Over-reaction in Macroeconomic Expectations

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

December 2017, Revised June 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 examines predictability of forecast errors from forecast revisions. We report two key findings: forecasters typically over-react to their individual news, while consensus forecasts under-react to average forecaster news. 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] Oxford Said Business School, Università Bocconi, Harvard University, and Harvard University. We thank Xavier Gabaix, Yuriy Gorodnichenko, Luigi Guiso, David Laibson, Jesse Shapiro, Paolo Surico, participants at the 2018 AEA meeting and seminar participants at EIEF, Ecole Politechnique, Harvard, and LBS for helpful comments. 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 (GA 647782). We thank Johan Cassell, Francesca Miserocchi, Johnny Tang, and especially Spencer Kwon and Weijie Zhang for 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. A growing body of work tests the Rational Expectations Hypothesis using survey data on the anticipations of households and professional forecasters. The evidence points to systematic departures from statistical optimality, which take the form of predictable forecast errors. Such departures have been documented in the cases of forecasting inflation and other macro variables (Coibion and Gorodnichenko 2012, 2015, henceforth CG, 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 (La Porta 1996, Bordalo, Gennaioli, La Porta and Shleifer 2017, henceforth BGLS), 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 (Hommes et al. 2004, Beshears et al. 2013, Frydman and Nave 2017, Landier, Ma, and Thesmar 2017).

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 and processing information entails significant material and cognitive costs. To economize on these costs, agents revise their expectations sporadically, or on the basis of selective news. As a consequence, expectations and decisions under-react to news relative to the case of unlimited information capacity. In a novel 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, Campbell and Shiller 1987, 1988). The leading psychological mechanism for 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 objectively rare. This approach is not inconsistent with limited information processing, but stresses that people infer too much from the information they attend to, however limited, so that 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, under what circumstances is each 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 of 16 macroeconomic variables, which include and expand 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 (four variables appear in both surveys), including forecasts of real economic activity, consumption, investment, unemployment, housing starts, government expenditures, as well as multiple interest rates. We examine both consensus and individual level forecasts. SPF data are publicly available; Blue Chip data were purchased and hand-coded for the earlier part of the sample.

Section 3 describes the patterns of over- and under-reaction in different series. We follow CG's methodology of measuring a forecaster's news by their forecast revision, and of using this forecast revision to predict the forecast error, computed as the difference between the realization and the 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 also individual ones. The consequences of aggregating forecasts turn out to be crucial for understanding their properties.

For the case of consensus forecasts, we confirm the CG findings of under-reaction: the average forecast revision positively predicts the average future forecast error 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 to the consensus results, at the level of the individual forecaster over-reaction is the norm, under-reaction the exception. These results are robust to several potential sources of predictability, including forecaster heterogeneity, small sample bias, measurement error, nonstandard loss functions, and non-normality of shocks.

In Section 4 we propose a model that reconciles these seemingly contradictory findings. 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. In 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.[2] We then allow for over-reaction by assuming that, in processing the noisy signal, agents are swayed by the representativeness heuristic.

To formalize this heuristic we use the Gennaioli and Shleifer (2010) model, originally proposed to describe lab experiments on probabilistic judgments but 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 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 overweighs recent news. As in earlier work, we call expectations distorted by representativeness "diagnostic."

[2] This setup can also capture a setting in which different forecasters observe different news (stemming for instance from their use of different models or different information sources, CG 2012).

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 a negative correlation between his forecast error and his earlier forecast revision. At the consensus level, however, the econometrician may still detect a positive correlation between the forecast error and the consensus revision provided the distortion caused by representativeness is not too strong. The reason is that, while over-reacting to their own signal, individual forecasters do not react to the signals observed by others. 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 is misleading. Even if all forecasters over-react, as under diagnostic expectations, consensus forecasts may point to under-reaction simply because different analysts over-react in different directions to partial information. In Section 5, then, we assess whether individual forecasts are consistent with a key prediction of diagnostic expectations, the "kernel of truth" property, which is the idea that expectations exaggerate true patterns in the data. This property yields testable predictions both across different series and in the time series of individual variables. These predictions help distinguish our model from mechanical extrapolation (and possibly over-reaction), such as adaptive expectations.

We present cross sectional tests in Section 5.1. We show first that, upon receiving news, individual forecast revisions are stronger for more persistent variables. This finding is consistent with diagnostic expectations, but not with adaptive expectations, where the updating rule is fixed. We then show that the individual-level CG coefficients of over-reaction documented in Section 3 are close to zero for series that are very persistent. This finding is once again in line with diagnostic expectations: as persistence increases, rational forecast revisions are more volatile, reducing the scope for over-reaction.

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 that the kernel of truth property implies 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. Intuitively, diagnostic expectations exaggerate both short-term momentum and long-term reversals. We find that these predictions are borne out in the data. 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 estimate the structural parameters of our baseline model using the simulated method of moments. We find the diagnostic parameter θ is significantly positive for 17 out of 20 series, varying between 0.2 and 1.5, and broadly consistent with estimates we obtained in other work using different methods and in different contexts (BGS 2017, BGLS 2018). We estimate a small but significantly negative θ for one series, unemployment. These results suggest that over-reaction is sizable: the predictable component of the forecast error is commensurate with the size of the rational response to news.

This paper documents the prevalence of over-reaction to news in individual macroeconomic forecasts and reconciles this finding with under-reaction in the consensus using a model of diagnostic expectations. There have been other approaches to similar phenomena. One is adaptive expectations; we show that the diagnostic expectations model has better psychological foundations and fits the data better. Another approach is Natural Expectations (Fuster, Laibson, and Mendel, 2010). In this model, forecast errors arise because agents fit a simple AR(1) model through a series that may have longer lags. Diagnostic expectations share some predictions with natural expectations, but also make distinctive predictions, such as over-reaction to long-term reversals, that more closely reflect the data.[3]

In general, and beyond Natural Expectations, predictable forecast errors may reflect model mis-specification, and not just over-reaction to information. Even econometricians find it difficult to find the best specification of many macroeconomic series. Of course, model misspecification is consistent with

[3] A large literature considers how incentives may distort professional forecasters' stated expectations. Ottaviani and Sorensen (2006) point out that if forecasters compete in an accuracy contest with particular rules (winner-take-all), they overweigh private information. In contrast, Fuhrer (2017) argues that in the context of SPF, individual forecast revisions can be negatively predicted from past deviations relative to consensus. Kohlhas and Walther (2018) model apparent under- and over-reaction to information as a byproduct of an asymmetric loss function. We return to these ideas in our robustness tests in Section 3.2, and in our tests of the kernel of truth in Section 5.

forecasters over-reacting to data pointing to representative outcomes of their models. Furthermore, the evidence in support of the kernel of truth suggests that forecasters pay attention to key features of reality and exaggerate them in their forecasts. Learning may help explain persistence of errors, but these errors may well be due to over-reaction to news.

Diagnostic expectations are also related to the idea of overconfidence, in the sense of overprecision of beliefs, which implies an exaggerated reaction to private signals (Daniel, Hirshleifer, and Subrahmanyam 1998, Moore and Healy 2008). Overconfidence has been used to explain excess volatility in prices of both asset and goods (Barber and Odean 2001, Benigno and Kourantasias 2018). In independent work, Broer and Kohlhas (2018) explore the role of overconfidence in driving individual level over-reaction in forecasts for GDP and inflation. We later return to the difference between overconfidence and our model. At the same time, we stress that diagnostic expectations describe beliefs and over-reaction in a wide range of settings, both in the lab and in the field, including those where overconfidence can be ruled out (such as in cases where information is common and public). We think that developing portable models that are applicable in very different domains is a key step in identifying robust departures from rationality.

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).[4] 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

[4] 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.

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, we use Blue Chip forecasts from the end-of-quarter month survey (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.[5]

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

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
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

We use an annual forecast horizon. For 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, consistent with the forecasters' information sets. Blue Chip reports forecasts of quarterly growth rates, so we add up these forecasts 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 B 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.[6]

[6] 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 our results, which are robust to considering only forecasters who have both forecasts and forecast revisions.

Data on Actual Outcomes. The values 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.[7] For a given quarter, we proxy the forecasters' information set as the latest estimates available by the time of the forecast. 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 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,[8] and the quarterly share of forecasters with positive revisions.

Table 2. Summary Statistics

Mean and standard deviation of main variables. All values are in percentages. Panel A shows the statistics for 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	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

[7] This choice is due to the fact that, when forecasters make forecasts in quarter t, only initial releases of macro variables in quarter $t-1$ are available.
[8] We categorize a forecaster as making no revision if he 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 no-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.

Real GDP (BC)	quarter *t*-1 to end of quarter t+3	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
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. Macro analysts do not seem to have asymmetric loss functions that systematically bias their forecasts in a given direction. Second, 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. 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 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 that the geographical location of forecasters influences their predictions of 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.

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

Many tests of the rational expectations hypothesis assess whether forecast errors can be predicted using information available at the time the forecast is 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.

CG (2015) address this problem with forecast revisions. Denote by $x_{t+h|t}$ the h-periods ahead forecast made at time t about the future value x_{t+h} of a variable. Denote by $x_{t+h|t-1}$ the forecast of the same variable in the previous period. The h-periods ahead forecast revision at t is given by $FR_{t,h} = \left(x_{t+h|t} - x_{t+h|t-1}\right)$, or the one period change in the forecast about x_{t+h}. This revision captures the information that the forecasters have observed and used to update their forecast. The extent to which forecasters under- or over-react to information can then be assessed by estimating the regression:

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

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 the forecast under-reacts to information, we expect $\beta_1 > 0$. To see why, 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 $\mathbb{E}_t(x_{t+h} - x_{t+h|t}) > 0$. The converse holds if negative information is received: the downward revision is insufficient, predicting a negative error. Under-reaction implies that the forecast error should be positively correlated with the forecast revision.

By the same logic, when the forecast over-reacts to information we should expect $\beta_1 < 0$. Indeed, over-reaction means that after positive information $FR_{t,h} > 0$ the forecast is too optimistic, so the forecast error is negative $\mathbb{E}_t(x_{t+h} - x_{t+h|t}) < 0$. On the other hand, after negative information $FR_{t,h} < 0$ it is too pessimistic, so the error is positive $\mathbb{E}_t(x_{t+h} - x_{t+h|t}) > 0$. 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. The consensus forecast $x_{t+h|t}$ is 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. Similarly, $FR_{t,h}$ is the h-periods ahead "consensus information" or forecast revision. CG estimate (1) for the GDP price deflator (PGDP_SPF) at a horizon $h = 3$ and find $\beta_1 = 1.2$, 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 $h = 3$,[9] 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 (2) 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 (and our point estimate for inflation

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

forecasts coincides with CG's). While results for the other SPF series are not directly comparable (since CG 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 over-reaction 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 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. Over-reaction of individual forecasters may thus be attenuated by heterogeneity and aggregation.

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	-025	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:

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

Superscript p on the coefficients recognizes that we are pooling individual level data. The logic of the test, however, does not change: $\beta^p_1 > 0$ indicates that the average analyst under-reacts to his own information, while $\beta^p_1 < 0$ indicates that the average analyst over-reacts.[10]

Columns (4) through (6) of Table 3 report the results of estimating Equation (2). Surprisingly, the picture is essentially reversed from the consensus: 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 (13 out of 16 variables), 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

[10] 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.

investment display neither over-nor under-reaction (β_1^p 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. The overall message of Table 3 is clear: at the level of the individual forecaster, over-reaction is the norm.

In sum, a fascinating picture emerges from these tests. At the consensus level, expectations typically under-react. At the individual level, they typically over-react. We conclude this section with a number of robustness checks. In Section 4, we present a model capable of reconciling these patterns.

3.1 Robustness Checks

Predictability of forecast errors might arise from features of the data unrelated to individuals' under- or over-reaction to news. We next show that our results are robust to many such confounds.

Small Samples. Our individual level estimates may be contaminated by a small sample problem. Finite-sample Stambaugh bias exists in time series regressions (Kendall 1954, Stambaugh 1999) and panel regressions with fixed effects (Nickell, 1981). The fact that baseline individual-level regressions in Table 3 do not have fixed effects assuages the concern (Hjalmarsson 2008). Adding fixed effects does not change the results much, indicating that the bias, even if present, 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).

Heterogeneity among Forecasters. Coibion and Gorodnichenko (2015) point out that heterogeneity among 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. To assess this problem, we perform forecaster level regressions, focusing on forecasters with at least 10 observations. Table C1 in Appendix C compares the median coefficient from forecaster level regressions to the coefficients from pooled individual level regressions from Table 3. The coefficients are very similar, suggesting that the observed over-reaction describes the median forecaster. On average across series, we estimate a negative β_1^p for two thirds of the forecasters. In some series, nearly every forecaster over-reacts while in other series the distribution of β_1^ps is more balanced. Finding the sources of forecaster heterogeneity is an open problem for future work.

Asymmetric Loss Functions. Another concern with our findings is that forecast errors reflect not cognitive limitations but analysts' biased incentives. Of course, an analyst's objective is difficult to observe. Here we discuss the implications of several analyst loss functions proposed in the literature.

With an asymmetric loss function (Capistran and Timmerman 2009), the over-reaction pattern in Table 3 may be generated by a combination of: i) a higher cost of over- than under-predicting, and ii) suitably time varying volatility (Pesaran and Weale 2006). However, an asymmetric loss function, would generate an average forecast error, namely average pessimism, under i) and ii). In the data, forecasts are not systematically upward or downward biased on average. The consensus forecast errors are both very small and insignificant (Table 2, panel A). This is also true for individual forecast errors: we fail to reject that the average error is different from zero for about 60% of forecasters for the macroeconomic variables.[11]

Another source of bias in reported expectations is that individuals may follow consensus forecasts (Morris and Shin 2002, Fuhrer 2017). 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 rational forecast and $\tilde{x}_{t+h|t}$ is the average contemporaneous forecast with 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, in line with intuition, following consensus forecasts leads to individual level under-reaction, contrary to our findings.[12]

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 $x^i_{t+h|t+j}$. To assess the relevance of this mechanism in the data, note that in forecast revisions for the current quarter ($h = 0$), forecast smoothing reduces the current revisions, creating under-reaction. This prediction is contradicted by the data: negative predictability prevails even at this horizon (Table A2).

Finally, analysts may appear to over-react to news because the series they predict exhibit fat tailed noise. For instance, after observing an anomalously large signal, analysts may disproportionally attribute it to fundamentals rather than noise, becoming too optimistic. In our data, both fundamentals and forecast revisions have high kurtosis. To assess the relevance of this mechanism, in Appendix E we

[11] No clear pattern emerges among the forecasters whose average error is significantly different from zero, consistent 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 the secular decline in rates over the time period we examine.

[12] 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 $cov\left(\widetilde{FE}^i_{t+h,t}, \widetilde{FR}^i_{t+h,t}\right) > 0$ follows from $cov\left(FE^i_{t+h,t}, FR^i_{t+h,t}\right) = 0$ and $cov\left(FE_{t+h|t}, FR_{t+h|t}\right) > 0$ under noisy rational expectations, together with $cov\left(FE^i_{t+h,t}, FR_{t+h|t}\right), cov\left(FE_{t+h|t}, FR^i_{t+h,t}\right) > 0$.

consider a learning setting with fat tailed shocks. Here the Kalman filter is no longer optimal, and the particle filter must be used. We find that when forecasts are produced using the particle filter, individual forecast errors are uncorrelated with forecast revisions, even with fat-tailed shocks. We also show that individual level over-reaction can be explained using a modified particle filter that, just as in the model in Section 4, overweighs the likelihood of representative future outcomes. Because fat tails do not appear to affect our qualitative results, we maintain the more tractable assumption of normality.

4. Diagnostic Expectations

We present a model that reconciles under-reaction of consensus expectations with over-reaction of individual level expectations. At each time t, the target of forecasts is a state 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 noise, 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. This AR(1) setting, also considered by CG (2015), yields convenient closed form predictions. In Section 6 we examine the AR(2) case.

This setup accommodates several interpretations. In CG (2015), unobservability of x_t stems from rational inattention (Sims 2003, Woodford 2003). Forecasters could in principle observe x_t but doing so is too costly, so 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. When we discuss this model, we sometimes refer to it as Rational Inattention.[13]

[13] As CG show, the same predictions are obtained if rational inattention is modelled à la Mankiw and Reis (2002), where agents observe the same information but only sporadically revise their predictions.

Another interpretation of Equations (3) and (4) is that forecasters have heterogeneous information about the future value of a macroeconomic variable. Here the forecaster specific shock ϵ_t^i captures the fact that different forecasters use different models or pieces of evidence. For professional forecasters, the latter interpretation is perhaps more compelling, since their job is to be attentive to, and to predict, the variables in question. Under both interpretations, forecasters form their revisions through rational updating on the basis of noisy signals. We thus refer to both interpretations as "Noisy Rational Expectations".

A Bayesian, or rational, forecaster enters period t carrying from the previous period beliefs about the current state x_t summarized by a probability density $f\left(x_t|S_{t-1}^i\right)$, 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:

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

Equation (5) iteratively defines the forecaster's beliefs. Given normal shocks, $f\left(x_t|S_t^i\right)$ is described by the Kalman filter. A rational forecaster estimates the current state at $x_{t|t}^i = \int xf\left(x|S_t^i\right)dx$ and forecasts future values using the AR(1) structure, so $x_{t+h|t}^i = \rho^h x_{t|t}^i$.

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), who apply Diagnostic Expectations to a (diagnostic) Kalman Filter, we define the representativeness of a state x at t as the likelihood ratio:

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

State x 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 representative states and deflates the probability of unrepresentative ones. 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 effect on probability assessments, has been shown to unify well-known laboratory 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 real world phenomena such as stereotyping (BCGS 2016), self-confidence (BCGS 2017), and expectation formation in financial markets (BGS 2018, BGLS 2017). Here we assess whether this same structure can shed light on errors in forecasting macroeconomic variables.

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 $\frac{\Sigma\sigma_\epsilon^2}{\Sigma+\sigma_\epsilon^2}$ and by a time varying mean $x_{t|t}^{i,\theta}$ where:*

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

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

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 Σ/σ_ϵ^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). 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. Equation (8) also highlights that diagnostic expectations do not create an average bias, but do create excess volatility. Indeed, 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, which gives rise to excess volatility.

In contrast to traditional departures from rationality such as adaptive expectations, diagnostic expectations are forward-looking in that they depend on the 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 tail, generating excessive optimism. As we show in Sections 5 and 6, the kernel of truth property offers testable predictions on how updating and forecast errors should change as the process becomes more persistent or when it is influenced by longer AR(2) lags. Critically, these predictions can be tested against conventional mechanical models of extrapolation such as adaptive expectations.

Equation (8) is reminiscent of overconfidence. Overconfidence is often modeled as inflating the signal to noise ratio of private information, which also causes over-reaction but of a different form from Equation (8).[14] As we previously noted, diagnostic expectations have proven capable of explaining over-reaction to news even in settings in which private information, and hence overconfidence, is not at play. As we show in Corollary 1 below, this property offers implications that may be used to distinguish these

[14] In particular, under overconfidence the Kalman gain K is inflated to $K_o > K$ and the current overconfident estimate $x_{t|t}^{i,o}$ updates on the past overconfident estimate and not on the past rational estimate as in Equation (8). Formally, $x_{t|t}^{i,o}$ is iteratively defined by $x_{t|t}^{i,o} = x_{t|t-1}^{i,o} + K_o\left(s_t^i - x_{t|t-1}^{i,o}\right)$.

models. We stress that our goal is not to tease these theories apart, but rather to study whether over-reaction to information can account for predictability of errors in consensus and individual forecasts.

Consider next the implications of Diagnostic Expectations for forecasts and forecast errors. Define the consensus diagnostic forecast of x_{t+h} at time t as

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

so that the Diagnostic forecast error and revision are respectively given by $x_{t+h} - x^{\theta}_{t+h|t}$ and $x^{\theta}_{t+h|t} - x^{\theta}_{t+h|t-1}$. In Appendix A, 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:*

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

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

where $g(\sigma_{\epsilon}^2, \Sigma, \rho, \theta) > 0$ *is a function of parameters. Thus, 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 Filter reconciles positive consensus coefficients with negative individual level coefficients, consistent with the patterns in Section 3. Intuitively, over-reaction of individual analysts to their own information, implies a negative pooled coefficient $\beta_1^p < 0$. At the same time, analysts do not react at all to the information received by other analysts (which they do not observe). This effect can create under-reaction of consensus to average information, particularly if $\sigma_{\epsilon}^2/\Sigma$ is high. Indeed, if information is very noisy, not using the signals observed by other forecasters entails a large loss of information in the aggregate. The condition $\theta < \sigma_{\epsilon}^2/\Sigma$ is intuitive: it requires the diagnostic Kalman gain $(1+\theta)\frac{\Sigma}{\Sigma+\sigma_{\epsilon}^2}$ to be smaller than 1. This means that, as long as individual forecasters filter news to some extent, consensus forecasts exhibit under-reaction, even if each analyst discounts his information too little.

In contrast 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, $\beta_1^p < 0$. In that model, because forecasters optimally use the limited information at their disposal, their forecast error is uncorrelated with their own forecast revision. As is evident from Equations (9) and (10), when $\theta = 0$ there is no individual-level predictability, 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 parameters characterizing the data generating process $(\sigma_\epsilon^2, \Sigma, \rho, \theta)$. In particular, stronger persistence ρ reduces individual over-reaction, in the sense that it pushes the pooled coefficient β_1^p toward zero.

In many real world settings, forecasters observe not only private but also public signals. Assume then that each analyst observes, in addition to the private signal s_t^i, a public signal $s_t = x_t + v_t$, where v_t is also normal with variance σ_v^2. In this case, the diagnostic estimate uses both the private and the public signal according to their informativeness. We then obtain:

Corollary 1 *Suppose that $\theta \in (0, \sigma_\epsilon^2/\Sigma)$. Then, increasing the precision $1/\sigma_v^2$ of the public signal while holding constant the total precision $(1/\sigma_\epsilon^2 + 1/\sigma_v^2)$ of the private and the public signals: i) leaves the pooled coefficient β^p unchanged, and ii) lowers the consensus coefficient β.*

When a higher share of information comes from a public signal, the information of different forecasters is more correlated, so that individual forecasts incorporate a larger share of the information available to others. As a result, the consensus forecast exhibits less under-reaction, or possibly even over-reaction. This may explain why in financial market variables such as interest rates we detect less consensus under-reaction than in most other series: market prices act as public signals that correlate to a significant extent the information set of different forecasters. This result cannot be obtained from overconfidence, which features overweighting of private signals relative to public signals (Daniel et al. 1998), and thus predicts less over-reaction when more information is public.

Table 4 summarizes the predictions of three departures from rational expectations for the tests of Section 3. These include: Noisy Rational Expectations (or Rational Inattention), 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 features of the data generating process.

Table 4.

Model	Consensus	Individual	Updating
Noisy Rational	under-reaction	no predictability	depends on fundamentals
Diagnostic	consistent with under-reaction	over-reaction	depends on fundamentals
Mechanical / Adaptive	Undetermined	under-reaction for persistent series	does not depend on fundamentals

The sign switch between consensus and individual coefficient we documented for 9 out of 20 series (and 8 out of 16 variables) is consistent with diagnostic expectations but not with noisy rational expectations. 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 $\beta_1^p \approx 0$. 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, $\beta_1, \beta_1^p > 0$. This pattern may be accounted for by 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 first run a cross sectional test based on the persistence of the different series, which allows us to check whether analysts are backward looking as with Adaptive Expectations or forward looking as in Diagnostic Expectations. We then turn to the auto-correlation structure of the time series and assess whether, for series that feature hump-shaped dynamics, expectations over-react both to short-term persistence but also to longer-term reversals.

5.1 Persistence Tests

Under Noisy Rational and Diagnostic Expectations forecast revision at t satisfies:

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

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:

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

where μ is a positive constant independent of ρ.[15]

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,[16] and we use the exact time period for which the forecasts are available.[17] We estimate the following individual level regression:

$$x^i_{t+3|t} - x^i_{t+3|t-1} = \gamma^p_o + \gamma^p_1\left(x^i_{t+2|t} - x^i_{t+2|t-1}\right) + \epsilon^i_{t+3}$$

We estimate the same regression at the consensus level, which yields coefficients estimates γ_0 and γ_1. By averaging this equation, it is easy to see that consensus forecasts should satisfy the same condition. We then regress the slope coefficients γ^p_1 and γ_1 on the estimated persistence $\hat{\rho}$ of each series.

The results are reported in Figure 1 Panel A. At both the individual and the consensus level, the more persistent series display larger forecast revisions. While we only have 20 series, the correlation is

[15] This formula is based on the Error-Learning model, a generalization of adaptive expectations for longer horizons (Pesaran and Weale 2006). This model postulates $x^i_{t+s|t} - x^i_{t+s|t-1} = \mu_s\left(x_t - x^i_{t|t-1}\right)$, so that $\mu = \mu_h/\mu_{h-1}$.

[16] 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.

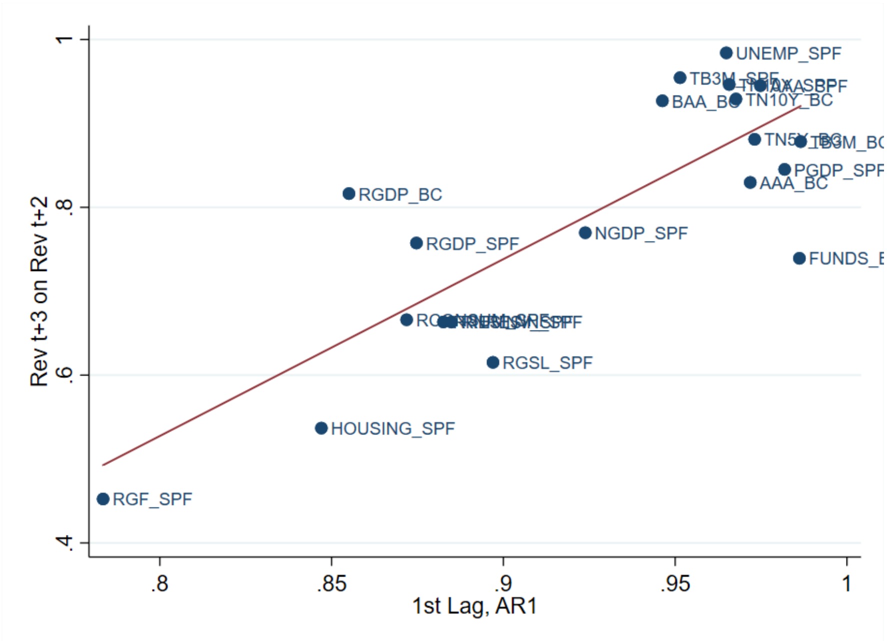
[17] Thus the properties of the actuals can be slightly different for the same variable from SPF and BlueChip (e.g. real GDP growth in SPF and Blue Chip), as these two datasets generally span different time periods.

statistically different from zero with a p-value less than 0.001.[18] In line with forward-looking models, forecasters update more aggressively for more persistent series. This evidence is inconsistent with adaptive expectations, where forecasters update mechanically, without taking persistence into account. 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, shown in Figure 1 Panel B.

Figure 1. Properties of Forecast Revisions and Actuals

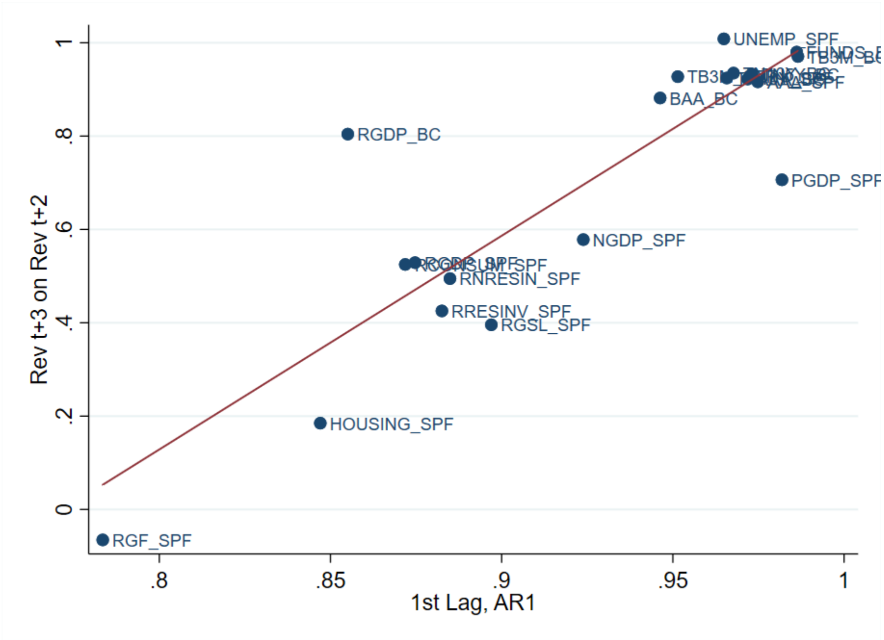
In Panel A, the y-axis is the coefficient γ_1^p from regression $x^i_{t+3|t} - x^i_{t+3|t-1} = \gamma_o^p + \gamma_1^p\left(x^i_{t+2|t} - x^i_{t+2|t-1}\right) + \epsilon^i_{t+3}$. The x-axis is the persistence measured from an AR(1) regression of the actuals corresponding to the forecasts. For each variable, 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



Panel B. Consensus Coefficients

[18] The results in Figure 1 and 2 also obtain 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).

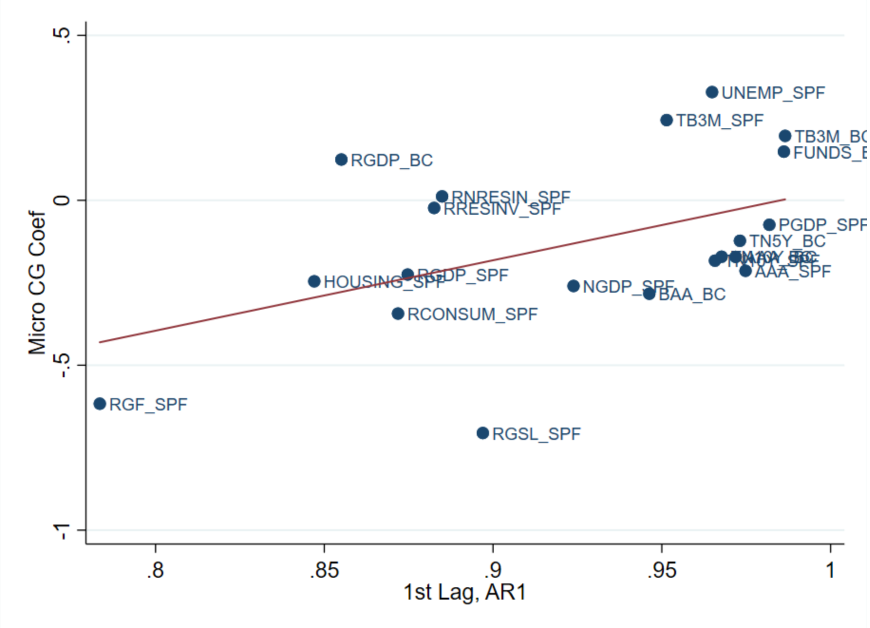


Another strategy is to assess the correlation between the persistence of a series and the CG coefficient of reaction to news. Diagnostic Expectations do not have clear predictions at the consensus level: the coefficient $(\sigma_\epsilon^2 - \theta\Sigma)g(\sigma_\epsilon^2, \Sigma, \rho, \theta)$ in Equation (10) can be either decreasing or increasing in persistence ρ, depending on parameter values. On the other hand, Equation (11) says that the individual CG coefficient should increase, i.e. get closer to zero, as 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. Of course, under Noisy Rational Expectations individual coefficients should be zero, so they should be uncorrelated with the persistence of fundamentals.

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 allow the forecasted series to be 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. This also allows us to compare these approaches to a model of Natural Expectations proposed by Fuster, Laibson and Mendel (2010), in which agents forecast and AR(2) process "as if" it was AR(1) in changes.

5.2.1 Diagnostic Expectations with AR(2) Processes

Suppose that forecasters seek to forecast an AR(2) process:

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

If $\rho_2 > 0$ and $\rho_1 < 0$, the variable displays short-term momentum and a long-term reversal. Each forecaster now observes two signals, one about the current state $s_{t,t}^i = x_t + \epsilon_t^i$ and another about the past state $s_{t-1,t}^{i} = x_{t-1} + v_t^i$. The presence of two signals implies that the current forecast revisions for x_{t+1} and x_{t+2} are not perfectly collinear, which is necessary for out test.

The diagnostic forecasts about $t+1$ and $t+2$ overweigh each signal (this is proved in Appendix A), so that forecast revisions are excessive. The diagnostic forecast of x_{t+3} is then a linear combination of the forecasts of x_{t+2} and x_{t+1} with weights given by the autoregressive parameters ρ_1 and ρ_2:

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

This decomposition of a long term forecast into shorter term ones suggests a way to test for overreaction, generalizing Equation (2) to AR(2). To do so, simply predict forecast errors in the long term using forecast revisions about shorter term:

$$x_{t+3} - x_{t+3|t}^{i} = \delta_0^p + \delta_2^p FR_{t,t+2}^i + \delta_1^p FR_{t,t+1}^i + \epsilon_{t,t+h}, \tag{12}$$

where $FR_{t,t+1}^i$ and $FR_{t,t+2}^i$ stand for the surveyed forecast revisions at for $t+1$ and $t+2$, respectively. Under Diagnostic Expectations, estimates of (12) satisfy the following property.

Proposition 3. *Under the Diagnostic Kalman filter, the estimated coefficients $\hat{\delta}_1^p$ and $\hat{\delta}_2^p$ in Equation* (12) *are proportional to the negative of the AR(2) coefficients:*

$$\hat{\delta}_1^p \propto -\rho_1\theta, \tag{13}$$

$$\hat{\delta}_2^p \propto -\rho_2\theta. \tag{14}$$

Once again, under rational expectations ($\theta = 0$) individual forecast errors cannot be predicted from any forecast revisions. Diagnostic expectations instead imply that the coefficients should be non-zero, with flipped signs relative to the data generating process. This is due to the kernel of truth. Over-reaction to short term momentum, $\rho_2 > 0$, implies that upward forecast revisions about x_{t+2} lead to exaggerated optimism about x_{t+3} and thus negative forecast errors. This yields $\hat{\delta}_2^p < 0$. On the other hand, over-reaction to long-term reversal, $\rho_1 < 0$, implies that upward forecast revisions about x_{t+1} lead to exaggerated pessimism about x_{t+3} and thus positive forecast errors. This yields $\hat{\delta}_1^p > 0$.

Proposition 3 also 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 short- and medium-term forecasts x_{t+1} and x_{t+2}. The former creates excess pessimism, the latter excess optimism. If the first effect is strong, the test of Section 3 may detect excess pessimism after good news, creating a false impression of under-reaction.

Before moving to the data, we briefly discuss Natural Expectations, which have been proposed to account for expectations errors in AR(2) settings. Under Natural Expectations, forecasts are based on an AR(1) process in changes $(x_{t+1} - x_t) = \varphi(x_t - x_{t-1}) + v_{t+1}$ with fitted coefficient $\varphi = (\rho_1 - \rho_2 - 1)/2$. For stationary processes, Natural Expectations exaggerate the short run persistence of the series while dampening long-term reversals.[19] Under Natural Expectations, Equation (12) cannot be estimated, because in this model the two forecast revisions are perfectly collinear. It seems clear, though, that our model's prediction of over-reaction to long term reversals ($\hat{\delta}_1^p > 0$) is against the spirit of Natural Expectations.

In the remainder of the section, we test the predictions of Proposition 3.

5.2.2 AR(1) vs AR(2) Dynamics

As a first step, we assess which of our 16 variables is 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. We fit a quarterly AR(2) process for our 20 series. Figure 4 below plots the estimates for ρ_1 and ρ_2.[20] As before, the actuals have the same format as the

---

[19] 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 $0 < \varphi < |\rho_2|$.

[20] 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.

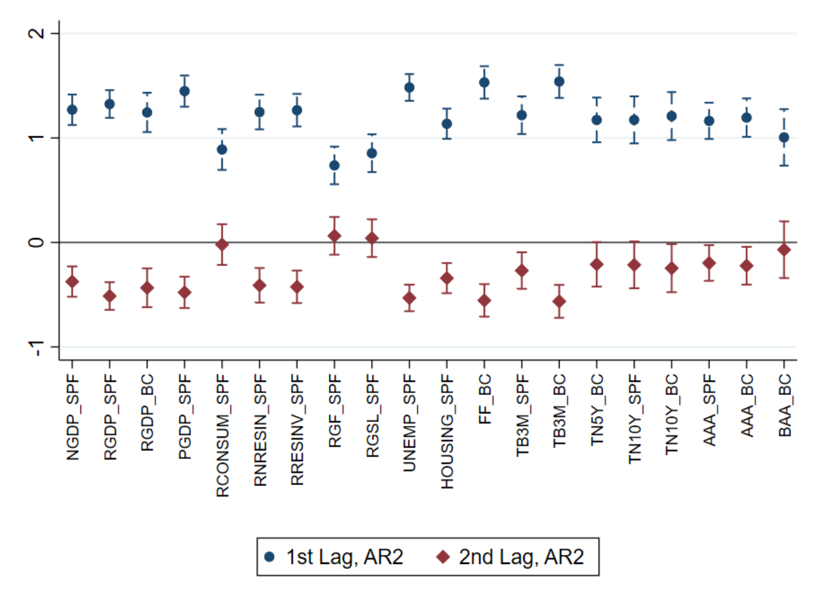
forecast variables, and for each series the regression covers the time period when the forecast data are available.

The signs of coefficients point to a positive momentum at short horizons ($\rho_2 > 0$) for all series, and to long-run reversals ($\rho_1 < 0$) for most series, the remaining ones having $\rho_1 \sim 0$.[21] To assess which dynamics better describe the series, we compare the AR(2) estimates to the AR(1) estimates from Section 5.1. Table 6 below shows the Bayesian Information Criterion (BIC) score associated with each fit.

For the majority of series, AR(2) is 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.[22] In sum, hump shaped dynamics are a key feature of several series.

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.



[21] 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.

[22] 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).

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}	ΔBIC_{2-1}	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 (12). Since our AR(2) series 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, $\hat{\delta}_1^p > 0$.

Figure 5 shows, for each relevant series, the forecast error regression coefficients $\hat{\delta}_2^p$ and $\hat{\delta}_1^p$ obtained from estimating Equation (12) with pooled individual data. Table 7 reports these 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 these 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.

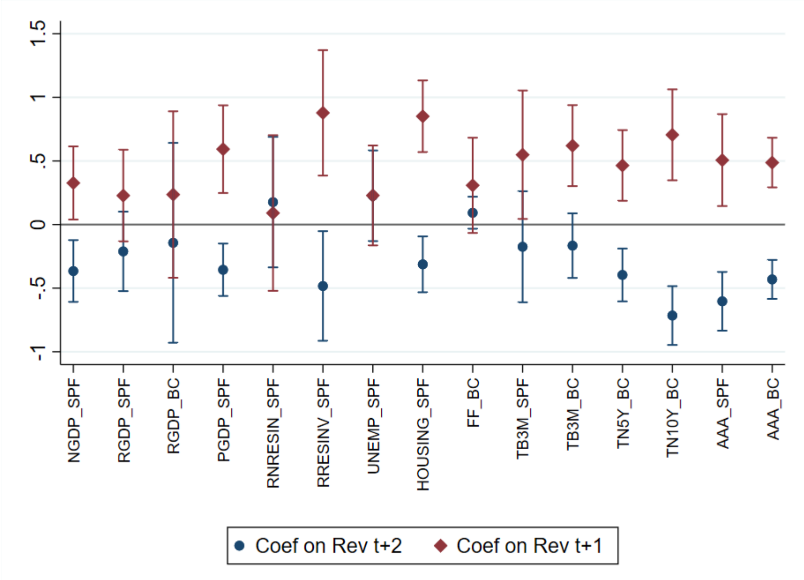


Table 7. Coefficients in CG Regression AR(2) Version

Coefficients $\delta_2^p$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

These results are consistent with kernel of truth but are 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).[23] Overall, then, the AR(2) analysis confirms and perhaps strengthens the evidence for over-reaction in the data. 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 in the more appropriate AR(2) case.

6. Model Estimation

We next use the simulated method of moments to estimate the parameters of the Diagnostic Expectations model. Appendix D provides supporting material as described below.

We assume forecasters describe each series k as either an AR(1) or an AR(2) process following Table 6, using the vector of estimated fundamental parameters $(\rho_{1,k}, \sigma_{u,k})$ or $(\rho_{1,k}, \rho_{2,k}, \sigma_{u,k})$

[23] Beshears et al. (2013) present a laboratory test of the hypothesis that forecasts neglect long term reversals in time series data. Their subjects recognize reversals when they occur within ten periods, but not when it occurs in fifty periods. Our results are consistent with their findings because reversals occurs within three quarters.

respectively (see Figure 4 and Appendix D, Table D1 for the estimates). Expectations depend not only on fundamentals but also on the measurement noise $\sigma_{\varepsilon,k}$ and the diagnostic parameter θ_k, which are latent parameters to estimate. For now, we take θ_k to be common to all forecasters but allow it to vary across series.

We estimate $\sigma_{\varepsilon,k}$ and θ_k by matching two key moments of the expectations data: the variance of the forecast errors, $\sigma^2_{FE,k} = var_{i,t}\left(FE^i_{k,t}\right)$, and the variance of forecast revisions, $\sigma^2_{FR,k} = var_{i,t}\left(FR^i_{k,t}\right)$, computed across time and forecasters. We choose these moments because they can be measured directly from the data with reasonable precision. We then use the Coibion Gorodnichencko coefficients – our key measure of over-reaction – to assess the performance of the estimated model.

Formally, for each series x^k_t of actuals and given parameters $(\theta_k, \sigma_{\varepsilon,k})$, we simulate time series of signals $s^{i,k}_t = x^k_t + \epsilon^{i,k}_t$ where $\epsilon^{i,k}_t$ is drawn from $\mathcal{N}\left(0, \sigma^2_{\varepsilon,k}\right)$ i.i.d. across time and forecasters. We mimic the structure of the survey data, generating measurements for the exact period that each forecaster follows a given series (we drop forecasters with fewer than ten observations). We then compute the expectations of each simulated forecaster iteratively, by inserting $s^{i,k}_t$ and θ_k in Equation (8) for the AR(1) processes and its generalization Equation (D1) for AR(2) processes. From the individual expectations, we recover individual forecast revisions and forecast errors. It is then straightforward to compute the model implied variances of forecast errors $\widehat{\sigma^2_{FE,k}}$ and of forecast revisions $\widehat{\sigma^2_{FR,k}}$. Finally, we define the loss function:

$$l\left(\theta_k, \sigma_{\varepsilon,k}\right) = \left(\sigma^2_{FE,k} - \widehat{\sigma^2_{FE,k}}(\theta, \sigma_\epsilon)\right)^2 + \left(\sigma^2_{FR,k} - \widehat{\sigma^2_{FR,k}}(\theta, \sigma_\epsilon)\right)^2$$

Our estimates are then defined by:

$$\left(\theta^*_k, \sigma^*_{\varepsilon,k}\right) = \underset{(\theta,\sigma_\epsilon)}{\operatorname{argmin}}\, l\left(\theta_k, \sigma_{\varepsilon,k}\right)$$

To obtain confidence intervals for our estimates, we repeat the process using 60 bootstrap samples (with replacement) from the panel of forecasters.

Table 8 summarizes the estimation results for series in both SPF and Blue Chip. For 17 out of the 20 series, we estimate a significantly positive θ, varying roughly between 0.2 and 1.5 (except for State & Local Government Consumption, which appears to be an outlier). For the Federal Funds rate and the 3M Treasury rate (BC), two closely related series, we estimate a θ of zero. Finally, for Unemployment we estimate a small but significant negative θ.

Overall, the simulated method of moments strengthens the finding of overreaction. Crucially, we can now quantitatively assess parameter estimates. Our estimates of θ exhibit tight confidence intervals, with an arithmetic average of 0.6. Estimates of standard deviation of noise σ_ϵ, normalized by the standard deviation of shocks σ_u, show more variation across series and are less precisely estimated.

Table 8. SMM Estimates of θ and σ_ϵ

This table shows the estimates of θ and σ_ϵ, as well as the 95% confidence interval based on block bootstrap (bootstrapping forecasters with replacement). The standard deviation of the noise σ_ϵ is normalized by the standard deviation of innovations in the actual process σ_u.

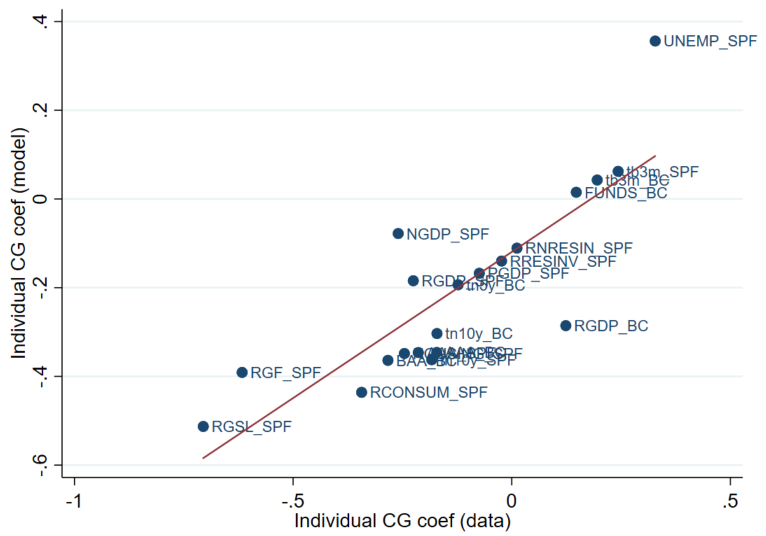
	θ	95% CI	σ_ϵ/σ_u	95% CI
Nominal GDP (SPF)	0.21	(0.06, 0.43)	0.45	(0.10, 1.08)
Real GDP (SPF)	0.51	(0.09, 0.87)	0.79	(0.34, 1.00)
Real GDP (BC)	0.34	(0.11, 0.58)	1.39	(0.58, 2.00)
GDP Price Index Inflation (SPF)	0.45	(0.12, 0.84)	3.18	(2.32, 4.00)
Real Consumption (SPF)	1.56	(0.95, 2)	3.56	(2.25, 4)
Real Non-Residential Investment (SPF)	0.35	(0.19, 0.57)	1.46	(1.03, 2.08)
Real Residential Investment (SPF)	0.28	(0.16, 0.45)	1.37	(0.82, 2.00)
Real Federal Government Consumption (SPF)	1.18	(0.8, 1.55)	1.66	(1, 2.4)
Real State & Local Govt Consumption (SPF)	2.80	(1.3, 3.9)	4.81	(3.74, 5)
Housing Start (SPF)	1.00	(0.54, 1.61)	1.81	(1.00, 3.36)
Unemployment (SPF)	-0.25	(-0.67, -0.08)	0.57	(0.01, 1.01)
Fed Funds Rate (BC)	-0.02	(-0.10, 0.06)	1.17	(0.77, 1.62)
3M Treasury Rate (SPF)	0.18	(0.11, 0.21)	1.11	(0.93, 1.43)
3M Treasury Rate (BC)	0.01	(-0.03, 0.09)	1.86	(1.44, 2.29)
5Y Treasury Rate (BC)	0.37	(0.32, 0.42)	2.19	(1.84, 2.61)
10Y Treasury Rate (SPF)	0.59	(0.5, 0.6)	2.91	(2.7, 3)
10Y Treasury Rate (BC)	0.55	(0.5, 0.6)	2.36	(1.95, 2.75)
AAA Corporate Bond Rate (SPF)	0.63	(0.50, 0.79)	4.60	(3.95, 5.21)
AAA Corporate Bond Rate (BC)	0.71	(0.60, 0.85)	4.85	(4.10, 5.60)
BAA Corporate Bond Rate (BC)	0.73	(0.64, 0.8)	2.63	(2.3, 3)

Interestingly, the estimates for θ are in line with the estimate from BGS (2018), who obtain $\theta = 0.9$ for expectations data on credit spreads, and with the estimate from BGLS (2018) who obtain $\theta = 1.2$ for expectations data on firm level earnings' growth. In the current exercise estimates are a bit lower, but we use a different method. In particular, allowing for AR(2) specifications, which we did not do in our previous work, reduces the estimates of θ (using AR(1) actuals we obtain an average θ of 0.8). To give a sense of the magnitude, a $\theta \approx 1$ means that forecasters' reaction to news is roughly twice as large as the rational expectations benchmark. In BGLS (2018), we find that this magnitude of θ can account for the observed 12% annual return spread between stocks analysts are pessimistic about and stocks they are optimistic about.

Having estimated $\theta_k, \sigma_{\epsilon,k}$ we can assess the performance of the model by running the standard CG regressions (Equation 2) using the simulated data. Consistent with Proposition 2, the series with highest estimated θ, such as State & Local Government Consumption, Federal Consumption and BAA, tend to exhibit over-reaction already at the consensus level. Figure 6 compares the predicted individual CG coefficients from the estimated model with those estimated from the survey data. The correlation between the two sets of coefficients is about 0.83 (p-value of 0.00).

Figure 6. Individual CG Coefficients using Estimated θ and σ_ϵ

The figure plots individual CG coefficients in the model (with estimated θ and σ_ϵ) in the y-axis and CG coefficients in the survey data in the x-axis. The correlation between the two variables is 0.83 (p-value of 0.00).



If we run consensus regression with the simulated data, we also find a positive correlation between simulated and estimated consensus coefficients, but lower than in the individual case (0.28 vs 0.83, see Appendix D Figure D1). The lower correlation reflects the fact that, unlike individual level coefficients, consensus coefficients are highly dependent on the magnitude of measurement noise $\sigma_{\epsilon,k}$, which is imprecisely estimated in Table 8. We also simulate the model allowing for analyst-specific representativeness and noise coefficients $(\theta_k^i, \sigma_{\epsilon,k}^i)$. Table D2 in Appendix D reports the median estimate $(\theta_k^i, \sigma_{\epsilon,k}^i)$, which confirms our previous results. The estimated θ_k^i are highly positively correlated across series: individuals who over-react more in forecasting certain series also tend to over-react more in forecasting other series; see Table D3.

In sum, our model can quantitatively account for the evidence with a diagnostic parameter that is qualitatively sizable and in the ballpark of previous estimates performed using somewhat different methods and different datasets.

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, as previously shown by CG (2015). We then applied a psychologically founded model of belief formation, diagnostic expectations, to these data. We showed that diagnostic expectations generate over-reaction in individual forecasts, but if different forecasters see different information and/or use different models, the consensus forecast may exhibit under-reaction. The model thus reconciles these seemingly opposite patterns in the data, and yields many additional predictions, which are also supported in the data.

The ubiquity of over-reaction in individual macroeconomic forecasts helps reconcile several findings 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 the opposite, but as we show this is likely to be a consequence of aggregation. In our estimates of the diagnostic parameter, the predictable component of individual forecast errors entailed by over-reaction is comparable in magnitude to the rational response to news.

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 (1976) for assuming that people are entirely backward looking. With diagnostic expectations, forecasters are forward looking but their judgment is distorted by representativeness. Diagnostic expectations can thus serve as a micro-foundation of adaptive expectations and extrapolation, so that the latter may reflect the former at a crude level. At the same time, the kernel of truth property of diagnostic expectations produces exact predictions as to when we would see over-reaction in forecasts, and by how much, which becomes important in some contexts such as credit cycles (Bordalo, Gennaioli, and Shleifer 2018).

A final benefit of our approach is that it enables us to document and reconcile 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.

We leave 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 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 *aggregate* 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. The ability of optimists to leverage their bets may cause asset prices to be vulnerable to their overreaction, potentially creating misallocations and fragility (Geanakoplos 2010). Second, news may sometimes 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 entails aggregate over-reaction, as shown by our analysis of public signals. There is 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 an open question.

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.

Barber, and Terence Odean. "Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment." Quarterly Journal of Economics 116, no. 1 (2001): 261 – 292.

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.

Benigno, Pierpaolo, and Anastasios Karantounias. "Overconfidence, Subjective Perception and Pricing Behavior." Working paper, 2018.

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* 73, no. 1 (2018): 199-227.

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

Broer, Tobias, and Alexandre Kohlhas. "Forecaster Mis-Behavior." Working Paper (2018).

Campbell, John, and Robert Shiller. "Stock Prices, Earnings, and Expected Dividends." Journal of Finance 43, no. 3 (1987): 661 – 676.

Campbell, John, and Robert Shiller. “Cointegration and Tests of Present Value Models.” Journal of Political Economy 95, no. 5 (1987): 1062 – 1088.

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 Over-react?” *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.

Kohlhas, Alexandre, and Ansgar Walther. “Asymmetric Attention.” Working Paper (2018).

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.

Lucas, Robert. "Econometric Policy Evaluation: A Critique". *Carnegie-Rochester Conference Series on Public Policy* 1 (1976): 19–46.

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. Proofs

Proposition 1. The data generating process is $x_t = \rho x_{t-1} + u_t$, where $u_t \sim \mathcal{N}(0, \sigma_u^2)$ i.i.d. over time. Forecaster i observes a noisy signal $s_t^i = x_t + \epsilon_t^i$, where $\epsilon_t^i \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is i.i.d. analyst specific noise. Rational expectations are obtained iteratively:

$$f\left(x_t | S_t^i\right) = f\left(x_t | S_{t-1}^i\right) \frac{f\left(s_t^i | x_t\right)}{f\left(s_t^i\right)}$$

The rational estimate thus follows $f\left(x_t | S_t^i\right) \sim \mathcal{N}\left(x_{t|t}^i, \frac{\Sigma_{t|t-1}\sigma_\epsilon^2}{\Sigma_{t|t-1}+\sigma_\epsilon^2}\right)$ with

$$x_{t|t}^i = x_{t|t-1}^i + \frac{\Sigma_{t|t-1}}{\Sigma_{t|t-1} + \sigma_\epsilon^2}\left(s_t^i - x_{t|t-1}^i\right),$$

where $\Sigma_{t|t-1}$ is the variance of the prior $f\left(x_t | S_{t-1}^i\right)$. The variance of $f\left(x_{t+1} | S_t^i\right)$ is:

$$\Sigma_{t+1|t} \equiv var_t(\rho x_t + u_{t+1}) = \rho^2 \frac{\Sigma_{t|t-1}\sigma_\epsilon^2}{\Sigma_{t|t-1} + \sigma_\epsilon^2} + \sigma_u^2,$$

so that the steady state variance $\Sigma = \Sigma_{t+1|t} = \Sigma_{t|t-1}$ is equal to:

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

Beliefs about the current state are then described by $f\left(x_t | S_t^i\right) \sim \mathcal{N}\left(x_{t|t}^i, \frac{\Sigma\sigma_\epsilon^2}{\Sigma+\sigma_\epsilon^2}\right)$, where:

$$x_{t|t}^i = x_{t|t-1}^i + \frac{\Sigma}{\Sigma + \sigma_\epsilon^2}\left(s_t^i - x_{t|t-1}^i\right)$$

Let us now construct diagnostic expectations. For $s_t^i = x_{t|t-1}^i$ we have $x_{t|t}^i = x_{t|t-1}^i = \rho x_{t-1|t-1}^i$, so that $f\left(x_t | S_{t-1}^i \cup \{x_{t|t-1}^i\}\right) \sim \mathcal{N}\left(\rho x_{t-1|t-1}^i, \frac{\Sigma\sigma_\epsilon^2}{\Sigma+\sigma_\epsilon^2}\right)$. In light of the definition of diagnostic expectations in Equation (7), we have that the diagnostic distribution $f^\theta\left(x_t | S_t^i\right)$ fulfils:

$$\ln f^\theta\left(x_t | S_t^i\right) \propto -\frac{\left(x_t - x_{t|t}^i\right)^2}{2\frac{\Sigma\sigma_\epsilon^2}{\Sigma + \sigma_\epsilon^2}} - \theta \frac{\left(x_t - x_{t|t}^i\right)^2 - \left(x_t - x_{t|t-1}^i\right)^2}{2\frac{\Sigma\sigma_\epsilon^2}{\Sigma + \sigma_\epsilon^2}}$$

$$= -\frac{1}{2\frac{\Sigma\sigma_\epsilon^2}{\Sigma + \sigma_\epsilon^2}}\left[x_t^2 - 2x_t\left(x_{t|t}^i + \theta\left(x_{t|t}^i - x_{t|t-1}^i\right)\right) + \left(x_{t|t}^i\right)^2(1+\theta) - \theta\left(x_{t|t-1}^i\right)^2\right]$$

Given the normalization $\int f^{\theta}(x|S_t^i)dx = 1$, we find $f^{\theta}(x_t|S_t^i) \sim \mathcal{N}\left(x_{t|t}^{i,\theta}, \frac{\Sigma\sigma_\epsilon^2}{\Sigma+\sigma_\epsilon^2}\right)$ with $x_{t|t}^{i,\theta} = x_{t|t}^i + \theta\left(x_{t|t}^i - x_{t|t-1}^i\right)$. Using the definition of the Kalman filter $x_{t|t}^i$ we can write:

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

Proposition 2. Denote by $K = \Sigma/(\Sigma+\sigma_\epsilon^2)$ the Kalman gain. The rational consensus estimate for the current state is then equal to $\int x_{t|t}^i \, di \equiv x_{t|t} = x_{t|t-1} + K\left(x_t - x_{t|t-1}\right)$.

The consensus estimation error under rationality is then equal to $x_t - x_{t|t} = \frac{1-K}{K}(x_{t|t} - x_{t|t-1})$. The diagnostic filter for an individual analyst is equal to $x_{t|t}^{i,\theta} = x_{t|t}^i + \theta(x_{t|t}^i - x_{t|t-1}^i)$, which implies a consensus equation $x_{t|t}^{\theta} = x_{t|t} + \theta(x_{t|t} - x_{t|t-1})$. We thus have:

$$x_t - x_{t|t}^{\theta} = \left(\frac{1-K}{K} - \theta\right)\left(x_{t|t} - x_{t|t-1}\right).$$

Note, in addition, that the diagnostic consensus forecast revision is equal to:

$$x_{t|t}^{\theta} - x_{t|t-1}^{\theta} = (1+\theta)\left(x_{t|t} - x_{t|t-1}\right) - \theta\rho\left(x_{t-1|t-1} - x_{t-1|t-2}\right).$$

Therefore, the consensus CG coefficient is given by:

$$\beta = \frac{cov\left(x_{t+h} - x_{t+h|t}^{\theta}, x_{t+h|t}^{\theta} - x_{t+h|t-1}^{\theta}\right)}{var\left(x_{t+h|t}^{\theta} - x_{t+h|t-1}^{\theta}\right)}$$

$$= \left(\frac{1-K}{K} - \theta\right) \cdot \frac{cov\left[x_{t|t} - x_{t|t-1}, (1+\theta)\left(x_{t|t} - x_{t|t-1}\right) - \theta\rho\left(x_{t-1|t-1} - x_{t-1|t-2}\right)\right]}{var\left[(1+\theta)\left(x_{t|t} - x_{t|t-1}\right) - \theta\rho\left(x_{t-1|t-1} - x_{t-1|t-2}\right)\right]}.$$

Where we have that:

$$\begin{aligned} &cov\left[x_{t|t} - x_{t|t-1}, (1+\theta)\left(x_{t|t} - x_{t|t-1}\right) - \theta\rho\left(x_{t-1|t-1} - x_{t-1|t-2}\right)\right] \\ &\qquad = (1+\theta)var\left(x_{t|t} - x_{t|t-1}\right) - \theta\rho cov\left(x_{t|t} - x_{t|t-1}, x_{t-1|t-1} - x_{t-1|t-2}\right), \end{aligned}$$

and

$$\begin{aligned} &var\left[(1+\theta)\left(x_{t|t} - x_{t|t-1}\right) - \theta\rho\left(x_{t-1|t-1} - x_{t-1|t-2}\right)\right] \\ &\qquad = \left[(1+\theta)^2 + \theta^2\rho^2\right]var\left(x_{t|t} - x_{t|t-1}\right) \\ &\qquad - 2\theta(1+\theta)\rho cov\left(x_{t|t} - x_{t|t-1}, x_{t-1|t-1} - x_{t-1|t-2}\right). \end{aligned}$$

To compute the covariance between adjacent rational revisions, note that $x_{t|t} = x_{t|t-1} + K(x_t - x_{t|t-1})$ and $x_{t|t-1} = x_{t|t-2} + K(\rho x_{t-1} - x_{t|t-2})$ imply that:

$$x_{t|t} - x_{t|t-1} = (1-K)\rho\left(x_{t-1|t-1} - x_{t-1|t-2}\right) + Ku_t.$$

As a result,

$$cov\left(x_{t|t} - x_{t|t-1}, x_{t-1|t-1} - x_{t-1|t-2}\right) = (1-K)\rho \cdot var\left(x_{t|t} - x_{t|t-1}\right)$$

Therefore:

$$\beta = \left(\frac{1-K}{K} - \theta\right) \cdot \frac{(1+\theta) - \theta\rho^2(1-K)}{[(1+\theta)^2 + \theta^2\rho^2] - 2\theta(1+\theta)\rho^2(1-K)},$$

which is positive if and only if $1 - K > \theta K$, namely, $\theta < \sigma_\epsilon^2/\Sigma$.

Consider individual level forecasts. The coefficient (at the individual level) of regressing forecast error on forecast revision is equal to:

$$\beta^p = \frac{cov\left(x_{t+h} - x_{t+h|t}^{i,\theta}, x_{t+h|t}^{i,\theta} - x_{t+h|t-1}^{i,\theta}\right)}{var\left(x_{t+h|t}^{i,\theta} - x_{t+h|t-1}^{i,\theta}\right)} =$$

$$\frac{cov\left(x_{t|t} - x_{t|t}^{i,\theta}, x_{t|t}^{i,\theta} - x_{t|t-1}^{i,\theta}\right)}{var\left(x_{t|t}^{i,\theta} - x_{t|t-1}^{i,\theta}\right)}$$

where $x_{t|t}^{i,\theta} - x_{t|t-1}^{i,\theta} = (1+\theta)\left(x_{t|t}^{i} - x_{t|t-1}^{i}\right) - \theta\rho\left(x_{t-1|t-1}^{i} - x_{t-1|t-2}^{i}\right)$. Because at the individual level $cov\left(x_{t|t}^{i} - x_{t|t-1}^{i}, x_{t|t-1}^{i} - x_{t|t-2}^{i}\right) = 0$, we immediately have that:

$$\beta^p = -\frac{\theta(1+\theta)}{(1+\theta)^2 + \rho^2\theta^2}.$$

■

Corollary 1. Denote by p_i the precision of the private signal, by p the precision of the public signal, by p_f the precision of the lagged rational forecast $x_{t|t-1}^{i}$. The diagnostic filter at time t is:

$$x_{t|t}^{i,\theta} = x_{t|t-1}^{i} + (1+\theta)\frac{p_i}{p_i + p + p_f}\left(s_t^i - x_{t|t-1}^{i}\right) + (1+\theta)\frac{p}{p_i + p + p_f}\left(s_t - x_{t|t-1}^{i}\right).$$

The precision p_f of the forecast depends on the sum of the precisions $(p_i + p)$ and hence stays constant as we vary the relative precision of the public versus private signal.

Denote the Kalman gains as $K_1 = \frac{p_i}{p_i+p+p_f}$ and $K_2 = \frac{p}{p_i+p+p_f}$, and $K = K_1 + K_2$. The consensus Kalman filter can then be written as $x_{t|t} = x_{t|t-1} + K\left(x_t - x_{t|t-1}\right) + K_2 v_t$, while the diagnostic filter can be written as $x_{t|t}^{\theta} = x_{t|t} + \theta\left(x_{t|t} - x_{t|t-1}\right)$. The consensus coefficient is then:

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

Consider first the numerator. Denote by $FR_t \equiv x_{t|t} - x_{t|t-1}$ the revision of the rational forecast of x_t between t and $t-1$. Then:

$$x_t - x_{t|t}^{\theta} = \left(\frac{1-K}{K} - \theta\right) FR_t - \frac{K_2}{K} v_t,$$

$$x_{t|t}^{\theta} - x_{t|t-1}^{\theta} = (1+\theta) FR_t - \theta\rho FR_{t-1}.$$

The difference between $x_{t|t} = x_{t|t-1} + K\left(x_t - x_{t|t-1}\right) + K_2 v_t$ and $x_{t|t-1} = x_{t|t-2} + K\big(\rho x_{t-1} - x_{t|t-2}\big) + K_2 \rho v_{t-1}$ reads:

$$FR_t = (1-K)\rho FR_{t-1} + K u_t + K_2 (v_t - \rho v_{t-1}),$$

which in turn implies:

$$cov(FR_t, FR_{t-1}) = (1-K)\rho \cdot var(FR_t) - \rho K_2^2 \sigma_v^2. \qquad (A.1)$$

It is also immediate to find that:

$$var(FR_t) = \frac{K^2 \sigma_u^2 + [(1+\rho^2) - 2\rho^2(1-K)] K_2^2 \sigma_v^2}{1 - [(1-K)\rho]^2}.$$

The numerator of the CG coefficient is then equal to:

$$cov\left(x_t - x_{t|t}^{\theta}, x_{t|t}^{\theta} - x_{t|t-1}^{\theta}\right) = \left(\frac{1-K}{K} - \theta\right) cov[FR_t, (1+\theta)FR_t - \theta\rho FR_{t-1}] - \frac{K_2}{K}(1+\theta) K_2 \sigma_v^2$$

$$= \left(\frac{1-K}{K} - \theta\right)\left[[1 + \theta - \theta\rho^2(1-K)] var(FR_t) + \theta\rho^2 K_2^2 \sigma_v^2\right] - \frac{(1+\theta) K_2^2 \sigma_v^2}{K} \qquad (A.2)$$

The denominator of the CG coefficient equals:

$$var\left(x_{t|t}^{\theta} - x_{t|t-1}^{\theta}\right) = var[(1+\theta)FR_t - \theta\rho FR_{t-1}]$$
$$= [(1+\theta)^2 + \theta^2\rho^2] var(FR_t) - 2\theta(1+\theta)\rho cov(FR_t, FR_{t-1})$$

which implies that:

$$\frac{var\left(x^{\theta}_{t|t}-x^{\theta}_{t|t-1}\right)}{[(1+\theta)^2+\theta^2\rho^2]}+\frac{2\theta(1+\theta)\rho}{[(1+\theta)^2+\theta^2\rho^2]}cov(FR_t,FR_{t-1})=var(FR_t). \qquad (A.3)$$

Putting $(A.3)$ together with $(A.1)$ one obtains:

$$cov(FR_t,FR_{t-1})=$$
$$=\frac{(1-K)\rho var\left(x^{\theta}_{t|t}-x^{\theta}_{t|t-1}\right)}{\left[1-\frac{2\theta(1-K)(1+\theta)\rho^2}{[(1+\theta)^2+\theta^2\rho^2]}\right][(1+\theta)^2+\theta^2\rho^2]}-\frac{\rho K_2^2\sigma_v^2}{\left[1-\frac{2\theta(1-K)(1+\theta)\rho^2}{[(1+\theta)^2+\theta^2\rho^2]}\right]} \quad (A.4)$$

Using Equations $(A.2)$ and $(A.4)$ we find:

$$\begin{aligned}
&cov\left(x_t-x^{\theta}_{t|t},x^{\theta}_{t|t}-x^{\theta}_{t|t-1}\right)\\
&\qquad=\left(\frac{1-K}{K}-\theta\right)\left[(1+\theta)\frac{var\left(x^{\theta}_{t|t}-x^{\theta}_{t|t-1}\right)}{(1+\theta)^2+\theta^2\rho^2}\right.\\
&\qquad\left.+\theta\rho\left(\frac{2(1+\theta)^2}{(1+\theta)^2+\theta^2\rho^2}-1\right)cov(FR_t,FR_{t-1})\right]-\frac{(1+\theta)K_2^2\sigma_v^2}{K}=
\end{aligned}$$

$$=\beta_\infty var\left(x^{\theta}_{t|t}-x^{\theta}_{t|t-1}\right)-K_2^2\sigma_v^2\left[\frac{\theta\rho^2\left(\frac{1-K}{K}-\theta\right)\left(\frac{2(1+\theta)^2}{(1+\theta)^2+\theta^2\rho^2}-1\right)}{\left[1-\frac{2\theta(1-K)(1+\theta)\rho^2}{(1+\theta)^2+\theta^2\rho^2}\right]}+\frac{(1+\theta)}{K}\right],$$

where β_∞ is the consensus coefficient obtained when the public signal is fully uninformative, namely $\sigma_u^2\to\infty$ and thus $K_2\to 0$. On the other hand using equation (A.3) this can be rewritten as:

$$var\left(x^{\theta}_{t|t}-x^{\theta}_{t|t-1}\right)=\frac{[(1+\theta)^2+\theta^2\rho^2-2\theta(1+\theta)(1-K)\rho^2]K^2\sigma_u^2}{1-[(1-K)\rho]^2}+AK_2^2\sigma_v^2,$$

where A is a suitable positive coefficient. The CG coefficient is then equal to:

$$\frac{cov\left(x_t-x^{\theta}_{t|t},x^{\theta}_{t|t}-x^{\theta}_{t|t-1}\right)}{var\left(x^{\theta}_{t|t}-x^{\theta}_{t|t-1}\right)}=\beta_\infty-\frac{\left[\frac{\theta\rho^2\left(\frac{1-K}{K}-\theta\right)\left(\frac{2(1+\theta)^2}{(1+\theta)^2+\theta^2\rho^2}-1\right)}{1-\frac{2\theta(1-K)(1+\theta)\rho^2}{(1+\theta)^2+\theta^2\rho^2}}+\frac{(1+\theta)}{K}\right]K_2^2\sigma_v^2}{\frac{[(1+\theta)^2+\theta^2\rho^2-2\theta(1+\theta)(1-K)\rho^2]K^2\sigma_u^2}{1-[(1-K)\rho]^2}+AK_2^2\sigma_v^2}.$$

For given total informativeness K, the above expression falls in the precision of the public signal, namely as K_2^2 grows, if and only if:

$$\left[\frac{\theta\rho^2\left(\frac{1-K}{K}-\theta\right)\left(\frac{2(1+\theta)^2}{(1+\theta)^2+\theta^2\rho^2}-1\right)}{1-\frac{2\theta(1-K)(1+\theta)\rho^2}{(1+\theta)^2+\theta^2\rho^2}}+\frac{(1+\theta)}{K}\right]>0.$$

A sufficient condition for this to hold is that $\left(\frac{1-K}{K}-\theta\right)>0$, which is equivalent to $\beta_\infty>0$.

■

Proof of Proposition 3

The diagnostic expectation at time t about $t+3$ is given by:

$$x_{t+3|t}^{i,\theta}=x_{t+3|t}^{i}+\theta FR_{t+3|t}^{i},$$

where $FR_{t+3|t}^{i}=\left(x_{t+3|t}^{i}-x_{t+3|t-1}^{i}\right)$ The diagnostic forecast revision $FR_{t+3|t}^{i,\theta}=\left(x_{t+3|t}^{i,\theta}-x_{t+3|t-1}^{i,\theta}\right)$ is therefore equal to:

$$FR_{t+3|t}^{i,\theta}=(1+\theta)FR_{t+3|t}^{i}-\theta FR_{t+3|t-1}^{i}.$$

The diagnostic forecast error $FE_{t+3|t}^{i,\theta}\equiv x_{t+3}-x_{t+3|t}^{i,\theta}$ is equal to:

$$FE_{t+3|t}^{i,\theta}=u_{t+3}-\theta FR_{t+3|t}^{i},$$

where u_{t+3} is white noise. We then have:

$$cov\left(FE_{t+3|t}^{i,\theta},FR_{t+3|t}^{i,\theta}\right)=-\theta cov\left(FR_{t+3|t}^{i},(1+\theta)FR_{t+3|t}^{i}-\theta FR_{t+3|t-1}^{i}\right)$$

$$=-\theta(1+\theta)var\left(FR_{t+3|t}^{i}\right)$$

$$var\left(FR_{t+3|t}^{i,\theta}\right)=(1+\theta)^2var\left(FR_{t+3|t}^{i}\right)+\theta^2var\left(FR_{t+3|t-1}^{i}\right).$$

As a result, the relationship between forecast error and forecast revision is equal to:

$$FE_{t+3|t}^{i,\theta}=-\frac{\theta(1+\theta)}{(1+\theta)^2+\theta^2\frac{var\left(FR_{t+3|t-1}^{i}\right)}{var\left(FR_{t+3|t}^{i}\right)}}FR_{t+3|t}^{i,\theta}+v_{t+3}$$

By plugging Equation (13) in the text, we obtain:

$$FE_{t+3|t}^{i}=-\frac{\rho_2\theta(1+\theta)}{(1+\theta)^2+\theta^2\frac{var\left(FR_{t+3|t-1}^{i}\right)}{var\left(FR_{t+3|t}^{i}\right)}}FR_{t+2|t}^{i}-\frac{\rho_1\theta(1+\theta)}{(1+\theta)^2+\theta^2\frac{var\left(FR_{t+3|t-1}^{i}\right)}{var\left(FR_{t+3|t}^{i}\right)}}FR_{t+1|t}^{i}+v_{t+3},$$

If $FR_{t+2|t}^{i}$ and $FR_{t+1|t}^{i}$ are not collinear, the above equation can be estimated and it satisfies the prediction of Proposition 3. To conclude the proof, we therefore need to prove non-collinearity. Recall that the state follows AR(2) dynamics:

$$x_{t+1} = ax_t + bx_{t-1} + u_{t+1},$$

At time t, the agent observes two signals, one about the current state, $s_t^i = x_t + \epsilon_t^i$, and one about the past state $z_t^i = s_{t-1,t}^i = x_{t-1} + v_t^i$. Signals ϵ_t^i and v_t^i are normal with precision ϵ and v. At time t, the agent forms estimates about x_t and x_{t-1}. He then combines them to forecast about x_{t+k}, $k \geq 1$.

To ease notation we drop superscripts i from the noise and the signals and subscript t from the signals. Conditional on the signals, the density of the current state $f(x_t, x_{t-1}|s_t, z_t)$ satisfies:

$$-\ln f \propto \epsilon(1-\varphi^2)(s_t - x_t)^2 + v(1-\varphi^2)(z_t - x_{t-1})^2 + \left(x_t - x_{t|t-1}\right)^2 p + \left(x_{t-1} - x_{t-1|t-1}\right)^2 q$$
$$- 2\varphi\sqrt{pq}\left(x_t - x_{t|t-1}\right)\left(x_{t-1} - x_{t-1|t-1}\right)$$

where p is the precision of x_t, q is the precision of x_{t-1}, and φ is their correlation.

Maximizing the likelihood f with respect to x_t and x_{t-1} yields the first order conditions:

$$-2\epsilon(1-\varphi^2)\left(s_t - x_{t|t}\right) + 2p\left(x_{t|t} - x_{t|t-1}\right) - 2\varphi\sqrt{pq}\left(x_{t-1|t} - x_{t-1|t-1}\right) = 0$$

$$-2v(1-\varphi^2)\left(z_t - x_{t-1|t}\right) + 2q\left(x_{t-1|t} - x_{t-1|t-1}\right) - 2\varphi\sqrt{pq}\left(x_{t|t} - x_{t|t-1}\right) = 0$$

which identify the conditional estimates (the Kalman filter):

$$x_{t|t} = \frac{(1-\varphi^2)\frac{\epsilon}{p}s_t + x_{t|t-1} + \varphi\sqrt{\frac{q}{p}}FR_{t-1|t}}{(1-\varphi^2)\frac{\epsilon}{p} + 1},$$

$$x_{t-1|t} = \frac{(1-\varphi^2)\frac{v}{q}z_t + x_{t-1|t-1} + \varphi\sqrt{\frac{p}{q}}FR_{t|t}}{(1-\varphi^2)\frac{v}{q} + 1},$$

Where $FR_{s|t}$ is the forecast revision at t for x_s. This further implies that:

$$FR_{t|t} = \frac{(1-\varphi^2)\frac{\epsilon}{p}\left(s_t - x_{t|t-1}\right) + \varphi\sqrt{\frac{q}{p}}FR_{t-1|t}}{(1-\varphi^2)\frac{\epsilon}{p} + 1},$$

$$FR_{t-1|t} = \frac{(1-\varphi^2)\frac{v}{q}\left(z_t - x_{t-1|t-1}\right) + \varphi\sqrt{\frac{p}{q}}FR_{t|t}}{(1-\varphi^2)\frac{v}{q} + 1}.$$

These equations imply that, provided $\varphi < 1$, the forecast revisions $FR_{t|t}$ and $FR_{t-1|t}$ are linearly independent combinations of the news $s_t - x_{t|t-1}$ and $z_t - x_{t-1|t-1}$:

$$FR_{t|t} = \frac{\left[(1-\varphi^2)\frac{v}{q}+1\right]\frac{\epsilon}{p}\left(s_t - x_{t|t-1}\right) + \varphi\sqrt{\frac{1}{qp}}v\left(z_t - x_{t-1|t-1}\right)}{\left[(1-\varphi^2)\frac{v}{q}+1\right]\frac{\epsilon}{p} + \frac{v}{q} + 1},$$

$$FR_{t-1|t} = \frac{\left[(1-\varphi^2)\frac{\epsilon}{p}+1\right]\frac{v}{q}\left(z_t - x_{t-1|t-1}\right) + \varphi\sqrt{\frac{1}{qp}}\epsilon\left(s_t - x_{t|t-1}\right)}{\left[(1-\varphi^2)\frac{\epsilon}{p}+1\right]\frac{v}{q} + \frac{\epsilon}{p} + 1}.$$

Therefore, $FR^i_{t|t}$ and $FR^i_{t-1|t}$ are not collinear. Since $FR^i_{t+1|t} = aFR^i_{t|t} + bFR^i_{t-1|t}$ and $FR^i_{t+2|t} = (a^2 + b)FR^i_{t|t} + abFR^i_{t-1|t}$, we conclude that $FR^i_{t+2|t}$ and $FR^i_{t+1|t}$ are not collinear.

■

B. Variable Definitions

For each variable, we report the source survey, the survey time, the survey question, and the definitions of forecast variable, revision variable, and actuals.

1. NGDP_SPF

- Variable: Nominal GDP. Source: SPF.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of nominal GDP in the current quarter and the next 4 quarters.
- Forecast: 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: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of real GDP in the current quarter and the next 4 quarters.
- Forecast: 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: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.

- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: End of the middle month in the quarter/beginning of the last month in the quarter.
- question: Real GDP growth (annualized rate) in the current quarter and the next 4 to 5 quarters.
- Forecast: 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: $\frac{F_t(z_t+z_{t+1}+z_{t+2}+z_{t+3})}{4} - \frac{F_{t-1}(z_t+z_{t+1}+z_{t+2}+z_{t+3})}{4}$.
- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of GDP price deflator in the current quarter and the next 4 quarters.
- Forecast: 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: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of real consumption in the current quarter and the next 4 quarters.
- Forecast: 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: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of real non-residential investment in the current quarter and the next 4 quarters.

- Forecast: 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: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of real residential investment in the current quarter and the next 4 quarters.
- Forecast: 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: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of real federal government consumption in the current quarter and the next 4 quarters.
- Forecast: 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: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of real state and local government consumption in the current quarter and the next 4 quarters.
- Forecast: 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: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average unemployment rate in the current quarter and the next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of housing starts in the current quarter and the next 4 quarters.
- Forecast: 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: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$.
- Actuals: $\frac{x_{t+3}}{x_{t-1}} - 1$, 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.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average federal funds rate in the current quarter and the next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: x_{t+3}.

13. TB3M_SPF
- Variable: 3-month Treasury rate. Source: SPF.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average 3-month Treasury rate in the current quarter and next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: x_{t+3}.

14. TB3M_BC

- Variable: 3-month Treasury rate. Source: Blue Chip.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average 3-month Treasury rate in the current quarter and next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: x_{t+3}.

15. TN5Y_BC

- Variable: 5-year Treasury rate. Source: Blue Chip.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average 5-year Treasury rate in the current quarter and the next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: x_{t+3}.

16. TN10Y_SPF

- Variable: 10-year Treasury rate. Source: SPF.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average 10-year Treasury rate in the current quarter and next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: x_{t+3}.

17. TN10Y_BC

- Variable: 10-year Treasury rate. Source: Blue Chip.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average 10-year Treasury rate in the current quarter and next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: x_{t+3}.

18. AAA_SPF

- Variable: AAA corporate bond rate. Source: SPF.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average AAA corporate bond rate in the current quarter and next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: x_{t+3}.

19. AAA_BC

- Variable: AAA corporate bond rate. Source: Blue Chip.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average AAA corporate bond rate in the current quarter and next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: x_{t+3}.

20. BAA_BC

- Variable: BAA corporate bond rate. Source: Blue Chip.
- time: Around the 3rd week of the middle month in the quarter.
- question: The level of average BAA corporate bond rate in the current quarter and next 4 quarters.
- Forecast: 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: $F_t x_{t+3} - F_{t-1} x_{t+3}$.
- Actuals: x_{t+3}.

C. Robustness Checks

Table C1. 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 C2. 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

D. Model Estimation: supporting information

Table D1. Estimates of AR(1) and AR(2) fundamentals

The Table shows the relevant autocorrelation parameters and variance of shocks for the AR(1) and AR(2) specifications of each time series.

	AR(1)		AR(2)		
variable	ρ	σ_u	ρ_1	ρ_2	σ_u
NGDP	0.92	1.08	1.27	-0.37	1.00
RGDP	0.87	1.12	1.33	-0.51	0.96
RGDPBC	0.86	0.01	1.24	-0.43	0.01
PGDP	0.98	0.49	1.45	-0.48	0.43
RCONSUM	0.87	0.72	0.89	-0.02	0.72
RNRESIN	0.88	3.43	1.25	-0.41	3.14
RRESINV	0.88	5.68	1.27	-0.42	5.01
RGF	0.78	2.83	0.74	0.06	2.82
RGSL	0.90	0.77	0.85	0.04	0.77
HOUSING	0.85	11.80	1.14	-0.34	11.12
UNEMP	0.96	0.37	1.48	-0.53	0.31
tb3m	0.95	0.58	1.22	-0.27	0.55
tb3mBC	0.99	0.45	1.54	-0.56	0.37
tn5yBC	0.97	0.44	1.17	-0.21	0.42
tn10y	0.97	0.38	1.17	-0.21	0.37
tn10yBC	0.97	0.38	1.21	-0.25	0.37
AAA	0.97	0.38	1.16	-0.20	0.36
AAABC	0.97	0.33	1.19	-0.22	0.32
BAABC	0.95	0.37	1.01	-0.07	0.37
FUNDSBC	0.99	0.50	1.53	-0.55	0.42

Kalman inference for AR(1) processes was described in the text, see Equations (8,9). We now describe Kalman inference for an AR(2) process. The state variable is a vector $\vec{x}_t = (x_t, x_{t-1})$ which evolves according to $\vec{x}_t = A\vec{x}_{t-1} + W_t$, with transition matrix $A = \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix}$ and disturbance $W_t = \begin{bmatrix} u_t & 0 \\ 0 & 0 \end{bmatrix}$ with $u_t \sim \mathcal{N}(0, \sigma_u^2)$ i.i.d. across time. The observation equation is $s_t = C\vec{x}_t + \epsilon_t$ with $C = [1 \quad 0]$ and $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$ i.i.d. across time. The Kalman filter can then be written:

$$x_{t|t}^{i,\theta} = x_{t|t-1}^{i} + (1+\theta)\frac{\Sigma_{11}}{\Sigma_{11} + \sigma_\epsilon^2}\left(s_t^i - \rho_1 x_{t-1|t-1}^i - \rho_2 x_{t-2|t-1}^i\right),$$

where Σ_{11} is the first entry of the steady state variance matrix of beliefs at $t-1$ about x_t, which is given by the following condition:

$$\Sigma = \mathrm{A}\Sigma\mathrm{A}^T + W - A\Sigma\mathrm{C}(\mathrm{C}^T\Sigma\mathrm{C} + \sigma_\epsilon^2)^{-1}\mathrm{C}^T\Sigma\mathrm{A}^T$$

where $W = \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & 0 \end{bmatrix}$.

Table 8 in Section 6 presents the pooled estimates of the latent parameters θ_k and $\sigma_{\epsilon,k}$ that were allowed to vary by series k but not by individual forecaster. Table D2 shows the median estimates of these parameters at the individual level.

Table D2. Model Estimation Results by Forecaster

This table shows the median of individual-level θ^i and σ_ϵ^i (normalized by σ_u) estimates, as well as the CG coefficients in the model with estimated θ^i and σ_ϵ^i .

	Median θ^i	Median $\sigma_\epsilon^i/\sigma_u$	Individual CG	Consensus CG
Nominal GDP (SPF)	0.32	1.08	-0.20	0.29
Real GDP (SPF)	0.69	0.78	-0.26	0.10
Real GDP (BC)	0.63	1.43	-0.30	0.35
GDP Price Index Inflation (SPF)	0.59	3.42	-0.25	1.06
Real Consumption (SPF)	1.74	4.17	-0.45	1.09
Real Non-Residential Investment (SPF)	0.44	1.55	-0.15	1.15
Real Residential Investment (SPF)	0.42	1.68	-0.22	0.90
Real Federal Government Consumption (SPF)	1.57	3.00	-0.43	0.58
Real State&Local Govt Consumption (SPF)	3.54	6.08	-0.52	0.12
Housing Start (SPF)	1.37	2.11	-0.42	0.60
Unemployment (SPF)	-0.17	0.67	0.26	1.00
Fed Funds Rate (BC)	-0.01	1.24	-0.04	0.61
3M Treasury Rate (SPF)	0.21	1.60	0.04	1.18
3M Treasury Rate (BC)	-0.03	1.87	0.01	1.08
5Y Treasury Rate (BC)	0.37	2.49	-0.21	1.10
10Y Treasury Rate (SPF)	0.37	2.89	-0.35	0.74
10Y Treasury Rate (BC)	0.26	2.74	-0.30	0.88
AAA Corporate Bond Rate (SPF)	0.63	5.21	-0.36	1.20
AAA Corporate Bond Rate (BC)	0.76	5.20	-0.35	1.47
BAA Corporate Bond Rate (BC)	0.78	2.83	-0.39	0.69

Table D3 shows that there is a consistent correlation between individual level estimates of θ^i across series.

Table D3. Rank correlations for θ^i

Panel A: SPF series

	NGDP	RGDP	PGDP	RCONSUM	RNRESINV	RRESINV	RGF	RGSL	UNEMP	HOUSING	tb3m	tn10y
RGDP	0.48											
	0.000											
PGDP	-0.04	0.00										
	0.747	0.976										
RCONSUM	0.13	0.15	-0.20									
	0.311	0.250	0.122									
RNRESINV	0.41	0.34	-0.20	0.37								
	0.001	0.008	0.127	0.003								
RRESINV	0.29	0.13	-0.07	0.31	0.25							
	0.023	0.326	0.571	0.013	0.048							
RGF	0.14	0.19	0.04	-0.21	-0.10	-0.19						
	0.269	0.146	0.748	0.107	0.454	0.136						
RGSL	-0.06	0.16	0.08	0.11	0.29	0.05	-0.27					
	0.649	0.223	0.527	0.382	0.024	0.719	0.036					
UNEMP	-0.18	-0.10	0.04	-0.10	-0.07	-0.01	0.23	0.08				
	0.159	0.443	0.754	0.427	0.581	0.913	0.079	0.557				
HOUSING	0.08	-0.03	-0.09	0.16	0.18	0.45	0.02	0.09	0.03			
	0.518	0.822	0.487	0.217	0.170	0.000	0.848	0.498	0.814			
tb3m	0.15	0.22	-0.01	0.08	0.18	0.07	0.09	-0.10	0.03	0.04		
	0.233	0.087	0.944	0.537	0.158	0.609	0.484	0.426	0.791	0.732		
tn10y	-0.14	0.07	-0.02	-0.21	-0.11	-0.19	0.08	-0.05	0.08	-0.20	-0.08	
	0.286	0.598	0.875	0.096	0.381	0.148	0.519	0.680	0.558	0.123	0.537	
AAA	0.15	0.13	0.21	0.15	0.29	0.14	0.11	-0.03	-0.02	0.04	0.36	-0.09
	0.249	0.300	0.102	0.236	0.021	0.295	0.407	0.791	0.898	0.745	0.004	0.498

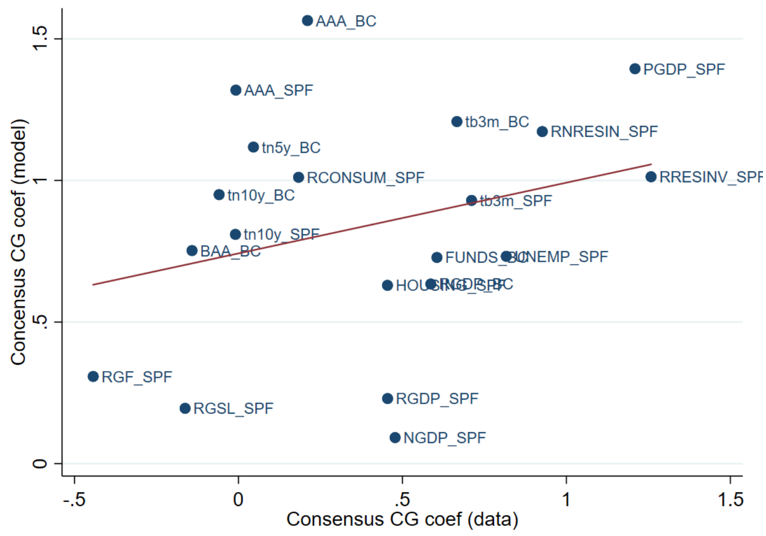
Panel B: Blue Chip series

	RGDPBC	FFBC	tb3mBC	tn5yBC	tn10yBC	AAABC
FFBC	0.13					
	0.306					
tb3mBC	0.10	0.54				
	0.450	0.000				
tb5yBC	0.15	0.45	0.37			
	0.243	0.000	0.003			
tn10yBC	-0.32	0.02	-0.01	0.02		
	0.010	0.876	0.956	0.863		
AAABC	-0.12	0.08	-0.03	0.15	0.20	
	0.346	0.530	0.808	0.247	0.122	
BAABC	0.26	0.40	0.52	0.22	0.10	-0.11
	0.041	0.001	0.000	0.083	0.424	0.376

Figure 6 in the text showed the model-predicted individual level CG coefficients were strongly correlated with those estimated in the pooled regressions. Figure D1 shows the corresponding predictions for the consensus CG coefficients.

Figure D1. Consensus CG Coefficients using Estimated θ and σ_ϵ

The figure plots consensus CG coefficients in the model (with estimated θ and σ_ϵ) in the y-axis and CG coefficients in the survey data in the x-axis. The correlation between the two variables is 0.28 (p-value of 0.22).



E. 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.

E.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 : $\frac{1}{N}\sum_{n=1}^{N} w(x_n) \to \int \frac{\tilde{p}(x)}{q(x)} \cdot q(x)dx = \int \tilde{p}(x)\, dx = Z$. 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 $p(x_n|x_{n-1} = x^i{}_{n-1}, s_n)$.

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

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.

E.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 E1 *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: $R(x_n|s_{1:n}) = \frac{p(x_n|\, s_{1:n})}{p(x_n|\, s_{1:n-1}, s^*)} = \frac{p(s_n|s_{1:n-1}, x_n) \cdot p(x_n|\, s_{1:n-1})}{p(s_n \mid s_{1:n-1})} \cdot (\frac{p(s^* \mid s_{1:n-1}) \cdot p(x_n|\, s_{1:n-1})}{p(s^*|\, s_{1:n-1})})^{-1}$.

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:

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

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

E.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 A3 shows, the diagnostic expectations over-reacts to news, relative to the rational benchmark.

We then check the results of the CG forecast error on forecast revision regressions. Figure A4 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 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.

Figure E1. 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.

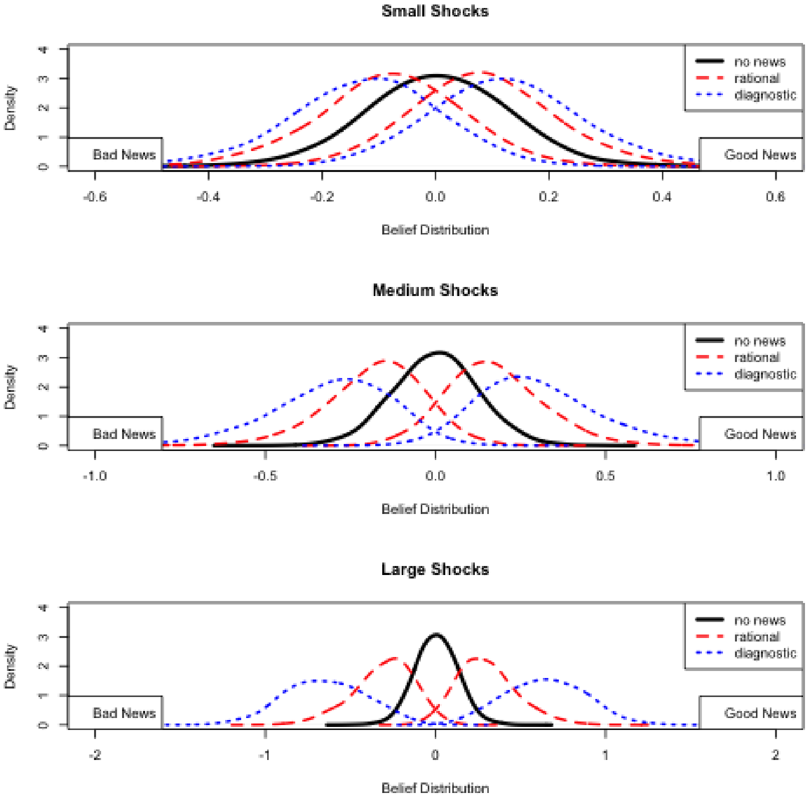
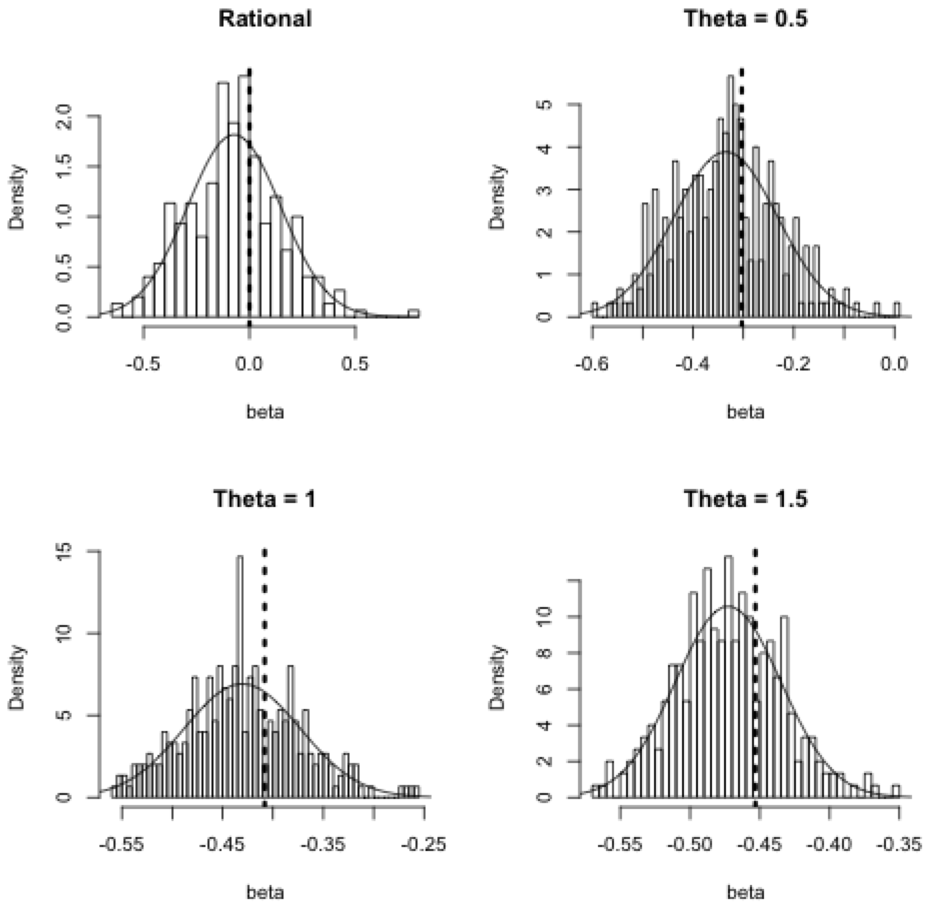
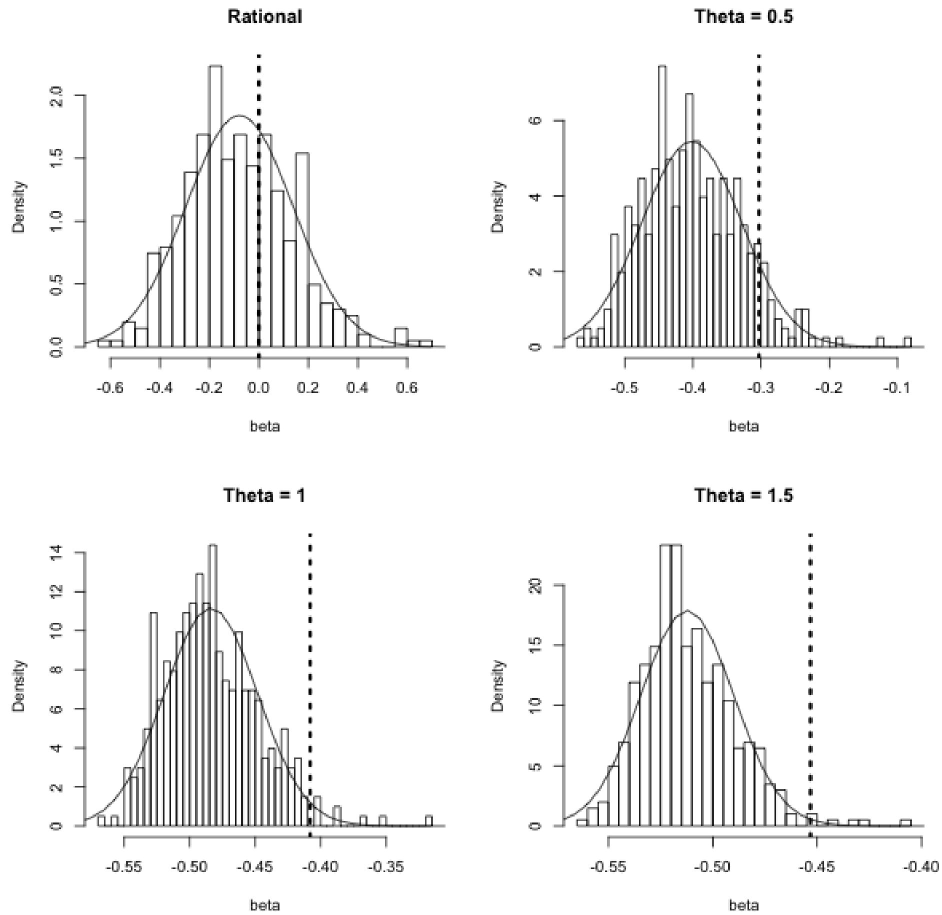


Figure E2. Individual CG Coefficients with Normal and Fat-Tailed Shocks

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

Finally, Figure A5 replicates the results for the contrast between regressions using individual and consensus forecasts. The general qualitative result is that there is much less over-reaction in consensus opinion, or even under-reaction for some cases. Under-reaction occurs when the noise ${\sigma_\epsilon}^2$is sufficiently high and individual over-reaction 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 E3. 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.