# Why Are Inflation Forecasts Sticky? Theory and Application to France and Germany*

Frédérique Bec,[a] Raouf Boucekkine,[b] and Caroline Jardet[c]
[a]Thema, CY Cergy Paris University, and CREST, Paris
[b]Centre for Unframed Thinking (CUT),
Rennes School of Business, France
[c]Banque de France, DGSEI-DECI, Paris

This paper proposes to adapt the model of pricing decisions developed by Alvarez, Lippi, and Paciello (2011) to the decision process of forecasters. The model features both a fixed cost of announcing a revised forecast and a fixed cost of updating the information set and adapting the forecast accordingly. Basically, the former fixed communication costs determine state dependence, which implies that the forecaster changes its forecast only when it is far enough from the optimal forecast, i.e., beyond a fixed threshold; the latter fixed information costs determine time dependence, which implies that the forecaster updates its information set only every other $T$ periods, where $T$ is optimally chosen. We show that survey data of inflation forecast updates as well as the last known monthly inflation rates can be used to estimate the threshold implied by the theoretical model. This threshold estimate is then crucial to uncover the existence of both types of costs as well as an upper bound of

*We are grateful to two anonymous referees and the Editor (Òscar Jordà) for their helpful comments and suggestions. Previous versions of this work have also benefited from many fruitful remarks by Anurag Banerjee, Mélika Ben Salem, Guillaume Chevillon, Annabelle De Gaye, Neil Ericsson, Fabian Gouret, David Hendry, Hervé Le Bihan, Francesco Lippi, Sarah Mouabbi, Muriel Roger, Pierre Sicsic, and Mark Watson, as well as by participants at the CIRET 2016 conference in Copenhagen, the 8th French Econometrics Conference, the 25th Symposium of SNDE in Cergy-Pontoise, the Durham University Economics Department's seminar, the 19th OxMetrics Conference in Paris, and the Banque De France seminar. F. Bec acknowledges support by Labex MME-DII (ANR-11-LBX-0023-01). The authors thank Isabelle Capelle and Céline Garnier for helpful research assistance. Of course, all remaining errors are ours. This paper reflects our opinion and does not necessarily express the views of the Banque de France. Author e-mails: bec@ensae.fr; raouf.boucekkine@rennes-sb.com; caroline.jardet@banque-france.fr.

the optimal time between two information observations. French and German data suggest that the maximum optimal time to next observation is six months, while the observation cost is at most twice as large as the communication cost.

JEL Codes: C23, D8, E31.

## 1. Introduction

Recent empirical evidence from forecast surveys panel data reveals forecasts stickiness: Coibion (2010), Coibion and Gorodnichenko (2012, 2015), Andrade and Le Bihan (2013), Dovern (2013), Loungani, Stekler, and Tamirisa (2013), and Dovern et al. (2015), among others, all point to the evidence that forecasters fail to systematically update their forecasts and/or their information set. This is also the case in our monthly forecasts survey data set, where only a proportion of around 43 percent of professional forecasters update their forecast between two consecutive months. Similar orders of magnitude are found from monthly survey data by Dovern (2013) and Dovern et al. (2015) for forecasts of real GDP growth. Using quarterly survey data from the European Survey of Professional Forecasters, Andrade and Le Bihan (2013) find that only 75 percent of these forecasters update their one-year or two-year forecasts each quarter. This figure is also very close to the proportion of forecasters who update every three months in our sample. As stressed by these authors, even if the proportion of updaters is much larger from micro-level survey data than from aggregate forecasts data, it still remains remarkable that professionals do not systematically update their forecasts.

Models with information rigidities have been proposed to solve this forecast stickiness puzzle. Basically, they can be divided into two strands of research: one which considers information stickiness and another one which considers information noisiness. In the first branch, made popular by Mankiw and Reis (2002), agents face fixed costs to acquiring and processing information and therefore update their information sets infrequently. In the second branch, initiated by Woodford (2002) and Sims (2003), agents are continuously updating their information set, but they receive noisy information about underlying macroeconomic conditions. However, both types of information rigidity models fail to reproduce the degree of stickiness observed in forecasts data (see, e.g., Andrade and Le Bihan 2013). Furthermore, they imply a decision to update the forecast which is mainly time dependent, in contradiction with some empirical evidence of state dependence found in the data by, e.g., Dovern (2013), Loungani, Stekler, and Tamirisa (2013), or Coibion and Gorodnichenko (2015).[1] This state dependence is confirmed by the results of the special questionnaire for participants in the European Central Bank (ECB) Survey of Professional Forecasters (SPF) conducted and published every five years by ECB. This special survey aims at exploring the forecast processes and methodologies underlying the contributions made to the regular quarterly SPF. In its third edition,[2] published in 2019, an average of 73 percent of respondents mentioned "internal timetable" as the determinant of the timing of their forecast updates. Many forecasters, however, reported that data releases are also an important trigger of a full update of their (short-term) forecasts. Some of them also provided qualitative comments explaining that if new data materially affected their view on the economy, they would react by updating their forecasts sooner than their regular timetable might imply.

Our first contribution to this literature is to introduce and evaluate what we consider as the missing ingredient in these models: a communication cost. Indeed, a forecaster can pay the observation cost (and refresh his/her forecast) but choose not to announce or communicate officially this new forecast due to the cost it involves in terms of meetings, reports, interviews, discussions with the hierarchy as well as with the customers, etc. We claim that taking into account this additional cost, as small as it might be, reinforces the forecast stickiness and generates state dependence in the decision rules.

We first elaborate on the model developed by Alvarez, Lippi, and Paciello (2011) to design a theoretical framework allowing for

---

[1] Evidence of both state and time dependence is also found by Magnani, Gorry, and Oprea (2016) from laboratory decision experiment data.

[2] See the document "Results of the Third Special Questionnaire for Participants in the ECB Survey of Professional Forecasters," downloadable from the ECB website: https://www.ecb.europa.eu/stats/ecb_surveys/survey_of_professional_forecasters.

optimal inflation forecasting in the presence of both observation and communication costs. Even though our framework is similar to Alvarez, Lippi, and Paciello (2011), there are two main differences. The first one comes from the fact that the timing of the next observation is bounded by the forecast horizon. There is no upper bound of this sort in Alvarez, Lippi, and Paciello (2011), because their problem is about price setters who do not bear any time horizon constraint. However, this is a pure control constraint, the simplest one from the mathematical point of view, and we show in our empirical section that it is never binding in practice.[3] The second one has to do with the underlying stochastic structure. In Alvarez, Lippi, and Paciello (2011), the latter concerns nominal price law of motion, while in our setting, we focus on the inflation forecast law of motion. Therefore, the underlying stochastic structure has to be adapted.

Our original setup can generate optimal forecast stickiness under rational expectations hypothesis. In this model, to make a long story short, fixed communication costs generate state dependence, which implies that the forecaster announces his/her forecast change only when it is large enough, i.e., larger than a given threshold; fixed information costs cause time dependence, which implies that the forecaster updates its information set only every other $T$ periods, where $T$ is optimally chosen by the forecaster. Finally, we also show how the threshold can be used to evaluate the relative importance of observation and communication costs.

A second contribution of this paper is to provide an estimate of the latter. To this end we use survey data of inflation forecast updates as well as last-known monthly inflation rates to estimate the threshold implied by the theoretical model. As will be shown below, this threshold estimate is in turn crucial to uncover the existence of both types of costs as well as an upper bound of the optimal time between two information observations. The main difficulty is that we only observe forecast changes which are announced—in other words, the communication activity. Indeed, no direct data on the forecasters' observation activity is available, which rules out the direct evaluation of its cost. To circumvent this issue, a logit model is fitted to

[3]Without loss of generality, we shall omit it hereafter to unburden the presentation.

the forecast's revision probability as a function of the time elapsed since the last revision as well as a proxy for the forecast gap to capture the state-dependent dimension. In this model, the threshold is endogenously determined as the one maximizing the likelihood function. Our French and German data suggest that the optimal time to next observation is about six months, while the observation cost is at most twice as large as the update communication cost.

The theoretical model is sketched in Section 2, along with its main implications and our empirical testing strategy. Section 3 presents the data. Section 4 describes the threshold estimation method and reports the results of the estimated threshold logit models. Their implications regarding the relevance of our theoretical model are then discussed, while Section 5 concludes.

## 2. A Theoretical Model of Forecast Formation

This section proposes a non-technical description of our model, while more details can be found in the appendix. After a brief presentation of the main assumptions, we proceed with the description of the decision maker's behavior. We then turn to the model's properties and finally present our testing strategy.

### *2.1 Preamble*

The following notation will be used throughout the paper. Time is continuous and $t \in (0, 1)$. The object to forecast is the annual inflation rate, $\pi(1)$. $\pi_f(t)$ is its forecast value at time $t$, with $0 < t \leq 1$, of the announced forecast of $\pi(1)$. $\pi^*_f(t)$ denotes the forecast target, namely the optimal forecast which would prevail if the information set was up to date in a frictionless setup. The "forecast gap," denoted $\tilde{\pi}_f(t)$, is defined as the difference $\pi_f(t) - \pi^*_f(t)$.

Assumption 1. *The law of motion of the target is subject to i.i.d. innovations, i.e., a Brownian motion (BM) without drift, so in continuous time that is $d\pi^*(t) = \sigma dW(t)$ where $W(t)$ is a standard BM and $\sigma^2$ the variance per unit of time.*

This assumption is in line with, e.g., Stock and Watson (2007)'s finding that the univariate inflation process is well characterized by a unit root. Note also that we depart from Alvarez, Lippi, and Paciello (2011), who assume that the law of motion of the targeted price has a time drift ($\mu \geq 0$).[4] While non-zero drifts make sense in their framework (as they are concerned with price setting and the time drift corresponds indeed to inflation), they do not in ours.

Assumption 2. *The instantaneous loss faced by the professional forecaster is a quadratic function of the forecast gap* $\tilde{\pi}_f(t)^2$.

Thus, the more the forecast deviates from the optimal forecast, the greater the loss incurred by the forecaster or, more generally, for the institution from which the forecast originates. This loss can be interpreted in several ways. Firstly, it may represent a loss in terms of reputation that the institution producing the forecast incurs when providing sub-optimal forecast. Of course, judging whether a forecast is optimal or not is not obvious as it is not observable. While agents do not expect the forecast to be exactly equal to the true ex post realized value, they however may expect that it should be close to it. So agents should be able to form an opinion on the quality/optimality of an institution's forecasts by looking at its ex post forecast errors. Therefore, an institution that would provide forecasts very different to ex post realization, notably in comparison with other institutions, could suffer from a loss in credibility and reputation. Secondly, one may think that these forecasts feed partly economic and financial decisions made by these institutions. Therefore, a sub-optimal forecast could lead to sub-optimal decision-making and induce an additional loss (financial and/or reputational) for the institution.

Assumption 3. *Each time the forecaster observes the state* $\pi_f^*(t)$*, it incurs a fixed observation cost* $\theta \in R^+$*. This cost includes, for instance, the time it takes in terms of hours worked to collect and process all relevant information, and the time it takes to implement forecasting models.*

Assumption 4. *It is costly to communicate about the revised forecast, to announce it "officially." Each revision announcement incurs*

[4]See Equation (1) in Alvarez, Lippi, and Paciello (2011, p. 1917).

*a fixed communication cost $\psi \in R^+$. This cost includes, e.g., the extra writing and publication of reports involved by the external communication process with the public, customers, and media. Indeed it is necessary to justify thoroughly the revision of the forecast, even more so if the forecaster is the only one who publishes a significant forecast revision. Actually, this revision is likely to attract a lot of attention. In addition, we assume that this cost also includes the loss of forecaster's credibility implied by too frequent forecast revisions.*

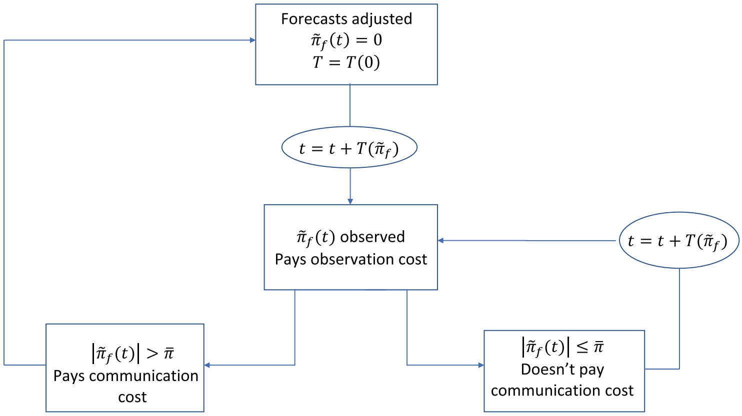
## *2.2 Sketch of the Model*

The problem considered here is the one of a decisionmaker (a forecasting unit) subject to the observation and communication costs, as defined in Assumptions 3 and 4 above. This is an adaptation of Alvarez, Lippi, and Paciello (2011)'s model of firms price setting to the forecast formation matter.[5] More precisely, the forecaster tracks a target, the optimal forecast $\pi_f^*(t)$, which by Assumption 1 is a BM. By Assumption 2, the forecaster's losses are quadratic in the forecast gap $\tilde{\pi}_f(t)$. Figure 1 illustrates the forecaster's problem, where $\bar{\pi} \in \mathbb{R}^+$ is a threshold value and $T(\cdot)$ is the optimal time until next observation as a function of the forecast gap $\tilde{\pi}_f(t)$—to be described later in Figure 2.

In this setting, an observation is defined as the forecasters' act to pay a fixed cost $\theta$ and retrieve the observation on the current $\pi_f^*(t)$ and hence on $\tilde{\pi}_f(t)$ too. This is illustrated in the middle of Figure 1. The forecaster's decision then depends on the size of the forecast gap, measured by $|\tilde{\pi}_f(t)|$, compared with the threshold $\bar{\pi}$. If the size of the gap is greater than $\bar{\pi}$, they choose to pay another fixed cost $\psi$ in order to communicate on a forecast revision, namely to adjust $\pi_f(t)$, as indicated in the left part of Figure 1. Since the target has no drift (i.e., $\pi_f^*(t)$ is a unit root until the next forecast revision announcement), it is straightforward that the optimal target revision occurs at the same time as the actual observation of the target. In other words, the forecast gap is set back to zero as noted in the top box of Figure 1.

[5]It turns out to be a special case with no trend in the forecast target and, by contrast with Alvarez, Lippi, and Paciello (2011)'s model, our problem has a finite horizon.

**Figure 1. Forecaster's Problem**

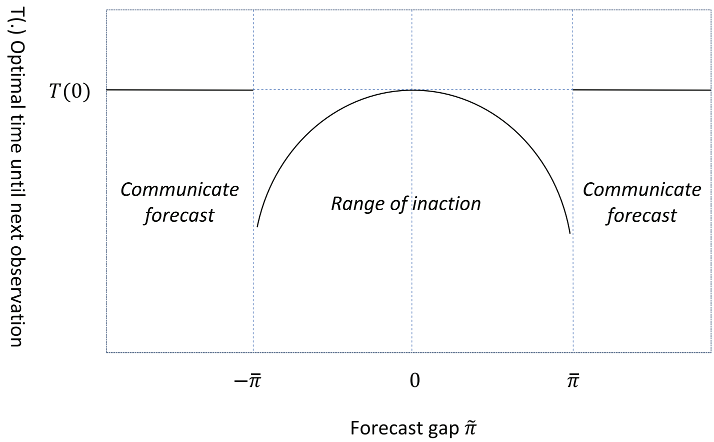



Otherwise, if the size of the gap is smaller than $\bar{\pi}$, they do not communicate, as it is not worth paying the fixed cost $\psi$. Instead, they wait until they choose to observe again. Therefore, the presence of a communication cost implies that not all observations are followed by a forecast revision announcement. Technically, conditionally on observing the gap, there is an Ss interval for gaps where inaction is optimal, here defined by $(-\bar{\pi}, \bar{\pi})$. This is illustrated in the bottom right box of Figure 1. In this case, the decisionmaker has to plan optimally the time to next observation by choosing $T(\tilde{\pi}_f)$.

As illustrated in Figure 2, which represents this optimal time until next observation as a function of the forecast gap, when the forecaster pays the communication cost and closes the gap (so that $\tilde{\pi}_f(t) = 0$), the time to next observation reaches its maximum, $T(0)$. Then, as innovations accumulate in $\pi^*_f(t)$, the forecast gap approaches one of the boundaries of the inaction set, either $\bar{\pi}$ or $-\bar{\pi}$, on the x-axis of Figure 2.[6] As the forecast gap widens, a new observation by the forecaster becomes more and more likely. As a consequence, the relationship between the optimal time to next observation and the forecast gap within the inaction set is hump-shaped. As soon as one of these boundaries is hit, the communication

[6]The symmetry around zero stems from the quadratic loss function of the forecaster, Assumption 2.

**Figure 2. Time Elapsed Until Next Observation**




cost is paid and the gap is closed, which drives the optimal time to next observation back to its maximum $T(0)$. So the optimal decision involves both a time-dependent component (the time elapsed since the last observation) and a state-dependent decision (the forecast gap at the last observation).

## *2.3 The Model's Main Properties*

As exposed to a larger extent in Appendix A, the optimal control problem of the forecaster is to choose the time until next observation, $0 < T \leq 1$, as well as the number and size of forecast revisions between two observations. This model's main properties are summarized below. More properties are derived in the appendix.

PROPERTY 1. *Optimal number and timing of forecast updates announcements:*

*(i) There is at most one forecast revision communication between two observations, and it occurs if and only if $|\tilde{\pi}_f| > \overline{\pi}$, i.e., when the forecast gap if large enough to compensate for the communication cost (bottom left box of Figure 1).*

*(ii) If there is a forecast revision communication, it is immediate after observation and it closes the gap:* $\pi_f = \pi_f^*$*, so that* $\tilde{\pi}_f = 0$ *(top box of Figure 1).*

See Appendix B for more details.

From Property 1, it follows straightforwardly that a forecast revision communication will take place if and only if the observed forecast gap is greater than the threshold $\overline{\pi}$, in absolute value. Let $\lambda(\tilde{\pi}_f)$ denote the indicator variable, which takes on value one if a forecast revision communication occurs and zero otherwise. Accordingly, it is defined as a function of the threshold $\overline{\pi}$ in Equation (1) below.

PROPERTY 2. *Optimal forecast revision communication occurrences:*

$$\lambda(\tilde{\pi}_f) = \begin{cases} 0 & \text{if } \tilde{\pi}_f \in (-\overline{\pi}, \overline{\pi}) \\ 1 & \text{otherwise.} \end{cases} \tag{1}$$

See Appendix A for more details. So, $\lambda(\tilde{\pi}_f)$ is state dependent. Yet, since $\tilde{\pi}_f$ is a function of $T(0)$, the maximum time elapsed until next observation as represented in Figure 2—see Proposition D.1 in Appendix D—$\lambda(\tilde{\pi}_f)$ is time dependent too. Indeed, the forecast gap is more likely to be bridged as the time elapsed until next observation approaches its maximum because, as long as it is not corrected, it behaves like a Brownian motion and, as such, it accumulates innovations—see Equation (A.2) in Appendix A. This in turn makes the forecast gap more likely to cross one of the inaction zone boundaries.

Then, the optimal time between two observations, $T(\tilde{\pi}_f)$, can be shown to be also a threshold regime-switching process with the forecast gap as the switching variable. The optimal rule for the time of the next observation of the forecast gap is given below.

PROPERTY 3. *Optimal time between two observations:*

$$T(\tilde{\pi}_f) = \begin{cases} T(0) - \left(\frac{\tilde{\pi}_f}{\sigma}\right)^2 + o(|\tilde{\pi}_f^3|) & \text{if } \tilde{\pi}_f \in (-\overline{\pi}, \overline{\pi}) \\ T(0), \quad \text{otherwise.} \end{cases} \tag{2}$$

See Appendix C for more details.

Moreover, as can be seen above, within the inaction set where $\tilde{\pi}_f \in (-\overline{\pi}, \overline{\pi})$, the closer to the threshold $|\overline{\pi}|$, the shorter the time to next observation: This is the theoretical ground of Figure 2.

Finally, as noticed by Alvarez, Lippi, and Paciello (2011), there is a rather simple approximate relation between the ratio of the observation to communication frequencies and the ratio of $\psi$ to $\theta$.[7]

Property 4. *The relative costs approximate evaluation:*

$$\mu = \frac{\psi}{\theta} \simeq \left(\frac{n_o}{n_c} - 1\right)^2, \tag{3}$$

*where $n_o$ and $n_c$ denote the observation and communication frequencies, respectively.*

Equation (3) is instrumental to ground the relevance of our approach, as the coexistence of both observation and communication costs implies that the costs ratio is strictly positive and finite. Indeed, if there is no communication cost, then this ratio is zero, while if there is no observation cost, it goes to infinity. The relation between the costs ratio $\mu$ and the ratio of frequencies $n_o/n_c$ is qualitatively intuitive. When the ratio $\mu$ is zero, which corresponds to the zero communication cost case, the forecaster will communicate at each observation, leading to $n_o/n_c = 1$. As $\mu$ rises, the ratio $n_o/n_c$ also goes up: When the communication cost increases more than the observation's, the forecaster will observe more frequently than she will communicate, implying a rising frequencies ratio $n_o/n_c$. On a more quantitative ground, the functional relation between $\mu$ and the ratio of frequencies is approximately quadratic within the analytical framework of Alvarez, Lippi, and Paciello (2011). In other words, $n_o/n_c - 1$ is well approximated by the square root of $\mu$ (so the elasticity with respect to the ratio of costs is near 1/2). As shown in Alvarez, Lippi, and Paciello (2011), the lower the communication cost, the better the approximation.

As stressed above, the costs ratio is zero if the absence of communication cost, and it tends to infinity if there is no observation cost.

[7] See the discussion in Section VI.A, pp. 1936–38, in Alvarez, Lippi, and Paciello (2011).

In practice, the latter case should not be observed, as it requires that either $n_c = 0$ or $n_o \to \infty$. Nevertheless, an arbitrarily large value of the ratio could be found for arbitrarily low (respectively, large) value of $n_c$ (respectively, $n_o$) and would question the presence of an observation cost. By contrast, a ratio close to zero cannot be ruled out a priori from an empirical point of view, as $n_c$ could be found equal to $n_0$, hence questioning the presence of a communication cost. Therefore, we believe that the empirical evaluation of this costs ratio is important, as it is likely to provide information about the relevance of our theoretical setting.

We will show here below how these properties can be used to estimate the inflation gap threshold which triggers action, $\bar{\pi}$, along with the maximum optimal time to next observation, $T(0)$. In turn, as explained in the next subsection, this allows to infer the observation to communication costs ratio.

### *2.4 Our Estimation Strategy*

Our estimation strategy runs in two main steps.

Firstly, we fit a threshold logit regression based on Equation (1), in order to estimate the forecast update communication probability as a function of the forecast gap and time dependence. However, as the forecast gap is not observable, we need to use an "observable" proxy instead in order to achieve estimation of the model. The choice of this proxy will be discussed in Section 3. Finally, this first step allow us to obtain the maximum-likelihood estimate of the threshold, $\bar{\pi}$.

Secondly, with the threshold estimate from the first step, it is then possible to check the existence of both types of costs grounding time and state dependence of the probability to announce a forecast revision. The computation of the communication frequency, $n_c$ in Equation (3), is straightforward: It is simply the share of forecasters who update in the sample. The computation of the observation frequency, $n_o$, is more tricky. This is the share of observers, which includes the ones who communicate, i.e., $n_c$,[8] and the ones who observe and do not communicate a revised forecast.

[8] One can safely assume that when a forecast update is announced, it has first been preceded by an observation.

It is necessary to circumvent the main caveat of our data: The observation activity of the forecasters is not observable.[9] Hence, it is impossible to disentangle among the ones who do not communicate the ones who have observed and the ones who haven't. However we propose to compute $n_o$ by considering the share of the ones who do not update, conditionally to $\tilde{\pi}_f < \overline{\pi}$, i.e., the forecast gap is lesser than the threshold. This is where an estimate of $\overline{\pi}$ is clearly required. In fact, this measure of $n_0$ should overestimate the true value of the observation frequency because in addition to those who observe but do not adjust for good reasons (the ones we want to count), this population also includes those who just do not observe. Unfortunately, these two categories are observationally equivalent with the data at hand: They both correspond to $\lambda(\tilde{\pi}_f) = 0$ *and* $\tilde{\pi}_f < \overline{\pi}$. Consequently, our computation of $n_o$ will overestimate its true value.

## 3. Data

### *3.1 Forecasts Updates Data*

The empirical analysis below is based on forecasts of the annual inflation rate from the monthly survey data set compiled by Consensus Economics Inc. and made by private and public economic institutions such as banks and research institutes. Since Consensus Economics Inc. reports forecasts of the current and next calendar years, the data set is a three-dimensional panel: forecasters indexed by $i$, target years indexed by $t$, and horizons indexed by $h$. For each target year, the data set contains a sequence of 24 forecasts from each forecaster made from January of the year before the target year to December of the target year. The latter will be labeled $h = 0$ since this corresponds to nowcasting. Consequently, the former will correspond to $h = 23$.

Let $\pi_{i,t|h}$ denote agent $i$ forecast of year $t$ inflation rate, made $h$ months before December of year $t$. The unconditional probability of updating a forecast for horizon $h$, denoted $\lambda(i, h)$ after Dovern (2013), is given by the probability that a forecast for the same

[9] Indeed, from forecasts data, the only activity which is tracked is the forecast revision communication.

**Table 1. Descriptive Statistics of $\hat{\lambda}(h)$**

| | **Average** | **Min. [$h$]** | **Max. [$h$]** | **$h$ = 22** | **$h$ = 0** |
|---|---|---|---|---|---|
| France | 45.0% | 27.4% [16] | 63.3% [9] | 28.8% | 32.4% |
| Germany | 41.7% | 29.9% [19] | 61.5% [12] | 30.0% | 34.2% |

object, e.g., year $t$ inflation rate, is revised by institution $i$ between two consecutive forecast horizons: $Pr(\pi_{i,t|h} \neq \pi_{i,t|h+1})$. As noted by Dovern (2013), assuming that the probability is the same for each institution, it can be estimated for each horizon as

$$\hat{\lambda}(h) = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{N_{t|h}} \sum_{i=1}^{N_{t|h}} 1[\pi_{i,t|h} \neq \pi_{i,t|h+1}], \tag{4}$$

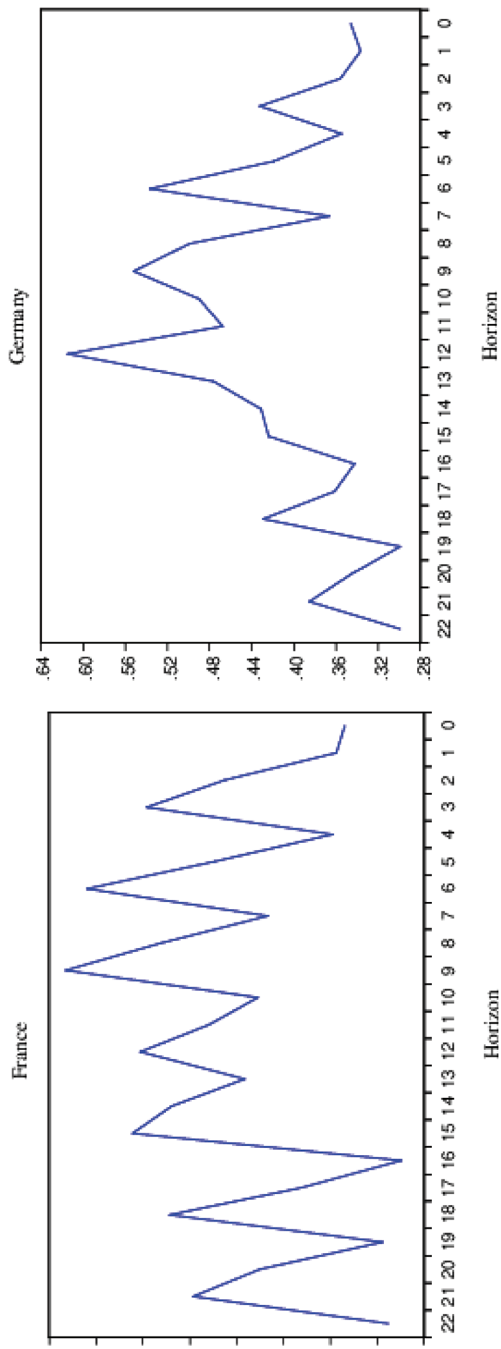
where $T$ is the total number of target years in the sample, $N_{t|h}$ is the number of observed forecasts for target year $t$ with forecast horizon $h$, and $1[.]$ is the indicator function which takes on value one if the condition into brackets is verified and zero otherwise.

Our data set includes yearly inflation rate forecasts for the current and next years, made monthly by individual professional forecasters since January 1998 for target years from 2000 to 2015. In the subsequent analysis, we consider country-specific forecast data. This guarantees indeed that the series to forecast, namely the national inflation rate, is homogenous across forecasting units. Moreover, this allows to reveal similarities and differences in forecasting behaviors across countries, if any. For France and Germany, the sample includes, respectively, 36 and 51 forecasters. Table 1 reports the main descriptive statistics of the average probabilities of updating between two consecutive months, the expression for which is given in Equation (4).[10] Figure 3 plots these update probabilities as a function of the forecast horizon.

These basic statistics look quite homogenous amongst countries. On average over all horizons, 41.7 percent (Germany) to 45 percent

[10]The panel is heavily unbalanced since a large part of the “individuals” have given their forecast in an irregular manner. Consequently, a significant part of forecasts observations could not be used to compute revisions because they are adjacent to missing values. As a result, we are left with a total of 4,990 usable points for France and 8,202 for Germany.

**Figure 3. Average Revision Probability as a Function of the Forecast Horizon**



(France) of forecasters revise their forecasts between two consecutive months. This result is slightly less than Dovern's (2013) average estimation of nearly 50 percent over 14 advanced economies: The European professional forecasters considered in our study are not the most attentive ones among advanced countries. Nevertheless, this implies a degree of information rigidity which is much lower than the one obtained from the aggregate (or average/median) forecast. For instance, Coibion and Gorodnichenko (2012) find that the average inflation forecast across 40 U.S. agents surveyed by SPF is updated every six to seven quarters.[11] Similarly, high degrees of information rigidity is found by Coibion (2010). Our data show that 43.3 percent of the professional forecasters update their forecasts as frequently as every month.[12]

As can be seen from Figure 3, there is no obvious linear trend in the revision probability as horizon increases or decreases, which differs from Dovern's [2013] finding based on GDP growth forecast survey data. It can be seen that attentiveness increases until $h = 12$ (Germany) and $h = 9$ (France) and decreases afterwards. The pattern of French update rate shows pronounced seasonality with clear quarterly peaks in March, June, September, and December. This could stem from two reasons. First, it is likely that the forecasting exercise is done every quarter instead of every month. Second, over the period considered in this analysis, the French National Institute of Statistics and Economics Studies releases its first estimate of quarterly real GDP growth 45 days after the end of the quarter, which is around the middle of the second month of each quarter. Since the monthly Consensus Economics survey is completed before the twelfth day of each month, this piece of information is not available when producing the forecast of the quarter's second month. It is incorporated in the quarter's third month instead, enhancing the revision probability. This quarterly pattern is also present in

[11] By contrast, in our French sample for instance, the update share reaches 78 percent when the probability of updating at least once over the last three months is considered.

[12] The relative frequencies of individual-specific unconditional probability of updating can be obtained by computing the average of $1[\pi_{i,t|h} \neq \pi_{i,t|h+1}]$ across target years and forecast horizons for each forecaster. To save space, descriptive statistics of these $\hat{\lambda}(i)$'s are not reported here but are available upon request. Note that their mean is again found to be close to 45 percent.

Germany, but to a lesser extent and mainly during the target year itself, i.e., for $h = 12$, 9, 6, and 3. These seasonal patterns will be captured by the inclusion of a horizon fixed effect in our regression equation.

### *3.2 Threshold and Explanatory Variables*

According to the model presented in Section 2, some variables must be allowed to switch from an inaction regime—defined by $\tilde{\pi}_f \in (-\bar{\pi}, \bar{\pi})$—to an action regime elsewhere. Let us gather in a vector denoted $X^S_{i,t}$ the variables allowed to switch. As discussed in Section 2—see Equation (2) above and Equation (D.9) in Appendix D—the dynamics of the time elapsed between two observations ($T(\tilde{\pi}_f)$) and of the forecast gap itself ($\tilde{\pi}_f$) are expected to switch according to the size of the forecast gap. In our empirical counterpart, the time dependence ($T(\tilde{\pi}_f)$) will be captured by a dummy variable denoted D(d=m), for $m = 2, 3 \ldots$, which is equal to one if the unit's last revision occurred $m$ months ago and zero otherwise. Since very few observations belong to each D(d=m) for $m \geq 8$, the latter are gathered into D(d≥8).[13]

Then, let us turn to the threshold variable, namely the forecast gap. Unfortunately, this is not observable, since the optimal forecast, from which it is evaluated, is unobservable. As mentioned earlier, to circumvent that issue, we use an "observable" proxy instead. This proxy should be available at the time of the forecast exercise. Hence, ex post realized annual inflation cannot be used for that purpose, as it is known only at the end of the target year. By contrast, the last release of monthly inflation rate is available, and even if the optimal forecast should depend on a lot of factors, we can reasonably assume that updates of optimal forecast should at least incorporate this new information. Under the strong assumption that forecasts of monthly inflation rate are constant, large values of realized monthly inflation rate should trigger large values of the forecast gap. Therefore, in the following, we consider the last known monthly inflation rate, seasonally adjusted and weighted by its mechanical contribution to the annual inflation rate forecast, as a proxy for the unobservable forecast gap. As it is released at the

[13]We have checked that this simplification does not affect our conclusions.

end of last month/very beginning of the current month (Eurostat release), it is considered with one month lag and denoted $\pi_m(-1)$ hereafter. From our theoretical model, we expect updates to be triggered only by large changes in the state variable: $|\pi_m(-1)|$. Accordingly, a switching function denoted $s_t$ is defined as the indicator function $s_t = 1(|\pi_m(-1)| > \hat{\bar{\pi}})$, taking on value one for absolute values of $\pi_m(-1)$ greater than the estimated threshold $\hat{\bar{\pi}}$ and zero otherwise. Finally, this allows to distinguish regressors in the outer regime, $s_t X_{i,t}^S$, and their analogues in the inaction band, $(1-s_t)X_{i,t}^S$, with $X_{i,t}^S = (D(d=1), D(d=2), \cdots, D(d=8), \pi_m(-1))$. As will be seen in the next section, the threshold will be estimated from a grid search set defined by the 25 percent and 75 percent quantiles of $|\pi_m(-1)|$ as its lower and upper bounds.

The variables with coefficients that are assumed to be the same across regimes, gathered in $X_{i,t}^0$, aim at controlling for various effects which have been shown to influence the forecast updating behavior in previous empirical studies.[14] Firstly, we control for the seasonal pattern of the forecast updates—noticed in Figure 3—by introducing a forecast horizon fixed effect,[15] denoted $c_h$. Then, in order to take into account a business cycle effect, we introduce a target year fixed effect, denoted $c_y$. Thirdly, we introduce the percentage of forecasting units which have revised their forecast last month so as to capture a "cascade" effect ($Sha(-1)$), as described, e.g., in Banerjee (1992) or Graham (1999). Fourthly, a variable called $|Dev(-1)|$, which measures the previous period deviation of the unit's forecast from the average in absolute value, is considered to capture a possible "herding" effect.[16] Finally, an individual fixed effect, $c_i$, is also introduced. As a total, on top of the horizon, target year, and individual fixed effects, the vector $X_{i,t}^0 = (Sha(-1), |Dev(-1)|)$ is introduced in the non-switching part of the regressors.

[14] See, for instance, Dovern (2013, 2015) and references therein.

[15] This horizon fixed effect completely captures a potential institutional pattern of updating within a given quarter, so that dummies indicating the second and third month of a quarter would be completely redundant. This is also the case for a dummy indicating August, when most people are having their summer vacation in the countries considered.

[16] See, e.g., Lamont (2002) or Pons-Novell (2003).

## 4. Empirical Testing of Our Model's Properties

### 4.1 The Threshold Logit Model

The empirical testing of our model's properties in terms of time and state dependence of the inflation forecast updates will rely on the estimation of threshold binary choice models for panel data. More precisely, the traditional random- or fixed-effects logit models are considered to analyze the conditional probability of forecasts revisions. Using a latent variable framework, where $\lambda^*_{i,t,h}$ is a continuous variable that is not observed, it may be written as

$$\lambda^*_{i,t,h} = s_t X^S_{i,t}\beta_{out} + (1 - s_t)X^S_{i,t}\beta_{in} + X^0_{i,t}\beta + c_i + c_y + c_h + u_{it} \quad (5)$$

$$s_t = 1(|\tilde{\pi}_f(t)| > \bar{\pi}) \quad (6)$$

$$\lambda_{i,t,h} = 1(\lambda^*_{i,t,h} > 0), \quad (7)$$

where $s_t$ is the regime-switching indicator function presented in the previous section. It is equal to one when the forecast gap is greater than a threshold $\bar{\pi}$ and zero otherwise. $X^S_{i,t}$ is the vector including the regressors which coefficients are allowed to switch across regimes, and $\beta_{out}$ and $\beta_{in}$ are the corresponding switching coefficients in the outer and inner regimes, respectively—the latter being the so-called inaction set defined by small $|\tilde{\pi}_f(t)|$'s. $\beta$ is the vector of non-switching coefficients corresponding to the explanatory variables gathered in $X^0_{i,t}$. $c_j$, for $j = i, y, h$, are unobserved individual, target year, and forecast horizon fixed effects, respectively. Then, $\lambda_{i,t,h}$ is an indicator for forecast updates and

$$Pr(\lambda_{i,t,h} = 1|X_{i,t}, c_i) = G(s_t X^S_{i,t}\beta^S_{out} + (1 - s_t)X^S_{i,t}\beta^S_{in} + X^0_{i,t}\beta + c_i + c_y + c_h),$$

where $G(.)$ is the cumulative distribution function (CDF) for the logit model and the standard normal CDF for the probit model. In the random-effect (RE) version, it is assumed that $(X^S_{i,t}, X^0_{i,t})$ and the $c_j$'s, for $j = i, y, h$, are independent and that the latter have a Gaussian distribution with zero mean and constant variance so that it is possible to estimate the model by maximum-likelihood techniques, integrating out the unobserved $c_j$'s. The assumption of independence between $(X^S_{i,t}, X^0_{i,t})$ and the $c_j$'s is relaxed in the fixed-effect (FE) version, and, in this case, the conditional maximum-likelihood estimator is used.[17] The threshold estimate $\hat{\tilde{\pi}}$ is obtained by grid search of all possible values of $|\tilde{\pi}_f(t)|$ over $\Gamma = [\gamma_{25\%}, \gamma_{75\%}]$, as the one which maximizes the log-likelihood of the threshold logit model described above. Here, $\gamma_{25\%}$ and $\gamma_{75\%}$ denote the 25 percent and 75 percent quantiles of $|\tilde{\pi}_f(t)|$.[18]

## *4.2 Estimation Results*

The estimation results are reported in Table 2. The models' estimates presented below include both horizon, target year, and individual fixed effects: The corresponding Hausman statistic p-value strongly rejects the null that both RE and FE models yields the same estimates.

The average probability of update predicted by the model is rather close to the observed one in both countries, even though slightly overestimated: It is found to be 46.6 percent in France and 44 percent in Germany, whereas their observed counterparts are 45 percent and 41.7 percent, respectively. So, this model correctly accounts for a slightly larger proportion of updaters in France compared with Germany.

Let us begin with the non-switching variables, at the bottom panel of the regressors list. It appears that the update probability is increased by the cascade effect (variable $Sha(-1)$), and particularly so in the German case where the estimated coefficient is 0.219, and it is significantly different from zero at the 1 percent level. In France, this estimated coefficient is 0.08, and it is significantly different from zero at the 10 percent level only. Contrarily, the herding

[17] All estimations have been performed using version 14.1 of Stata software. In order to estimate a robust variance–covariance matrix, a clustered sandwich estimator is used to allow for intragroup correlation, relaxing the usual requirement that the observations be independent. In other words, the observations are independent across groups but not necessarily within groups.

[18] These quantiles, which make the search set $\Gamma$ smaller than what is typically considered in time-series threshold models applications, have been chosen so as to save computation time since we do not expect the inaction zone to be very large. Indeed, recall that almost 45 percent of our forecasters change their forecast between two consecutive months. Hence, the threshold has to be crossed quite often.

**Table 2. Threshold Logit Models with Fixed Effects**

| | **FR** | | **GE** | |
|---|---|---|---|---|
| | **Coefficient** | **p-value** | **Coefficient** | **p-value** |
| $\lvert\pi_m(-1)\rvert > \hat{\bar{\pi}}$: | | | | |
| $s_t$D(d=2) | 0.025 | (0.353) | 0.013 | (0.362) |
| $s_t$D(d=3) | 0.009 | (0.740) | 0.067*** | (0.004) |
| $s_t$D(d=4) | 0.047 | (0.350) | 0.036 | (0.415) |
| $s_t$D(d=5) | –0.036 | (0.564) | –0.011 | (0.771) |
| $s_t$D(d=6) | 0.192** | (0.030) | 0.076 | (0.186) |
| $s_t$D(d=7) | 0.084 | (0.562) | –0.108 | (0.281) |
| $s_t$D(d≥8) | 0.194 | (0.175) | –0.024 | (0.779) |
| $s_t\lvert\pi_m(-1)\rvert$ | 0.219** | (0.045) | 0.350*** | (0.000) |
| $\lvert\pi_m(-1)\rvert \leq \hat{\bar{\pi}}$: | | | | |
| $(1 - s_t)$D(d=2) | –0.004 | (0.847) | 0.012 | (0.505) |
| $(1 - s_t)$D(d=3) | 0.119*** | (0.000) | 0.133*** | (0.000) |
| $(1 - s_t)$D(d=4) | –0.065 | (0.215) | –0.040 | (0.248) |
| $(1 - s_t)$D(d=5) | –0.063 | (0.364) | –0.072* | (0.076) |
| $(1 - s_t)$D(d=6) | 0.332*** | (0.000) | 0.172*** | (0.000) |
| $(1 - s_t)$D(d=7) | 0.064 | (0.701) | –0.24 | (0.738) |
| $(1 - s_t)$D(d≥8) | –0.155 | (0.494) | –0.080 | (0.220) |
| $(1 - s_t)\lvert\pi_m(-1)\rvert$ | 0.580 | (0.292) | –0.551** | (0.050) |
| $Sha(-1)$ | 0.080* | (0.083) | 0.219*** | (0.000) |
| $\lvert Dev(-1)\rvert$ | 0.348*** | (0.000) | 0.030 | (0.594) |
| Horizon FE | Yes | | Yes | |
| Target Year FE | Yes | | Yes | |
| Individual FE | Yes | | Yes | |
| $\hat{\bar{\pi}}$ (Quantile) | 0.064 (54%) | | 0.085 (57%) | |
| $\hat{\bar{\lambda}}$, Predicted Prob. | 46.6% | | 44% | |
| Log-Likelihood | –3,187.13 | | –5,176.56 | |
| No. Obs. | 4,990 | | 8,202 | |
| No. $i$ | 36 | | 51 | |

**Note:** Numbers are marginal effects computed at sample means for continuous covariates. For dummies, they show the effect of moving from one discrete state to the other one. $z$-statistics p-values, given in parentheses, are computed using a robust variance-covariance matrix (from a clustered sandwich estimator). *, **, and *** denote 10, 5, and 1 percent significance levels, respectively.

effect (variable $|Dev(-1)|$) is found to be much stronger in France, with an estimated coefficient of 0.35, than in Germany, where it is not even significantly different from zero. So, in France, the probability to update is significantly increased when a forecaster's previous forecast is far from last-period average forecast.[19]

Let us now turn to the threshold effect. First, it is worth noticing that the threshold models' likelihoods are improved compared with their non-switching analogues, reported in Appendix E. As can be seen there, their log-likelihoods are −3,195.6 and −5,188.1 for France and Germany, respectively. So as to test for the presence of a threshold effect, we have performed a SupWald test of the non-threshold model under the null versus our threshold model under the alternative. Since the threshold is a nuisance parameter under the non-switching null, the test statistics is not distributed as a chi-squared. Consequently, the SupWald p-value is instead calculated by bootstrap following the fixed-regressor bootstrap method of Hansen (1996, 2000b) or Hansen and Seo (2002). We ran 1,000 simulations, each with only 25 equally spaced thresholds values between the 25 percent and 75 percent quantiles of $|\tilde{\pi}_f(t)|$ under the alternative.[20] It turns out that the null is rejected at the 6 percent level for Germany and at the 31 percent level for France. This result for France might be partly due to the smaller sample size available in this country than in Germany (about 60 percent of the German sample size). Accordingly, French results should be interpreted with caution. Then, the threshold estimates correspond to the 54 percent and 57 percent quantiles of $|\Pi_m(-1)|$ distribution in France and Germany, respectively. This means that the forecast gap does not need to be very large to trigger more updates among the forecast units of our sample: The gap may just be slightly above the median.

[19]Even though the comparison with Dovern (2013)'s results is not straightforward, as this author considers GDP forecasts from a cross-country survey data, it is worth noting that the estimates of our non-threshold regressions reported in Appendix E support his findings regarding the influence of the forecast horizon, the last revisions occurring three and six months ago (D(d=3) and D(d=6)), as well as the "herding" ($|Dev(-1)|$) and "cascade" ($Sha(-1)$) effects.

[20]A more precise grid search of the threshold values would probably result in slightly lower p-values, but our actual choice seems to be a good compromise between accuracy and computational time burden concerns.

Regarding the time dependence of updates, empirical evidence supporting this view comes from the significant increase in the probability of updating if the last revision has been made three or six months ago. With our theoretical model's property that forecast update is triggered immediately after observation if the forecast gap is large enough, this gives an upper bound estimate for $T(0)$ of six months in both regimes, as predicted by the theoretical model. The six-month pattern is even stronger than the three-month one. This is particularly true in France, where the probability of update is increased by 19.2 percent if the last revision was six months ago, compared with 1 percent if it occurred three months ago in the outer regime. This result holds despite the inclusion of a forecast horizon fixed effect, hence revealing a pure time dependence effect which is not exclusively imputable to, e.g., forecasters' institutional framework. Also, as expected given the model (see Figure 2), the optimal time to next observation can be shorter in the inaction band, as the estimated coefficients associated with $(1-s_t)$D(d=3) are larger than the ones related to $s_t$D(d=3).[21]

Finally, our empirical results also support the state dependence of the forecast update decision rule. Actually, the update probability is significantly increased when the forecast gap crosses the threshold, by 22 percent in France and 35 percent in Germany. Contrarily, this probability is strongly decreased in the aptly named inaction band for Germany (−55 percent). French updates are not significantly influenced by $(1 - s_t)|\pi_m(-1)|$, as the $z$-test p-value is close to 30 percent. Again, these conclusions are robust to the random-effect version of the logit model.

### *4.3 Empirical Evidence of Coexistence of Observation and Adjustment Costs*

We turn now to Property 4 of the theoretical model so as to evaluate the relative costs. Recall that the ratio of adjustment to observation

[21] The conclusions of these models remain unchanged from random-effect logit model estimation and also if a probit model is substituted for the logit model. The exclusion of institutions that provide fewer than 20 usable observations does not affect the conclusions either. These results are not reported to save space but are available upon request.

**Table 3. An Evaluation of Relative Costs**

| | $\hat{\hat{\pi}}$ | $n_o$ | $n_c$ | $\psi/\theta$ |
|---|---|---|---|---|
| FR | 0.064 | 76.6% | 45.0% | 49.1% |
| | [0.046;0.074] | [71.8%;79.7%] | — | [35.4%;59.3%] |
| GE | 0.085 | 75.9% | 41.7% | 67.1% |
| | [0.081;0.106] | [76.6%;82.1%] | — | [70.1%;93.7%] |

costs, $\psi/\theta$, is a non-linear function of observations to adjustment frequencies ratio, the former being calculated thanks to the threshold estimates obtained in Table 2, along the way described in Section 2. Using these approximations, we obtain $\psi/\theta = 49.1$ percent in France and 67.1 percent in Germany, as reported in Table 3. In order to calculate the 95 percent confidence interval of these estimates, the method developed in Hansen (1999, 2000a) is used. More precisely, for all threshold values included in the grid search interval, i.e., $\forall|\tilde{\pi}_f(t)| \in \Gamma = [\gamma_{25\%}, \gamma_{75\%}]$, his proposed likelihood ratio test $LR_1(\gamma)$ is calculated: it compares the likelihood obtained with the estimated threshold $\hat{\hat{\pi}}$ reported in Table 2 with the ones obtained with all other values in $\Gamma$. The asymptotic distribution of $LR_1(\gamma)$ is shown to be highly non-standard but free of nuisance parameters and its critical value at the 5 percent level is 7.35 (see Hansen 1999, 2000a). Hence, the "no-rejection region" of confidence level 95 percent is the set of values $\gamma$ such that $LR_1(\gamma) \leq 7.35$. With this confidence interval for the threshold estimate at hand, it is straightforward to compute the $n_o$ and $\psi/\theta$ boundaries accordingly.[22] These confidence intervals are given in square brackets in Table 3.

It is worth noting that this ratio is neither close to zero nor arbitrarily large, hence confirming coexistence of both costs: As can be seen from Table 3, the smallest lower boundary of the confidence interval for the costs ratio is 35.4 percent in France, while the highest upper boundary is 93.7 percent in Germany. In addition, the costs ratio is smaller than one in both countries, meaning that observation costs dominate communication costs. Even though applied to price-setting rules, conclusions obtained by Zbaracki et al.

[22]Remember that $n_c$ is observed, hence assumed known and fixed.

(2004) have the same flavor as ours in the sense that the managerial costs of a large U.S. industrial manufacturer (including information gathering, decisionmaking, and internal communication costs) are found to be larger than its menu costs (comparable to the external communication costs of our model).

## 5. Conclusion

This paper first develops a theoretical model of forecasts formation which incorporates separate observation and communication costs. As a result, the forecast update decision rule is found to be both time and state dependent. The model's main properties for the forecasts update process are the following. First, the forecast update is communicated immediately after observation if the forecast gap upon observation is large enough in absolute value. Second, the forecast revision communication is a non-linear function of the forecast gap: For small gaps $\in (-\overline{\pi}, \overline{\pi})$, there is no communication, whereas there is one as soon as the absolute value of the gap exceeds $\overline{\pi}$. If so, the gap is closed by the update. Third, the optimal time between two observations is also a non-linear function of the forecast gap, and the closer to the boundaries $-\overline{\pi}$ or $\overline{\pi}$, the sooner the next observation is. Fourth, the time between two observations reaches a maximum when the gap is closed.[23]

The time and state dependence of the observation and forecast revision communication implied by this model are then tested using inflation forecast updates of professional forecasters from recent Consensus Economics panel data for France and Germany. For this purpose, conditional forecast revision communication probabilities are estimated from binary choice models incorporating regime-switching features. This allows to estimate the threshold value $\overline{\pi}$ along all other parameters.

Our findings clearly support time-dependent updates, a result which is compatible with the observation cost assumption as introduced by Mankiw and Reis (2002). Indeed, they point to a strong positive effect when the last update has occurred three and/or six

[23]This maximum time increases in both observation and communication costs; see Proposition C.2 in Appendix C.

months ago, even after controlling for the institution's periodic forecast framework (quarterly or biannual). This gives a maximum of six months for the optimal time between two observations. This confirms Coibion and Gorodnichenko (2015)'s estimate of the average duration between information updates.

Evidence of update state dependence is also provided. In fact, a strong positive and significant effect is found on updates when the forecast gap is larger than the estimated threshold, as proxied by the last known monthly inflation rate, weighted by its mechanical contribution to the yearly inflation rate forecast.

Finally, our results confirm the co-existence of both types of costs with a forecast communication cost smaller than the observation cost. This result is very much in line with Zbaracki et al. (2004)'s findings for firms' price-setting decision rules.

All in all, this work is a first attempt to explain why not communicating forecast updates every month can be optimal for professional forecasters. The theoretical model grounding it is of course highly stylized. We believe that it is desirable to improve this model so as to make it closer to the true forecaster's environment and hence to test it directly instead of estimating a reduced-form regression. This is on our research agenda.

## Appendix A. The Optimal Control Problem

Suppose we are in a base period, say $t = 0$, where the forecast is equal to $\pi_f(0)$. From there on, we seek to forecast the inflation rate at date $t = 1$. As stated in Assumption 2, the instantaneous quadratic loss function faced by a representative professional forecaster is $\tilde{\pi}_f(t)^2$. As long as a new forecast has not been publicly adjusted and communicated, $\pi_f(t)$ is constant, equal to $\pi_f(0)$. It is assumed that the forecaster's objective is to produce the best possible forecast, which amounts to minimize the distance between his forecast and the optimal forecast. The latter is a BM, according to Assumption 1.

We assume that at time $t = 0$, the forecaster pays the cost $\theta$ and observes $\pi_f^*(0)$. Then, until the next observation of the target forecast, $E(\pi_f^*(h)|\mathcal{I}_0) = \pi_f^*(0)$, where $\mathcal{I}_0$ is the information set observed at time $t = 0$. Indeed, with $\pi_f^*(0) = E(\pi(1)|\mathcal{I}_0)$, where $\pi(1)$ is the inflation rate at $t = 1$, we have

$$E\left(\pi_f^*(h)|\mathcal{I}_0\right) = E\left(E\left(\pi(1)|\mathcal{I}_h\right)|\mathcal{I}_0\right) = E\left(\pi(1)|\mathcal{I}_0\right) \\ = \pi_f^*(0). \tag{A.1}$$

The target forecast dynamics is given by Assumption 1.

We define the "uncontrolled" forecast gap as the forecast gap $\tilde{\pi}_f(t)$ between two observations of $\pi_f^*(t)$ and before a new adjustment of $\pi_f(t)$. It follows that starting from $\tilde{\pi}_f(0)$ at time $t = 0$, the uncontrolled forecast gap evolves as

$$d\tilde{\pi}(t) = -\sigma dW(t). \tag{A.2}$$

As will be shown below, at optimum there will be adjustments at finite time in accordance with the loss-minimization objective, which will make the forecast gap process globally stationary.

In this setup, the problem of the forecaster is to choose the time elapsed until the next observation of the target forecast, $0 < T \leq 1$, which will cost $\theta$, as well as the number of forecast update communications between two observations (at $t = 0$ and $T$, respectively), $J \in N$, occurring at successive dates $0 \leq t_1, t_2, \ldots, t_J < T$, each one incurring an adjustment cost $\psi$. He also chooses the size of the forecast update so that the expected value of the forecast gap on adjustment is $\hat{\pi}_{f,j}$, with $j = 1, \ldots, J$.

As it is stated, the problem is formally similar to the one treated by Alvarez, Lippi, and Paciello (2011). There are two differences:

(i) The first difference comes from the fact that the timing of the next observation, that is the choice of $T$, is bounded by the forecast horizon, 1. There is no upper bound of this sort in Alvarez, Lippi, and Paciello (2011). However, this is a pure control constraint, the simplest one from the mathematical point of view, and we show in our empirical section that it is never binding in practice. Without loss of generality, we shall omit it hereafter to unburden the presentation.

(ii) The second concerns the law of motion assumed by Alvarez, Lippi, and Paciello (2011) for the target: It has a time drift ($\mu \geq 0$), which also shows up in the law of motion of the analogous stochastic gap.[24] While non-zero drifts make sense in their framework (they are studying *price* setting), they do not in ours (focused on *inflation forecast* behavior). As a result, from this point of view, our problem corresponds to a special case in Alvarez, Lippi, and Paciello (2011), explicitly treated in their Section V, pp. 1928–34.

Henceforth, we shall use the same formalism and methodology as in Alvarez, Lippi, and Paciello (2011). Let $V(\tilde{\pi}_f)$ denote the value function of the forecaster at the time of an observation of the forecast gap $\tilde{\pi}_f$, and $V_J(\tilde{\pi}_f)$ the best value that the forecaster can reach by making $J$ forecast updates between observations. Then, note that

$$V(\tilde{\pi}_f) = \min_{J \geq 0} V_J(\tilde{\pi}_f), \quad \text{so that} \quad J^\star(\tilde{\pi}_f) = \arg\min_{J \geq 0} V_J(\tilde{\pi}_f), \qquad \text{(A.3)}$$

where $J^\star(\tilde{\pi}_f)$ is the optimal number of forecast updates. The corresponding Bellman equations can be written accordingly. To ease the exposition, suppose that $J \in \{0, 1\}$, that is, at most one forecast update is optimal. This will turn out to be true for our forecast formation model. Let's denote $\tilde{\pi}_f = \tilde{\pi}_f(0)$ the value of the forecast gap at time $t = 0$.

For $J = 0$, the conditional value function for the forecaster problem writes

$$\begin{aligned} V_0(\tilde{\pi}_f) = \theta + \min_T \int_0^T & e^{-\rho t}(\tilde{\pi}_f^2 + \sigma^2 t)dt \\ & + e^{-\rho T} \int_{-\infty}^{\infty} V(\tilde{\pi}_f - s\sigma\sqrt{T})dN(s), \end{aligned} \qquad \text{(A.4)}$$

where $s$ is a standard normal. The first component of the right member of $V_0(\tilde{\pi}_f)$ is of course the information observation cost, while the second one is the time $t = 0$ expected loss between $t = 0$ and $T$. In the third (continuation) component, $N(.)$ is the probability density function of a Gaussian distribution.

[24]See Equation (1) in Alvarez, Lippi, and Paciello (2011, p. 1917).

For $J = 1$ it becomes (with the simplified notation of the adjusted forecast at time $t_1$ as $\hat{\pi}_f = \hat{\pi}_{f,1}$)

$$V_1(\tilde{\pi}_f) = \theta + \min_{T,\hat{\pi}_f,t_1} \int_0^{t_1} e^{-\rho t}\tilde{\pi}_f^2 dt + \int_0^T e^{-\rho t}\sigma^2 t dt$$
$$+ e^{-\rho t_1}\left[\psi + \int_0^{T-t_1} e^{-\rho t}\hat{\pi}_f^2\right] dt$$
$$+ e^{-\rho T}\int_{-\infty}^{\infty} V(\hat{\pi}_f - s\sigma\sqrt{T})dN(s). \qquad (A.5)$$

The unconditional value function is hence given by

$$V(\tilde{\pi}_f) = min\{V_0(\tilde{\pi}_f), V_1(\tilde{\pi}_f)\}. \qquad (A.6)$$

Because of the quadratic functions involved, it is possible to provide with a partial analytical characterization of the solutions to these Bellman equations.[25] Let us first focus on the optimal forecast revisions before turning to the optimal time between observations.

## Appendix B. The Optimal Number and Timing of Forecast Updates Announcements

Using Proposition 1, p. 1921, in Alvarez, Lippi, and Paciello (2011), Proposition B.1 below can be established straightforwardly.[26]

PROPOSITION B.1. *Let $\theta > 0$, $\psi > 0$ and $\sigma > 0$.*

*(i)* $J^\star(\tilde{\pi}_f) \leq 1, \forall\tilde{\pi}_f \in R$.

*(ii) If $\tilde{\pi}_f$ is such that $J^\star(\tilde{\pi}_f) = 1$, then the forecast update is full and occurs instantaneously ($\hat{\pi}_f = 0$ and $t_1 = 0$).*

[25] Alvarez, Lippi, and Paciello (2011) do solve the more general case with positive time drift in Equation (2). As mentioned above, only the zero drift case is relevant here.

[26] Actually, Alvarez, Lippi, and Paciello (2011) get the same characterization even if the law of motion of their forecast target has a time drift, provided its absolute value is small enough. Our case—zero time drift—is indeed a very elementary special case.

This proposition means that for the quadratic loss function and the forecast gap process defined in Equation (A.2), there will be at most one forecast update communication between two forecast target observations and, if there is one, it will be fully and instantaneously adjusted to the updated forecast target. The intuition of this result is as follows. If $\psi$ and $\sigma$ are strictly positive, it is immediate to see that in case of inflation forecast adjustment, $\hat{\pi}_f = 0$. Then the optimal choice of the forecaster is to reset the forecast gap to zero. Indeed, it follows from Equation (A.2) that the expected value of the forecast gap remains at zero between observations. Consequently, there are no gains to expect from a forecast update without the new pieces of information acquired thanks to an observation. By contrast, due to the fixed communication cost $\psi > 0$, the losses involved by a forecast update release are strictly positive. So, it is not optimal to adjust the forecast between information set updates. Using the same type of arguments, one can also readily get why the update communication, if any, should not only be full but also instantaneous. Indeed, if $t_1 > 0$, then the forecaster will incur losses in the time interval $[0; t_1[$. Because the presence of "menu costs" typically induces that adjustments only occur if the gaps are sufficiently large, delaying cannot be optimal, as it involves starting the adjustment at such values of the gap: The corresponding trajectory of losses is clearly dominated by a trajectory where adjustment is made instantaneously. Hence $\hat{\pi}_f = 0$ and $t_1 = 0$. A first result of our theoretical model of forecasts formation is that forecasts communication plans, i.e., progressive release of forecasts' revisions between two updates of the information set, are not optimal. The new forecast release, if any, is full and instantaneous at the observation time. Notice that the resulting Bellman equation (A.5) can be rewritten as

$$\begin{aligned} V_1 = \psi + \theta + \min_{T, \hat{\pi}_f} \int_0^T e^{-\rho t} (\hat{\pi}_f^2 + \sigma^2 t) dt \\ + e^{-\rho T} \int_{-\infty}^{\infty} V(\hat{\pi}_f - s\sigma\sqrt{T}) dN(s). \end{aligned} \tag{B.1}$$

Note that $V_1$ does not depend on $\tilde{\pi}_f$ since $t_1 = 0$. Hence the value function is given by

$$V(\tilde{\pi}_f) = min\{V_0(\tilde{\pi}_f), V_1\}. \tag{B.2}$$

One can see immediately that the value function $V$ is symmetric around $\tilde{\pi}_f = 0$ and increasing for $|\tilde{\pi}_f| < \bar{\pi}$, where $\bar{\pi}$ is a threshold value such that $V_1 > V_0(\tilde{\pi})$ for $\tilde{\pi}_f \in (-\bar{\pi}, \bar{\pi})$. Hence $(-\bar{\pi}, \bar{\pi})$ defines the range of inaction in which no forecast adjustment is communicated. It means that when $\tilde{\pi}_f$ is smaller than $\bar{\pi}$ in absolute value, then the inflation forecast is not adjusted. Since $V$ is not differentiable at $\tilde{\pi}_f = \bar{\pi}$, the value function is discontinuous, non-smooth at the boundaries of the inaction band. These properties are summarized in the next proposition:[27]

PROPOSITION B.2. *The value function $V$ is symmetric around $\tilde{\pi}_f = 0$, and $V$ is strictly increasing in $\tilde{\pi}_f$ for $0 < \tilde{\pi}_f < \bar{\pi}$. $V'(\tilde{\pi}) = 0$ for $\tilde{\pi}_f > \bar{\pi}$ and $V$ is not differentiable at $\tilde{\pi}_f = \bar{\pi}$.*

## Appendix C. The Optimal Time between Observations

Since $\hat{\pi}_f = 0$ at the forecaster optimum, then the optimal time between observations, conditional on adjustment, is $T(0)$. It is possible to characterize much more closely function $T(.)$: Indeed, Alvarez, Lippi, and Paciello (2011) show that it has a maximum at $\tilde{\pi}_f = 0$ and is symmetric around 0 with an inverted U-shape.[28] The function $T(\cdot)$ is discontinuous and not differentiable at $\tilde{\pi}_f = \overline{\pi}$ , as stated in Proposition C.1.[29]

PROPOSITION C.1. *As $\rho \downarrow 0$, the optimal rule for the time of the next observation of the forecast gap is given by Equation (2).*

Indeed, Equation (2) is obtained from a third-order expansion of $T(.)$ around zero, which requires the condition $\rho \downarrow 0$ appearing in Proposition C.1. Note that the optimal rule for the time of the

[27] The proof of Proposition B.2 follows the same steps as the proof of Proposition 3, p. 1928, and Lemma 1, p. 1945, in Alvarez, Lippi, and Paciello (2011).

[28] See Alvarez, Lippi, and Paciello (2011), Proposition 4, p. 1929.

[29] Because the proof of Proposition C.1 follows the same steps as the proof of Proposition 4, p. 1929, in Alvarez, Lippi, and Paciello (2011), we refer the reader to the latter in order to lighten the presentation.

next observation is a discontinuous, threshold function of the forecast gap: it switches from $T(0)$ for all $|\tilde{\pi}_f| > \overline{\pi}$ to a function of $\tilde{\pi}_f$ otherwise. It is also worth noticing that if there is no forecast gap after an observation, i.e., $\tilde{\pi}_f = 0$, then T(0) is optimal. Finally, $T(\tilde{\pi}_f)$ decreases with $|\tilde{\pi}_f|$. When close to the boundaries of the range of inaction, the forecaster plans an early observation since the target is likely to cross the threshold.

Finally, the next proposition can easily be shown to hold[30]:

Proposition C.2. *Let $\theta > 0$, $\psi > 0$ and $\sigma > 0$, $\frac{\psi}{\theta} < 5.5$. As $\rho \downarrow 0$, there exists a unique solution for $T(0)$ and $\overline{\pi}$ and it is such that*

*(i) $T(0)$ is increasing in $\psi$ and $\theta$*

*(ii) $\overline{\pi}$ is increasing in $\psi$ and decreasing in $\theta$.*

As expected, the time to the next observation after a forecast update, $T(0)$, is increasing in the observation cost, and the width of the inaction band defined by $(-\overline{\pi}, \overline{\pi})$ is increasing in the adjustment cost.

## Appendix D. The Optimal Forecast Gap Dynamics

The decision rules described by the threshold $\overline{\pi}$ and the function $T(\cdot)$ imply a stationary Markov process for the dynamics of the forecast gap on observation and before communication. It follows from Proposition B.1 in Appendix B that the forecast gap adjustment rule $\Delta(\tilde{\pi}_f)$ is zero in the inaction zone and $-\tilde{\pi}_f$ otherwise. Assume that the next observation occurs in $T' = T(\tilde{\pi}_f)$ periods and the corresponding forecast gap is $\tilde{\pi}'_f$. Then, the proposition below follows immediately from Proposition C.1.[31]

Proposition D.1. *Optimal forecast gap dynamics:*

$$\tilde{\pi}'_f = \begin{cases} \tilde{\pi}_f - s\sigma\sqrt{T(0) - \left(\frac{\tilde{\pi}_f}{\sigma}\right)^2}, & \text{if } \tilde{\pi}_f \in (-\overline{\pi}, \overline{\pi}) \\ -s\sigma\sqrt{T(0)}, & \text{otherwise.} \end{cases} \tag{D.1}$$

[30]Here, we use Proposition 5, p. 1931, in Alvarez, Lippi, and Paciello (2011) as well as on results reported p. 1932 therein.

[31]See also Equation (16), p. 1934, in Alvarez, Lippi, and Paciello (2011).

Again, this is a threshold regime-switching process: If the forecast gap is large enough, i.e., larger than $\overline{\pi}$ in absolute value, then it is corrected by setting the forecast to its target and the remaining gap is entirely imputable to unexpected shocks. $\tilde{\pi}_f'$ increases with $|\tilde{\pi}_f|$. When close to the boundaries of the range of inaction but still inside it, the forecast gap reaches its largest values until it hits the boundary and then goes back to its minimum. This is the Ss rule.

## Appendix E. Logit Models Estimates without Threshold Effect

**Table E.1. Logit Models: Estimates without Threshold Effect**

| | FR | | GE | |
|---|---|---|---|---|
| | Coefficient | p-value | Coefficient | p-value |
| D(d=2) | 0.010 | (0.541) | 0.013 | (0.218) |
| D(d=3) | 0.062*** | (0.009) | 0.101*** | (0.000) |
| D(d=4) | –0.012 | (0.713) | –0.002 | (0.937) |
| D(d=5) | –0.048 | (0.369) | –0.033 | (0.314) |
| D(d=6) | 0.245*** | (0.001) | 0.126*** | (0.002) |
| D(d=7) | 0.077 | (0.537) | –0.701 | (0.297) |
| D(d≥8) | 0.069 | (0.473) | –0.043 | (0.529) |
| $\|\pi_m(-1)\|$ | 0.183** | (0.026) | 0.428*** | (0.000) |
| $Sha(-1)$ | 0.089** | (0.050) | 0.197*** | (0.000) |
| $\|Dev(-1)\|$ | 0.354*** | (0.000) | 0.029 | (0.611) |
| Horizon FE | Yes | | Yes | |
| Target Year FE | Yes | | Yes | |
| Individual FE | Yes | | Yes | |
| No. Obs. | 4,990 | | 8,202 | |
| Log-Likelihood | –3,195.57 | | –5,188.12 | |
| No. $i$ | 36 | | 51 | |

**Note:** Numbers are marginal effects computed at sample means for continuous covariates. For dummies, they show the effect of moving from one discrete state to the other one. $z$-statistics p-values, given in parentheses, are computed using a robust variance-covariance matrix (from a clustered sandwich estimator). *, **, and *** denote 10, 5, and 1 percent significance levels, respectively.

## References


Alvarez, F., F. Lippi, and L. Paciello. 2011. "Optimal Price Setting With Observation and Menu Costs." *Quarterly Journal of Economics* 126 (4): 1909–60.

Andrade, P., and H. Le Bihan. 2013. "Inattentive Professional Forecasters." *Journal of Monetary Economics* 60 (8): 967–82.

Banerjee, A. 1992. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics* 107 (3): 797–817.

Coibion, O. 2010. "Testing the Sticky Information Phillips Curve." *Review of Economics and Statistics* 92 (1): 87–101.

Coibion, O., and Y. Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy* 120 (1): 116–59.

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

Dovern, J. 2013. "When Are GDP Forecasts Updated? Evidence from a Large International Panel." *Economics Letters* 120 (3): 521–24.

Dovern, J., U. Fritsche, P. Loungani, and N. Tamirisa. 2015. "Information Rigidities: Comparing Average and Individual Forecasts for a Large International Panel." *International Journal of Forecasting* 31 (1): 144–54.

Graham, J. 1999. "Herding among Investment Newsletters: Theory and Evidence." *Journal of Finance* 54 (1): 237–68.

Hansen, B. 1996. "Inference when a Nuisance Parameter Is Not Identified Under the Null Hypothesis." *Econometrica* 64 (2): 413–30.

———. 1999. "Threshold Effects in Non-dynamic Panels: Estimation, Testing, and Inference." *Journal of Econometrics* 93 (2): 345–68.

———. 2000a. "Sample Splitting and Threshold Estimation." *Econometrica* 68 (3): 575–604.

———. 2000b. "Testing for Structural Change in Conditional Model." *Journal of Econometrics* 97 (1): 93–115.

Hansen, B., and B. Seo. 2002. "Testing for Two-Regime Threshold Cointegration in Vector Error-Correction Models." *Journal of Econometrics* 110 (2): 293–318.

Lamont, O. 2002. "Macroeconomic Forecasts and Microeconomic Forecasters." *Journal of Economic Behavior and Organization* 48 (3): 265–80.

Loungani, P., H. Stekler, and N. Tamirisa. 2013. "Information Rigidity in Growth Forecasts: Some Cross-Country Evidence." *International Journal of Forecasting* 29 (4): 605–21.

Magnani, J., A. Gorry, and R. Oprea. 2016. "Time and State Dependence in an Ss Decision Experiment." *American Economic Journal: Macroeconomics* 8 (1): 285–310.

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

Pons-Novell, J. 2003. "Strategic Bias, Herding Behaviour and Economic Forecasts." *Journal of Forecasting* 22 (1): 67–77.

Sims, C. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3): 665–90.

Stock, J., and M. Watson. 2007. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking* 39 (s1): 3–33.

Woodford, M. 2002. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information and Expectations in Modern Macroeconomics,* ed. P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, 25–58. Princeton University Press.

Zbaracki, M., M. Ritson, D. Levy, S. Dutta, and M. Bergen. 2004. "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets." *Review of Economics and Statistics* 86 (2): 514–33.