Models, Inattention and Expectation Updates*

Raffaella GIACOMINI † Vasiliki SKRETA ‡ Javier TURÉN §

December 12, 2015

Abstract

We formulate a theory of expectation 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. This simple, "micro-founded" theory could enhance the explanatory power of macroeconomic and finance models.

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

JEL CLASSIFICATION: E27, E37, D80, D83.

*to be added

†UCL. r.giacomini@ucl.ac.uk

‡UCL. vskreta@gmail.com

§UCL. javier.roman.12@ucl.ac.uk

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? What information do agents incorporate: publicly available data relevant to the variable, or 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 theory fits remarkably well the dynamic evolution of both forecast accuracy and disagreement in the data and our empirical findings support the following narrative: Agents use heterogeneous models to form an initial forecast and have heterogeneous degrees of attention. Attentive agents update using Bayes' rule and, rather than herd, they incorporate public information about the forecasted variable, to which they attribute homogeneous weights. In normal times, agents have high faith in their models but their faith sharply decreases during crises. Although we do not explicitly model agents' behaviour during crises, our results support the hypothesis that during the 2008-2009 crisis agents discarded their initial model and no longer updated using Bayes' rule.

The ECFC survey dataset consist of a panel of approximately 75 professional forecasters who forecast US annual (year-on-year) CPI inflation (among other variables) for the years 2007 to 2014 during the 18 months before the release of the figure. We thus analyze the updates of fixed-event forecasts as the forecast horizon shrinks and more information becomes available. The survey has some unique features that make it well-suited for answering the above questions: participants can update their forecasts at *any* time and as often as they like and, upon logging onto the terminal, they observe in real time the consensus and each others' forecasts. The period 2007-2014 is rather turbulent, as Figure 1 suggests, so forecasting US inflation is not a trivial task.

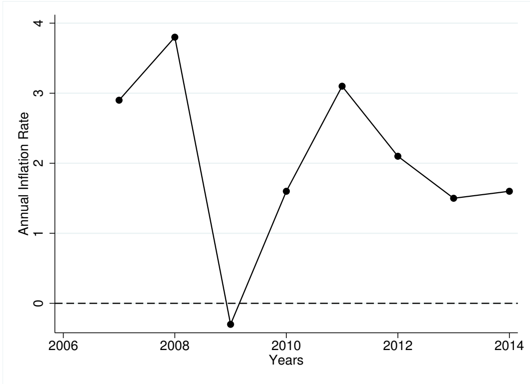


Figure 1: Annual US inflation 2007-2014

Our theory has three key elements: (i) not all agents update (some may be inattentive); (ii) agents may use heterogeneous forecasting models that can differ from year to year but remain fixed over the updating period (iii) agents who are attentive use Bayes' rule to incorporate public information and assign homogeneous weights to it. Figure 2 shows that our theory fits remarkably well the dynamics of disagreement (standard deviation of forecasts across agents) *and* of accuracy (Root Mean Squared Error, henceforth RMSE) in the data, except for the crisis years 2008-2009:

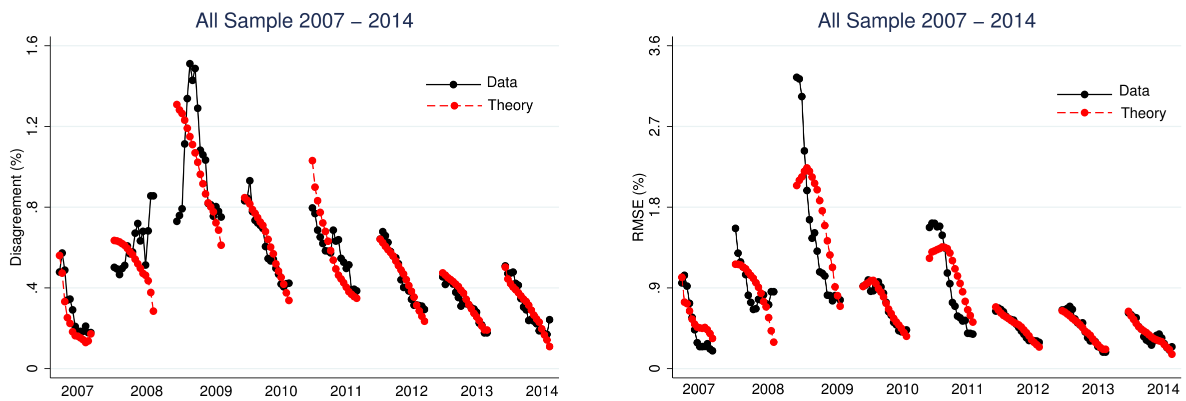


Figure 2: Fit of theory for disagreement (left) and RMSE (right)

The ingredients of our theory are motivated by the empirical regularities. First, agents in our dataset are inattentive: Given that in the ECFC survey we observe the exact moment updates occur, we can obtain an empirical proxy for attention. We find that on average only about 40%-50% of agents update at least once a month and the frequency of updates varies across agents and over the updating period, but is similar across years, including during the crisis. Updaters (those who revised) are more accurate than non-updaters, and inattention can explain why consensus forecasts violate rationality tests. Allowing for inattention thus appears to be an important ingredient when attempting to explain the patterns of accuracy in the data.

The second ingredient of our 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. 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 our theory allows 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 maybe be biased for a number of different reasons that we don't explicitly model. Because our theory focuses on forecast updates, any individual-specific bias that is constant over the updating period does not affect our analysis. Once we allow for heterogeneity in agents' attention (which can vary across agents and over time) and heterogeneity in agents' models (which can vary across agents and from year to year but remain constant during the updating period), our theory postulates that agents respond homogeneously to the arrival of information.

The final ingredient of our theory is the assumption that, as new information arrives, agents who are attentive update their previous forecast using Bayes' rule. What information do agents incorporate when they update? 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 as the most relevant for annual inflation: 1. the release of monthly CPI and 2. 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 information, but also available private 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. So what source of information–the monthly release of CPI or the consensus forecast–fits the data better? Our dataset is uniquely suited to address this question, since agents can observe the current consensus forecast at each moment they decide to update. In our baseline theory we assume that information is related to monthly CPI. More precisely, agents use an autoregressive model of order one (AR(1)) 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 could capture 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. We also consider a variation of the theory where information is the consensus forecast (agents "herd").

Structural estimation of our theory shows that the version where agents herd does not fit the data well, whereas the baseline theory where information is the monthly release of CPI fits remarkably well the observed dynamics of both accuracy and disagreement in normal times. Focusing on the baseline theory, we obtain estimates of the structural parameters, which include a parameter capturing agents' "faith" in their model.[1] We find that in normal times agents place high faith in their model, even though it may produce quite distinct forecasts compared to others. However, agents's faith drops dramatically during the crisis years 2008-2009 and our theory has trouble fitting the dynamics of disagreement and accuracy (see Figure 2). Interestingly, for the crisis years the ECFC agents are, for the most part, more accurate than our theory would have predicted. A possible explanation for this finding is that our theory assumes that agents use the same model over the updating period but this is a bad idea during a turmoil since there is often a paradigm shift. The fact that in the data the initial forecasts for the crisis years are highly inaccurate but their accuracy improves sharply during the forecasting period suggests that agents in our sample may have discarded their initial model for a new one and thus obtained a dramatic improvement in accuracy. Our theory does not allow for such model-shifting during the updating period. Still, our empirical estimates succeed in capturing this paradigm shift by showing that agents' "faith" in their model decreases from 199.42 for 2007 to 0.0001 for 2008![2]

By estimating a number of variations of our baseline theory we assess the relative importance of its key ingredients. We find that heterogeneity in models is crucial, as eliminating this source of heterogeneity would worsen the fit of our theory over all years by 156% for disagreement and by 4% for accuracy. In contrast, by eliminating inattention the fit of our theory would decrease by 37% for disagreement and by 15% for accuracy (see Table 4).

To evaluate how our theory fares compared to existing theories of information rigidities, we estimate some leading alternative theories using our new data. For *normal* years the fit of our theory is better than that of Mankiw and Reis (2002)'s "sticky information" theory (whose fit is worse by 581.6% for disagreement and 162.3% for accuracy), Sims (2003)'s "noisy information" theory (whose fit is worse by 117.4% for disagreement and 12.7% for accuracy), Andrade and

[1]Formally, "faith" is the precision of the initial forecast (inverse of the variance), assumed to be constant across agents.

[2]The paradigm shift we suspect is consistent with the findings of Bassetto, Messer and Ostrowski (2013) who show that "the recovery of inflation in 2009/2010 occurred precisely at the only time (since 1985) in which the statistical models considered here would predict sharp disinflation, that is, inflation went up at the time at which the models would most strongly predict that it should go down."

Le Bihan (2013)'s "sticky-noisy" theory (whose fit is worse by 61.8% for disagreement and 35.1% for accuracy) and Patton and Timmermann (2010)'s "long-run means" theory (whose fit is worse by 114.5% for disagreement and 30.9% for accuracy).[3]

For the *crisis* years, the fit of our theory remains better than that of alternative theories, apart from two instances: The first is Andrade and Le Bihan (2013)'s "sticky-noisy" theory, which fits disagreement and accuracy better (it outperforms ours by 21.6% and 14.8%, respectively). The reason for this is that our theory assumes that agents' models do not change over the updating period and this is what drives the disagreement. In Andrade and Le Bihan (2013)'s "sticky-noisy" theory the disagreement is driven by the noise in the information, which varies over time. The second is Patton and Timmermann (2010)'s "long-run means" theory, which fits better than ours for both accuracy and disagreement (it outperforms our theory by 38.7% and by 40.6% respectively). The key difference between our baseline and Patton and Timmermann (2010)'s theory (besides the fact that our theory has inattention and their theory has noisy information) is that they assume that agents assign ad-hoc (data-driven) weights to information, so agents in their theory do not follow Bayes' rule. This difference in fit suggests that Bayesian updating fits remarkably well in normal times, while not that well during the crisis when, perhaps, agents do not have adequate estimates about the precision of their information nor of their previous forecast. Remarkably, as we already noted, our theory with Bayesian updating fits better than Patton and Timmermann (2010)'s theory both accuracy and disagreement during non-crisis years despite the fact that Patton and Timmermann (2010) choose the weights to ex-post rationalize the data.

Summing up, the structural estimation of our theory, as well as the fit comparisons with alternative theories support the following conclusion: In normal times, agents place high faith in the heterogeneous models they used to produce their initial forecast. When agents are attentive, they update their forecast by incorporating information directly linked to the forecasted variable using Bayes' rule and assign homogeneous weights to it. During crises, agents lose faith in their initial model and no longer behave like Bayesians.

This paper is related to several theoretical and empirical literatures in economics: the literature on expectation formation (rational and behavioural), the literature that highlights the importance of information rigidities in expectations, and the literature on social learning. Our main contribution to the first two strands is to 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 learning, something that no existing theory has been able to accomplish so far. Lucas (1972) and Kydland and Prescott (1982), emphasize the presence of information frictions as a constraint faced by agents. The existence of these frictions prevents

[3] See Table 6

agents from observing or acquiring all the relevant information when making their decisions. While the full-information rational expectations (FIRE) hypothesis cannot be reconciled with the stylized facts in survey data, information rigidities can explain why the consensus forecast violates the FIRE hypothesis. The presence of these rigidities is backed by ample empirical evidence: Coibion and Gorodnichenko (2012) document the existence of information rigidities in survey data leveraging shocks and Coibion and Gorodnichenko (2015) propose an empirical specification to quantify these rigidities.

A further contribution to the literature on information rigidities is that we show that inattention is important in explaining the data, whereas (at least for inflation expectations) assuming that agents observe noisy signals is not empirically relevant. The theoretical foundations of inattention are in the seminal works by Sims (2003), Mankiw and Reis (2002) and Reis (2006). The empirical performance of these theories is investigated by Andrade and Le Bihan (2013) and Andrade, Crump, Eusepi and Moench (2014). Our theory builds on the theoretical underpinnings of this empirical work and, in addition to inattention, it incorporates the possibility that agents have heterogeneous models, a feature that was found to be important by Patton and Timmermann (2010) and Lahiri and Sheng (2008) in explaining disagreement in survey data. Patton and Timmermann (2010) is the first contribution that formalises a theory of forecast formation with heterogeneous priors, but, in contrast to this paper, it does so using an ad-hoc behavioural learning rule. Our empirical results support their finding that heterogeneity in the initial forecasts is a crucial ingredient of a theory that can fit forecast disagreement. Our theory in addition shows the importance of accounting for inattention and, remarkably, it fits the dynamics of both disagreement and accuracy while maintaining Bayesian learning. It is also worth mentioning an early empirical contribution by Caskey (1985), who builds a Bayesian learning model that matches the mean Livingston forecasts from 1958-1982. Compared to this paper, we match the observed *dynamics* of accuracy and disagreement of forecast updates.

Two key features of our theory that set it apart from previous contributions is that, once we account for inattention and heterogeneous models, our agents use Bayesian learning and behave in a homogeneous fashion. These features are in apparent contrast with Manzan (2011) and Lahiri and Sheng (2008), who, respectively, show that professional forecasters violate Bayesian learning 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 theory agents are inattentive and that only attentive agents are Bayesian.

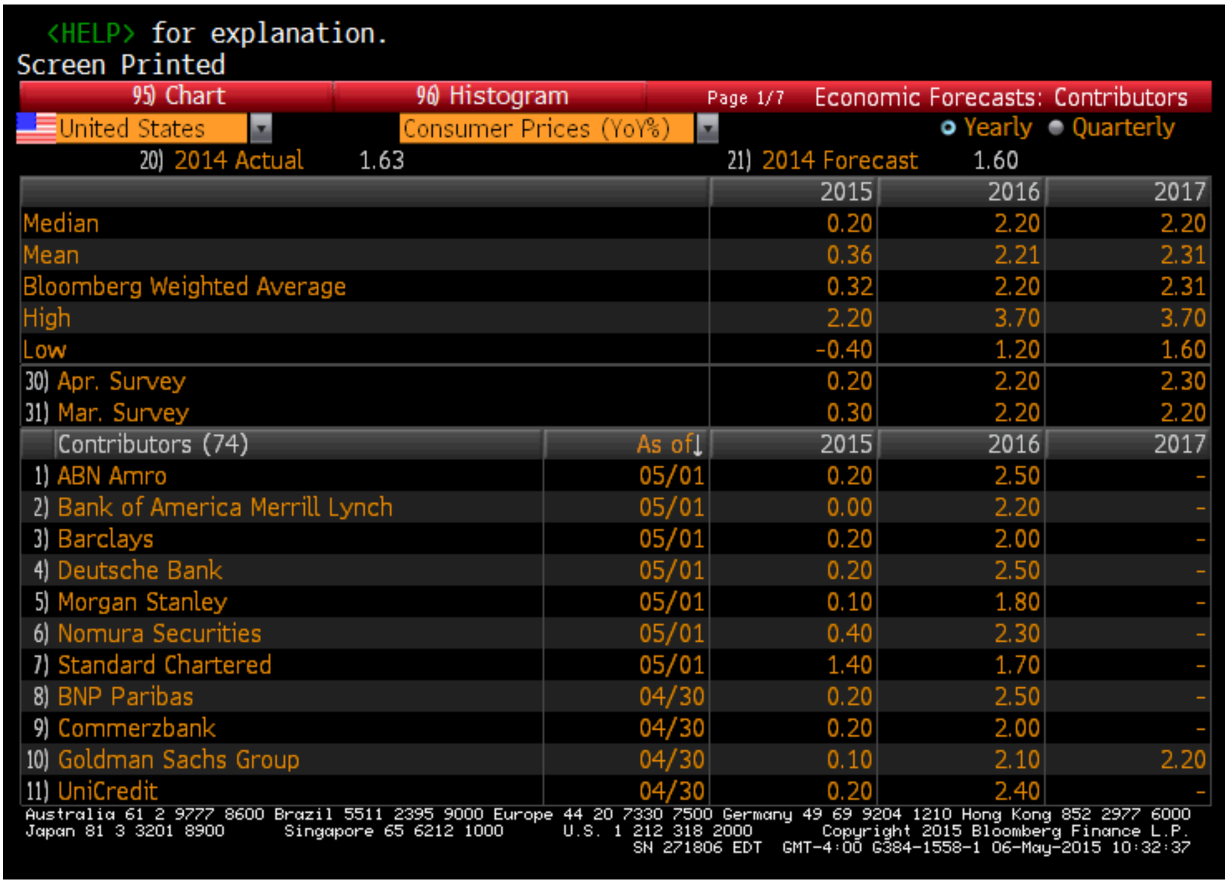
To our knowledge, this paper is the first to investigate what *source* of information agents use by empirically evaluating whether information relevant to the variable or herding fits better the observed patterns of accuracy and disagreement. Our findings suggest that agents trust their models and do not copy each other despite the fact that in our survey it is easy to do so.

This result is in contrast with other social or professional settings where agents tend to copy each other (see, e.g., the seminal contribution Bikhchandani, Hirshleifer and Welch (1992)).

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 presents stylized facts about forecast updates and their relationship with accuracy and disagreement. Section 4 describes our theory of expectation formation. The structural estimation and fit of our theory is described in Section 5. In Section 6 we investigate whether agents are herding by building and structurally estimating a variation of our theory where updaters pay attention to the consensus forecast. A number of further variations of the baseline theory (homogeneous models, full attention, noisy information, non-Bayesian learning) are described and estimated in Section 7. In Section 8 we briefly describe a number of leading theories of dynamic expectation formation with rigidities, estimate them and compare their fit to that of our baseline theory. 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 a forecast 18 months before the end of the year.

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%

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 (for simplicity, in the remainder of the paper we will not make this distinction explicit). 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: Based on the information provided by Bloomberg in private communications, 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 larger 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 this data publicly available.

3 Stylized Facts

Fact 1: Agents are inattentive: Figure 4 plots the proportion of agents that updated (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. On average over all years 44% of agents revise each month, whereas this number goes to up to 49% during the crisis.

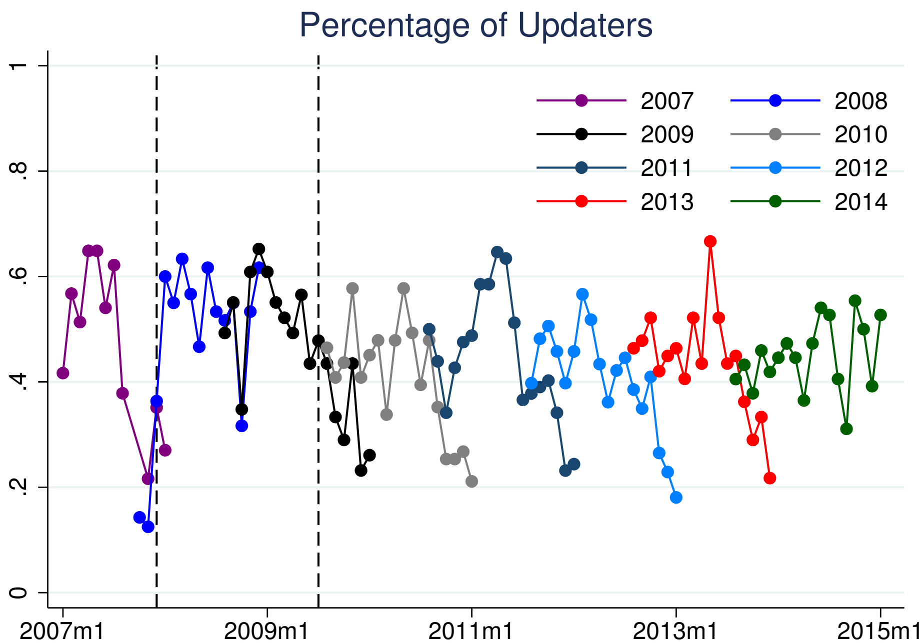


Figure 4: Monthly updating frequencies

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

updated 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 $r_{i,h} = 1$.

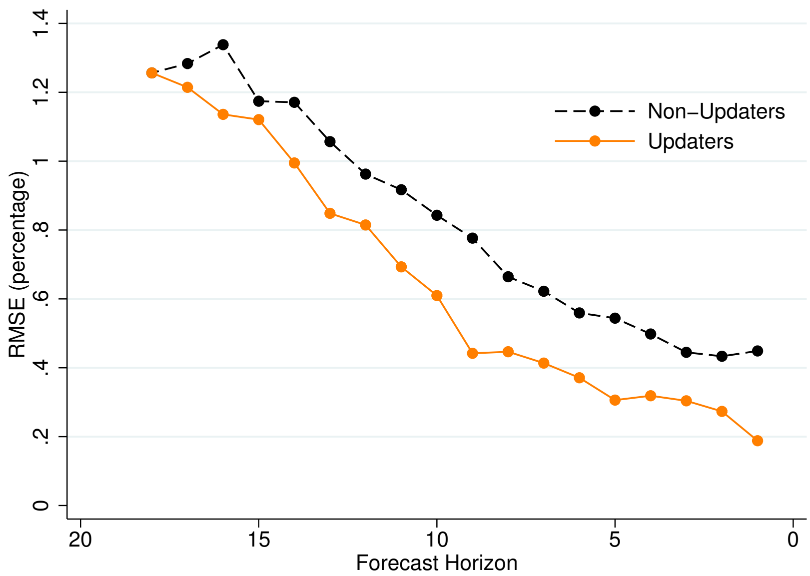


Figure 5: RMSE: average across years

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.[4]

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

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

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 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. 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.

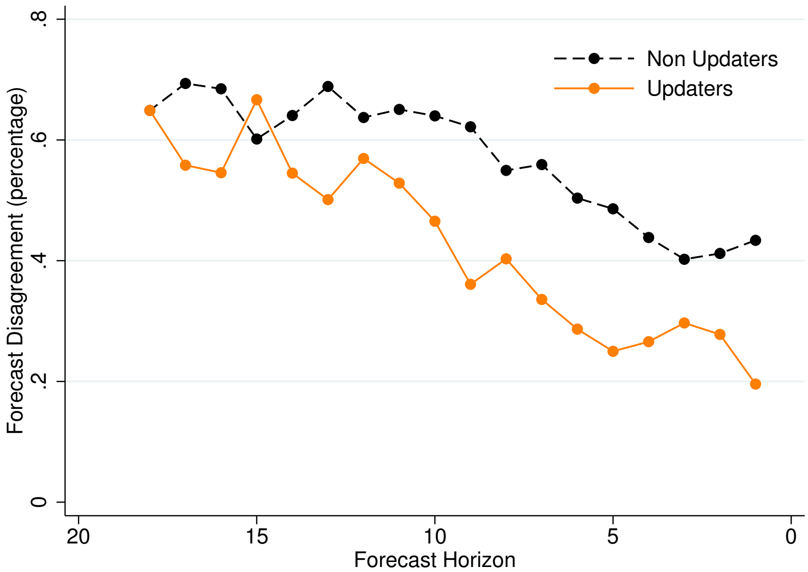


Figure 6: Disagreement: average across years

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 expectation updating and later establish that it fits remarkably well the observed dynamics of disagreement and accuracy in the data. 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.[5] 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

[5]Since we are describing the procedure for a *specific* year, we do not explicit index the calendar year.

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

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

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

The weight on the public signal is

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

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

Public signal: What source of public information do agents use when updating? In reality, there is a variety of publicly, but also privately, observed signals that may induce a forecaster to update: official statistics, news articles, central bank announcements and so forth. Among those sources of information the release of monthly CPI is particularly relevant since it can be used to construct year-on-year monthly inflation x_h, from which we can obtain a measure of annual inflation. This is clearly something that agents should pay attention to, since, from $h = 11, \ldots, 1$ it is (roughly) 1/12 of the variable they are forecasting. In what follows, we assume that the public signal is monthly CPI. Assuming that agents use only this signal and no additional information turns out to fit the data very well.

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

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

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:

$$x_h = c_i + \phi x_{h+1} + v_h, \tag{7}$$

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

[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

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 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.

5 Structural Estimation of Baseline Theory and Results

In this section we describe the structural estimation of our theory. In Sections 6 and 7 we use the same procedure to estimate a number of alternative specifications. These exercises allow us not only to establish the superiority of this baseline specification but also to get a sense of the relative contribution of each feature of the theory to the model's fit.

We seek to estimate the following parameters $\theta = (\lambda_h, \phi, \sigma_v^2, \mu, \sigma_\mu^2, a_{18})$ in order to get the theoretically predicted RMSE and disagreement to be as close as possible to their empirical counterparts.

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 a $N(\mu, \sigma_\mu^2)$. 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 data from FRED.[7] 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.[8]

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. 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 $h = 1, 3, 5, 7, 9, 12, 14, 16, 18$. The chosen moments are the disagreement and RMSE at different horizons h, delivering a total of 18 moments. We provide some further details of the estimation procedure in Appendix B.

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

[8] 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.

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 ρ 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-identified 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.

Our theory fits remarkably well the non-crisis years, but it has trouble fitting disagreement and accuracy during the crisis. Disagreement in the data is higher compared to what the model predicts. The theory predicts a higher RMSE compared to the data. However, the estimates can help shed light on why this is the case: our theory assumed that agents "stick" to their initial model. This is a bad idea during a turmoil since there is often a paradigm shift. The sharp decrease in RMSE is the data during the crisis 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 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.

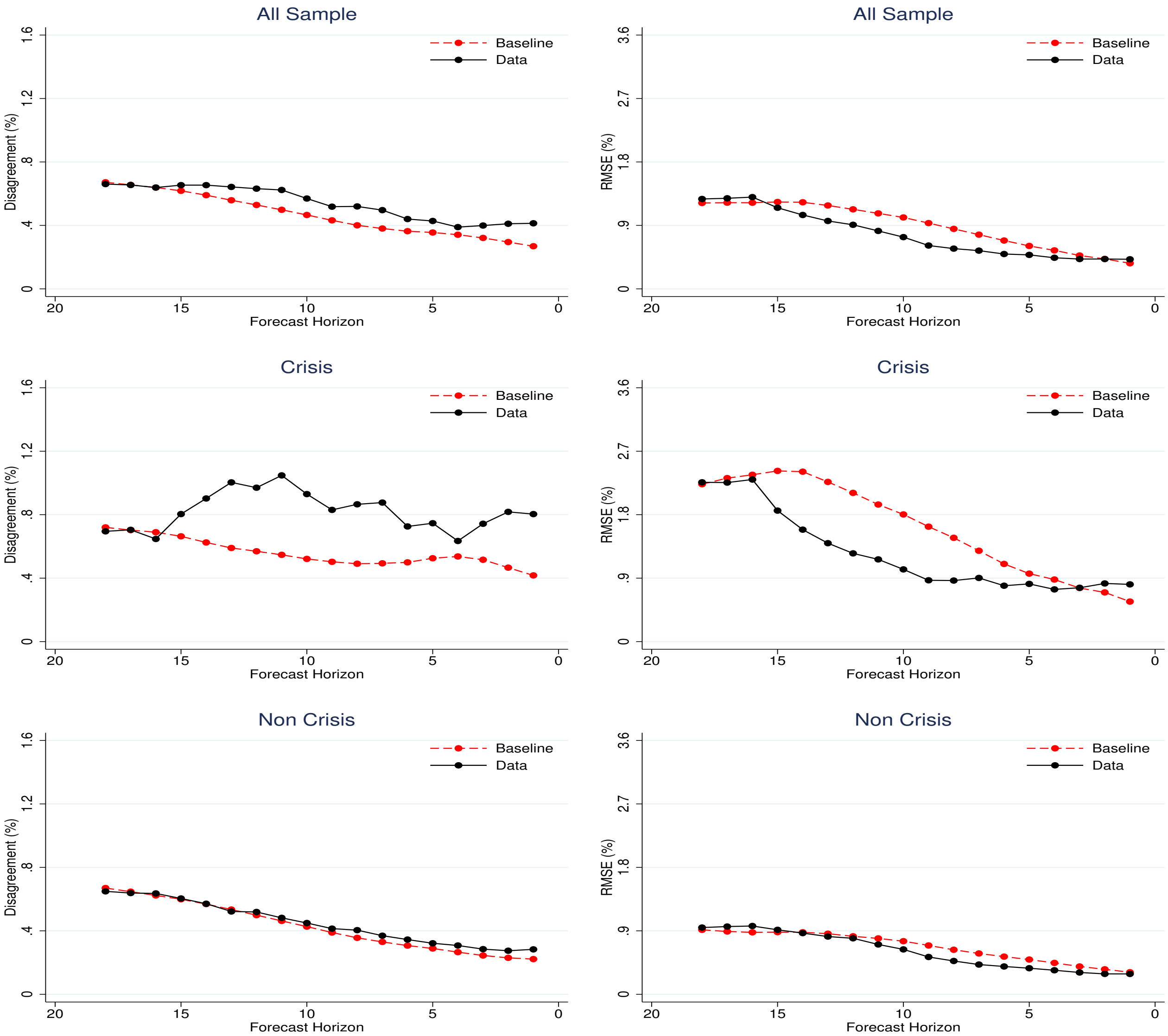


Figure 7: Fit of baseline theory by subsamples. Disagreement and RMSE

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.

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.

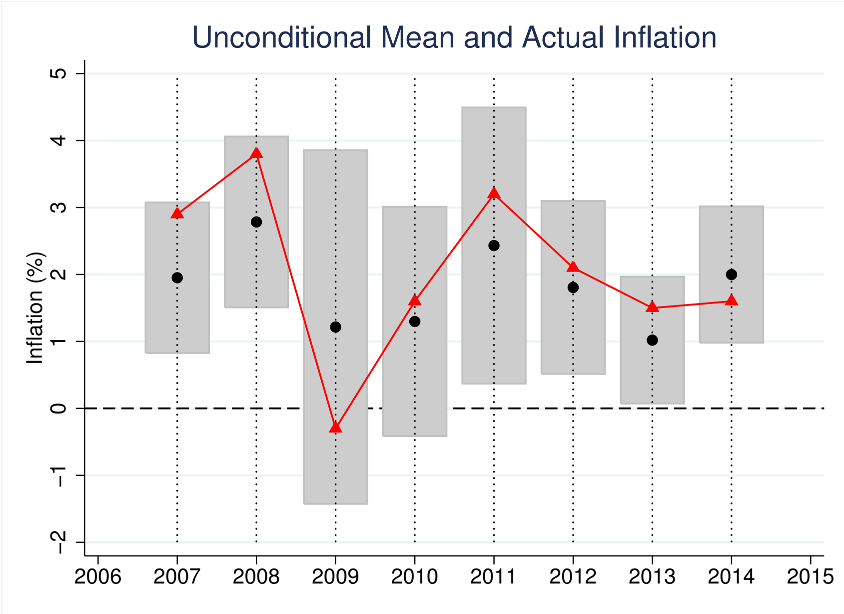


Figure 8: Estimated μ_i distribution and actual inflation

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 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 are visible to users of the Bloomberg terminal, and thus are known to our agents when they decide to update. We leverage this feature to build a variant of our 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 (the mean forecast), 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.[9]

We model "herding" by assuming that the common signal $z_{i,h} = z_h$ is the current month's

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

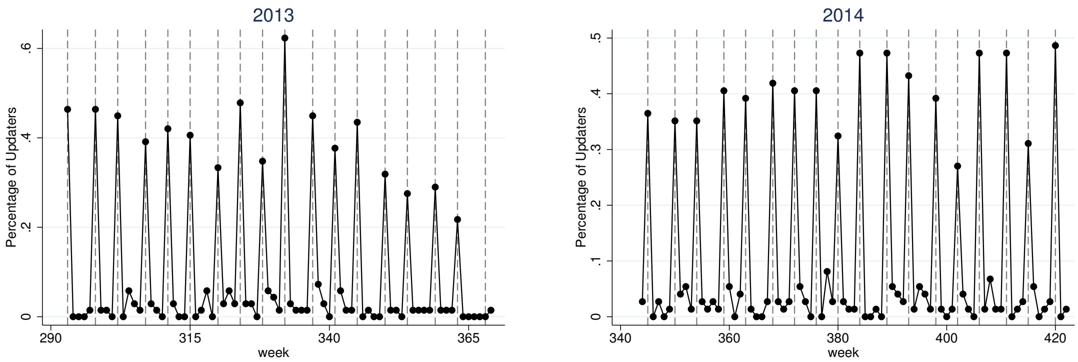


Figure 9: Weekly updating frequencies for years 2013 (left) and 2014 (right)

consensus forecast, which is compatible with the finding that the consensus forecast in the ECFC survey is an unbiased forecast of annual inflation (as shown in Table 2). Since there are several days in a month when revisions could potentially take place, it is not clear a priori when to measure the consensus forecast. However, one interesting pattern that emerges when measuring the fraction of updaters at the weekly frequency (see Figure 9)[10] 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 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

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

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 10. Table 4 further reports a summary measure of the relative fit of the "herding" theory compared to the baseline theory (among other variations). Figure 10 and Table 4 show that the fit of the "herding" theory is always sizeably worse than the fit of the baseline theory for both accuracy and disagreement and for crisis and non-crisis years.

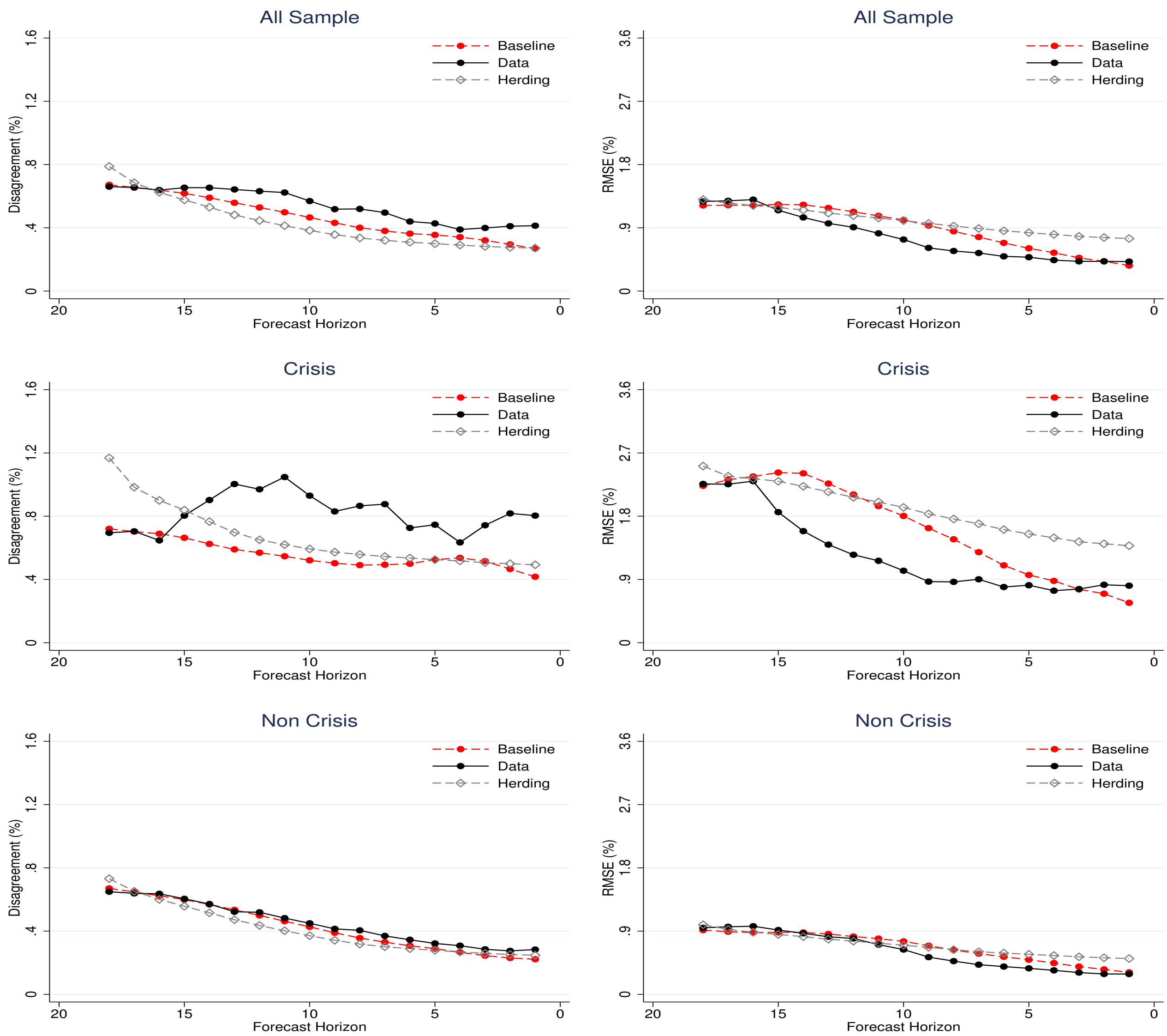


Figure 10: Fit of herding theory. Disagreement and RMSE

7 Importance of Different Ingredients of our Theory

As described in Section 4 our theory has three main features: i) inattention ii) heterogeneous models and iii) Bayesian updating using a public signal that is either monthly CPI or the consensus forecast. 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 Model. 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 space-state 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 $\epsilon_{i,h} = y - \widehat{z}_{i,h}$:

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

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

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 $P_{h|h}$.[11] 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.

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

S

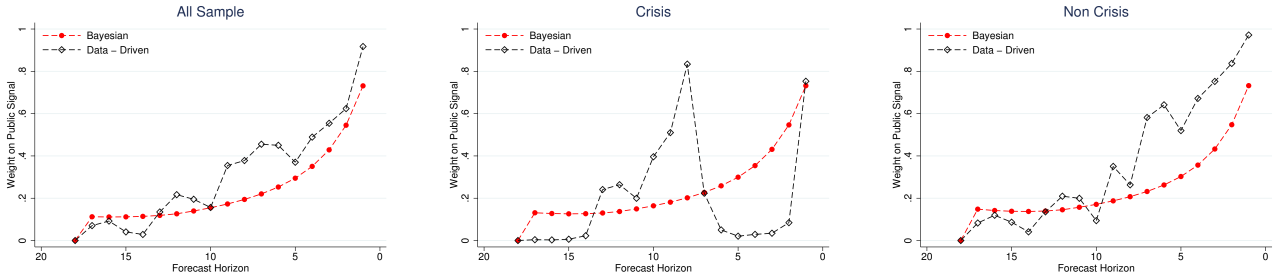


Figure 11: **Estimated Weights:** Baseline (Bayesian) and Data-Driven Weights

Structural Estimation: To estimate these alternative specifications we follow a procedure analogous to the one used to estimate the baseline we 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 C.1 we present all results graphically by plotting the theoretical and empirical RMSE and disagreement.

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. The fit of the "Herding" theory is worse than our baseline (by 72% for disagreement and 52% for accuracy), so monthly CPI fits the data better. 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 11, we further report the estimated data-driven weights, and we can see that the Bayesian weights assumed by our baseline theory are relatively close to the data-driven weights for all years and, in particular, for the non-crisis years. These findings provide empirical

Table 4: Relative Fit (%), Nested Models

	All years		Crisis		Non Crisis	
Models	Disag	Accuracy	Disag	Accuracy	Disag	Accuracy
Baseline	0	0	0	0	0	0
Herding	72.6	52.3	1.1	53.1	97.9	43.7
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.

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.

8 Comparison with Leading Alternative Theories

How does our theory fare compared to leading theories of expectation updating with information rigidities? To evaluate the relative performance we estimate these theories using the same methodology as that used to estimate our theory and compare their relative fit. We first briefly summarize the alternative theories we consider. Then, in Section 5 we estimate and compare the relative fit of these with our theory.

8.1 Leading Theories of Expectation Updates

Sticky Information. Originally proposed by Mankiw and Reis (2002), this theory involves inattentive agents that 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 $c_i = c$, which implies no disagreement among updaters.

Noisy Information. This theory is based on Sims (2003) and Mackowiak and Wiederholt (2009), and it assumes that all agents are constantly tracking new information, (i.e. "no inattention") but they observe a noisy signal as in equation (12) and use the Kalman filter to extract the information. Agents have homogeneous models so: $c_i = c, \ \ \forall i$. Then, at each horizon $h = 18 \ldots, 1$, agent's i forecast is equal to filtered public signal given by (13) and (14).[12]

Sticky-Noisy Theory. As originally stressed by Andrade and Le Bihan (2013), the presence of both inattention and disagreement within updaters could be reconciled by allowing for a hybrid Sticky-Noisy theory. 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 disagreement in long run means, 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:

$$\widehat{y}_{i,h} = w_h \mu_i + (1 - w_h) E[y|s_{i,h}] \tag{18}$$

$$w_h = \frac{E[e_{ih}^2]}{\kappa^2 + E[e_{ih}^2]} \tag{19}$$

$$e_{ih} \equiv y - E[y|s_{i,h}]. \tag{20}$$

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. 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 forecast increases. As before, κ is a free parameter, estimated to determine the degree of shrinkage.

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 with our new data. To do so, we estimate them using the procedure described in Section 5. All the non-nested specifications, except

[12] 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	✓			✓		✓

for Patton and Timmermann's theory, assume homogeneous models across agents, i.e. $\mu_i \equiv \mu$. Hence, except for this latter case, the distribution of models is again degenerate. For the **"sticky information"** model, the updating rule is given by (17) using the same random mechanism to determine the updaters as described in in Section 5. Since this theory implies no disagreement within updaters and at $h = 18$ every agent has to provide a forecast, we have zero disagreement at the beginning of the horizon. In the **"noisy information"** theory (and all the remaining theories) there is full attention, so every agent updates their forecast based on (13) and (14) with constant μ. Although the theory assumes homogeneous models, the noisy signal induces disagreement (12). The simulation for the **"sticky noisy"** model is simply a combination of the two previous procedures, where we replace z_h on equation (17) by $\widehat{z}_h$ as given by (13) and (14). Finally, for **"Patton and Timmermann's"** theory we first draw different μ_i as before and then let agents update their forecasts as (18) - (20), where $E(y|s_{i,h}) = \widehat{z}_h$ is again given by (13) and (14).[13]

Table 6 summarizes the results about their relative performance. For *non-crisis* years the fit of our theory is significantly better than that of the "sticky information" model (that model's fit is worse by 581.6% for disagreement and 162.3% for accuracy), of the noisy information model (that model's fit is worse by 117.4% for disagreement and 12.7% for accuracy), of the "sticky-noisy" model (that model's fit is worse by 61.8% for disagreement and 35.1% for accuracy) and of Patton and Timmermann (2010)'s "long-run means" model (that model's fit is worse by 114.5% for disagreement and 30.9% for accuracy).

For the *crisis* years, the fit of our model remains better apart from two instances: The

[13] Recall that although (13) starts from $h = 17$, this equation also holds for $h = 18$. This is because none of the alternative theories makes assumptions about the initial forecast.

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.

"sticky-noisy" model fits disagreement and accuracy better (our model is worse by 21.6% and 14.8%, respectively). The reason for this is that, in our theory, what drives the disagreement is the heterogeneous but time-invariant models of agents. In Andrade and Le Bihan (2013)'s "sticky-noisy" model disagreement is driven by the noise in the information which is time-variant and changes from crises to non-crises years. Coming to the second instance, Patton and Timmermann (2010)'s "long-run means" model fits better than ours in both accuracy and disagreement (our model is worse by 38.7% and by 40.6% respectively). The key difference between our baseline and Patton and Timmermann (2010)'s approach is that they assume that agents assign ad-hoc (data-driven) weights to information and to their previous forecast, so agents in their model do not follow Bayes' rule. This difference in fit suggests that Bayesian updating fits remarkably well in "normal times," while not that well during the crisis when, perhaps, agents do not have adequate estimates about the precision of their information nor of their model. Remarkably, as we already noted, our model with Bayesian updating fits better than Patton and Timmermann (2010)'s model both accuracy and disagreement during non-crisis years despite the fact that they choose the weights on information to ex-post rationalize the data.

Summing up, the structural estimation of our theory, as well as the aforementioned 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 crises agents lose faith in their model, and become uncertain about how to interpret new information.

9 Conclusions

Expectations are a key determinant of economic decisions. How economic agents form expectations is a key building block of macro and finance models and 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. Our 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. Our theory is simple, assumes rational learning and it empirically outperforms most leading theories of information rigidities. Due to its simplicity it can be incorporated in larger economic theories studying phenomena that depend on the dynamics of expectation formation. It would be interesting to evaluate in the future whether such models with more “micro-founded” expectations processes have enhanced explanatory power.

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:

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

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

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

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 Estimation Procedure

The SMM estimation solves:

$$\underset{\widehat{\theta}}{\text{Min}} \;\; \boldsymbol{G_h(\theta)'W_tG_h(\theta)}, \tag{23}$$

where

$$\boldsymbol{G_h(\theta)} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{m_{t,h}} - \frac{1}{\tau T}\sum_{t=1}^{\tau T} \boldsymbol{m_{t,h}(\theta)}$$

and where the observed accuracy and disagreement for different years t and horizons h are collected in $\boldsymbol{m_{t,h}}$ while their simulated theoretical counterparts are collected in $\boldsymbol{m_{t,h}(\theta)}$.

Here $\boldsymbol{G_h(\theta)}$ is a $q \times 1$ vector of moments, and $\boldsymbol{W_t}$ corresponds to a $q \times q$ positive definite weighting matrix. We let $\boldsymbol{W_t}$ equal the identity matrix, so we assign the same weight to all moments. The asymptotic distribution of the SMM estimator when $T \to \infty$ is

$$\sqrt{T}(\widehat{\theta} - \theta) \to N(\mathbf{0}, (1 + 1/\tau)(\boldsymbol{D'W^{-1}D})^{-1}), \tag{24}$$

where $\boldsymbol{D} = E(\partial \boldsymbol{m_{t,h}(\theta)}/\partial\theta)$ is the Jacobian evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{W}$ corresponds to the optimal weighting matrix. Based on (24), we obtain the standard deviation of the estimated parameters by computing the variance and covariance matrix of the moments $\boldsymbol{G_h(\theta)}$ given the estimated parameters, $\widehat{\boldsymbol{W_t}}$. Additionally, based on this matrix the test of over-identifying restrictions (J-test) is given by:

$$T\,(1 + 1/\tau)\,\boldsymbol{G_h(\widehat{\theta})}'\widehat{\boldsymbol{W_t}}\boldsymbol{G_h(\widehat{\theta})} \sim \chi^2_{q-p}. \tag{25}$$

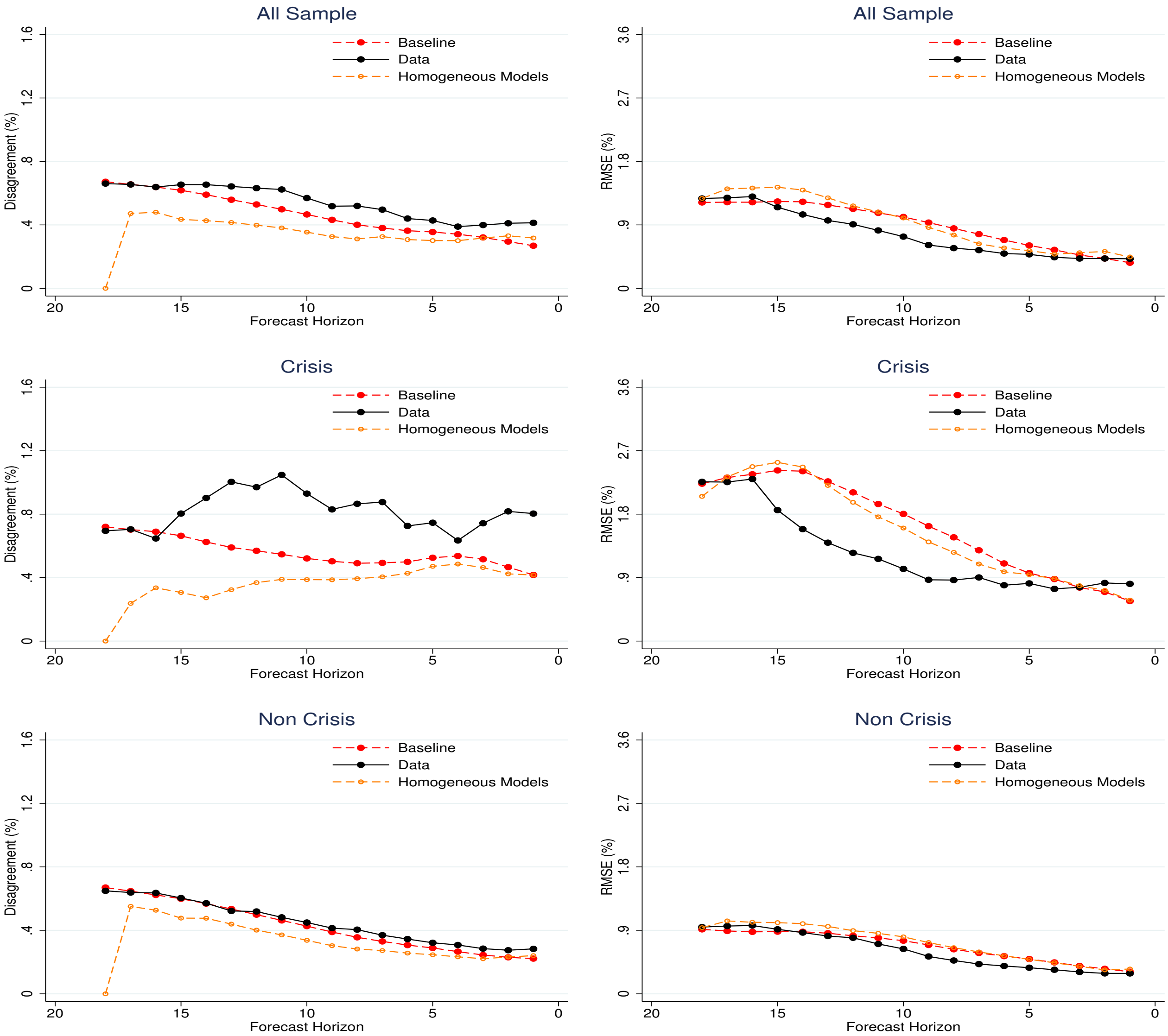


Figure 12: **Homogeneous Models:** Disagreement and Accuracy

C Further Empirical Results

We report graphically the fit the variants of the baseline theory and of the alternative theories discussed in Section 8. 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. For the variation of the baseline theory with data-driven weights we further report the estimated weights compared to the Bayesian weights.

C.1 The Fit of Variations of Baseline Theory

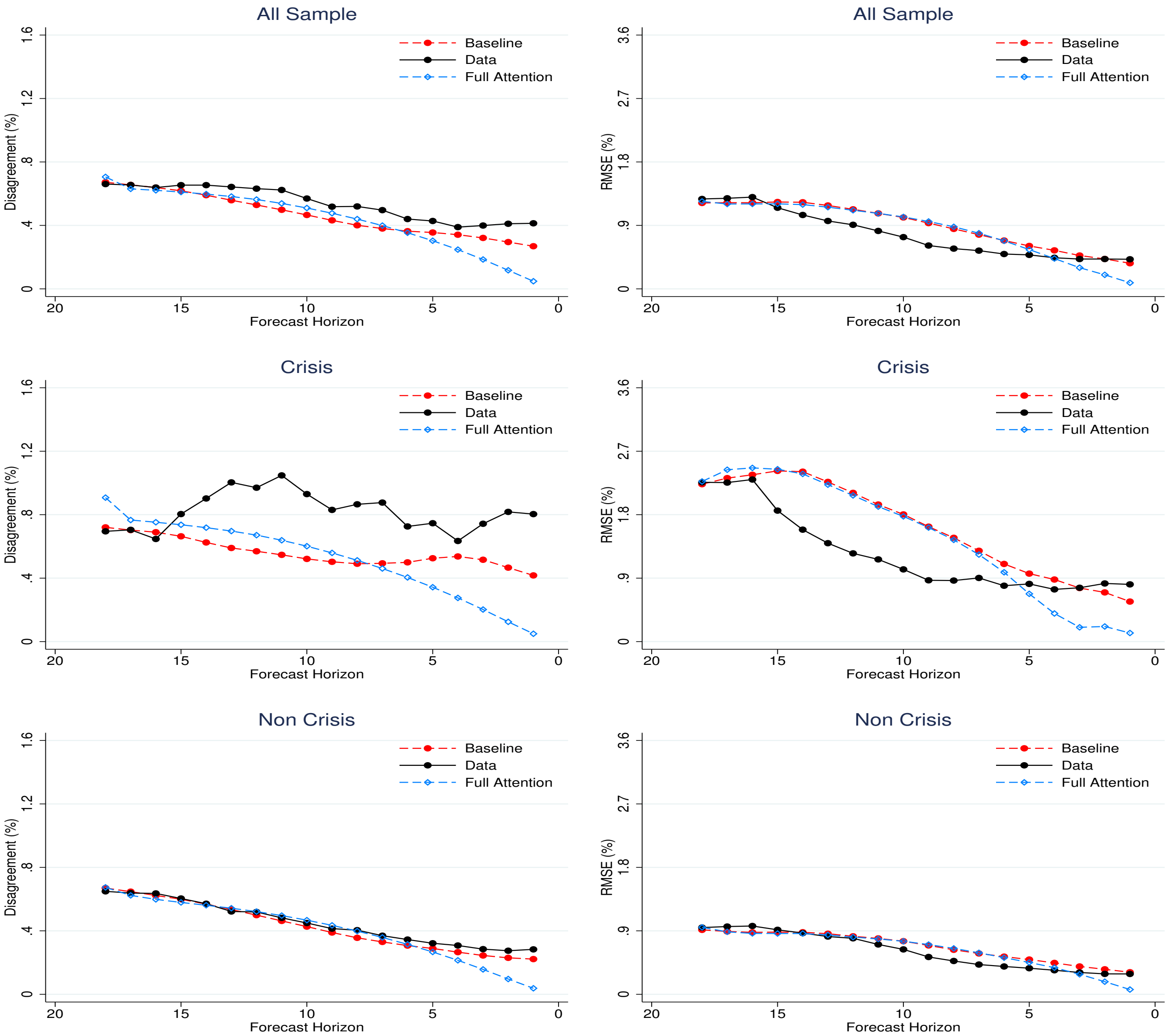


Figure 13: **Full Attention:** Disagreement and Accuracy

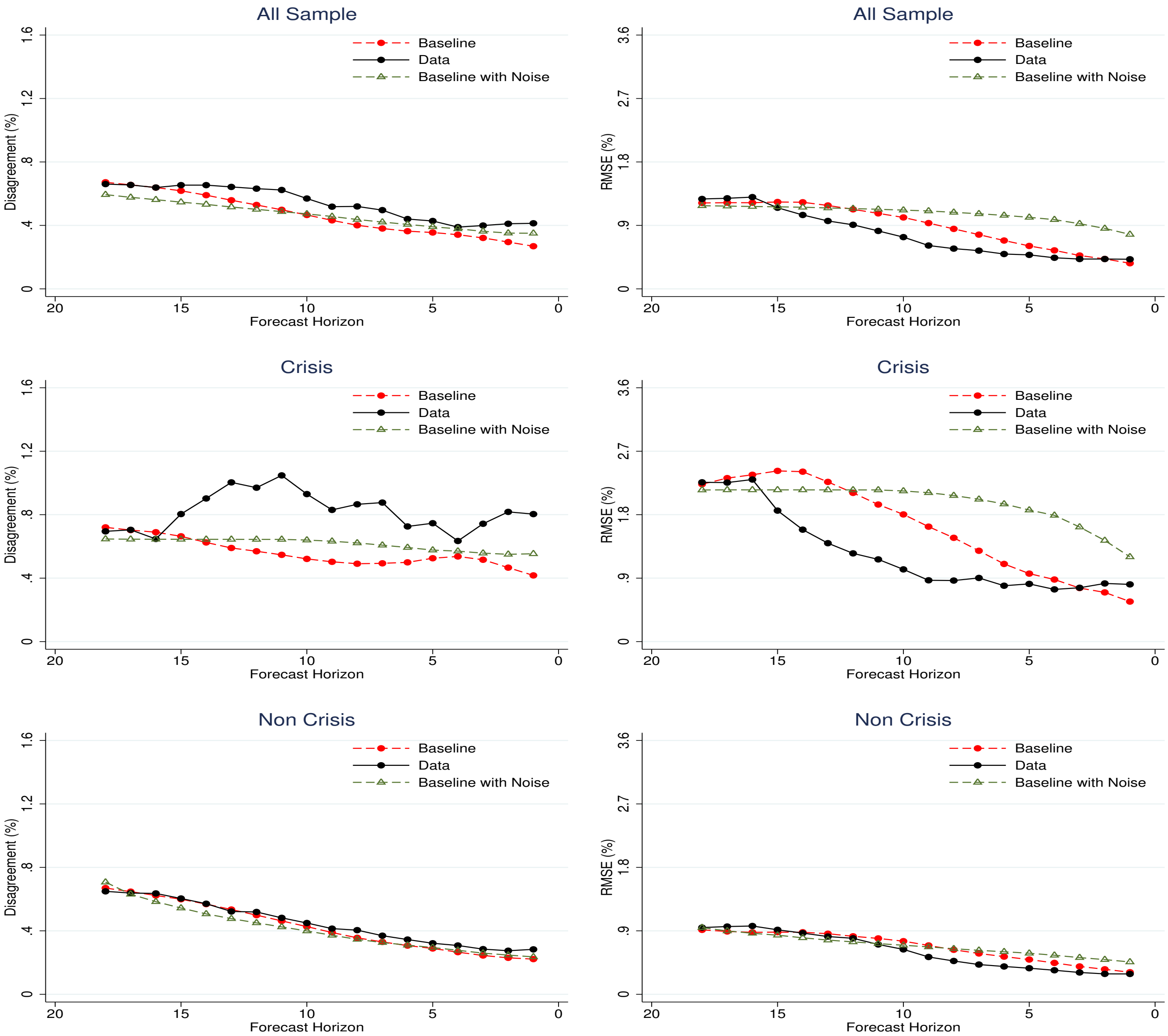


Figure 14: **Baseline with Noise:** Disagreement and Accuracy

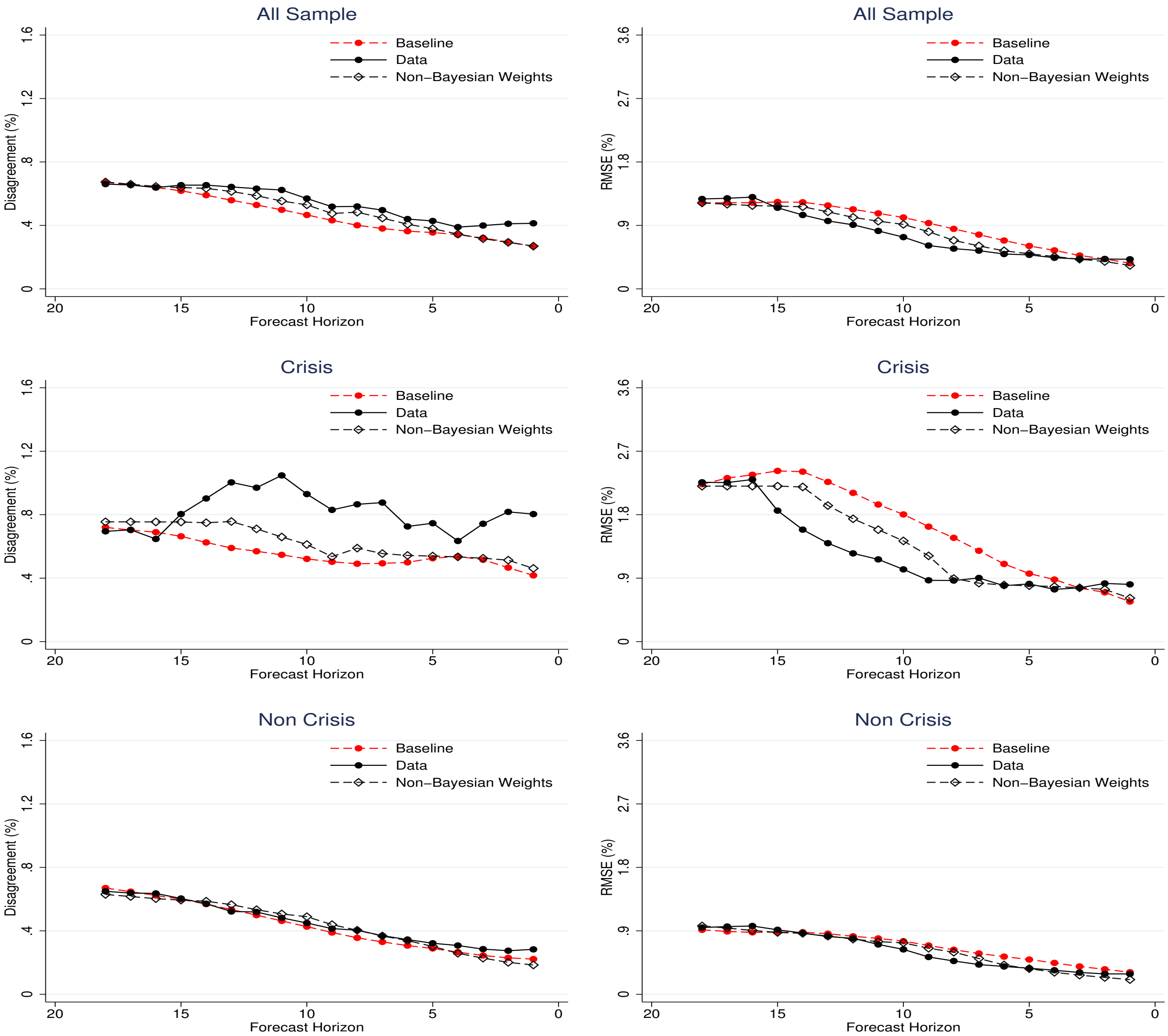


Figure 15: **Data-Driven Weights:** Disagreement and Accuracy

C.2 The Fit of Alternative Theories

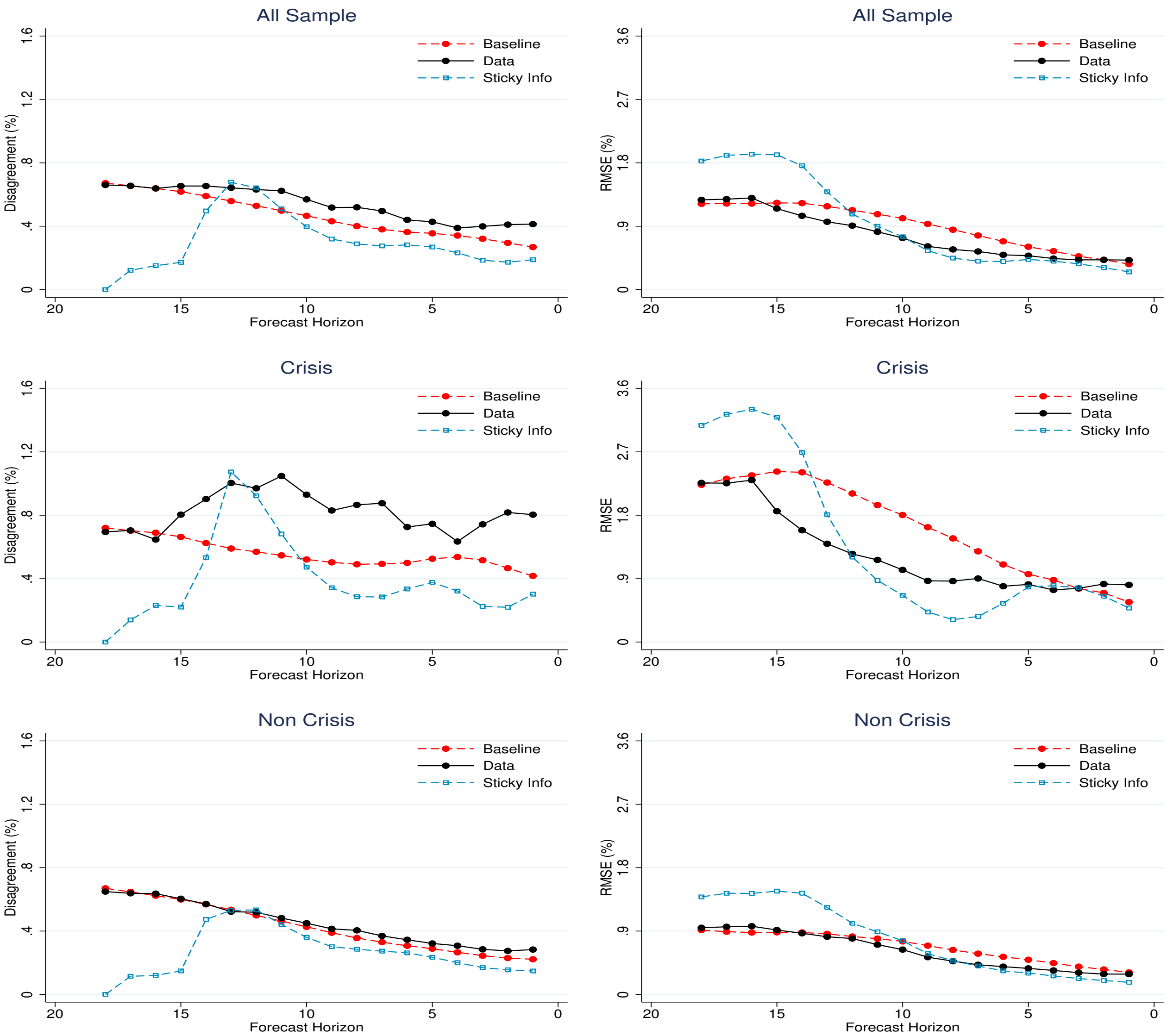


Figure 16: **Sticky Information:** Disagreement and Accuracy

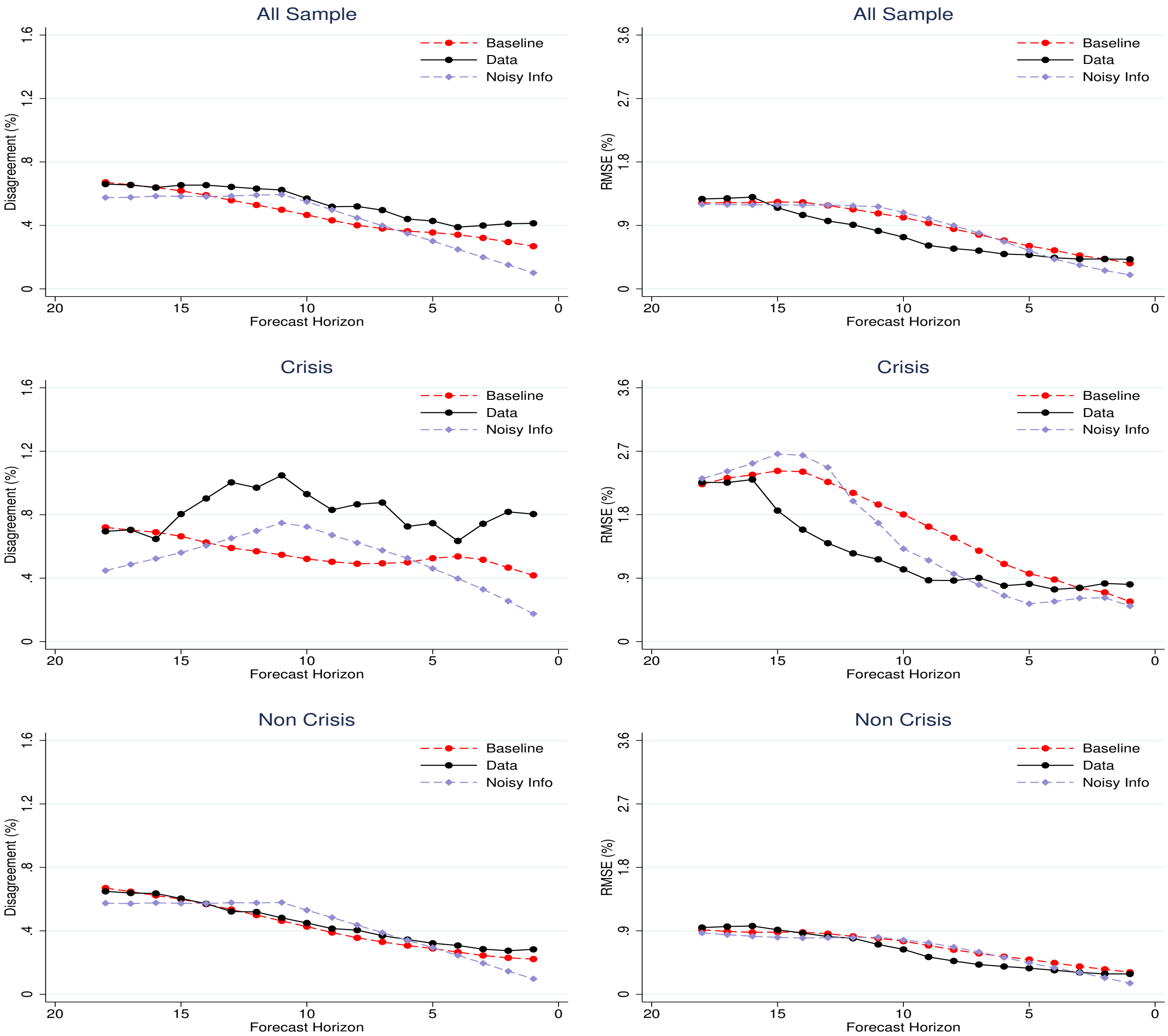


Figure 17: **Noisy Information:** Disagreement and Accuracy

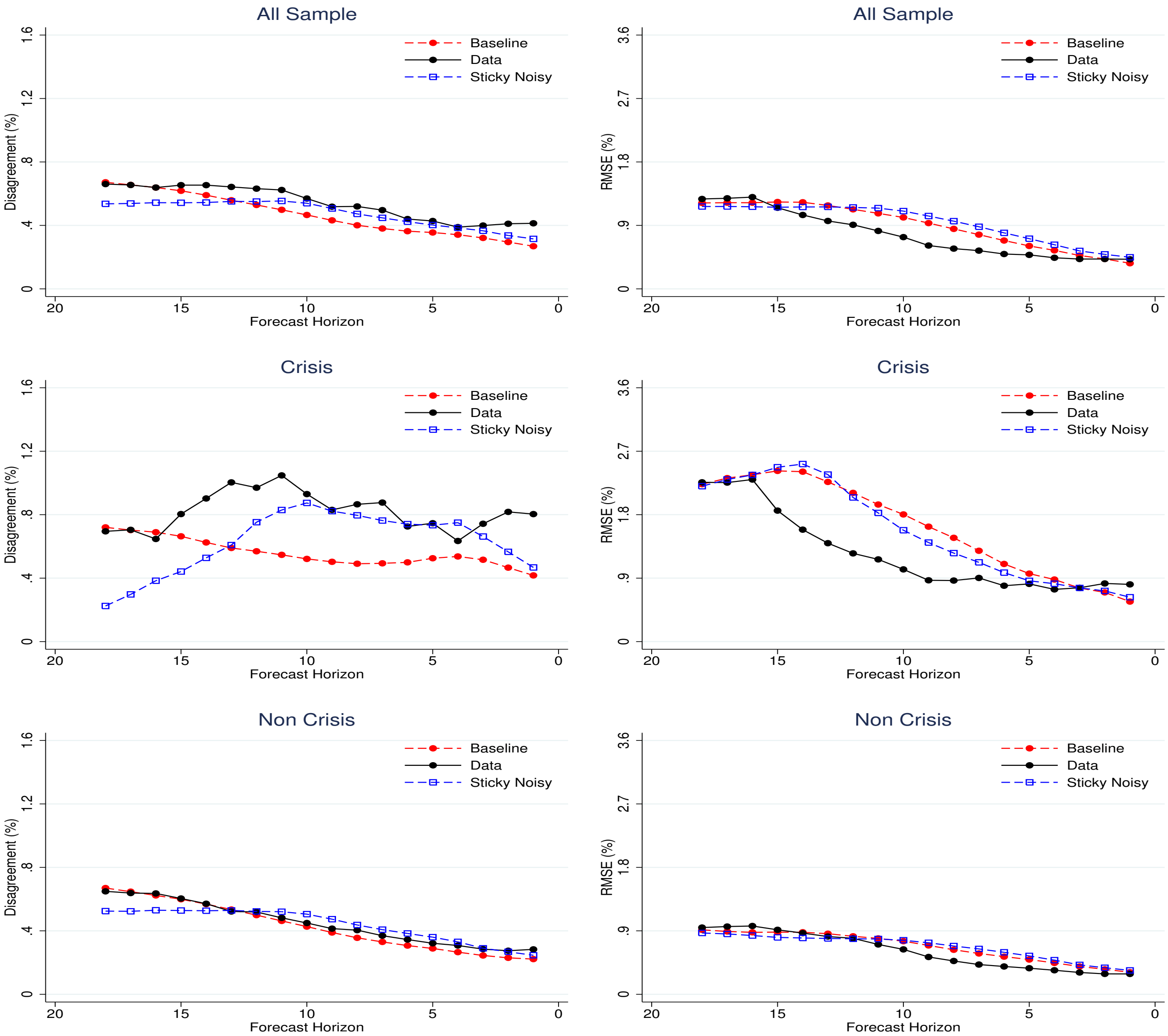


Figure 18: **Sticky Noisy:** Disagreement and Accuracy

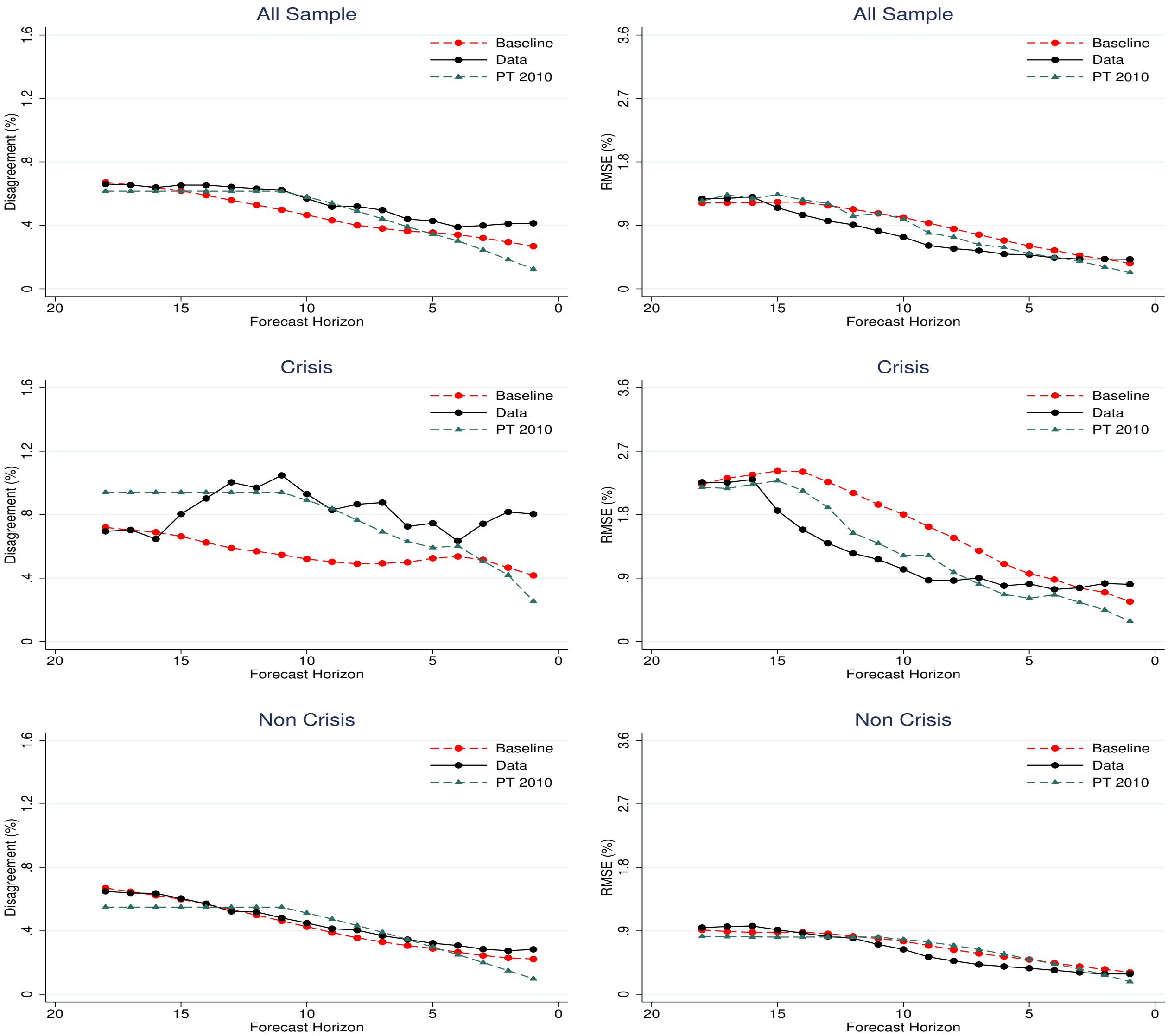


Figure 19: **PT 2010:** Disagreement and Accuracy

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, P., R. Crump, S. Eusepi, and E. Moench, "Fundamental disagreement," Working papers 524, Banque de France 2014.

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.

Bassetto, Marco, Todd Messer, and Christine Ostrowski, "Forecasting inflation and the Great Recession," *Economic Perspectives*, 2013, (Q III), 79–106.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, 1992, pp. 992–1026.

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.

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.

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

Kydland, Finn E and Edward C Prescott, "Time to build and aggregate fluctuations," *Econometrica: Journal of the Econometric Society*, 1982, pp. 1345–1370.

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

Lucas, Robert E, "Expectations and the Neutrality of Money," *Journal of economic theory*, 1972, *4* (2), 103–124.

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

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.

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.

_ **and** _ , "Predictability of Output Growth and Inflation: A Multi-Horizon Survey Approach," *Journal of Business & Economic Statistics*, 2011, *29* (3).

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

Reis, Ricardo, "Inattentive producers," *The Review of Economic Studies*, 2006, *73* (3), 793–821.

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.