no. 5 (1996): 1715-1742.

Landier, Augustin, Yueran Ma, and David Thesmar. “New Experimental Evidence on Expectations Formation.” Working Paper (2017).

López-Salido, David, Jeremy C. Stein, and Egon Zakrajšek. “Credit Market Sentiment and the Business Cycle.” *Quarterly Journal of Economics* 132, no. 3 (2017): 1373-1426.

Mankiw, N. Gregory, and Ricardo Reis. “Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve.” *Quarterly Journal of Economics* 117, no. 4 (2002): 1295-1328.

Morris, Stephen, and Hyun Song Shin. "Social Value of Public Information." *American Economic Review* 92, no. 5 (2002): 1521-1534.

Nickell, Stephen. "Biases in Dynamic Models with Fixed Effects." *Econometrica* (1981): 1417-1426.

Pesaran, M. Hashem, and Martin Weale. "Survey Expectations." *Handbook of Economic Forecasting* (2006): 715-776.

Woodford, Michael. “Imperfect Common Knowledge and the Effects of Monetary Policy.” In *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, edited by Philippe Aghion, Roman Frydman, Joseph Stiglitz, and Michael Woodford, 25–58 (2003). Princeton: Princeton University Press.

Shiller, Robert J. “The Use of Volatility Measures in Assessing Market Efficiency.” *Journal of Finance* 36, no. 2 (1981): 291-304.

Sims, Christopher A. “Rational Inattention and Monetary Economics.” *Handbook of Monetary Economics* 3 (2010): 155-181.

Stambaugh, Robert F. "Predictive Regressions." *Journal of Financial Economics* 54, no. 3 (1999): 375-421.

Tversky, Amos, and Daniel Kahneman. “Judgment under Uncertainty: Heuristics and Biases.” *Science* 185, no. 4157 (1974): 1124-1131.

Appendix

A. Variable Definitions

1. NGDP_SPF

- Variable: Nominal GDP. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of nominal GDP in the current quarter and the next 4 quarters.
- Forecast variable: Nominal GDP growth from end of quarter *t*-1 to end of quarter *t*+3 $\frac{F_t x_{t+3}}{x_{t-1}} - 1$, where *t* is the quarter of forecast and x is the level of GDP in a given quarter; x_{t-1} uses the initial release of actual value in quarter *t*-1, which is available by the time of the forecast in quarter *t*.
- Revision variable: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial release of x_{t+3}. published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

2. RGDP_SPF

- Variable: Real GDP. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of real GDP in the current quarter and the next 4 quarters.
- Forecast variable: Real GDP growth from end of quarter *t*-1 to end of quarter *t*+3 $\frac{F_t x_{t+3}}{x_{t-1}} - 1$, where *t* is the quarter of forecast and x is the level of GDP in a given quarter; x_{t-1} uses the initial release of actual value in quarter *t*-1, which is available by the time of the forecast in quarter *t*.
- Revision variable: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

3. RGDP_BC

- Variable: Real GDP. Source: Blue Chip.
- Survey time: End of the middle month in the quarter/beginning of the last month in the quarter.
- Survey question: Real GDP growth (annualized rate) in the current quarter and the next 4 to 5 quarters.
- Forecast variable: Real GDP growth from end of quarter *t*-1 to end of quarter *t*+3 $F_t(z_t + z_{t+1} + z_{t+2} + z_{t+3})/4$, where *t* is the quarter of forecast and z_t is the annualized quarterly GDP growth in quarter *t*.
- Revision variable: $\frac{F_t(z_t+z_{t+1}+z_{t+2}+z_{t+3})}{4} - \frac{F_{t-1}(z_t+z_{t+1}+z_{t+2}+z_{t+3})}{4}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

4. PGDP_SPF

- Variable: GDP price deflator. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of GDP price deflator in the current quarter and the next 4 quarters.
- Forecast variable: GDP price deflator inflation from end of quarter *t*-1 to end of quarter *t*+3 $\frac{F_t x_{t+3}}{x_{t-1}} - 1$, where *t* is the quarter of forecast and x is the level of GDP price deflator in a given quarter; x_{t-1} uses the initial release of actual value in quarter *t*-1, which is available by the time of the forecast in quarter *t*.
- Revision variable: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

5. RCONSUM_SPF

- Variable: Real consumption. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of real consumption in the current quarter and the next 4 quarters.
- Forecast variable: Growth of real consumption from end of quarter *t*-1 to end of quarter *t*+3 $\frac{F_t x_{t+3}}{x_{t-1}} - 1$, where *t* is the quarter of forecast and x is the level of real consumption in a given quarter; x_{t-1} uses the initial release of actual value in quarter *t*-1, which is available by the time of the forecast in quarter *t*.
- Revision variable: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

6. RNRESIN_SPF

- Variable: Real non-residential investment. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of real non-residential investment in the current quarter and the next 4 quarters.
- Forecast variable: Growth of real non-residential investment from end of quarter *t*-1 to end of quarter *t*+3 $\frac{F_t x_{t+3}}{x_{t-1}} - 1$, where *t* is the quarter of forecast and x is the level of real non-residential investment in a given quarter; x_{t-1} uses the initial release of actual value in quarter *t*-1, which is available by the time of the forecast in quarter *t*.
- Revision variable: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

7. RRESIN_SPF

- Variable: Real residential investment. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.

- Survey question: The level of real residential investment in the current quarter and the next 4 quarters.
- Forecast variable: Growth of real residential investment from end of quarter *t*-1 to end of quarter *t*+3 $\frac{F_t x_{t+3}}{x_{t-1}} - 1$, where *t* is the quarter of forecast and x is the level of real residential investment in a given quarter; x_{t-1} uses the initial release of actual value in quarter *t*-1, which is available by the time of the forecast in quarter *t*.
- Revision variable: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

8. RGF_SPF

- Variable: Real federal government consumption. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of real federal government consumption in the current quarter and the next 4 quarters.
- Forecast variable: Growth of real federal government consumption from end of quarter *t*-1 to end of quarter *t*+3 $\frac{F_t x_{t+3}}{x_{t-1}} - 1$, where *t* is the quarter of forecast and x is the level of real federal government consumption in a given quarter; x_{t-1} uses the initial release of actual value in quarter *t*-1, which is available by the time of the forecast in quarter *t*.
- Revision variable: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

9. RGSL_SPF

- Variable: Real state and local government consumption. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of real state and local government consumption in the current quarter and the next 4 quarters.
- Forecast variable: Growth of real state and local government consumption from end of quarter *t*-1 to end of quarter *t*+3 $\frac{F_t x_{t+3}}{x_{t-1}} - 1$, where *t* is the quarter of forecast and x is the level of real state and local government consumption in a given quarter; x_{t-1} uses the initial release of actual value in quarter *t*-1, which is available by the time of the forecast in quarter *t*.
- Revision variable: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

10. UNEMP_SPF

- Variable: Unemployment rate. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average unemployment rate in the current quarter and the next 4 quarters.

- Forecast variable: Average quarterly unemployment rate in quarter *t*+3 $F_t x_{t+3}$, where t is the quarter of forecast and x is the level of unemployment rate in a given quarter.
- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4.

11. HOUSING_SPF

- Variable: Housing starts. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of housing starts in the current quarter and the next 4 quarters.
- Forecast variable: Growth of housing starts from quarter t-1 to quarter *t*+3 $\frac{F_t x_{t+3}}{x_{t-1}} - 1$, where *t* is the quarter of forecast and x is the level of housing starts in a given quarter; x_{t-1} uses the initial release of actual value in quarter *t*-1, which is available by the time of the forecast in quarter *t*.
- Revision variable: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actual variable: $\frac{x_{t+3}}{x_{t-1}} - 1$, using real time macro data: initial realease of x_{t+3} published in quarter *t*+4 and initial release of x_{t-1} published in quarter *t*.

12. FF_BC

- Variable: Federal funds rate. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average federal funds rate in the current quarter and the next 4 quarters.
- Forecast variable: Average quarterly 3-month federal funds rate in quarter *t*+3 $F_t x_{t+3}$, where *t* is the quarter of forecast and x is the level of federal funds rate in a given quarter.
- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}.

13. TB3M_SPF
- Variable: 3-month Treasury rate. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average 3-month Treasury rate in the current quarter and the next 4 quarters.
- Forecast variable: Average quarterly 3-month Treasury rate in quarter *t*+3 $F_t x_{t+3}$, where *t* is the quarter of forecast and x is the level of 3-month Treasury rate in a given quarter.
- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}.

14. TB3M_BC

- Variable: 3-month Treasury rate. Source: Blue Chip.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average 3-month Treasury rate in the current quarter and the next 4 quarters.
- Forecast variable: Average quarterly 3-month Treasury rate in quarter *t*+3 $F_t x_{t+3}$, where *t* is the quarter of forecast and x is the level of 3-month Treasury rate in a given quarter.

- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}.

15. TN5Y_BC

- Variable: 5-year Treasury rate. Source: Blue Chip.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average 5-year Treasury rate in the current quarter and the next 4 quarters.
- Forecast variable: Average quarterly 5-year Treasury rate in quarter *t*+3 $F_t x_{t+3}$, where *t* is the quarter of forecast and x is the level of 5-year Treasury rate in a given quarter.
- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}.

16. TN10Y_SPF

- Variable: 10-year Treasury rate. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average 10-year Treasury rate in the current quarter and the next 4 quarters.
- Forecast variable: Average quarterly 10-year Treasury rate in quarter *t*+3 $F_t x_{t+3}$, where *t* is the quarter of forecast and x is the level of 10-year Treasury rate in a given quarter.
- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}.

17. TN10Y_BC

- Variable: 10-year Treasury rate. Source: Blue Chip.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average 10-year Treasury rate in the current quarter and the next 4 quarters.
- Forecast variable: Average quarterly 10-year Treasury rate in quarter *t*+3 $F_t x_{t+3}$, where *t* is the quarter of forecast and x is the level of 10-year Treasury rate in a given quarter.
- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}.

18. AAA_SPF

- Variable: AAA corporate bond rate. Source: SPF.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average AAA corporate bond rate in the current quarter and the next 4 quarters.
- Forecast variable: Average quarterly AAA corporate bond rate in quarter *t*+3 $F_t x_{t+3}$, where *t* is the quarter of forecast and x is the level of AAA corporate bond rate in a given quarter.
- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}.

19. AAA_BC

- Variable: AAA corporate bond rate. Source: Blue Chip.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average AAA corporate bond rate in the current quarter and the next 4 quarters.
- Forecast variable: Average quarterly AAA corporate bond rate in quarter *t*+3 $F_t x_{t+3}$, where *t* is the quarter of forecast and x is the level of AAA corporate bond rate in a given quarter.
- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}.

20. BAA_BC

- Variable: BAA corporate bond rate. Source: Blue Chip.
- Survey time: Around the 3rd week of the middle month in the quarter.
- Survey question: The level of average BAA corporate bond rate in the current quarter and the next 4 quarters.
- Forecast variable: Average quarterly BAA corporate bond rate in quarter *t*+3 $F_t x_{t+3}$, where *t* is the quarter of forecast and x is the level of BAA corporate bond rate in a given quarter.
- Revision variable: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actual variable: x_{t+3}.

B. Robustness Checks

Table A1. Forecaster-by-Forecaster CG Regressions

Column "Pooled" shows the pooled panel CG regressions at the individual level (same as Table 3 column (4)). Column "By Forecaster (Median)" shows the median coefficient from forecaster-by-forecaster CG regressions; column "By Forecaster (%<0)" shows the fraction of forecasters where the coefficient is negative. For the forecaster-by-forecaster coefficients, we restrict to forecasters with at least 10 forecasts available.

Variable	Pooled	By Forecaster	
		Median	%<0
Nominal GDP (SPF)	-0.26	-0.14	0.63
Real GDP (SPF)	-0.23	-0.09	0.54
Real GDP (BC)	0.12	0.00	0.50
GDP Price Index Inflation (SPF)	-0.07	-0.11	0.57
Real Consumption (SPF)	-0.34	-0.20	0.83
Real Non-Residential Investment (SPF)	0.01	-0.20	0.58
Real Residential Investment (SPF)	-0.02	-0.32	0.64
Real Federal Government Consumption (SPF)	-0.62	-0.43	0.95
Real State&Local Govt Consumption (SPF)	-0.71	-0.50	0.91
Housing Start (SPF)	0.33	0.24	0.35
Unemployment (SPF)	-0.25	-0.19	0.73
Fed Funds Rate (BC)	0.15	0.21	0.27
3M Treasury Rate (SPF)	0.24	-0.02	0.51
3M Treasury Rate (BC)	0.20	0.20	0.28
5Y Treasury Rate (BC)	-0.12	-0.18	0.82
10Y Treasury Rate (SPF)	-0.18	-0.18	0.58
10Y Treasury Rate (BC)	-0.17	-0.29	0.86
AAA Corporate Bond Rate (SPF)	-0.21	-0.35	0.85
AAA Corporate Bond Rate (BC)	-0.17	-0.28	0.84
BAA Corporate Bond Rate (BC)	-0.28	-0.34	0.95

Table A2. Last Forecast Revision

The Table shows the pooled panel CG regressions at the consensus and individual levels (pooled panel regression) for horizon $h = 0$ (same as Table 3 columns 1, 2, 4, and 5).

Variable	β_1	t-stat	β_1^p	t-stat
Nominal GDP (SPF)	-0.05	-1.03	-0.14	-2.35
Real GDP (SPF)	0.06	1.01	-0.06	-1.15
Real GDP (BC)	0.16	1.04	-0.05	-0.54
GDP Price Index Inflation (SPF)	-0.01	-0.14	-0.10	-2.14
Real Consumption (SPF)	-0.12	-1.62	-0.23	-3.59
Real Non-Residential Investment (SPF)	0.03	0.50	-0.06	-0.85
Real Residential Investment (SPF)	0.23	3.74	0.04	0.99
Real Federal Government Consumption (SPF)	-0.08	-0.74	-0.22	-3.58
Real State&Local Govt Consumption (SPF)	-0.18	-2.84	-0.26	-3.33
Housing Start (SPF)	0.23	6.55	0.03	1.20

Unemployment (SPF)	0.42	5.95	0.09	2.09
Fed Funds Rate (BC)	-0.03	-0.89	-0.11	-2.25
3M Treasury Rate (SPF)	0.17	7.30	0.00	0.21
3M Treasury Rate (BC)	0.01	0.40	-0.18	-2.01
5Y Treasury Rate (BC)	0.12	3.27	0.00	0.04
10Y Treasury Rate (SPF)	0.15	3.34	-0.05	-1.86
10Y Treasury Rate (BC)	0.04	1.50	-0.01	-0.52
AAA Corporate Bond Rate (SPF)	0.07	1.29	-0.10	-2.15
AAA Corporate Bond Rate (BC)	-0.10	-2.46	-0.16	-4.74
BAA Corporate Bond Rate (BC)	0.04	1.26	-0.09	-3.43

C. Non-Normal Shocks and Particle Filtering

In the main text, we assume that both the innovations of the latent process, u_t, and the measurement error for each expert, ϵ_t, follow normal distributions. In this case, as all the posterior distributions are normal, the Kalman filter provides the closed form expression for the posterior densities. However, many processes for macro and financial variables may have heavy tails and more closely follow, for example, a t-distribution. In this case, while the point estimates of the Kalman filter still minimize mean-squared error (MSE), the mean and covariance estimates of the filter are no longer sufficient to determine the posterior distribution. Given that our formulation of diagnostic expectations involves a reweighting of the likelihood function, we require more than the posterior mean and variance to properly compute the diagnostic expectation distribution. In the following, we relax the normality assumption and verify the model predictions with a fat-tailed t-distribution.

C.1 Particle Filtering: Motivation and Set-Up

We start with the processes in Equations (3) and (4):

$$s_t^i = x_t + \epsilon_t^i, \quad x_t = \rho x_{t-1} + u_t$$

where x_t is the fundamental process and s_t^i is forecaster i's noisy signal. In Section 3, the shocks to these processes are assumed to be normal. In the following, we analyze the case where u_t follows a t-distribution.

Since the t-distribution is no longer conjugate to normal noise, one can no longer get closed form solutions. Instead, we draw from the posterior distribution in a Monte Carlo approach using the particle filter, a popular algorithm for simulating Bayesian inference on Hidden Markov Models (Doucet and Johansen 2011). We first briefly describe this approach, then formulate the application to diagnostic expectations, and finally show the simulation results for the CG forecast error on forecast revision regressions.

Particle filtering builds on the idea of importance sampling. Specifically, suppose we wish to estimate the expectation of $f(x)$, where x is distributed according to p; we are not able to sample from p, or in general unable to compute its precise density, but can compute p up to a proportionality constant:

$p(x) = \frac{1}{Z}\tilde{p}(x)$, where $Z = \int \tilde{p}(x)\, dx$ is the (unknown) normalizing constant. If we can sample from an arbitrary density q, we can use the following importance sampling mechanism to indirectly sample from p: for each sample from q, x_n, compute the importance weight $w_n = \frac{\tilde{p}(x_n)}{q(x_n)}$ and resample from $x_n$according to probability proportional to the weights. It is easy to see that the average of the weights estimates the proportionality factor Z : $\frac{1}{N}\sum_{n=1}^{N} w(x_n) \rightarrow \int \frac{\tilde{p}(x)}{q(x)} \cdot q(x)dx = \int \tilde{p}(x)\, dx = Z$. Consequently, one can easily derive that the resampled x_n converge in distribution to p: given any measurable function ϕ, the expectation of $\phi(x)$for the resampled x converges to $E_P\phi$:

$$\int \sum_{i=1}^{N} \phi(x_i)\, \frac{w(x_i)}{N}\frac{q(x_{1:N})}{Z}\, dx_{1:N} = \frac{1}{Z}\frac{1}{N}\sum_{i=1}^{N} \int \phi(x_i)\, \frac{\tilde{p}(x_i)}{q(x_i)} q(x_i) q(x_{-i})\, dx_{1:N} = \frac{1}{N}\sum_{i=1}^{N} E_p[\phi(x)] = E_p\phi$$

The algorithm above, called the sample-importance resample (SIR) algorithm, can be summarized in the following steps:

1. Sample N particles from q, denoted as $x_{1:N}$
2. For each x_i, compute$w_i = \frac{\tilde{p}(x_i)}{q(x_i)}$.
3. Resample according to probability $\propto w_i$

For the hidden Markov Process model, the above idea generalizes to give us a quick algorithm to sample from the filtering density $p(x_n|s_{1:n})$. Like the Kalman filter, the idea is to proceed inductively, using the following forward equation:

$$p(x_n|s_{1:n}) = \frac{g(s_n|x_n)\, p(x_n|s_{1:n-1})}{p(s_n|\, s_{1:n-1})} = \frac{\int g(s_n|x_n)\, f(x_n|x_{n-1})\, p(x_{n-1}|s_{1:n-1}) ds_{1:n-1} dx_{n-1}}{p(s_n|s_{1:n-1})}$$

By induction, suppose that we have samples from the previous filtered distribution $p(x_{n-1}|s_{1:n-1})$. Now, given a (conditional) proposal $q(x_n|x_{n-1}, s_{1:n})$for each sample, the recursive equality above suggests the resampling weights: $w(x_n \mid x_{n-1}) = \frac{g(s_n|x_n) f(x_n|x_{n-1})}{q(x_n \mid x_{n-1}, s_{1:n})}$. For the base case, where we have only seen the data point s_1, our filtered density $p(x_1|s_1)$is the standard Bayesian posterior, which can be sampled via importance sampling.

The particle filter algorithm refers to this extension of the SIR algorithm to the sequential setting:

1. At time n = 1, generate Ni.i.d. samples from a default proposal q.
2. Compute for each sample the weights $w(x_i) = \frac{\mu(x_i)\, g(s_1 \mid x_i)}{q(x_i)}$
3. Resample according to the weights, and store the sample.
4. For $n \geq 2$: for each x_{n-1}^i in the sample, propose x_n^i according to $q(x_n|x_{n-1} = x^i{}_{n-1}, s_{1:n})$
5. Compute for each x_n^i the weights $w(x_n{}^i) = \frac{g(s_n|x_n{}^i)\, f(x_n{}^i|x_{n-1}{}^i)}{q(x_n|x_{n-1}=x^i{}_{n-1}, s_{1:n})}$
6. Resample according to the weights, save as x_n^i.

Finally, we need to specify the proposal density $q(x_n|x_{n-1} = x^i{}_{n-1}, s_{1:n})$. It is well-known that the optimal proposal density should be the conditional distribution $p(x_n|x_{n-1} = x^i{}_{n-1}, s_n)$. If the latent Markov process is a simple AR(1) process with normal innovation, one can analytically derive the optimal proposal density $p(x_n|x_{n-1} = x^i{}_{n-1}, s_n)$.

$$x_n|x_{n-1}, s_n \sim N(\frac{\sigma_\epsilon{}^2}{\sigma_\epsilon{}^2 + \sigma_u{}^2}\rho\, x_{n-1} + \frac{\sigma_u{}^2}{\sigma_\epsilon{}^2 + \sigma_u{}^2}\, s_n, \frac{\sigma_\epsilon{}^2\ \ \sigma_u{}^2}{\sigma_\epsilon{}^2 + \sigma_u{}^2}) = N(\bar{\mu}, \bar{\Sigma})$$

While this result is only precise for normal processes, we shall still use $\bar{\mu}, \bar{\Sigma}$as location and scale parameters for our proposal, which is now t-distributed. If the original innovations have $d$degrees of freedom, our proposal will have $\frac{d+2}{2}$degrees of freedom, which have much thicker tails.

C.2 Application to Diagnostic Expectations

To analyze the case of diagnostic expectations, we only need to re-adjust the resampling weights by a simple likelihood ratio, as given by the following proposition:

Proposition A1 *Let $s^*(s_{1:n-1})$be the predictive expectation of s_n given $s_{1:n-1}$. The representativeness $R(x_n|s_{1:n}) = \frac{p(x_n|\, s_{1:n})}{p(x_n|\, s_{1:n-1}, s^*)}$can be simplified to the likelihood ratio$\frac{g(s_n|x_n)}{g(s^*|x_n)}$, up to a proportionality constant independent of x_n.*

Proof. By Bayes' rule: $R(x_n|s_{1:n}) = \frac{p(x_n|\, s_{1:n})}{p(x_n|\, s_{1:n-1}, s^*)} = \frac{p(s_n|s_{1:n-1}, x_n)\, \cdot p(x_n|\, s_{1:n-1})}{p(s_n\,|\, s_{1:n-1})} \cdot$
$(\frac{p(s^*\,|\, s_{1:n-1})\, \cdot p(x_n|\, s_{1:n-1})}{p(s^*|\, s_{1:n-1})})^{-1}$.

Due to the Markov property, $p(s_n|s_{1:n-1}, x_n) = g(s_n|x_n)$ and $p(s_n = s^*\,|s_{1:n-1}, x_n) = g(s^*|x_n)$. Plugging this in, we obtain:

$$R(x_n\,|s_{1:n}) = \frac{g(s_n\,|\,x_n)\cdot p(x_n|s_{1:n-1})}{p(s_n|s_{1:n-1})}\cdot(\frac{g(s^*|x_n)\cdot p(x_n\,|\,s_{1:n-1})}{p(s^*|s_{1:n-1})})^{-1} = \frac{g(s_n|x_n)}{g(s^*\,|x_n)}\cdot\frac{p(s^*|s_{1:n-1})}{p(s_n|s_{1:n-1})}$$

The latter term $\frac{p(s^*|s_{1:n-1})}{p(s_n|s_{1:n-1})}$ is constant with respect to x_n, as desired.

As we have assumed that g is a normal density, the likelihood ratio simplifies to:

$$R(x_n|s_{1:n}) \propto exp(-\frac{(x_n - s_n)^2}{2{\sigma_\epsilon}^2} + \frac{(x_n - s^*\;\;)^2}{2{\sigma_\epsilon}^2}) = exp(\frac{(s_n - s^*)x_n}{{\sigma_\epsilon}^2})$$

Hence, if the observed signal s_n is greater than s^*(a positive news), the forecaster puts an exponentially heavier weight on larger values of x_n, and for negative news, he overweights smaller values of x_n, which is in line with over-reaction to most recent news.

With the particle filter, we get the exponential reweighting by multiplying to the original weights $w({x_n}^i) = \frac{g(s_n|{x_n}^i)\, f({x_n}^i|{x_{n-1}}^i)}{q(x_n|x_{n-1}={x^i}_{n-1}, s_{1:n})}$ with the extra exponential factor $exp(\frac{(s_n - s^*)x_n}{{\sigma_\epsilon}^2})$. As with the basic particle filter algorithm discussed above, we need to specify our proposal density q to target regions of high density. We would like to target $\tilde{q} \propto exp(\frac{(s_n - s^*)x_n}{{\sigma_\epsilon}^2})p(x_n|x_{n-1,}s_n)$, which we estimate by first assuming the normal model. Given that $x_n|x_{n-1}, s_n \sim N(\bar{\mu}, \bar{\Sigma})$ in the normal model, the diagnostic expectations is given by a shift of the posterior density by $\frac{\theta\cdot\bar{\Sigma}\cdot(s_n - s^*)}{{\sigma_\epsilon}^2}$. Thus we set the location and scale parameter of our proposals as $\mu_{diag} = \bar{\mu} + \frac{\theta\cdot\bar{\Sigma}(s_n - s^*)}{{\sigma_\epsilon}^2}$, $\Sigma_{diag} = \bar{\Sigma}$, where $\bar{\mu}, \bar{\Sigma}$ are the location and scale parameters for our original proposal. As before, we have $df_q = \frac{df\, +\, 2}{2}$ to ensure that our proposal has heavier tails than the target distribution. To summarize, the algorithm is as follows:

1. From the original particle filter, estimate $s^* = \rho\mu_{n-1}$, with μ_{n-1} our predictive mean $E[x_{n-1}|\, s_{1:n-1}]$, estimated by the mean of our particles $x^i{}_{n-1}$.

2. Propose according to a t-distribution with location parameter $\mu_{diag} = \bar{\mu} + \frac{\theta \cdot \bar{\Sigma}(s_n - s^*)}{\sigma_\epsilon{}^2}$, $\Sigma_{diag} = \bar{\Sigma}$, $df_q = \frac{df + 2}{2}$.

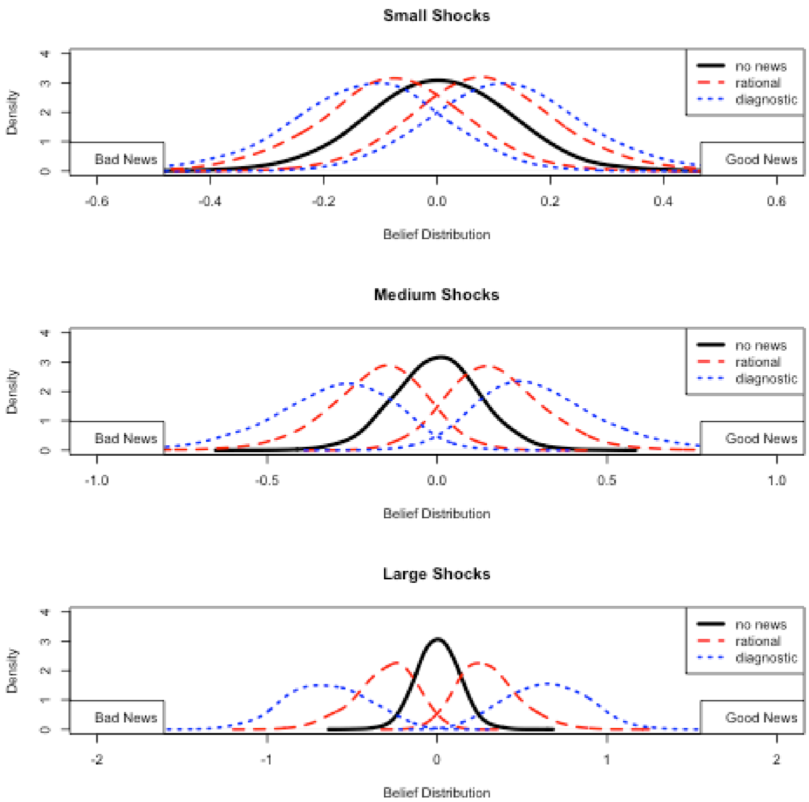
3. For each proposal, resample with weights$w_{diag}(x_n|x_{n-1,}s_n) =$ $\frac{g(s_n|x_n{}^i)\, f(x_n{}^i|x_{n-1}{}^i)}{q(x_n|x_{n-1}=x^i{}_{n-1}, s_{1:n})} exp(\frac{(s_n - s^*)x_n}{\sigma_\epsilon{}^2})$

C.3 Results

In the simulations below, we set $\rho = 0.9, \sigma_u = 0.2, \sigma_\epsilon = 0.2$, and $0 \leq \theta \leq 1.5$. We find that the basic qualitative characteristics of diagnostic expectations are robust to heavy tails. As Figure A1 shows, the diagnostic expectations overreacts to news, relative to the rational benchmark.

Figure A1. Response to News under Rational and Diagnostic Expectations

This plot shows the belief distribution in reponse to news. The black line plots the distribution with no news. The dashed red line plots the distribution in response to news with rational expectations. The dotted blue line plots the distribution in reponse to news with diagnostic expectations.

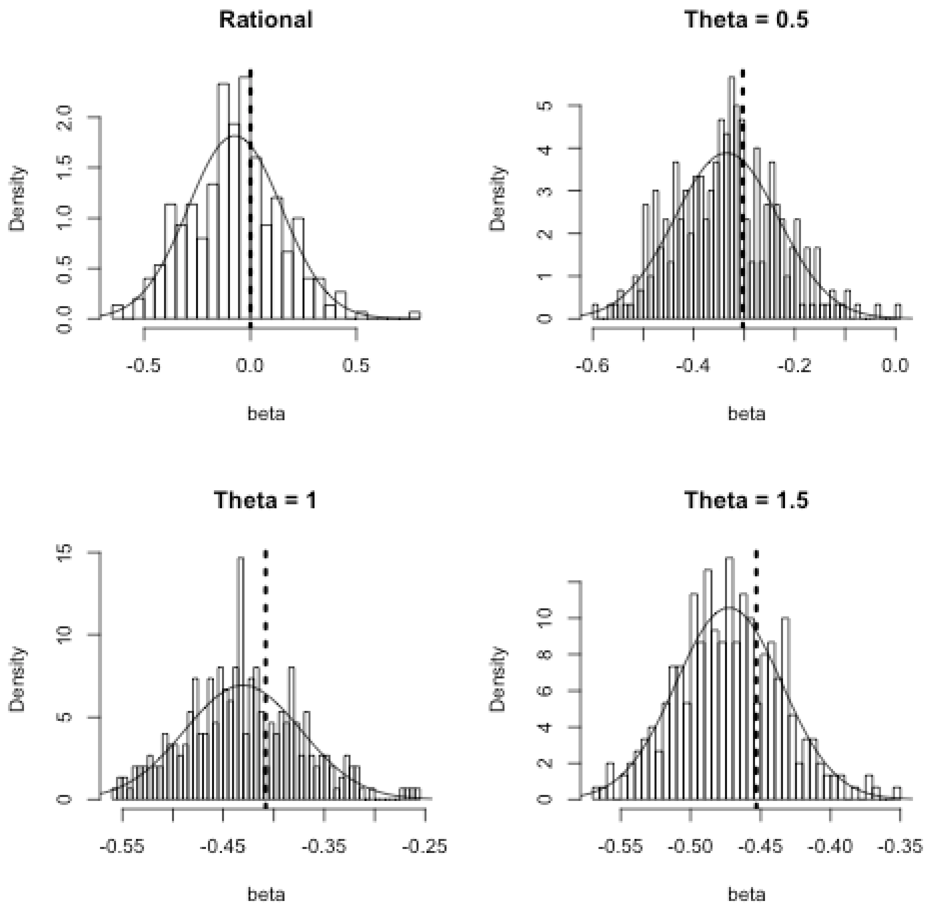


We then check the results of the CG forecast error on forecast revision regressions. Figure A2 shows the distribution of bootstrapped regression coefficients. Panel A first checks the case with normal shocks, the particle filter simulation agrees with the predicted coefficients $-\frac{\theta(1+\theta)}{(1+\theta)^2+\theta^2\rho^2}$ using the Kalman filter. Panel B then shows the case where the shocks are heavy-tailed. We see that the coefficients for the heavy-tailed shocks are more negative compared to the predicted values.

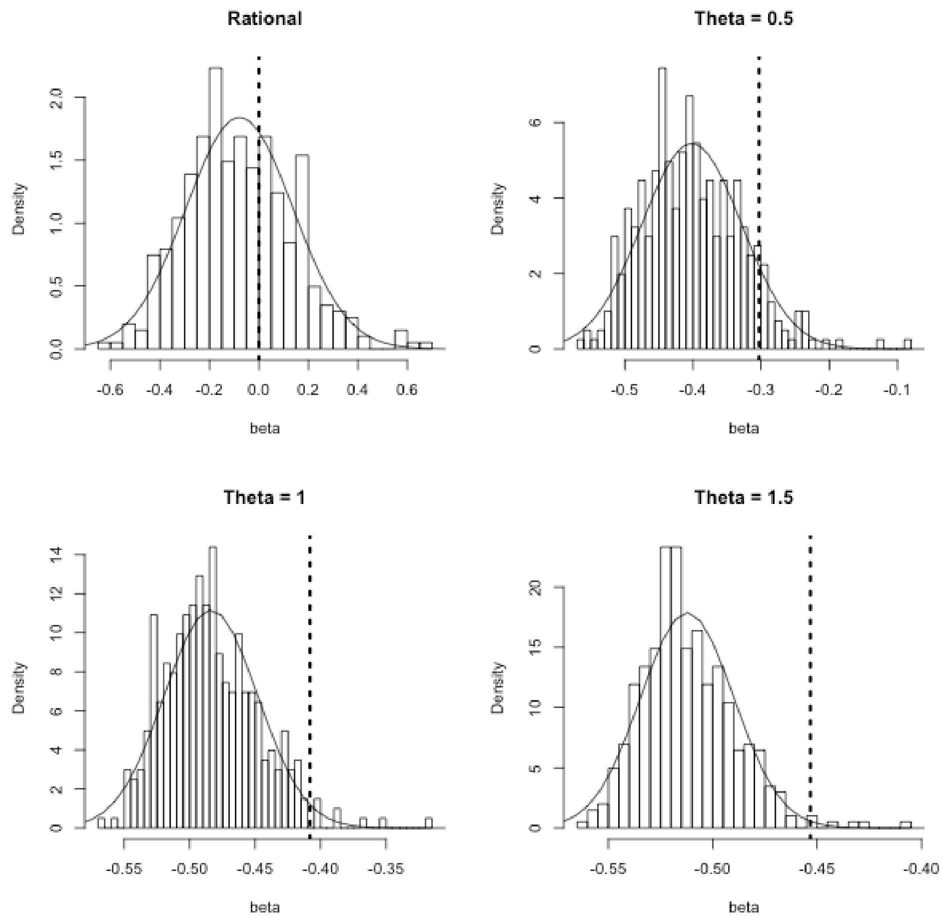
Figure A2. Individual CG Coefficients with Normal and Fat-Tailed Shocks

This plot shows the distribution of coefficients from individual level (pooled panel) CG regressions. Panel A analyzes the case for normal shocks and Panel B analyzes the case for fat-tailed shocks, both using the particle filter. Each simulation has 80 time periods and each plot shows the coefficients from 300 simulations. The dashed vertical line indicates $-\frac{\theta(1+\theta)}{(1+\theta)^2+\theta^2\rho^2}$, which is the coefficient predicted by normal shocks and Kalman filtering.

Panel A. Normal Shocks



Panel B. Heavy-Tailed Shocks, df = 2.5



for the normal case. Specifically, as the rational posterior exhibits heavier tail, the exponential reweighting of the diagnostic expectation results in greater mass located on the extreme values of the exponential weight, and hence greater shift in the diagnostic expectation. This effect is only present for diagnostic expectations — for rational expectations i.e. $\theta = 0$, we do not observe a divergence between normal and fat-tailed distributions.

Finally, Figure A3 replicates the results for the contrast between regressions using individual and consensus forecasts. The general qualitative result is that there is much less overreaction in consensus opinion, or even underreaction for some cases. Underreaction occurs when the noise ${\sigma_\epsilon}^2$is sufficiently high and individual overreaction parameter θ is sufficiently low. Figure A3 plots the case where $\sigma_\epsilon = 1, \theta = 0.1$, which shows robustly positive consensus regression coefficients for 20 forecasters.

Figure A3. Individual vs. Consensus Diagnostic Expectations

This plot shows the distribution of coefficients from individual level (pooled panel) and consensus CG regressions, using fat-tailed shocks and particle filtering. The left panel shows the coefficients from pooled individual level regressions, and the right panel shows the coefficients from consensus regressions. Each draw has 40 forecasters and 80 time periods; there are 300 draws.