Incentive-driven Inattention*

Wagner Piazza GAGLIANONE † Raffaella GIACOMINI ‡
João Victor ISSLER § Vasiliki SKRETA ¶

June 5, 2020

Abstract

"Rational inattention" is becoming increasingly prominent in economic modeling, but there is little empirical evidence for its central premise–that the choice of attention results from a cost-benefit optimization. Observational data typically do not allow researchers to infer attention choices from observables. We fill this gap in the literature by exploiting a unique dataset of professional forecasters who update their inflation forecasts at days of their choice. In the data we observe how many forecasters update (extensive margin of updating), the magnitude of the update (intensive margin), and the objective of optimization (forecast accuracy). There are also "shifters" in incentives: A contest that increases the benefit of accurate forecasting, and the release of official data that reduces the cost of processing information. These features allow us to link observables to attention and incentive parameters. We structurally estimate a model where the decision to update and the magnitude of the update are endogenous and the latter is the outcome of a rational inattention optimization. The empirical findings provide support for the key implication of rational inattention that information-processing efforts react to changing incentives. Counterfactuals reveal that accuracy is maximized if the contest date coincides with the release of information, aligning higher benefits with lower costs of attention.

Keywords: Rational Inattention; Contest; Incentives; Structural Estimation; Survey Design.

JEL Classification: E27, E37, D80, D83.

---

*We are grateful to two anonymous referees and to the editor for excellent comments and suggestions. Many thanks to Javier Turen and to Francisco Lima for helpful research assistance and to the Investor Relations and Special Studies Department (Gerin) of the Banco Central do Brasil (BCB) for kindly providing the data used in this paper. We thank Oli Coibion, Philipp Kircher, Imran Rasul, Colin Stewart, Nico Voigtländer, and participants at various workshops and seminars for useful comments and suggestions. We also acknowledge the support from CeMMAP, CNPq, FAPERJ, INCT and FGV on different grants. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Banco Central do Brasil.

†Research Department, Banco Central do Brasil. Email: wagner.gaglianone@bcb.gov.br

‡University College London, Department of Economics. Email: r.giacomini@ucl.ac.uk

§Graduate School of Economics and Finance – FGV EPGE, Getulio Vargas Foundation. Email: joao.issler@fgv.br

¶UT Austin, University College London & CEPR. Email: vskreta@gmail.com

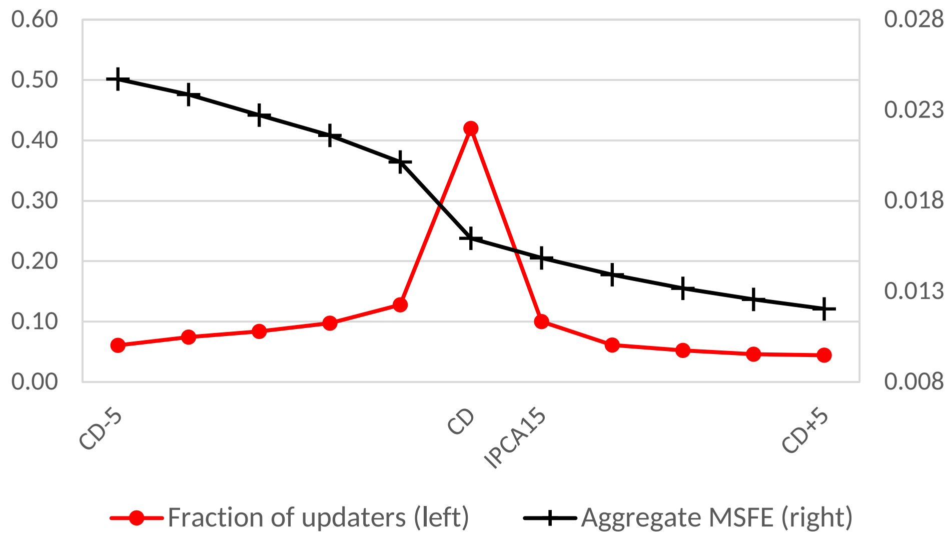
1 Introduction

Rational inattention (Sims, 2003) is the leading theoretical framework for endogenizing information frictions in economic models. Its central premise is that agents allocate their limited attention budget optimally when making decisions. A fast-growing theoretical literature models attention as an optimal cost-benefit decision in contexts ranging from price-setting (Mackowiak and Wiederholt, 2009) and investment choice (Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2016), to dynamic stochastic general equilibrium models (Maćkowiak and Wiederholt, 2015; Afrouzi, 2019). Yet, there is little empirical evidence that the choice of attention results from a cost-benefit optimization. One reason for the lack of evidence is the difficulty of observing attention choices, as well as the costs and benefits (i.e., the incentives) in observational data. This paper fills the gap in the literature by exploiting a unique dataset that enables us to identify the incentive parameters from observable decisions. We document new stylized facts and structurally estimate a rational inattention model that can explain the patterns in the data. Our results lend empirical credence to the emerging consensus on modeling inattention as an incentive-driven optimal decision.

The panel dataset we study is the little-known *Focus Survey* of professional forecasters maintained by the Central Bank of Brazil. We observe the decisions of participants to update their forecasts of current month's inflation (nowcasts) in Brazil's consumer price index, the IPCA. The data is unique for several reasons. First, forecasters can update any time they choose, so we can study both the decision to update and the magnitude of the update as endogenous decisions.[1] This allows us to study variations in both the extensive margin of updating (how many people update) and in the intensive margin (by how much they update), which matters for policy and survey design. Second, the objective function of the optimizing updating decision (the forecast accuracy) is observable in the data. This allows us to link the observables to the underlying incentive parameters, through optimal attention choices. Third, in addition to the informal incentives that it shares with other surveys analyzed in the literature, the survey has two "shifters" in the benefits and costs of processing information. The benefit

[1]This is in contrast to other common surveys, where forecasters are sampled at exogenously determined and infrequent times–e.g. monthly in the Consensus Forecasters or black Chip Analysts or quarterly in the Survey of Professional Forecasters.

Figure 1. Contest versus Information Release



Notes: The figure shows the fraction of forecasters in the *Focus Survey* who update their nowcast of inflation on a five-day window around the contest (CD) and the information release (IPCA15) days, averaged over all months in the dataset. It also shows the aggregate MSFE, which is the average across forecasters of the individual Mean Squared Forecast Errors. The individual MSFE is the squared difference between the nowcast associated with each forecaster on that day and the realization of inflation for that month, averaged over all months. Accuracy is the negative of the MSFE.

shifter is a monthly contest that ranks participants based on the accuracy of their (most recent) forecast on a specific day. The cost shifter is the release of official information about inflation, the IPCA15 inflation, which occurs the day after the contest and which largely overlaps with the variable that agents seek to forecast.[2] These incentive shifters allow us to identify the cost and benefit parameters from shifts in the observables.

Figure 1 illustrates a striking empirical pattern in the raw data. On the contest day we see a large increase in both the fraction of updaters (from about 10% to 42%) and in the aggregate accuracy improvements (a sharp fall in the aggregate mean-squared forecast error, henceforth MSFE). In contrast, the information release on the day after the contest appears to have no aggregate effect: the fraction of updaters and the aggregate accuracy improvements on the IPCA15 day are similar to those on any other non-contest day. Panel regressions confirm that the contest is the strongest driver of the decision to update, and that it also improves updaters'

[2]The IPCA15 measures inflation from the 15th of the previous month to the 15th of the current month, whereas the IPCA measures inflation between the first and the last day of the current month.

accuracy. This suggests that the gain in aggregate accuracy on the contest day in Figure 1 is due to changes in both extensive and intensive margins of updating. The structural model we develop allows us to quantify their relative contribution. A possible reason why the information release has no visible effect in Figure 1 is that almost no agents who update on the contest update again the day of information release, so they are not exploiting the additional accuracy improvement that would be available on that day. This raises the question whether the contest mis-aligns updates that would have otherwise occurred the day of information release. Our structural model can be used to investigate this and other questions related to survey design.

Motivated by the empirical findings,[3] we develop a decision-theoretic model of rational inattention where agents have limited resources to allocate to produce an accuracy-maximizing forecast. The baseline model assumes that each month, agents use a realistic statistical model (an Autoregressive Moving Average–ARMA) to produce an initial forecast. On each subsequent day, they make two decisions: 1) whether to update their forecast and 2) how much attention to allocate to process information in order to update. The first decision is driven by a fixed cost reflecting the *opportunity cost* of time devoted to forecast updating. The second decision is the result of a rational inattention optimization problem that depends on the information available to the agent and on the information-processing cost.

To clarify the role of fixed costs–which lead to "stickiness" in the sense of Mankiw and Reis (2002) – versus information-processing costs–which lead to "noisyness" in the sense of Sims (2003) and Mackowiak and Wiederholt (2009)–we introduce three variations of the baseline model. Variation 1 is a pure "sticky" model that switches off the attention decision and assumes that there is only a fixed cost of updating, but otherwise agents get perfect signals. Variation 2 is a pure "noisy" model in which agents only decide how much attention to allocate and updaters are agents who exert positive effort. Variation 3–the "restricted baseline" model–has both decisions that are now *linked* through a common parameter, namely the marginal benefit of increasing accuracy. The restricted baseline model has the plausible implication that higher stakes are associated with both a larger revision and a larger probability of updating.

There are two dynamic dimensions in our setting: the month-to-month problem of forecast-

[3] In addition, based on the stylized facts reported in Section 4, we rule out strategic behaviour and self-selection based on ability on the contest day.

ing inflation at the beginning of the month, and the within-month problem of updating the forecast. We use results from aggregation of ARMA processes to make the statistical model for the initial forecast compatible with the statistical model used for the update. The statistical model allows us to decompose the forecast accuracy into a component that depends on the resolution of uncertainty as the forecast horizon decreases and a component that depends on attention. We can then obtain optimal attention as a function of the forecast horizon and the incentive parameters. The baseline model and its variations yield simple and intuitive analytical expressions for the theoretical counterparts of the observables in the data–the fraction of updaters and the accuracy of the update–as functions of the model's parameters.

We structurally estimate all variations of the model by Simulated Method of Moments, in order to match the *joint* dynamics of updating frequencies and aggregate accuracy reported in Figure 1. For model validation, we consider the J-test of overidentifying restrictions and compare the estimates of the ARMA parameters–which are not matched in the estimation–to those estimated on Brazilian data.

Both the sticky model and the noisy model are rejected. This shows that, indeed, we need a model with both fixed costs and information-processing costs to fit the data. The baseline model passes the J-test but has the drawback that it is heavily parameterized and its parameters are only identified up to scale. The restricted baseline model passes the J-test, is parsimonious and its parameters are identified. The external validity of the restricted baseline model is further reinforced by the fact that the estimates of the ARMA parameters are almost identical to those implied by Brazil's inflation data.

Overall, the estimation results provide support for the key implication of rational inattention that agents increase their information processing efforts when the incentives of doing so increase. First, the estimates of the models that fit the data imply that agents increase their information-processing efforts on the contest day, when the incentives of producing an accurate forecast are higher. Second, none of the models fits the data when assuming that agents do not increase attention in response to the contest.

We use the estimated restricted baseline model to perform counterfactuals. First, we quantify how changes in aggregate accuracy are affected by changes in the extensive and intensive margins. We find that 70% of the accuracy improvement on the contest day that is visible in Figure 1 is due to more agents updating (the extensive margin) and 30% to agents paying more attention (the intensive margin).

We further perform counterfactuals to investigate alternative survey designs. First, we find that holding the contest on any given day of the month would result in an accuracy improvement from the previous day that is 3 to 4 times larger than it would be without the contest. Second, we show that the optimal contest day is the IPCA15 day. On this day the increase in benefits from the contest is amplified by the availability of low-cost information. Finally, we investigate the extent to which the contest mis-aligns updates from the more "natural" IPCA15 day. We find that without the contest average accuracy is worse, even though most updates happen on the IPCA15 day. This underscores that the coordinated updates that occur because of the contest are crucial for the survey's aggregate accuracy.

Our models rely on two main simplifying assumptions, which lead to elegant and easily interpretable analytical expressions for the key counterparts of the observables in the data. The first is that we separate the decision of whether to update from the decision of how much attention to allocate to processing information. We model the latter as a problem of rational inattention, while the former boils down to comparing the opportunity cost of time to a fixed threshold.[4] In the restricted baseline model, both decisions are endogeneous, as they depend on the same incentive parameters. The second assumption concerns the information available to agents who update. We assume that updaters have access to *past* public signals that are more accurate than any past private signals. This reduces the dynamic rational inattention problem to a sequence of static problems: An agent only needs to process information for the current day, which depends on the current benefit and cost parameters.[5] One could easily relax these assumptions and obtain numerical solutions. However, this would sacrifice analytical

[4]We believe this does not diminish the contribution of the paper in terms of providing empirical credence to models of rational inattention, as the same conclusion would have emerged by focusing only on the second decision. As discussed by Woodford (2009), applying rational inattention to a timing decision like the first decision here would present additional challenges, so attempting to solve a joint rational inattention optimization for the two decisions would substantially complicate the analysis without necessarily adding new insights. See Section 5.3 for further discussion.

[5]In the likely presence of both public and private information, we argue that this assumption is more realistic than the opposite extreme assumption that updaters only rely on their past, less accurate, private signals. Relaxing the assumption would imply that the accuracy of updaters depends on past updating decisions. We show that this prediction is not supported by the data. See Section 5.3 for further discussion.

interpretability while, at the same time, require additional assumptions that are not necessarily supported by the data.

The paper is organized as follows. Section 2 relates the paper to the literature; Section 3 discusses the data; Section 4 presents the empirical findings. Section 5 presents the models; Section 6 discusses the key assumptions; Section 7 presents the structural estimation results; and Section 8 the counterfactual analysis. Section 9 concludes.

2 Related Literature and Contribution

This paper makes several contributions to the literature.

We follow the recent theoretical literature on rational inattention (Maćkowiak and Wiederholt, 2015) in endogenizing the attention choice, as opposed to assuming that the bound is exogenous as in, e.g., Mackowiak and Wiederholt (2009), Caplin and Dean (2015), and Steiner, Stewart, and Matějka (2017). By exploiting a unique dataset to show that the choice of attention is driven by cost-benefit considerations, we provide an empirical foundation for this increasingly prominent theoretical mechanism in models of rational inattention.

This paper brings a rational inattention model to the data, contributing to a literature still in its infancy. The typical approach to validation of rational inattention models involves calibrating and deriving testable implications of the model. We go beyond the existing literature by structurally estimating the model. We can do this because our data allows us to overcome some of the challenges typically encountered in validating these models using observational data.[6] One such key challenge is the difficulty in separately identifying the unobservable attention from the (usually also unobservable) prior uncertainty. Caplin, Leahy, and Matejka (2016) and Csaba (2018) make important steps towards overcoming this challenge in the context of discrete choice analysis. In our data we can separately identify the two components.

We consider expectations data, similarly to the literature studying the role of information

[6]There are important experimental studies, however, including Cheremukhin, Popova, Tutino et al. (2011), Caplin and Dean (2013), Dean and Neligh (2017), Martin (2016) and Cavallo, Cruces, and Perez-Truglia (2017).

frictions in expectation formation (see the survey of Woodford, 2013). Coibion and Gorodnichenko (2012) provide evidence of such frictions in expectations data, and Coibion and Gorodnichenko (2015) find empirical support for some predictions of models of exogenous information frictions. The theoretical part of this paper takes a step forward by endogenizing the information frictions that are assumed exogenous in this literature. One of the empirical findings that we document is that the measure of information frictions in Coibion and Gorodnichenko (2015) depends on the incentives. This has more general implications for models of information rigidities: For example, in a model of agents who take actions infrequently (firms updating prices or households buying houses), it suggests that one should calibrate information rigidity of home buyers or price changers to information rigidity conditional on high incentives as opposed to the information rigidity of the average agent.[7]

Fuhrer (2018) finds evidence of "irrationality" in surveys of professional forecasters as forecast errors are correlated with previously available information. While our analysis is not directly comparable, as our focus is on analyzing attention changes in response to changing incentives, we conduct Mincer and Zarnowitz (1969)'s forecast rationality tests (reported in Table A.3 in Appendix A). We find little evidence of irrationality in our data, once considering forecasts produced on the contest day or on the day of information release and once computing the consensus forecasts only for updaters.

Our findings show that a forecasting contest improves accuracy. Marinovic, Ottaviani, and Sørensen (2013) study theoretically the effect of a forecasting contest in a strategic model. They find that the effect of the contest on forecast accuracy can be ambiguous. Forecasters in our dataset seem to ignore the strategic component and focus on overcoming the information barriers in order to provide accurate forecasts.[8]

The link that we document between attention and incentives speaks to the broader question of what are the productivity drivers in economics (see the survey of Syverson, 2011)). Lazear (2000) studies the effect of monetary incentives on output, while Shearer (2004) shows how the structure of compensation contracts affects productivity using data from a field experiment.

[7]We thank an anonymous referee for suggesting this analysis and interpretation of the findings.

[8]This was also confirmed in personal interviews with some participants. One reason why participants seem not to act strategically may be that the survey is confidential and anonymous.

In the psychology literature, Reeve, Olson, and Cole (1985) consider the role of incentives and competition in motivation and performance. In the same spirit, Glaeser, Hillis, Kominers, and Luca (2016) argue that tournaments can be a cost-effective tool to outsource public services. Recent work, see DellaVigna and Pope (2018), investigates determinants of effort on forecasts in a field experiment. The detailed role of various behavioral biases and how they interact with the economic environment is surveyed in Gabaix (2019). Viewed from this broader perspective, our study contributes to establish a clear link between incentives and performance.[9]

3 Data

Our panel data are from the Central Bank of Brazil's (BCB) survey of professional forecasters, the *Focus Survey.* We study forecasts–nowcasts–of current month's inflation in the consumer price index (IPCA), which is the official inflation measure and the target of monetary policy at the BCB.[10] The panel includes all forecasters who provide forecasts that are confirmed or updated within 30 days from the first forecast considered. The panel is unbalanced since not all forecasters participate each month and the number of participants is generally increasing over time. It consists of forecasts for a given month that each participant can produce every working day of the month, starting from January 8th, 2004 to January 8th, 2015, amounting to a total of 2,751 daily forecasts for 132 months, with an average of 85.3 forecasters.[11] We treat months and the forecasts associated with them as events. Each event entails a decreasing-horizon forecasting problem. The events are connected by inflation, which is a continuous process over the whole sample. Table A.1 in the Appendix reports summary statistics. Forecasts are of similar average magnitude as actual inflation. Forecasters, on average, revise slightly upwards. The standard deviation of revisions is 0.09, which is 38% of the standard deviation of inflation, indicating that revisions are of low-to-moderate magnitude.

[9]Note that in our setting the effects of the incentive are observable and easily measurable compared to other settings where the effects of workers' effort on output are confounded by other factors. Moreover, the stark variation in incentives from one day to another is unique to our data.

[10]The inflation index measures the change in prices of a fixed set of goods and services. The price research is done daily, covering thirteen cities in Brazil. Inflation is closely monitored by economic agents in Brazil for tracking monetary policy and also because it is used to index treasury bonds, in wage negotiations, and as an adjustment for certain contractually regulated prices.

[11]We start the sample in 2004 because there were too few participants prior to this year.

Forecast Updates: The BCB provides forecasters with a software (the *Market Expectation System*) that they can access any time. Any time a forecaster logs in the system, she can change a forecast or confirm it. For forecasters who do not log in, the system copies the previous forecasts. We say that a forecast is *updated* if the forecaster changed the forecast. We do not consider the negligible number of cases where forecasters logged in and confirmed their previous forecast. Because we equate forecast updating with a changed forecast, we cannot capture the possibility that forecasters process information but decide not to change the forecast. In this case, we may be over-estimating the *level* of informational rigidity, however this would not affect our analysis since its focus is on *changes* in informational rigidity.[12]

Incentives to Update Outside the Contest Day: As for the other surveys of professional forecasters typically considered in the literature (the "black Chip," the "Consensus" or the Fed's "Survey of Professional Forecasters"), we don't know what are the incentives for participation: All we know is that in the *Focus Survey* about 10% of forecasters update on any non-contest day (see Figure 1). We can speculate that this is due to several 'low-stakes" incentives built into the survey that encourage regular participation. First, every Monday the BCB publishes the highly visible in the media "Focus-Market Readout."[13] The readout only considers forecasts that were updated during the previous thirty days. Second, forecasters who are inactive for more than thirty days are removed from the system. Those who remain inactive for six months are blocked from the system and are not considered for the contest, and need to formally request a renewal if they wish to resume participation. Third, some of the active participants are invited to BCB meetings to provide opinions about the economic outlook.

The Contest: The contest represents a clear "high-stakes" incentive. Every month, upon the release of the realization of the variable, the forecasters are ranked based on the accuracy

[12]Our assumption can be considered realistic if forecasts are based on estimated models. Consider, for example, the literature on nowcasting in macroeconomics (e.g., Giannone, Reichlin, and Small (2008)) which involves updating forecasts to incorporate new information as soon as it becomes available. In this approach the arrival of new information always results in a new forecast. If the forecasts are not based on models and/or involve judgment, then it is possible that a forecaster may decide not to update even after processing new information, but we cannot identify this occurrence in our data.

[13]The readout reports key aggregate statistics from the *Focus Survey* based on data collected at 5 PM of the previous Friday. See Marques (2013) for further details.

of the forecast that was on the Market Expectation System on the pre-announced day of the previous month, the *contest day*. The names of the five most accurate forecasters (institutions) according to the absolute forecast error are published on the BCB website. The contest dates are announced before the beginning of each calendar year. The contest is highly valued by the survey participants and the top-five forecasting institutions usually advertize their contest accomplishments on their websites or advertising material.[14] Figure 2 shows as an example the outcome of the monthly contest for February 2017.[15]

Information Releases: The main information release is the monthly release of IPCA15 inflation, which measures inflation between the 15th of the current month and the 15th of the previous month.[16] The date of release of the IPCA15 changes from month to month, but it is always the day after the contest. In the panel regressions, we further consider two other types of information releases that are not directly linked to inflation: the date release of the minutes of the meeting of the BCB Monetary Policy Committee (MPC), which occurs less frequently and at irregular times, and the day of release of the monthly employment survey (PME), published by The Brazilian Institute of Geography and Statistics.

Forecast Timeline: Forecasters know the dates of the contest and data releases in advance. There are no survey-related activities on holidays. The contest and all data releases happen on workdays. Thus, only workdays are considered throughout our analysis. The number of workdays in the month (i.e., the duration of the forecasting period) and the timing of the contest can vary across months. The chronology of relevant events within a representative month is depicted in Figure 3: the first forecast for February's inflation can be given on the day of release of the IPCA for January, which occurs most often on the 8th of February. The

[14] There is anecdotal evidence (confirmed to us by *Focus* participants) that winning institutions reward successful forecasters with bonuses. TOP 5 forecasters have their name on the Central Bank homepage, refreshed every month (see https://www.bcb.gov.br/en/monetarypolicy/focustop5monthly), and, every year, the best 5 forecasters receive a plaque and overall market recognition in the most prestigious Central Bank conference of that year (i.e., the "Focus Survey Top Five Award" at *The Annual Inflation Targeting Conference* of the BCB, see https://www.bcb.gov.br/en/about/events/1).

[15] See http://www4.bcb.gov.br/pec/gci/ingl/focus/top5.asp for further details about the contest.

[16] The components of the IPCA15 index are released on the same day and there are no other releases of information about the components of IPCA15 during the month. There are other consumer inflation indexes released by other institutions on different days, but they consider different goods and weights.

Figure 2. Example of Contest Outcome

BANCO CENTRAL DO BRASIL
Gerin

Top 5 Forecasting Institutions - February 2017

March 10, 2017

The Investor Relations and Special Studies Department (Gerin) has announced the Top 5 forecasting institutions for February 2017.

Table 1
Top 5 Forecasting Institutions - Short-Run
February 2017

	IPCA	Deviation
1	Flag Gestora de Recursos	0.0717
1	Petros Fundação de Seguridade Social -	0.0717
3	Quantitas Asset Management	0.0833
4	Banco Bradesco S.A.	0.0852
5	ICAP Brasil	0.0867

	IGP-DI	Deviation
1	Banco Itaú S.A.	0.0883
2	BBM Investimentos	0.1217
2	SPX Capital	0.1217
4	Haitong Banco de Investimento do Brasil	0.1317
5	J. Safra Asset Management	0.1350
5	Verde Asset Management	0.1350

	IGP-M	Deviation
1	LCA Consultores S/C Ltda.	0.0683
2	Haitong Banco de Investimento do Brasil	0.0733
3	Icatu Vanguarda Administração de Recursos	0.0833
4	Banco Itaú S.A.	0.0883
5	Banco Fibra S.A.	0.0983

	Exchange Rate	Deviation
1	Telefônica / Vivo	0.0691
2	Rosenberg & Associados S/C Ltda.	0.0693
3	BB DTVM S.A.	0.0741
4	Tendências Consultoria Integrada	0.0746
5	Banco do Brasil S.A.	0.0751

	Over Selic	Deviation
1	Barclays Capital	0.0417
1	Bozano Gestão de Recursos	0.0417
1	CSHG Gauss	0.0417
1	M. Safra	0.0417
5	Banco do Brasil S.A.	0.0625
5	Banco Itaú S.A.	0.0625
5	Banco Original do Agronegócio	0.0625
5	Brasilprev Seguros e Previdência S.A.	0.0625
5	BW Gestão de Investimentos Ltda.	0.0625
5	Caixa Asset	0.0625
5	Daiwa Asset Management	0.0625
5	Deutsche Bank - Banco Alemão S.A.	0.0625
5	Fapes - BNDES	0.0625
5	Flag Gestora de Recursos	0.0625
5	Ibiuna Investimentos Ltda.	0.0625
5	Icatu Vanguarda Administração de Recursos	0.0625
5	Kondor Admin. e Gest. de Rec. Financ. Ltda.	0.0625
5	MCM Consultores	0.0625
5	PREVI Caixa Previd Funci Banco Brasil	0.0625
5	Quantitas Asset Management	0.0625
5	Quest Investimentos Ltda.	0.0625
5	Santander Asset Management	0.0625
5	Sul America Investimentos	0.0625
5	Vintage Investimentos	0.0625

Information for unrestricted disclosure. It is not intended to bind Banco Central do Brasil in its monetary or foreign exchange policy actions. Questions and comments to gerin@bcb.gov.br

contest most often takes place on the 21st of February which is always the day *before* the release of IPCA15 inflation (measuring inflation between the 15th of February and 15th of January). Forecasters can provide a new forecast on each working day between the 8th of February and the day of the release of IPCA for February–(most often) the 8th of March.

Survey Participants: Participants include non-financial institutions, commercial banks, asset-management firms and consulting firms.

Figure 3. Example of Forecast Timeline

forecast begins	contest day	IPCA15 released	IPCA released and used to evaluate forecasts
8 Feb	21 Feb	22 Feb	8 March

Confidentiality and Anonymity: The data are proprietary and the identity of the forecasters is not known to us nor is it revealed to the public or to other survey participants, except for the winners of the contest (see Figure 2)[17]. Survey participants do not observe the forecasts of other institutions. Besides publishing the names and accuracy of the top-five contest winners once a month, the BCB only releases summary statistics of the forecasts (the median, mean, standard deviation, coefficient of variation, maximum and minimum forecasts).

4 Stylized Facts

In this section we document new stylized facts about the drivers of forecast updates and accuracy improvements. Summary statistics contained in Table A.2 in Appendix A show that, on average, 42.0% of forecasters update on the contest day vs. 7.5% on non-contest days. The average forecast error is smaller on the contest day.

We next formally analyze updating and accuracy at the *individual* level, by means of panel regressions. We then consider the *aggregate* dynamic behavior of updates and accuracy around the contest day and around days of information releases.

Drivers of the Decision to Update: We first consider the extensive margin of updating (how many forecasters update) by estimating a panel logit model for forecast updates:

$$\Pr\left(z_{it} = 1 \,|x_{it}\right) = G\left(\alpha_i + x_{it}'\beta\right), \tag{1}$$

[17]There is no record so far of top-five survey participants who refused to have their name disclosed.

where G is the logistic function and

$$z_{it} = \begin{cases} 1 & \text{if forecaster } i \text{ updates on day } t \\ 0 & \text{otherwise.} \end{cases}$$

The regressors x_{it} include dummy variables for the day of the contest (d_t^{CD}), the day of release of the IPCA15 (d_t^{IPCA15}), the day before the contest (d_t^{CD-1}), the day of release of the MPC minutes (d_t^{MPC}) and the day of release of the monthly employment survey (d_t^{PME}). We also include an interaction dummy ($d_t^{CD} \times d_t^{\text{beforeCD}}$), where d_t^{beforeCD} equals 1 if the forecaster updates on any of the five days before the contest. Other regressors are dummy variables for Mondays and Fridays and the $EMBI_{t-1}$, the Emerging Markets Bond Index Plus for Brazil (EMBI+BR)–a measure of uncertainty on the previous day.[18]

Table 1 reports the coefficient estimates and the average marginal effects (in square brackets). The table shows that the contest is not the only driver of updates, but it has the largest effect on the extensive margin of updating (how many people update): The probability of updating goes up by 38.9 percentage points (p.p.) on the contest.[19] There is also a "contest anticipation" effect with a 18.8 p.p. increase in the probability of updating one day before the contest. The release of information has a smaller effect (the IPCA15 is associated with a 12.1 p.p. increase and the MPC with a 5.9 p.p. increase). From column (5), we see that the probability of updating on both the contest day and within the five days before is about 9 p.p. lower than the probability of updating only on the contest day (within this window). The Friday dummy is significant, which may reflect the importance of, in this case more informal, incentives in the survey, as summary statistics about the forecasts collected on Fridays are released on the following Monday as part of the Focus-Market Readout. Forecasters are more likely to update when there is higher uncertainty, as indicated by the coefficient for $EMBI_{t-1}$. While the magnitude of this effect is small, this finding is consistent with one of the main predictions of rational inattention models.

Column (3) in Table 1 adds as a regressor the day of release of the unemployment index

[18]The contest day coincides with the MPC date in 6% of the months and with the PME release day in 19% of the months, so there is no concern of near multicollinearity in this respect in the regression.

[19]We focus the discussion on marginal effects.

Table 1. Drivers of the Decision to Update

Regressors	**Logit Fixed Effect Coefficients**				
	(1)	(2)	(3)	(4)	(5)
d_t^{CD-1}	0.910***	0.889***	0.887***	0.869***	0.888***
	(0.031)	(0.031)	(0.031)	(0.032)	(0.031)
	[0.203]	[0.188]	[0.188]	[0.152]	[0.188]
d_t^{CD}	2.608***	2.715***	2.738***	2.673***	2.847***
	(0.023)	(0.024)	(0.024)	(0.037)	(0.028)
	[0.417]	[0.389]	[0.392]	[0.297]	[0.396]
d_t^{IPCA15}	0.538***	0.547***	0.574***	0.766***	0.545***
	(0.034)	(0.035)	(0.035)	(0.036)	(0.035)
	[0.125]	[0.121]	[0.127]	[0.137]	[0.120]
d_t^{MPC}	0.104**	0.259***	0.283***	0.269***	0.261***
	(0.042)	(0.043)	(0.043)	(0.044)	(0.043)
	[0.025]	[0.059]	[0.064]	[0.053]	[0.059]
$EMBI_{t-1}$	-	0.024**	0.026***	0.033***	0.024***
		(0.010)	(0.010)	(0.010)	(0.009)
		[0.006]	[0.006]	[0.007]	[0.005]
d_t^{MON}	-	0.382***	0.370***	0.404***	0.380***
		(0.021)	(0.021)	(0.022)	(0.021)
		[0.087]	[0.085]	[0.079]	[0.086]
d_t^{FRI}	-	0.494***	0.486***	0.484***	0.493***
		(0.022)	(0.022)	(0.022)	(0.021)
		[0.112]	[0.111]	[0.094]	[0.112]
d_t^{PME}	-	-	-0.167***	-	-
			(0.037)		
			[-0.039]		
$duration_{i,t}$	-	-	-	0.070***	-
				(0.002)	
				[0.014]	
$duration_{i,t} \times d_t^{CD}$	-	-	-	0.013***	-
				(0.004)	
				[0.003]	
$d_t^{CD} \times d_t^{\text{beforeCD}}$	-	-	-	-	$-.378$***
					(0.043)
					[-0.088]
Log likelihood	-55258.9	-52989.0	-52978.7	-50312.6	-52951.42

Notes: Model for the probability that a forecaster updates on day t. Sample from January 8th, 2004 to January 8th, 2015. Number of observations (model 1) = 228,157. Robust standard errors in parentheses. ***, ** and * indicate, respectively, significance at the 1%, 5% and 10% level. Average marginal effects are in square brackets.

(PME), to assess whether forecasters react to information releases that are not directly related to the variable they are forecasting (for example, because they are updating other forecasts). The table shows that there is a significant, small *decrease* in the probability of updating on PME release days, suggesting that forecasters respond to releases of information, but only if it is relevant for the variable they are forecasting.

Column (4) in Table 1 indicates a statistically significant but small effect of the duration between updates on the decision to update (captured by adding the regressor $duration_{i,t}$, which, for agent i, measures the number of days between her update on day t and her previous update). This is expected, since the survey encourages regular participation. The interaction between the duration and the contest dummy is also significant but its marginal effect is negligible, and the main conclusions of the regression are preserved. This suggests that we cannot attribute changes in updating between normal and contest days to systematic shifts in observable characteristics such as previous updating behaviour.

Drivers of Accuracy Improvements for Updaters: We next analyze the drivers of accuracy improvements, conditional on updating. While we expect to find that information releases improve accuracy, it is less clear a priori whether the contest would affect the accuracy of updaters. We show below that the contest not only induces more forecasters to update, but also makes them exert more "effort" into producing accurate forecasts. This provides supportive evidence in favour of a main prediction of rational inattention: that attention/effort increases when the incentives are higher.

Analyzing accuracy improvements is complicated by the fact that there are confounding factors that cause time variation in forecast accuracy and in the effects we want to investigate. First, the decreasing-horizon setting means that the forecast accuracy is expected to improve during the month and that the accuracy improvement is not necessarily constant. Second, the contest and IPCA15 days fall on different dates each month so they are also associated with different horizons from month to month. Finally, the accuracy improvement by construction depends on the time between updates. To partly control for these factors, we add the duration and the forecast horizon ($horizon_t$) as regressors.

Table 2. Drivers of Accuracy Improvements for Updaters

Regressors	**Panel Fixed Effect Coefficients**	
	(1)	(2)
d_t^{CD-1}	-3.694	-3.616
	(3.227)	(3.289)
d_t^{CD}	7.132**	6.334**
	(3.017)	(3.219)
d_t^{IPCA15}	33.656***	35.299***
	(4.839)	(5.002)
$d_t^{IPCA15+1}$	35.272***	33.549***
	(5.647)	(5.766)
d_t^{MPC}	20.660***	22.847***
	(5.585)	(5.536)
$duration_{i,t}$	2.423***	2.381***
	(0.197)	(0.198)
$horizon_t$	−0.490***	−0.490***
	(0.169)	(0.171)
$EMBI_{t-1}$	-	-0.995
		(0.835)
d_t^{MON}	-	4.797*
		(2.693)
d_t^{FRI}	-	3.968
		(2.679)
$constant$	26.015***	24.150***
	(2.375)	(2.613)

Notes: Dependent variable is minus the change in the log of the squared forecast error for an *updater* on day t, relative to the previous update. Sample from January 8th, 2004 to January 8th, 2015. Number of observations (model 1) = 26,911. Standard errors in parentheses. ***, ** and * indicate, respectively, significance at the 1%, 5% and 10% level.

We consider only observations for which agent i updated on day t and estimate the following panel regression:

$$ln(e_{it-1}^2) - ln(e_{it}^2) = \alpha_i + x_{it}'\beta + u_{it}, \tag{2}$$

where e_{it} denotes the forecast error for forecaster i on day t and e_{it-1} the forecast error on the day that forecaster i previously updated. The regressors x_{it} are reported in the first column of Table 2.

Table 2 reports the estimation results.[20] Column (1) confirms the expected finding that

[20]The estimates in Table 2 can also be expressed in terms of absolute forecast errors (AFE). The table

information releases are associated with accuracy improvements, as the coefficients for IPCA15 and MPC are large and significant. Perhaps more surprisingly, it shows that the contest also has an effect, as accuracy improvements go up by 7.1 p.p. on the contest day. In contrast, column (2) shows that other variables that according to Table 1 are associated with an increased probability of updating–Mondays, Fridays and the EMBI–are not associated with an increase in accuracy improvements (except for the coefficient for the Monday dummy, which is however only significant at the 10% level).

Aggregate Dynamics of Updates and Accuracy: We now focus on the aggregate dynamics of updates and accuracy around the contest day and around days of information releases (the MPC meetings and the IPCA15). Since these dates change across months, we consider a window of five days around these days. The left panel of Figure 4 is the same as Figure 1, and the right panel considers a window of five days around the MPC day. The figure shows that the main driver of updates and accuracy improvements at the aggregate level is the contest, as there are no visible similar changes on the days associated with information releases. The figure also confirms the finding from the panel regressions that forecasters update outside the contest and information release days (about 10% of forecasters update on each non-contest day).

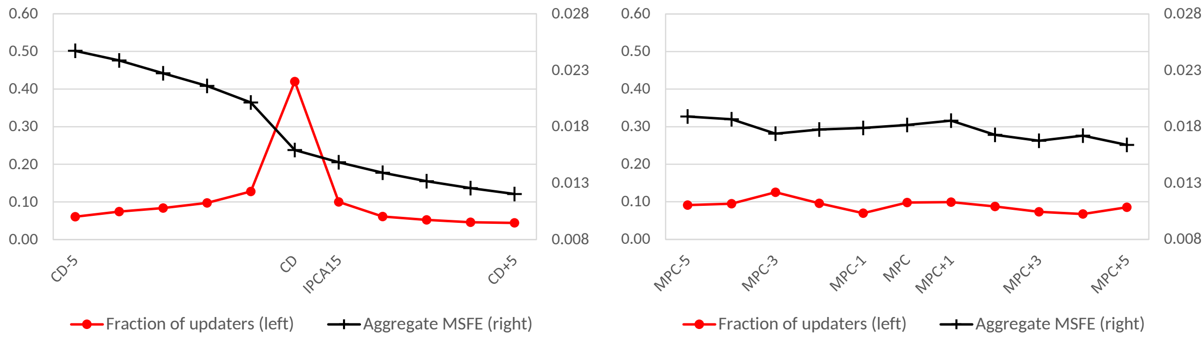
The conclusion from the left panel of Figure 4 is that, although there is a small asymmetry in updating behavior before and after the contest, the fraction of updaters is approximately constant on non-contest days, but it rises substantially on the contest day. The $MSFE$ declines as the forecast horizon shrinks, which is an expected consequence of the natural resolution of uncertainty leading up to the revelation of the forecasted variable. The effect of the contest is to induce a sizable level shift downwards in the $MSFE$ curve, resulting in a much larger improvement in accuracy on the contest day (and consequently for the rest of the month), relative to the (approximately constant) improvement we see on any other day.

The documented jump in aggregate accuracy on the contest day could be caused by both changes in the extensive margin (if more forecasters update, their average accuracy is higher)

---

measures accuracy as minus the change in the log of the squared forecast error (SFE) relative to the previous update. Using the fact that $\log(SFE) = 2\log(AFE)$, the contest effect estimate of 7.1, for example, translates into an accuracy improvement of approximately 3.6 p.p. in terms of absolute forecasts errors. In words, the absolute deviation of the forecast from realized inflation, on average, is reduced by 3.6 p.p. on the contest day.

Figure 4. Dynamics of Updates and Aggregate MSFE



Notes: : Daily evolution of the fraction of updaters and aggregate $MSFE$ around the contest and IPCA15 day (left graph) and around the MPC day (right graph).

and in the intensive margin (each forecaster may be putting more effort into obtaining an accurate forecast). The estimated structural model allows us to decompose the aggregate accuracy improvement on the contest day into the contribution of changes along both margins.

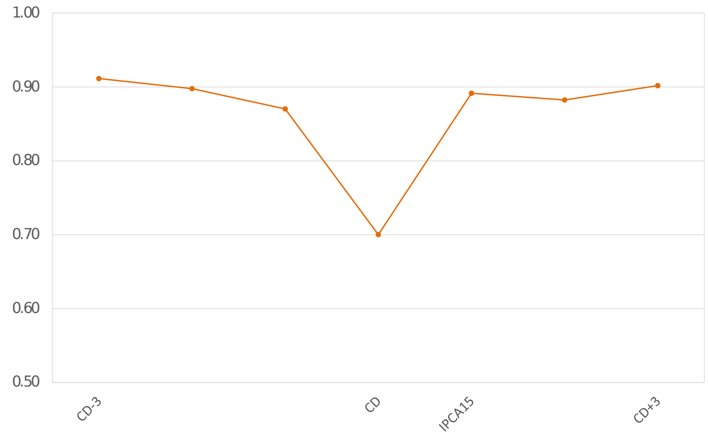
Our findings suggest that the contest may be crowding-out updates on other days. Since the contest is the day before, few forecasters update on the IPCA15 day, even though doing so improves accuracy. In fact we computed that the cases when a forecaster updates on both the contest day and the IPCA15 day constitute only 0.65% of the sample. Another counterfactual exercise allows us to shed light onto a potential crowding-out effect of the contest on aggregate accuracy.

Incentive-Dependent Information Rigidities: We conduct the analysis in Coibion and Gorodnichenko (2015) on our data, to investigate if their measure of information rigidity depends on whether the incentives are high (as on the contest day) or low (as on regular days).

We estimate regression (10) in Coibion and Gorodnichenko (2015) separately for each day t around the contest day.[21] We then compute the implied degree of information rigidity, $\frac{\beta_t}{(1+\beta_t)}$ with β_t the slope parameter in the regression for day t, and report it in Figure 5. The figure

[21]For each given day t in a window around the contest day (CD), $t = CD-3, ..., CD+3$, the regression is $y_m - F_t y_m = c_t + \beta_t (F_t y_m - F_{t-1} y_m) + error_m$, $m = 1, ..., 132$, where y_m is inflation in month m, $F_t y_m$ is the consensus forecast (the mean across forecasters) of y_m made on day t of that month and $F_{t-1} y_m$ is the consensus forecast on day $t-1$.

Figure 5. Coibion and Gorodnichenko (2015) estimates of information rigidity



Notes: The figure reports the estimated degree of information rigidity, as in Coibion and Gorodnichenko (2015), on each day around the contest.

shows that the degree of information rigidity clearly depends on the incentives, as it is much lower on the contest day than on other days.

Forecaster Heterogeneity: The observable dimensions of heterogeneity in our data are the updating behavior and the forecast accuracy across forecasters and over time. To investigate whether they are driven by time-invariant fixed effects (e.g., forecasters who are always frequent updaters and/or the most accurate) we compute measures of mobility in the cross-sectional distributions of both updating probability and accuracy. Specifically, we consider forecasters who participated in the last two years of the sample and compute the normalized trace measure of Shorrocks (1978), by dividing the cross-sectional distribution of average MSFE or total number of updates during each year into 5, 10 or 20 quantiles and computing transition probabilities among the different quantiles.

Table 3 shows that the indexes are close to 1 (which corresponds to perfect mobility).[22] This suggests that there is significant mobility in the distribution of both accuracy and updating probability across forecasters. This result supports the conclusion that our findings are not pri-

[22] The index represents transition probabilities among different quantiles of the distribution of either accuracy or updating frequencies between two consecutive years. For example, the value 0.749 for the MSFE for 5 quantiles implies that a forecaster has a 74.9% probability of moving to a different quantile of accuracy (out of 5) the following year.

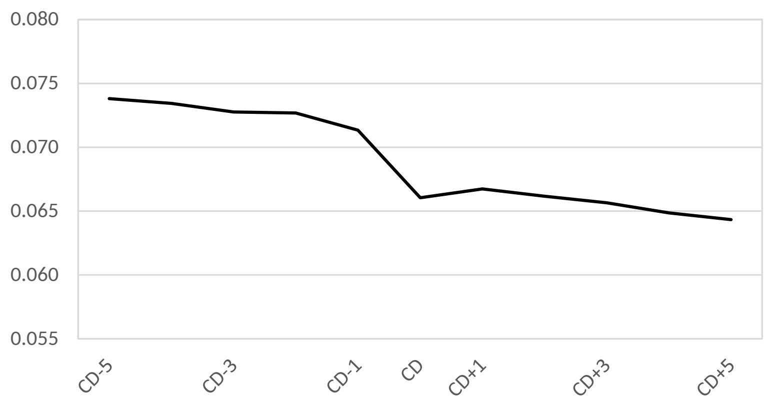
Table 3. Mobility Index of Shorrocks (1978)

Number of Quantiles	Frequency of Updates	$MSFE$
5	0.691	0.749
10	0.760	0.793
20	0.872	0.886

Notes: Index based on the last two years of the sample and on dividing the cross-sectional distribution of the yearly frequency of updates and the average MSFE over the year into different quantiles. Immobility=0 and perfect mobility=1.

marily driven by time-invariant heterogeneity, and it leads us to rule out the potential presence of *positive selection* on the contest day that may induce more able forecasters to update on that day. The general conclusion also motivates our modelling of heterogeneity in the theory that we develop below: We assume that forecasters face heterogeneous incentive parameters that result in both heterogeneous updating histories and heterogeneous attention choices.

Figure 6. Forecast Disagreement (Standard Deviation of Forecasts)



Forecaster's Objective and Strategic Behaviour: The empirical facts document that the contest affects both the updating decision and the accuracy of forecasts. So a reasonable objective function is that forecasters seek to win the contest and maybe have another objective on other days. In the simpler contest with one prize studied in Marinovic et al. (2013), the probability of winning increases the more accurate is a forecast but, also, the more it differs from the competing forecasts: If, say, N forecasters submit the same forecast, the probability that one of them wins is 1/N. Marinovic et al. (2013) argue that this encourages strategic forecasters to put more weight on private signals, compared to what would have been optimal in a decision-theoretic setup, leading to an increase in disagreement among forecasters on the contest day. This is however inconsistent with the data (see Figure 6), as the disagreement *decreases* on the contest day. We take this as evidence that forecasters in our data do not behave strategically (this was confirmed in conversation by a few participants). Based on these observations, we find it more suitable (as well as tractable) to employ a decision-theoretic setup, rather than a game-theoretic one, and assume that a forecaster's objective is to maximize accuracy.

5 Theory

We build on the theory of rational inattention (Sims, 2003) to link the observable dynamics of updating decisions to the available information and to the unobservable cost and benefit parameters that are driving agents' decisions. The theory is inspired by the way forecasters behave in reality: They use state-of-the art statistical models in which they input both publicly available information and privately collected and processed signals.

The theory is also motivated by the empirical regularities: The fact that forecasters do not update every day suggests that they face fixed costs of updating. Moreover, the patterns of accuracy improvements suggest that forecasters might face time-varying costs of processing information in order to produce a forecast. In the model, forecasters choose the amount of mental capacity–henceforth "attention"–to employ in order to formulate a forecast that is accuracy-maximizing given those resources. This optimal amount depends on the time-varying costs of processing information. We build a general framework that allows for two separate decisions, namely whether or not to update and how much attention to devote to processing information. In the estimation, we consider different variations of the baseline model that shut down one of the two decisions at a time or link the two decisions to each other.

5.1 Baseline Model

Our baseline model is a model of sticky and noisy information, but with endogenous information.

The model assumes that, at the beginning of each month all agents produce an initial forecast. On each subsequent day t, agent i makes two decisions: (i) whether or not to update and (ii) how much attention, k_{it}, to devote to collecting/processing information in order to update the forecast. Thus, attention is *endogenous*.[23] The following two subsections discuss the two decisions separately.

5.1.1 Decision of Whether to Update

Agents face not only explicit costs of processing information, but also opportunity costs of time or mental effort—e.g. consultants need to travel, employees at financial institutions have meetings or other inflexible work obligations. Let C_{it}^F denote the *"fixed" cost* of updating–that is, a cost that has to be incurred regardless of the level of "attention" choice. If agent i has a fixed cost below some constant $\bar{C}$, so

$$C_{it}^F < \bar{C} \tag{3}$$

she finds it worthwhile to update on day t.

5.1.2 Decision of How Much Attention to Allocate

In addition to the fixed cost of updating, agents also face information-processing costs, c_{it}. We assume that agents update in order to improve accuracy, or, equivalently, reduce Mean Squared Forecast Error (MSFE). The MSFE for agent i on day t is defined as:

$$MSFE_{it} = E[(y_m - f_{it})^2], \tag{4}$$

where f_{it} is agent i's forecast of monthly inflation y_m produced on day t.

[23]In contrast, in leading dynamic rational inattention models (e.g., Steiner et al., 2017) the decision-maker faces some exogenously *fixed* capacity and, given this, chooses and commits at $t = 0$ to a full contingent plan of which signals to observe in each period.

At the beginning of day t an agent chooses how much attention k_{it} to allocate in order to minimize the sum of future MSFEs, conditional on the fixed cost being less than $\bar{C}$. The MSFE on day t can in principle depend on all past attention choices as well as the current choice. Let $k_i^t = \{k_{i1}, \ldots, k_{it}\}$ denote this sequence of choices. The optimization problem for agent i on day t can thus be stated as follows:[24]

$$\min_{k_{i\tau}} \sum_{\tau=t}^{T} \left(\frac{1}{2\ln 2} MSFE_{i\tau}(k_i^\tau) + c_{i\tau} k_{i\tau} \right) \mathbb{1}_{C_{i\tau}^F < \bar{C}}, \tag{5}$$

where T is the number of working days in the month, $\mathbb{1}$ is the indicator function and $c_{i\tau}$ is the marginal cost of attention on day τ.

In what follows, we describe how to obtain an analytical expression for $MSFE_{i\tau}$ in equation (5) in terms of attention and then solve the optimization problem. We proceed in five steps: Step 1 specifies agents' statistical model; Step 2 specifies agents' information set–the information they need to process to feed into their statistical model; Step 3 derives an analytical expression of $MSFE_{i\tau}$ which depends on the statistical model and the precision of information; Step 4 expresses precision and ultimately $MSFE_{i\tau}$ as a function of attention and Step 5 derives an expression for optimal attention.

Step 1. Specifying a Statistical Model: Monthly inflation y_m–the difference between the log of the price index at the end of the current month and that at the end of the previous month–can be written as the sum of daily inflation x_t (the difference between the logs of the prices on days t and $t-1$):

$$y_m = \sum_{t=1}^{T} x_t. \tag{6}$$

We assume that agents model monthly inflation as an ARMA, which implies (using results from temporal aggregation of ARMA models, e.g., Amemiya and Wu, 1972) that daily inflation is also an ARMA, and the orders and parameters of the two models can be related analytically. In the case of Brazil, the ARMA that best fits monthly inflation according to the BIC is an

[24] The division by $2ln2$ in the minimization problem is a normalization.

ARMA(1,1):

$$y_m = a + \psi y_{m-1} + v_m + \theta v_{m-1}, v_m \sim i.i.d.\mathcal{N}(0, \sigma_v^2), \tag{7}$$

which implies that daily inflation is an AR(1):

$$x_t = b + \phi x_{t-1} + \varepsilon_t, \varepsilon_t \sim i.i.d.\mathcal{N}(0, \sigma_\varepsilon^2), \tag{8}$$

with $b = \frac{a(1-\psi^{1/T})}{T(1-\psi)}$, $\phi = \psi^{1/T}$ and $\sigma_v^2 = (1 + (1+\phi)^2 + ... + (1 + \phi + \phi^2 + ... + \phi^{T-1})^2)\sigma_\varepsilon^2$.

Note that there are two dynamic dimensions in our setting: the month-to-month problem of forecasting current-month inflation at the beginning of the month, and the within-month problem of updating the forecast. We make the two problems coherent by assuming that the initial forecast is based on the ARMA(1,1) in (7) and the updates are based on the AR(1) in (8). Our main focus here is on the within-month dynamic problem.

Step 2: Specifying Agents' Information Set: During the forecasting period, agents can not only collect and process private information, but also have access to public information. We make the following assumptions.

Assumption 1 (Public Signals) On day t, the public signal contains *past* values of daily inflation:[25]

$$s_p^t = \{x_{t-1}, x_{t-2}, ...\}. \tag{9}$$

Assumption 2 (Private Signals) On day t, *current* daily inflation x_t is not observed but agent i can obtain a noisy signal s_{it} about it.

Assumptions 1 and 2 imply that the agent's information set on day t is:

$$s_i^t \equiv (s_{it}, s_p^t) = \{s_{it}, x_{t-1}, x_{t-2}, ...\}. \tag{10}$$

The precision of the current private signal is endogenous and depends on how much attention the agent decides to allocate to information gathering and processing.

[25]This assumes perfect observability of past inflation realizations, but measurement error could be easily accommodated. This would, on the one hand, add a parameter that would give more flexibility in the estimation, but, on the other, would require additional assumptions.

Assumptions 1 and 2 are crucial in making the optimization problem tractable. The next two steps show that these assumptions imply that $MSFE_{i\tau}(k_i^\tau)$ in (5) is only a function of current attention $k_{i\tau}$ and not of the entire sequence of past choices. This implies that the agent's problem is equivalent to a myopic choice and the dynamic problem turns into a sequence of static problems. We discuss and motivate these assumptions in Section 6.

Step 3. Forecast Updates and MSFE: The optimal forecast is the conditional expectation of y_m, based on the information set available to the agent. The initial forecast is based on the ARMA(1,1) model in equation (7) and is given by $E\left[y_m|y_{m-1}, y_{m-2}, ...\right]$. This corresponds to an initial MSFE that is constant across agents and months:

$$MSFE_0 = \sigma_v^2 = [1 + (1+\phi)^2 + ... + (1+\phi+\phi^2+...+\phi^{T-1})^2]\sigma_\varepsilon^2. \tag{11}$$

On a given day $1 \leq t \leq T$ of the month, the forecast update is the conditional expectation of y_m based on each agent's information set s_i^t, $E\left[y_m|s_i^t\right]$. Combining (10), (6) and (8) implies:

$$E\left[y_m|s_i^t\right] = \sum_{j=1}^{t-1} x_j + \sum_{j=t}^{T}\left(\frac{b(1-\phi^{j-t})}{1-\phi} + \phi^{j-t}E\left[x_t|s_i^t\right]\right), \tag{12}$$

so that $MSFE_{it} = E\left[\left(y_m - E\left[y_m|s_i^t\right]\right)^2\right] = \left[1 + (1+\phi)^2 + ... + \left(1+\phi+...+\phi^{T-1-t}\right)^2\right]\sigma_\varepsilon^2 + \left(\sum_{j=t}^{T}\phi^{j-t}\right)^2 E\left[(x_t - E[x_t|s_i^t])^2\right]$, or

$$MSFE_{it} = \left[1 + (1+\phi)^2 + ... + \left(1+\phi+...+\phi^{T-1-t}\right)^2\right]\sigma_\varepsilon^2 + \left(\sum_{j=t}^{T}\phi^{j-t}\right)^2 E\left[\sigma^2_{x_t|s_i^t}\right], \tag{13}$$

where $\sigma^2_{x_t|s_i^t}$ denotes the conditional variance of x_t. Note that Assumptions 1 and 2 imply that the MSFE depends only on the conditional variance of the current signal and not on past private signals.

Step 4. MSFE as a Function of Attention: We now express the expectation $E\left[\sigma^2_{x_t|s_i^t}\right]$ in equation (13) in terms of attention. Following the rational inattention literature, the *additional* information content of the signal s_{it} is captured by the relative conditional entropy based on the information sets with and without the signal (respectively (s_{it}, s_p^t) and s_p^t):

$$I(x_t; s_{it}|s_p^t) = H(x_t|s_p^t) - E_{s_{it}}[H(x_t|s_{it}, s_p^t)|s_p^t] \leq k_{it}. \tag{14}$$

The cost of information is then $c_{it}I(x_t; s_{it}|s_p^t)$ and in the formulation (5) we have anticipated that the constraint binds as agents exhaust "attention," k_{it} (see also Wiederholt (2010)).

In our Gaussian-quadratic objective framework it is well-known that the optimal distribution of signals is normal. Then, under the assumption that x_t and s_i^t have a joint normal distribution, the conditional entropy of $x_t|s_i^t$ is: $H(x_t|s_i^t) = \frac{1}{2}\log_2(2\pi e\sigma^2_{x_t|s_i^t})$. The inequality (14) holds with equality because the agent exhausts all capacity. The agent chooses the distribution of s_{it} so that $\sigma^2_{x_t|s_p^t,s_{it}}$ (which is $\sigma^2_{x_t|s_i^t}$) is the same for each signal realization s_{it} and, hence, $E\left[\sigma^2_{x_t|s_i^t}\right] = \sigma^2_{x_t|s_i^t}$. This implies that $\frac{\sigma^2_{x_t|s_p^t}}{\sigma^2_{x_t|s_p^t,s_{it}}} = 2^{2k_{it}}$ or

$$E\left[\sigma^2_{x_t|s_i^t}\right] = \sigma^2_{x_t|s_i^t} = \sigma^2_{x_t|s_p^t}(2^{2k_{it}})^{-1}. \tag{15}$$

By substituting (15) into (13) and using the fact the AR(1) model implies $\sigma^2_{x_t|s_p^t} = \sigma^2_\varepsilon$, we obtain:

$$MSFE_{it} = \left[1 + (1+\phi)^2 + ... + \left(1 + \phi + ... + \phi^{T-1-t}\right)^2\right]\sigma^2_\varepsilon + \left(\sum_{j=t}^{T}\phi^{j-t}\right)^2 \sigma^2_\varepsilon(2^{2k_{it}})^{-1}. \tag{16}$$

Equation (16) shows that the $MSFE_{it}$ depends only on current attention k_{it} and not on past attention choices and it has two components: The first is common across agents and captures the resolution of uncertainty due to the public signal. The second depends on how much attention each agent allocates to obtaining a better signal for current-day inflation (i.e., on the choice of k_{it}), and on how this feeds into the monthly forecast.

Step 5. Optimal Attention: As usual, we solve the problem of sequential decisions backwards. Consider agent i's problem at the last period $t = T$:

$$\min_{k_{iT}} \left(\frac{MSFE_{iT}}{2\ln 2} + c_{iT}k_{iT}\right)\mathbb{1}_{C^F_{iT}<\bar{C}} \text{ subject to } k_{iT} \geq 0, \text{ (15) and (16)}. \tag{17}$$

The agent can only control the part of the $MSFE$ in (16) that depends on collecting information about the current daily signal, so optimal attention solves:

$$\min_{k_{iT}} \left(\frac{1}{2\ln 2} \sigma_\varepsilon^2 \left(2^{2k_{iT}}\right)^{-1} + c_{iT} k_{iT} \right) \mathbb{1}_{C_{iT}^F < \bar{C}} \text{ s.t. } k_{iT} \geq 0.$$

Differentiating with respect to k_{iT} and rearranging gives optimal attention as:[26]

$$k_{iT}^* = \begin{cases} \frac{1}{2}\log_2 \left(\frac{\sigma_\varepsilon^2}{c_{iT}} \right) & \text{if } c_{iT} < \sigma_\varepsilon^2 \text{ and } C_{iT}^F < \bar{C} \\ 0 & \text{otherwise.} \end{cases}$$

As anticipated, our assumption that the public signal is the realization of the variable implies that it is by construction weakly more precise than any past private signal which would require infinite attention to be as precise. The agent knows that t's choice of attention does not affect $t+1$'s decision, because the prior uncertainty that she reduces the following day is not based on past private signals, but on the more precise public signal. Hence, the agent's dynamic problem turns into a sequence of static problems. In particular, at the beginning of day t an agent solves:

$$\min_{k_{it}} \left(\frac{\left(\sum_{j=t}^{T} \phi^{j-t}\right)^2}{2\ln 2} \sigma_\varepsilon^2 \left(2^{2k_{it}}\right)^{-1} + c_{it} k_{it} \right) \mathbb{1}_{C_{it}^F < \bar{C}} \text{ s.t. } k_{it} \geq 0, \tag{18}$$

which gives

$$k_{it}^* = \begin{cases} \frac{1}{2}\log_2 \left(\frac{\left(\sum_{j=t}^{T} \phi^{j-t}\right)^2}{c_{it}} \sigma_\varepsilon^2 \right) & \text{if } c_{it} < \left(\sum_{j=t}^{T} \phi^{j-t}\right)^2 \sigma_\varepsilon^2 \text{ and } C_{it}^F < \bar{C} \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

The formula implies that attention is higher the smaller the marginal cost of attention, the earlier the day is in the month, and the larger the prior variance of the signal (measured by σ_ε^2). Optimal attention varies over time because of two reasons. The first has to do with the resolution of uncertainty due to the revelation of the public signal, which is common across

[26] Note that if $C_{iT}^F > \bar{C}$ the objective is constant and equal to zero so any choice of k is optimal, and we choose $k_{iT}^* = 0$.

agents (this is captured by the time-varying component $\left(\sum_{j=t}^{T}\phi^{j-t}\right)^2$). The second is due to agents possibly facing different costs c_{it} on different days.

5.1.3 Model-Implied Moments

The moments we seek to match are the fraction of updaters and the aggregate MSFE. In this model, updaters are agents who pay positive attention, which occurs when the condition in the first line of equation (19) is satisfied, so the fraction of updaters out of N agents on day t is given by:

$$\lambda_t = \frac{\left\{\#\ i \text{ s.t. } c_{it} < \left(\sum_{j=t}^{T}\phi^{j-t}\right)^2 \sigma_\varepsilon^2 \text{ and } C_{it}^F < \bar{C}\right\}}{N}, \tag{20}$$

and the optimal MSFE of agent i on day t – obtained by substituting optimal attention in (19) into (16) – is given by:

$$MSFE_{it}^* = \begin{cases} \left[1 + (1+\phi)^2 + ... + \left(1+\phi+...+\phi^{T-1-t}\right)^2\right]\sigma_\varepsilon^2 + c_{it} & \text{if } i \text{ updates} \\ MSFE_{it-1}^* \text{ otherwise} \end{cases}. \tag{21}$$

We next consider a number of variations of this model.

5.2 Variation 1: Sticky Model

In this variation of the model there is no cost of processing information since updaters – agents whose fixed cost satisfies (3) – receive a perfect signal of today's inflation.

This model thus implies that the fraction of updaters on day t is

$$\lambda_t = \frac{\left\{\#\ i \text{ s.t. } C_{it}^F < \bar{C}\right\}}{N}, \tag{22}$$

and the optimal MSFE of updaters is as in (13), but with the last term equal to zero, so that:

$$MSFE_{it}^{*} = \begin{cases} \left[1 + (1+\phi)^2 + ... + \left(1 + \phi + ... + \phi^{T-1-t}\right)^2\right]\sigma_{\varepsilon}^2 & \text{if } i \text{ updates} \\ MSFE_{it-1}^{*} \text{ otherwise} \end{cases}. \tag{23}$$

5.3 Variation 2: Noisy Model

In this variation there are no fixed costs, so $C_{it}^F = 0$ for all i, t, but only information-processing costs c_{it}. In this variation updaters are agents who pay positive attention, given by:

$$k_{it}^{*} = \begin{cases} \frac{1}{2}\log_2\left(\frac{\left(\sum_{j=t}^{T}\phi^{j-t}\right)^2}{c_{it}}\sigma_{\varepsilon}^2\right) & \text{if } \frac{\left(\sum_{j=t}^{T}\phi^{j-t}\right)^2}{c_{it}}\sigma_{\varepsilon}^2 > 1 \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

This model thus implies that the fraction of updaters on day t is:

$$\lambda_t = \frac{\left\{\# \; i \text{ s.t. } c_{it} < \left(\sum_{j=t}^{T}\phi^{j-t}\right)^2\sigma_{\varepsilon}^2\right\}}{N}, \tag{25}$$

and the MSFE is as in the baseline model:

$$MSFE_{it}^{*} = \begin{cases} \left[1 + (1+\phi)^2 + ... + \left(1 + \phi + ... + \phi^{T-1-t}\right)^2\right]\sigma_{\varepsilon}^2 + c_{it} & \text{if } i \text{ updates} \\ MSFE_{it-1}^{*} \text{ otherwise} \end{cases}. \tag{26}$$

5.4 Variation 3: Restricted Baseline Model

In this variation we micro-found the fixed cost C_{it}^F, which is the fixed cost of time, and link it to the information processing cost c_{it}. Let w_{it}^o denote the marginal benefit of time devoted to activities other than forecasting and let w_{it} be the marginal benefit of increasing forecast accuracy for agent i at day t.[27] The opportunity cost of time is:

$$C_{it}^F \equiv \frac{w_{it}^o}{w_{it}}. \tag{27}$$

[27] More precisely, w_{it}^o stands for the highest marginal benefit of time on any of the activities other than forecasting. Its purpose is to capture the opportunity cost of attention.

The attention choice is as before, except that the parameter w_{it} multiplies MSFE in the optimization problem (18). Then, following analogous arguments as in the baseline model, the dynamic attention allocation problem boils down to a sequence of static ones and at the beginning of day t an agent solves:

$$\min_{k_{it}} \left(\frac{w_{it}}{2\ln 2} \left(\sum_{j=t}^{T} \phi^{j-t} \right)^2 \sigma_\varepsilon^2 \left(2^{2k_{it}}\right)^{-1} + c_{it} k_{it} \right) \mathbb{1}_{\frac{w_{it}^o}{w_{it}} < \bar{C}} \text{ s.t. } k_{it} \geq 0, \tag{28}$$

which gives

$$k_{it}^* = \begin{cases} \frac{1}{2}\log_2 \left(\frac{w_{it}}{c_{it}} \left(\sum_{j=t}^{T} \phi^{j-t} \right)^2 \sigma_\varepsilon^2 \right) & \text{if } \frac{w_{it}}{c_{it}} \left(\sum_{j=t}^{T} \phi^{j-t} \right)^2 \sigma_\varepsilon^2 > 1 \text{ and } \frac{w_{it}^o}{w_{it}} < \bar{C} \\ 0 & \text{otherwise.} \end{cases} \tag{29}$$

The fraction of updaters is then given by

$$\lambda_t = \frac{\left\{ \#\ i \text{ s.t. } \frac{w_{it}}{c_{it}} \left(\sum_{j=t}^{T} \phi^{j-t} \right)^2 \sigma_\varepsilon^2 > 1 \text{ and } \frac{w_{it}^o}{w_{it}} < \bar{C} \right\}}{N}, \tag{30}$$

and the optimal MSFE is:

$$MSFE_{it}^* = \begin{cases} \left[1 + (1+\phi)^2 + ... + \left(1 + \phi + ... + \phi^{T-1-t}\right)^2 \right] \sigma_\varepsilon^2 + \frac{c_{it}}{w_{it}} & \text{if } i \text{ updates} \\ MSFE_{it-1}^* \text{ otherwise} & \end{cases}. \tag{31}$$

Note that in this version of the model the decision to update is linked to the size of the revision in a tractable way: The assumption that the same benefit drives both the decision to update and the choice of attention implies that an increase in w_{it} is associated with both a "larger" revision (manifested as a larger improvement in MSFE) and a larger probability of updating.

6 Discussion of Assumptions

For the sake of tractability, our models rely on three main simplifying assumptions.

The first simplification is that we consider a decision-theoretic rather than a strategic setting.

This assumption buys tractability and is well-grounded, because the empirical regularities are at odds with the key predictions of strategic models. In particular, the motive to win the contest would imply that forecasters put more weight on private signals /idiosyncratic sources of information rather than public information available to everyone. This incentive would tend to attenuate forecast dispersion ("disagreement") on the contest day. This is something that we don't see in the data, as we document in Figure 6. Interviews with forecasters revealed that all they try to do on the contest day is to provide the most accurate forecast they can. That is – they don't strategize. This could be perhaps because there are many forecasters and the type of strategizing that the theory suggests does not pay off. Perhaps, also, it is harder to think how to strategize in a multi-prize contest like the one used in this survey, where the five most accurate forecasts are named. Multi-prize contests are very difficult to analyze theoretically and, to the best of our knowledge, the paper by Ottaviani and Sørensen (2003) focuses primarily on a winner-take-all contest and has not been followed up by a study of multi-prize contests. The analysis is quite complex, but the prediction of an increase in disagreement is still valid, thus further suggesting that strategizing is probably not present in our data (and that, even if it were, it would have no empirical implications).

The second simplification is the assumption that the decision of whether or not to update boils down to comparing the opportunity cost of time to a fixed threshold. Similar decision rules are considered in the literature on price setting with menu costs (e.g., Midrigan, 2011; Gertler and Leahy, 2008), where it is shown that a rule of the form "update price if the menu cost is below some fixed threshold" is a good approximation to the optimal decision rule. We believe this simplification does not diminish the contribution of the paper in terms of providing empirical credence to models of rational inattention, as the same conclusions would have emerged by focusing only on the intensive margin of updating and validating this part of the model using the patterns of accuracy in the data. As discussed by Woodford (2009) in a different context, applying rational inattention to a timing decision presents additional challenges, so attempting to solve a joint rational inattention optimization for the two decisions would substantially complicate the analysis without necessarily adding new insights. We consider both decisions in this paper for quantitative realism and because both extensive and intensive margins matter for policy and survey design. Our estimation findings also show that the models that only consider one of the two decisions at a time (i.e., the sticky and the noisy models) do not fit the data.

The third simplification is that updaters rely on past public information rather than past private signals (Assumptions 1 and 2). Arguably, Assumptions 1 and 2 are more realistic and plausible than the opposite extreme of assuming that updaters only use past private signals: If we want the theory to be generally applicable, we should note that our survey is an exception in its high frequency–typical surveys are collected monthly or quarterly. When updating at these frequencies, agents surely have access to past official monthly or quarterly data releases. It is doubtful that agents would rather use a sequence of incomplete and less accurate private signals instead of the complete set of accurate public information. Even in our high-frequency case, there is plausibly public information about past daily inflation, for example daily releases of gasoline prices. These assumptions make the model tractable by turning the dynamic rational inattention problem into a sequence of static decisions. Otherwise, the choice of attention would depend on all past and current cost and benefit parameters.

The prediction of our model that the accuracy of updaters on a given day only depends on the incentives on that day and not on past updating choices is supported by the data. First, Table 2 shows that there is no accuracy improvement for updaters the day before the contest, whereas in a dynamic setting we should see higher attention (so higher accuracy) before the contest as participants expect to leverage this attention on the contest day. Further, the results in Table 4 – a panel regression for the accuracy of updaters considering the same regressors as Table 2 – show that the duration between updates (which is linked to past updating choices), has no significant effect on accuracy at the 5% confidence level.

Finally, in our model the "attention cost" takes the usual form $c \times I(x_t; s_{it}|s_p^t)$, where the mutual information $I(x_t; s_{it}|s_p^t)$ is bounded above by the "attention choice" k. In principle, we can allow for $c\left(I(x_t; s_{it}|s_p^t)\right)$, where c is increasing and convex.[28] This would imply that the optimal choice k_{it} would involve the curvature of the cost function. Given the nature of the data, we cannot estimate the cost's curvature for the following reasons: In our setting fixed costs (opportunity costs of time) seem important as a model with only information processing

[28] Cost functions based on mutual information are a special case of those more general cost functions discussed in Hébert and Woodford (2017), Caplin, Dean, and Leahy (2017), Pomatto, Strack, and Tamuz (2020), where the cost depends on the prior and the posterior.

Table 4. Drivers of Accuracy for Updaters

Regressors	Panel Fixed Effect Coefficients
d_t^{CD-1}	0.00049
	(0.00048)
d_t^{CD}	0.00208***
	(0.00035)
d_t^{IPCA15}	0.00331***
	(0.00045)
$d_t^{IPCA15+1}$	0.00467***
	(0.00039)
d_t^{MPC}	-0.00069
	(0.00054)
$duration_{i,t}$	-0.00006*
	(0.00004)
$horizon_t$	-0.00087***
	(0.00003)
$constant$	-0.00200***
	(0.00035)

Notes: Dependent variable is minus the squared forecast error for agent i on day t. Sample from January 8th, 2004 to January 8th, 2015. Number of observations 31,319. Standard errors in parentheses. ***, ** and * indicate, respectively, significance at the 1%, 5% and 10% level.

costs is rejected. See Appendix B. Also, it is not just the (unscaled) marginal cost of processing information that matters: Forecasters' benefit from accuracy varies as we see updates before and after the contest day and many more updates on the contest day. Lastly, we cannot leverage intertemporal links to make inferences about the curvature of the cost function because, as the data suggest, (and also confirmed through informal interviews) forecasters behave in a myopic way and focus on the current task. In other words, they optimize attention k given current day's costs.[29]

[29] We find comparable average reduction in MSFE per updater before and after the contest day (0.0119 and 0.0123, respectively) suggesting that forecasters do not think ahead and do not internalize that more attention today (some day before the contest) leads to lower attention costs tomorrow. The patterns in the data suggest that they seem to put similar attention before and after the contest day.

7 Estimation

In this section we report estimation results for the restricted baseline model, which is our preferred model and the one we use for the counterfactuals. We also summarize the estimation results for the other versions of the model, while relegating the details of these results to Appendix B. We estimate the models by Simulated Method of Moments (SMM) (e.g., Gourieroux and Monfort, 1996; Duffie and Singleton, 1993; Ruge-Murcia, 2012), which involves matching empirical moments with their theoretical counterparts. The moments we match are the average (over different months) fraction of updaters and the aggregate $MSFE$, which is the average (over different months) of the average $MSFE$ across agents computed each day. We report results for a 3-day window around the contest day.[30] The simulation is based on τM months and N agents, where $M = 132$ and $N = 85$ as in the data, and τ is an arbitrary number of replications.[31] Due to the high correlation among moments we use a diagonal weighting matrix in the SMM estimation.

7.1 Restricted Baseline Model

We impose normalizations that deliver a parsimoniously parameterized model where all parameters are identified.

7.1.1 Parameters

We first assume that the information processing cost is constant across agents and during the month, except on the IPCA15 day:

$$c_{it} = \begin{cases} c_{IPCA15} & \text{if } t = IPCA15 \\ c & \text{otherwise.} \end{cases}$$

[30] Results are robust to considering different window lengths; however note that larger windows run the risk of going outside the current month, as for some months in the sample the contest day falls early or late in the month.

[31] Following Duffie and Singleton (1993), the requirement is $\tau M \to \infty$ as $M \to \infty$. We set $\tau = 5$.

Regarding the first-stage decision, we first set $\bar{C} = 1$. To eliminate w_{it}^o as a free parameter, we first assume that the fixed cost ($C_{it}^F = \frac{w_{it}^o}{w_{it}}$) is small relative to the information-processing cost c_{it}, so an agent who pays the fixed cost (i.e., $w_{it} > w_{it}^o$) also finds it worthwhile to pay the information-processing cost and choose positive attention (i.e., $k_{it}^* > 0$). The normalization

$$w_{it}^o = \begin{cases} \frac{c_{IPCA15}}{\sigma_\varepsilon^2} & \text{if } t = IPCA15 \\ \frac{c}{\sigma_\varepsilon^2} & \text{otherwise} \end{cases} \quad (32)$$

satisfies this assumption because the first condition for a positive k_{it}^* in (29) – which is equivalent to $w_{it} > \frac{1}{\left(\sum_{j=t}^{T} \phi^{j-t}\right)^2} \frac{c_{it}}{\sigma_\varepsilon^2}$ – always holds if $w_{it} > w_{it}^o$ and w_{it}^o is given by (32).

We model the heterogeneity in the benefits by assuming that the cross-sectional distribution for w_{it} is a truncated normal $TN(\mu_t^w, \sigma_w^2)$. The most parsimonious parameterization for μ_t^w that fits the data is the following step function:

$$\mu_t^w = \begin{cases} \mu_1^w & \text{if } t \leq \text{ CD-2} \\ \mu_{CD-1}^w & \text{if } t = CD - 1 \\ \mu_{CD}^w & \text{if } t = CD \\ \mu_2^w & \text{if } t \geq \text{ CD +1} \end{cases} \quad (33)$$

The parameters of the restricted baseline model are $\theta = \left(\mu_1^w, \mu_{CD-1}^w, \mu_{CD}^w, \mu_2^w, \sigma_w, c, c_{IPCA15}, \phi, \sigma_\varepsilon\right)$.

7.1.2 Theoretical and Empirical Moments

The model-implied moments are obtained as follows. Every month the initial $MSFE$ for all agents is (11). On every subsequent working day $t = 1, ..., T$ of the month, each agent receives a random draw of the benefit parameter w_{it} from a $TN(\mu_t^w, \sigma_w^2)$, with μ_t^w as in (33). The fraction of updaters on day t can be computed analytically as the probability that $w_{it} > w_{it}^o$ implied by the truncated normal:

$$\lambda_t = \begin{cases} P\left(w_{it} > \frac{c}{\sigma_\varepsilon^2}\right) & \text{if } t \neq \text{ IPCA15} \\ P\left(w_{it} > \frac{c_{IPCA15}}{\sigma_\varepsilon^2}\right) & \text{if } t = IPCA15. \end{cases} \quad (34)$$

The $MSFE$ is given by:

$$MSFE^*_{it} = \begin{cases} \left[1 + (1+\phi)^2 + ... + \left(1+\phi+...+\phi^{T-1-t}\right)^2\right]\sigma^2_\varepsilon + \frac{c}{w_{it}} & \text{if } w_{it} > \frac{c}{\sigma^2_\varepsilon},\ t \neq IPCA15 \\ \left[1 + (1+\phi)^2 + ... + \left(1+\phi+...+\phi^{T-1-t}\right)^2\right]\sigma^2_\varepsilon + \frac{c_{IPCA15}}{w_{it}} & \text{if } w_{it} > \frac{c_{IPCA15}}{\sigma^2_\varepsilon},\ t = IPCA15 \\ MSFE^*_{it-1} \text{ otherwise.} \end{cases} \tag{35}$$

The same simulation is repeated for all months, changing only the number of working days T and the date of the contest to match those in the corresponding month in the data.

7.1.3 Estimation results

The estimation results for the restricted baseline model are in Table 5 and Figure 7 shows the fit of this model. The model passes the J-test of overidentifying restrictions with a p-value of 0.49. The estimates confirm that there is an approximately constant incentive for participation to the survey on "normal" days but the contest provides an additional benefit. The information-processing cost is lower on the IPCA15 day. As a test of the external validity of the model, we compare the predictions of the model for the AR(1) parameters – which are not matched in the estimation - to those estimated using actual data. Remarkably, the model-based estimate are very close to the estimates in Brazilian inflation data:[32] The autoregressive coefficient ϕ equals .922 in the model and .963 in the data; the error standard deviation σ_ε equals 5.0E-03 in the model and 3.36E-03 in the data. In unreported results, we tested the restrictions $\mu_t = \mu$, $\mu^w_1 = \mu^w_2$, $\mu^w_{CD-1} = \mu^w_{CD}$, $c_{IPCA15} = c$, and found that each one makes the model fail the J-test. We conclude that time variation in the incentive parameters around the contest day is necessary to explain the patterns in the data.

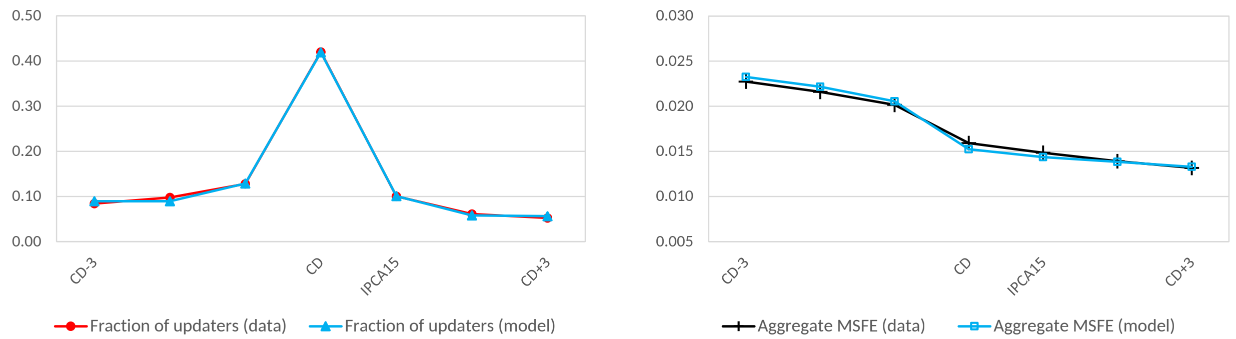
[32] We obtain the estimates of the AR(1) parameter for (unobservable) daily inflation by estimating an ARMA(1,1) on observable monthly inflation data in Brazil from January 2004 to December 2014 and assuming 21 working days in each month, then using the formulas after equation (8) to back out the AR(1) parameters.

Table 5. Restricted Baseline Model Estimation

Parameter	Estimate	Standard Error	Parameter	Estimate	Standard Error
μ_1^w	0.476	4.1E-18	c_{IPCA15}	1.34E-05	1.4E-13
μ_{CD-1}^w	0.488	4.0E-18	c	1.38E-05	1.4E-13
μ_{CD}^w	0.542	3.6E-18	ϕ	0.925	2.1E-18
μ_2^w	0.463	4.2E-18	σ_ε	0.005	3.9E-16
σ_w	0.058	3.4E-17			

Note: p-value of the J-test = 0.49.

Figure 7. Restricted Baseline Model Predictions versus Data. Fraction of Updaters (Left) and Aggregate MSFE (Right)



7.2 Summary of Estimation Results

In Appendix B, we show that the sticky and noisy models are rejected by the data, while the baseline model (marginally) passes the J-test but only by allowing for substantial time-variation in both fixed costs and information-processing costs. Overall, the estimation results provide support for the key implication of rational inattention that agents increase their information processing efforts when the incentives of doing so are higher. First, the estimates of the models that fit the data (the baseline and the restricted baseline) imply that agents increase their information-processing efforts on the contest day, compared to normal days. Second, all models are rejected when assuming that agents do not increase attention in response to the contest (which corresponds to the restriction that costs and/or benefits are constant over time).

The fact that the sticky model is rejected implies that higher incentives do no just make forecasters update in greater numbers, but also make updaters increase their information-processing efforts. The fact that a "pure" rational inattention model (the noisy model) is rejected suggests that there are both fixed costs and information-processing costs at play in our data.

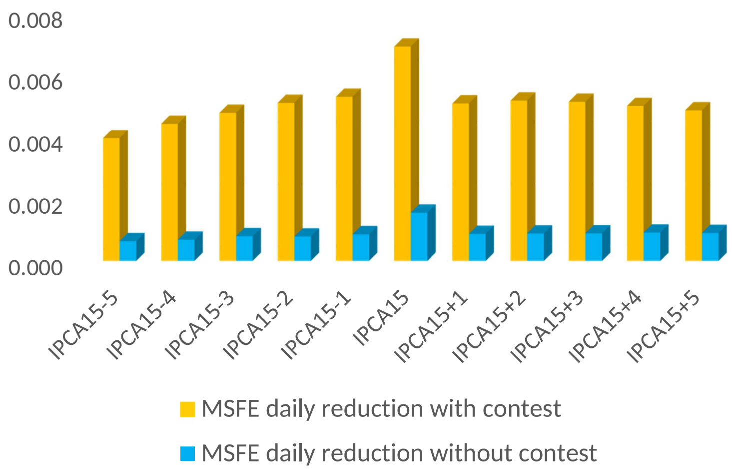
8 Counterfactuals

In this section we perform a number of counterfactual analyses, using the estimates of the restricted baseline model reported in Table 5.

Aggregate Accuracy and Changes in Extensive versus Intensive Margins: The estimates from Table 5 imply that the average $MSFE$ across agents falls from 0.0205 the day before the contest to 0.0152 on the contest day, which is due to both an increase in the number of updaters and to a shift in the benefit distribution across agents (so agents who update put more effort). We then assume that the number of updaters remains the same as before the contest day, but that they receive draws from the shifted distribution of benefits that characterizes the contest day. This would make the $MSFE$ fall to 0.0189, which implies that 30% of the accuracy improvement on the contest day is due to agents paying more attention (intensive margin) and 70% to more agents updating (extensive margin).

Quantifying the Value of the Contest: To assess the value of the contest, we let the contest day fall on each possible day in a five-day window around the IPCA15 and generate counterfactual $MSFE$s as in (31), using the estimates from Table 5. Figure 8 reports the reduction in the average $MSFE$ across agents on each potential contest day (relative to the previous day) and compares it to the reduction in average $MSFE$ that one would observe between two consecutive days in the absence of the contest. We assume that the mean benefit without a contest would be the constant benefit we now only observe after the contest day, i.e., $\mu_t^w = \mu_2^w$ from Table 5 for all t. Figure 8 shows that having the contest on any day has a very large effect on accuracy. The largest accuracy improvement is obtained by having the contest on the IPCA15 day. It amounts to a 347% accuracy improvement relative to not having a contest[33] and a 31% improvement relative to having the contest the day before, as it is now in the survey.

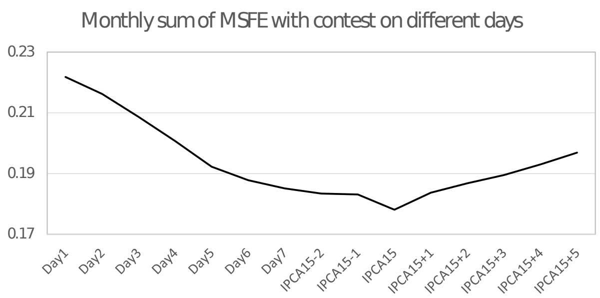
Figure 8. MSFE Daily Reduction With and Without the Contest on Different Days

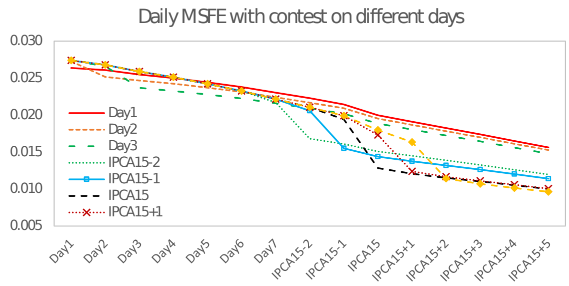


Optimal Timing of the Contest: We next determine the optimal timing of the contest. The right panel of Figure 9 shows the counterfactual *daily* average $MSFE$ when the contest is on different days, starting from the first day of the month and ending five days after the contest day. The daily $MSFE$ decreases almost linearly, and the contest induces a downward shift in the line. The left panel of Figure 9 plots the counterfactual *cumulative* $MSFE$ (the monthly sum of the average $MSFE$). The figure shows that the optimal timing in terms of both daily and cumulative improvements in accuracy is the IPCA15. The percentage improvement in cumulative $MSFE$ of having the contest on the IPCA15 instead of on the day before (as it is currently in the survey) would be 3%. Note that the cumulative $MSFE$ is a U-shaped curve around the IPCA15. Intuitively, this is because there is a tradeoff between holding the contest earlier in the month, when agents observe fewer past signals but there are more days left to

[33] Even without a contest the IPCA15 day would benefit from a larger MSFE reduction than other days, due to the lower cost.

Figure 9. Cumulative MSFE (Left) and MSFE (Right) with Contest on Different Days



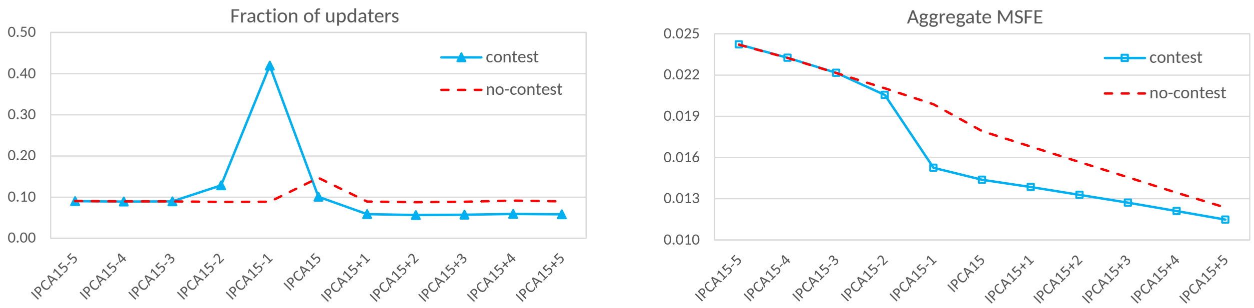


lower the path of the $MSFE$, and holding it later, when more signals are observed but there are fewer remaining days to lower the $MSFE$.

Contest versus No-Contest: We finally investigate the extent to which the contest crowds-out updates–or more precisely mis-aligns updates from the more "natural" IPCA15 day when it is cheaper to process information. One challenge we face is that we do not know how the total number of updates within a month would change in the absence of the contest. In what follows we assume that this number stays constant. This stacks the cards in favour of the no-contest scenario, since one would expect fewer updates without the contest. Holding the total number of updates in a month fixed, we use the model to back out the parameter μ (constant on each day of the month but possibly varying from month to month) that would deliver the same number of updates as in the estimated model. We maintain the same estimates for the other parameters as in Table 5. Then, the IPCA15 is the date with the smallest cost-to-benefit ratio.

Using this counterfactual parameter we simulate the fraction of updaters and the average MSFE for each agent and compare them to those obtained in the presence of the contest. Figure 10 reports the results. The results are eye-opening: Accuracy is worse without the contest, even though most updates happen on the IPCA15 and there are more updates in the days after the IPCA15 than in the presence of the contest. This underscores that the coordinated updates that occur because of the contest are important for the survey's average accuracy.

Figure 10. Fraction of Updaters and Accuracy With and Without the Contest.



9 Conclusions

This paper connects two important ideas in economics: that attention is limited and that incentives matter. We analysed panel data from a unique survey of professional forecasters where the forecast updating decisions of participants are observable and incentive "shifters" are present. The empirical findings are consistent with a rational inattention model in which forecast updating in general, and attention in particular, respond to changes in incentives.

Kacperczyk et al. (2016) conclude "*While information choices have consequences for real outcomes that are poorly understood because they are difficult to measure.*" Our model has predictions for the observables in our data and thus ties information choices to outcomes. The empirical patterns we document and the counterfactuals underscore the importance competition for accuracy. The role of competition among forecasters on the quality of forecasts is also underlined in the influential book Tetlock and Gardner (2016) in the context of forecasting election outcomes. We show that a contest makes more forecasters participate and each forecaster put more effort, resulting in an increase in both individual and aggregate accuracy. Aligning the contest day with information releases would lead to even higher accuracy gains. Our findings can be of interest to central banks and private institutions that run surveys of professional forecasters. Such surveys are increasingly becoming a key input in economic and policy decisions by governments and firms. Despite that many policy institutions worldwide have been running surveys for years and private sector surveys are a thriving and growing industry,[34] little attention has been paid in the literature to how survey design affects forecast quality.

While our model is tailored to a forecast updating setting, ~~we believe that~~ the finding that the amount of attention decision-makers devote to information-processing varies with the incentives carries through to other settings, such as investment or purchasing decisions by household and firms. Afrouzi (2019) shows that firms in more concentrated industries (higher market power) have a higher incentive to track competitors' beliefs. Anecdotal evidence suggests that consumers do spend more time researching purchases when the welfare impact is larger. For a somewhat amusing example, think of coffee lovers who tend to spend more time researching coffee than other drinks! In our setting the incentive takes the form of a contest and it proves a powerful driver of accuracy gains and participation. This finding suggests that contests for best expert opinions, choices of employees' savings, retirement portfolios or investment choices could encourage more attention and active participation.

[34]Interestingly, most private firms focus on nowcasts (short-term) forecasts, as we do in this paper.

A Additional Data Analyses

Table A.1. Summary Statistics: Inflation vs. Survey's Forecasts and Forecast Updates

Variable	Minimum	Average	Maximum	Std. Deviation
Inflation	-0.21	0.46	0.92	0.24
Survey Forecast	-0.30	0.45	1.20	0.21
Forecast Update	-0.55	0.01	0.51	0.09

Note: Sample from January 8th, 2004 to January 8th, 2015. The vast majority of forecasters (95% in our sample) report forecasts with two decimal points.

Table A.2. Summary Statistics: Contest vs. Non-contest days

Variable	Contest day	Non-contest day
Average frequency of updating	0.4201	0.0748
Average size of updates	0.0028	0.0007
Average size of forecast errors	0.0100	0.0122

Note: Sample from January 8th, 2004 to January 8th, 2015.

Table A.3. Mincer and Zarnowitz (1969) rationality tests

(a) Individual forecasts	N	NR	% Rational=NR/N
Contest day	176	161	91%
IPCA15 release day	176	162	92%
(b) Consensus	$\widehat{\alpha}$	$\widehat{\beta}$	p-value
Contest day	-0.04523 (0.02508)	1.12517 (0.05308)	0.0363
IPCA15 release day	-0.04505 (0.02315)	1.12414 (0.04930)	0.0256
(c) Consensus updaters	$\widehat{\alpha}$	$\widehat{\beta}$	p-value
Contest day	-0.01903 (0.01960)	1.05070 (0.04039)	0.4142
IPCA15 release day	-0.01291 (0.01621)	1.04046 (0.03322)	0.3413

Note: Robust standard error in parentheses. Rationality tests based on the null hypothesis $H_0: \alpha = 0, \beta = 1$. N is the number of forecasters with at least 10 observations. NR is the number of rational forecasters (i.e., with p-value>0.05 for the individual rationality test).

B Estimation of Alternative Models

B.1 Baseline Model

We set $\bar{C} = 1$ and parameterize the cross-sectional distribution of the cost parameters as two independent truncated normals: $C_{it}^F \sim TN(\mu_t^F, 0.0001)$ and $c_{it} \sim TN(\mu_t^c, 0.0005)$.[35] We then seek to restrict the time-variation in μ_t^F and μ_t^c, in order to find a parameterization that is as parsimonious as possible. After estimating more parsimonious parameterizations which were rejected by the J-test, the following step functions delivered a model that fits the data:

$$\mu_t^F = \begin{cases} \mu_1^F & \text{if } t \leq \text{CD-2} \\ \mu_{CD-1}^F & \text{if } t = CD - 1 \\ \mu_{CD}^F & \text{if } t = CD \\ \mu_2^F & \text{if t} \geq CD + 1 \end{cases} \tag{36}$$

$$\mu_t^c = \begin{cases} \mu_1^c & \text{if t} \leq CD - 2 \\ \mu_{CD-1}^c & \text{if } t = CD - 1 \\ \mu_{CD}^c & \text{if } t = CD \\ \mu_{IPCA15}^c & \text{if } t = CD + 1 \\ \mu_2^c & \text{if t} \geq CD + 2 \end{cases} \tag{37}$$

The parameters of this model are $\theta = (\mu_1^F, \mu_{CD-1}^F, \mu_{CD}^F, \mu_2^F, \mu_1^c, \mu_{CD-1}^c, \mu_{CD}^c, \mu_{IPCA15}^c, \mu_2^c, \phi, \sigma_\varepsilon)$. The model-implied moments are obtained as in the restricted baseline model except that on day $t = 1, ..., T$ of the month, each agent receives a random draw of the costs c_{it}^F from a $TN(\mu_t^F, 0.0001)$ and c_{it} from a $TN(\mu_t^c, 0.0005)$. If the cost draws satisfy the condition in (20), with $\bar{C} = 1$, agent i is an updater and her $MSFE_{it}$ is given by the first line of (21). Agents who don't update keep their previous $MSFE$, $MSFE_{it-1}$.

The estimation results for the baseline model are in Table 6 and Figure 11 shows the fit of this model. Note that in the baseline model the magnitude of μ_t^F and μ_t^c is not meaningful as these parameters are only identified up to scale, but we can draw conclusions about their time variation around the contest day. The model passes the J-test of overidentifying restrictions with a p-value of 0.09, and, in unreported results, we find that further restrictions on the parameters make the model no longer pass the J-test. We thus conclude that time variation in both the fixed cost of updating and in the information processing cost is necessary to explain the patterns in the data. In particular, the contest induces a sizeable reduction in the cost of processing information relative to other days.

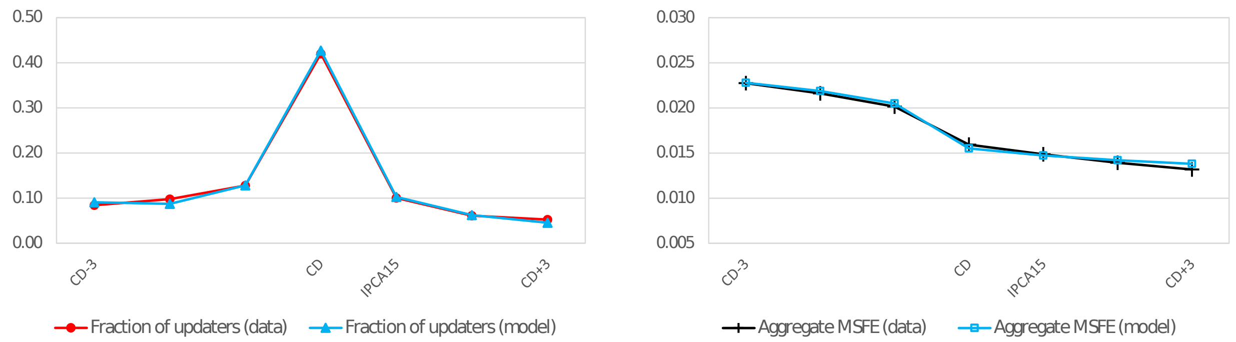
[35] The variances of these distributions are not identifiable and thus arbitrary. The reported values worked well in the simulations.

Table 6. Baseline Model Estimation

Parameter	Estimate	Standard Error	Parameter	Estimate	Standard Error
μ_1^F	1.00012	3.1E-05	μ_1^c	0.00145	2.2E-07
μ_{CD-1}^F	1.00010	1.1E-04	μ_{CD-1}^c	0.00131	6.8E-07
μ_{CD}^F	1.00002	4.0E-05	μ_{CD}^c	0.00040	3.8E-07
μ_2^F	1.00009	7.4E-05	μ_{IPCA15}^c	0.00138	3.3E-07
ϕ	0.90123	1.6E-07	μ_2^c	0.00156	3.7E-07
σ_ε	0.00556	5.2E-06			

Note: p-value of the J-test = 0.0963.

Figure 11. Baseline Model Predictions versus Data. Fraction of Updaters (Left) and Aggregate MSFE (Right)



B.2 Sticky Model

To establish a direct comparison with the baseline model, we consider the same parameterization for the fixed cost: $\bar{C} = 1$ and $C_{it}^F \sim TN(\mu_t^F, 0.0001)$, with μ_t^F as in (36). The parameters of the sticky information model are thus $\theta = (\mu_1^F, \mu_{CD-1}^F, \mu_{CD}^F, \mu_2^F, \phi, \sigma_\varepsilon)$. The model-implied moments are obtained as in the restricted baseline model, except that on day t an agent receives a random draw only of the fixed cost C_{it}^F from a $TN(\mu_t^F, 0.0001)$. Agent i is an updater if the draw satisfies the condition in (22), with $\bar{C} = 1$, and the $MSFE$ evolution is given by (23).

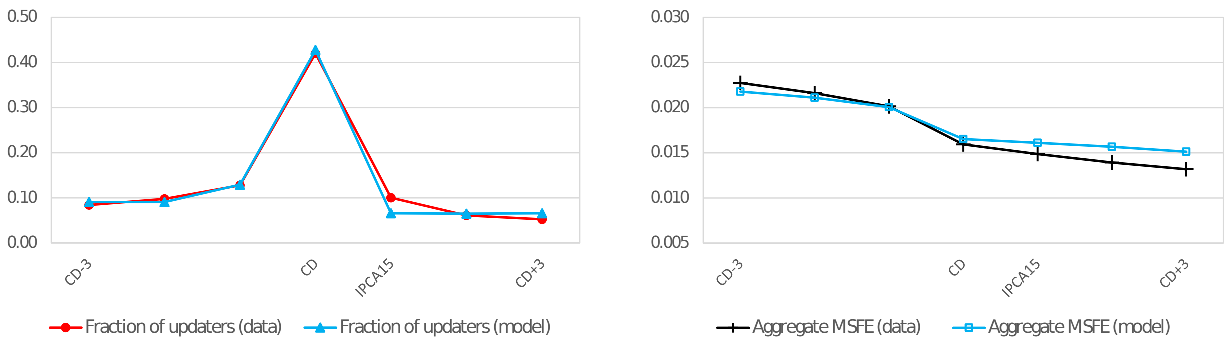
The estimation results for the sticky model are in Table 7 and Figure 12 shows the fit of this model. The sticky model is rejected by the J-test for overidentifying restriction and does not fit the patterns in the data.

Table 7. Sticky Model Estimation

Parameter	Estimate	Standard Error
μ_1^F	1.00013	3.8E-05
μ_{CD-1}^F	1.00011	1.8E-04
μ_{CD}^F	1.00002	4.7E-05
μ_2^F	1.00015	6.3E-05
ϕ	0.65134	9.3E-08
σ_ε	0.01316	1.8E-06

Note: p-value of the J-test = 6.44E-04.

Figure 12. Sticky Model Predictions versus Data. Fraction of Updaters (Left) and Aggregate MSFE (Right)



B.3 Noisy Model

To establish a direct comparison with the baseline model, we consider the same parameterization for the information processing cost: $c_{it} \sim TN(\mu_t^c, 0.0005)$, with μ_t^c as in (37). The parameters of the endogenous noisy information model are thus $\theta = (\mu_1^c, \mu_{CD-1}^c, \mu_{CD}^c, \mu_{IPCA15}^c, \mu_2^c, \phi, \sigma_\varepsilon)$. The model-implied moments are obtained as in the restricted baseline model, except that on day t an agent receives a random draw only of the information processing cost c_{it} from a $TN(\mu_t^c, 0.0005)$. Agent i is an updater if the draw satisfies the condition in (25), and the $MSFE$ evolution is given by (26).

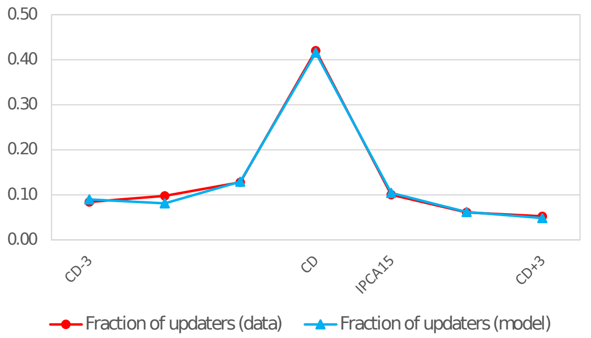
The estimation results for the noisy model are in Table 8 and Figure 13 shows the fit of this model. Both show that the model is rejected by the J-test and is clearly unable to fit the patterns in the data.

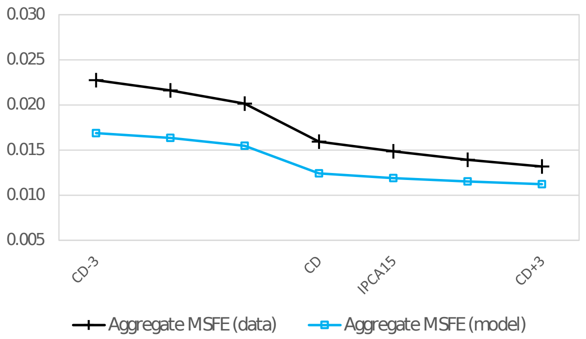
Table 8. Noisy Model Estimation

Parameter	Estimate	Standard Error
μ_1^c	0.00200	3.7E-05
μ_{CD-1}^c	0.00182	7.0E-06
μ_{CD}^c	0.00130	1.0E-05
μ_{IPCA15}^c	0.00178	7.3E-06
μ_2^c	0.00185	7.0E-06
ϕ	0.83908	1.5E-08
σ_ε	0.00634	1.1E-06

Note: p-value of the J-test = 1.41E-05.

Figure 13. Noisy Model Predictions versus Data. Fraction of Updaters (Left) and Aggregate MSFE (Right)





References

Afrouzi, H. (2019): "Strategic inattention, inflation dynamics and the non-neutrality of money," *Columbia University (unpublished paper)*.

Amemiya, T. and R. Y. Wu (1972): "The effect of aggregation on prediction in the autoregressive model," *Journal of the American Statistical Association*, 67, 628–632.

Caplin, A. and M. Dean (2013): "Behavioral implications of rational inattention with shannon entropy," Tech. rep., National Bureau of Economic Research.

——— (2015): "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 105, 2183–2203.

Caplin, A., M. Dean, and J. Leahy (2017): "Rationally inattentive behavior: Characterizing and generalizing Shannon entropy," Tech. rep., National Bureau of Economic Research.

Caplin, A., J. Leahy, and F. Matejka (2016): "Rational Inattention and Inference from Market Share Data," *Working Paper.*

Cavallo, A., G. Cruces, and R. Perez-Truglia (2017): "Inflation expectations, learning, and supermarket prices: Evidence from survey experiments," *American Economic Journal: Macroeconomics*, 9, 1–35.

Cheremukhin, A., A. Popova, A. Tutino, et al. (2011): "Experimental evidence on rational inattention," *Federal Reserve Bank of Dallas Working Paper*, 1112.

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

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

Csaba, D. (2018): "Attentional Complements," Tech. rep., Working Paper, NYU.

Dean, M. and N. Neligh (2017): "Experimental tests of rational inattention," Tech. rep., Working Paper, Columbia University.

DellaVigna, S. and D. Pope (2018): "What motivates effort? Evidence and expert forecasts," *The Review of Economic Studies*, 85, 1029–1069.

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

Fuhrer, J. C. (2018): "Intrinsic expectations persistence: Evidence from professional and household survey expectations," *FRB of Boston Working Paper.*

Gabaix, X. (2019): "Behavioral inattention," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, 261–343.

Gertler, M. and J. Leahy (2008): "A Phillips curve with an Ss foundation," *Journal of Political Economy*, 116, 533–572.

Giannone, D., L. Reichlin, and D. Small (2008): "Nowcasting: The real-time informational content of macroeconomic data," *Journal of Monetary Economics*, 55, 665–676.

Glaeser, E. L., A. Hillis, S. D. Kominers, and M. Luca (2016): "Crowdsourcing city government: Using tournaments to improve inspection accuracy," *American Economic Review*, 106, 114–18.

Gourieroux, C. and A. Monfort (1996): *Simulation-Based Econometric Methods*, Oxford University Press.

Hébert, B. and M. Woodford (2017): "Rational inattention and sequential information sampling," Tech. rep., National Bureau of Economic Research.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2016): "A rational theory of mutual funds' attention allocation," *Econometrica*, 84, 571–626.

Lazear, E. P. (2000): "Performance Pay and Productivity," *The American Economic Review*, 90, 1346–1361.

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

Maćkowiak, B. and M. Wiederholt (2015): "Business cycle dynamics under rational inattention," *The Review of Economic Studies*, 82, 1502–1532.

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

Marinovic, I., M. Ottaviani, and P. N. Sørensen (2013): "Forecasters' objectives and strategies," *Handbook of economic forecasting*, 2, 691–720.

Marques, A. B. C. (2013): "Central Bank of Brazil's market expectations system: a tool for monetary policy," *IFC Bulletins - Bank for International Settlements*, 36, 304–324.

Martin, D. (2016): "Rational inattention in games: experimental evidence," Tech. rep.

MIDRIGAN, V. (2011): "Menu costs, multiproduct firms, and aggregate fluctuations," *Econometrica*, 79, 1139–1180.

OTTAVIANI, M. AND P. SØRENSEN (2003): "Forecasting and rank order contests," *University of Copenhagen.*

POMATTO, L., P. STRACK, AND O. TAMUZ (2020): "The cost of information," *American Economic Review, forthcoming.*

REEVE, J., B. C. OLSON, AND S. G. COLE (1985): "Motivation and performance: Two consequences of winning and losing in competition," *Motivation and Emotion*, 9, 291–298.

RUGE-MURCIA, F. (2012): "Estimating nonlinear DSGE models by the simulated method of moments: With an application to business cycles," *Journal of Economic Dynamics and Control*, 36, 914–938.

SHEARER, B. (2004): "Piece rates, fixed wages and incentives: Evidence from a field experiment," *The Review of Economic Studies*, 71, 513–534.

SHORROCKS, A. F. (1978): "The measurement of mobility," *Econometrica: Journal of the Econometric Society*, 1013–1024.

SIMS, C. A. (2003): "Implications of rational inattention," *Journal of Monetary Economics*, 50, 665–690.

STEINER, J., C. STEWART, AND F. MATĚJKA (2017): "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," *Econometrica*, 85, 521–553.

SYVERSON, C. (2011): "What determines productivity?" *Journal of Economic Literature*, 49, 326–365.

TETLOCK, P. E. AND D. GARDNER (2016): *Superforecasting: The art and science of prediction*, Random House.

WIEDERHOLT, M. (2010): "Rational inattention," *The New Palgrave Dictionary of Economics (Online Edition ed.).*

Woodford, M. (2009): "Information-constrained state-dependent pricing," *Journal of Monetary Economics*, 56, S100–S124.

——— (2013): "Macroeconomic analysis without the rational expectations hypothesis," Tech. rep., National Bureau of Economic Research.