NBER WORKING PAPER SERIES

IMPERFECT MACROECONOMIC EXPECTATIONS: EVIDENCE AND THEORY

George-Marios Angeletos
Zhen Huo
Karthik A. Sastry

Working Paper 27308
http://www.nber.org/papers/w27308

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2020

This paper was prepared for the 2020 NBER Macroeconomics Annual. We owe special thanks to our discussants, Ricardo Reis and Jessica Watcher, and the organizers, Martin Eichenbaum and Erik Hurst. We also thank Fabrice Collard for help with some of the material in Section 5. We finally thank Jeffrey Campbell, Olivier Coibion, Nicola Gennaioli, Cosmin Ilut, Yueran Ma, and Alp Simsek for useful comments. Angeletos acknowledges the financial support of the National Science Foundation (Award #1757198). There is no other financial interest or conflict to disclose. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Imperfect Macroeconomic Expectations: Evidence and Theory
George-Marios Angeletos, Zhen Huo, and Karthik A. Sastry
NBER Working Paper No. 27308
June 2020
JEL No. E03,E3,E7,G02

ABSTRACT

We document a new fact about expectations: in response to the main shocks driving the business cycle, expectations under-react initially but over-shoot later on. We show how previous, seemingly conflicting, evidence can be understood as different facets of this fact. We finally explain what the cumulated evidence means for macroeconomic theory. There is little support for theories emphasizing under-extrapolation or two close cousins of it, cognitive discounting and level-K thinking. Instead, the evidence favors the combination of dispersed, noisy information and over-extrapolation.

George-Marios Angeletos
Department of Economics, E52-530
MIT
50 Memorial Drive
Cambridge, MA 02142
and NBER
angelet@mit.edu

Zhen Huo
Department of Economics
Yale University
New Haven, CT 06520
zhen.huo@yale.edu

Karthik A. Sastry
Department of Economics, MIT
77 Massachusetts Avenue E52-300
Cambridge, MA 02139
ksastry@mit.edu

1 Introduction

The rational expectations hypothesis is a bedrock of modern macroeconomics. It is often combined with a strong, complementary hypothesis that all data about the state of the economy is common knowledge. But an explosion of recent theoretical and empirical work has questioned both premises. This has pushed the discipline back toward reckoning with the "wilderness" of alternative models for expectations formation and equilibrium (as Sargent, 2001, paraphrasing Sims, 1980, famously put it).

One strand of the literature emphasizes informational frictions, which are sometimes rich enough to blur the boundary between the rational and non-rational.[1] Moving strictly beyond the rational model, some authors emphasize biases to over-extrapolate the past,[2] while others advocate for two close cousins of under-extrapolation, cognitive discounting and level-K thinking.[3] Another strand emphasizes overconfidence in various information sources, or prioritization of those that seem "representative."[4]

What does survey evidence on expectations tell us within the space of these alternative hypotheses? And what kind of evidence is most useful for building macroeconomic models and providing guidance about counterfactual scenarios?

In the hopes of answering these questions, and helping identify "where we are in the wilderness," this article uses a simple but flexible framework to accomplish the following goals: to draw a variety of recent theoretical and empirical contributions under a common umbrella; to guide a new, more informative, empirical strategy; and to select among competing theories of "imperfect expectations" in macroeconomics.

Our main empirical finding is *initial under-reaction* of beliefs in response to shocks followed by *delayed over-reaction*. Both unemployment and inflation expectations have an initially sluggish response to the shocks that drive most of the business-cycle variation in these variables. But over medium horizons, forecasts tend to over-shoot the actual outcomes.

This pattern speaks in favor of models that combine two key mechanisms: dispersed, noisy information and over-extrapolation. The former leaves room for theories emphasizing higher-order beliefs. The latter points in the opposite direction of cognitive discounting and level-K thinking, two concepts that, at least for our purposes, are close cousins of under-extrapolation.

We also demonstrate why our empirical strategy is more informative, at least vis-a-vis the class of theories under consideration, than previous alternatives. And we explain how our findings help resolve the apparent inconsistency between three previous empirical findings, which indeed serve as our starting point.

Understanding prior, seemingly conflicting, evidence. Previous empirical studies of expectations have often relied on simple regressions or correlations between actual outcomes and their forecasts in surveys.[5]

[1] This includes works on rational inattention (Sims, 2003, 2010; Mackowiak and Wiederholt, 2009; Matejka, 2015), sticky information (Mankiw and Reis, 2002; Kiley, 2007), and higher-order uncertainty (Morris and Shin, 2002, 2006; Woodford, 2003; Nimark, 2008; Angeletos and Lian, 2016, 2018).

[2] Gennaioli, Ma, and Shleifer (2015); Fuster, Laibson, and Mendel (2010); Guo and Wachter (2019)

[3] Gabaix 2020; Garcıa-Schmidt and Woodford 2019; Farhi and Werning 2019; Iovino and Sergeyev 2017.

[4] Bordalo, Gennaioli, and Shleifer (2017); Kohlhas and Broer (2019).

[5] This applies to the papers cited below, as well as Andrade and Le Bihan (2013), Gennaioli, Ma, and Shleifer (2015), Kohlhas and Broer (2019), and Fuhrer (2018). See also the discussion of Coibion and Gorodnichenko (2012) in Section 5 and Appendix B.

In Section 3, we revisit three such previously documented facts, henceforth referred to as Facts 1-3:

F1. For both unemployment and inflation, aggregate forecast errors are positively related to lagged aggregate forecast revisions, as in Coibion and Gorodnichenko (2015), or CG hereafter. This pattern suggests that aggregate forecasts *under-react* to aggregate news.

F2. The opposite pattern is often present at individual-level forecasts: as previously shown in Bordalo, Gennaioli, Ma, and Shleifer (2018), or BGMS hereafter, individual forecasts appear to *over-react* to own revisions (in the case of inflation, although not in the case of unemployment).[6]

F3. Finally, the following pattern, first noted in Kohlhas and Walther (2018), or KW hereafter, points toward *over-reaction* even at the aggregate level: aggregate forecast errors are positively correlated with the actual levels of unemployment and inflation.

These facts elude a simple, unified explanation. Do beliefs in the data under-react to innovations, as predicted by theories emphasizing informational frictions, higher-order uncertainty, cognitive discounting and level-K thinking? Or do they over-react, suggesting an entirely different mechanism?

To provide a clearer picture, we turn to theory. In Section 4, we introduce the "PE version" of our framework. Like the related empirical literature, this abstracts from the equilibrium fixed point between expectations and outcomes. But it allows for two key mechanisms: dispersed noisy information and over-extrapolation. A third mechanism, over-confidence, is also nested but turns out to be rather inessential.

The combination of dispersed information and over-extrapolation makes a sharp prediction for the impulse response functions (IRFs) of the average forecasts and forecast errors to aggregate shocks. In the first few periods after a shock occurs, the informational friction guarantees that forecasts under-react. But as time passes and learning kicks in, this friction dies out and over-extrapolation takes over, guaranteeing that forecast eventually over-react. The most telling feature of the combination of the two mechanisms is therefore a reversal of sign in the IRF of the average forecast errors.

The regressions underlying Facts 1 and 3 can be described as different weighted averages of this IRF. The one in CG happens to put more weight on the early portion of this IRF, where errors are positively correlated with past revisions due to dispersed information, while that in KW happens to put more weight on the later portion, where errors are negatively correlated with outcomes due to over-extrapolation. This resolves the apparent conflict between the form of under-reaction documented in CG and the form of over-reaction documented in KW, but perhaps most importantly underscores the difficulty in interpreting and using this kind of evidence. A similar point applies to the BGMS evidence, or Fact 2.

Focusing on impulse response functions (IRFs). Under the lens of our analysis, a superior empirical strategy emerges: the IRFs of the average forecasts and the average forecast errors to aggregate shocks provide strictly more information than the aforementioned empirical strategies and are also more easily interpretable. This leads to our main empirical contribution, which appears in Section 5 and which is to

[6] For inflation forecasts, the same pattern has been independently documented in Kohlhas and Broer (2019). BGMS offer a comprehensive investigation across variables, surveys, and empirical methods.

show that the hypothesized pattern of "sign reversal" in the response of forecast errors holds true in the data. We summarize this below as Fact 4:

F4. Consider two shocks, one that accounts for most of the business-cycle variation in unemployment and other macroeconomic quantities, and another that accounts for most of the business-cycle variation in inflation.[7] Construct the IRFs of the average *forecasts* of unemployment and inflation to the corresponding shocks. In both cases, average forecasts are initially under-react before over-shooting later on, or predicting larger and longer-lasting effects of the shock than those that occur.

For the reasons already explained, Fact 4 *alone* helps nail down the "right" combination of frictions under the lens of our framework: to match this fact, it is necessary and sufficient to combine over-extrapolation with a sufficiently large informational friction. And since this combination implies Facts 1-3, Fact 4 subsumes them and serves as a "sufficient statistic" for the counterfactuals of interest (more on this below).

We provide additional evidence for each of the two mechanisms as follows. First, we show that the subjective persistence, as revealed by the term structure of subjective expectations, is larger than the objective persistence, as measured by the impulse response of the outcome. And second, we show that the forecasts revisions of one one agent help predict the forecast errors of other agents. The former fact speaks directly to over-extrapolation, the latter to not only noisy but also dispersed, or private, information.

From PE to GE. In Section 6, we incorporate a GE feedback between expectations and outcomes. This part of our paper, which builds on the methods of Angeletos and Huo (2019), lets us accomplish four goals. First, we extend our lessons about the "right" model of beliefs to a broader GE context. [8] Second, we connect level-K thinking and cognitive discounting to the GE implications of under-extrapolation, and spell out the empirical content of these theories vis-a-vis expectations data. Third, we clarify how the causal effect of the belief distortions on macroeconomic outcomes depends parameters that determine the relative strength of PE and GE effects, such as the marginal propensity to consume. Finally, we quantify these distortions in a three-equation New Keynesian model.

The bottom line. The combination of old and new evidence we marshal in this paper offers, not only support for theories emphasizing informational frictions and higher-order uncertainty, but also guidance on what type of departure from full rationality seems most relevant in the business cycle context. In particular, we argue that over-extrapolation is needed in order to not only reconcile the previous, seemingly conflicting evidence of CG, KW and BGMS, but also account for the eventual overshooting in the response of the average forecasts we have documented here.

Conversely, we have ruled out theories that rely heavily on under-extrapolation of the present to the future, whether in the simple PE form of under-estimating the persistence of an exogenous fundamental or in the related GE forms of cognitive discounting and level-K thinking. These mechanisms are at odds both with the dynamic overshooting of the average forecasts documented here and with the over-reaction of individual forecasts documented in Bordalo et al. (2018) and Kohlhas and Broer (2019).

[7] These shocks are described at the end of Section 2 and are obtained from Angeletos, Collard, and Dellas (2019).

[8] This echoes lessons from Angeletos and Lian (2018), Angeletos and Huo (2019), and Farhi and Werning (2019).

The same is true for adaptive expectations insofar as the latter means systematic anchoring of current expectations to past outcomes. Adaptive expectations can generate a similar "stickiness" or sluggishness in the response of average forecasts to aggregate shocks as that generated by dispersed, noisy information. But only the latter helps account for why such stickiness is absent in the response of individual forecasts to individual news, or why individual forecast errors are predictable by the past information of others. This echoes a broader lesson of our analysis, which is to highlight how the similarities or differences of the properties of the individual and average forecast errors help disentangle mechanisms.

Over-extrapolation in finance and macro. Our main empirical finding echoes a literature in finance documenting a similar pattern—slow initial reaction and subsequent over-reaction—in individual stock prices (De Bondt and Thaler, 1985; Cutler, Poterba, and Summers, 1991; Lakonishok, Shleifer, and Vishny, 1994). Theoretical work such as Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), and Hong and Stein (1999) provide parsimonious interpretations which combine tentative initial reactions with medium-run over-reaction due to over-extrapolation. More recently, Greenwood and Shleifer (2014) and Gennaioli, Ma, and Shleifer (2015) demonstrate patterns in survey expectations of stock returns and firm earnings that are also suggestive of over-extrapolation.

We complement these works in three ways. First, we provide the first, to the best of our knowledge, evidence of over-extrapolation in expectations of unemployment and inflation. Second, we propose and implement a new empirical strategy, in terms of the IRFs of forecast errors to identified aggregate shocks, and explain why this strategy is best suited to guide theory. And third, we show how to combine over-extrapolation and dispersed, noisy information in a GE setting. Both our empirical strategy and our GE tools could find applications in finance in the future.

Other related literature. We distill the essence of a diverse set of theories of expectation formation, and use survey evidence to evaluate their potential relevance for business cycles. But we do *not* address related laboratory evidence (e.g., Nagel, 1995; Dean and Neligh, 2017; Landier, Ma, and Thesmar, 2019) and field experiments (Coibion, Gorodnichenko, and Kumar, 2018; Coibion, Gorodnichenko, and Ropele, 2019).

We also leave out of the analysis a variety of other plausible theories, which help explain different types of data. These include wishful thinking (e.g., Brunnermeier and Parker, 2005; Caplin and Leahy, 2019); over-weighting of personal experience (e.g., Malmendier and Nagel, 2016; D'Acunto, Malmendier, Ospina, and Weber, 2019; Das, Kuhnen, and Nagel, 2020); heterogeneous priors (e.g., Caballero and Simsek, 2017; Geanakoplos, 2010); adaptive learning (e.g., Eusepi and Preston, 2011; Evans and Honkapohja, 2001; Sargent, 2001); uncertainty shocks (e.g., Bloom, 2009a; Baker, Bloom, and Davis, 2016); robustness and ambiguity (e.g., Hansen and Sargent, 2012; Ilut and Schneider, 2014; Bhandari, Borovička, and Ho, 2019); non-Bayesian belief contagion (e.g., Carroll, 2001; Burnside, Eichenbaum, and Rebelo, 2016); and other plausible departures from the fully rational model (e.g., Gabaix, 2019; Molavi, 2019; Woodford, 2018).

Finally, we do not address questions relating to *optimal* macroeconomic policy in the presence of informational frictions (e.g., Angeletos and La'O, 2020; Lorenzoni, 2010; Paciello and Wiederholt, 2014) and mis-specified beliefs (e.g., Adam and Woodford, 2012; Angeletos and Sastry, 2020; Caballero and Simsek, 2019; Gabaix, 2020).

2 Data and Measurement

We focus on two macroeconomic outcomes: unemployment and inflation. We now review the exact data sources we use for forecasts and realized outcomes of these variables.

Forecasts from the Survey of Professional Forecasters. Our main dataset for forecasts is the Survey of Professional Forecasters (SPF), a panel survey of about 40 experts from industry, government, and academia, currently administered by the Federal Reserve Bank of Philadelphia. Every quarter, each survey respondent is asked for point-estimate projections of the civilian unemployment rate and the GDP deflator, among several macro aggregates. Our main sample runs from 1968.Q4 to 2017.Q4.

Whenever our analysis requires requires aggregate (or "consensus") forecasts, we use the *median* forecast of the object of interest (e.g., unemployment or inflation at a given horizon). Using the median instead of the mean in standard in the related empirical literature. The rationale is that it alleviates concerns about outliers and/or data-entry errors, which could be quite influential in the 40-forecaster cross section, from driving the results. That said, our main empirical finding is robust to using the mean instead of the median.

For the individual-level results, where concerns about outliers are even more relevant, we always trim observations in forecast errors and revisions that are plus or minus 4 times the inter-quartile range from the median, where both reference values are calculated over the entire sample.[9]

Other survey sources. Although our main analysis focuses on the SPF, we provide corroborating evidence from two additional survey datasets. The first is the Blue Chip Economic Indicators Survey, a privately-operated professional forecast with a similar scale and scope to the SPF. We use Blue Chip data from 1980 to 2017 and focus on the reported "consensus forecast" for unemployment and GDP deflator.[10] The second source is the University of Michigan Survey of Consumers, which is (for our purposes) a repeated cross-section of about 500 members of the "general public" contacted by phone. Like with the Blue Chip survey, we focus on end-of-quarter waves. We take the Michigan survey inflation forecast as the median response to the question about price increases.[11] We code also a forecast for the growth rate of unemployment based on a question about whether unemployment will increase or decrease over the coming twelve months.[12] For this measure we take the cross-sectional mean, which corresponds to a "consensus forecast" about the sign of the growth rate of unemployment.

[9] For context, in a Gaussian distribution, the probability of an observation so far in the tails is about 6.8×10^{-8}. Nonetheless, in the sample of three-quarter ahead inflation forecast errors, there are 57 such observations out of 7,438 forecaster-quarter observations, or about 10^6 times the aforementioned probability. All of these outliers involve forecast errors greater than 5.37 percentage points and often appear to be typos (an extra digit).

[10] This dataset is available at the monthly frequency, so we use end-of-quarter forecasts (i.e., those made in March, June, September, and December) for comparability with the SPF. The reported inflation forecasts in the Blue Chip consensus are actually quarter-to-quarter, so we construct the consensus estimate of longer-horizon inflation as the "chained consensus" rather than the "consensus of chained inflation."

[11] The exact question is the following: "By about what percent do you expect prices to go (up/down) on the average, during the next 12 months?" Respondents can key in a response rounded to the nearest whole number.

[12] The exact question is the following: "How about people out of work during the coming 12 months. Do you think that there will be more unemployment than now, about the same, or less?" There are three responses, as indicated in the question.

Macro data (and vintages thereof). Our unemployment measure u_t is the average BLS unemployment rate in a given quarter t. Our inflation measure π_t is the annualized percentage increase in GDP or GNP deflator over the four quarters up to t.[13] For the corresponding forecast data, our "default" choice of horizon is $k = 3$, in line with the main specification of CG, but we explore other choices for robustness.

In our replication of CG, BGMS and KW in Section 3, we use *first-vintage* macro data for consistency with these works.[14] However, such measurement is not necessarily the right one vis-a-vis theory. If agents are forecasting the *actual* levels of unemployment and inflation, the econometrician should use the *final-release* data. We will thus verify the robustness of the relevant facts to the use of final-release data.

We finally use *final-release* data in our study of IRFs in Section 5 both for the above reason and for consistency with the main macro time-series literature. But once again, we consider the opposite measurement (in this case, first-vintage data in place of final-release data) for robustness.

Shocks. Our study of IRFs requires the use of identified shocks. For our main exercises, we borrow two such shocks from Angeletos, Collard, and Dellas (2019): their "main business cycle shock," which accounts for the bulk of the business-cycle co-movements in unemployment, hours worked, output, consumption, investment; and a nearly-orthogonal shock that accounts for most of the fluctuations in inflation. A description of these shocks and the rationale for using them are provided in Section 5. For robustness, we also consider other, more "standard," shocks, such as a technology shock identified as in Galí (1999).

3 A Puzzling Empirical Backdrop: Under-reaction or Over-reaction?

This section reviews three stylized facts about macroeconomic forecasts. One of them suggests that expectations under-react to news. The other two point in the opposite direction. The apparent contradiction paves the way for the theoretical exercise and the empirical strategy we undertake in the subsequent sections: we will eventually argue that there is a "better" way to think about the issue both in the theory and in the data.

3.1 Fact 1: *under*-reaction in average forecasts

Coibion and Gorodnichenko (2015), henceforth CG, test for a departure from full-information rational expectations by estimating the predictability of professionals' aggregate ("consensus") forecast errors using information in previous forecast revisions.

Let $\bar{\mathbb{E}}_t[x_{t+3}]$ denote the median expectation of variable x_{t+3} (either unemployment or inflation) measured at time t. Let $\bar{\mathbb{E}}_{t-1}[x_{t+3}]$ be the median forecast at time $t-1$.[15] The associated *forecast error* from time t is $\text{Error}_t \equiv x_{t+3} - \bar{\mathbb{E}}_t[x_{t+3}]$, suppressing notation for the variable x and the forecast horizon, and the

[13] The ambiguity between GDP and GNP matches the fact that the Survey of Professional Forecasters changed its main target variable from GNP (and the deflator thereof) to GDP (and the deflator thereof) starting in 1992.

[14] We take all vintage data series from the Philadelphia Fed's website: https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/

[15] In the data, we prefer to use the median to limit the influence of outliers and/or data entry errors. But results with the mean are essentially identical. In the theory, means and medians coincide because we let all variables and signals be Normally distributed.

Table 1: Predicting Aggregate Forecast Errors with Revisions, from (1)

	(1)	(2)	(3)	(4)
	Unemployment		Inflation	
	1968-2017	1984-2017	1968-2017	1984-2017
Revision_t (K_{CG})	0.741	0.809	1.528	0.292
	(0.232)	(0.305)	(0.418)	(0.191)
R^2	0.111	0.159	0.278	0.016
N	191	136	190	135

Notes: The dataset is the Survey of Professional Forecasters and the observation is a quarter between Q4-1968 and Q4-2017. All regressions include a constant. The forecast horizon is 3 quarters. Standard errors are HAC-robust, with a Bartlett kernel and lag length equal to 4 quarters. The data used for outcomes are first-release ("vintage").

forecast revision is $\text{Revision}_t \equiv \bar{\mathbb{E}}_t[x_{t+3}] - \bar{\mathbb{E}}_{t-1}[x_{t+3}]$. CG run the following regression that projects aggregate forecast errors onto aggregate forecast revisions:

$$\text{Error}_t = \alpha + K_{\text{CG}} \cdot \text{Revision}_t + u_t \quad (1)$$

where K_{CG}, in shorthand notation that references the authors, is the main object of interest.

Table 1 reports results from estimating (1) at the horizon $k = 3$ for both unemployment and inflation in our data. We report results over the full sample 1968-2017 (columns 1 and 3), and also over a restricted sample after 1984 (columns 2 and 4). We may believe *a priori* that the latter is a more consistent and "stationary" regime for the US macroeconomy (i.e., after the oil crisis and Volcker disinflation).

Like the original authors, we find in all specifications a point estimate of $K_{\text{CG}} > 0$: when professional forecasters, in aggregate, revise upward their estimation of unemployment or inflation, they on average always "undershoot" the eventual truth. For inflation, we find the predictability is considerably lower on the restricted sample, which underscores the large influence of the aforementioned key events for US inflation expectations. Appendix Table A.1 shows robustness along a number of dimensions including (i) using different forecast horizons; (ii) putting final release data in place of the vintage data; and (iii) using forecasts from the Blue Chip Economic Indicators survey. All findings, including the differences across older and newer samples, are very similar to those reported in Table 1.

The finding of $K_{\text{CG}} > 0$ rejects full-information rational expectations: since $\text{Revision}_{t,k}$ is necessarily known to the representative agent at time t, it should not be systematically predict that agent's forecast error at $t+1$ if that agent is rational.[16] But note that it provides ambiguous evidence on the separate hypotheses of *informational frictions* versus *non-rationality*. In particular, the fact is just as consistent with a population of rational but heterogeneously informed agents (as indeed Coibion and Gorodnichenko, 2015, propose in their paper) as it is with a representative *irrational* agent who systematically under-reacts to

[16] An auxiliary assumption in this context, which we will not question throughout the analysis, is "perfect recall": a rational, Bayesian agent who forgets past information (like last period's forecast) could make such a predictable error.

Table 2: Predicting Individual Forecast Errors with Revisions, from (2)

	(1)	(2)	(3)	(4)
	Unemployment		Inflation	
	1968-2017	1984-2017	1968-2017	1984-2017
$\text{Revision}_{i,t}$ (K_{BGMS})	0.321	0.398	0.143	-0.263
	(0.107)	(0.149)	(0.123)	(0.054)
R^2	0.028	0.052	0.005	0.025
N	5,383	3,769	5,147	3,643

Notes: The observation is a forecaster by quarter between Q4-1968 and Q4-2017. The forecast horizon is 3 quarters. Standard errors are clustered two-way by forecaster ID and time period. Both errors and revisions are winsorized over the sample to restrict to 4 times the inter-quartile range away from the median. The data used for outcomes are first-release ("vintage").

news because of a behavioral bias (as indeed Gabaix, 2020, proposes in his own paper). Similarly, an "old-fashioned" model of adaptive expectations can also generate the fact. It is only by combining this fact with the additional fact reported next that we can start disentangling the role of informational frictions and mis-specified beliefs.

3.2 Fact 2: *over*-reaction in individual forecasts

To probe further the need for irrationality to explain the data, recent papers by Bordalo et al. (2018), Fuhrer (2018), and Kohlhas and Broer (2019) have studied forecast error patterns at the individual level in the professional forecasts. Let $\text{Error}_{i,t} \equiv x_{t+3} - \mathbb{E}_{i,t}[x_{t+3}]$ and $\text{Revision}_{i,t,} \equiv \mathbb{E}_{i,t}[x_{t+3}] - \mathbb{E}_{i,t-1}[x_{t+3}]$ denote forecast errors and revisions for a particular forecaster, indexed by i, at the baseline horizon $k = 3$. Each of the aforementioned studies estimates the following regression that translates (1) to the individual level:

$$\text{Error}_{i,t} = \alpha + K_{\text{BGMS}} \cdot \text{Revision}_{i,t} + u_{i,t} \tag{2}$$

where the object of interest K_{BGMS}, named in shorthand reference to the authors of Bordalo et al. (2018), is the individual-level analogue to K_{CG}. Regardless of the information structure, individual-level rationality imposes $K_{\text{BGMS}} = 0$.

In columns 1 and 3 of Table 2, we provide estimates of the individual-level regression (2) in the SPF over the full sample for our two variables of interest, unemployment and inflation. Columns 2 and 4 of the same table conduct the analysis on the sub-sample from 1984 to the present. Results for different horizons and data choices (vintage versus final) are similar and reported in Appendix Table A.2.

For unemployment, we find substantial evidence that $K_{BGMS} > 0$ over the full and restricted sample period. And for inflation, we find imprecise evidence that $K_{\text{BGMS}} > 0$ over the full sample, which includes the 1970s and Volcker disinflation, but strong evidence of $K_{\text{BGMS}} < 0$ in the "more stationary" environment post 1984.

Table 3: Predicting Aggregate Forecast Errors with Recent Outcomes, from (3)

	(1)	(2)	(3)	(4)
	Unemployment		Inflation	
	1968-2017	1984-2017	1968-2017	1984-2017
x_t (K_{KW})	-0.061	-0.036	0.111	-0.068
	(0.056)	(0.038)	(0.075)	(0.068)
R^2	0.016	0.007	0.058	0.012
N	194	136	193	135

Notes: The dataset is the Survey of Professional Forecasters and the observation is a quarter between Q4-1968 and Q4-2017. All regressions include a constant. The forecast horizon is 3 quarters. Standard errors are HAC-robust, with a Bartlett kernel and lag length equal to 4 quarters. The data used for outcomes are first-release ("vintage").

BGMS argue that a negative relation between revisions and subsequent errors, or $K_{\text{BGMS}} < 0$, is a robust feature of the forecasts of various macroeconomic variables. A closer look at their findings yields a more nuanced picture. But if we take for granted their thesis, we have that macroeconomic forecasts appear to over-react at the individual level at the same that they appear to under-react at the aggregate level.

We reinforce this apparent contradiction below. But we also invite the reader to keep the following basic insight in mind: while the CG evidence confounds the effects of informational frictions and non-rationality, the BGMS evidence speaks exclusively to the latter. We will leverage on this insight later to argue that the *gap* between the CG and the BGMS evidence speaks to the role of informational frictions.

3.3 Fact 3: *over*-reaction in aggregate forecasts

Facts 1 and 2 by themselves may suggest that going to the individual-level data is necessary, if not sufficient, to see evidence of over-reaction. But a recent paper by Kohlhas and Walther (2018) calls into question this view by presenting an additional moment: the slope of forecast errors in current *realizations* of the variable, as measured in the following regression:

$$\text{Error}_t = \alpha + K_{\text{KW}} \cdot x_t + u_t \tag{3}$$

In our implementation, Error_t is the error in three-quarter-ahead forecast of unemployment or annual inflation, and x_t is the current (forecasting) period's realization of one or the other.

Clearly, $K_{\text{KW}} \neq 0$ is inconsistent with full-information rational expectations. It is also hard to square with the CG evidence. More heuristically, in a world of "sluggish" expectations, we may expect $K_{\text{KW}} > 0$, or a positive correlation between today's realization and the direction of the forecast error k periods out.

Table 3 reports results from estimating (3). For unemployment we find weak evidence supporting the hypothesis that $K_{\text{KW}} < 0$. The results for inflation depend once again on whether we want to consider data from the 1970s and early 80s. In the whole sample, the evidence is more supportive of $K_{\text{KW}} > 0$. But in the more recent sample period, for inflation too we find weak evidence of $K_{\text{KW}} < 0$. Appendix Table A.3

probes robustness to different data choices and sub-samples and uncovers broadly consistent results. KW provide evidence of $K_{\text{KW}} < 0$ for forecasts of other variables, such as GDP growth.[17]

All in all, there is a good case for $K_{\text{KW}} < 0$ in the data. This is consistent with a world of over-reactive expectations: as an example, if agents are forecasting unemployment to be too high in recessions (high x_t, negative forecast error) and too low in booms (low x_t, positive forecast error), then we may naturally get $K_{\text{KW}} < 0$. But in such a world we would also expect $K_{\text{CG}} < 0$, which is not what we found earlier. This reinforces the puzzle: the picture for over- or under-reaction is unclear even if we focus on the properties of aggregate forecasts.

4 A Simple Model

In this section we introduce a simplified version of our framework, which combines dispersed noisy information with misspecified beliefs but abstracts from the fixed point between expectations and outcomes. We use this to reconcile Facts 1-3, but also, and more importantly, to pave the way to our preferred empirical strategy, which we in turn implement in the next section.

4.1 Primitives

Let $\{x_t\}$ be a stochastic process that a group of agents, indexed by $i \in [0,1]$, are trying to forecast (e.g., unemployment or inflation). Ideally we want to think of x_t as endogenous to the agents' behavior. But for now, to put the focus only on the expectations formation process, we assume that x_t follows an exogenous AR(1) process with Gaussian errors. That is,

$$x = \frac{1}{1-\rho\mathbb{L}}\varepsilon_t, \tag{4}$$

where $\rho \in (0,1)$ parameterizes the persistence of the process, $\varepsilon_t \sim N(0,1)$ is a Gaussian innovation, and $\mathbb{L}$ denotes the lag operator (i.e., $\mathbb{L}x_t = x_{t-1}$).

An agent's observation of x_t is contaminated with idiosyncratic noise. That is, each agent in period t observes a signal $s_{i,t}$ given by

$$s_{i,t} = x_t + \frac{u_{i,t}}{\sqrt{\tau}}, \tag{5}$$

where τ measures precision and $u_{i,t} \sim_{iid} N(0,1)$ is idiosyncratic Gaussian noise. As in a large literature, we can think of this noise either literally, as the product of dispersed noisy information (Lucas, 1972; Morris and Shin, 2002), or metaphorically, as a representation of rational inattention and imperfect perception (Sims, 2003, 2010; Woodford, 2003; Mankiw and Reis, 2002).

[17] One discrepancy between our implementation of KW and the original one is that these authors apply an HP filter to x_t. We prefer not to do so because it complicates the mapping to the theory: as the filtered value of x_t is a function of realizations after t, finding $K_{\text{KW}} \neq 0$ does not necessarily reject full-information, rational expectations under their approach, whereas it does under ours. That said, the big picture is the same. And in Subsection 5.4 and Table 4 we will use an instrumental-variables method that conditions on pre-determined data (identified shocks) to achieve a similar goal of extracting the business cycle component of variation in x_t.

We depart from this literature by adding two forms of irrationality, or belief misspecification. First, whereas the true process of the private signal is given by (5), agents perceive this process to be

$$s_{i,t} = x_t + \frac{u_{i,t}}{\sqrt{\hat{\tau}}} \tag{6}$$

for some *perceived* precision $\hat{\tau} > 0$ that may differ from τ. And second, whereas the true process from x_t is given by (4), agents perceive this process to be

$$x_t = \frac{1}{1-\hat{\rho}\mathbb{L}}\varepsilon_t \tag{7}$$

for some *perceived* persistence $\hat{\rho}$ which may differ from ρ.

The case $\hat{\tau} > \tau$ captures overconfidence: each agent thinks their information is better than it truly is. The opposite case, $\hat{\tau} < \tau$, captures underconfidence. Moore and Healy (2008) provide a representative review of the experimental psychological evidence for such biases. Their broad conclusion is that overconfidence is consistently prevalent for reported beliefs in the laboratory, but that the extent of effects can be context-specific. Kohlhas and Broer (2019) and Bordalo et al. (2018) use, respectively, $\hat{\tau} > \tau$ and a close variant of it to reconcile Facts 1 and 2.[18] We will nest this possibility in the subsequent analysis but also show that $\hat{\rho} > \rho$ serves the same goal while also matching Facts 3. And we will provide additional evidence in favor of $\hat{\rho} > \rho$ in the form of our (not yet introduced) Fact 4 about dynamic over-shooting.

The case $\hat{\rho} > \rho$ encodes an over-extrapolation of today's state to tomorrow, while $\hat{\rho} < \rho$ encodes under-extrapolation. Both narratives are appealing in different economic contexts. On the one hand, Greenwood and Shleifer (2014) and Gennaioli, Ma, and Shleifer (2015) argue that over-extrapolation is evident both in stock-market expectations and in expectations of firms' sales forecasts; see also Guo and Wachter (2019) for how a simple model with over-extrapolation over dividend growth can explain a variety of asset-price phenomena. On the other hand, level-K thinking (Garcıa-Schmidt and Woodford, 2019; Farhi and Werning, 2019) and cognitive discounting (Gabaix, 2020) are "close in spirit" to the opposite scenario, $\hat{\rho} < \rho$, because they cause agents to be under-estimate the (endogenous or exogenous) response of future outcomes to current innovations. We will make this connection formal in Section 6.4, once we extend the analysis to a GE context and properly nest these models.

As anticipated in the Introduction and will become clear in the sequel, only the second type of misspecification ($\hat{\rho} \neq \rho$) is strictly needed for our main purposes. The first type ($\hat{\tau} \neq \tau$) is nevertheless useful for two complementary reasons: it enlarges the set of theories nested in, or proxied by, our framework; and it helps clarify which evidence is most directly relevant in the GE context of Section 6.

4.2 Facts 1, 2, and 3 in the Model

The structure introduced above yields a highly tractable, finite ARMA representation of the individual and average forecasts, which can be found in Lemma 1 in the Appendix. This in turns allows a simple, closed-form characterization of the theoretical counterparts of the regressions reviewed in Section 3.

[18] The variant used in Bordalo et al. (2018) is motivated by a broader concept, "diagnostic expectations," the precise formal content of which varies across applications. It is the *specific* formalization employed in Bordalo et al. (2018) that is very similar to over-confidence; this similarity is evident in the modified Kalman filter that is at the core of that paper.

Proposition 1 (Regression coefficients in the theory)**.** *The theoretical counterparts of the coefficients of regressions* (1), (2) *and* (3) *are given by the following:*

$$K_{CG} = \mathcal{K}_{CG}(\hat{\tau},\rho,\hat{\rho}) \equiv \kappa_1 \hat{\tau}^{-1} - \kappa_2(\hat{\rho}-\rho) \tag{8}$$

$$K_{KW} = \mathcal{K}_{KW}(\hat{\tau},\rho,\hat{\rho}) \equiv \kappa_3 \hat{\tau}^{-1} - \kappa_4(\hat{\rho}-\rho) \tag{9}$$

$$K_{BGMS} = \mathcal{K}_{BGMS}(\tau,\hat{\tau},\rho,\hat{\rho}) \equiv -\kappa_5(\hat{\tau}-\tau) - \kappa_6(\hat{\rho}-\rho) \tag{10}$$

for some scalars $\kappa_1,...,\kappa_6$ that depend on the deeper parameters but are necessarily positive. In particular, $\kappa_1,\kappa_2,\kappa_3$, and κ_4 are functions only of $(\hat{\tau},\rho,\hat{\rho})$, whereas κ_5 and κ_6 depend also on τ.

Let us unpack these expressions. First of all, note that the actual precision, τ, enters only the BGMS coefficient. That is, the moments of the average forecasts do *not* depend on the true level of noise, conditional on the perceived noise. The latter dictates how each agent's forecasts responds to her information, and hence also how the average forecasts respond to the underlying shocks. The actual idiosyncratic noise, instead, washes out at the aggregate level.

Consider next condition (8), which characterizes the CG coefficient. With rational expectations, which herein means $\hat{\tau} = \tau$ and $\rho = \hat{\rho}$, K_{CG} is merely a monotone transformation of the level of noise. In particular,

$$K_{\text{CG}} = \frac{1-g}{g} \tag{11}$$

where $g \in (0,1)$ is the Kalman gain.[19] This is the structural interpretation given in CG.

Our result qualifies this structural interpretation in two ways. First, if we maintain $\rho = \hat{\rho}$ but allow $\hat{\tau} \neq \tau$, we get an analogue of (11) with a *subjective* Kalman gain $\hat{g}$ in place of its objective counterpart. That is, even in the absence of over-extrapolation, the CG coefficient tells us something about the *subjective* level of noise, which does not have to coincide with the *objective* level. Second, if we allow $\hat{\rho} \neq \rho$, we now have that K_{CG} confounds two mechanisms: a high value for K_{CG} could be evidence of either large informational friction or large under-extrapolation. Or, a low value for K_{CG} could hide a large information friction if there is also large over-extrapolation. Indeed, K_{CG} could even be *negative.*

Consider next condition (9), which characterizes the KW coefficient. The informational friction and the over-extrapolation enter this coefficient in a qualitatively similar was they enter the CG coefficient. The former contributes towards $K_{\text{KW}} > 0$, the latter towards $K_{\text{KW}} < 0$. The logic is *exactly* the same as that for the CG coefficient. What is subtle is the possibility that the two forces balance out in such a way that the one coefficient is negative at the same time that the other is positive, a point we revisit below.

Finally, consider condition (10), which characterizes the BGMS coefficient. When $\hat{\tau} = \tau$ and $\rho = \hat{\rho}$, $K_{\text{BGMS}} = 0$. This is an example of the more general property that, under rational expectations, an individual's forecast error is unpredictable by his own past information. Away from this benchmark, both overconfidence ($\hat{\tau} > \tau$) and over-extrapolation ($\hat{\rho} > \rho$) contribute towards $K_{\text{BGMS}} < 0$. In the presence of over-extrapolation, agents over-estimate the effect of any given innovation today on future outcomes. In

[19]As in the textbook version of the Kalman filter, g is such that $\mathbb{E}_{i,t}[z_t] = (1-g)\mathbb{E}_{i,t-1}[z_t] + g s_{i,t}$ and is an increasing and continuous function of τ, with $g \to 0$ as $\tau \to 0$ and $g \to 1$ as $\tau \to \infty$.

the presence of over-confidence, they get this effect right but over-estimate the precision of the signal they receive about the innovation. In both cases, they make a systematic mistake in the direction of over-estimating the informational content of their current signal about the future outcome, and this mistake manifests as $K_{\text{BGMS}} < 0$. The converse is true for underconfidence or under-extrapolation.

4.3 The right combination of belief distortions

Let us summarize the lesson for two versions of our model that are familiar from the literature but fail to match Facts 1-3:

Corollary 1. *The following two cases are inconsistent with Facts 1-3:*

(i) Noisy but rational expectations: $\hat{\tau} = \tau < \infty$ *and* $\hat{\rho} = \rho$ *implies* $K_{CG} > 0$, $K_{KW} > 0$, *and* $K_{BGMS} = 0$

(ii) Noiseless but extrapolative expectations: $\tau = \hat{\tau} \to \infty$ *and* $\hat{\rho} \neq \rho$ *implies* $sign(K_{CG}) = sign(K_{KW}) = sign(K_{BGMS}) = sign(\rho - \hat{\rho})$.

The first case stylizes a large literature on informational frictions, and is precisely the case considered in Coibion and Gorodnichenko (2015). This case counterfactually forces K_{CG} and K_{KW} to be the same sign, a restriction first pointed out in Kohlhas and Walther (2018), because there is only a single "dampening" force coming from noisy expectations. Moreover, this case cannot accommodate $K_{\text{BGMS}} \neq 0$ because forecasters remain individually rational.

The second model is an entirely "behavioral" one that admits a mis-calibrated representative agent, as in Gabaix (2016, 2020). But switching from an under-extrapolative model ($\hat{\rho} < \rho$) to an over-extrapolative model ($\hat{\rho} > \rho$) must necessarily flip *all three signs* for the aforementioned moments. Thus it too cannot match the patterns observed so far.

Let us now turn to the scenario that best accounts for the evidence.

Corollary 2 (Matching Facts 1-3)**.** *The combination of informational friction and over-extrapolation is necessary and sufficient for all three facts in the following sense:*

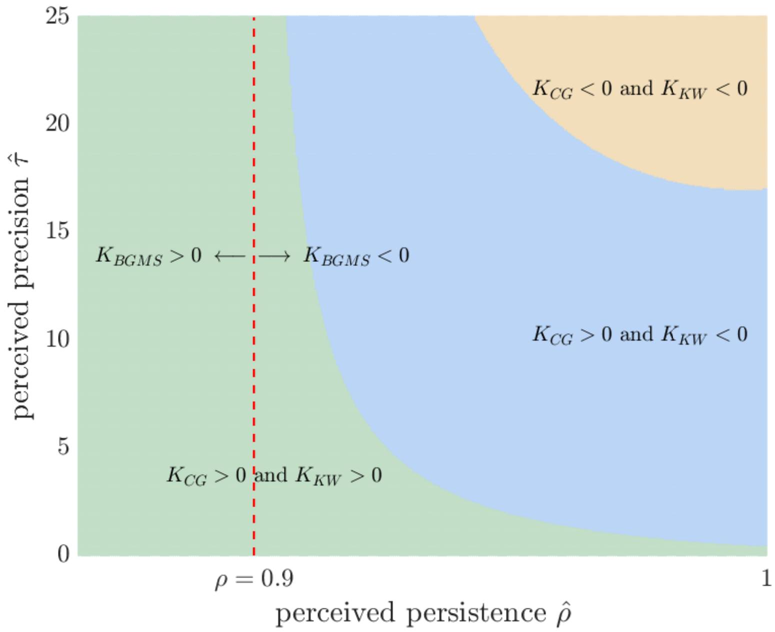
(i) $K_{CG} > 0$ *and* $K_{KW} < 0$ *only if* $0 < \hat{\tau} < \infty$ *and* $\hat{\rho} > \rho > 0$,

(ii) There exists an open set of parameter values, with $0 < \tau \leq \hat{\tau} < \infty$ *and* $\hat{\rho} > \rho > 0$, *such that* $K_{CG} > 0$, $K_{KW} < 0$, *and* $K_{BGMS} < 0$.

Figure 1 illustrates this by plotting the model's "sign predictions" for the three coefficients in the $(\hat{\tau}, \hat{\rho})$ space. For this picture, we set $\rho = 0.90$, which is illustrative but immaterial to the overall pattern. We also restrict $\tau = \hat{\tau}$, that is, we assume away both over- and under-confidence. The blue region identifies the combinations of $\hat{\tau}$ and $\hat{\rho}$ that match all three facts qualitatively.

What happens if we let $\tau \neq \hat{\tau}$? The green, blue and orange regions remain intact, and so does the mapping from the specific values of K_{CG} and K_{KW} to the corresponding values of $\hat{\tau}$ and $\hat{\rho}$. This is because the stochastic properties of the average forecasts depend merely on the perceived level of noise and the degree of over-extrapolation, not only the actual level. What changes as we vary τ, or equivalently the degree of over-confidence, is only the position of the vertical red line, and along with it the specific value of this "free" parameter needed to match a specific value for K_{BGMS}.

Figure 1: The Regression Coefficients K_{CG}, K_{KW}, and K_{BGMS} in the Theory



This suggests a simple, recursive, identification strategy: first, calibrate ρ to actual process of unemployment or inflation; next, identify $\hat{\tau}$ and $\hat{\rho}$ jointly from K_{CG} and K_{KW}; finally, identify τ form K_{BGMS}. Appendix Table A.4 implements this strategy and reports the specific values of the model parameters that *quantitatively* match the evidence reported before. But both this identification strategy and Corollary 2 suffer from the same basic problem: it is unclear *how* the theory produces at once under-reaction in the sense of $K_{\text{CG}} > 0$ and over-reaction in the sense of $K_{KW} < 0$. We cut the Gordian knot in the next subsection by proposing a different, more transparent, way of connecting the theory and the data.

We end this subsection with a bibliographical note. Kohlhas and Walther (2018) offer a different resolution to Facts 1 and 3 (i.e., $K_{\text{CG}} > 0 > K_{\text{KW}}$) than that presented. This alternative preserves rational expectations by allowing for asymmetric attention to procyclical and countercyclical components of the forecasted outcome. But it imposes $K_{\text{BGMS}} = 0$, failing to match Fact 2, and it does not square with Fact 4, the new evidence we provide in Section 5. Kohlhas and Broer (2019) match Fact 2 by introducing over-confidence, and Bordalo et al. (2018) achieve the same with a variant ("diagnostic expectations") that is formally similar to over-confidence. But neither of these papers addresses Fact 3 and 4.

4.4 A more informative approach: impulse response functions (IRFs)

Our intuition about the various forces behind Facts 1-3, and particularly the tension between K_{CG} and K_{KW}, had a dynamic flavor which was collapsed to essentially static moments. Indeed our derivation of Proposition 1 quite literally involved starting with a moving-average form of each stochastic process and then computing static correlations. Let us now explore more directly what we would learn from observing directly the dynamic response of forecast errors in response to shocks.

Proposition 2 (IRF of Forecast Errors)**.** *Let $\{\zeta_k\}_{k=1}^{\infty}$ be the Impulse Response Function (IRF) of the average, one-step-ahead, forecast error. That is, for all $k \geq 1$,*

$$\zeta_k \equiv \frac{\partial\left(x_{t+k} - \overline{\mathbb{E}}_{t+k-1}[x_{t+k}]\right)}{\partial \varepsilon_t}$$

is the k-th coefficient in the moving-average representation of the average forecast error.[20]

(i) If $\hat{\rho} < \rho$, or agents under-extrapolate, then $\zeta_k > 0$ for all $k \geq 1$.

(ii) If $\hat{\rho} > \rho$ and $\hat{\tau}$ is small enough relative to $\hat{\rho} - \rho$, or agents over-extrapolate and learning is slow enough, then $\zeta_k > 0$ for $1 \leq k < k_{IRF}$ and $\zeta_k < 0$ for $k > k_{IRF}$, for some $k_{IRF} \in (1, \infty)$.

(iii) Finally, if $\hat{\rho} > \rho$ but $\hat{\tau}$ is large enough relative to $\hat{\rho} - \rho$, or agents over-extrapolate but learning is fast, then $\zeta_k < 0$ for all $k \geq 1$.

Corollary 3 (Delayed overshooting)**.** *The IRF of the average forecast errors starts positive but eventually switches negative if and only if there is both over-extrapolation and sufficiently large informational friction.*

A sign-switch in the impulse response of forecast errors to a macro shock is "smoking-gun" evidence for a combination of noise and over-extrapolation. A complementary lesson is that the point at which the sign-switch occurs provides a gauge of the relative importance of the two mechanism: the slower the learning relative to over-extrapolation, the longer it takes for the sign-switch to occur.

This is, in our view, easier to interpret than our previous strategy of comparing K_{CG} with K_{KW} because it gets to the heart of the economic question: at what point are economic agents sufficiently informed about an economic event (i.e., particular shock) such that their model mis-specification becomes the dominant explanation for any errors?

Figure 2 illustrates these patterns by plotting the IRFs of outcomes and forecasts (left column) and forecast errors and revisions (right column) in two scenarios: a benchmark without over-extrapolation (top row), and a variant with (bottom row). The key observation is that *only* with the combination of slow learning and over-extrapolation can the theory generate a sign reversal for the aggregate forecast errors, or average forecasts that undershoot initially and overshoot later on.

Now, to drive home the connection to K_{CG} and K_{KW}, consider the MA representations of the forecast errors, the forecast revisions, and the actual outcome:

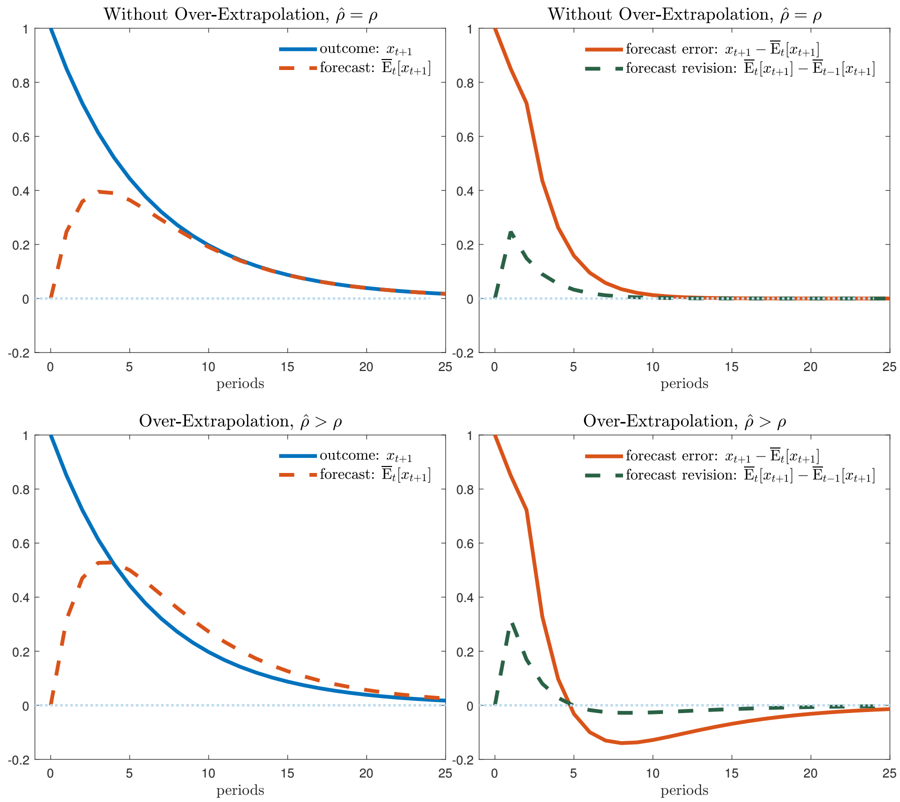
$$\text{Error}_{t,t+1} = \sum_{k=0}^{\infty} \zeta_k \varepsilon_{t+1-k} \qquad \text{Revision}_t = \sum_{k=0}^{\infty} f_k \varepsilon_{t-k} \qquad \text{Outcome}_t = \sum_{k=0}^{\infty} \rho^k \varepsilon_{t-k}.$$

where $\{\zeta_k\}$ and $\{f_k\}$ are the IRFs of, respectively, the average forecast errors and the average forecast revisions. Using these representations, the coefficient of regression (1) can be expressed as

$$K_{\text{CG}} = \frac{\text{Cov}(\text{Error}_{t,t+1}, \text{Revision}_t)}{\text{Var}(\text{Revision}_t)} = \frac{\sum_{k=0}^{\infty} \zeta_{k+1} \cdot f_k}{\sum_{k=0}^{\infty} {f_k}^2} \tag{12}$$

[20] We exclude ζ_0 from this statement because it is mechanically 1. Also, all our theoretical statements focus on one-step-ahead forecasts for expositional simplicity, but our empirical implementations of the theory use the exact counterparts of the objects constructed in the data (e.g., as the three-quarter-ahead forecasts of annualized inflation).

Figure 2: IRFs of Aggregate Forecasts and Errors in the Theory



and similarly the coefficient of regression (2) can be expressed as

$$K_{\mathrm{KW}} = \frac{\mathrm{Cov}(\mathrm{Error}_{t,t+1}, \mathrm{Outcome}_t)}{\mathrm{Var}(\mathrm{Outcome}_t)} = \frac{\sum_{k=0}^{\infty} \zeta_{k+1} \cdot \rho^k}{\sum_{k=0}^{\infty} \rho^{2k}} \tag{13}$$

This makes clear that K_{CG} and K_{KW} are, up to rescaling, equal to the dot-products of the IRF of the forecast errors with the IRF of, respectively, the revisions and the outcome.

What does this look like? Consider the bottom row of Figure 2, which corresponds to the combination of noise and over-extrapolation. The dynamic response of the forecast errors, or the red, solid IRF in the bottom-right subfigure, exhibits the reversal property we noted earlier: forecast errors switch from positive to negative after a while. A similar reversal is also present in the forecast revisions; see the green, dashed line in the same subfigure. It follows that the dot-product of these two IRFs contains more positive terms than the dot-product of either one of them with that of the outcome, which is given by the blue, solid line in the bottom-left subfigure. This helps explain why $K_{\mathrm{CG}} > 0$ at the same time that $K_{\mathrm{KW}} < 0$.

Apart from resolving the "mystery" behind the different signs of K_{CG} and K_{KW}, this exercise also underscores that the IRFs of the forecast errors contain strictly more information about the dynamic properties of beliefs than any of the these regression coefficients. Either one of these coefficients offers a confusing picture by averaging under- and over-reaction across different horizons. The IRFs let one see when exactly beliefs under-shoot and when exactly they over-shoot.

5 A New Fact: Delayed Over-shooting

We now go after what the theory has identified as the most useful moment to characterize imperfect expectations: the dynamic response of forecasts and forecast errors to shocks. Here we corroborate the hypothesis derived above and uncover a consistent pattern of *initial under-reaction and delayed over-shooting* in the response of forecast errors to shocks. This, at least in the context of the last section's analysis, is "smoking gun" evidence of a combination of noisy information and over-extrapolation.

5.1 Methodology

We start with the details of the empirical implementation.

Identified shocks. As anticipated in Section 2, we consider two empirical shocks, both borrowed from Angeletos, Collard, and Dellas (2019).[21]

The first shock, which these authors call "main business cycle shock," is constructed by maximizing its contribution to the business cycle variation in unemployment and is found to have the following properties: it encapsulates strong positive co-movement in employment, output, investment, and consumption *only* over the business cycle; it is nearly indistinguishable, in terms of IRFs and variance contributions, to the shocks identified by targeting any of the aforementioned variables; it has a negligible footprint on TFP at all horizons; it has a small to modest footprint on inflation. It can thus be interpreted as an non- or mildly-inflationary demand shock, which drives the bulk of the business cycle in the data.[22]

The second shock is identified by maximizing its contribution to the business cycle variation in inflation and it is found to have a negative but very small footprint on real quantities and zero footprint on TFP. It is thus akin to the kind of markup or cost-push shocks the DSGE literature uses to account for the bulk of the inflation fluctuations in the data.[23]

We denote the two shocks, respectively, as $(\varepsilon_t^D, \varepsilon_t^S)$ for "demand" and "supply". Whether these shocks, or any other SVAR-based shocks, are "truly" structural is largely a philosophical question and certainly beyond the scope of the present paper. For our purposes, the appeal of the particular shocks compared to others found in the literature (e.g., Galí, 1999; Sims and Zha, 2006) is that they drive a significant component of the business-cycle variation in macroeconomic activity and inflation. There is thus a good chance that they also drive a significant component of the corresponding variation in real-world expectations.

[21] The empirical strategy taken in that paper builds on the max-share approach (Uhlig, 2003; Barsky and Sims, 2011) but is guided by the following goal: providing a parsimonious representation of the business cycle in terms of one dominant shock. To this goal, Angeletos, Collard, and Dellas (2019) run a VAR on a set of ten or more key macroeconomic variables that includes the two variables we focus on here, the rate of unemployment and the rate of inflation. They then compile a collection of multiple shocks, each identified by maximizing its contribution to the volatility of a particular variable over a particular frequency band, and they draw lessons from comparing the empirical footprint of *all* these shocks.

[22] Angeletos, Collard, and Dellas (2019) further show that this shock in the data is closely related to the following counterparts in models: the investment-specific demand shock in Justiniano, Primiceri, and Tambalotti (2010), the risk shock in Christiano, Motto, and Rostagno (2014), and the confidence shock in Angeletos, Collard, and Dellas (2018).

[23] The two shocks are not constructed to be orthogonal to one another, but are very close to being so in the data.

Main specification: ARMA-IV. To estimate dynamic responses to the aforementioned shocks, we consider two different empirical strategies.

The first is to estimate the IRFs via a parsimonious, instrumental-variables ARMA(P,K) representation. In particular, we estimate the following regression:

$$z_t = \alpha + \sum_{p=1}^{P} \gamma_p \cdot z^{\text{IV}}_{t-p} + \sum_{k=0}^{K} \beta_k \cdot \varepsilon_{t-k} + u_t \tag{14}$$

Depending on the variable whose dynamic response we want to look at, z_t is the actual outcome (unemployment or inflation), the relevant forecast, or the corresponding forecast error. In all cases, $\varepsilon_t \in \{\varepsilon^D_t, \varepsilon^S_t\}$ is one of the aforementioned two shocks drawn from Angeletos, Collard, and Dellas (2019). Finally, for $p \in \{1, ..., P\}$, z^{IV}_{t-p} are the lagged values of z_t *instrumented* by the lagged values of ε_t.[24] This IV approach recovers the *conditional* dynamic responses to the structural shock under consideration—intuitively, how z_t moves when driven by the shock process of interest. We will call this method the "ARMA-IV" estimation.

By estimating (14) for outcomes (e.g., z_t equal to that quarter's unemployment rate or the past four quarters' inflation rate), we can generate dynamic impulse response coefficients $(\beta_{\text{out},h})_{h=0}^{H}$ as functions of $(\beta_0, (\gamma_p)_{p=1}^{P})$. For forecasts, we can do the same thing with z_t equal to the forecast in period t (e.g,. $\bar{\mathbb{E}}_t[u_{t+3}]$ and $\bar{\mathbb{E}}_t[\pi_{t+3,t-1}]$): estimate the impulse response coefficients $(\tilde{\beta}_{\text{fc},h})_{h=0}^{H}$ and then "re-index" these coefficients to line up with the realized outcomes. More specifically, we generate $(\beta_{\text{fc},h})_{h=0}^{H}$ such that $\beta_{\text{fc},h} = 0$ for $h < 3$ (effectively imposing unpredictability of the shocks), and $\beta_{\text{fc},h} = \tilde{\beta}_{\text{fc},h-3}$ for $h \geq 3$. Finally, we can construct the IRF of the forecast errors either by taking the difference between the IRF of the outcome and the forecasts, or by repeating the aforementioned procedure with x_t being the average forecast error.

In all cases, we construct standard errors for the coefficients that are heteroskedasticity and autocorrelation robust (HAC) with a 4-quarter Bartlett kernel; and then use the delta method to calculate standard errors for the impulse response functions. All reported error bands are 68% confidence intervals ($\pm 1 \cdot \text{SE}$).

Local projection. Our main strategy strives for parsimony by requiring the IRFs to accept a low-dimension ARMA representation as in (14). But we can also estimate impulse responses directly using the projection method of Jordà (2005). In this case, the estimating equation, for each horizon $0 \leq h \leq H$, is the following:

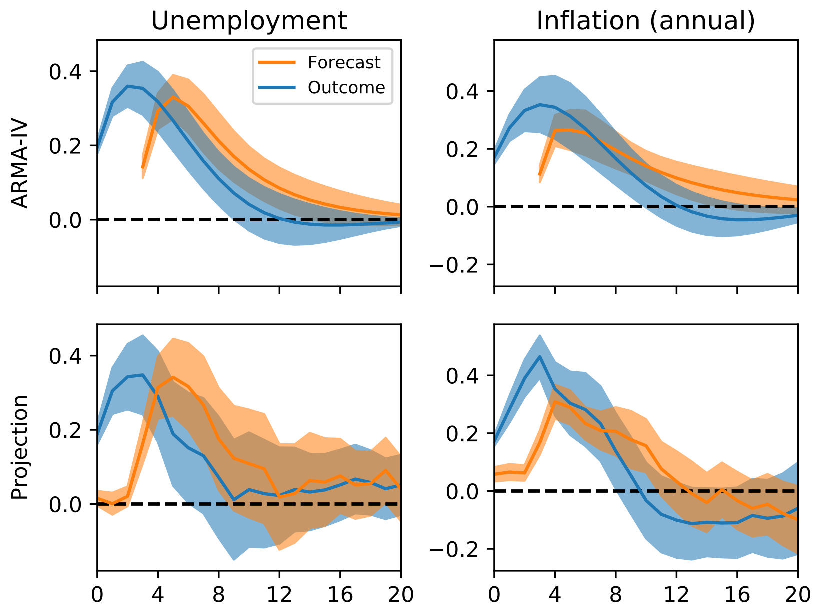
$$z_{t+h} = \alpha_h + \beta_h \cdot \varepsilon_t + \gamma' W_t + u_{t+h} \tag{15}$$

where $(\beta_h)_{h=0}^{H}$ trace out the dynamic response of the outcome, W_t is a vector of control variables, and γ are the coefficients on these controls. Consistently, across specifications, we include the lagged outcome x_{t-1} and the lagged forecast $\bar{\mathbb{E}}_{t-k-1}[x_{t-1}]$ as control variables. Conceptually, as long as these controls are orthogonal to the shock ε_t, these should not affect the population estimate we get of the impulse response parameters; but their inclusion may help with small-sample precision. We find overall that results are not sensitive to choices of controls. Standard errors are constructed in the same, aforementioned way.

Finally, we set $k = 3$ quarters as the forecast horizon, in line with what we did in Section 3, and we set $H = 20$ quarters as the maximum period for tracing out IRFs.

[24] The first-stage equation is given, in vector form, by $Z_{t-1} = \eta + \mathscr{E}'_{t-K-1}\Theta + e_t$, where $Z_{t-1} \equiv (z_{t-p})_{p=1}^{P}$, $\mathscr{E}_{t-K-1} \equiv (\varepsilon_{t-K-j})_{j=K+1}^{J}$ and $J - K \geq P$. Our main specifications use $P = 3$, $K = 1$, and $J = 9$ (i.e., 8 instruments for 3 regressors). But the results are robust to $P = 2$ and $P = 4$, as well as to different J and K.

Figure 3: Dynamic Responses: Outcomes and Forecasts



Notes: The sample period is Q1 1968 to Q4 2017. The shaded areas are 68% confidence intervals based on HAC standard errors with a Bartlett kernel and 4 lags. The x-axis denotes quarters from the shock, starting at 0. In the first column the outcome is u_t and the forecast is $\bar{\mathbb{E}}_{t-3}[u_t]$; in the second column the outcome is $\pi_{t,t-4}$, or annual inflation, and the forecast is $\bar{\mathbb{E}}_{t-3}[\pi_{t,t-4}]$.

5.2 The fact: dynamic over-shooting

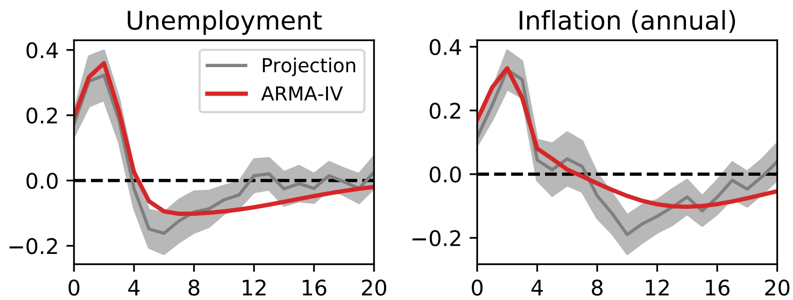
Figure 3 shows, in a two-by-two grid, the main impulse response estimates.[25] In the first column, we show the dynamic response of unemployment and median forecasts thereof to the demand shock ε_t^D. The first row shows the instrumented ARMA method of equation (14) , and the second row shows the projection method of (15). For both methods, we "align" the forecast responses such that, at a given vertical slice of the plot, the outcome and forecast responses are measured over the same horizon, and the difference thereof is a measure of the response of forecast errors. In the second column, we plot the same for the response of one-year-average inflation to the supply shock ε_t^S.

The consistent pattern across specifications is an initially delayed, and then over-persistent response of forecasts to the shock. Consider, as an illustration, the response of unemployment and forecasts thereof to ε_t^D. Unemployment spikes around quarter 3 in both estimation methods before reverting back to its long-run mean. The point-estimate is extremely close to zero by $t = 12$ in both cases.

Now consider the response of forecasts at $t = 3$ in the plot. These are forecasts *made* at $t = 0$, when the very first macro data (e.g., BLS reports) from $t = 0$ become available. Forecasted unemployment immediately spikes and begins to decay over the next 5-6 quarters. Forecasters remain convinced there are

[25] We report first-stage F statistics for the ARMA-IV estimates in Table A.5. These are low with respect to the reference values suggested by Stock and Yogo (2005) which is part of the reason that we also consider alternative estimation methods including the linear projection and multi-variate linear model.

Figure 4: Dynamic Responses: Forecast Errors



Notes: The sample period is Q1 1968 to Q4 2017. The shaded areas are 68% confidence intervals based on HAC standard errors with a Bartlett kernel and 4 lags. The x-axis denotes quarters from the shock, starting at 0. In the first column the outcome is u_t and the forecast is $\bar{\mathbb{E}}_{t-3}[u_t]$; in the second column the outcome is $\pi_{t,t-4}$, or annual inflation, and the forecast is $\bar{\mathbb{E}}_{t-3}[\pi_{t,t-4}]$.

adverse demand conditions, when in reality conditions have reverted back to the mean. A similar, and indeed more dramatic, pattern is visible in the response of inflation to the supply shock (second row). And these patterns look qualitatively and quantitatively quite similar with both the smooth, ARMA estimates (left column) and the unrestricted projection regression estimates (right column).

Figure 4 shows this overshooting pattern more clearly in terms of the impulse response of forecast errors. For both the ARMA and projection methods, this is obtained by taking the difference of the previous estimates for outcomes and forecasts. For both unemployment and inflation, we find evidence that forecast errors start positive and then turn negative at longer horizons. The estimated "crossing points" of the forecast errors response with 0, using the ARMA method, are $K^u_{\text{IRF}} = 4.14$ and $K^\pi_{\text{IRF}} = 6.43$, respectively.[26]

Finally, in the left panel of Appendix Figure A.1, we complete the picture with the "off-diagonal" impulse responses of inflation to the demand shock and unemployment to the supply shock. The former is weakly inflationary at longer horizons and the latter weakly contractionary at medium horizons. And in both cases we have modest evidence of the over-shooting pattern of interest.

5.3 Robustness and extensions

Sample choice. We conduct a number of initial robustness checks, mirroring those in Section 3, related to measurement and sample choice. The middle panel of Figure A.1 recreate the regression results in the SPF, back again with modern data, in the sample 1984-2017. As discussed previously, we might think of the post-Volcker and post-oil-crisis data as a more "consistently stationary" regime for forecasters trying to model the world. We find largely the same patterns in forecast errors. The right panel of Figure A.1 recreates the main analysis with forecast data from Blue Chip Economic Indicators over the shorter available sample (1980-2017) and again finds the same patterns.[27]

[26] The corresponding estimates from the projection regressions are 4.87 and 7.79.

[27] We also replicate all SPF and Blue Chip findings with vintage data and find similar results (not reported for brevity).

General public. Carroll (2003) and others have argued that the forecasts of professional forecasters are in general good proxies for those of the general public. But does this apply to our *particular* finding?

To address this question, we look at the University of Michigan Survey of Consumer Sentiment and construct an "unemployment expectation" using the survey's question about whether unemployment will go up, stay the same, or go down over the next 12 months. We code a variable $\bar{\mathbb{E}}_t[\text{UnempUp}_{t+4}]$ that averages the "up" responses, and code a data equivalent UnempUp_{t+4} using the BEA unemployment rate.[28] For inflation, we use the survey's estimate for inflation over the next 12 months.[29]

Appendix Figure A.2 shows the results from projecting our business cycle shocks on these variables using (15). The left panel shows the response of the UnempUp variable and forecasts thereof to ε_t^D. The Michigan survey expectations perk up slightly *before* the shock hits (i.e.,. for $t < 4$) and then spike one quarter "too late." We see further evidence that the general public is also particularly unable to forecast the "mean-reverting" part of the shock, or the eventual downward trend in unemployment.

The right panel shows the response of the response of GDP deflator and the annual inflation expectation of the Michigan survey to ε_t^S. Here, responses are much too noisy to pick out an obvious "peak response." Again, there is some weak evidence of anticipation, and at quarters 10 and onward evidence of some over-extrapolation of recent price trends.

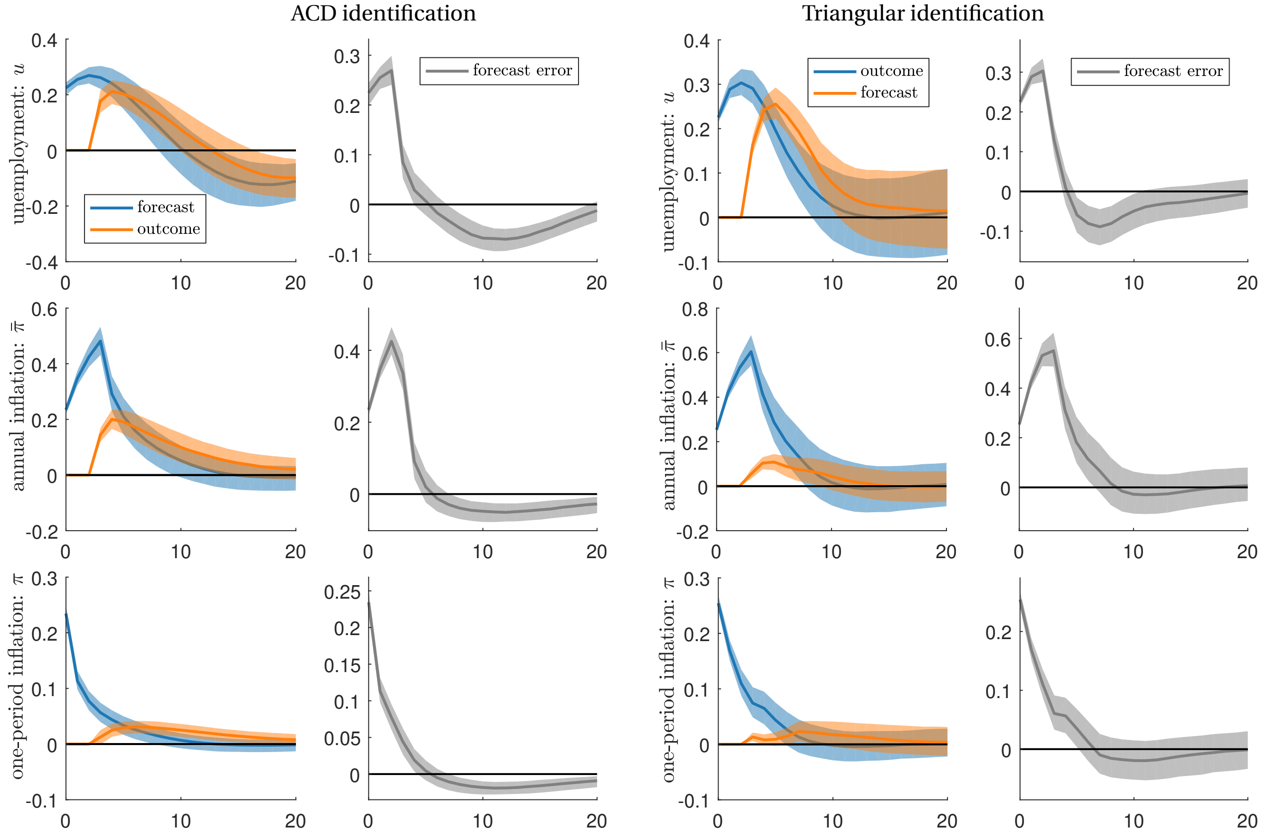
Other shocks of interest. An appealing feature of using projections and the ARMA-IV method is that it is easy to combine with auxiliary identification techniques, without fully specifying a multivariate model and considering the problem of jointly identifying many shocks. To illustrate this property, and probe the robustness of our results to other candidate "supply and demand" shocks from the macroeconomics literature, Appendix Figure A.3 replicates our main analysis for three different shocks: a technology shock à la Galí (1999), normalized here to be inflationary and contractionary; an oil price shock à la Hamilton (1996); and the investment-specific shock extracted from the DSGE model of Justiniano, Primiceri, and Tambalotti (2010). The former two are variations of "supply shocks" (to productivity or input costs), and we show the response of inflation; the last is like a demand shock, and we show the response of unemployment. In all cases we see evidence of the overshooting pattern.

Methods for estimating dynamics. Our ARMA-IV method resembles the method suggested by Romer and Romer (2004) and applied by Coibion and Gorodnichenko (2012) in their study of how forecast errors respond to structural shocks. That method estimates an empirical ARMA process like (14) via ordinary least squares ("ARMA-OLS"). It therefore uses *unconditional* auto-covariance properties to pin down dynamics. Our prior is that, in a world of very different, shock-specific dynamics (induced, for instance, by differential persistence in the driving process or differential ability to learn about these shocks), the ARMA-OLS method could give mis-leading results. Indeed, in our replication of a key result from Coibion and Gorodnichenko (2012), the response of inflation and forecast errors thereof to technology shocks, we find evidence of our overshooting patterns when we use both our ARMA-IV method and a local projection. Ap-

[28] Results are similar if we treat a different portion (e.g., 1/2 or all) of the "about the same" responses as corresponding to "up."

[29] For consistency with the previous analysis, we compare this to data on the GDP deflator, even though this is almost certainly not a perfect match for the price variable households have in mind when answering the survey.

Figure 5: Dynamic Responses in a Structural VAR



Notes: The sample period is Q4 1968 to Q4 2017. The x-axis denotes quarters from the shock (starting at 0). The shaded areas are 68% high-posterior-density regions and the point estimate is the posterior median. In the first row the outcome is u_t and the forecast is $\bar{\mathbb{E}}_{t-3}[u_t]$; in the second row the outcome is $\pi_{t,t-4}$, or annual inflation, and the forecast is $\bar{\mathbb{E}}_{t-3}[\pi_{t,t-4}]$; and in the last row, the outcome is $\pi_{t,t-1}$, or one-quarter inflation, and the forecast is $\bar{\mathbb{E}}_{t-3}[\pi_{t,t-1}]$. The columns show results from a "max share" identification and a triangular identification, respectively; see the main text for details.

pendix B unpacks the differences in methodology and demonstrates why the particular implementation in Coibion and Gorodnichenko (2012) makes it impossible to see the over-shooting patterns uncovered here: the forecast errors are therein *restricted* to be uniformly positive.

Two complementary SVARs. Another option for estimating complex dynamics, of course, is to jointly estimate a multivariate model. We estimate a 13-variable VAR comprised of the ten key macroeconomic variables from Angeletos, Collard, and Dellas (2019) plus three forecast variables of interest: the three-period-ahead unemployment forecast, the three-period-ahead annual inflation forecast, and the three-period-ahead quarterly inflation forecast.[30] We apply the same Bayesian inference procedure as that pa-

[30]The ten macro variables are the following: real GDP, real investment, real consumption, labor hours, the labor share, the Federal Funds Rate, the rate of change in the GDP deflator, labor productivity in the non-farm business sector, and utilization-adjusted TFP. Full variable descriptions and data construction discussion is Angeletos, Collard, and Dellas (2019). The forecast variables are the three-step-ahead unemployment and inflation forecasts from the SPF. The sample period is Q4 1968 to Q4 2017.

per, including prior specification and posterior sampling procedures, and replicate their identification of shocks that target the "max share" of the business-cycle variation in unemployment and inflation.[31]

The left panel of Figure 5 shows the results. In the first row, we show the response of unemployment, forecasts thereof, and forecast errors to the "unemployment shock." This can be compared directly to the first column of Figures 3 and 4, and largely agrees about the potential for large and persistent "over-shooting" in forecast errors. The second and third row show the response of outcomes and forecasts to the inflation shock in the same SVAR model, but with different forecast horizons and transformations of the outcome variable (annual averages in Row 2 versus quarter-to-quarter rates in Row 3). Here we find quantitatively smaller effects per period, but also very persistent ones.[32]

In the right panel of Figure 5, we show the results of two different "Cholesky" identifications based on triangular short-run restrictions (ordering unemployment or inflation first). We find strong corroborating evidence of over-shooting for unemployment but only very weak evidence for inflation. Hence, Fact 4 for inflation seems to rely on filtering the "right" variation in the data. A related point is made below in the context of Fact 3, where we show that this fact is reinforced when we focus on the shock that drives the variation of inflation at business-cycle frequencies.

Dispersions, not means. Motivated by our framework, which emphasizes only mean forecasts, we have not looked at the dynamics of the dispersion in forecasts. In Appendix Figure A.5, we re-estimate (15) using the cross-sectional inter-quartile range of forecasts as the outcome.[33] There is a rough pattern of dispersion spiking on impact of shocks, particularly in the "diagonal" responses. A "cheap" way to accommodate this fact, which echoes Mankiw, Reis, and Wolfers (2004), in our framework is to let τ be time-varying while maintaining $\hat{\tau}$ fixed; this allows dispersion to vary without affecting at all the joint dynamics of average forecasts and aggregate outcomes. The more interesting possibility that time-variation in the levels of uncertainty and disagreement influence aggregate behavior (e.g., Bloom, 2009b; Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry, 2018) is left outside our analysis.

5.4 Three complementary tests

The impulse response evidence, combined with the discussion in Section 4, suggest we are heading toward a model that includes both incomplete information and over-extrapolation. Here, before proceeding to determine the implications of such a theory, we organize three additional tests that independently corroborate our main story.

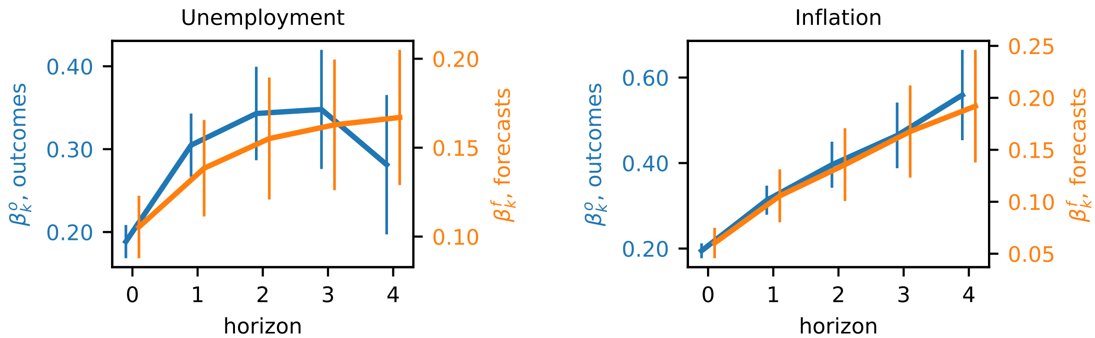
The "term structure" of forecasts. The impulse response functions plotted show forecasts of a *constant horizon* at different dates after the shock. But they do not show a forecaster's belief at any fixed date about how macro outcomes will behave in the future, which could perhaps offer the most direct evidence of over- or under- extrapolation.

[31] We are grateful to Fabrice Collard for help with this replication.

[32] In Appendix Figure A.4, we show the "off-diagonal" impulse responses of unemployment to the supply shock and inflation to the demand shock. They, too, show evidence of the overshooting.

[33] Results are similar using cross-sectional standard deviations, but the IQR method seems safer in the presence of outliers.

Figure 6: The Term Structure of Forecasts and Outcomes



Notes: The sample period is Q4 1968 to Q4 2017. The x-axis denotes quarters from the shock or horizon of forecast (starting at 0). The lines are one-standard-error bars. The orange lines plot the terms structure of forecasts, or β_k^f from (16), and the blue lines show the response of outcomes, or β_k^o from (17).

We can estimate a version of this in the SPF data, for forecasts up to 4 quarters out; forecasts of longer horizons are unavailable for the full sample. We thus consider the following "slice" of the projection regressions for forecasted variable and the realized outcome at horizons $k \in \{0, \ldots, 4\}$:

$$\bar{\mathbb{E}}_t[x_{t+k}] = \alpha_k + \beta_k^f \cdot \varepsilon_t + \gamma \cdot W_t + u_{t+k} \tag{16}$$

$$x_{t+k} = \alpha_k^o + \beta_k^o \cdot \varepsilon_t + \gamma^o \cdot W_t + u_{t+k} \tag{17}$$

We run these specifications for x equal to unemployment and inflation, and for ε equal to the corresponding shock. For consistency, we use the same control as those used earlier in projection (15).

The coefficients of interest are (β_k^f, β_k^o), which reveal the persistence of outcomes and forecasts. If $\beta_k^f < \beta_k^o$, which we have already verified for $k = 3$, we know agents under-react on impact. If β_k^o is much more persistent across k than β_k^f, this is also evidence of over-extrapolation right at the impact of the shock—that is, agents end up being *more correct* about impacts further in the future because their over-extrapolation partially cancels out their under-reaction.

Figure 6 plots the results, showing the values of β_k^o and β_k^f on the left and right scales, respectively. By comparing the left and right scales, we see that forecasts at all horizons under-react. But by comparing the blue line to the orange one, we also see that, in the case of unemployment, agents expect the effect of the shock to persist longer than it actually does.

A "structural" version of Facts 1 and 3. Now that we have committed to some notion of what "business cycle variation" we want to map to the model's shocks, we should recognize that a more precise analogue to Facts 1 and 3 (and our argument linking these Facts to the IRFs) would involve sub-setting to this particular variation. We thus revisit those regressions with an appropriate instrumental variables strategy.[34]

[34] Note, of course, that there is no clear "filtered" or "shock-specific" version of Fact 2 that can be estimated in the data. Once one subsets to aggregate variation, the BGMS regression essentially replicates the CG regression.

Table 4: Forecast Error Predictability with Business-Cycle Variation

Regression		Unemployment 1968-2017	Unemployment 1984-2017	Inflation 1968-2017	Inflation 1984-2017
	Revision$_t$ (K_{CG})	0.585	0.867	1.460	0.511
		(0.393)	(0.270)	(0.521)	(0.358)
(1)	First-stage F	7.527	4.736	3.517	5.047
	N	189	130	188	130
	OLS Estimate	0.741	0.809	1.528	0.292
	x_t (K_{KW})	-0.260	-0.073	0.085	-0.642
		(0.144)	(0.086)	(0.125)	(0.328)
(3)	First-stage F	2.671	5.560	1.697	1.513
	N	191	136	190	135
	OLS Estimate	-0.061	-0.036	0.111	-0.068

Notes: The dataset is the Survey of Professional Forecasters and the observation is a quarter between Q4-1968 and Q4-2017. All regressions include a constant. The forecast horizon is 3 quarters. Standard errors are HAC-robust, with a Bartlett kernel and lag length equal to 4 quarters. The data used for outcomes are first-release ("vintage"). The F values are a (multivariate extension of) the Kleibergen and Paap (2006) rk statistic and can be compared with critical values reported in Stock and Yogo (2005) for given levels of tolerated bias.

Concretely, we estimate versions of (1) and (3) using the current and six lags of the corresponding shock as an instrument for, respectively, Revision$_t$ and x_t. Table 4 reports the results.

For Fact 1, we find similar values to the OLS estimates for unemployment, but slightly larger, and more stable, values for inflation. This is consistent with the relative stability of our IRFs, as demonstrated in the comparison of the left and middle panels of Figure A.1.For Fact 3, we find more negative values that the OLS estimates. This underscores that the aspect of over-shooting that is picked up in the KW regression is more pronounced when focusing on the relevant business-cycle variation.

Auxiliary support for dispersed, private information. By embracing over-extrapolation, we have committed to a model in which some noisy perception is necessary to capture the initial sluggishness in expectations and the related Fact 1. Appendix C shows more clearly how to use a hybrid regression of individual and aggregate predictability (Facts 1 and 2) to test for noisy signals in this class of models. The underlying idea is that the difference between the CG and BGMS regression coefficients speaks directly to dispersed private information per se: if aggregate forecast errors are strongly positively related to past aggregate revisions, and if in addition one's forecast errors are negatively, or less positively, related to his own past revision, it has to be that one's forecast error is positively related to *others'* past revision. In other words, there is evidence supporting the hypothesis of "forecasting the forecasts of others."[35]

[35]Another strand of evidence comes from recent field-experiment work in Coibion, Gorodnichenko, and Kumar (2018) and Coibion, Gorodnichenko, and Ropele (2019).

6 Imperfect Expectations in GE

We now put the accumulated evidence about expectations to work in an equilibrium macroeconomic model. This serves four purposes. First, we verify that our main conclusions about identifying specific frictions via survey data are robust to considering the GE fixed point. Second, we demonstrate how our evidence can speak to intrinsically equilibrium theories of expectations formation, which had no place in our earlier PE analysis. Third, we clarify how the causal effects of the imperfect expectations in such a context depend on parameters that regulate GE feedbacks, such as the marginal propensity to consume, or the slope of the Keynesian cross. And finally, we offer a proof-of-concept calibration exercise which demonstrates how our evidence can offer not only theoretical but also quantitative guidance.

6.1 Primitives

Consider the New Keynesian model with no capital and perfectly rigid prices. Let y_t be output and c_t be consumption, where all quantities are in log deviations. The market clearing condition is $y_t = c_t$ and output is purely demand-determined given a fixed path of nominal interest rates (which also equal real interest rates).[36]

When agents have different and potentially irrational expectations, aggregate demand can no more be represented by the Euler equation of a representative consumer. Following the same steps as in Angeletos and Lian (2018), one can instead obtain the following "modern" version of the Keynesian cross:

$$c_t = \beta \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t[\xi_{t+1}^d - \varsigma r_{t+k}] + (1-\beta) \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t[y_{t+k}] \tag{18}$$

where $\bar{\mathbb{E}}_t$ denotes the average expectation in period t, β is the subjective discount factor, ς is the EIS, r_t is the nominal (also real) interest rate, and ξ_t^d is a demand (preference) shock. This condition follows from aggregating the log-linearized optimal consumption function and aggregating. The second term captures the consumers' present discounted value income, as in the Permanent Income Hypothesis (PIH).

To see more clearly how (18) captures the Keynesian cross, let $Y = \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t[y_{t+k}]$ be the average, possibly irrational, expectation of permanent income. We can then read (18) as $c = a + bY$, where $a \equiv \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t[\xi_t^d - \varsigma r_t]$ is the intercept of the Keynesian cross and $b \equiv (1-\beta)$ is its slope, or equivalently the marginal propensity to consume out of income (MPC).

For our purposes, it is therefore best to replace β in condition (18) with $1 - \text{mpc}$, treat mpc as a primitive parameter, and think of $\text{mpc} \approx 0.3$ as an empirically plausible benchmark. This is further justified in Angeletos and Huo (2019) by drawing a connection between a heterogeneous-agent variant of the present framework and the HANK literature.[37]

Let $\xi_t \equiv \xi_t^d - \varsigma r_t$ denote the "total demand shock" relative to steady state; formally, this is a rescaling of the deviation of interest rates from the natural rate. We close the model by letting ξ_t be an exogenous

[36] For completeness, assume that labor is supplied to meet final demand; and that a competitive, representative firm operates a linear production technology with constant productivity to produce the homogeneous final good.

[37] See also the related OLG versions of the New Keynesian model in Piergallini (2007), Del Negro, Giannoni, and Patterson (2015) and Farhi and Werning (2019).

AR(1) process with persistence ρ and Gaussian one-step-ahead innovations ε_t:

$$\xi_t = \frac{1}{1-\rho\mathbb{L}}\varepsilon_t \tag{19}$$

A positive $\varepsilon_t$can be either an expansionary monetary policy or an expansionary demand shock.

Like in Section 4, we let each consumer observe only a noisy Gaussian private signal of ξ_t, the true precision of which is given by $\tau > 0$:

$$s_{i,t} = \xi_t + \frac{u_{i,t}}{\sqrt{\tau}} \qquad\qquad u_{i,t} \sim_{iid} N(0,1) \tag{20}$$

And we let consumers' subjective perception of the precision of their information and of the persistence of the underlying impulse be, respectively, some $\hat{\tau} > 0$ and some $\hat{\rho} \in (0,1)$, which may differ from the objective counterparts.

6.2 Solving and characterizing the fixed point

As already mentioned, Angeletos and Huo (2019) have solved the fixed point of a similar model as ours, under the restrictions $\hat{\rho} = \rho$ and $\hat{\tau} = \tau$. The following two propositions extend their results to the present environment and offer a simple description of how the frictions influence macro dynamics.

Proposition 3. *An equilibrium to this model exists, is unique, and admits a finite ARMA representation for the aggregate outcome and the average forecasts.*

Proposition 4 (As-if Representation)**.** *There exist functions Ω_f and Ω_b such that the unique equilibrium dynamics of the imperfect-expectations economy is the same as that of a perfect-expectations counterpart with the following Euler equation:*

$$c_t = -\varsigma r_t + \omega_f \mathbb{E}_t^*[c_{t+1}] + \omega_b c_{t-1} + \xi_t^d \tag{21}$$

where $\omega_f = \Omega_f(\hat{\tau}, \rho, \hat{\rho}, mpc)$, $\omega_b = \Omega_b(\hat{\tau}, \rho, \hat{\rho}; mpc)$, and $\mathbb{E}_t^*$ is the rational, full-information, expectation operator. Furthermore, $\omega_b > 0$ if and only if $\hat{\tau} < \infty$, and $\omega_f < 1$ if and only if either $\hat{\rho} < \rho$ or $\hat{\tau}$ is small enough relative to $\hat{\rho} - \rho$.

This result offers a bridge to simple representative-agent macro models: ω_b resembles habit persistence, ω_f represents a form of myopia (if $\omega_f < 1$) or hyperopia (if $\omega_f > 1$). The economy with noisy perception ($\hat{\tau} < \infty$ and $\tau < \infty$) but no over-extrapolation ($\hat{\rho} = \rho$) features both myopia ($\omega_f < 1$) and anchoring ($\omega_b > 0$). At the other extreme, if we shut down noisy perception (i.e., take $\hat{\tau} = \tau \to \infty$), we find that over-extrapolation alone maps to hyperopia ($\omega_f > 1$) and under-extrapolation alone maps to myopia ($\omega_f < 1$), but neither by itself produces anchoring ($\omega_b = 0$).

The case of most interest, over-extrapolation combined with *sufficiently* large noise, maps to $\omega_f < 1$ and $\omega_b > 0$. The former dampens the economy's response to innovations, and to news about the future. The latter plays a similar role as habit persistence in consumption—or, if we translate the results to other contexts, as adjustment costs to investment, price-indexation in the NKPC, or momentum in asset prices (see Angeletos and Huo, 2019, and references therein). Finally, compared to the versions of these mechanisms found in the DSGE literature, the ones obtained here have two distinctive qualities. First, they are endogenous to policy, market structures and GE multipliers. And second, they are disciplined by the provided evidence on expectations.[38]

6.3 Facts 1-4 in GE

Let us next focus on the connection with the data, which is the main contribution of our paper. The following result verifies that *all* our main insights from Section 5 go through modulo additional dependence on GE feedback:

Proposition 5 (Facts 1-4 in GE). *Corollaries 1, 2 and 3 go through in the GE context, but all moments now depend also on the MPC.*

That is: (i) we can still show that the pure noisy rational expectations and pure mis-specification models are insufficient to describe the observed moments; (ii) we can still select the combination of noise and over-extrapolation as the “right” model; and (iii) we still have that delayed over-shooting in IRFs is smoking-gun evidence of the combination of noise and over-extrapolation; but (iv) we now must condition all inference on additional information about the extent of GE feedback in the economy.

6.4 Imperfect reasoning versus imperfect expectations

A claim we have made earlier, and now have the tools to formalize, is that certain models of *imperfect reasoning in equilibrium* work in similar ways as our more mechanical model of under-extrapolation and are therefore not a good fit for the data. We consider, in particular, three such models:

1. **Dogmatic higher-order doubts.** Assume that each consumer observes ξ_t with probability 1 but attaches only probability $q \in (0,1)$ that any other consumers also observes ξ_t; with the remaining probability, any other agent is expected to have her belief about ξ_t reset to the prior. Such a model is the main specification in Angeletos and Sastry's (2020) work on forward guidance at the ZLB. It captures the same kind of inertia in forward-looking higher-order beliefs and the same consequent forms of myopia and GE attention as those featured in Angeletos and Lian (2018) and our own GE setting, but replaces the informational friction with a systematic bias in beliefs. It therefore builds a bridge to the following two models, which introduce similar biases.

2. **Level-K thinking.** Assume that a consumer of “level 1” perfectly observes ξ_t but assumes all others consumers a default action $c^d_{i,t} = 0$; an agent of level 2 also perfectly observes ξ_t but assumes all other agents play the level-1 action; and this definition recursively extends up to order K, for some finite $K > 2$. Such models have been used to explain the sluggish, and often incomplete, convergence to

[38] Maćkowiak and Wiederholt (2015) and Afrouzi and Yang (2019) argue a version of the first point by focusing on the endogeneity of attention, or of τ in our framework. Here, we instead emphasize the endogeneity of ω_f and ω_b on the MPC and other GE parameters for *given* τ, as in Angeletos and Huo (2019).

Nash equilibrium play in laboratory settings (e.g., Nagel, 1995) and, more recently, agents' expectations formation about "unconventional" policy (e.g., Garcıa-Schmidt and Woodford, 2019; Farhi and Werning, 2019; Iovino and Sergeyev, 2017).

3. **Cognitive discounting** (Gabaix, 2020). Agents have misspecified priors about the processes of the exogenous state and the endogenous aggregate spending. In particular, whenever that the actual laws of motion are

$$\xi_t = \rho \xi_{t-1} + \varepsilon_t \qquad \text{and} \qquad y_t = R y_{t-1} + D\varepsilon_t,$$

for some constants R and D (to be determined as part of the solution), the agents believe that

$$\xi_t = m\rho \xi_{t-1} + \varepsilon_t \qquad \text{and} \qquad y_t = mR y_{t-1} + D\varepsilon_t,$$

for some exogenous scalar $m \in (0,1)$ that represents the degree of "cognitive discounting" applied when the consumers contemplate the future.

These models have different methodological underpinnings. But they induce essentially the same distortion in beliefs. In particular, it is easy to show that all three models impose that the average *subjective* expectation and the corresponding *rational* expectation are connected by the following restriction:

$$\bar{\mathbb{E}}_t[y_{t+1}] = d \cdot \mathbb{E}_t^*[y_{t+1}],$$

where $d \in (0,1)$ is a scalar that depends on the "deep" parameter $\zeta \in \{q, K, m\}$ of the respective model. This scalar measures how much consumers underestimate the future response in the behavior of others and, equivalently, the future response of y_t. What differs is the reason for $d < 1$: underestimating the knowledge of others (Model 1), underestimating the rationality of others (Model 2), or applying a behavioral discount to the future (Model 3).

Now note that the form of under-extrapolation accommodated in our framework plays the same role as well. Indeed, if we shut down noisy perception, we can show that

$$\bar{\mathbb{E}}_t[y_{t+1}] = \frac{\hat{\rho}}{\rho}\mathbb{E}_t^*[y_{t+1}].$$

It follows that, for any of the aforementioned three models, we can find a value of $\hat{\rho}$ less than ρ such that our model implies the same effective friction in the expectations. The next Proposition verifies that this logic carries over to the entire set of predictions about outcomes and forecasts. The Corollary spells out the relevant empirical implications.

Proposition 6. *For any the three models described above and any value for the corresponding parameter $\zeta \in \{q, K, m\}$, there exists some $\hat{\rho} = f(\rho, \zeta, mpc) < \rho$, such that the outcomes of the original model is observationally equivalent to our own model without noise ($\tau = \hat{\tau} = \infty$) and under-extrapolation ($\hat{\rho} < \rho$).*

Corollary 4. *For any the three models described above, the following properties hold: $K_{CG} = K_{BGMS} > 0$, $K_{KW} > 0$, and the IRF of the average forecast errors is uniformly positive. That is, these models are at odds with Facts 2, 3 and 4.*

All these models have consumers under-estimate the future response of others, which in turn impact behavior in a similar way as an under-estimation of the persistence of ξ_t. The only subtle difference between them is whether the belief mis-specification operates through both PE and GE considerations, or only through GE. For the first two models (noiseless higher-order doubts and level K thinking), because the extent of the friction is tied closely to the extent of strategic interaction, the replicating $\hat{\rho}$ will be a function of the MPC.

Our observation that pure under-extrapolation cannot explain the business-cycle macro data on imperfect expectations (Corollary 1 and Proposition 5) thus extends to the aforementioned GE dampening models as well. First, each model restricts $K_{\text{CG}} = K_{\text{BGMS}}$, or it fails to provide a reason why the forecast errors of one are predictable by the information of others. And second, even in variants that add some noisy perception and that could thus help match the CG and BGMS evidence, none of these theories could explain the observed over-extrapolation in impulse response functions.[39]

The obvious caveat to this conclusion is that it only applies to the particular evidence we have considered here and may not extend to other contexts. Another caveat is that this conclusion is modulated by parsimony: given the evidence at hand, we cannot reject the hypothesis that agents over-extrapolate the aggregate shocks and at the same time are shallow thinkers with respect to GE.

6.5 A quantitative assessment

Let us now illustrate how our empirical findings can help quantify the GE effects of the documented mechanisms, using the New Keynesian model as our laboratory economy.[40]

To speak jointly to data on output (unemployment) and inflation, let us first extend the model to allow for partially flexible prices. This involves adding a block of non-competitive intermediate goods firms who operate a linear technology and reset prices with probability $1-\theta \in (0,1)$ and a final goods firm that competitively combines these goods with a constant-returns-to-scale technology. As in Angeletos and Lian (2018) and Angeletos and Huo (2019), the model's three equations can be expressed as follows:

$$c_{i,t} = \mathbb{E}_{i,t}\left[\sum_{k=0}^{\infty}(1-\text{mpc})^{k+1}\left[\xi^d_{t+k} - \varsigma(i_{t+k} - \pi_{t+k+1})\right] + \text{mpc}\sum_{k=0}^{\infty}(1-\text{mpc})^k c_{t+k}\right]$$

$$\pi_{i,t} = \mathbb{E}_{i,t}\left[\theta\sum_{k=0}^{\infty}(\beta\theta)^k \kappa(c_t + \xi^s_t) + (1-\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \pi_{t+k}\right]$$

$$i_t = \phi_\pi \pi_t$$

The first equation is the Dynamic IS Curve, modified to allow for informational frictions and mis-specified beliefs along the lines discussed earlier. i_t is the nominal interest rate, π_t is inflation, and ξ^d_t is a preference

[39]Furthermore, in a survey of firms designed to shed light on related issues, Coibion, Gorodnichenko, Kumar, and Ryngaert (2018) find support for informational frictions but no relation between measured level-K thinking and expectations.

[40]Previous works such as Mankiw and Reis (2007), Maćkowiak and Wiederholt (2015) and Melosi (2016) have also sought to quantify the macroeconomic effects of informational frictions in the baseline New Keynesian model, but have not disciplined the exercise with the expectations evidence we consider here. Bordalo, Gennaioli, Shleifer, and Terry (2019), on the other hand, quantify the role of over-extrapolation in an RBC model with credit friction, but they abstract from informational frictions and do not address the particular patterns of the expectations of inflation and unemployment on which we focus.

Table 5: Model Parameters

(a) exogenously fixed

Parameter	Description	Value
θ	1 - probability of price reset	0.6
κ	Slope of Phillips Curve	0.02
β	Discount factor	0.99
mpc	Marginal propensity to consume	0.3
ς	IES	1.0
ϕ	Policy rule slope	1.5

(b) calibrated

	ρ	$\hat{\rho}$	$\hat{\tau}$
Demand shock	0.80	0.95	0.36
Supply shock	0.57	0.82	0.15

shock, which maps to our empirical demand shock. The second equation is the corresponding modification of the NKPC. κ is its slope with respect to the real marginal cost, θ is the Calvo parameter (one minus the probability of resetting prices), and ξ_t^s is a cost-push shock, which maps to our empirical supply shock. The third and final equation is the rule for monetary policy, in which ϕ_π is the slope in current inflation.

We close the model by specifying the shock processes and the belief structures in the same way as before. Using the methods of Angeletos and Huo (2019), we then analytically solve for the equilibrium responses of inflation and consumption as functions of two sets of parameters: the "familiar" parameters (ς, mpc, $\beta, \theta, \kappa, \phi_\pi$); the actual and perceived persistence of the shocks ($\rho, \hat{\rho}$); and the perceived precision ($\hat{\tau}$). For the reasons already explained, the actual precision (τ) does not enter the determination of either the aggregate outcomes or the average expectations thereof.

To connect the model to the data, we interpret π_t as the quarterly rate of inflation and the negative of y_t as the quarterly rate of unemployment. The first choice requires no justification. The second one is based on the logic that, in our model, y_t coincides with the output gap, which in turn is closely related to unemployment both in richer models and in the data. We next fix the model's behavioral and policy parameters to conventional values, as shown in left panel of Table 5. We finally pick, for each shock, the values of $\rho, \hat{\rho}$ and $\hat{\tau}$ so as to match as well as possible the key evidence reported in Section 5—that is, the IRF of outcomes and forecasts in Figures 3 and 4 (ARMA-IV method), as well as the "term structure" of forecasts in Figure 6. These moments provide the most direct evidence of the forces we have in mind, as discussed in Section 4. This procedure yields the parameters values seen in the right panel of Table 5.

Figure 7 illustrates the model's fit vis-a-vis the empirical IRFs seen earlier in Figure 3. The fit is quite good in the context of the demand shock, but mediocre in the context of the supply shock. This underscores that, although the model has the right *qualitative* ingredients, its *quantitative* performance is not automatic: there is no abundance of degrees of freedom.

We henceforth focus on the demand shock and study two counterfactuals. In the one, we shut down the over-extrapolation, isolating the role of the information friction. In the second, we shut down both frictions, recovering the textbook New Keynesian model. These counterfactuals are illustrated in, respectively, the second and third column of Figure 8. The first column is the full model, with both frictions.

Figure 7: Model vs Data

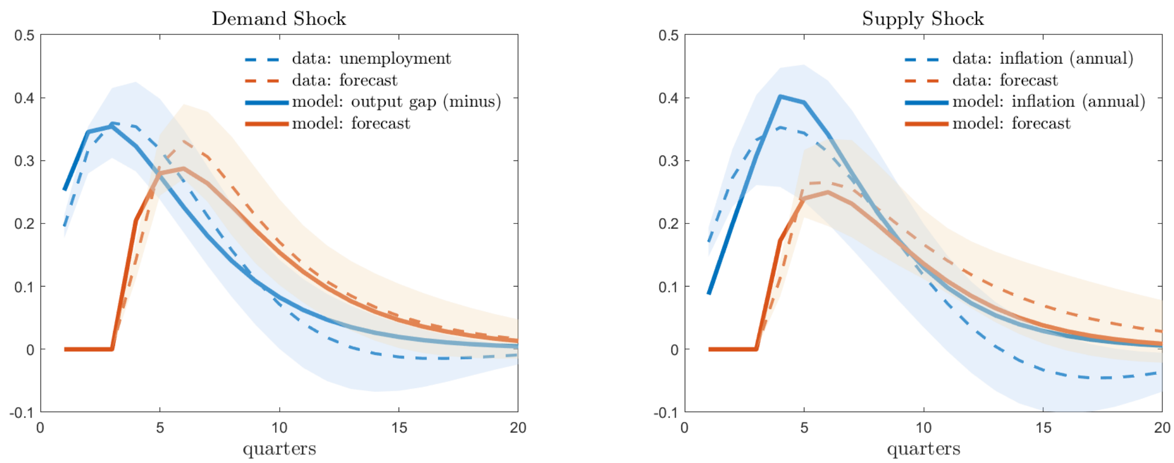
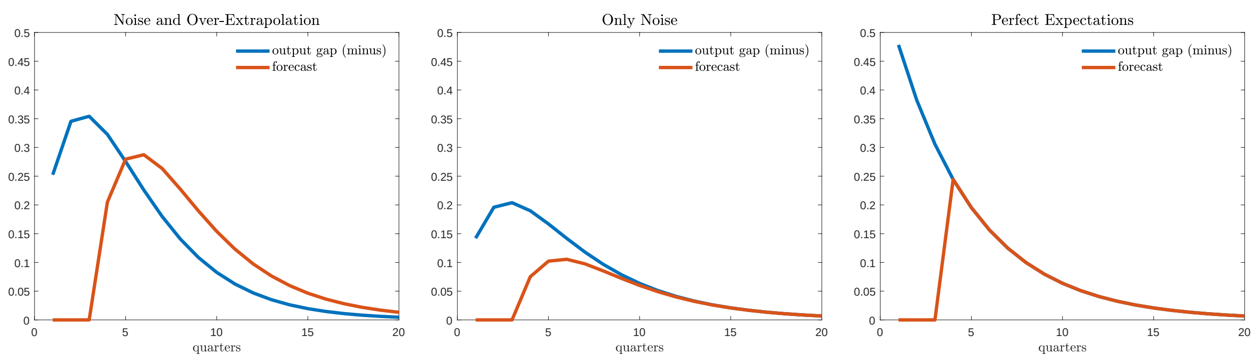


Figure 8: Counterfactuals for Demand Shock



By comparing the second column to the third one, we see that the informational friction alone is the source of both significant dampening and significant persistence relative to the frictionless benchmark. Compared to the textbook model, the informational friction—calibrated to the evidence presented in this paper—decreases the impact of the demand shock on the output gap by about 50% and its impact on inflation by about 75%. As for the induced persistence, it is quantitatively comparable to that obtained in richer DSGE model with the use of habit persistence in consumption and the hybrid version of the NKPC.

This echoes the common message of a large literature on information frictions (e.g., Woodford, 2003; Mankiw and Reis, 2007; Nimark, 2008; Maćkowiak and Wiederholt, 2015). The added value here is that we have disciplined the theory with expectations evidence (as in Angeletos and Huo, 2019) and that we have accommodated over-extrapolation. Without it, the model fails to capture Fact 4: as seen in the second column of Figure 8, the forecasts in the noise-only model do not overshoot.

By comparing the first column to the second one, we then see that the main effect of over-extrapolation on actual outcomes is to amplify their responses to the shock. And while the over-shooting looks "small" in terms of the size of the forecast errors, the aforementioned amplification is sizable for two reasons. First, a small difference between $\hat{\rho}$ and ρ translates to a large difference in the kind of discounted present values that consumer spending and firm pricing. And second, any such belief mistake gets amplified at the aggregate level by GE feedback.

Needless to say, these counterfactuals should not be taken too seriously. They do, however, illustrate the potential value of accommodating the mechanisms and the evidence presented here in richer models.

We close this section with the following note. So far, we have utilized only evidence on *average* forecasts, ignoring the kind of individual-level evidence that was the focus of BGMS. This is because the BGMS regression coefficient only helps pin down a "residual" parameter (τ) that does not enter the dynamics of either the aggregate outcomes or the average forecasts. The BGMS evidence therefore has no (independent) effect on the counterfactuals conducted above.[41] As for the CG and KW evidence, they were subsumed by our evidence about IRFs.

7 Conclusion

Where are we in the "wilderness" of imperfect expectations? This paper organized theory and survey evidence to answer this question, taking into account both the possibility for multiple competing distortions in expectations formation and the equilibrium fixed point between expectations and outcomes.

We proposed and implemented a new empirical strategy: estimation of the impulse response function of the average forecast errors of unemployment and inflation to the shocks that drive most of the business-cycle variation in these variables. We explained why *in theory* this strategy is both more informative and more easily interpretable than alternatives found in previous empirical studies. And we demonstrated *in practice* how the information extracted via our strategy helps achieve three goals: resolve the apparent conflict between the empirical findings of previous works; help select among multiple competing theories of expectation formation; and serve as a "sufficient statistic" for quantitative purposes.

The main empirical finding was a form delayed over-shooting in expectations: following any shock, forecasts appear to under-react for the first few quarters but over-shoot later on. The main lesson for theory was that, at least with the class of models considered, the data require the combination of a sizable informational friction and a behavioral tendency to over-extrapolate the macroeconomic dynamics.

Theories that emphasize under-extrapolation or closely related mechanisms, such as cognitive discounting and level-K thinking, were shown to be at odds not only with the new fact documented here but also with the individual-level evidence on expectations documented in Bordalo, Gennaioli, Ma, and Shleifer (2018). At the same time, we echoed Angeletos and Huo's (2019) point that such *individual*-level evidence may not be strictly needed for the purpose of quantifying the overall effect on the macroeconomic dynamics: in the class of models considered, our evidence about *average* forecasts served as "sufficient statistics" for the counterfactuals of interest.

We conclude with few notes on future research that would further solidify our understanding of macro belief dynamics and further the research program outlined in this article.

[41] The BGMS evidence was nevertheless useful in corroborating the case for over-extrapolation and, conversely, in ruling out theories that resemble under-extrapolation. In this sense, it remains useful for selecting the "right" model of beliefs. But it could be dispensed with in our counterfactuals.

Learning foundations. A question we have not tried to answer, at any point in this article, is where agents' subjective model of the world comes from. More specifically, why would agents think that business cycle shocks to demand or marginal costs have a higher persistence than they really do?

Our analysis simplifies the matter greatly by having agents put a dogmatic belief that the true persistence is $\hat{\rho}$. One might hypothesize that standard results on the learning foundations of rational expectations equilibria, extended to our setting, would rule out convergence of beliefs to $\hat{\rho} \neq \rho$ given a reasonable (non-degenerate) prior on $\hat{\rho}$ (Marcet and Sargent, 1989; Evans and Honkapohja, 2001). That said, if we extended the model to make the one-dimensional, AR(1) representation of fundamentals only an imperfect approximation of a richer truth underlying truth, we may observe convergence to $\hat{\rho} \neq \rho$ (Molavi, 2019).

A variant story involves rational confusion between transitory and permanent shocks. Such confusion may cause agents to respond to the transitory shock *as if* they incorrectly perceive its persistence to be higher than the true. That is, such confusion can produce a rational form of over-extrapolation. But it also predicts that agents ought to under-extrapolate the effects of the permanent shock. Of course an auxiliary prediction is that agents underreact to the permanent changes. We found no support for this prediction when we looked at the dynamic responses of forecasts to a technology shock identified as in Galí (1999). But a more thorough quantitative analysis along these lines is an interesting angle for future work.

Non-stationary environments. Agents may also be using the wrong model because the underlying structure of the economy is changing underneath them. Two examples of this stick out.

The first involves the long-run changes in the behavior of US inflation, especially after the 1970s and early 1980s. The empirical findings in Section 3 suggest more severe over-extrapolation for inflation in the modern period, in which inflation itself is less persistent. This could be consistent with agents' perceiving some "shadow" of the more ferocious shocks and/or timid policy response of the earlier period.[42]

A second important event in our sample is the extended stint at the Zero Lower Bound during the Great Recession. A number of authors have postulated that this unfamiliar and extreme event may have caused agents to "throw out" their conventional models, justifying more dramatic departures from rational expectations (Angeletos and Sastry, 2020; Farhi and Werning, 2019; Gabaix, 2020; Garcıa-Schmidt and Woodford, 2019; Iovino and Sergeyev, 2017). We are sympathetic to this view and not insistent that our conclusions need to apply for expectations at the ZLB. There is more work to be done in investigating exactly for what counterfactuals and policy changes our empirical findings may provide good guidance.

The "right" expectations data. This paper, like much of the related empirical literature, has relied primarily on surveys of professional forecasters and analysts, because of data availability and quality. We provided corroborating evidence from the University of Michigan Consumer Sentiment survey, but the imprecise measure of the relevant expectations in that survey precluded an equally sharp exercise as that based on SPF and Blue Chip data. The ideal implementation of our approach, which we leave for the

[42]This explanation relates to a rich literature looking for statistical break points in volatility and/or policy in modern history (Sargent, 2001; Primiceri, 2005; Sargent, Williams, and Zha, 2006; Sims and Zha, 2006). It is also natural within the GE theory presented here and in Angeletos and Huo (2019): in this context, the information-driven persistence in inflation is modulated by policy and, more specifically, decreased by a steeper Taylor rule.

future, requires sufficiently long time series of the expectations of consumers and firms, not only about macroeconomic outcomes, but also about the objects that matter more directly to their behavior, such as consumers' own income and firms' own sales.

Applications to finance. The co-existence of under-reaction and over-extrapolation is a classic fact for many asset prices (De Bondt and Thaler, 1985; Cutler, Poterba, and Summers, 1991; Lakonishok, Shleifer, and Vishny, 1994). Our findings thus represent a step toward unifying our understanding of imperfect expectations in both macroeconomics and finance. An interesting possibility for future work is to replicate our impulse response evidence with dividends or earnings and expectations thereof, to determine if a similar structural interpretation (noise plus over-reaction) holds true in this domain and also provides useful predictions for stock price dynamics.

References

Adam, Klaus and Michael Woodford. 2012. "Robustly optimal monetary policy in a microfounded New Keynesian model." *Journal of Monetary Economics* 59 (5):468 – 487.

Afrouzi, Hassan and Choongryul Yang. 2019. "Dynamic Rational Inattention and the Phillips Curve." mimeo.

Andrade, Philippe and Hervé Le Bihan. 2013. "Inattentive professional forecasters." *Journal of Monetary Economics* 60 (8):967–982.

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas. 2018. "Quantifying Confidence." *Econometrica* 86 (5):1689–1726.

———. 2019. "Business Cycle Anatomy." *American Economic Review* conditionallyaccepted.

Angeletos, George-Marios and Zhen Huo. 2019. "Myopia and Anchoring." *MIT mimeo* .

Angeletos, George-Marios and Jennifer La'O. 2020. "Optimal Monetary Policy with Informational Frictions." *Journal of Political Economy* 128 (3).

Angeletos, George-Marios and Chen Lian. 2016. "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination." *Handbook of Macroeconomics* 2:1065–1240.

———. 2018. "Forward Guidance without Common Knowledge." *American Economic Review* 108 (9):2477–2512.

Angeletos, George-Marios and Karthik A. Sastry. 2020. "Managing Expectations: Instruments vs. Targets." *MIT mimeo* .

Baker, Scott R., Nicholas Bloom, and Steven J. Davis. 2016. "Measuring Economic Policy Uncertainty." *Quarterly Journal of Economics* 131 (4):1593–1636.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. 1998. "A model of investor sentiment." *Journal of Financial Economics* 49 (3):307 – 343.

Barsky, Robert B. and Eric R. Sims. 2011. "News shocks and business cycles." *Journal of Monetary Economics* 58 (3):273 – 289.

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho. 2019. "Survey Data and Subjective Beliefs in Business Cycle Models." University of Minnesota, New York University and Federal Reserve Bank of Richmond.

Bloom, Nicholas. 2009a. "The impact of uncertainty shocks." *Econometrica* 77 (3):623–685.

———. 2009b. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3):623–685.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* forthcoming.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2018. "Over-reaction in Macroeconomic Expectations." *NBER Working Paper No. 24932* .

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2017. "Diagnostic Expectations and Credit Cycles." *The Journal of Finance* 73 (1):199–227.

Bordalo, Pedro, Nicola Gennaioli, Andrei Shleifer, and Stephen J. Terry. 2019. "Real Credit Cycles." *mimeo* .

Brunnermeier, Markus K. and Jonathan A. Parker. 2005. "Optimal Expectations." *American Economic Review* 95 (4):1092–1118.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 2016. "Understanding Booms and Busts in Housing Markets." *Journal of Political Economy* 124 (4):1088–1147.

Caballero, Ricardo and Alp Simsek. 2017. "A Risk-centric Model of Demand Recessions and Macroprudential Policy." miméo, MIT.

Caballero, Ricardo J. and Alp Simsek. 2019. "Monetary Policy with Opinionated Markets." miméo, MIT.

Caplin, Andrew and John V Leahy. 2019. "Wishful Thinking." Working Paper 25707, National Bureau of Economic Research.

Carroll, Christopher D. 2001. "The Epidemiology of Macroeconomic Expectations." Working Paper 8695, National Bureau of Economic Research.

Carroll, Christopher D. 2003. "Macroeconomic Expectations of Households and Professional Forecasters." *Quarterly Journal of Economics* 118 (1):269–298.

Christiano, Lawrence J., Roberto Motto, and Massimo Rostagno. 2014. "Risk Shocks." *American Economic Review* 104 (1):27–65.

Coibion, Olivier and Yuriy Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy* 120 (1):116–159.

———. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105 (8):2644–78.

Coibion, Olivier, Yuriy Gorodnichenko, and Saten Kumar. 2018. "How Do Firms Form Their Expectations? New Survey Evidence." *American Economic Review* 108 (9):2671–2713.

Coibion, Olivier, Yuriy Gorodnichenko, Saten Kumar, and Jane Ryngaert. 2018. "Do You Know That I Know That You Know...? Higher-Order Beliefs in Survey Data." *NBER Working Paper No. 24987* .

Coibion, Olivier, Yuriy Gorodnichenko, and Tiziano Ropele. 2019. "Inflation Expectations and Firm Decisions: New Causal Evidence*." *The Quarterly Journal of Economics* 135 (1):165–219.

Cutler, David M., James M. Poterba, and Lawrence H. Summers. 1991. "Speculative Dynamics." *The Review of Economic Studies* 58 (3):529–546.

D'Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber. 2019. "Exposure to Daily Price Changes and Inflation Expectations." Working Paper 26237, National Bureau of Economic Research.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. 1998. "Investor Psychology and Security Market Under- and Overreactions." *The Journal of Finance* 53 (6):1839–1885.

Das, Sreyoshi, Camelia M Kuhnen, and Stefan Nagel. 2020. "Socioeconomic status and macroeconomic expectations." *The Review of Financial Studies* 33 (1):395–432.

De Bondt, Werner FM and Richard Thaler. 1985. "Does the stock market overreact?" *The Journal of Finance* 40 (3):793–805.

Dean, Mark and Nate Leigh Neligh. 2017. "Experimental tests of rational inattention." .

Del Negro, Marco, Marc P Giannoni, and Christina Patterson. 2015. "The Forward Guidance Puzzle." *FRB of New York mimeo* .

Eusepi, Stefano and Bruce Preston. 2011. "Expectations, Learning, and Business Cycle Fluctuations." *American Economic Review* 101 (6):2844–72.

Evans, George W and Seppo Honkapohja. 2001. *Learning and expectations in macroeconomics.* Princeton University Press.

Farhi, Emmanuel and Iván Werning. 2019. "Monetary Policy, Bounded Rationality, and Incomplete Markets." *American Economic Review* 109 (11):3887–3928.

Fuhrer, Jeffrey C. 2018. "Intrinsic expectations persistence: evidence from professional and household survey expectations." Working Papers 18-9, Federal Reserve Bank of Boston.

Fuster, Andreas, David Laibson, and Brock Mendel. 2010. "Natural Expectations and Macroeconomic Fluctuations." *Journal of Economic Perspectives* 24 (4):67–84.

Gabaix, Xavier. 2016. "Behavioral Macroeconomics Via Sparse Dynamic Programming." *NBER Working Paper No. 21848* .

———. 2019. "Behavioral inattention." In *Handbook of Behavioral Economics: Foundations and Applications*, vol. 2, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, chap. 4. North-Holland, 261 – 343.

———. 2020. "A Behavioral New Keynesian Model." *American Economic Review* forthcoming.

Galí, Jordi. 1999. "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review* 89 (1):249–271.

Garcıa-Schmidt, Mariana and Michael Woodford. 2019. "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis." *American Economic Review* 109 (1):86–120.

Geanakoplos, John. 2010. "The leverage cycle." In *NBER Macroeconomics Annual 2009, Volume 24*. University of Chicago Press, 1–65.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer. 2015. "Expectations and investment." In *NBER Macroeconomics Annual 2015, Volume 30*. University of Chicago Press.

Greenwood, Robin and Andrei Shleifer. 2014. "Expectations of returns and expected returns." *Review of Financial Studies* 27 (3):714–746.

Guo, Hongye and Jessica A Wachter. 2019. "'Superstitious' Investors." Working Paper 25603, National Bureau of Economic Research.

Hamilton, James D. 1996. "This is what happened to the oil price-macroeconomy relationship." *Journal of Monetary Economics* 38 (2):215–220.

Hansen, Lars Peter and Thomas J. Sargent. 2012. "Three Types of Ambiguity." *Journal of Monetary Economics* 59 (5):422–445.

Hong, Harrison and Jeremy C. Stein. 1999. "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets." *The Journal of Finance* 54 (6):2143–2184.

Ilut, Cosmin L and Martin Schneider. 2014. "Ambiguous Business Cycles." *American Economic Review* 104 (8):2368–2399.

Iovino, Luigi and Dmitriy Sergeyev. 2017. "Quantitative Easing without Rational Expectations." *Work in progress* .

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1):161–182.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti. 2010. "Investment Shocks and Business Cycles." *Journal of Monetary Economics* 57 (2):132–145.

Kiley, Michael T. 2007. "A Quantitative Comparison of Sticky-Price and Sticky-Information Models of Price Setting." *Journal of Money, Credit and Banking* 39 (s1):101–125.

Kleibergen, Frank and Richard Paap. 2006. "Generalized reduced rank tests using the singular value decomposition." *Journal of econometrics* 133 (1):97–126.

Kohlhas, Alexandre and Tobias Broer. 2019. "Forecaster (Mis-)Behavior." *IIES miméo* .

Kohlhas, Alexandre and Ansgar Walther. 2018. "Asymmetric Attention." *IIES miméo* .

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. 1994. "Contrarian Investment, Extrapolation, and Risk." *The Journal of Finance* 49 (5):1541–1578.

Landier, Augustin, Yueran Ma, and David Thesmar. 2019. "Biases in expectations: Experimental evidence." *Available at SSRN 3046955* .

Lorenzoni, G. 2010. "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information." *The Review of Economic Studies* 77 (1):305–338.

Lucas, Robert E. Jr. 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4 (2):103–124.

Mackowiak, Bartosz and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review* 99 (3):769–803.

Maćkowiak, Bartosz and Mirko Wiederholt. 2015. "Business Cycle Dynamics under Rational Inattention." *The Review of Economic Studies* 82 (4):1502–1532.

Malmendier, Ulrike and Stefan Nagel. 2016. "Learning from inflation experiences." *The Quarterly Journal of Economics* 131 (1):53–87.

Mankiw, N Gregory and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* :1295–1328.

———. 2007. "Sticky information in general equilibrium." *Journal of the European Economic Association* 5 (2-3):603–613.

Mankiw, N Gregory, Ricardo Reis, and Justin Wolfers. 2004. "Disagreement about inflation expectations." In *NBER Macroeconomics Annual 2003, Volume 18*. The MIT Press, 209–270.

Marcet, Albert and Thomas J Sargent. 1989. "Convergence of least squares learning mechanisms in self-referential linear stochastic models." *Journal of Economic theory* 48 (2):337–368.

Matejka, Filip. 2015. "Rationally inattentive seller: Sales and discrete pricing." *The Review of Economic Studies* forthcoming.

Melosi, Leonardo. 2016. "Signalling effects of monetary policy." *The Review of Economic Studies* 84 (2):853–884.

Molavi, Pooya. 2019. "Macroeconomics with Learning and Misspecification: A General Theory and Applications." miméo, MIT.

Moore, Don A and Paul J Healy. 2008. "The trouble with overconfidence." *Psychological review* 115 (2):502.

Morris, Stephen and Hyun Song Shin. 2002. "Social Value of Public Information." *American Economic Review* 92 (5):1521–1534.

———. 2006. "Inertia of Forward-looking Expectations." *The American Economic Review* :152–157.

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *The American Economic Review* 85 (5):1313–1326.

Nimark, Kristoffer. 2008. "Dynamic Pricing and Imperfect Common Knowledge." *Journal of Monetary Economics* 55 (2):365–382.

Paciello, Luigi and Mirko Wiederholt. 2014. "Exogenous Information, Endogenous Information, and Optimal Monetary Policy." *Review of Economic Studies* 81 (1):356–388.

Piergallini, Alessandro. 2007. "Real Balance Effects and Monetary Policy." *Economic Inquiry* 44 (3):497–511.

Primiceri, Giorgio E. 2005. "Time Varying Structural Vector Autoregressions and Monetary Policy." *The Review of Economic Studies* 72 (3):821–852.

Romer, Christina D. and David H. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review* 94 (4):1055–1084.

Sargent, Thomas, Noah Williams, and Tao Zha. 2006. "Shocks and Government Beliefs: The Rise and Fall of American Inflation." *American Economic Review* 96 (4):1193–1224.

Sargent, Thomas J. 2001. *The conquest of American inflation*. Princeton University Press.

Sims, Christopher A. 1980. "Macroeconomics and Reality." *Econometrica* 48 (1):1–48.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3):665–690.

———. 2010. "Rational Inattention and Monetary Economics." *Handbook of Monetary Economics* 3:155–181.

Sims, Christopher A. and Tao Zha. 2006. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review* 96 (1):54–81.

Stock, James H. and Motohiro Yogo. 2005. *Testing for weak instruments in Linear Iv regression.* United Kingdom: Cambridge University Press, 80–108.

Uhlig, Harald. 2003. "What Moves Real GNP?" mimeo.

Woodford, Michael. 2003. "Imperfect Common Knowledge and the Effects of Monetary Policy." *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps* .

———. 2018. "Monetary Policy Analysis When Planning Horizons Are Finite." In *NBER Macroeconomics Annual 2018, Volume 33*. University of Chicago Press.

Appendices

A Extra Tables and Figures

Table A.1: Regression (1), Robustness to Data Choices and Horizons

		Unemployment			Inflation		
Sample	horizon =	1	2	3	1	2	3
Full, Vintage, SPF	Revision_t (K_{CG})	0.384	0.606	0.741	0.649	1.048	1.528
		(0.128)	(0.178)	(0.232)	(0.290)	(0.337)	(0.418)
	R^2	0.111	0.143	0.111	0.122	0.200	0.278
	N	196	196	191	195	195	190
1984-, Vintage, SPF	Revision_t (K_{CG})	0.385	0.657	0.809	-0.100	0.160	0.292
		(0.203)	(0.255)	(0.305)	(0.159)	(0.174)	(0.191)
	R^2	0.116	0.195	0.159	0.002	0.005	0.016
	N	136	136	136	135	135	135
Full, Final, SPF	Revision_t (K_{CG})	0.411	0.612	0.731	0.578	0.991	1.403
		(0.127)	(0.180)	(0.233)	(0.215)	(0.261)	(0.334)
	R^2	0.135	0.147	0.108	0.104	0.200	0.249
	N	199	198	192	199	198	192
1980-, Vintage, BC	Revision_t (K_{CG})	0.310	0.544	0.804	0.024	0.378	0.618
		(0.129)	(0.213)	(0.231)	(0.204)	(0.188)	(0.205)
	R^2	0.091	0.132	0.149	0.000	0.033	0.067
	N	151	151	150	150	150	149

Notes: All regressions include a constant. Standard errors are HAC-robust, with a Bartlett kernel and lag length equal to 4 quarters.

Table A.2: Regression (2), Robustness to Data Choices and Horizons

		Unemployment			Inflation		
Sample	horizon =	1	2	3	1	2	3
Full, Vintage, SPF	$\text{Revision}_{i,t}$ (K_{BGMS})	0.186	0.300	0.321	-0.100	0.024	0.143
		(0.077)	(0.094)	(0.107)	(0.084)	(0.098)	(0.123)
	R^2	0.029	0.042	0.028	0.004	0.000	0.005
	N	5,808	5,699	5,383	5,496	5,458	5,147
Full, Final, SPF	$\text{Revision}_{i,t}$ (K_{BGMS})	0.200	0.296	0.321	-0.100	0.056	0.179
		(0.075)	(0.090)	(0.106)	(0.075)	(0.091)	(0.122)
	R^2	0.035	0.042	0.028	0.004	0.001	0.006
	N	5,831	5,728	5,419	5,571	5,520	5,226

Notes: All regressions include a constant. Standard errors are clustered two-way by forecaster ID and time period. Both errors and revisions are winsorized over the sample to restrict to 4 times the inter-quartile range away from the median.

Table A.3: Regression (3), Robustness to Data Choices

		Forecast Error			
		Unemployment		Inflation	
Sample	data for error =	vintage	final	vintage	final
Full, Vintage, SPF	x_t (K_{KW})	–0.061	-0.061	0.111	0.097
		(0.056)	(0.056)	(0.075)	(0.066)
	R^2	0.016	0.016	0.058	0.047
	N	194	195	193	195
Full, Final, SPF	x_t (K_{KW})	-0.058	-0.058	0.117	0.115
		(0.056)	(0.056)	(0.078)	(0.069)
	R^2	0.014	0.009	0.062	0.063
	N	194	195	192	194

Notes: All regressions include a constant. Standard errors are HAC-robust, with a Bartlett kernel and lag length equal to 4 quarters.

Table A.4: Calibrating with Unconditional Moments

	Unemployment		Inflation	
	1968-2017	1984-2017	1968-2017	1984-2017
ρ	0.91		0.89	
K_{CG}	0.741	0.809	1.528	0.292
K_{BGMS}	0.321	0.398	0.143	-0.263
K_{KW}	-0.061	-0.036	0.111	-0.068
$\hat{\rho}$	0.972	0.966	0.893	0.947
$\hat{\tau}$	0.449	0.418	0.335	1.850
τ	2.028	2.231	0.464	0.693

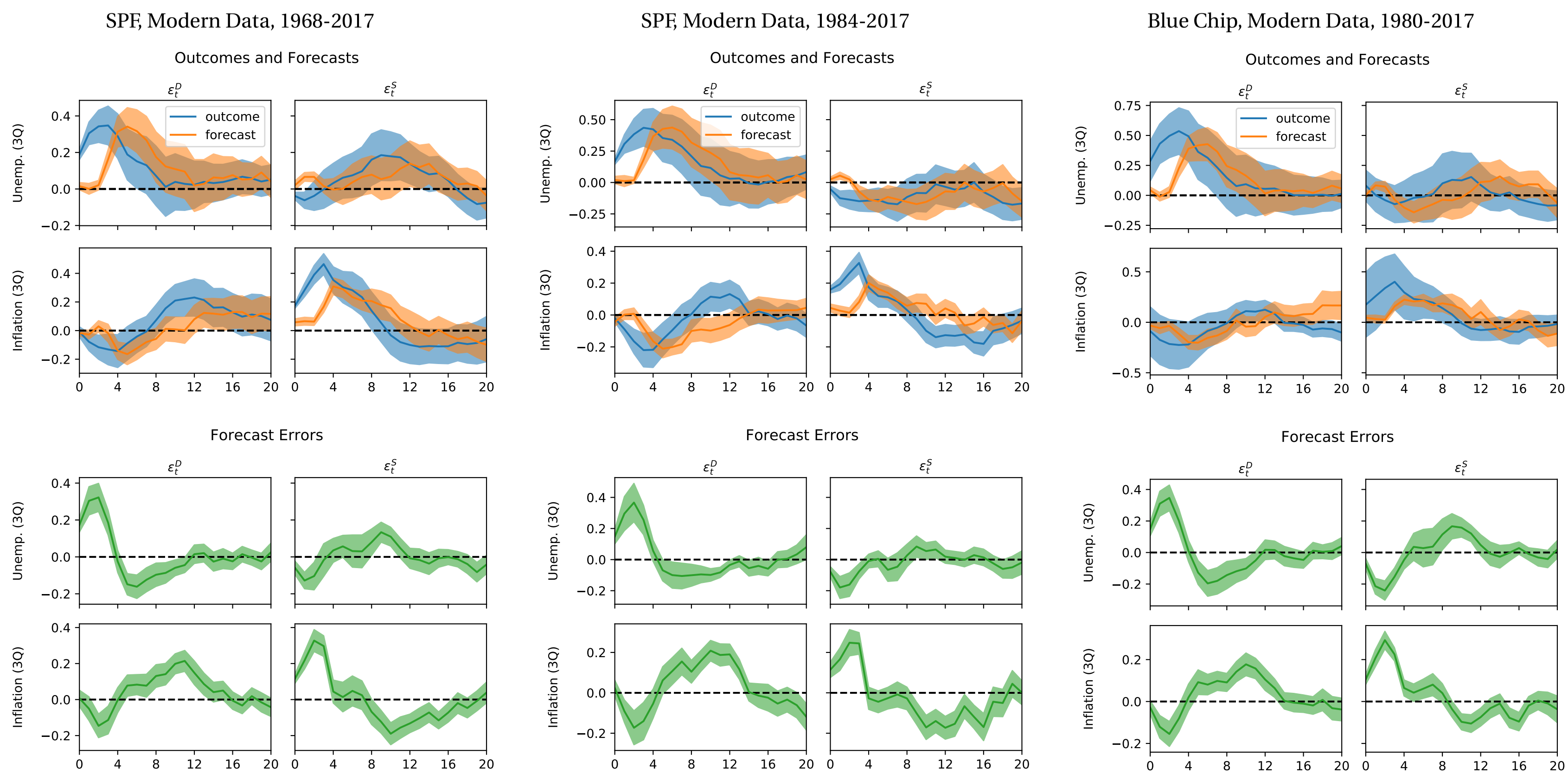
Notes: The persistence valeus come from band-pass filtered data on final outcomes over our sample.

Table A.5: First-stage F Statistics

	Unemployment		Inflation	
	Outcomes	Forecasts	Outcomes	Forecasts
N	188		188	
F	1.686	2.941	2.077	2.381
endogenous regressors	3		3	
instruments	8		8	

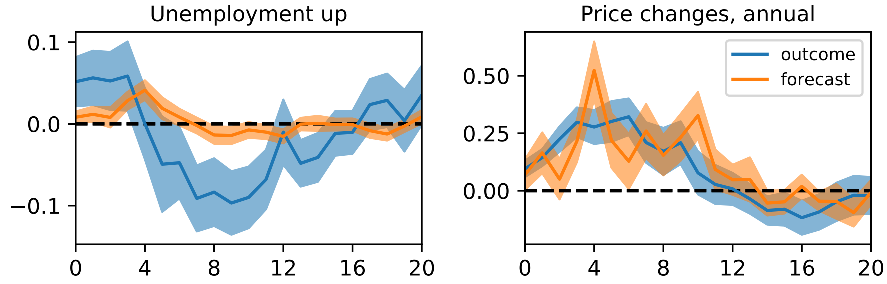
Notes: The F values are a (multivariate extension of) the Kleibergen and Paap (2006) rk statistic and can be compared with critical values reported in Stock and Yogo (2005) for given levels of tolerated bias.

Figure A.1: The Dynamic Response of Unemployment and Inflation, Robustness



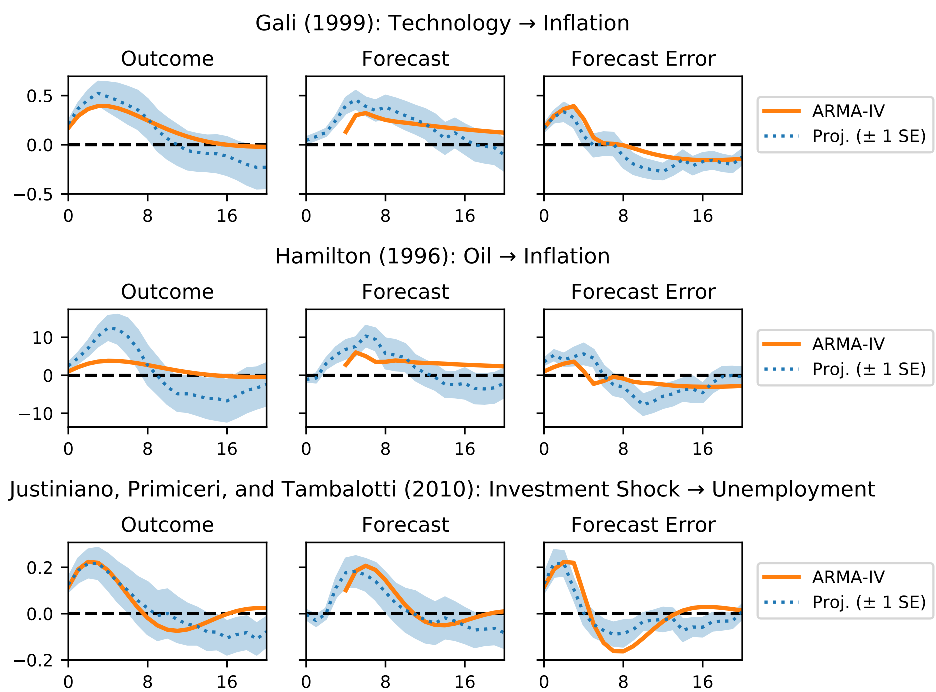
Notes: The sample period is Q4 1968 to Q4 2017. The shaded areas are 68% confidence intervals based on HAC standard errors with a Bartlett kernel and 4 lags. In the first row of each panel the outcome is u_t and the forecast is $\bar{\mathbb{E}}_{t-3}[u_t]$; in the second row the outcome is $\pi_{t,t-4}$, or annual inflation, and the forecast is $\bar{\mathbb{E}}_{t-3}[\pi_{t,t-4}]$.

Figure A.2: Dynamic Responses in the Michigan Survey



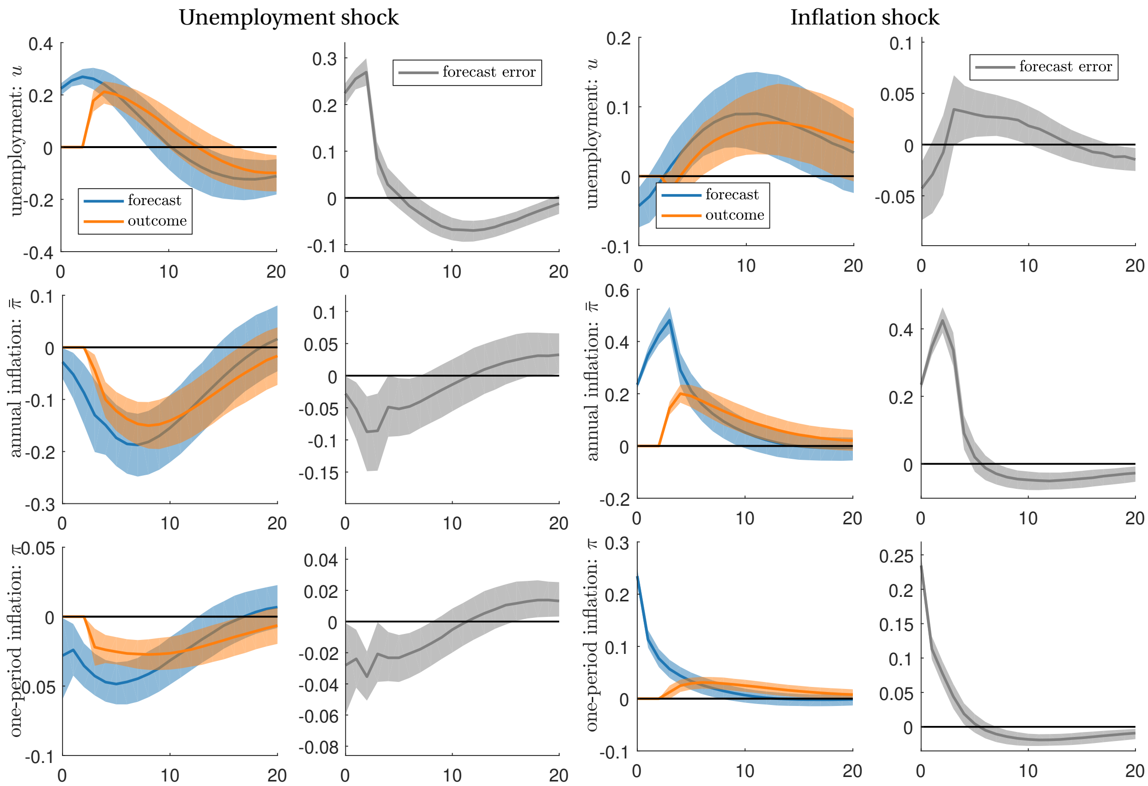
Notes: The sample period is Q1 1968 to Q4 2017. The shaded areas are 68% confidence intervals based on HAC standard errors with a Bartlett kernel and 4 lags. In the first plot the outcome is $\text{UnempUp}_{t,t-4}$ and the forecast is $\bar{\mathbb{E}}_{t-4}[\text{UnempUp}_{t,t-4}]$; in the second plot the outcome is $\pi_{t,t-4}$, or annual inflation, and the forecast is $\bar{\mathbb{E}}_{t-4}[\pi_{t,t-4}]$.

Figure A.3: Responses to Other Structural Shocks



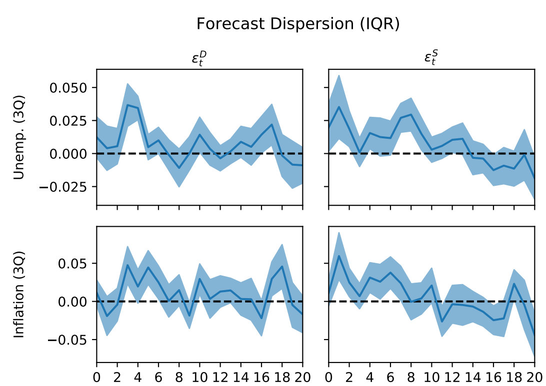
Notes: The sample period is Q4 1968 to Q4 2017. The x-axis denotes quarters from the shock (starting at 0). The shaded areas are 68% confidence intervals based on HAC standard errors with a Bartlett kernel and 4 lags. The first shock is a technology shock à la (Galí, 1999), as obtained from Coibion and Gorodnichenko (2012) and normalized to be inflationary and contractionary. The second is an oil shock à la Hamilton (1996), again obtained from Coibion and Gorodnichenko (2012). The third is the investment-specific shock of Justiniano, Primiceri, and Tambalotti (2010), updated to cover the full sample until 2017. See Appendix B for details.

Figure A.4: Dynamic Responses in the Angeletos, Collard, and Dellas (2019) SVAR, All Responses



Notes: The sample period is Q4 1968 to Q4 2017. The x-axis denotes quarters from the shock (starting at 0). The shaded areas are 68% high-posterior-density regions and the point estimate is the posterior median. In the first row the outcome is u_t and the forecast is $\bar{\mathbb{E}}_{t-3}[u_t]$; in the second row the outcome is $\pi_{t,t-4}$, or annual inflation, and the forecast is $\bar{\mathbb{E}}_{t-3}[\pi_{t,t-4}]$; and in the last row, the outcome is $\pi_{t,t-1}$, or one-quarter inflation, and the forecast is $\bar{\mathbb{E}}_{t-3}[\pi_{t,t-1}]$. The first column shows the response to a shock that maximizes the business-cycle variation in unemployment; the second for a shock that maximizes the business-cycle variation in GDP deflator inflation.

Figure A.5: Dynamic Response of Dispersion



Notes: The sample period is Q4 1968 to Q4 2017. The x-axis denotes quarters from the shock (starting at 0). The shaded areas are 68% confidence intervals based on HAC standard errors with a Bartlett kernel and 4 lags. The outcome variable is the cross-sectional interquartile range of forecasts in the SPF.

B Conditional vs. Unconditional Dynamics

Coibion and Gorodnichenko (2012) test models of expectations inertia by estimating the dynamic response of outcomes, forecasts, and forecast errors to shocks, just like this paper does in Section 5. But, while this paper and Coibion and Gorodnichenko (2012) agree about the initial under-reaction of professional forecasters to economic shocks, only the present paper finds robust evidence of the "over-shooting" that we characterize as *Fact 4*. What explains the differences in results, given that our analyses study similar data over a similar time period?

In this Appendix, we will show that a major difference is estimation methodology—and we will argue that our approach is preferable.

To this goal, we will focus on one main result from Coibion and Gorodnichenko (2012): that inflation expectations respond sluggishly to an inflationary negative supply shock. We will recreate this fact using the data directly provided in that paper for the strongest comparability, although these data are of course essentially identical to those used in our own main analysis.[43]

To identify a technology shock, the authors run a four-lag, three-variable VAR with labor productivity, the change in labor hours, and the (one-quarter-ahead) GDP deflator inflation and apply the long-run restrictions introduced by Galí (1999).[44] Finally, to make the shock inflationary like our main example shock is, we take the *negative* shock which corresponds to a technological contraction.

Their method. To estimate impulse responses, Coibion and Gorodnichenko (2012) apply the following method due to Romer and Romer (2004). For a given variable z_t (e.g., forecast errors), they estimate the empirical ARMA process via Ordinary Least Squares (OLS):

$$z_t = \alpha + \sum_{p=1}^{P} \gamma_p \cdot z_{t-p} + \sum_{k=0}^{K} \beta_k \cdot \varepsilon_{t-k} + u_t \tag{22}$$

where the $(\varepsilon_{t-k})_{k=0}^{K}$ are the identified shocks. The authors use information criteria to pick an optimal lag length combination (P, K). In the empirical application, for estimating the response of inflation, forecasts, and forecast errors to the technology shocks, they find that $K = 1$ and $P = 1$ uniformly fits the data the best subject to their chosen penalty for extra parameters.

But now note that $P = 1$ effectively *imposes* that the IRF of forecast errors cannot switch signs. Indeed, abstracting from MA term (which after all turns out to be small in their estimation), their method effectively imposes that the IRF of the average forecast error to the technology shock is that of the AR(1) process that best describes the *unconditional* dynamics of the average forecast errors.

Our method. The approach we take in Section 5 has two key differences. First, we fix a larger value of P (in our preferred specification, $P = 3$), in anticipation of the fact that the model may demand more complex dynamics than an AR(1). Second, we instrument for lagged values of z_t using past shocks. Intuitively, this isolates the possibility that dynamics may be "shock-specific" and not informed entirely by the unconditional auto-covariance patterns in z_t. This is to be expected if the data-generating process does in fact involve multiple shocks and/or variables, so thinking of the model as exactly a single-shock ARMA could be very inaccurate.

[43] There are only three salient differences. The first is that Coibion and Gorodnichenko (2012) use forecast means rather than medians as a measure of the aggregate. The second is that Coibion and Gorodnichenko (2012) measure expected annual inflation with the forecast of the 4-quarter-ahead price level relative to the now-cast of the (unreleased) current-quarter price level; whereas our main analysis uses three-quarters-ahead relative to the previous quarter. And the third is that their sample period runs from Q4 of 1974 to Q4 of 2007.

[44] The estimation period they use for this VAR covers Q2 1952 to Q3 2007.

For comparability with (22), we will estimate the following system of equations with two-stage least squares. The reduced-form equation is exactly (22) with $K = 1$ and $P = 3$ (to capture higher-order dynamics):

$$z_t = \alpha + \sum_{p=1}^{3} \gamma_p z_{t-p} + \sum_{k=0}^{1} \beta_k \varepsilon_{t-k} + u_t \tag{23}$$

The first-stage relates the lags of z_t with shocks before $t-1$. In vector form,

$$Z_{t-1} = \eta + \mathcal{E}'_{t-2}\Theta + e_t \tag{24}$$

with $Z_{t-1} = [z_{t-1}, z_{t-2}, z_{t-3}]$ and $\mathcal{E}_{t-2} = [\varepsilon_{t-j}]_{j=2}^{J}$. Like in the main text, we have $J = 9$, which means there are 8 instruments. Armed with these IV estimates of the γ and β coefficients, we can calculate an alternative impulse response.

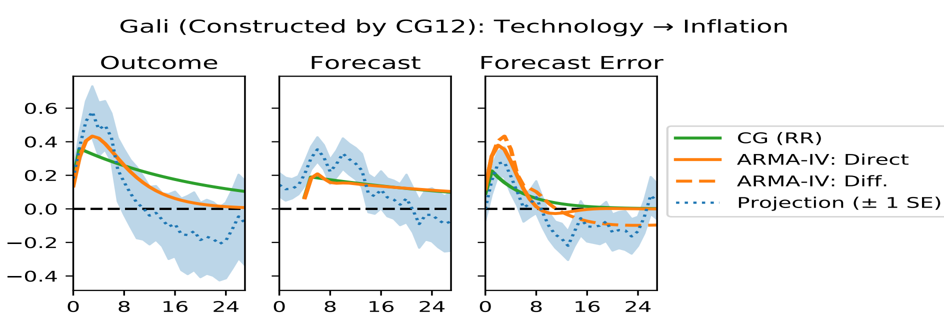
Local projections. Finally, we can also run the following local projection regression separately for each horizon h:

$$z_{t+h} = \alpha_h + \beta_{h,d} \cdot \varepsilon_t + \gamma' W_t + u_{t+h} \tag{25}$$

For controls W_t we will use the four lags each of labor productivity, the change in labor hours, and inflation that entered the original VAR. This is necessary, in the smaller sample, to make the estimated shock series truly orthogonal to lagged macro conditions.

Results. Figure A.6 compares the results, extended out to 28 quarters. Plotted in the blue dotted line, with a shaded 68% confidence interval, is the projection estimate of impulse responses for outcomes (left), forecasts (middle), and forecast errors. Plotted in green is the point estimate of the Coibion and Gorodnichenko (2012) method, or the estimate that comes from (22). Plotted in orange are the estimates from the IV method, or the combination of (23) and (24). And plotted in the orange dashed line is the difference between the orange lines for outcomes and forecasts, which is a different estimator for the response of forecast errors.

Figure A.6: Comparison of IRF Methods for Response to Technology Shock



The green lines in all cases are much more persistent than the projection responses. In the first and third case, in particular, the green lines smoothly and slowly converge back to zero. The unrestricted projection estimator, however, suggests that the response of inflation eventually turns negative (slightly, but not completely, offsetting the effects on the price level) and that the response of forecast errors also turns negative.

The ARMA-IV estimator, compared to the Coibion and Gorodnichenko (2012) method, gives a very similar response of forecast errors but a much less persistent response of the outcome. This estimation of the outcome IRF more closely matches the projection estimates. As such the "difference" estimator, or the dashed orange line in the third panel, shows evidence of over-extrapolation in the point estimate at moderate (>10 quarter) horizons. The ARMA-IV estimator directly applied to forecast errors, on the other hand, shows only modest evidence of overshooting.

Bottom line. A method that imposes uniform dynamics as if the data-generating process involved only one shock, like that introduced by Romer and Romer (2004) and adopted by Coibion and Gorodnichenko (2012), may provide a distorted picture of the *conditional* dynamics. The possible solutions include the "shock-specific" IV approach introduced here, a flexible local-projection, or a more structured multi-variate model. The trade-offs between these models involve robustness and small-sample efficiency.

C Noise and a Hybrid Regression

Proposition 1 underscored that, away from rational expectations, the CG regression coefficient is no more a measure of the informational friction alone: it is "contaminated" by the departure from the rationality. But the BGMS coefficient isolates the role of the latter. This suggests that the *gap* between the two coefficients ought to say something about the actual level of noise.

We next show how one can arrive at essentially the same answer with a "hybrid" of the CG and BGMS regressions. Let $(\mathcal{V}_{\text{ind}}, \mathcal{V}_{\text{agg}})$ respectively denote the variances of the idiosyncratic and aggregate forecast revisions.

Lemma. *The following regression holds in the theory:*

$$Error_{i,t,k} = K_{CG} \cdot Revision_{t,k} - K_{noise} \cdot \frac{\mathcal{V}_{agg}}{\mathcal{V}_{idio}} \cdot \Delta Revision_{i,t,k} + u_{i,t,k} \tag{26}$$

where $\Delta Revision_{i,t,k} \equiv Revision_{i,t,k} - Revision_{t,k}$ *and*

$$K_{noise} \equiv K_{CG} - \frac{\mathcal{V}_{ind}}{\mathcal{V}_{agg}} K_{BGMS} = \kappa_1 \tau^{-1} \tag{27}$$

From the perspective of this regression, K_{CG} measures the predictability in *individual* forecast errors attributed to the *common* component of the lagged forecast revisions, and K_{noise} the one attributed to the purely *idiosyncratic* components of the lagged forecast revisions. As already explained, the former confounds the effects of misspecification and information. The latter, which is again the gap between K_{CG} and K_{BGMS} appropriately rescaled, isolates the effect of the idiosyncratic noise.[45]

Table A.6 shows results from estimating the hybrid regression over the full and restricted samples for all horizons of forecast. Across these margins, the estimated value of K_{noise} is positive (and statistically different from zero). This is lines up with the following observation: if we go back to the results presented in Section 3 and the Appendix regarding Facts 1 and 2, we can readily verify that K_{BGMS} was consistently lower that K_{CG}, even in specifications where both were positive.[46]

Of course, as evident from the previous discussion, the hybrid regression does not provide *independent* information compared to Facts 1 and 2. The coefficients of the hybrid regression can be inferred from the original CG and BGMS regressions, and vice versa, up to small-sample differences between some moments.[47] What this regression however accomplishes is to *combine* Facts 1 and 2 in way that more clearly illustrates how the gap between K_{CG} and K_{BGMS}, or more precisely the object K_{noise} described above, provides the needed "correction" of the original CG coefficient. With rational expectations, K_{noise} coincides with K_{CG}. Away from that benchmark, K_{noise} partials out from K_{CG} the component due to irrationality. In both cases, K_{noise} isolates the effect of idiosyncratic noise.[48]

[45]To the best of our knowledge, the particular regression we propose here and the offered structural interpretation are novel. However, Fuhrer (2018) and Kohlhas and Broer (2019) contain a few empirical specifications that have a similar spirit, namely the separately test the extent to which aggregate-level and ind individual-level variables help predict forecast errors.

[46]The same seems to be true for almost all the specifications considered in Bordalo et al. (2018), including those regarding a variety of interest rates and spreads.

[47]To be precise, one also needs to compute $\mathcal{V}_{\text{idio}}$ and $\mathcal{V}_{\text{agg}}$, the variances of, respectively, the individual and aggregate forecast

Table A.6: The Hybrid Regression (26)

Sample		Unemployment 1	2	3	Inflation 1	2	3
Full	$\Delta\text{Revision}_{i,t,k}$ (-K_{noise})	-0.183	-0.189	-0.166	-0.422	-0.427	-0.346
		(0.035)	(0.043)	(0.043)	(0.047)	(0.036)	(0.042)
	$\text{Revision}_{t,k}$ (K_{agg})	0.441	0.6421	0.745	0.675	1.108	1.550
		(0.114)	(0.138)	(0.173)	(0.209)	(0.245)	(0.278)
	R^2	0.120	0.147	0.103	0.168	0.194	0.211
	N	5,808	5,699	5,383	5,496	5,458	5,147
Post 1984	$\Delta\text{Revision}_{i,t,k}$ (-K_{noise})	-0.217	-0.264	-0.162	-0.517	-0.481	-0.410
		(0.039)	(0.043)	(0.053)	(0.034)	(0.035)	(0.041)
	$\text{Revision}_{t,k}$ (K_{agg})	0.462	0.722	0.841	-0.070	0.179	0.412
		(0.159)	(0.183)	(0.210)	(0.185)	(0.178)	(0.180)
	R^2	0.136	0.195	0.152	0.106	0.085	0.072
	N	3,986	3,918	3,769	3,779	3,745	3,643

Notes: The observation is a forecaster by quarter between Q4-1968 and Q4-2017. Standard errors are clustered two- way by forecaster ID and time period. Both errors and revisions are winsorized over the sample to restrict to 4 times the inter-quartile range away from the median. The data used for outcomes are first-release ("vintage").

D Proofs

The following lemma, which is proved below, will help in proving the results in the main text:

Lemma 1. *The one-step-ahead forecasts obey*

$$\mathbb{E}_{i,t}[x_{t+1}] = \hat{\rho}\mathbb{E}_{i,t}[x_t] = (\hat{\rho}-\hat{\lambda})\frac{1}{1-\hat{\lambda}\mathbb{L}}s_{i,t} = (\hat{\rho}-\hat{\lambda})\frac{1}{1-\hat{\lambda}\mathbb{L}}\left(\frac{1}{1-\rho\mathbb{L}}\varepsilon_t + \tau^{-\frac{1}{2}}u_{i,t}\right)$$

The corresponding forecast errors obey

$$Error_{i,t} = x_{t+1} - \mathbb{E}_{i,t}[x_{t+1}] = \frac{1-\hat{\rho}L}{(1-\rho L)(1-\hat{\lambda}L)}\varepsilon_{t+1} - \frac{\hat{\rho}-\hat{\lambda}}{1-\hat{\lambda}\mathbb{L}}\tau^{-\frac{1}{2}}u_{i,t}$$

And finally the forecast revisions obey

$$Revision_{i,t} = \mathbb{E}_{i,t}[x_{t+1}] - \mathbb{E}_{i,t-1}[x_{t+1}] = \frac{(\hat{\rho}-\hat{\lambda})(1-\hat{\rho}L)}{(1-\rho L)(1-\hat{\lambda}L)}\varepsilon_t + \frac{(\hat{\rho}-\hat{\lambda})(1-\hat{\rho}L)}{1-\hat{\lambda}\mathbb{L}}\tau^{-\frac{1}{2}}u_{i,t}$$

Proof. The perceived signal process can be represented as

$$s_{i,t} = \mathbf{M}(\mathbb{L})\begin{bmatrix}\varepsilon_t \\ u_{i,t}\end{bmatrix}, \qquad \text{with} \qquad \mathbf{M}(\mathbb{L}) = \begin{bmatrix}\frac{1}{1-\hat{\rho}\mathbb{L}} & \hat{\tau}^{-\frac{1}{2}}\end{bmatrix}.$$

Let $B(\mathbb{L})$ denote the fundamental representation of the perceived signal process,[49] which is given by

$$B(\mathbb{L}) = \hat{\tau}^{-\frac{1}{2}}\sqrt{\frac{\hat{\rho}}{\hat{\lambda}}\frac{1-\hat{\lambda}\mathbb{L}}{1-\hat{\rho}\mathbb{L}}}, \quad \text{where} \quad \hat{\lambda} = \frac{1}{2}\left(\hat{\rho} + \frac{1+\hat{\tau}}{\hat{\rho}} + \sqrt{\left(\hat{\rho}+\frac{1+\hat{\tau}}{\hat{\rho}}\right)^2 - 4}\right).$$

revisions. But these variances are already implicit in the calculation of K_{BGMS} and K_{CG}.

[48] The following caveat applies to the adopted interpretation of K_{noise}. In the model we work with in this paper, idiosyncratic noise is the sole source of heterogeneity in beliefs: irrationality is a (possibly time-varying) fixed effect in the cross-section of the population. Without this restriction, K_{noise} may confound the effects of “rational” noise (due to idiosyncratic information) and “irrational” noise (due to idiosyncratic misspecification).

[49] $B(\mathbb{L})$ satisfies the requirement $B(\mathbb{L})B'(\mathbb{L}^{-1}) = M(\mathbb{L})M'(\mathbb{L}^{-1})$ and $B(\mathbb{L})$ is invertible.

It is useful to note that $\hat{\lambda} < \hat{\rho}$, and $\hat{\lambda}$ is decreasing in $\hat{\tau}$. By the Wiener-Hopf prediction formula, the individual forecast about x_t is

$$\mathbb{E}_{i,t}[x_t] = \left[\frac{1}{1-\hat{\rho}\mathbb{L}}\mathbf{M}'(\mathbb{L}^{-1})B(\mathbb{L}^{-1})^{-1}\right]_+ B(\mathbb{L})^{-1}s_{i,t} = \left(1-\frac{\hat{\lambda}}{\hat{\rho}}\right)\frac{1}{1-\hat{\lambda}\mathbb{L}}s_{i,t}.$$

Alternatively, this forecast rule can be written as

$$\mathbb{E}_{i,t}[x_t] = (1-\hat{g})\hat{\rho}\,\mathbb{E}_{i,t-1}[x_{t-1}] + \hat{g}\,s_{i,t},$$

which is a weighted average of the prior $\rho\mathbb{E}_{it-1}[x_{t-1}]$ and the new signal $s_{i,t}$, where the weight on the signal is the Kalman gain $\hat{g} = 1-\frac{\hat{\lambda}}{\hat{\rho}}$. In the equations above, note that only perceived $\hat{\rho}$ and $\hat{\tau}$ matter for how agents use their signals. The actual ρ and τ matter for how the signal $s_{i,t}$ evolves overtime.

Proof. Accordingly, the one-period ahead forecast is

$$\mathbb{E}_{i,t}[x_{t+1}] = \hat{\rho}\mathbb{E}_{i,t}[x_t] = (\hat{\rho}-\hat{\lambda})\frac{1}{1-\hat{\lambda}\mathbb{L}}s_{i,t} = (\hat{\rho}-\hat{\lambda})\frac{1}{1-\hat{\lambda}\mathbb{L}}\left(\frac{1}{1-\rho\mathbb{L}}\varepsilon_t + \tau^{-\frac{1}{2}}u_{i,t}\right).$$

The individual forecast error and revision are then straightforward to obtain:

$$\text{Error}_{i,t} = x_{t+1} - \mathbb{E}_{i,t}[x_{t+1}] = \frac{1-\hat{\rho}L}{(1-\rho L)(1-\hat{\lambda}L)}\varepsilon_{t+1} - \frac{\hat{\rho}-\hat{\lambda}}{1-\hat{\lambda}\mathbb{L}}\tau^{-\frac{1}{2}}u_{i,t},$$

$$\text{Revision}_{i,t} = \mathbb{E}_{i,t}[x_{t+1}] - \mathbb{E}_{i,t-1}[x_{t+1}] = \frac{(\hat{\rho}-\hat{\lambda})(1-\hat{\rho}L)}{(1-\rho L)(1-\hat{\lambda}L)}\varepsilon_t + \frac{(\hat{\rho}-\hat{\lambda})(1-\hat{\rho}L)}{1-\hat{\lambda}\mathbb{L}}\tau^{-\frac{1}{2}}u_{i,t}.$$

□

Proof of Proposition 1

Let $\mathcal{V}_{\text{ind}}$ denote the variance of $\text{Revision}_{i,t}$ and $\mathcal{V}_{\text{agg}}$ denote the variance of Revision_t. First consider the calculation of K_{CG}. We have

$$\begin{aligned}\text{Cov}(\text{Error}_t, \text{Revision}_t) =& \text{Cov}\left(\frac{1-\hat{\rho}L}{(1-\rho L)(1-\hat{\lambda}L)}\varepsilon_{t+1}, \frac{(\hat{\rho}-\hat{\lambda})(1-\hat{\rho}L)}{(1-\rho L)(1-\hat{\lambda}L)}\varepsilon_t\right)\\ =&(\hat{\rho}-\hat{\lambda})\left(\frac{\hat{\lambda}}{1-\hat{\lambda}^2} + (\rho-\hat{\rho})\frac{(1+\hat{\lambda}^2)(1-\rho^2)+(\hat{\lambda}+\rho)(\rho-\hat{\rho})}{(1-\hat{\lambda}^2)(1-\rho^2)(1-\hat{\lambda}\rho)}\right)\\ =&\hat{\tau}^{-1}\frac{(\hat{\rho}-\hat{\lambda})^2(1-\hat{\lambda}\hat{\rho})}{1-\hat{\lambda}^2} + (\rho-\hat{\rho})(\hat{\rho}-\hat{\lambda})\frac{(1+\hat{\lambda}^2)(1-\rho^2)+(\hat{\lambda}+\rho)(\rho-\hat{\rho})}{(1-\hat{\lambda}^2)(1-\rho^2)(1-\hat{\lambda}\rho)},\end{aligned}$$

which leads to

$$K_{CG} = \kappa_1\hat{\tau}^{-1} - \kappa_2(\hat{\rho}-\rho),$$

where

$$\kappa_1 = \frac{1}{\mathcal{V}_{\text{agg}}}\frac{(\hat{\rho}-\hat{\lambda})^2(1-\hat{\lambda}\hat{\rho})}{1-\hat{\lambda}^2}, \quad \kappa_2 = \frac{1}{\mathcal{V}_{\text{agg}}}(\hat{\rho}-\hat{\lambda})\frac{(1+\hat{\lambda}^2)(1-\rho^2)+(\hat{\lambda}+\rho)(\rho-\hat{\rho})}{(1-\hat{\lambda}^2)(1-\rho^2)(1-\hat{\lambda}\rho)}.$$

As $1 > \hat{\rho} > \hat{\lambda} > 0$, $\kappa_1 > 0$. To show that $\kappa_2 > 0$, it is equivalent to show that $(1+\hat{\lambda}^2)(1-\rho^2)+(\hat{\lambda}+\rho)(\rho-\hat{\rho}) > 0$. Given that $\hat{\rho} < 1$, it follows that

$$(1+\hat{\lambda}^2)(1-\rho^2)+(\hat{\lambda}+\rho)(\rho-\hat{\rho}) > (1+\hat{\lambda}^2)(1-\rho^2)+(\hat{\lambda}+\rho)(\rho-1) = (1-\rho)(1-\hat{\lambda}+\hat{\lambda}^2(1+\rho)) > 0.$$

Now turn to the calculation of K_{BGMS}. We have

$$\begin{aligned}
&\text{Cov}(\text{Error}_{i,t}, \text{Revision}_{i,t}) \\
=&\text{Cov}\left(\frac{1-\hat{\rho}\mathbb{L}}{(1-\rho\mathbb{L})(1-\hat{\lambda}\mathbb{L})}\varepsilon_{t+1}, \frac{(\hat{\rho}-\hat{\lambda})(1-\hat{\rho}\mathbb{L})}{(1-\rho\mathbb{L})(1-\hat{\lambda}\mathbb{L})}\varepsilon_t\right) + \text{Cov}\left(-\frac{\hat{\rho}-\hat{\lambda}}{1-\hat{\lambda}\mathbb{L}}\tau^{-\frac{1}{2}}u_{i,t}, \frac{(\hat{\rho}-\hat{\lambda})(1-\hat{\rho}\mathbb{L})}{1-\hat{\lambda}\mathbb{L}}\tau^{-\frac{1}{2}}u_{i,t}\right) \\
=&-(\hat{\rho}-\hat{\lambda})\frac{\hat{\lambda}}{(1-\hat{\lambda}^2)}\frac{\hat{\tau}-\tau}{\tau} + (\rho-\hat{\rho})(\hat{\rho}-\hat{\lambda})\frac{(1+\hat{\lambda}^2)(1-\rho^2)+(\hat{\lambda}+\rho)(\rho-\hat{\rho})}{(1-\hat{\lambda}^2)(1-\rho^2)(1-\hat{\lambda}\rho)}.
\end{aligned}$$

It follows that

$$K_{\text{BGMS}} = -\kappa_5(\tau^{-1}-\hat{\tau}^{-1}) + \kappa_6(\rho-\hat{\rho}),$$

where κ_5 and κ_6 are

$$\kappa_5 = \frac{1}{\mathcal{V}_{\text{ind}}}(\hat{\rho}-\hat{\lambda})\frac{\hat{\lambda}}{\tau(1-\hat{\lambda}^2)}, \quad \kappa_6 = \frac{1}{\mathcal{V}_{\text{ind}}}(\hat{\rho}-\hat{\lambda})\frac{(1+\hat{\lambda}^2)(1-\rho^2)+(\hat{\lambda}+\rho)(\rho-\hat{\rho})}{(1-\hat{\lambda}^2)(1-\rho^2)(1-\hat{\lambda}\rho)}.$$

Lastly, we look at K_{KW}. We have

$$\text{Cov}(\text{Error}_t, x_t) = \text{Cov}\left(\frac{1-\hat{\rho}\mathbb{L}}{(1-\rho\mathbb{L})(1-\hat{\lambda}\mathbb{L})}\varepsilon_{t+1}, \frac{1}{1-\rho\mathbb{L}}\varepsilon_t\right) = \frac{\hat{\tau}^{-1}(\hat{\rho}-\hat{\lambda})(1-\hat{\lambda}\hat{\rho})(1-\rho^2)+\rho-\hat{\rho}}{(1-\hat{\lambda}\rho)(1-\rho^2)},$$

which leads to

$$K_{\text{KW}} = \kappa_3\hat{\tau}^{-1} - \kappa_4(\hat{\rho}-\rho),$$

where κ_3 and κ_4 are

$$\kappa_3 = \frac{\hat{\rho}-\hat{\lambda}}{1-\rho^2}, \quad \kappa_4 = \frac{1}{(1-\hat{\lambda}\rho)(1-\rho^2)^2}.$$

Proof of Corollary 1

As $\hat{\tau} = \tau$ and $\hat{\rho} = \rho$, we have

$$\text{Cov}(\text{Error}_t, \text{Revision}_t) = \hat{\tau}^{-1}\frac{(\hat{\rho}-\hat{\lambda})^2(1-\hat{\lambda}\hat{\rho})}{1-\hat{\lambda}^2} > 0$$

$$\text{Cov}(\text{Error}_{i,t}, \text{Revision}_{i,t}) = 0 \qquad \text{and} \qquad \text{Cov}(\text{Error}_t, x_t) = \hat{\tau}^{-1}(\hat{\rho}-\hat{\lambda}) > 0,$$

which together imply

$$K_{\text{CG}} = \kappa_1\hat{\tau}^{-1} > 0, \quad K_{\text{BGMS}} = 0, \quad K_{\text{KW}} = \kappa_3\hat{\tau}^{-1} > 0.$$

As $\tau = \hat{\tau} \to \infty$, $\hat{\lambda} \to 0$ and it follows that

$$\kappa_1 \to \frac{\hat{\rho}^2}{\mathcal{V}_{\text{agg}}}, \quad \kappa_2 = \frac{1}{\mathcal{V}_{\text{agg}}}\hat{\rho}\frac{1-\rho\hat{\rho}}{1-\rho^2}, \quad \kappa_3 \to \frac{\hat{\rho}}{1-\rho^2}, \quad \kappa_4 = \frac{1}{(1-\rho^2)^2}, \quad \kappa_5 \to 0, \quad \kappa_6 = \frac{1}{\mathcal{V}_{\text{agg}}}\hat{\rho}\frac{1-\rho\hat{\rho}}{1-\rho^2}$$

As a result, the signs of the three regression coefficients are the same as the sign of $\rho - \hat{\rho}$.

Proof of Corollary 2

With $\hat{\rho} \leq \rho$, $K_{\text{KW}} > 0$, and therefore $\hat{\rho} > \rho$ is necessary to make $K_{\text{KW}} < 0$. With $\hat{\tau} = \infty$ and $\hat{\rho} > \rho$, both $K_{\text{CG}} < 0$ and $K_{\text{KW}} < 0$. Therefore, it is necessary to have both $\hat{\tau} < \infty$ and $\hat{\rho} > \rho$ to allow $K_{\text{CG}} > 0$ and $K_{\text{KW}} < 0$. The sufficiency part is established by the numerical example and a standard continuity argument.

Proof of Proposition 2

The law of motion of the average forecast error is given by

$$\text{Error}_t = \frac{1-\hat{\rho}\mathbb{L}}{(1-\rho\mathbb{L})(1-\hat{\lambda}\mathbb{L})}\varepsilon_{t+1} = \left(\frac{\rho-\hat{\rho}}{\rho-\hat{\lambda}}\frac{1}{1-\rho L} + \frac{\hat{\rho}-\hat{\lambda}}{\rho-\hat{\lambda}}\frac{1}{1-\hat{\lambda}L}\right)\varepsilon_{t+1}.$$

Suppose $\rho > \hat{\rho}$, then $\rho > \hat{\lambda}$. The coefficients of the two AR(1) terms are both positive, and the responses are therefore all positive.

Suppose $\rho < \hat{\rho}$. Consider the following continuous time version of the response

$$g(t) = \frac{\rho-\hat{\rho}}{\rho-\hat{\lambda}}\rho^t + \frac{\hat{\rho}-\hat{\lambda}}{\rho-\hat{\lambda}}\hat{\lambda}^t,$$

and $g(t) = \zeta_k$ when $t = k \in \{0,1,\ldots\}$. Note that: (1) $g(t)$ is negative when t is large enough (no matter $\rho > \hat{\lambda}$ or $\rho < \hat{\lambda}$); (2) when $t = 0$, $g(0) = 1 > 0$; (3) there is at most one root of $g(t)$. As a result, $\{\zeta_k\}_{k=1}^{\infty}$ eventually stay negative, but they might be positive or negative for k small enough.

The root of $g(t)$ is

$$k_{\text{IRF}} = \frac{\log(\hat{\rho}-\rho) - \log(\hat{\rho}-\hat{\lambda})}{\log\hat{\lambda} - \log\rho}.$$

To have $\{\zeta_k\}_{k=1}^{\infty}$ switch signs, it is necessary that $g(1) > 0$ and $\hat{\rho} > \rho$, which correspond to $g(1) = \rho + \hat{\lambda} - \hat{\rho} > 0$ and $\hat{\rho} > \rho$, or

$$\hat{\lambda} > \hat{\rho} - \rho \quad \text{and} \quad \hat{\rho} > \rho.$$

Finally, note that k_{IRF} is decreasing in $\hat{\rho}$ for given $\hat{\lambda}$, which verifies the claim in the main text that the magnitude of k_{IRF} reveals information about the relative importance of the two mechanisms.[50]

When $\hat{\rho} > \rho$ but $\hat{\lambda} > \hat{\rho} - \rho$, $g(1) < 0$ and the sequences $\{\zeta_k\}_{k=1}^{\infty}$ stay negative all the time.

Proof of Corollary 3

Follows directly from Proposition 2.

Proof of Proposition 3

Aggregate consumption satisfies the fixed point restriction

$$c_t = \sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t[\xi_{t+k}] + (1-\beta)\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t[c_{t+k+1}],$$

where we have used the market clearing condition $y_t = c_t$, and the assumption that agents observe y_t but do not extract information from it. This aggregate outcome is the outcome of the following beauty-contest game

$$c_{i,t} = \mathbb{E}_{i,t}[\xi_t] + \beta\mathbb{E}_{i,t}[c_{i,t+1}] + (1-\beta)\mathbb{E}_{i,t}[c_{t+1}].$$

[50] On a more technical level, note that, as written, k_{IRF} need not be an integer. It is indeed obtained from the continuous-time limit of the ARMA process that describes the average forecast error. But the result, as stated, holds for the true, discrete-time process. Also, a small caveat is that in the model with over-extrapolation ($\hat{\rho} > \rho$) but no noise ($\hat{\tau} \to \infty$), *right after* $t = 0$, the forecast error switches sign from positive (by construction, given that the data was hit by an unpredictable innovation) to negative (as a result of flawed reasoning). That is, $\zeta_0 = 1$ always, but we can have $\lim_{h\downarrow 0}\zeta_h < 0$. In the data, given that we properly observe some average of forecast errors between $t = 0$ and $t = 1$ as the "observation" at $t = 0$, we would expect to see impulse responses of uniform sign.

Denote the agent's equilibrium policy function as

$$c_{i,t} = h(\mathbb{L}) s_{i,t}$$

for some lag polynomial $h(\mathbb{L})$. The actual law of motion of aggregate outcome can then be expressed as follows

$$c_t = h(\mathbb{L})\xi_t = \frac{h(\mathbb{L})}{1-\rho\mathbb{L}}\varepsilon_t.$$

However, the perceived law of motion by consumers is

$$c_t = \frac{h(\mathbb{L})}{1-\hat{\rho}\mathbb{L}}\varepsilon_t.$$

As in the case where the outcome is given by the exogenous AR(1) process, the forecast about the fundamental is

$$\mathbb{E}_{i,t}[\xi_t] = \left(1 - \frac{\hat{\lambda}}{\hat{\rho}}\right)\frac{1}{1-\hat{\lambda}\mathbb{L}} s_{i,t} \equiv G_1(\mathbb{L}) s_{i,t}.$$

Consider the forecast of the future own and average actions. The perceived law of motion of $c_{i,t+1}$ and c_{t+1} are

$$c_{t+1} = \begin{bmatrix} \frac{h(\mathbb{L})}{L(1-\hat{\rho}\mathbb{L})} & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ u_{i,t} \end{bmatrix}, \quad c_{i,t+1} - c_{t+1} = \begin{bmatrix} 0 & \hat{\tau}^{-\frac{1}{2}} \frac{h(\mathbb{L})}{\mathbb{L}} \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ u_{i,t} \end{bmatrix},$$

and the forecasts are

$$\mathbb{E}_{i,t}\left[c_{t+1}\right] = G_2(\mathbb{L}) s_{i,t}, \qquad G_2(\mathbb{L}) \equiv \frac{\hat{\lambda}}{\hat{\rho}}\hat{\tau}\left(\frac{h(\mathbb{L})}{(1-\hat{\lambda}\mathbb{L})(\mathbb{L}-\hat{\lambda})} - \frac{h(\hat{\lambda})(1-\hat{\rho}\mathbb{L})}{(1-\hat{\rho}\hat{\lambda})(\mathbb{L}-\hat{\lambda})(1-\hat{\lambda}\mathbb{L})}\right),$$

$$\mathbb{E}_{i,t}\left[c_{i,t+1} - c_{t+1}\right] = G_3(\mathbb{L}) s_{i,t}, \qquad G_3(\mathbb{L}) \equiv \frac{\hat{\lambda}}{\hat{\rho}}\left(\frac{h(\mathbb{L})(\mathbb{L}-\hat{\rho})}{\mathbb{L}(\mathbb{L}-\hat{\lambda})} - \frac{h(\hat{\lambda})(\hat{\lambda}-\hat{\rho})}{\hat{\lambda}(\mathbb{L}-\hat{\lambda})} - \frac{\hat{\rho}}{\hat{\lambda}}\frac{h(0)}{\mathbb{L}}\right)\frac{1-\hat{\rho}\mathbb{L}}{1-\hat{\lambda}\mathbb{L}}$$

Recall that fixed point problem that characterizes the equilibrium is

$$c_{i,t} = \mathbb{E}_{i,t}[\xi_t] + \beta\mathbb{E}_{i,t}[c_{i,t+1}] + (1-\beta)\mathbb{E}_{i,t}[c_{t+1}].$$

We can replace the left-hand side with $h(\mathbb{L})s_{i,t}$. Using the results derived above, on the other hand, we can replace the right-hand side with $\left[G_1(\mathbb{L}) + G_2(\mathbb{L}) + \beta G_3(\mathbb{L})\right] s_{i,t}$. It follows that in equilibrium

$$h(\mathbb{L}) = G_1(\mathbb{L}) + G_2(\mathbb{L}) + \beta G_3(\mathbb{L}).$$

Equivalently, we need to find an analytic function $h(z)$ that solves

$$\begin{aligned} h(z) = {} & \frac{\hat{\lambda}}{\hat{\rho}}\hat{\tau}\frac{1}{1-\hat{\rho}\hat{\lambda}}\frac{1}{1-\hat{\lambda}z} + \frac{\hat{\lambda}}{\hat{\rho}}\hat{\tau}\left(\frac{h(z)}{(1-\hat{\lambda}z)(z-\hat{\lambda})} - \frac{h(\hat{\lambda})(1-\hat{\rho}z)}{(1-\hat{\rho}\hat{\lambda})(z-\hat{\lambda})(1-\hat{\lambda}z)}\right) \\ & + \beta\frac{\hat{\lambda}}{\hat{\rho}}\left(\frac{h(z)(z-\hat{\rho})}{z(z-\hat{\lambda})} - \frac{h(\hat{\lambda})(\hat{\lambda}-\hat{\rho})}{\hat{\lambda}(z-\hat{\lambda})} - \frac{\hat{\rho}}{\hat{\lambda}}\frac{h(0)}{z}\right)\frac{1-\hat{\rho}z}{1-\hat{\lambda}z}, \end{aligned}$$

which can be transformed as

$$\tilde{C}(z)h(z) = d(z; h(\hat{\lambda}), h(0))$$

where

$$\tilde{C}(z) \equiv z(1-\hat{\lambda}z)(z-\hat{\lambda}) - \frac{\hat{\lambda}}{\hat{\rho}}\left\{\beta(z-\hat{\rho})(1-\hat{\rho}z) + \hat{\tau}z\right\}$$

$$d(z; h(\hat{\lambda}), h(0)) \equiv \frac{\hat{\lambda}}{\hat{\rho}}\hat{\tau}\frac{1}{1-\hat{\rho}\hat{\lambda}}z(z-\hat{\lambda}) - \frac{1}{\hat{\rho}}\left(\hat{\tau}\frac{\hat{\lambda}}{1-\hat{\rho}\hat{\lambda}} + \beta(\hat{\lambda}-\hat{\rho})\right)z(1-\hat{\rho}z)h(\hat{\lambda}) - \beta(z-\hat{\lambda})(1-\hat{\rho}z)h(0)$$

Note that $\tilde{C}(z)$ is a cubic equation and therefore contains with three roots. We will verify later that there are two inside roots and one outside root. To make sure that $h(z)$ is an analytic function, we choose $h(0)$ and $h(\hat{\lambda})$ so that the two roots of $d(z; h(\hat{\lambda}), h(0))$ are the same as the two inside roots of $\tilde{C}(z)$. This pins down the constants $\{h(0), h(\hat{\lambda})\}$, and therefore the policy function $h(\mathbb{L})$ is

$$h(\mathbb{L}) = \left(1 - \frac{\vartheta}{\hat{\rho}}\right) \frac{1}{1-\hat{\rho}} \frac{1}{1-\vartheta\mathbb{L}},$$

where ϑ^{-1} is the root of $\tilde{C}(z)$ outside the unit circle.

Now we verify that $\tilde{C}(z)$ has two inside roots and one outside root. $\tilde{C}(z)$ can be rewritten as $\hat{\lambda}C(z)$ where

$$\begin{aligned} C(z) &= -z^3 + \left(\hat{\rho} + \frac{1}{\hat{\rho}} + \frac{1}{\hat{\rho}}\hat{\tau} + \beta\right) z^2 - \left(1 + \beta\left(\hat{\rho} + \frac{1}{\hat{\rho}}\right) + \frac{1}{\hat{\rho}}\hat{\tau}\right) z + \beta, \\ &= -z^3 + \left(\hat{\rho} + \frac{1}{\hat{\rho}} + \frac{1}{\hat{\rho}}\hat{\tau} + 1 - \text{mpc}\right) z^2 - \left(1 + (1-\text{mpc})\left(\hat{\rho} + \frac{1}{\hat{\rho}}\right) + \frac{1}{\hat{\rho}}\hat{\tau}\right) z + 1 - \text{mpc}. \end{aligned}$$

With the assumption that $1 > \text{mpc} > 0$, it is straightforward to verify that the following properties hold:

$$C(0) = 1 - \text{mpc} > 0, \quad C(\hat{\lambda}) = -\text{mpc}\frac{\hat{\tau}}{\hat{\rho}} < 0, \quad C(1) = \text{mpc}\left(\frac{1}{\hat{\rho}} + \hat{\rho} - 2\right) > 0.$$

Therefore, the three roots are all real, two of them are between 0 and 1, and the third one ϑ^{-1} is larger than 1.

To show that ϑ is less than $\hat{\rho}$, it is sufficient to show that

$$C\left(\frac{1}{\hat{\rho}}\right) = \frac{\hat{\tau}(1-\hat{\rho})}{\hat{\rho}^3} > 0.$$

Since $C(\vartheta^{-1}) = 0$, it has to be that ϑ^{-1} is larger than $\hat{\rho}^{-1}$, or $\vartheta < \hat{\rho}$.

Similarly, to show that ϑ is larger than $\hat{\lambda}$, it is sufficient to show that

$$C\left(\frac{1}{\hat{\lambda}}\right) = -\frac{\hat{\tau}(1-\text{mpc})\text{mpc}}{\hat{\rho}\hat{\lambda}^2} < 0.$$

Therefore, it has to be that $\vartheta > \hat{\lambda}$. In the proof the properties of the expectations, we will utilize the fact that $\vartheta \in (\hat{\lambda}, \hat{\rho})$.

In Angeletos and Huo (2019), the equilibrium policy rule is derived under $\rho = \hat{\rho}$ and $\tau = \hat{\tau}$. In the derivation above, note that $h(\mathbb{L})$ does not depend on ρ nor τ. The actual law of motion of $y_t = c_t$ will depend on ρ:

$$y_t = \frac{1}{1-\hat{\rho}}\left(1 - \frac{\vartheta}{\hat{\rho}}\right) \frac{1}{1-\vartheta\mathbb{L}} \frac{1}{1-\rho\mathbb{L}} \varepsilon_t.$$

On the other hand, the frictionless case is given by

$$y_t^* = \frac{1}{1-\rho} \frac{1}{1-\rho\mathbb{L}} \varepsilon_t.$$

Combining these two leads to

$$y_t = \left(1 - \frac{\vartheta}{\hat{\rho}}\right)\left(1 + \frac{\hat{\rho} - \rho}{1-\hat{\rho}}\right)\left(\frac{1}{1-\vartheta\mathbb{L}}\right) y_t^*.$$

Turn to the forecast of the future outcome. By the Wiener-Hopf prediction formula, the individual forecast is

$$\begin{aligned} \mathbb{E}_{i,t}[y_{t+1}] &= \left[\frac{1}{1-\hat{\rho}}\left(1 - \frac{\vartheta}{\hat{\rho}}\right) \frac{1}{1-\vartheta\mathbb{L}} \frac{1}{1-\rho\mathbb{L}} \mathbf{M}'(\mathbb{L}^{-1}) B(\mathbb{L}^{-1})^{-1}\right]_+ B(\mathbb{L})^{-1} s_{i,t}, \\ &= \frac{1}{1-\hat{\rho}}\left(1 - \frac{\vartheta}{\hat{\rho}}\right)\left(1 - \frac{\hat{\lambda}}{\hat{\rho}}\right) \frac{1}{1-\vartheta\hat{\lambda}} \frac{\hat{\rho} + \vartheta - \hat{\rho}\vartheta(\mathbb{L} + \hat{\lambda})}{(1-\vartheta\mathbb{L})(1-\hat{\lambda}\mathbb{L})} s_{i,t}, \end{aligned}$$

and the average forecast is

$$\overline{\mathbb{E}}_t[y_{t+1}] = \left(1 - \frac{\hat{\lambda}}{\hat{\rho}}\right) \frac{1}{1-\vartheta\hat{\lambda}} \frac{\hat{\rho} + \vartheta - \hat{\rho}\vartheta(\mathbb{L} + \hat{\lambda})}{(1-\vartheta\mathbb{L})(1-\hat{\lambda}\mathbb{L})} \left(1 - \frac{\vartheta}{\hat{\rho}}\right)\left(1 + \frac{\hat{\rho} - \rho}{1-\hat{\rho}}\right) y_t^*$$

Proof of Proposition 4

Denote $\kappa \equiv \left(1-\frac{\vartheta}{\hat{\rho}}\right)\left(1+\frac{\hat{\rho}-\rho}{1-\hat{\rho}}\right)\frac{1}{1-\rho}$. If $c_t = \kappa \frac{1}{(1-\vartheta L)(1-\rho \mathbb{L})}\varepsilon_t$ is the perfect-information outcome, it has to be that

$$
\begin{aligned}
c_t =& \xi_t + \omega_f \mathbb{E}_t^*[c_{t+1}] + \omega_b c_{t-1} \\
=& \frac{1}{1-\rho\mathbb{L}} + \omega_f \kappa \frac{\vartheta+\rho-\vartheta\rho\mathbb{L}}{(1-\vartheta\mathbb{L})(1-\rho\mathbb{L})} + \omega_b \kappa \frac{\mathbb{L}}{(1-\vartheta\mathbb{L})(1-\rho\mathbb{L})}
\end{aligned}
$$

where the right-hand side is simply the perfect information expectation of the behavioral equilibrium. This leads to

$$
\omega_f = \frac{\hat{\rho}^2-\vartheta}{(\vartheta+\rho)(\hat{\rho}-\vartheta)} \qquad \text{and} \qquad \omega_b = \frac{\vartheta(\rho(\hat{\rho}-\vartheta)+\vartheta\hat{\rho}(1-\hat{\rho}))}{(\vartheta+\rho)(\hat{\rho}-\vartheta)}.
$$

In the absence of informational friction ($\hat{\tau}\to\infty$), we have $\vartheta = 0$ and therefore $\omega_b = 0$ and $\omega_f = \hat{\rho}/\rho$. In its presence ($\hat{\tau} < \infty$), we have that $\vartheta > 0$ and $\omega_b > 0$ necessarily. When $\hat{\rho} < \rho$, we have

$$
\omega_f = \frac{\hat{\rho}^2-\vartheta}{(\vartheta+\rho)(\hat{\rho}-\vartheta)} < \frac{\hat{\rho}^2-\vartheta}{(\vartheta+\hat{\rho})(\hat{\rho}-\vartheta)} < \frac{\hat{\rho}^2-\vartheta^2}{(\vartheta+\hat{\rho})(\hat{\rho}-\vartheta)} = 1.
$$

Note that ϑ is decreasing in $\hat{\tau}$. With a very level of high informational friction ($\hat{\tau}\to 0$), we have $\vartheta = \hat{\rho}$. Particularly, when $\vartheta \in (\hat{\rho}^2, \hat{\rho})$, $\omega_f$is negative. Therefore, in order to show that $\omega_f < 1$when $\hat{\tau}$ is small, it is sufficient to show that ω_f is decreasing in ϑ. Note that

$$
\frac{\partial \omega_f}{\partial \vartheta} = \frac{\rho(\hat{\rho}^2-\hat{\rho})-\vartheta^2-\hat{\rho}^3+2\hat{\rho}^2\vartheta}{(\rho+\vartheta)^2(\hat{\rho}-\vartheta)^2},
$$

where the numerator is linear in ρ with a negative slope. To verify $\frac{\partial\omega_f}{\partial\vartheta} < 0$,we only need to show that the numerator is negative when $\rho = 0$, or $g(\vartheta) \equiv -\vartheta^2 - \hat{\rho}^3 + 2\hat{\rho}^2\vartheta < 0$. Note that $g(\vartheta)$ is maximized at $\vartheta = \hat{\rho}^2$, and $g(\hat{\rho}^2) = \hat{\rho}^4 - \hat{\rho}^3 < 0$, which completes the proof.

Proof of Proposition 5

Properties of Average Forecast Errors. The average forecast error is given by

$$
\begin{aligned}
y_{t+1} - \overline{\mathbb{E}}_t[y_{t+1}] =& \frac{1}{1-\hat{\rho}}\left(1-\frac{\vartheta}{\hat{\rho}}\right)\frac{1}{1-\vartheta\mathbb{L}}\xi_{t+1} - \frac{1}{1-\hat{\rho}}\left(1-\frac{\vartheta}{\hat{\rho}}\right)\left(1-\frac{\hat{\lambda}}{\hat{\rho}}\right)\frac{1}{1-\vartheta\hat{\lambda}}\frac{\hat{\rho}+\vartheta-\hat{\rho}\vartheta(\mathbb{L}+\hat{\lambda})}{(1-\vartheta\mathbb{L})(1-\hat{\lambda}\mathbb{L})}\xi_t \\
=& \frac{1}{1-\hat{\rho}}\left(1-\frac{\vartheta}{\hat{\rho}}\right)\left(\frac{\omega_1}{1-\vartheta\mathbb{L}} + \frac{\omega_2}{1-\rho\mathbb{L}} + \frac{\omega_3}{1-\hat{\lambda}\mathbb{L}}\right)\varepsilon_{t+1},
\end{aligned}
$$

where

$$
\omega_1 = \frac{\hat{\lambda}\vartheta(\hat{\rho}-\vartheta)(1-\hat{\rho}\vartheta)}{\hat{\rho}(\vartheta-\hat{\lambda})(1-\hat{\lambda}\vartheta)(\rho-\vartheta)} \quad \omega_2 = \frac{(\rho-\hat{\rho})(\hat{\lambda}\vartheta(1-\rho\hat{\rho})+\hat{\rho}(\rho-\vartheta))}{\hat{\rho}(\rho-\hat{\lambda})(1-\hat{\lambda}\vartheta)(\rho-\vartheta)}, \quad \omega_3 = 1-\omega_1-\omega_2.
$$

We use $\{\zeta_k\}_{k=0}^{\infty}$ to denote the IRF. The following properties hold:

1. When $\rho > \hat{\rho}$, $\zeta_k > 0$ for all $k \geq 0$.

   Note that if $\rho > \hat{\rho}$, it is also the case that $\rho > \hat{\rho} > \vartheta > \hat{\lambda}$. As a result, $\omega_1 > 0$ and $\omega_2 > 0$. Also note that $\zeta_k = \omega_1\vartheta^k + \omega_2\rho^k + \omega_3\hat{\lambda}^k$. It follows that

$$
\zeta_k > (\omega_1+\omega_2)\vartheta^k + \omega_3\hat{\lambda}^k = (\omega_1+\omega_2)(\vartheta^k - \hat{\lambda}^k) + \hat{\lambda}^k > 0.
$$

2. When $\rho < \hat{\rho}$, $\zeta_k < 0$ for k large enough.

   When k large enough, that the sign of ζ_k will be the same as the sign of ω_1 if $\vartheta > \rho$, and it will be the same as the sign of ω_2 if $\vartheta < \rho$. If $\vartheta > \rho$, $\omega_1 < 0$. If $\vartheta < \rho$, $\omega_2 < 0$. Therefore, the forecast error is negative in the long run.

3. When $\rho < \hat{\rho}$, there exists a threshold $\overline{\lambda}$ such that only if $\hat{\lambda} > \overline{\lambda}$, $\zeta_1 > 0$. That is, the forecast error does not immediately switches to negative only if learning is slow enough.

   A straightforward calculation yields

$$\zeta_1 = \omega_1 \vartheta + \omega_2 \rho + \omega_3 \hat{\lambda} = \frac{D(\hat{\lambda})}{\hat{\rho}(1-\hat{\lambda}\vartheta)}, \quad \text{where} \quad D(\hat{\lambda}) = \frac{(-\hat{\rho}\vartheta)\hat{\lambda}^2 + (\hat{\rho}^2\vartheta - \hat{\rho}\vartheta^2 - \rho\hat{\rho}\vartheta + \hat{\rho} + \vartheta)\hat{\lambda} - \hat{\rho}^2 + \rho\hat{\rho}}{\hat{\rho}(1-\hat{\lambda}\vartheta)}.$$

   The sign of ζ_1 is the same as the numerator $D(\hat{\lambda})$. Since $D(0) = \hat{\rho}(\rho - \hat{\rho}) < 0$, and $D(\hat{\rho}) = \hat{\rho}(\rho + \vartheta)(1 - \hat{\rho}\vartheta) > 0$, there exists $\overline{\lambda} \in (0, \hat{\rho})$ such that $D(\hat{\lambda}) > 0$ only if $\hat{\lambda} > \overline{\lambda}$.

Regression Coefficients. We now study the theoretical counterparts of K_{CG}, K_{BGMS} and K_{KW}.

Case 1: $\tau = \hat{\tau}$ and $\rho = \hat{\rho}$. We have already proved that the IRF of the forecast error is always positive. Because the IRF of the outcome is always positive, K_{KW} has to positive. By individual rationality, K_{BGMS} has to be zero. What remains is to prove that K_{CG} is positive.

As the outcome follows an AR(2) process, the individual forecast error and forecast revision are given by

$$y_{t+1} - \mathbb{E}_{i,t}[y_{t+1}] = \frac{1}{1-\rho}\left(1 - \frac{\vartheta}{\rho}\right)\left(g_1^{\varepsilon}(\mathbb{L})\varepsilon_{t+1} + g_1^{u}(\mathbb{L})u_{i,t}\right),$$
$$\mathbb{E}_{i,t}[y_{t+1}] - \mathbb{E}_{i,t-1}[y_{t+1}] = \frac{1}{1-\rho}\left(1 - \frac{\vartheta}{\rho}\right)\left(g_2^{\varepsilon}(\mathbb{L})\varepsilon_{t} + g_2^{u}(\mathbb{L})u_{i,t}\right),$$

where

$$g_1^{\varepsilon}(\mathbb{L}) = \frac{1}{(1-\vartheta\mathbb{L})(1-\rho\mathbb{L})} - \left(1 - \frac{\hat{\lambda}}{\rho}\right)\frac{1}{1-\vartheta\hat{\lambda}}\frac{\rho + \vartheta - \rho\vartheta(\mathbb{L} + \hat{\lambda})}{(1-\vartheta\mathbb{L})(1-\hat{\lambda}\mathbb{L})(1-\rho\mathbb{L})}\mathbb{L},$$
$$g_2^{\varepsilon}(\mathbb{L}) = \left(1 - \frac{\hat{\lambda}}{\rho}\right)\frac{1}{1-\vartheta\hat{\lambda}}\left(\frac{\rho + \vartheta - \rho\vartheta(\mathbb{L} + \hat{\lambda})}{(1-\vartheta\mathbb{L})(1-\hat{\lambda}\mathbb{L})(1-\rho\mathbb{L})}(1 - (\vartheta + \rho)\mathbb{L}) + \frac{\rho\vartheta(1 - \rho\vartheta\hat{\lambda}\mathbb{L})}{(1-\vartheta\mathbb{L})(1-\hat{\lambda}\mathbb{L})(1-\rho\mathbb{L})}\mathbb{L}\right),$$
$$g_1^{u}(\mathbb{L}) = -\left(1 - \frac{\hat{\lambda}}{\rho}\right)\frac{1}{1-\vartheta\hat{\lambda}}\frac{\rho + \vartheta - \rho\vartheta(\mathbb{L} + \hat{\lambda})}{(1-\vartheta\mathbb{L})(1-\hat{\lambda}\mathbb{L})}\tau^{-1},$$
$$g_2^{u}(\mathbb{L}) = \left(1 - \frac{\hat{\lambda}}{\rho}\right)\frac{1}{1-\vartheta\hat{\lambda}}\left(\frac{\rho + \vartheta - \rho\vartheta(\mathbb{L} + \hat{\lambda})}{(1-\vartheta\mathbb{L})(1-\hat{\lambda}\mathbb{L})}(1 - (\vartheta + \rho)\mathbb{L}) + \frac{\rho\vartheta(1 - \rho\vartheta\hat{\lambda}\mathbb{L})}{(1-\vartheta\mathbb{L})(1-\hat{\lambda}\mathbb{L})}\mathbb{L}\right)\tau^{-1}.$$

The covariance between individual forecast error and individual forecast revision is

$$\text{Cov}\left(\text{Error}_{i,t}, \text{Revision}_{i,t}\right) = \left(\frac{1}{1-\rho}\left(1 - \frac{\vartheta}{\rho}\right)\right)^2\left(\text{Cov}\left(g_1^{\varepsilon}(\mathbb{L})\varepsilon_{t+1}, g_2^{\varepsilon}(\mathbb{L})\varepsilon_t\right) + \text{Cov}\left(g_1^{u}(\mathbb{L})u_{i,t}, g_2^{u}(\mathbb{L})u_{i,t}\right)\right),$$

and a long but straightforward calculation yields the following expression:

$$\text{Cov}(g_1^{u}(\mathbb{L})u_{i,t}, g_2^{u}(\mathbb{L})u_{i,t}) = -\tau^{-1}\left(\left(1 - \frac{\hat{\lambda}}{\rho}\right)\frac{1}{1-\vartheta\hat{\lambda}}\right)^2\frac{1-\hat{\lambda}\rho}{(1-\hat{\lambda}\vartheta)(1-\hat{\lambda}^2)}\Delta,$$

where

$$\Delta \equiv \left(\vartheta^3\hat{\lambda}\left(1-\hat{\lambda}^2\right) - 3\vartheta\hat{\lambda}\left(1-\vartheta\hat{\lambda}\right) + \left(1-\vartheta^2\right)\right)\rho^2 - \left(\vartheta^3\left(1-\hat{\lambda}^2\right) + \vartheta\left(3\vartheta\hat{\lambda} - 2\right)\right)\rho + \vartheta^2.$$

With $\tau = \hat{\tau}$ and $\rho = \hat{\rho}$, agents are rational and $K_{\text{BGMS}} = 0$. That is, $\text{Cov}\left(\text{Error}_{i,t}, \text{Revision}_{i,t}\right) = 0$. Let us assume momentarily that $\Delta > 0$. It follows that

$$\begin{aligned}\text{Cov}(\text{Error}_t, \text{Revision}_t) &= \left(\frac{1}{1-\rho}\left(1 - \frac{\vartheta}{\rho}\right)\right)^2\text{Cov}\left(g_1^{\varepsilon}(\mathbb{L})\varepsilon_{t+1}, g_2^{\varepsilon}(\mathbb{L})\varepsilon_t\right) \\ &= -\left(\frac{1}{1-\rho}\left(1 - \frac{\vartheta}{\rho}\right)\right)^2\text{Cov}\left(g_1^{u}(\mathbb{L})u_{i,t}, g_2^{u}(\mathbb{L})u_{i,t}\right) > 0,\end{aligned}$$

which implies that $K_{\text{CG}} > 0$.

The argument is completed by the lemma below, which verifies that $\Delta > 0$ by mapping ρ to x, ϑ to y, and λ to z.

Lemma. When $x, y, z \in (0,1)$, the following inequality holds

$$\left(y^3 z\left(1-z^2\right)-3yz\left(1-yz\right)+\left(1-y^2\right)\right)x^2-\left(y^3\left(1-z^2\right)+y\left(3yz-2\right)\right)x+y^2>0$$

Proof. Recast the left hand side of the above inequality as a quadratic in x:

$$C(x)\equiv\left(y^3 z\left(1-z^2\right)-3yz\left(1-yz\right)+\left(1-y^2\right)\right)x^2-\left(y^3\left(1-z^2\right)+y\left(3yz-2\right)\right)x+y^2.$$

This has two real roots, $x = x_1$ and $x = x_2$, given by

$$x_1=-\frac{y}{1-yz}\qquad\text{and}\qquad x_2=-\frac{y}{y^2z^2-2yz-y^2+1}.$$

Clearly, given the assumption that $y, z \in (0,1)$, x_1 is negative and $C(0) = y^2 > 0$. If x_2 is negative, then $C(x)$ is positive when $x \in (0,1)$. If x_2 is positive, to guarantee that $C(x)$ is positive when $x \in (0,1)$, we need to show that $x_2 > 1$, which is equivalent to show that

$$y^2z^2-2yz+\left(y-y^2+1\right)>0$$

Define the following quadratic equation in z:

$$D(z)=y^2z^2-2yz+\left(y-y^2+1\right).$$

Its discriminant is $-4y^3\left(1-y\right)$, which is negative given that $y \in (0,1)$. Therefore, $D(z)$ is always positive, which in turn verifies $x_2 > 1$. □

Case 2: $\tau = \hat{\tau} = \infty$ and $\rho \neq \hat{\rho}$. If $\hat{\tau} = \tau = \infty$, then $\hat{\lambda} = \vartheta = 0$. In this case, all agents receive the same signal, and there is no distinction between $\mathbb{E}_{i,t}[\cdot]$ and $\overline{\mathbb{E}}_t[\cdot]$. It follows that $K_{\text{CG}} = K_{\text{BGMS}}$.

To derive the K_{BGMS}, note that

$$\begin{aligned}
y_{t+1}-\mathbb{E}_{i,t}[y_{t+1}]&=\frac{1}{1-\hat{\rho}}(\varepsilon_{t+1}+(\rho-\hat{\rho})y_t)\\
\mathbb{E}_{i,t}[y_{t+1}]-\mathbb{E}_{i,t-1}[y_{t+1}]&=\frac{1}{1-\hat{\rho}}\hat{\rho}(y_t-\hat{\rho}y_{t-1})
\end{aligned}$$

It follows that

$$K_{\text{BGMS}}=\frac{\hat{\rho}(1-\rho\hat{\rho})(\rho-\hat{\rho})}{(\hat{\rho}^2+\hat{\rho}^4-2\rho\hat{\rho}^3)}$$

Clearly, the sign of K_{BGMS} is the same as the sign of $\rho - \hat{\rho}$.

The sequence of the forecast error IRF $\{\zeta_k\}_{k=1}^{\infty}$ is given by

$$\zeta_k=\frac{1}{1-\hat{\rho}}\rho^{k-1}(\rho-\hat{\rho}),$$

which are either all positive or all negative. Since the IRF of the outcome is always positive, the sign of K_{KW} is the same as $\rho - \hat{\rho}$.

Case 3: $0 < \hat{\tau} < \infty$ and $\hat{\rho} > \rho > 0$. With $\hat{\tau} = \infty$, the signs of $K_{\text{CG}} > 0$ and $K_{\text{KW}} < 0$ are always the same as $\rho - \hat{\rho}$. Therefore, $0 < \hat{\tau} < \infty$ is necessary to allow $K_{\text{CG}} > 0$ and $K_{\text{KW}} < 0$.

With $\hat{\rho} \leq \rho$, the average forecast error is always positive, the IRF of the forecast error is always positive. Together with the fact that the IRF of the outcome is always positive, we have $K_{\text{KW}} > 0$. Therefore, $\hat{\rho} > \rho$ is necessary to allow $K_{\text{KW}} < 0$.

Proofs of Proposition 6 and Corollary 4

We first consider the case with "higher-order doubts". The recursive formulation of individual consumer i's consumption choice is

$$c_{i,t} = \mathbb{E}_t[\xi_t] + \beta\mathbb{E}_t[c_{i,t+1}] + (1-\beta)\mathbb{E}_t[c_{t+1}]$$

As ξ_t is perfectly observed by consumer i, we guess the policy function is

$$c_{i,t} = a\xi_t,$$

for some constant a.

Under the assumption that agent i believes that other agents observe the fundamental shock with probability q, it follows that

$$\mathbb{E}_{i,t}[c_{i,t+1}] = \mathbb{E}_{i,t}[a\xi_{t+1}] = a\rho\xi_t \qquad \mathbb{E}_{i,t}[\overline{\mathbb{E}}_t[\xi_t]] = q\xi_t, \qquad \mathbb{E}_{i,t}[c_{t+1}] = \mathbb{E}_{i,t}[\overline{\mathbb{E}}_t[a\xi_{t+1}]] = aq\rho\xi_t.$$

Substituting these expectations into consumers' optimal response leads to

$$a\xi_t = \xi_t + \beta a\rho\xi_t + (1-\beta)aq\rho\xi_t,$$

which further verifies our guess by setting the constant a as

$$a = \frac{1}{1-(\beta\rho+(1-\beta)q\rho)} < \frac{1}{1-\rho}.$$

In the economy without higher-order doubts but with mis-perceived $\hat{\rho}$, the aggregate outcome is

$$c_t = \frac{1}{1-\hat{\rho}}\xi_t.$$

The outcomes in the two economies are observationally equivalent iff

$$\frac{1}{1-\hat{\rho}} = \frac{1}{1-(\beta\rho+(1-\beta)q\rho)} \quad \rightarrow \quad \hat{\rho} = \rho-(1-\beta)\rho(1-q) < \rho$$

In terms of forecasts, in the economy with higher-order doubts,

$$\mathbb{E}_{i,t}[c_{t+1}] = \overline{\mathbb{E}}_t[c_{t+1}] = q\mathbb{E}^*_t[c_{t+1}].$$

where $\mathbb{E}^*_t[\cdot]$ is the perfect-information rational expectation operator.

Next, we consider the level-k thinking. The agents are assumed to observe the fundamental and to have the correct prior about its process but a mis-specified prior about the behavior of others: they are "level-k thinkers" for some finite integer $k \geq 0$. Level 0 agents are assumed to play $c_t = c^0_t \equiv 0$, for all t and for all ξ^t. Level 1 agents believe that other agents are level 0. They therefore play $c_t = c^1_t$, where c^1_t is given by the solution to

$$c^1_t = \xi_t + \beta\mathbb{E}_t[c^1_{t+1}]$$

Level 2 agents believe that other agents are level 1. They therefore choose $c_t = c^2_t$, where c^2_t is given by the solution to

$$c^2_t = \xi_t + \beta\mathbb{E}_t[c^2_{t+1}] + (1-\beta)\mathbb{E}_t[c^1_{t+1}].$$

Similarly, the aggregate outcome for level-k agent when $k > 0$ satisfies

$$c^k_t = \xi_t + \beta\mathbb{E}_t[c^k_{t+1}] + (1-\beta)\mathbb{E}_t[c^{k-1}_{t+1}].$$

We proceed by a guess-and-verify approach. Suppose that $c_t^k = a_k \xi_t$. Then for $k > 0$, a_k has the following recursive structure

$$a_k = 1 + \beta\rho a_k + (1-\beta)\rho a_{k-1}.$$

Using the fact $g_0 = 0$, we have for $k > 0$,

$$a_k = \frac{1}{1-\rho}\left(1 - \left(\frac{(1-\beta)\rho}{1-\beta\rho}\right)^k\right),$$

which has proved the conjecture.

Compared with the economy with mis-perceived $\hat{\rho}$, the aggregate outcomes are equivalent iff

$$\frac{1}{1-\hat{\rho}} = \frac{1}{1-\rho}\left(1 - \left(\frac{(1-\beta)\rho}{1-\beta\rho}\right)^k\right).$$

Since $\left(1 - \left(\frac{(1-\beta)\rho}{1-\beta\rho}\right)^k\right) < 1$, we have $\hat{\rho} < \rho$.

In terms of the forecast, in the level-k economy,

$$\mathbb{E}_{i,t}[c_{t+1}] = \overline{\mathbb{E}}_t[c_{t+1}] = a_{k-1}\rho\xi_t = \frac{a_{k-1}}{a_k}\mathbb{E}_t^*[c_{t+1}],$$

where $\frac{a_{k-1}}{a_k} < 1$.

Lastly, consider the cognitive discounting economy. We still proceed by a guess-and-verify approach. Suppose that the actual law of motion of c_t is

$$c_t = Rc_{t-1} + D\varepsilon_t,$$

and the perceived law of motion is

$$c_t = mRc_{t-1} + D\varepsilon_t.$$

Meanwhile, the perceived law of motion of ξ_t is

$$\xi_t = m\rho\xi_{t-} + \varepsilon_t.$$

Recall that the aggregate outcome is given by

$$c_t = \sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t[\xi_{t+k}] + (1-\beta)\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t[c_{t+k+1}].$$

Using the mis-specified priors, we have

$$c_t = \frac{1}{1-\beta m\rho}\xi_t + (1-\beta)\frac{mR}{1-\beta mR}c_t,$$

which leads to the actual law of motion of c_t as

$$c_t = \rho c_{t-1} + \frac{1-\beta mR}{1-mR}\frac{1}{1-\beta m\rho}\varepsilon_t.$$

To be consistent with our guess, we have

$$R = \rho, \quad D = \frac{1}{1-m\rho}.$$

Compared with the economy with mis-perceived $\hat{\rho}$, the aggregate outcomes are equivalent iff

$$\frac{1}{1-\hat{\rho}} = \frac{1}{1-m\rho}, \quad \rightarrow \quad \hat{\rho} = m\rho < \rho.$$

In terms of the forecast, in the cognitive-discounting economy,

$$\mathbb{E}_{i,t}[c_{t+1}] = \overline{\mathbb{E}}_t[c_{t+1}] = m\mathbb{E}_t^*[c_{t+1}].$$

In all the three economies (higher-order doubts, level-k, cognitive discounting), the individual forecast is the same as the average forecast about the aggregate outcome , and it follows that $K_{\text{CG}} = K_{\text{BGMS}}$. In addition, in all the three economies,

$$c_t = \varphi\xi_t, \quad \text{and} \quad \overline{\mathbb{E}}_t[c_{t+1}] = \zeta\mathbb{E}_t^*[c_{t+1}] = \zeta\rho\varphi\xi_t,$$

for some constant φ and $\zeta \in (0,1)$. Therefore, we have

$$\text{Cov}(\text{Error}_t, \text{Revision}_t) = \text{Cov}(\varphi\xi_{t+1} - \zeta\rho\varphi\xi_t, \zeta\rho\varphi\xi_t - \zeta^2\rho^2\varphi\xi_{t-1}) = \varphi^2\rho^2\zeta(1-\zeta)\frac{1-\zeta\rho^2}{1-\rho^2},$$

which implies $K_{\text{CG}} = K_{\text{BGMS}} > 0$.

In addition, the law of motion of the forecast error is

$$\text{Error}_t = \varphi\frac{1-\zeta\rho\mathbb{L}}{1-\rho\mathbb{L}}\varepsilon_{t+1} = \varphi\left((1-\zeta)\frac{1}{1-\rho L} + \zeta\right)\varepsilon_{t+1},$$

and the corresponding IRF is always positive given $\zeta \in (0,1)$.

Given that in all these economies the IRF of the outcomes are always positive and that the IRF of the forecast error is always positive, we know that K_{KW} in all of these economies have to be positive as well.

Proof of Lemma in Appendix C

We consider the case with $k = 1$. Note that average revision, Revision_t, and the idiosyncratic component of individual revision, $(\text{Revision}_{i,t} - \text{Revision}_t)$, are independent of each other. Therefore, the regression coefficient on the average forecast revision remains to be K_{CG}.

The covariance between individual forecast error and idiosyncratic revision component is

$$\begin{aligned}
\text{Cov}(\text{Error}_{i,t}, \text{Revision}_{i,t} - \text{Revision}_{i,t}) &= \text{Cov}\left(-\frac{\hat{\rho}-\hat{\lambda}}{1-\hat{\lambda}\mathbb{L}}u_{i,t}, \frac{\hat{\rho}-\hat{\lambda}}{1-\hat{\lambda}\mathbb{L}}(\tau^{-\frac{1}{2}}u_{i,t} - \hat{\rho}\tau^{-\frac{1}{2}}u_{i,t-1})\right) \\
&= -\tau^{-1}\frac{(\hat{\rho}-\hat{\lambda})^2(1-\hat{\lambda}\hat{\rho})}{1-\hat{\lambda}^2} \\
&= -\kappa_1\mathcal{V}_{\text{agg}}\tau^{-1}.
\end{aligned}$$

Denote the regression coefficient on $(\text{Revision}_{i,t} - \text{Revision}_t)$ as β. It follows that

$$\begin{aligned}
\beta &= \frac{\text{Cov}(\text{Error}_{i,t}, \text{Revision}_{i,t} - \text{Revision}_t)}{\mathcal{V}_{\text{idio}}} = \frac{\text{Cov}(\text{Error}_{i,t}, \text{Revision}_{i,t}) - \text{Cov}(\text{Error}_t, \text{Revision}_t)}{\mathcal{V}_{\text{idio}}} \\
&= \frac{\mathcal{V}_{\text{ind}}}{\mathcal{V}_{\text{idio}}}K_{\text{BGMS}} - \frac{\mathcal{V}_{\text{agg}}}{\mathcal{V}_{\text{idio}}}K_{\text{CG}},
\end{aligned}$$

and hence

$$K_{\text{noise}} = -\beta\frac{\mathcal{V}_{\text{idio}}}{\mathcal{V}_{\text{agg}}} = K_{\text{CG}} - \frac{\mathcal{V}_{\text{ind}}}{\mathcal{V}_{\text{agg}}}K_{\text{BGMS}}$$

Using the definitions of K_{CG} and K_{BGMS}, we then also have $K_{\text{noise}} = \kappa_1\tau^{-1}$. Because $\kappa_1$are independent of τ, K_{noise} is decreasing in τ, and vanishes when $\tau \to \infty$.