INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

515

2018

Models, Inattention and Bayesian Updates

Javier Turén, Raffaella Giacomini y Vasiliki Skreta

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

Models, Inattention and Bayesian Updates

Raffaella Giacomini * Vasiliki Skreta † Javier Turén ‡

June 8, 2018

Abstract

We formulate a theory of expectations updating that fits the dynamics of accuracy and disagreement in a new survey dataset where agents can update at any time while observing each other's expectations. Agents use heterogeneous models and can be inattentive but, when updating, they follow Bayes' rule and assign homogeneous weights to public information. Our empirical findings suggest that agents do not herd and, despite disagreement, they place high faith in their models, whereas during a crisis they lose this faith and undergo a paradigm shift. Bayesian updating fits the data well, but only in non-crisis years. Furthermore, we empirically evaluate this theory's relative strengths and weaknesses in both crisis- and non-crisis years vis-a-vis several leading alternatives and find that it fits better on average and in non-crisis years.

KEYWORDS: *Bayesian updating, Information rigidities, Heterogeneous agents, Expectation formation, Disagreement, Forecast accuracy, Herding.*

JEL CLASSIFICATION: E27, E37, D80, D83.

*University College London. r.giacomini@ucl.ac.uk

†University of Texas at Austin & University College London. vskreta@gmail.com

‡Pontificia Universidad Católica de Chile. jturen@uc.cl

1 Introduction

I see a critical need for basic research on expectations formation. Understanding how persons update their expectations with receipt of new information often is a prerequisite for credible use of econometric decision models to predict behavior. Charles Manski, Econometrica 2004.

"Expectations matter. [....] Yet how those expectations are formed, and how best to model this process, remains an open question." Olivier Coibion and Yuriy Gorodnichenko, AER 2015.

How do agents update their expectations about an economic variable? Are their updates rational, in the sense of being consistent with Bayes' rule? How heterogeneous are they? Do they copy each other ("herding")? During a crisis, do they stick to their models or are they willing to discard them and allow for a paradigm shift? We answer these questions by analyzing a previously unstudied dataset–*Bloomberg's ECFC survey of professional forecasters*–and by building a theory of expectation updating. Our goal is to explain the dynamic evolution of both forecast accuracy and disagreement in the data *simultaneously*, which can be challenging. While the theory we develop is simple and tightly parameterized, it fits the data remarkably well. The key ingredients of the theory are that agents use heterogeneous models to form an initial forecast and update infrequently. Importantly, when updating they use Bayes' rule to incorporate a public signal. This theory differs from the leading theories of expectations formation with information frictions, for example sticky information (Mankiw and Reis (2002)); noisy information (Sims (2003), Woodford (2003)); and heterogeneous priors (Patton and Timmermann (2010)). However, it is close enough to them to allow for an illuminating comparative evaluation of their relative strengths and weaknesses in fitting the data.

Because we focus on expectations updating, we can empirically investigate whether forecasters are Bayesians, and we find support for that. A key parameter of the theory is agents' "faith"[1] in their model. The estimates suggest that in non-crisis years, agents have high faith in their models but their faith sharply decreases during the 2008-2009 crisis. Although we do not explicitly model agents' behaviour during the crisis, the results support the hypothesis that agents discarded their initial model and did not follow Bayes' rule. The Bloomberg's ECFC survey is unique in that not only can participants update their forecasts at any point but also, upon logging onto the terminal, they can observe in real time the consensus and each others' forecasts. We leverage this feature to investigate whether forecasters "herd" and do not find support for it.

[1]Formally, the precision of the initial forecast (inverse of the variance).

The data consists of a panel of approximately 75 professionals forecasting US annual (year-on-year) CPI inflation for the years 2007 to 2014 during the 18 months before the release of the figure.

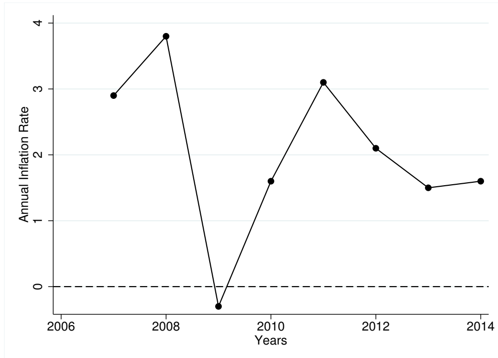


Figure 1: Annual US inflation 2007-2014

The period 2007-2014 is rather turbulent, as Figure 1 suggests, and this has implications for the patterns we are interested in. Figure 2 shows that our theory simultaneuosly fits the dynamics of disagreement (standard deviation of forecasts across agents) and accuracy (Root Mean Squared Error, henceforth RMSE) in the data, except for the crisis years 2008-2009.

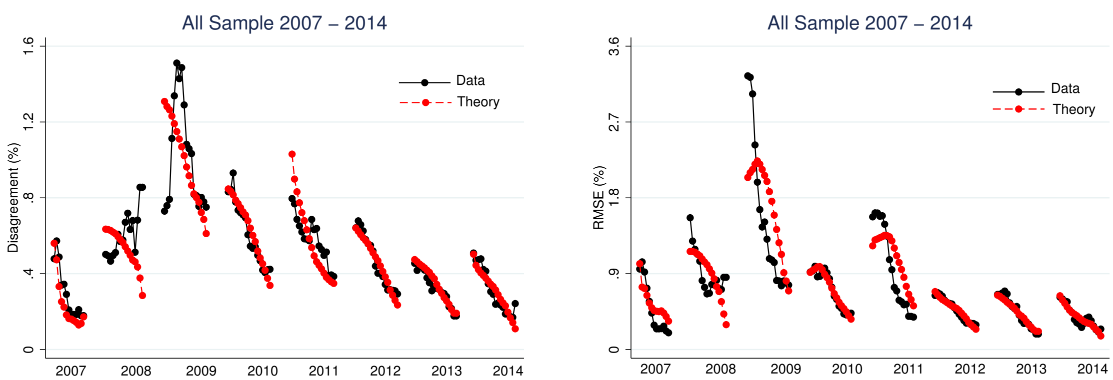


Figure 2: Fit of theory for disagreement (left) and RMSE (right)

The theory is motivated by the key empirical regularities in the data. First, agents are inattentive: On average only about 40%-50% of agents update at least once a month and this frequency varies across agents and over the updating period, but is similar across years, including during the crisis. Updaters are more accurate than non-updaters and by performing separate rationality tests for the two groups we find that inattention can explain why population consensus forecasts usually violate rationality tests (Pesaran and Weale (2006)).

The second ingredient of the baseline theory is motivated by the finding that inattention is not enough to account for the disagreement in the data, since even agents who update at the same time disagree (a prediction at odds with the sticky information theory). The fact that disagreement persists even towards the end of the calendar year, when most of the forecasted variable has been observed, suggests that there are deep idiosyncratic reasons why agents disagree: They could rely on different statistical models and/or have different industry- or individual-specific incentives, career concerns, pessimism/optimism biases, differences in private information, or different experience and skills. To capture these dimensions of heterogeneity parsimoniously, we allow agents to use heterogeneous models to produce an *initial forecast*. It is important to stress here that what we call "model" is a reduced-form way to capture both the fact that agents may use statistical models to produce a forecast as well as apply judgmental or incentive-driven corrections. In other words, individual forecasts may be biased for a number of different reasons that we don't explicitly model but as long as these biases are constant over the updating period they can be viewed as a "fixed effect" which does not affect the updates.

The main novelty of our theory is to combine the two aforementioned features with Bayesian updating and the possibility of herding: As new information arrives, agents who update use Bayes' rule to incorporate a *common* public signal. In reality, agents observe a variety of public but also private signals: official statistics, news articles, central bank announcements and so forth. Among these sources of information two stand out: The release of monthly CPI and the forecasts of other agents. Monthly CPI is clearly something that agents should pay attention to because it is a component of annual inflation, which is the variable they are forecasting. Similarly, other agents' forecasts as well as the consensus forecast are relevant, since they may not only include public, but also private information.[2] In the baseline theory we assume that information is related to monthly CPI.[3] We also consider the case where information is the consensus forecast (agents "herd").

By estimating a number of variations of the baseline theory we provide a quantitative assessment of the relative importance of its key ingredients. Furthermore, by estimating several competing theories we formally assess their performance in fitting the dynamic patters of accuracy and disagreement. Our main findings are as follows.

First, heterogeneity in models is crucial, as assuming homogeneous models would underestimate the observed disagreement and accuracy over the entire forecast horizon. Second, another important feature is inattention, which is confirmed by the fact that all theories that assume full attention underestimate disagreement. These two conclusions are in line with previous findings

[2]Indeed, consensus forecasts from various surveys have been documented to be remarkably accurate and are often used in policy, and more generally, in economic decision-making (Ang, Bekaert and Wei (2007)).

[3]More precisely, agents use an autoregressive model of order one (AR(1)) with normally distributed errors for year-on-year monthly inflation to produce an initial forecast for annual inflation and, when updating, they use the same model to translate the monthly CPI signal into a signal about annual inflation. The AR(1) model has agent-specific intercepts, which captures heterogeneous beliefs across agents about the long-run mean of inflation and/or any other agent-specific bias or forecast correction that remains constant over the updating period. The normality assumption implies a linear updating rule.

in the literature, and strengthen the case for including inattention and heterogeneous models in any realistic theory of expectation formation.

Our third finding is that Bayesian updating, and the implied weights agents assign to the common public signal fit remarkably well the dynamic updating patterns in non-crisis years. We find that the already good fit of the baseline theory in those years is not improved when substituting Bayesian weights with the data-driven weights of Patton and Timmermann (2010). The dynamic evolution of the estimated Bayesian and the data-driven weights is similar in non-crisis years: In both cases the weight on the public signal increases as the horizon shrinks. During the crisis years, however, the data-driven weights imply that agents assign the largest weight on information not at the end, but mid-way through the updating period. This provides empirical evidence in support of Ortoleva (2012)'s theory where decision-makers employ Bayes' rule in normal times but behave differently when they face "rare" events. Fourth, we show that the version of the model where agents "herd" worsens the fit, as it underestimates accuracy and disagreement over the entire updating period. Fifth, models with noisy information imply a shape with a "kink" at the beginning of the current year for the evolution of disagreement and accuracy which is not in the data. Finally, during the crisis, the baseline theory cannot match the sharp increase increase in both accuracy and disagreement mid-way through the updating period.

This paper contributes to the large and influential literature that empirically tests theories of informational rigidities (Mankiw, Reis and Wolfers (2004), Coibion and Gorodnichenko (2012), Coibion and Gorodnichenko (2015), Andrade and Le Bihan (2013)). Our main contribution to this literature is twofold. First, we provide a simple theory that fits the dynamics of both accuracy and disagreement in the data while remaining within the confines of a rational theory with Bayesian updating. Second, our estimation results highlight the key theoretical ingredients that are necessary for fitting the data and provide a quantitative assessment of the strengths and weaknesses of a number of leading theories.

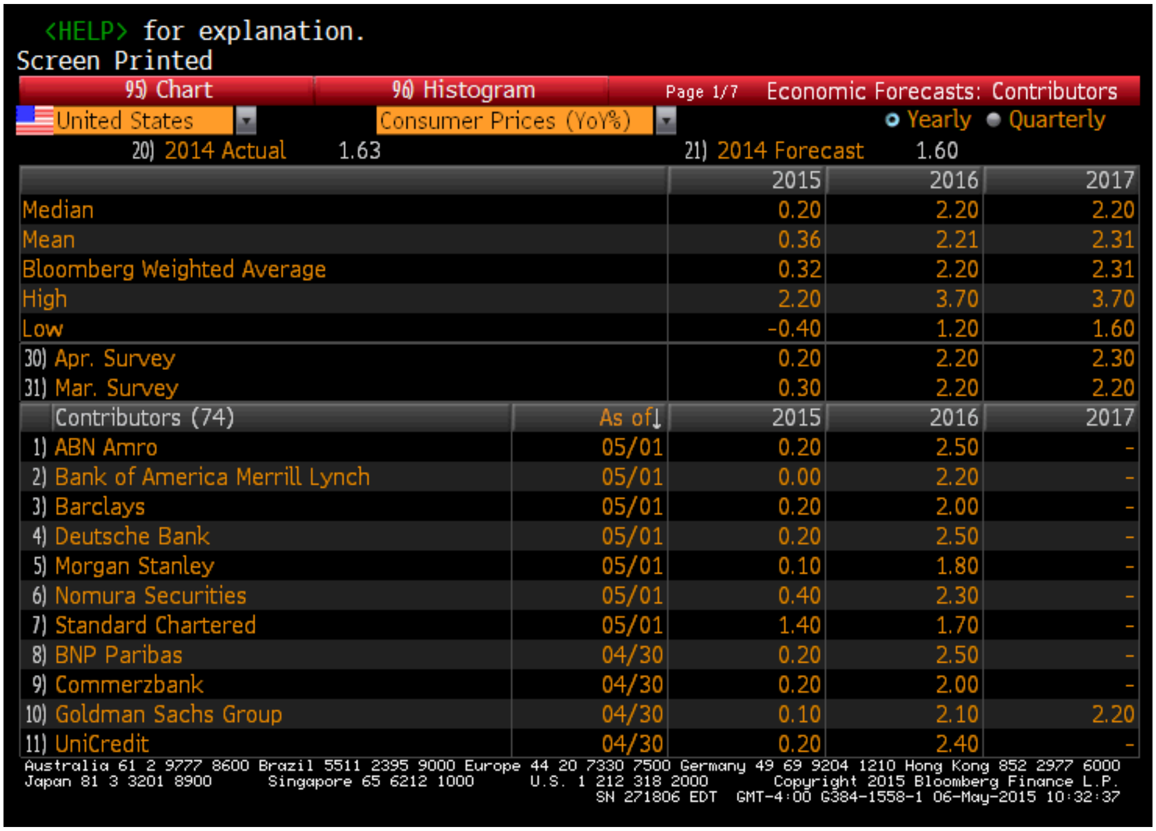
Our findings suggest that agents use Bayes rule and behave in a homogeneous fashion. In contrast, Manzan (2011) and Lahiri and Sheng (2008) respectively, show that professional forecasters violate Bayesian updating and that heterogeneous weights on information are needed to match the disagreement in the data. We can reconcile these findings by noting that in our baseline theory only attentive agents are Bayesian. Some evidence of Bayesian updating is also reported in Coibion, Gorodnichenko and Kumar (2015) who study a survey of firm's expectations and in an early empirical contribution by Caskey (1985).

We find that agents trust their models and do not copy each other, despite the fact that in our survey data it is easy to do so. This result contributes to the literatures investigating whether agents herd in professional settings: The empirical evidence on professional forecasters is mixed. Ehrbeck and Waldmann (1996) rules it out, while Laster, Bennett and Geoum (1999) reports evidence of herding among participants from banks and industrial corporations and Lamont (2002) among younger and less experienced forecasters.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 presents stylized facts. Section 4 describes the baseline theory. The results of structural estimation are in Section 5. A number of variations of the baseline theory (homogeneous models, full attention, noisy information, non-Bayesian learning) are described and estimated in Section 6. Section 7 discusses the herding variation and presents the results of its structural estimation. In Section 8 we consider a number of leading theories of expectation formation and present results of their estimation. Finally, Section 9 concludes.

2 The dataset: Bloomberg's ECFC survey

This paper analyzes updates of US annual (year-on-year) CPI inflation forecasts from a new dataset: the "Economic Forecasts ECFC" survey of professional forecasters conducted by Bloomberg. Users of the Bloomberg terminal can access it at any point in time (see Figure 3). Each forecast on the screen is associated with the name of the forecaster's institution as well as with the date of the last forecast update. For example, the screen shot on the 6th of May 2015 shows that Barclays last updated their forecast on May 1st. The screen also displays in the first row the consensus forecast updated in real time. To the best of our knowledge we are the first to analyze this survey.



<HELP> for explanation.
Screen Printed
95) Chart 96) Histogram Page 1/7 Economic Forecasts: Contributors
United States Consumer Prices (YoY%) Yearly Quarterly
20) 2014 Actual 1.63 21) 2014 Forecast 1.60

		2015	2016	2017
Median		0.20	2.20	2.20
Mean		0.36	2.21	2.31
Bloomberg Weighted Average		0.32	2.20	2.31
High		2.20	3.70	3.70
Low		-0.40	1.20	1.60
30) Apr. Survey		0.20	2.20	2.30
31) Mar. Survey		0.30	2.20	2.20
Contributors (74)	As of	2015	2016	2017
1) ABN Amro	05/01	0.20	2.50	-
2) Bank of America Merrill Lynch	05/01	0.00	2.20	-
3) Barclays	05/01	0.20	2.00	-
4) Deutsche Bank	05/01	0.20	2.50	-
5) Morgan Stanley	05/01	0.10	1.80	-
6) Nomura Securities	05/01	0.40	2.30	-
7) Standard Chartered	05/01	1.40	1.70	-
8) BNP Paribas	04/30	0.20	2.50	-
9) Commerzbank	04/30	0.20	2.00	-
10) Goldman Sachs Group	04/30	0.10	2.10	2.20
11) UniCredit	04/30	0.20	2.40	-

Australia 61 2 9777 8600 Brazil 5511 2395 9000 Europe 44 20 7330 7500 Germany 49 69 9204 1210 Hong Kong 852 2977 6000
Japan 81 3 3201 8900 Singapore 65 6212 1000 U.S. 1 212 318 2000 Copyright 2015 Bloomberg Finance L.P.
SN 271806 EDT GMT-4:00 G384-1558-1 06-May-2015 10:32:37

Figure 3: Bloomberg EFCF survey: a snapshot

Participants: A comparative overview of the ECFC survey with respect to other surveys of professional forecasters considered in the literature is provided in Table 1. The table shows that although the composition of the participants is comparable to other surveys the number of participants is larger, updates can happen anytime, and the most recent forecasts of other agents are visible to the participants at any point in time. Participants are based in different countries and can be divided into three main categories: financial institutions (private banks, investment institutions and large global investment banks), economic consulting firms, and a mixed group formed by others such as universities, research centers and governmental agencies. On average, the percentage of participants from financial institutions is higher for ECFC and Blue Chip surveys, reaching 60%. The percentage of economic consulting firms is relatively in line with other surveys at 26%, while the number of universities, research centers and other types of organizations is lower (14%) compared to the US SPF, Consensus Forecasts, Livingston and the ECB's SPF.

Forecasts: We focus on fixed-event forecasts of annual US inflation. For each year, we consider the subset of agents who provide an initial forecast 18 months before the end of

Table 1: Common Surveys of Professional Forecasters

	US SPF	Blue Chip	Consensus	Livingston	ECB SPF	HM Treasury	ECFC
Frequency	Q	M	M	Semi-annual	Q	M	Anytime
Participants	45	50	45	45	60	40	75
Anonymity	✓	×	×	✓	✓	×	×
Observability of consensus	×	×	×	×	×	×	✓
Financial Inst.	37%	64%	49%	50%	49%	52%	60%
Economic Cons.	29%	26%	24%	21%	22%	32%	26%
Univs. & Gov.	34%	10%	27%	29%	29%	16%	14%

the year.[4] For instance, for 2014 annual inflation, our sample contains the agents who start forecasting in July of 2013. The only exceptions are the initial years of the survey, 2007 and 2008, for which the initial forecasts were provided 13 months and 16 months before the end of the year, respectively. Our dataset contains the full history of forecast updates for all agents during the 18 months before the end of each year. Updates in the ECFC are irregularly spaced and more frequent than in other surveys, however they are not frequent enough to allow us to conduct the analysis at a daily or weekly frequency since there are few updates in any given week. We therefore analyze updates at the monthly frequency and consider the forecasts available on the terminal the last day of each month. This is motivated by the fact that our baseline theory assumes that forecast updates take into account the release of monthly CPI–which typically occurs around the 20th of the month–so sampling the forecasts at the end of the month ensures that this piece of information is in the agents' information set. For each agent we, therefore, have a sequence of 18 forecasts for each year. We index the horizon backwards, so for a given year the index $h = 18, ..., 1$ indicates that the forecast was produced h months before the end of the corresponding year.

Reward structure and incentives: The survey manager of Bloomberg explained in private communications that participants are not explicitly rewarded nor ranked based on their accuracy. Active participants are, however, more likely to be cited in survey-related Bloomberg news stories and newsletters. The fact that forecasts are not anonymous could provide a stronger incentive for accuracy compared to other surveys (Table 1).

Survey availability and uses: The survey is available on the Bloomberg terminals. In addition, its main results are divulged and published by Bloomberg in various newsletters, reports and media outlets such as the monthly *Bloomberg Briefs Economics Newsletter* that are available to a broad array of users. Bloomberg collects data on the number of hits ECFC gets and on the number of users that subscribe to ECFC-related publications, but does not make these data publicly available.

[4] We also remove outliers, defined as agents who provided fewer than 8 forecasts for either inflation, GDP, unemployment or the Federal Funds Rate during the forecasting period.

3 Stylized Facts

Fact 1: Agents are inattentive: Figure 4 plots the proportion of agents that update (changed their forecast) at each horizon and year in our sample. On average, 40% to 50% of participants update at least once a month, which is in line with the figures in Dovern (2013) for the monthly Consensus Economics survey and slightly below the 60% to 70% documented by Andrade and Le Bihan (2013) for the survey of the ECB, which is quarterly. Notice that the proportion of updaters varies within a year but does not change dramatically between crisis (2008-2009) and other years.

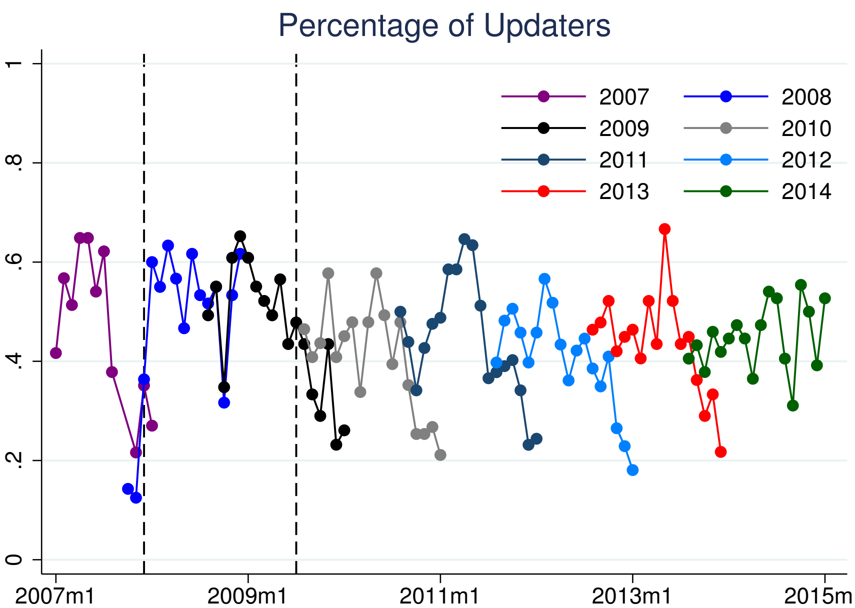


Figure 4: Monthly updating frequencies

The frequency with which agents update their forecast provides an empirical proxy for attention. Our historical dataset does not contain the time-stamps of when each forecast was submitted, so we can tell that a forecaster logged in the system only when she decided to *change* her forecast. Because the absence of an update could be also due to the decision not to revise the forecast after acquiring new information, our empirical measure can be viewed as a lower bound for attention. In order to measure attention we construct an indicator variable capturing whether agent i updates her previous forecast h months before the end of the year:

$$r_{i,h} = \begin{cases} 1 & \text{if } \widehat{y}_{i,h} \neq \widehat{y}_{i,h+1} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where $\widehat{y}_{i,h}$ is agent i's forecast of annual inflation for a given year measured on the last day of month h and $\widehat{y}_{i,h+1}$ is her forecast measured on the last day of the previous month $(h+1)$. We compute this indicator for each year 2007, ..., 2014 and for $h = 17, \ldots, 1$. Then, each month the set of *updaters* consists of the agents for which $r_{i,h} = 1$.

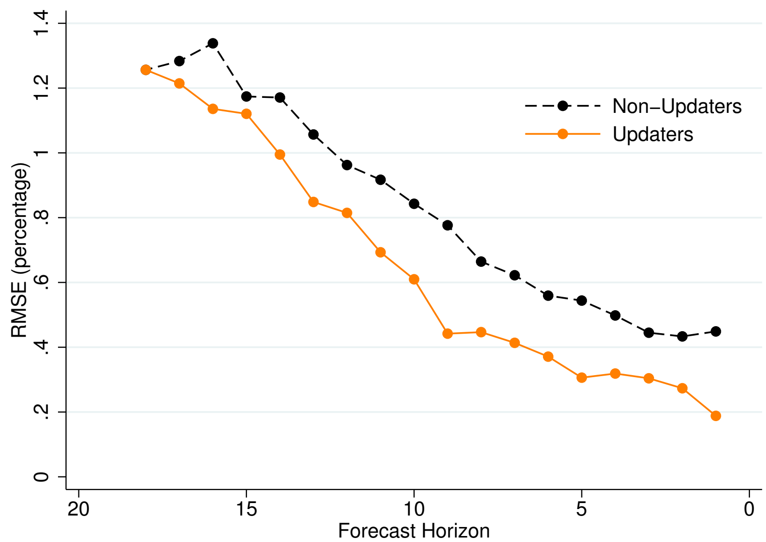


Figure 5: RMSE: average across years

Fact 2: Attention improves accuracy: Figure 5 presents the evolution of the Root Mean Square Error as the horizon shrinks from $h = 18$ to $h = 0$, averaged across the years 2007-2014. This measure is constructed using the consensus forecast separately for updaters and non-updaters at each forecast horizon and comparing it with the actual realization of annual inflation. As expected, as the forecast horizon is reduced, the accuracy of both groups improves since relevant new information accumulates over time. However, updaters are uniformly more accurate than non-updaters.[5]

Fact 3: Inattention can explain violations of Full Information Rational Expectations (FIRE): We now test whether the rationality of the consensus forecast for each horizon differs between updaters and non-updaters. As the time series dimension only includes eight observations for each horizon, the results of the tests should be interpreted with due caution. As standard, the idea is to test for predictability of forecast errors by a "Mincer-Zarnowitz" (MZ) regression:

$$y = \alpha + \beta \bar{y}_h + \epsilon, \tag{2}$$

where y is the realization of annual inflation for a given year and $\bar{y}_h$ is the consensus forecast

[5]Note that since in our sample all agents give a forecast at $h = 18$ the two lines coincide at horizon $h = 18$.

produced h months before the end of the year for each group (updaters, non-updaters). The consensus forecast is unbiased if $\alpha = 0$ and it is rational if $\alpha = 0$ and $\beta = 1$. Table 2 reports the p-values for the test of the aforementioned two hypotheses obtained by estimating regression (2) for each h separately across different years and considering HAC standard errors constructed using the Bartlett Kernel. The first two columns of Table 2 present the results using all participants, while the remaining columns reports the results for updaters and non-updaters.

Table 2: P-values of bias and rationality (MZ) tests for consensus forecast

	All Sample		Updaters		Non-Updaters	
Horizon	Bias	MZ	Bias	MZ	Bias	MZ
1	0.549	0.0000	0.577	0.1737	0.082	0.0000
2	0.611	0.0004	0.200	0.3365	0.102	0.0007
3	0.159	0.0946	0.211	0.0070	0.161	0.3825
4	0.241	0.5316	0.486	0.1614	0.227	0.3541
5	0.195	0.4158	0.295	0.0668	0.214	0.0628
6	0.238	0.4842	0.262	0.3406	0.269	0.2110
7	0.199	0.2420	0.513	0.1341	0.325	0.0053
8	0.602	0.1015	0.139	0.2957	0.708	0.0471
9	0.986	0.1204	0.574	0.6848	0.667	0.0069
10	0.855	0.0170	0.920	0.5145	0.681	0.0002
11	0.606	0.0041	0.701	0.4991	0.511	0.0003
12	0.336	0.1490	0.224	0.4697	0.558	0.1093
13	0.432	0.1214	0.199	0.4374	0.957	0.0779
14	0.888	0.2761	0.683	0.0696	0.498	0.7969
15	0.626	0.8849	0.772	0.7611	0.55	0.6798
16	0.589	0.0614	0.630	0.5078	0.543	0.0215
17	0.394	0.0000	0.435	0.0001	0.367	0.0001
18	0.379	0.0000	0.370	0.0000	0.85	0.0001

The results confirm previously established findings in the literature (Pesaran and Weale (2006)) that the consensus based on all agents (updaters and non-updaters) fails rationality tests at different horizons, at typical significance levels. Attention can partly explain such finding, as the consensus forecast for updaters does not violate rationality for the majority of forecast horizons, whereas the consensus forecast for non-updaters violates rationality for most horizons.

Fact 4: Attention differences cannot fully account for disagreement: Consistent with other surveys, there is large disagreement in the forecasts at all horizons (measured by the standard deviation of individual forecasts) not only for the entire cross-section of participants, but also when considering updaters. Figure 6 shows the evolution of disagreement at each forecast horizon, averaged across the years 2007-2014.

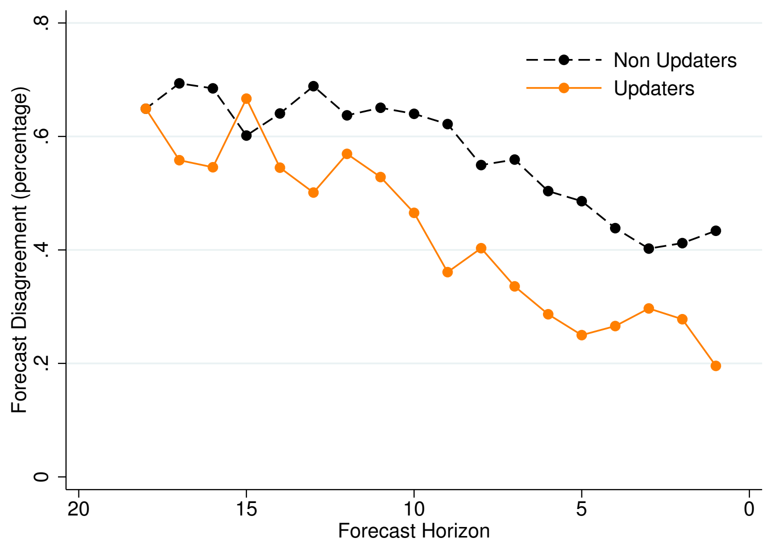


Figure 6: Disagreement: average across years

As expected, the dispersion of predictions decreases as the forecast horizon decreases, which is consistent with the fact that more information about annual inflation accumulates as the release date draws near. Updaters disagree less than non-updaters uniformly at all horizons but their disagreement is never zero. This implies that heterogeneity in attention is not enough to explain the persistent disagreement among agents. There could be many other reasons why attentive agents disagree: they use heterogeneous forecasting models, have different incentives (e.g. industry or individual-specific incentives), display pessimism or optimism biases, or have access to different sources information or experience/skills. In the following section, we build a theory where we endow agents with heterogeneous models. The notion of "model" in our theory is a reduced-form way to capture any of these dimensions of heterogeneity so long as they are time-invariant.

4 A Theory of Expectation Updating

We now build a rational theory of how agents update their expectations of inflation for a given year during the forecasting period. We should note that, since we are describing the procedure for a *specific* year, we allow any parameter to change from year to year, but for ease of notation we don't make this explicit. The theory has three key features: i) heterogeneity in models, ii) inattention and iii) Bayesian updating for attentive agents. We refer to this version of the theory as "**baseline**."

Heterogeneous Models: We assume that there is a population of "models". A model describes any procedure that an agent follows to produce a forecast such as running a statistical package, discussing with colleagues or applying any judgemental corrections to the forecasts produced by his software. Each model gives normally distributed forecasts with mean μ_i and precision $a_{i,18}$ as follows: $N\left(\mu_i, a_{i,18}^{-1}\right)$. The population of models is described by a normal $N\left(\mu, \sigma_\mu^2\right)$, so each μ_i is drawn from this distribution. Heterogeneity is captured by the fact that the mean μ_i is agent-specific. However, we assume that all agents assign the same precision $a_{i,18} = a_{18}$ so, in other words, they trust their methods is the same way. The precision a_{18} is a key parameter in our theory as it captures agents' faith in their model.

Initial Forecast: All agents start forecasting inflation for a given year 18 months before the end of that year. At that point, each agent draws a μ_i out of the population, and her initial forecast is $\widehat{y}_{i,18} = \mu_i$. Note that an implication of our assumptions is that agents can change the model from year to year but they keep it constant during the 18 months forecasting period.

Updates and Inattention: At each subsequent month $h = 17, ..., 1$ a fraction $1 - \lambda_h$ of agents update their forecast. The key assumption is that the fraction of updaters $1 - \lambda_h$ is exogenous and known (we set it equal to the observed frequency of updates depicted in Figure 4). Let $\mathcal{I}_h$ denote the set of updaters at time h.

Information Arrival and Bayesian Updating: Let y denote the true inflation for the year under consideration. Updaters observe a public signal $z_{i,h}$ about y with precision $b_{i,h}$:

$$\begin{aligned} z_{i,h} &= y + \varepsilon_{i,h}, \\ \varepsilon_{i,h} &\sim N(0, b_{i,h}^{-1}), \end{aligned}$$

and use Bayes' rule to update their forecast. Normality and Bayes' rule imply that the updated forecast is a linear combination of the previous forecast and the public signal where the weight on public informations is determined by the relative precision of the signal and the previous forecast. Thus, the forecast for the attentive agent i at month h - i.e., for $i \in \mathcal{I}_h$ - is a normal

$N(\widehat{y}_{i,h}, a_{i,h}^{-1})$ with

$$\widehat{y}_{i,h} = (1 - w_{i,h})\widehat{y}_{i,h+1} + w_{i,h} z_{i,h} \quad (3)$$

$$a_{i,h} = a_{i,h+1} + b_{i,h} = a_{18} + \sum_{j=h}^{17} b_{i,h}. \quad (4)$$

The weight on the public signal is

$$w_{i,h} \equiv \frac{b_{i,h}}{a_{i,h+1} + b_{i,h}} = \frac{b_{i,h}}{a_{18} + \sum_{j=h}^{17} b_{i,h}.}. \quad (5)$$

The forecast for non-attentive agents is the previous month's forecast, $\widehat{y}_{i,h} = \widehat{y}_{i,h+1}$.

Public signal: The release of monthly CPI is a key piece of information as it can be used to construct year-on-year monthly inflation x_h, from which we can obtain a measure of annual inflation. In what follows, we assume that the public signal is monthly CPI.

Agents use the current month's CPI to calculate year-on-year monthly inflation x_h, defined as

$$x_h = \frac{1}{12}\left(\frac{cpi_h - cpi_{h+12}}{cpi_{h+12}}\right), \forall h = 11, \ldots, 0. \quad (6)$$

They, then, employ their model to obtain a signal for annual inflation. Because annual inflation can be approximated by the sum of year-on-year monthly inflation:[6] $y \cong \sum_{h=0}^{11} x_h$, a model for x_h implies a model for y.

We assume that agents use an AR(1) model for year-on-year monthly inflation to translate x_h into a signal for annual inflation:

$$x_h \;\; = \;\; c_i + \phi x_{h+1} + v_h, \quad (7)$$

with

$$v_h \sim N(0, \sigma_v^2).$$

Note that we assume that the model for year-on-year monthly inflation has heterogeneous intercepts across agents. This, in turn, implies heterogeneous unconditional means of annual inflation: $\mu_i \equiv 12\left(\frac{c_i}{1-\phi}\right)$. The annual signal $z_{i,h}$ is the model-implied conditional mean of y

[6] Annual inflation, e.g. for year 2007, is $y = (\overline{cpi}_{2007} - \overline{cpi}_{2006})/\overline{cpi}_{2006}$ where $\overline{cpi}_{2007} = \frac{1}{12}\sum_{j=0}^{11} cpi_j$ and cpi_j is the Consumer Price Index measured j months before the end of year 2007

based on the information set available at month h, $z_{i,h} = E[y|x_h, x_{h+1}, ...]$:

$$z_{i,h} = \mu_i + \frac{\phi^{h-11}(1-\phi^{12})}{1-\phi}(x_h - c_i/(1-\phi)) \text{ for } h = 17, \ldots, 12 \tag{8}$$

$$z_{i,h} = \frac{h}{12}\mu_i + \frac{\phi(1-\phi^{h})}{1-\phi}(x_h - c_i/(1-\phi)) + \sum_{j=h}^{11} x_j \text{ for } h = 11, \ldots, 1. \tag{9}$$

The signal's precision is the inverse of the variance of the error $e_{i,h} = y - z_{i,h}$. Heterogeneity is only in the means (and so in $z_{i,h}$), so the precision b_h is the same across agents and it is a known function of ϕ and σ_v^2:

$$b_h^{-1} = \begin{cases} \frac{\sigma_v^2}{(1-\phi)^2}\left[12 - \frac{2\phi(1-\phi^{12})}{1-\phi} + \frac{\phi^2(1-\phi^{24})}{1-\phi^2}\right] + \frac{\phi^2(1-\phi^{12})^2(1-\phi^{2h-24})}{(1-\phi)^3(1+\phi)}\sigma_v^2 & \text{if } h \geq 12 \\ \frac{\sigma_v^2}{(1-\phi)^2}\left(h - \frac{2\phi(1-\phi^{h})}{1-\phi} + \frac{\phi^2(1-\phi^{2h})}{1-\phi^2}\right) & \text{if } h \leq 11. \end{cases} \tag{10}$$

Full details about the derivations of $z_{i,h}$ and b_h can be found in Appendix A.

For coherence, we assume that agents use the same model to interpret the monthly signal and to generate the initial forecast. We further assume that agent i's initial forecast at $h = 18$, $\widehat{y}_{i,18}$, is the unconditional mean of annual inflation implied by the AR(1) model for year-on-year monthly inflation, that is, $\mu_i \equiv 12\left(\frac{c_i}{1-\phi}\right)$. The latter assumption is motivated by the fact that the unconditional mean is the optimal forecast (for a quadratic loss) for any mean-reverting process at long horizons. In order to obtain a distribution across agents for μ_i: $\mu_i \sim N(\mu, \sigma_\mu^2)$ we assume that $c_i \sim N(c, \sigma_c^2)$, so that $\mu = 12\frac{c}{1-\phi}$ and $\sigma_\mu = 12\frac{\sigma_c}{1-\phi}$.

Homogeneous weight on signal: An implication of our assumptions is that when agents update they assign homogeneous weight on the public signal. To see this, note that the signal's precision in equation (4) only depends on common parameters and the initial precision a_{18} is assumed to be constant across agents, so the weight on public information in (3) is homogeneous across agents:

$$w_h = \frac{b_h}{a_{18} + \sum_{j=h}^{17} b_j}. \tag{11}$$

In conclusion, given that all attentive agents attach the same weight to information, the only sources of heterogeneity in our theory are heterogeneity in attention and heterogeneity in the long-run mean of inflation implied by agents' models.

In Section 5 that follows, we describe the structural estimation of this theory. In Sections 6 and 7 we use the same procedure to estimate a number of different specifications, which allows us to quantify the relative contribution of each feature of the theory to the fit. In Section 8 we compare the fit of this with the ones of alternative theories considered in the literature. These results establish that the baseline specification fits better on average.

5 Structural Estimation of Baseline Theory and Results

For each year separately, we estimate the following parameters $\theta = (\phi, \sigma_v^2, \mu, \sigma_\mu^2, a_{18})$ in order to get the theoretically predicted RMSE and disagreement to be as close as possible to their empirical counterparts.[7].

We compute the disagreement and RMSE implied by the theory as follows: For a given year, we draw the initial forecasts for 75 agents from $N(\mu, \sigma_\mu^2)$. For horizon $h = 17$, we randomly draw a fraction $1 - \lambda_h$ of agents, where $1 - \lambda_h$ is the proportion of updaters observed in the data at horizon h for that year. The forecast for the non-updaters equals the initial forecast. The forecast for the updaters is obtained using equations (3)-(5) where the public signal $z_{i,h}$ is derived using equations (8) and (9), x_h is year-on-year monthly inflation (given by (6)) for that year and horizon, computed using actual data from FRED.[8] The public signal's precision is given by (10). We repeat the same procedure for each of the subsequent horizons $h = 16, ..., 1$, where now the forecast for a non-updater at horizon h is equal to his forecast at horizon $h+1$. At each horizon h, we use the simulated forecasts to compute the disagreement across all agents and the RMSE for the consensus forecast. The RMSE is calculated by comparing the consensus forecast to the realization of annual inflation for the year under consideration. We repeat the simulation for $\tau = 100$ replicas for each year, which overall yields τT different series for disagreement and accuracy, where $T = 8$ years.[9]

We estimate the structural parameters by Simulated Method of Moments (SMM) as in Gourieroux and Monfort (1996), Duffie and Singleton (1993), Ruge-Murcia (2012), which amounts to matching empirical moments with their theory-implied counterparts. We set the weighting matrix equal to the identity matrix. Based on the evidence studying the small sample bias of GMM estimators with a large number of moment conditions discussed in, e.g., Tauchen (1986) and Altonji and Segal (1996), we restrict attention to a subset of horizons

[7]We calibrate the remaining parameter of the model, λ_h, using the proportion of updaters in the data

[8]The data corresponds to the historical series of the Consumer Price Index for all items. Available at: https://research.stlouisfed.org/fred2/series/CPIAUCSL

[9]Although the choice of the number of replications (τ) is arbitrary, following Duffie and Singleton (1993), the idea is to have a simulation sample, τT, generated for a sample size of T years, where $\tau T \to \infty$ as $T \to \infty$. Moreover, as stressed by Gourieroux and Monfort (1996), when the number of replications tends to infinity, the SMM estimator coincides with GMM.

$h = 1, 3, 5, 7, 9, 12, 14, 16, 18$. The chosen moments are the disagreement and RMSE at different horizons h, delivering a total of 18 moments.

We start by describing the estimated parameters and discuss the findings and proceed with a number of figures that illustrate the theory's fit of the data. Table 3 presents all the estimated parameters. The first part of the table contains the estimated parameters when moments are averaged over all years (2007-2014), then over only crisis years (2008-2009) and finally over all years excluding the crisis years, referred to as non-crisis years. We also separately estimate parameters for each year (using the exact number of total agents in each year) and present those in the second part of the table. The AR(1) coefficient ϕ and the precision a_{18} of the initial forecast are almost always significant. The noise attached to the public signal (σ_v) is significant for several years and particularly higher during the crisis years, in line with the results in Nimark (2014). The results also suggest that the initial forecast precision parameter a_{18}, measuring forecaster's faith in their model, is large and significant in all years, but drops dramatically during the years of the crisis. Finally, from the estimated ϕ, c_i and σ_c we calculate the long-run mean (μ, σ_μ) depicted in Figure 8. The last column of Table 3 reports the p-values of the test of over-identifying restrictions (J-Test).

Figure 7 presents the fit of the baseline theory for both disagreement and accuracy. Again, following the sequence described above, we first illustrate the fit over the entire sample, then we focus on crisis years 2008-2009 and, finally, present the results for non-crisis years. We see that the fit is much better for non-crisis years. Disagreement is higher during crises than for normal years. Although the disagreement starts at similar levels in both crisis- and non- crisis years, its behaviour during the forecast horizon is completely different. While for the non-crisis years disagreement always decreases, during the crisis it goes up, reaching a maximum at the 11 months horizon, and then it starts decreasing and stays relatively constant. Likewise, the RMSE is much higher during crisis- than non-crisis years. Figure 2 present the fit of the estimation on a year by year basis. The fit is very close for all non-crisis years.

Overall, the baseline theory fits remarkably well patterns in the data for the non-crisis years, but it underpredicts disagreement and accuracy during the crisis. The parameter estimates of the model can help shed light on why this is the case: our theory assumed that agents "stick" to their initial model, but this is a bad idea during a turmoil since there is often a paradigm shift. The sharp decrease in RMSE in the data half-way through the updating period for the crisis years suggests that agents in our sample discarded their "old" model for a new one and obtained a dramatic improvement in accuracy. Our basic theory does not allow for such mid-year model-shifting. Our empirical estimates give us a measure of agents faith in their model. This is the parameter a_{18}–the precision of the prior model. Our estimates in Table 3 show

Table 3: Estimated Parameters, Baseline Theory

Parameters	ϕ	σ_v	a_{18}	μ	σ_μ	**J Test**
Samples						
All years	0.9982	0.0042	333.31	2.091	0.6758	0.9969
	(0.2016)	(0.005)	(0.00001)	(1.0564)	(0.0921)	
Crisis	0.9983	0.024	8.5486	2.8623	0.7246	1.0000
	(0.4943)	(0.0669)	(0.0001)	(1.7153)	(0.1942)	
Non Crisis	0.9982	0.0036	321.78	1.836	0.6709	0.9991
	(0.1328)	(0.0035)	(0.00001)	(0.1547)	(0.0654)	
Years						
2014	0.9166	0.0014	266.31	1.9987	0.5089	1.0000
	(0.00003)	(0.00002)	(0.00001)	(0.00005)	(0.00006)	
2013	0.9975	0.0061	365.53	1.019	0.4733	1.0000
	(0.0669)	(0.0004)	(0.00001)	(0.0772)	(0.0275)	
2012	0.9529	0.0052	361.27	1.8063	0.6454	1.0000
	(0.0004)	(0.00001)	(0.00001)	(0.00014)	(0.00005)	
2011	0.9416	0.0001	110.38	2.4316	1.031	1.0000
	(0.0326)	(0.091)	(0.00001)	(0.1676)	(0.0296)	
2010	0.9397	0.0074	198.38	1.2981	0.8562	1.0000
	(0.00016)	(0.00006)	(0.00001)	(0.00012)	(0.00002)	
2009	0.9991	0.0318	16.791	1.2151	1.3206	1.0000
	(0.6589)	(0.044)	(0.00001)	(0.4892)	(1.2815)	
2008	0.6058	0.2003	0.0001	2.7841	0.6379	0.9994
	(1.861)	(0.0167)	(16.798)	(0.0902)	(0.0887)	
2007	0.9978	0.0001	199.42	1.951	0.562	0.9949
	(0.1051)	(0.0353)	(0.00001)	(0.1057)	(0.0185)	

Notes: Estimated parameters of the baseline theory based on the Simulated Method of Moments, matching accuracy and disagreement on average over the different sample of years and separately for each year. The J-Test column shows the p-values of the test of over-identifying restrictions.

that agents' faith in their model drops from 321.78 for non-crisis years to 0.0001 in 2008! This provides significant empirical evidence that agents lose faith in their models during turbulent years.

Our results give for each year an estimate of the distribution across agents of the long-run mean of annual inflation implied by the agents' models. We report this distribution in Figure 8 along with the actual realizations of annual inflation.

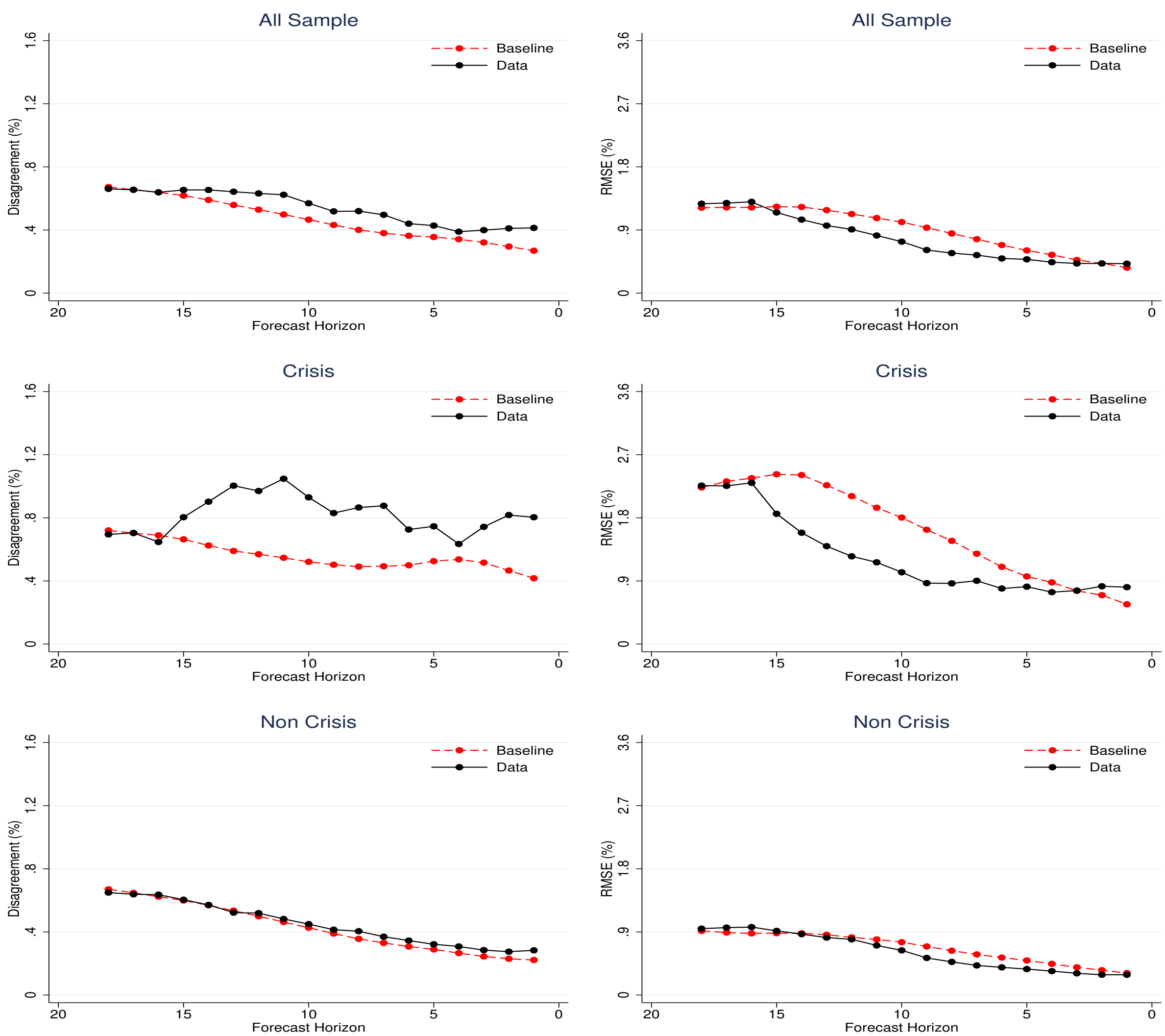


Figure 7: Fit of baseline theory by subsamples. Disagreement and RMSE

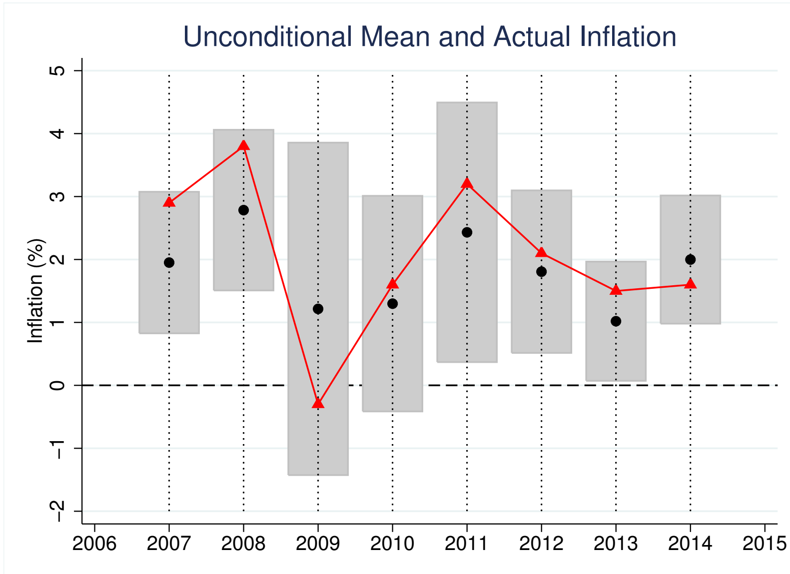


Figure 8: Estimated μ_i distribution and actual inflation

The dots in Figure 8 represent the estimated long run mean of inflation (μ) averaged across agents for each year. The grey boxes are the confidence intervals for the mean (computed as the estimated mean plus/minus two times the estimated standard deviation across agents, σ_μ). The red line represents the realizations of annual inflation for each year. In all years, the estimates of μ are between 1% and 3% and actual inflation is always within the confidence intervals. The higher uncertainty during the crisis is reflected in a wider confidence interval for 2009.

6 Importance of Different Ingredients of our Theory

As described in Section 4 the baseline theory has three main features: i) inattention ii) heterogeneous models and iii) Bayesian updating using monthly CPI as a public signal. In order to evaluate the extent to which each feature is essential in matching the patterns of accuracy and disagreement in the data, we estimate a number of variations of our theory:

Homogeneous Models. All agents use the same model, so that $\sigma_c^2 = 0$ in (7), which implies that $\mu_i = \mu, \ \forall i$.

Full Attention. All agents update their forecasts, i.e. $\lambda_h = 0$ at all horizons h.

Baseline with Noise. Updaters observe a noisy signal of monthly CPI. So, instead of x_h, agents observe a signal $s_{i,h}$, where:

$$s_{i,h} = x_h + u_{i,h} \tag{12}$$

$$u_{i,h} \sim N(0, \sigma_u^2) \quad \forall i, h$$

$$E[u_{i,h}, v_\tau] = 0 \quad \forall i, h, \tau.$$

Since agents do not fully observe the actual realization of x_h, they have to filter out the noise to produce a forecast.

The model for x_h remains as in equation (7) so we can write the model in state-space form as: $\widetilde{x}_h = \phi \widetilde{x}_{h+1} + v_h$ and $s_{i,h} = \widetilde{x}_h + u_{i,h}$, with $\widetilde{x}_h = x_h - c_i/(1-\phi)$. Agents compute the optimal estimate of the unobserved monthly CPI at each horizon h using the Kalman filter: Letting $\widehat{x}_{i,h|h} \equiv E[\widetilde{x}_h | s_{i,h}, s_{i,h+1}, ...]$ and $P_{i,h|h} \equiv Var[\widetilde{x}_h | s_{i,h}, s_{i,h+1}, ...]$, we have:

$$\begin{aligned} \widehat{x}_{i,h|h} &= \widehat{x}_{i,h|h+1} + P_{i,h|h+1}(P_{i,h|h+1} + \sigma_u^2)^{-1}(s_{i,h} - \widehat{x}_{i,h|h+1}) \\ P_{i,h|h} &= P_{i,h|h+1} - P_{i,h|h+1}(P_{i,h|h+1} + \sigma_u^2)^{-1} P_{i,h|h+1}, \end{aligned}$$

where $\widehat{x}_{i,h|h+1} = \phi \widehat{x}_{i,h+1|h+1}$ and $P_{i,h|h+1} = \phi^2 P_{i,h+1|h+1} + \sigma_v^2$. In this version of the theory there are therefore two sources of heterogeneity: agents observe different signals and they interpret them according to different models. Relative to the baseline theory, the public signal (previously described by (8) and (9)) is now $\widehat{z}_{i,h}$:

$$\widehat{z}_{i,h} = \mu_i + \frac{\phi^{h-11}(1-\phi^{12})}{1-\phi} \widehat{x}_{i,h|h} \quad \text{for } h = 17, \ldots, 12 \tag{13}$$

$$\widehat{z}_{i,h} = \frac{h+1}{12}\mu_i + \frac{\phi(1-\phi^h)}{1-\phi} \widehat{x}_{i,h|h} + \widehat{x}_{i,h|h} + \sum_{j=h+1}^{11} x_j \quad \text{for } h = 11, \ldots, 1. \tag{14}$$

For numerical stability we initialize the Kalman filter 150 months before the 18 months forecast horizon and set the initial forecast and variance for agent i to $\widehat{x}_{i,167|168} = \phi(x_{168} - c_i/(1-\phi))$ and $P_{i,167|168} = \sigma_v^2/(1-\phi^2)$, so that there is no noise at the initial month and the initial variance is the unconditional variance of $\widetilde{x}_h$. This implies that the filter's accuracy is constant across agents, $P_{i,h|h} = P_{h|h}$.

The signal's precision is as before given by the inverse of the forecast error $\epsilon_{i,h} = y - \widehat{z}_{i,h}$:

$$E[\epsilon_{i,h}^2] = \frac{\phi^{2h-22}(1-\phi^{12})^2}{(1-\phi)^2} P_{h|h} + \sigma_v^2 \sum_{j=0}^{11} \frac{(1-\phi^{j+1})^2}{(1-\phi)^2} + \sigma_v^2 \sum_{j=12}^{h-1} \frac{\phi^{2j-22}(1-\phi^{12})^2}{(1-\phi)^2}, h \geq 12 \quad (15)$$

$$= \left(\frac{1-\phi^{h+1}}{1-\phi}\right)^2 P_{h|h} + \sigma_v^2 \sum_{j=0}^{h-1} \frac{1-\phi^{j+1}}{1-\phi}, h \leq 11. \quad (16)$$

Now the signal's precision $\widehat{b}_h^{-1} = E[\epsilon_{i,h}^2]$ is not only a function of the unforecastable part of y, but it also depends on the accuracy of the filter $P_{h|h}$.[10] The fact that the precision is constant across agents implies that all agents attach the same weight to the public signal. Specifically, b_h is now replaced by $\widehat{b}_h$ and the weights are as in equation (11).

Baseline with Data-Driven Weights. In this variation of the baseline theory we replace the Bayesian weights with the data-driven weights used in Patton and Timmermann (2010):

$$w_h = 1 - \frac{E[e_{i,h}^2]}{\kappa^2 + E[e_{i,h}^2]},$$

where $e_{i,h} \equiv y - z_{i,h}$ and $E[.]$ is the average across agents. The parameter κ corresponds to a parameter that governs the degree of shrinkage. Note that, since y represents the actual inflation rate, the weights are an *ex-post* measure of the precision of the public signal $z_{i,h}$. Therefore, by construction, more accurate public information implies a higher weight attached to the public signal.

Structural Estimation: To estimate these alternative specifications we follow a procedure analogous to the one used to estimate the baseline, with a number of small modifications. We seek to estimate the following parameters $\theta = (\lambda_h, \phi, \sigma_v^2, \mu, \sigma_\mu^2, a_{18})$.

When simulating the *"homogeneous models"* version of the theory we assume that there is a common $\mu_i \equiv \mu$ for all agents, so all agents have the same first forecast and there is no disagreement at $h = 18$. For the *"full attention"* version of the theory everything is as in the baseline case apart from the fact that there is no inattention so $\lambda_h = 0$. When simulating the *"baseline with noise"*, we follow the same procedure as in the baseline theory but now the updaters rely on the public signals (13) and (14) while the weights depend on the overall signal's precision given by (15) and (16). This version of the theory has an additional parameter σ_u^2. For the *"baseline with data-driven weights"*, the weights are computed using the actual realizations of annual inflation and κ is an additional parameter to estimate. In Appendix B we

[10]The second part of the expression for $E[\epsilon_{i,h}^2]$ is the same as in the baseline model.

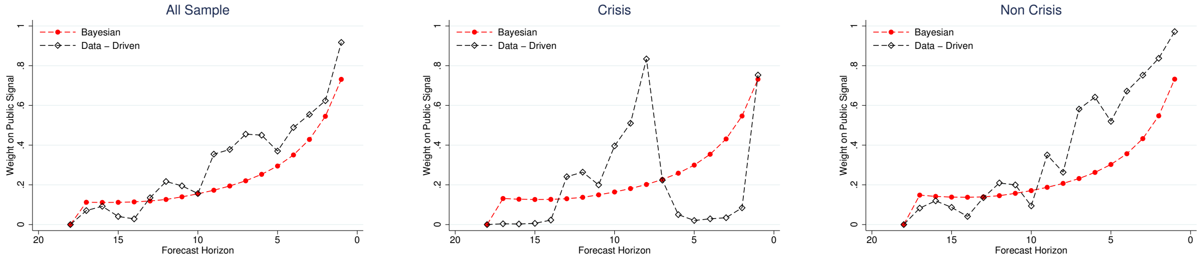


Figure 9: **Estimated Weights:** Baseline (Bayesian) and Data-Driven Weights

present all results graphically by plotting the theoretical and empirical RMSE and disagreement.

Table 4: Relative Fit (%), Variations of Baseline Theory

	All years		Crisis		Non Crisis	
Models	Disag	Accuracy	Disag	Accuracy	Disag	Accuracy
Baseline	0	0	0	0	0	0
Homogeneous Models	156.2	4.1	71.8	-11.8	346.6	31.6
Full Attention	37.8	15.1	26.7	20.3	89	4.5
Baseline with Noise	0.9	117.5	-23.4	80.9	51.9	40.2
Baseline with Data-Driven Weights	-39.2	-44.8	-19.2	-49.6	12.04	-43.8

Notes: The table reports the relative fit (%) of the alternative theories and the baseline theory. The fit of theory j is measured by the cumulative absolute deviation between theoretical and empirical moments: $CAE_j = \sum_{h=1}^{H} |y_h(\theta) - y_h|$ where $y_h(\theta)$ is the theoretical disagreement or accuracy at horizon h and y_h its empirical counterpart. The relative fit is the percentage change of CAE_j with respect to the baseline theory CAE_b: $(CAE_j - CAE_b)/CAE_b$. Hence, a positive (negative) entry indicates an inferior (superior) fit of theory j with respect to the baseline theory.

In addition to the figures showing the model-based and empirical evolution of disagreement and accuracy, we further provide a measure of the overall fit of each theory based on the cumulative absolute deviation of the model-based moments relative to the empirical moments. The relative performance is given by the percentage change of this measure under the alternative theories, with respect to our baseline model. Table 4 compares the fit of a number of variations of the baseline theory, with positive numbers denoting the percentage deterioration in fit. These results highlight the role of each of the key ingredients of the theory in fitting the data. Heterogeneity in models is a crucial aspect of our theory, as eliminating this source of heterogeneity would worsen the fit over all years by 156% for disagreement and by 4% for accuracy. In contrast, eliminating inattention worsens the fit by 37% for disagreement and by 15% for accuracy. Adding noise to the Baseline model worsens the fit by 0.9% for disagreement and 117% for accuracy. As expected, data-driven weights always improve the fit of accuracy, however Bayes' rule outperforms these data-driven weights in matching disagreement in non-crisis years! In Figure 9, we further report the estimated data-driven weights, and we can see that the Bayesian weights are similar to the data-driven weights for the non-crisis years.

7 Are Agents Herding?

An important feature of the ECFC survey is that the most recently updated forecasts of other agents, as well as the consensus forecast (the mean forecast), are visible to users of the Bloomberg terminal, and thus are known to our agents when they update. We leverage this feature to build a variant of the baseline theory that investigates whether forecast updates in our dataset are more plausibly driven by herding than by the arrival of information in the form of monthly CPI release as in the baseline case. The forecasts of others, as well as the consensus, are relevant for a forecaster, since they may include private as well as public information. Indeed, consensus forecasts from various surveys have been documented to be remarkably accurate and are often used in policy, and, more generally, in economic decision-making.[11]

We model "herding" by assuming that the common signal z_h is the current month's consensus forecast. Note, that doing so satisfies our assumption that z_h is an unbiased forecast of annual inflation because, as shown in Table 2, the consensus forecast in the ECFC survey satisfies this property. Since there are several days in a month when revisions could potentially take place, it is not a priori clear when to measure the consensus forecast. However, one interesting pattern that emerges when measuring the fraction of updaters at the weekly frequency (see Figure 10)[12] is that almost all updates in a month happen in a given week. Why is that? This may seem puzzling at first, but after talking to the survey managers at Bloomberg, we found out that Bloomberg sends a monthly email reminder to the survey participants asking them to update their forecast. The reminder is usually sent sometime during the first ten days of the month, but the exact date changes from month to month. This reminder provides a focal date for each month after which most updates take place, and it motivates us to assume that almost all updaters in a month have observed the consensus forecast on the day of the email reminder.

In the "herding" variation of our theory, we thus assume that the public signal $z_{i,h} = z_h$ is the current month's consensus forecast measured on the day of Bloomberg's email reminder. As an empirical estimate of the signal's precision we use the inverse of the variance of forecasts across agents measured on the same day, which is also observable to the agents and is common

[11]For instance, Ang et al. (2007) document that forecasts from surveys of expectations are more accurate than predictions based on macro variables or asset prices.

[12]Figure 4 depicts the fraction of updaters at the monthly frequency

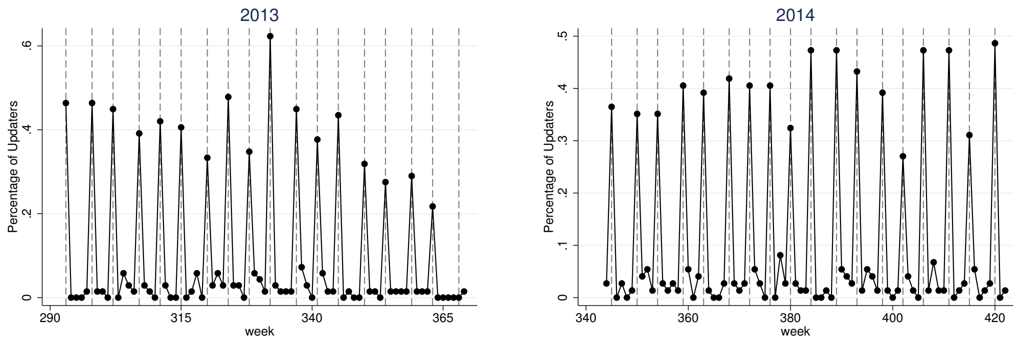


Figure 10: Weekly updating frequencies for years 2013 (left) and 2014 (right)

across agents $b_{i,h} = b_h$. Since the initial precision of the forecasts a_{18}, is assumed to be constant across agents, in the "herding" theory the weights on public information in (3) are homogeneous across agents, as in the baseline version (recall equation (11)).

The parameters for the "herding" theory are $\theta = (\lambda_h, \mu, \sigma_\mu^2, a_{18})$. We compute the disagreement and RMSE implied by the theory as in the baseline case. The forecast for the updaters is obtained using equations (3)-(5) where $z_{i,h} = z_h$ is now the consensus forecast measured on the day of the Bloomberg email reminder in month h and $b_{i,h}$ is the inverse of the variance of forecasts across agents measured on the same day. The forecast for the non-updaters equals the previous month's forecast. We present the fit of the herding theory in Figure 11. Table 4 further reports a summary measure of the relative fit of the "herding" theory compared to the baseline theory (among other variations). Figure 11 and Table 4 show that the fit of the "herding" theory is worse than the fit of the baseline theory for both accuracy and disagreement and for crisis and non-crisis years. According to the same metric as in Table 4 the fit of the "Herding" theory is worse than the baseline by 72% for disagreement and 52% for accuracy, so monthly CPI fits the data better.

8 Comparison with Leading Alternative Theories

Theories of expectation updating with information rigidities have been very influential in macro research in recent years. It is then important to evaluate their relative strengths and weaknesses in fitting the data. The purpose of this section is to compare the baseline theory to competing theories. We first briefly summarize the alternatives we consider and in Section 8.2 we report the estimation results. Appendix C contains the figures showing how each of these theories fits the dynamic patterns of accuracy and disagreement in the data.

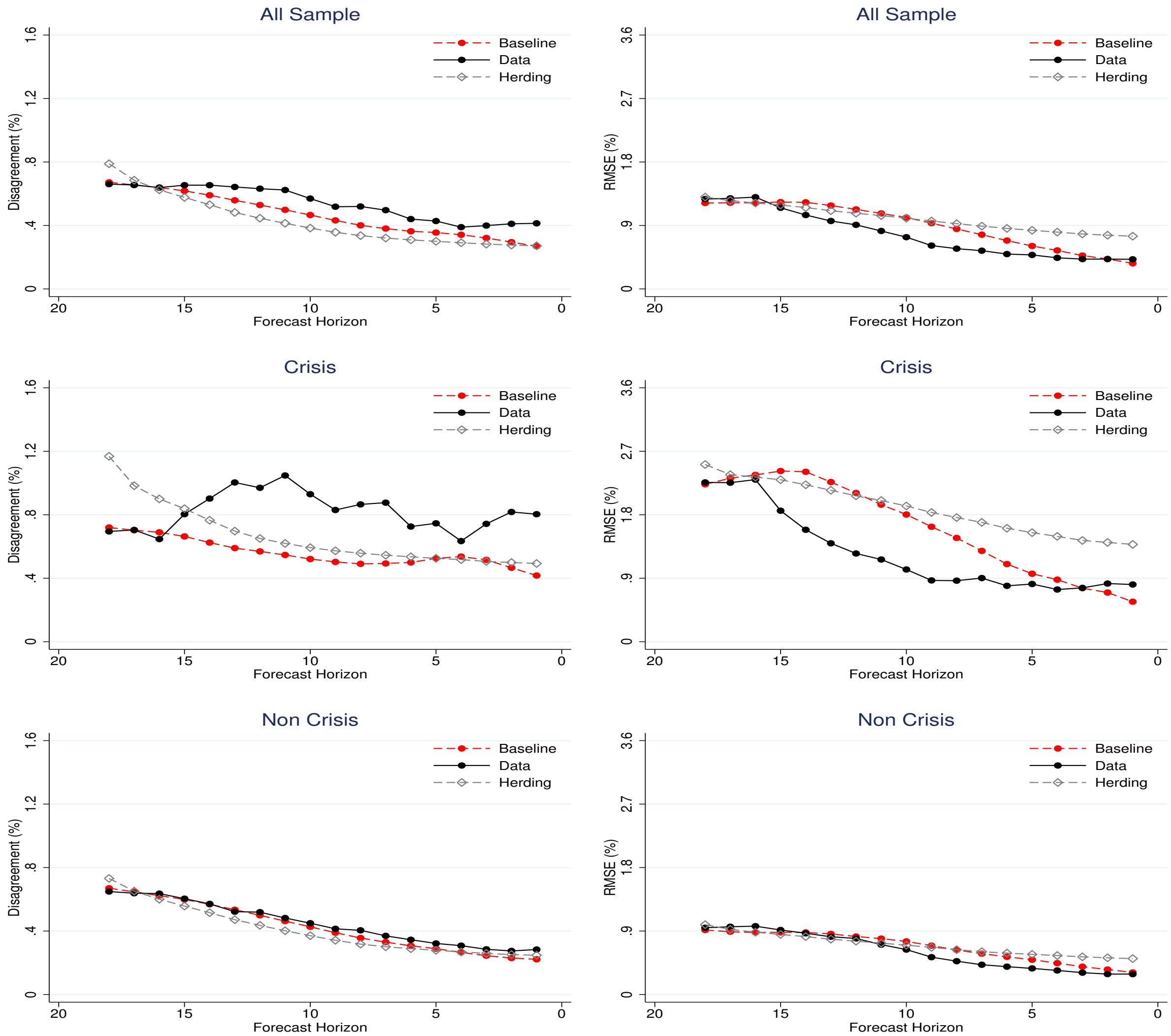


Figure 11: Fit of herding theory. Disagreement and RMSE

8.1 Leading Theories of Expectation Updating

Sticky Information. Originally proposed by Mankiw and Reis (2002), this theory involves inattentive agents who update their information set at infrequent times. In contrast with our theory, all agents have the same model and updaters assign weight 1 to the public signal. In other words, forecast updates are:

$$\widehat{y}_{i,h} = \begin{cases} z_h & \text{if } i \in \mathcal{I}_h \\ \widehat{y}_{i,h+1} & \text{otherwise,} \end{cases} \tag{17}$$

where $\mathcal{I}_h$ is the set of updaters at time h and z_h the public signal about annual inflation as in equations (8) and (9). There is no heterogeneity in models, so $c_i = c$.

Noisy Information. This theory is based on Woodford (2003) and Coibion and Gorodnichenko (2012), and it assumes that all agents are constantly tracking new information, but they observe a noisy signal as in equation (12) and use the Kalman filter to extract the information. Agents have homogeneous models so again $c_i = c$. Although the theory assumes homogeneous models, the noisy signal induces disagreement (recall (12)). Then, at each horizon $h = 18 \ldots, 1$, agent's i forecast is equal to the filtered public signal given by (13) and (14).[13]

Sticky-Noisy Theory. Based on Andrade and Le Bihan (2013), in this theory, a fraction $1 - \lambda_h$ of agents observe a *noisy* public signal. Those who update their forecasts do so in the same way described in the "Noisy Information" theory.

Patton and Timmermann's Theory. The theory in Patton and Timmermann (2010) allows for heterogeneous models, but assumes that all agents are attentive and there is noise in the public signal. This theory assumes that agent i's forecast at horizon h is given by a weighted average between the optimal forecast extracted from the Kalman Filter and the heterogenous long run means, μ_i, as follows:

$$
\begin{aligned}
\widehat{y}_{i,h} &= w_h \mu_i + (1 - w_h) E[y|s_{i,h}] && (18)\\
w_h &= \frac{E[e_{ih}^2]}{\kappa^2 + E[e_{ih}^2]} && (19)\\
e_{ih} &\equiv y - E[y|s_{i,h}]. && (20)
\end{aligned}
$$

In contrast to our theory, in Patton and Timmermann (2010) the weights ($w_h = \frac{E[e_{ih}^2]}{\kappa^2 + E[e_{ih}^2]}$) are not Bayesian and are given by an ad-hoc specification chosen to match the data (by choosing the κ). The assumed function for the weights has the property that as the *filtered* signal $E[y|s_{i,h}]$ becomes more accurate, the weight attached to the signal increases.

Table 5 summarizes the key features of each of the aforementioned theories.

8.2 Alternative Theories: Fit Comparisons

Here we compare the fit of the alternative theories. To do so, we estimate them using the procedure described in Section 5.

Appendix C shows all plots of the theoretical and empirical evolution of RMSE and disagreement. Table 6 provides summary measures of the overall fit. The figures and table suggest that the baseline theory has the best overall fit, in particular for non-crisis years. For these years, all competing theories imply that the evolution of disagreement and accuracy is tent-shaped with

[13]Since the Noisy Information model does not assume a specific initial forecast (as in the baseline model), the expression for $\widehat{z}_{i,h}$ holds even at $h = 18$.

Table 5: Alternative Theories

Models	Het. Models	Inattention	Bayes Rule	Signal: Monthly CPI	Signal: Consensus	Noisy Info
Baseline	✓	✓	✓	✓		
Herding	✓	✓	✓		✓	
Homogeneous Models		✓	✓	✓		
Full attention	✓		✓	✓		
Baseline with Noise	✓	✓	✓	✓		✓
Sticky Info.(MR 2002)		✓		✓		
Noisy Info. (CG 2012)				✓		✓
Sticky - Noisy (AL 2013)		✓		✓		✓
PT 2010	✓			✓		✓

Table 6: Relative Fit (%), Alternative Theories

	Crisis		Non Crisis	
	Disagreement	Accuracy	Disagreement	Accuracy
Sticky Info	64.8	15.6	581.6	162.3
Noisy Info	5	6.2	117.4	12.7
Sticky Noisy	-21.6	-14.8	61.8	35.1
PT 2010	-38.7	-40.6	114.5	30.9

The table reports the relative fit (%) of the alternative theories and the baseline theory. The fit of theory j is measured by the cumulative absolute deviation between theoretical and empirical moments: $CAE_j = \sum_{h=1}^{H} |y_h(\theta) - y_h|$ where $y_h(\theta)$ is the theoretical disagreement or accuracy at horizon h and y_h its empirical counterpart. The relative fit is the percentage change of CAE_j with respect to the baseline theory CAE_b: $(CAE_j - CAE_b)/CAE_b$. Hence, a positive (negative) entry indicates an inferior (superior) fit of theory j with respect to the baseline theory.

a peak around horizon 11, which does not match the decreasing shape observed in the data. The sticky-information clearly underpredicts disagreement and accuracy at long horizons.

For the *crisis* years, the fit of our theory remains better except for the sticky-noisy theory and Patton and Timmermann (2010)'s theory. The superior fit of the sticky-noisy theory is mostly due to its ability to match the increase in disagreement mid-way through the updating period observed in the crisis years, whereas the baseline theory is not able to replicate this jump in disagreement, and, in fact, it implies decreasing disagreement. The superior fit of Patton and Timmermann (2010)'s theory is in great part due to its ability to replicate better than the baseline model the sharp improvement in accuracy observed after the beginning of the updating period during the crisis. This is likely due to the non-Bayesian nature of the weights on information assigned by Patton and Timmermann (2010)'s theory, which, as seen in Figure

9, for the crisis years have a spike mid-way through the updating period, a pattern that the Bayesian weights in the baseline theory cannot generate.

Summing up, the structural estimation of our theory, as well as the fit comparisons with alternative theories support the following consistent message: In normal times, agents use heterogeneous models to which they attach high faith, and, when updating, they incorporate information directly linked to the forecasted variable using Bayes' rule. During the crisis agents lose faith in their model, and no longer behave like Bayesians.

9 Conclusions

Expectations are a key determinant of economic decisions and a building block of macro and finance models. The increasing role of expectation manipulation in monetary policy makes it important not only to understand how expectations are formed and how they evolve over time, but also how they respond to communication of public information. A large literature in macroeconomics studies stochastic dynamics systems in order to predict the dynamic evolution of macroeconomic variables such as consumption, income and investment. In line with earlier works, this paper shows that information rigidities and heterogeneous models are key elements of any theory of expectation formation that can explain the dynamic patterns of accuracy and disagreement observed in surveys of professional forecasters. However, it departs from previous contributions by showing that Bayesian updating by agents who pay attention to public information that is directly linked to the variable they are forecasting (in our case the monthly price index for forecasting annual inflation) fits the data remarkably well in non-crisis years. Our theory is simple, assumes rational updating and it empirically outperforms most leading theories of information rigidities.

A Baseline Model Derivations

Optimal Forecast: We derive the conditional mean of annual inflation at different horizons h. The forecasted variable y is approximately equal to the last twelve realizations of x_h:

$$y = \sum_{h=0}^{11} x_h.$$

Since $x_h = c_i + \phi x_{h+1} + v_h$, the unconditional mean of x_h is $\widetilde{\mu}_i \equiv \frac{c_i}{1-\phi}$ and:

$$\begin{aligned}
x_0 &= \widetilde{\mu}_i + \phi^{12}(x_{12} - \widetilde{\mu}_i) + \sum_{j=0}^{11} \phi^j v_j \\
x_1 &= \widetilde{\mu}_i + \phi^{11}(x_{12} - \widetilde{\mu}_i) + \sum_{j=0}^{10} \phi^j v_{j+1} \\
\dots \\
x_{11} &= \widetilde{\mu}_i + \phi(x_{12} - \widetilde{\mu}_i) + v_{11},
\end{aligned}$$

which implies

$$y = 12\widetilde{\mu}_i + \frac{\phi(1-\phi^{12})}{1-\phi}(x_{12} - \widetilde{\mu}_i) + \sum_{j=0}^{11} \frac{1-\phi^{j+1}}{1-\phi} v_j.$$

At horizon $h \geq 12$ we have:

$$y = 12\widetilde{\mu}_i + \frac{\phi^{h-11}(1-\phi^{12})}{1-\phi}(x_h - \widetilde{\mu}_i) + \sum_{j=0}^{11} \frac{1-\phi^{j+1}}{1-\phi} v_j + \sum_{j=12}^{h-1} \frac{\phi^{j-11}(1-\phi^{12})}{1-\phi} v_j. \quad (21)$$

Within the target year 11 - h realizations of x_h are observed. Hence, for $h \leq 11$:

$$y = h\widetilde{\mu}_i + \frac{\phi(1-\phi^h)}{1-\phi}(x_h - \widetilde{\mu}_i) + \sum_{j=h}^{11} x_j + \sum_{j=0}^{h+1} \frac{1-\phi^{j+1}}{1-\phi} v_j. \quad (22)$$

The optimal forecast for the annual inflation rate at horizon $h \geq 12$ and $h \leq 11$ is equal to the conditional expectation of equations (21) and (22) respectively, i.e. $z_{i,h} = E[y_t | x_h, x_{h+1}, \dots]$. Given the definition for μ_i we obtain equations (8) and (9).

Signal's Precision: Agents evaluate the signal's precision trough the lens of their models. Given the signal $z_{i,h} = y + \varepsilon_{i,h}$ and the expressions for y and $z_{i,h}$ the variance of the forecast error at horizon $h \geq 12$ is derived as follows:

$$\begin{aligned}
\varepsilon_{i,h} &= y - z_{i,h} \\
&= \sum_{j=0}^{11} \frac{1-\phi^{j+1}}{1-\phi} v_j + \sum_{j=12}^{h-1} \frac{\phi^{j-11}(1-\phi^{12})}{1-\phi} v_j \\
E[\varepsilon_{i,h}^2] &= \sigma_v^2 \sum_{j=0}^{11} \frac{(1-\phi^{j+1})^2}{(1-\phi)^2} + \sigma_v^2 \sum_{j=12}^{h-1} \frac{\phi^{2j-22}(1-\phi^{12})^2}{(1-\phi)^2}.
\end{aligned}$$

The first expression on the right hand side of the previous equation is:

$$
\begin{aligned}
\sigma_v^2 \sum_{j=0}^{11} \frac{(1-\phi^{j+1})^2}{(1-\phi)^2} &= \frac{\sigma_v^2}{(1-\phi)^2} \sum_{j=0}^{11} (1-\phi^{j+1})^2 \\
&= \frac{\sigma_v^2}{(1-\phi)^2} [(1-\phi)^2 + (1-\phi^2)^2 + \cdots + (1-\phi^{12})^2] \\
&= \frac{\sigma_v^2}{(1-\phi)^2} [12 - 2(\phi + \phi^2 + \cdots + \phi^{12}) + (\phi^2 + \phi^4 + \cdots + \phi^{24})] \\
&= \frac{\sigma_v^2}{(1-\phi)^2} \left[12 - \frac{2\phi(1-\phi^{12})}{1-\phi} + \frac{\phi^2(1-\phi^{24})}{1-\phi^2} \right].
\end{aligned}
$$

The second expression is:

$$
\begin{aligned}
\sigma_v^2 \sum_{j=12}^{h-1} \frac{\phi^{2j-22}(1-\phi^{12})^2}{(1-\phi)^2} &= \sigma_v^2 \frac{(1-\phi^{12})^2}{(1-\phi)^2} \sum_{j=12}^{h-1} \phi^{2j-22} \\
&= \sigma_v^2 \frac{(1-\phi^{12})^2}{(1-\phi)^2} (\phi^2 + \phi^4 + \cdots + \phi^{2h-2-22}) \\
&= \sigma_v^2 \frac{(1-\phi^{12})^2}{(1-\phi)^2} \frac{\phi^2(1-\phi^{2h-24})}{1-\phi^2} \\
&= \sigma_v^2 \frac{\phi^2(1-\phi^{12})^2(1-\phi^{2h-24})}{(1-\phi)^3(1+\phi)}.
\end{aligned}
$$

Summing the two expressions we obtain the expression for b_h^{-1} for $h \geq 12$, while relying on the same derivation as in the first equation, we get the expression for $h \leq 11$.

Baseline with Noise: The variance of the forecast error for $h \geq 12$ is:

$$
\begin{aligned}
\epsilon_{i,h} &= y - z_{i,h} \\
&= \frac{\phi^{h-11}(1-\phi^{12})}{1-\phi}(x_h - \widehat{x}_{i,h}) + \sum_{j=0}^{11} \frac{1-\phi^{j+1}}{1-\phi} v_j + \sum_{j=12}^{h-1} \frac{\phi^{j-11}(1-\phi^{12})}{1-\phi} v_j \\
E[\epsilon_{i,h}^2] &= \frac{\phi^{2h-22}(1-\phi^{12})^2}{(1-\phi)^2} E(x_h - \widehat{x}_{i,h})^2 + \sigma_v^2 \sum_{j=0}^{11} \frac{(1-\phi^{j+1})^2}{(1-\phi)^2} + \sigma_v^2 \sum_{j=12}^{h-1} \frac{\phi^{2j-22}(1-\phi^{12})^2}{(1-\phi)^2}.
\end{aligned}
$$

The last two expressions on the right hand side of $E[\epsilon_{i,h}^2]$ are the same as in the baseline case. The same argument can be used for $h \leq 11$:

$$
\begin{aligned}
\epsilon_{i,h} &= \left(\frac{\phi(1-\phi^h)}{1-\phi}+1\right)(x_h-\widehat{x}_{i,h})+\sum_{j=0}^{h+1}\frac{1-\phi^{j+1}}{1-\phi}v_j \\
E[\epsilon_{i,h}^2] &= \left(\frac{1-\phi^{h+1}}{1-\phi}\right)^2 E(x_h-\widehat{x}_{i,h})^2+\sigma_v^2\sum_{j=0}^{11}\frac{(1-\phi^{j+1})^2}{(1-\phi)^2}.
\end{aligned}
$$

Hence, the variance of the forecast error also depends on $P_{i,h|h} \equiv E(x_h - \widehat{x}_{i,h})^2$ which is the mean squared error of the Kalman Filter. All agents solve the same recursion based on the same parameters, so that P_{ih} =P$_h$. Hence, these two expressions above give the signal's precision in the baseline model with noise $\widehat{b}_h^{-1}$.

B The Fit of Variations of Baseline Theory

We report graphically the fit the variants of the baseline theory. The RMSE and disagreement from the data are computed as averages over all years and separately for crisis and non-crisis years. We estimate each of the theories to fit the data on average and we perform separate estimations for crisis and non-crisis years.

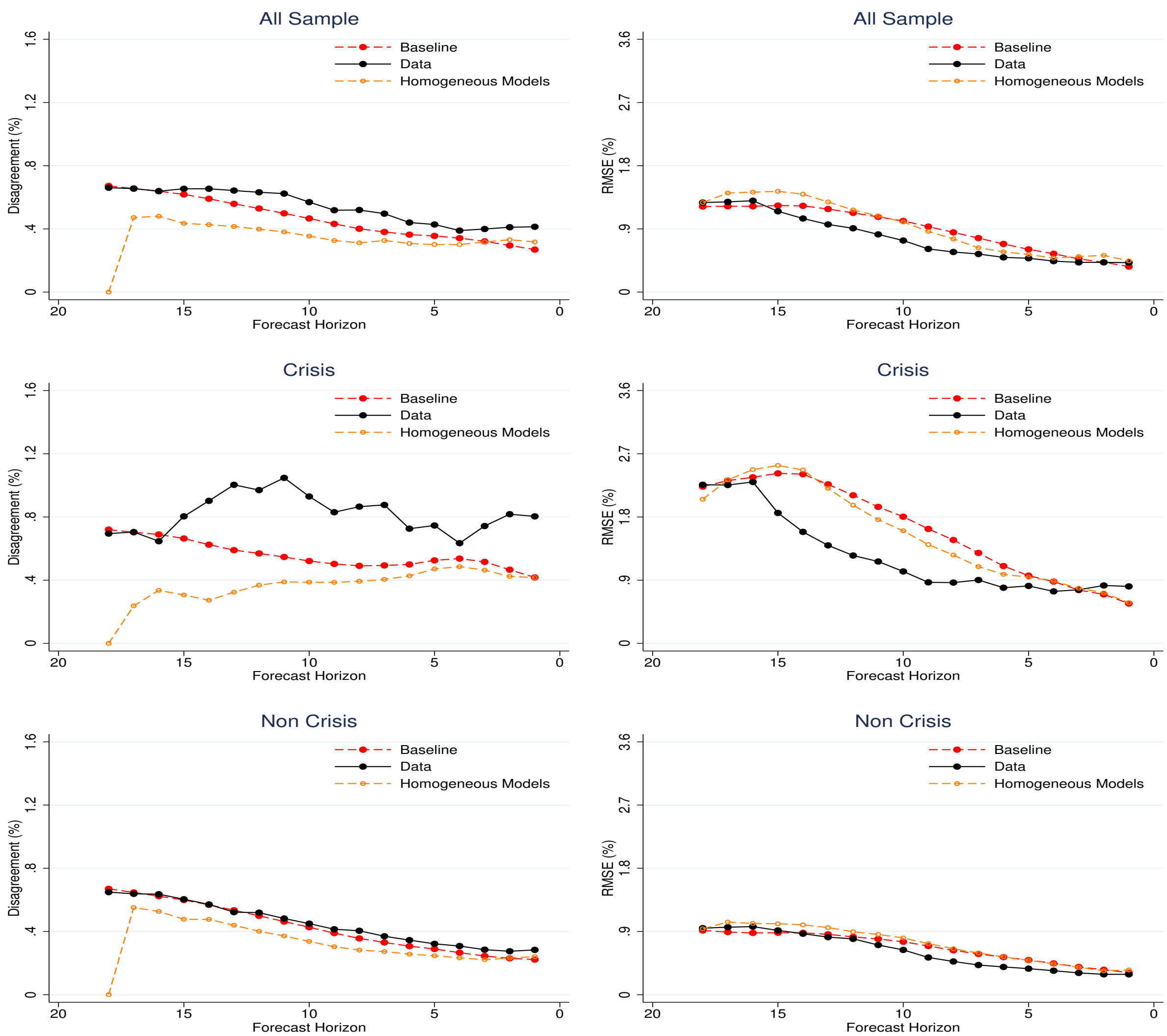


Figure 12: **Homogeneous Models:** Disagreement and Accuracy

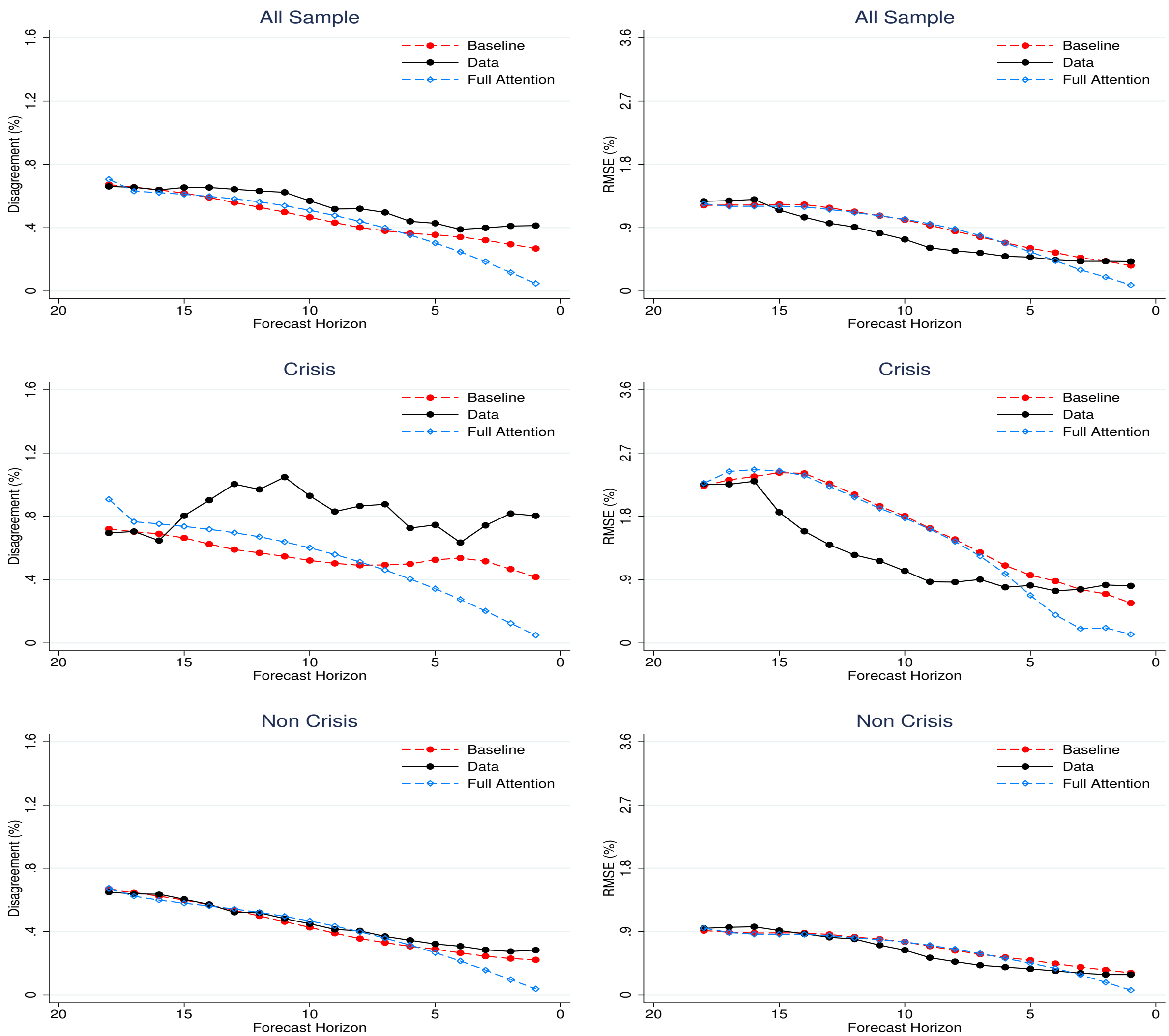


Figure 13: **Full Attention:** Disagreement and Accuracy

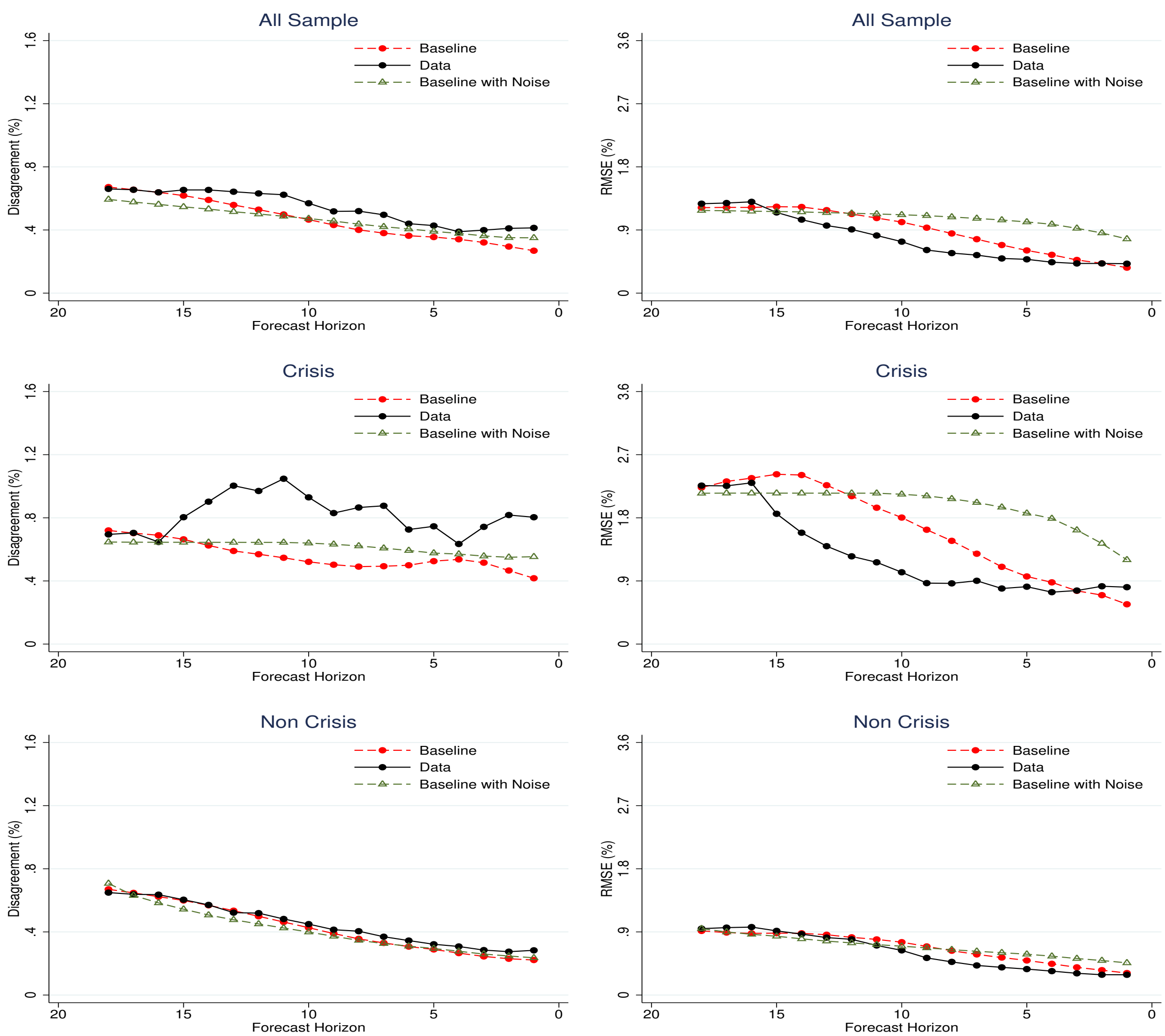


Figure 14: **Baseline with Noise:** Disagreement and Accuracy

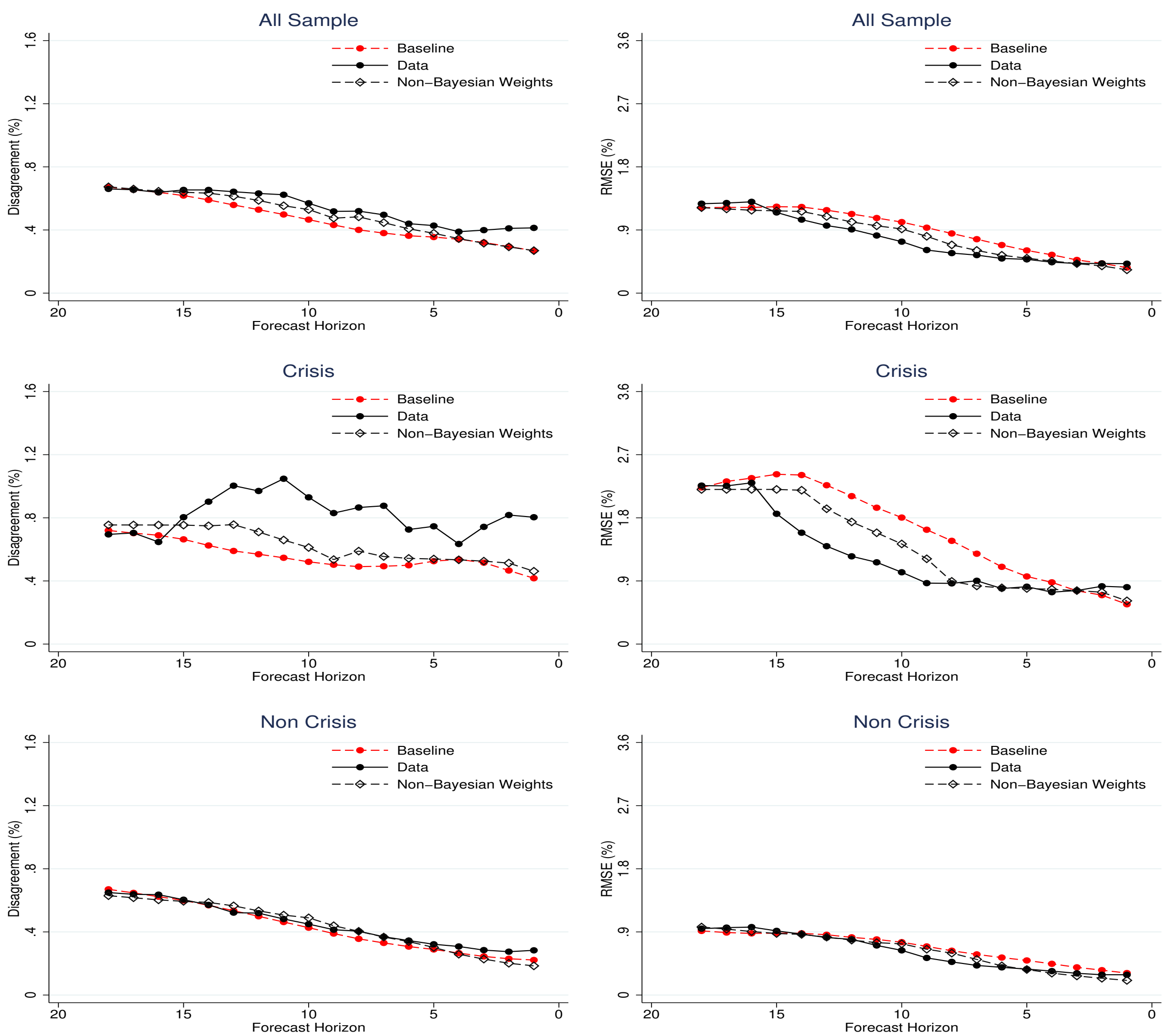


Figure 15: **Data-Driven Weights:** Disagreement and Accuracy

C Appendix: The Fit of Alternative Theories

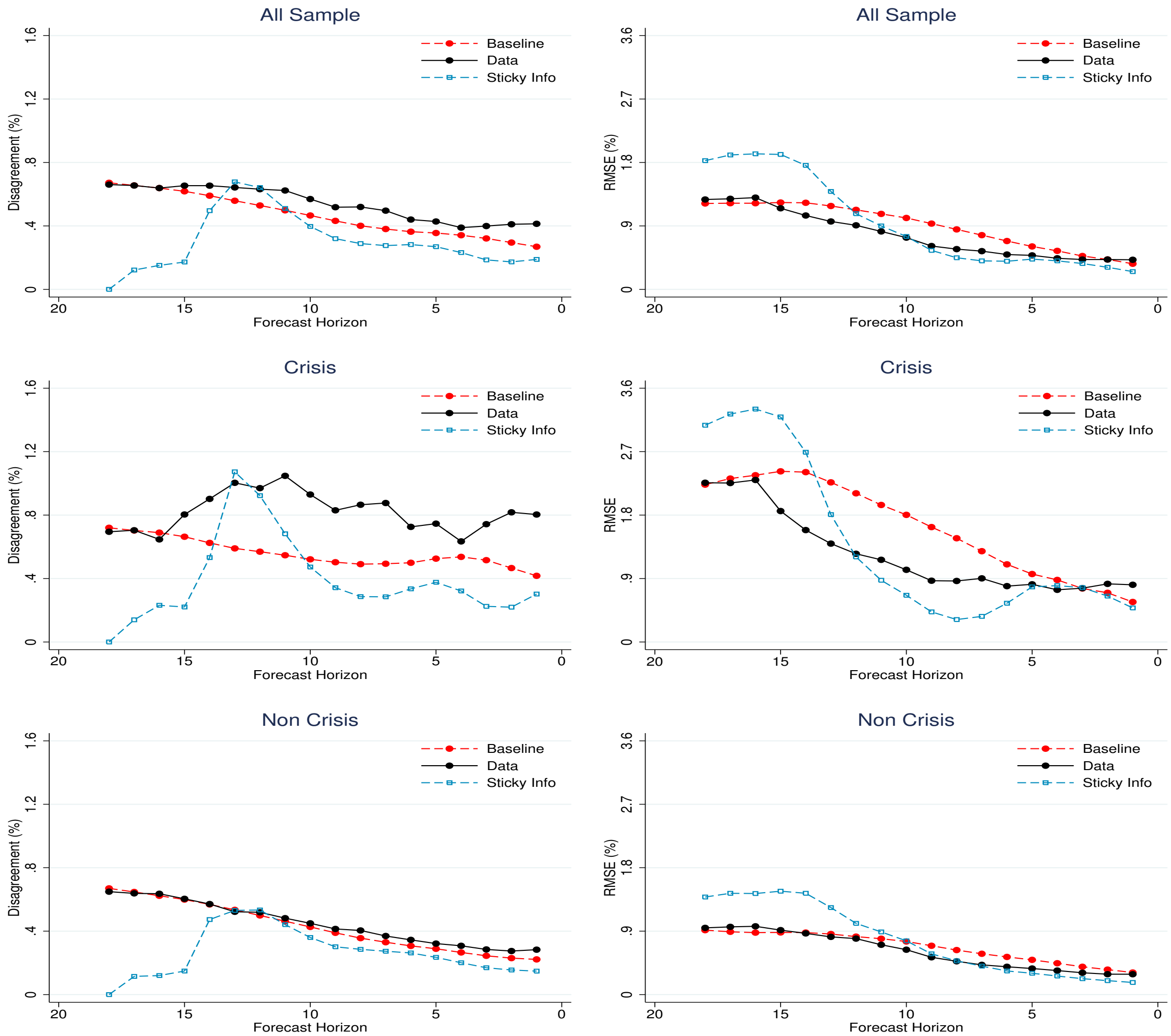


Figure 16: **Sticky Information:** Disagreement and Accuracy

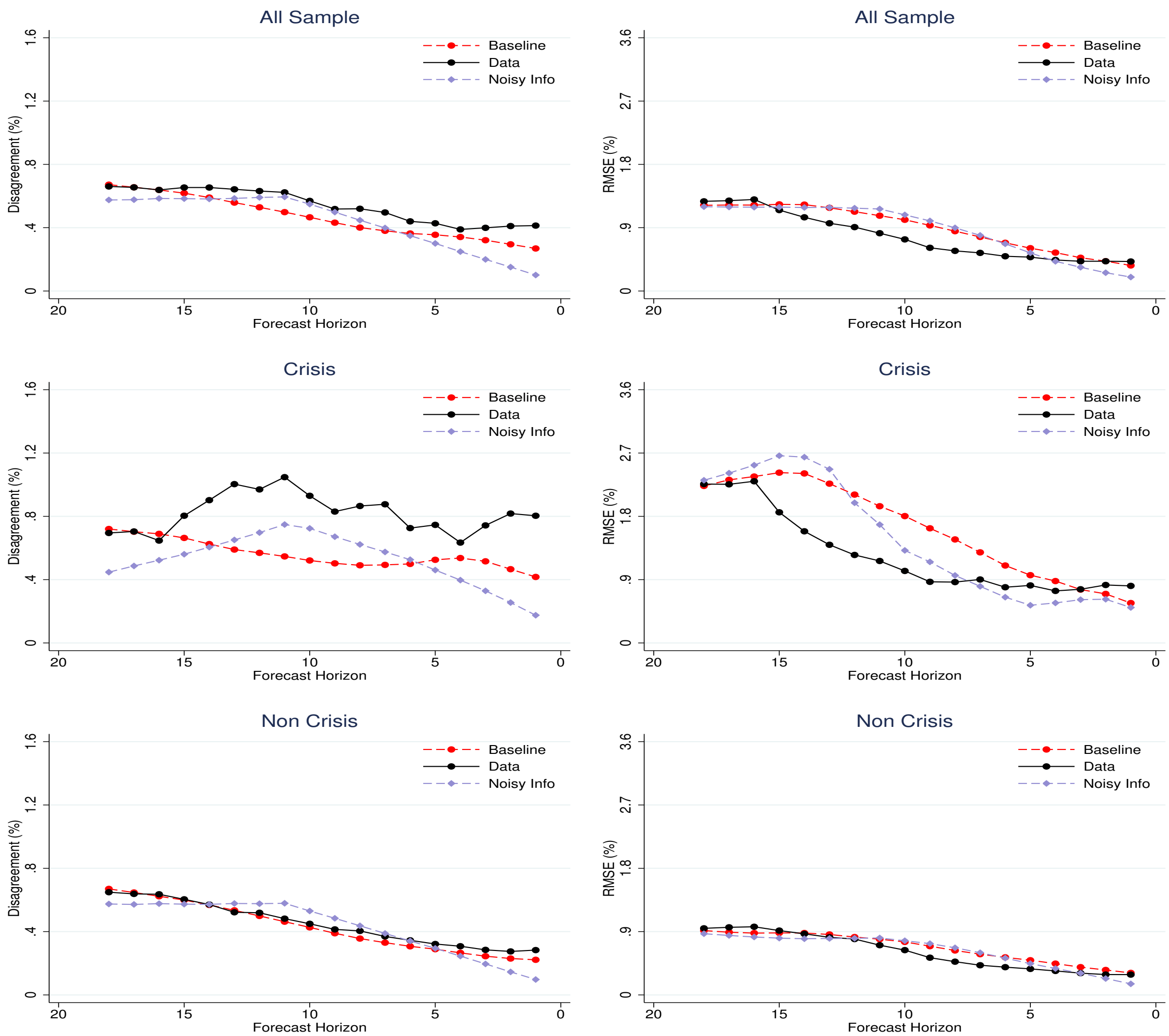


Figure 17: **Noisy Information:** Disagreement and Accuracy

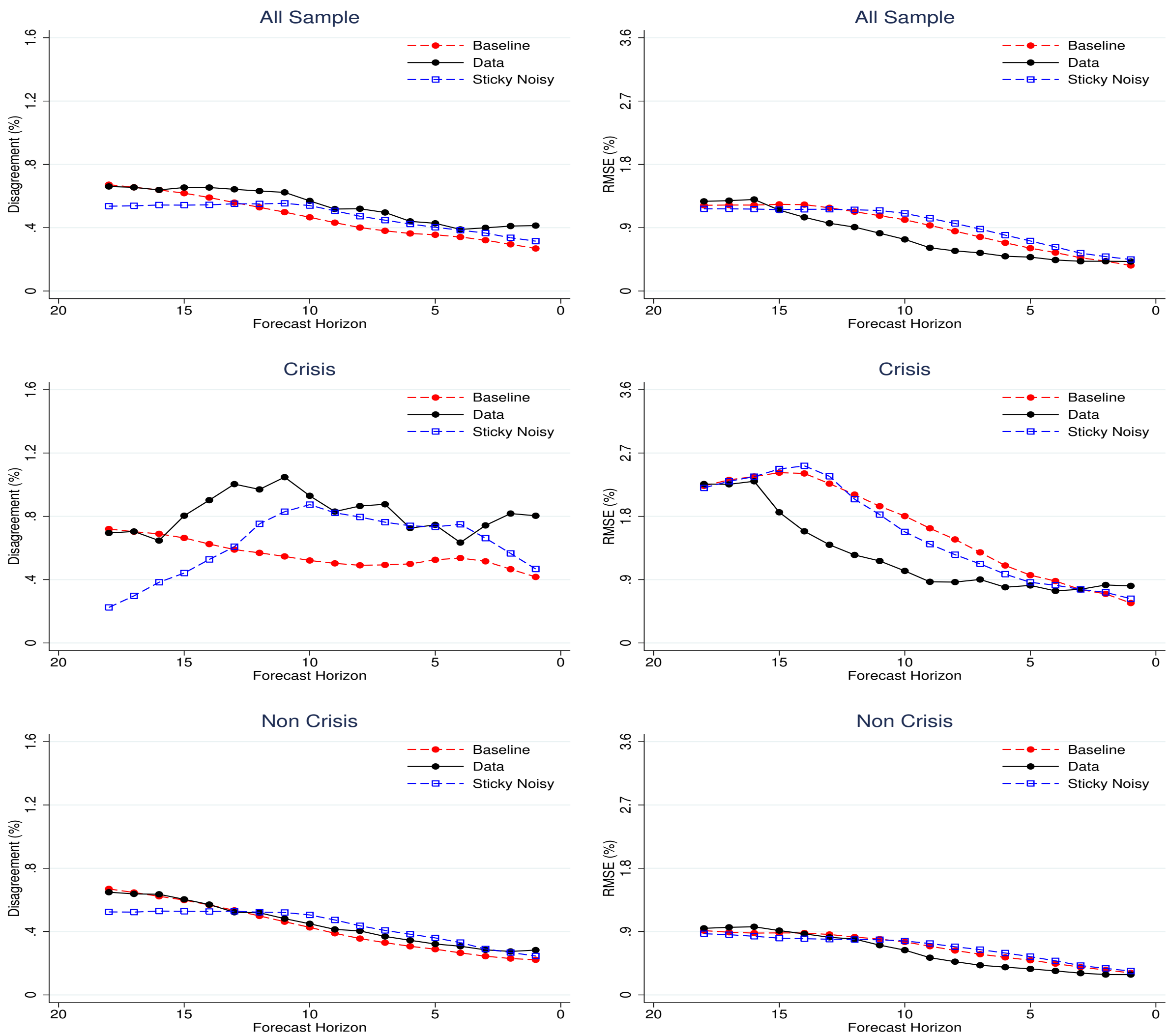


Figure 18: **Sticky Noisy:** Disagreement and Accuracy

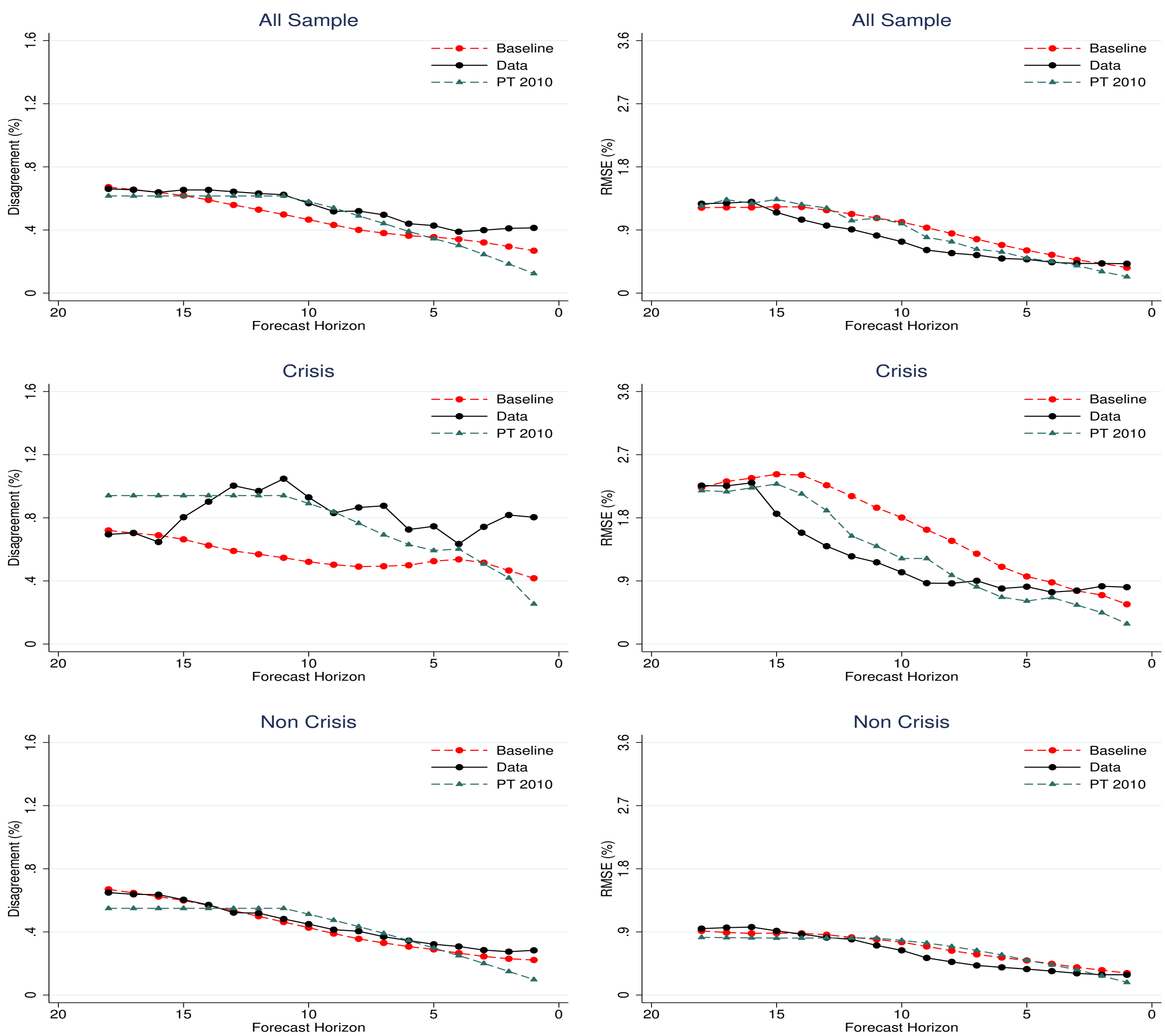


Figure 19: **PT 2010:** Disagreement and Accuracy

References

Altonji, Joseph G. and Lewis M. Segal, "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Bussiness and Economic Statistics*, 1996, *14* (3), 353–366.

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

Ang, Andrew, Geert Bekaert, and Min Wei, "Do Macro Variables, Asset Markets or Surveys Forecast Inflation Better?," *Journal of Monetary Economics*, 2007, *54* (4), 1163 – 1212.

Caskey, John, "Modeling the formation of price expectations: A Bayesian approach," *The American Economic Review*, 1985, pp. 768–776.

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

_ and _ , "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, 2015, *105* (8), 2644 – 2678.

_ , _ , and Saten Kumar, "How do firms form their expectations? new survey evidence," Technical Report, National Bureau of Economic Research 2015.

Dovern, Jonas, "When are GDP forecasts updated? Evidence from a large international panel," *Economics Letters*, 2013, *120* (3), 521–524.

Duffie, Darrell and Kenneth J Singleton, "Simulated Moment Estimation of Markov Models of Asset Prices," *Econometrica*, 1993, *61* (4), 929–952.

Ehrbeck, Tilman and Robert Waldmann, "Why are professional forecasters biased? Agency versus behavioral explanations," *The Quarterly Journal of Economics*, 1996, pp. 21–40.

Gourieroux, Christian and Alain Monfort, *Simulation-Based Econometric Methods*, Oxford University Press, 1996.

Lahiri, Kajal and Xuguang Sheng, "Evolution of forecast disagreement in a Bayesian learning model," *Journal of Econometrics*, 2008, *144* (2), 325–340.

Lamont, Owen A, "Macroeconomic forecasts and microeconomic forecasters," *Journal of economic behavior & organization*, 2002, *48* (3), 265–280.

Laster, David, Paul Bennett, and In Sun Geoum, "Rational Bias In Macroeconomic Forecasts," *The Quarterly Journal of Economics*, February 1999, *114* (1), 293–318.

Mankiw, N Gregory and Ricardo Reis, "Sticky Information versus Sticky Prices: A proposal to replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, 2002, *117* (4), 1295–1328.

_ , _ , and Justin Wolfers, “Disagreement about inflation expectations,” in “NBER Macroeconomics Annual 2003, Volume 18,” The MIT Press, 2004, pp. 209–270.

Manzan, Sebastiano, “Differential interpretation in the Survey of Professional Forecasters,” *Journal of money, Credit and Banking*, 2011, *43* (5), 993–1017.

Nimark, Kristoffer, “Man-Bites-Dog Business Cycles,” *The American Economic Review*, 2014, *104* (8), 2320–2367.

Ortoleva, Pietro, “Modeling the change of paradigm: Non-Bayesian reactions to unexpected news,” *The American Economic Review*, 2012, *102* (6), 2410–2436.

Patton, Andrew J and Allan Timmermann, “Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion,” *Journal of Monetary Economics*, 2010, *57* (7), 803–820.

Pesaran, M Hashem and Martin Weale, “Survey expectations,” *Handbook of economic forecasting*, 2006, *1*, 715–776.

Ruge-Murcia, Francisco, “Estimating nonlinear DSGE models by the simulated method of moments: With an application to business cycles,” *Journal of Economic Dynamics and Control*, 2012, *36* (6), 914–938.

Sims, Christopher A, “Implications of rational inattention,” *Journal of monetary Economics*, 2003, *50* (3), 665–690.

Tauchen, George, “Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained from Financial Market Data,” *Journal of Bussiness and Economic Statistics*, 1986, *4* (4), 397–416.

Woodford, Michael, *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Princeton: Princeton University Press, 2003.