Reslow, André

Working Paper

Inefficient use of competitors' forecasts?

Working Paper, No. 2019:9

Provided in Cooperation with:
Department of Economics, Uppsala University

Suggested Citation: Reslow, André (2019) : Inefficient use of competitors' forecasts?, Working Paper, No. 2019:9, Uppsala University, Department of Economics, Uppsala, https://nbn-resolving.de/urn:nbn:se:uu:diva-395740

This Version is available at:
https://hdl.handle.net/10419/224113

UPPSALA
UNIVERSITET

Department of Economics
Working Paper 2019:9

Inefficient Use of Competitors' Forecasts?

André Reslow

Department of Economics
Uppsala University
Box 513
751 20 Uppsala
Sweden

Working Paper 2019:9
October 2019
ISSN 1653-6975

Inefficient Use of Competitors' Forecasts?

André Reslow

Inefficient Use of Competitors' Forecasts?*

André Reslow†

October 2019

Abstract

This paper assesses to what extent forecasters make efficient use of competitors' forecasts. Using a panel of forecasters, I find that forecasters *underuse* information from their competitors in their forecasts for current and next year's annual GDP growth and inflation. The results also show that forecasters increase the attention to their competitors as the forecast horizon decreases. In a model of noisy information with fixed target forecasts, I confirm the empirical results of underuse of competitors' information. I also extend the model to include a revision cost and show how this can explain the observed inefficiency and observed horizon dynamics. Using the same model framework, I also rule out overconfidence as the main explanation of the observed behavior.

Keywords: Forecast Behavior, Efficiency, Revision Cost, Forecast Smoothing, Overconfidence

JEL Classification: C53, D82, E17, E37

*I would like to thank Mikael Carlsson, Jesper Lindé and Erik Öberg for extensive feedback, and Daniel Bougt, Tobias Broer, Fabio Canova, David Cesarini, Davide Cipullo, Nils Gottfries, Sofia Hernnäs, Isaiah Hull, Sebastian Jävervall, Karl Walentin, and the participants in the seminars held at Uppsala University and Sveriges Riksbank for their dialogue and comments. Financial support from Handelsbankens Forskningsstiftelser is gratefully acknowledged. The forecast data used in this paper is copyright protected against redistribution but can be purchased from the provider. The views expressed in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of Sveriges Riksbank.

†Dept. of Economics, Uppsala University and Sveriges Riksbank. E-mail: andre.reslow@nek.uu.se

1 Introduction

In a world with incomplete and private information, an important source of information for a macroeconomic forecaster is the forecasts of other forecasters. The starting point of this paper is to assess to what extent macroeconomic forecasters make efficient use of forecasts from their competitors. I use the concept of weak and strong forecast efficiency presented in Nordhaus (1987) and augment the weak concept to include competitor information in addition to own lagged information. I also deploy a simple model of noisy information to help us understand forecasters' optimal and actual behavior. Starting from the baseline model, I show that a revision cost together with horizon discounting can explain the observed actual behavior that differs from the suggested optimal behavior.

The empirical strategy makes use of forecaster firm-level panel data from Consensus Economics. The data used covers 10 countries: Canada, France, Germany, Italy, Japan, Netherlands, Spain, Sweden, the United Kingdom, and the United States. The data span forecasts made between 1995 and 2017. The forecasts refer to the current and next year's annual growth rates such that we have fixed target forecasts where multiple forecasts are made for the same outcome at a decreasing forecast horizon. The main analysis focuses on forecasts for GDP growth and inflation (consumer prices). Using this dynamic panel of 591 forecasters, I find that forecasters *underuse* competitor information in their forecasts for current and next year's annual GDP growth and annual inflation. The forecasters appear to put a sub-optimal high weight on their own lagged forecast and too low a weight on the work by their competitors. This results in forecasts that are too persistent. The empirical results also show that the forecasters behave differently at different horizons. As the forecast horizon decreases, the forecasters increase the attention to their competitors—but they are still far from an optimal level. I do not find any sizeable heterogeneity regarding the variable, country, or year during which the forecast is produced.

The main contribution of this paper is twofold. First, the specific focus on the *relative* use between forecasters' own lagged forecasts and those of others and how this differs between horizons. Second, how this behavior can be explained in a noisy information framework by introducing revision costs and horizon discounting. The empirical analysis in this paper is related to work by, for example, Batchelor and Dua (1992), who find that most U.S. forecasters in the Blue Chip panel could have improved their forecasts if they were less attached to their own past forecast when making forecast revisions. The analysis is also related to Lahiri and Sheng (2008), who estimate a Bayesian learning model with heterogeneity aimed at explaining forecast disagreement and its evolution over horizons.

The analysis also relates to more recent work by, among others, Coibion and Gorodnichenko (2015) and Dovern et al. (2015), both of whom estimate how forecasters react to new information. In line with Coibion and Gorodnichenko (2015) and Dovern et al. (2015), I also find, as a robustness check, that both the average (consensus mean) forecast and individual forecasts on average under-respond to new information. As a result, too much weight is attached to the own past forecasts.[1]

In line with the empirical analysis, a baseline model with noisy and private information predicts that, in order for a forecaster to try to minimize forecast errors, it is optimal to give considerable attention to the work by competitors and consequently put a high weight on the forecast from others. The suggestion that it is optimal to pool information from many forecasters is commonly raised in forecasting. For example, Zarnowitz (1984) finds that mean forecasts are on average more accurate over time than individual forecasts.[2] This also resembles the work on herd behavior by, for example, Banerjee (1992), but with some differences. In the Banerjee (1992) model, herding is doing what everyone else is doing even if private information suggests doing something different. In the model in this paper, forecasters still put positive weight on private information.[3] I consider two versions of the model: one model that has fully informed forecasters and one that has "rule of thumb" forecasters. In the fully informed model, the forecaster filters out signals from competitors to produce the most efficient forecast. The "rule of thumb" model is used to impose some form of limitation on the forecasters' ability. The "rule of thumb" that forecasters use will be such that they use lagged aggregate forecasts instead of filtering out individual signals.

The two models differ in terms of efficiency, but both models provide the same prediction regarding relative attention to own versus others' lagged forecasts. Both models are noisy information models that introduce a signal extraction problem for Bayesian updating forecasters. In this type of model (see for example Lucas (1973), Woodford (2001) and Sims (2003)), agents continuously update their information sets and their beliefs via signal extraction. The information structure in the models matches the specific structure of fixed target forecasts that is used in the empirical analysis. The baseline models do not have any form of strategic trade-off for forecasters and, hence, assume that

[1] Other related papers include Ager et al. (2009), who study accuracy and efficiency of consensus forecasts, and Isiklar et al. (2006), who study forecast revision to see how quickly forecasters incorporate news in a cross-country setting. Deschamps and Ioannidis (2013) study forecast revisions and find that forecasters underreact to new information. In addition, Chen and Jiang (2006) find that corporate analysts place larger, rather than efficient, weights on their private information when they forecast corporate earnings.

[2] One ongoing example of this result is the work by Nate Silver and the FiveThirtyEight polling aggregation website, which focuses on predicting outcomes in politics and sports. See also the book *The Signal and the Noise* by Silver (2012).

[3] See Lamont (2002), Gallo et al. (2002), Bernhardt et al. (2006) and Rülke et al. (2016) for more on herd behavior among forecasters.

forecasters objective is pure forecast error minimization.

To explain the observed actual (inefficient) behavior, I extend the baseline model. I did this to consider alternatives to pure inefficiency. I also rule out one pure cognitive bias by considering forecasters to be overconfident in their ability (signal precision), which leads to a behavior response of overweighting private information. I find it unlikely that overconfidence can be the sole explanation for the observed behavior. For overconfidence to explain the observed behavior, we need forecasters to think they are around 100 times better than their competitors.

As an alternative, I propose an explicit cost from forecast revisions to capture a credibility loss from making large revisions. Given that we observe forecast smoothing in the data, I find it intuitive to consider a "backward-looking" extension to the model. A revision cost and horizon discounting, can together generate a strategic trade-off between horizons and allow us to match the observed horizon heterogeneity in data. The idea behind the revision cost is that large revisions motivated by work from competitors are costly since they could hurt a forecaster's credibility.

There are other types of strategic behavior among macroeconomic forecasters in the forecasting literature. For example, Laster et al. (1999) introduce a trade-off problem for the forecasters where they have to balance accuracy (minimizing forecast errors) and publicity. Ottaviani and Sørensen (2006) and Marinovic et al. (2013) generate strategic behavior by considering a forecast contest where forecasters benefit from being the only correct one. In a political economy setting, Cipullo and Reslow (2019) show that it is optimal for forecasters with economic interests and voter influence to publish biased forecasts before a referendum to influence voting outcomes. Hence, the notion of strategic behavior in addition to forecast error minimization among macroeconomic forecasters is not novel.

Why should we care about the behavior of forecasters? In economics, we often assume forward-looking individuals. If we assume that some of them make use of macroeconomic forecasts from professional institutions to form their expectations, this might have implications for the economy. One recent example of this is Tanaka et al. (2018), who use Japanese firm data to show that firms' GDP forecasts are positively and significantly associated with firms' input choices, such as investment and employment. Hence, inefficient forecasts can result in inefficient firm decisions. Also, in more general terms, all users of forecasts should have a strong interest in knowing about the quality of the forecasts and any alternative motives that could produce inefficient or biased forecasts.

This paper proceeds as follows. Section 2 describes the data at hand. Section 3 outlines the empirical strategy. Section 4 presents the results, and Section 5 investigates potential heterogeneities.

A baseline model is established in Section 6 while Section 7 extends the model to understand the observed inefficient behavior. Finally, Section 8 concludes.

2 Data

This paper utilizes forecaster firm-level forecast data from a monthly survey by Consensus Economics.[4] The data used in this paper covers 10 countries—Canada, France, Germany, Italy, Japan, Netherlands, Spain, Sweden, the United Kingdom and the United States and span forecasts made between 1995 and 2017. The forecasts refer to forecasts for the current and next year's annual growth rates, where, for example, the forecasters were asked in each month during 2017 to forecast the 2017 and 2018 annual growth rates. The main analysis is done using forecasts for two variables: GDP growth and inflation (consumer prices). This results in a dataset that contains forecasts for the GDP growth and inflation in 10 countries, thus creating 20 different outcome series. For the UK, the inflation forecasts refer to both RPIX and HICP (CPI). The RPIX covers the entire period, but towards the end of the sample many forecasters produce a forecast for both measures while others change entirely to HICP. I make use of all available price forecasts, adding one outcome series.

GDP growth and inflation outcome data are collected from the OECD database. For inflation, I use the latest available numbers; for GDP growth, I use the first available release (often referred to as real-time data). GDP is often subject to large revisions, and we often use real-time data when evaluating forecast performance for variables that are subject to large revisions.[5] Outcome data for RPIX is collected from the UK Office for National Statistics.

Table 1 reports firm number descriptive statistics. The sample consists of a total of 591 forecasting firms.[6] Some firms produce a forecast for more than one country. In these cases, the firms are treated as separate entities and are defined at the country level. Hence, the country label identifies the target outcome that is forecasted. Many of the 591 firms are not observed during the entire sample period, and the panel is therefore unbalanced. The columns *Mean*, *Median*, *Min*, and *Max* in Table 1 report the firm number statistics from the monthly survey in such a way that the total mean of 19.8 is an indication that each survey month has on average 19.8 forecasters. The firms in the survey are almost exclusively private firms such as financial institutions (banks) and private research or consulting institutions.

[4] The data is copyright protected against redistribution but can be purchased at http://www.consensuseconomics.com.
[5] Real-time GDP growth data is constructed using the OECD revision database.
[6] Some firm name entries in the raw data have been corrected due to spelling and naming errors.

Table 2 reports descriptive statistics for the forecasts, while Tables A1 and A2 in the appendix report descriptive statistics for the outcomes and forecast errors. From the data, a total of 179,310 observations can be used. From Table 2, we see that the number of observations differs between countries; the most observations are from Germany, the USA, and the UK, and the fewest are from the Netherlands. The average mean GDP growth forecast of 1.9 and the average mean inflation forecast of 1.8 are similar, while the standard deviation in the GDP growth forecasts is slightly higher. The number of observations is well-balanced between the two variables.

Table 1: Firm Number Statistics

Country	N	Mean	Median	Min	Max
Canada	46	14.6	15	8	19
France	57	18.7	18	10	26
Germany	63	26.8	27	19	32
Italy	55	14.5	14	4	21
Japan	72	19.5	19	10	25
Netherlands	42	9.9	9	5	15
Spain	45	14.2	14	5	19
Sweden	42	12.9	13	7	17
UK	94	24.4	24	16	37
USA	75	25.0	25	15	32
Total	591	19.8	20	4	37

Notes: Forecasters from the monthly survey by Consensus Economics between 1995 and 2017. N refers to the total number of firms in the entire sample, while mean, median, min and max refer to firms by survey month.

Table 2: Forecast Statistics

		GDP growth			**Inflation**		
Country	Obs.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Canada	14,872	7,457	2.4	1.0	7,415	1.9	0.5
France	18,600	9,331	1.6	1.0	9,269	1.4	0.6
Germany	27,481	13,739	1.6	1.2	13,742	1.6	0.6
Italy	14,336	7,228	1.1	1.3	7,108	1.9	1.0
Japan	18,450	9,409	1.2	1.5	9,041	0.2	0.9
Netherlands	9,586	4,959	1.7	1.3	4,627	1.9	0.8
Spain	13,927	7,011	2.0	1.7	6,916	2.2	1.2
Sweden	12,885	6,469	2.4	1.3	6,416	1.6	0.9
UK	23,663	12,539	2.0	1.2	11,124	2.7	0.8
USA	25,510	12,778	2.6	1.1	12,732	2.2	0.9
Total	179,310	90,920	1.9	1.4	88,390	1.8	1.0

Notes: Forecasts from the monthly survey by Consensus Economics between 1995 and 2017. In each month, forecasters are asked to forecast the current and next year's annual average growth rate. This results in 24 potential forecast origins (and horizons) for each year's outcome.

3 Empirical Strategy

Nordhaus (1987) presents a concept of weak and strong forecast efficiency. Consider x_t to be the target outcome and F_{ith} to be the forecast for x_t made by forecaster i, at a horizon of h periods before the realization of x_t. Define *strong efficiency* to be if $\mathbb{E}_{ith}[(x_t - F_{ith})^2|I_{ith}]$ is minimized, where I_{ith} is all information available to individual i, h periods before the realization of x_t. From the minimization argument, it also holds that the forecasts are unbiased: $\mathbb{E}_{ith}[x_t - F_{ith}|I_{ith}] = 0$. The overall intuition of efficiency is simple: available information should not explain the forecast errors.[7] Nordhaus (1987) argues that we should consider a weaker concept where we instead of the full information set I_{ith} only consider the forecaster's own past forecast. Define *weak efficiency* to be if $\mathbb{E}_{ith}[(x_t - F_{ith})^2|F_{ith+1}]$ is minimized, where F_{ith+1} is the own past (lagged) forecasts. Weak efficiency is nested in strong efficiency and is, of course, a much weaker concept. However, an appealing feature about the weaker test is that it can be hard to define the information set I_{ith} in the case of strong efficiency. Claiming that the past forecast was in the information set is not that strong of an assumption.

If we continue to assume that forecasters know their own past forecasts and further assume that they know the past forecasts of their competitors, $Z_{ith+1} = \frac{1}{n-1}\sum_{j\neq i} F_{jth+1}$, then, we have an augmented weak efficiency definition of minimization of $\mathbb{E}_{ith}[(x_t - F_{ith})^2|F_{ith+1}, Z_{ith+1}]$, which also implies that $\mathbb{E}_{ith}[x_t - F_{ith}|F_{ith+1}, Z_{ith+1}] = 0$.[8] Estimating

$$x_t - F_{ith} = \beta_0 + \beta_1 F_{ith+1} + \beta_2 Z_{ith+1} + \varepsilon_{ith} \tag{1}$$

should then yield $\beta_0 = \beta_1 = \beta_2 = 0$ to not rule out efficiency.[9] $\beta_2 > 0$ can be interpreted as underuse of competitors' information, while $\beta_2 < 0$ can be interpreted as overuse of competitors' information. To understand this interpretation, decompose equation (1) in two equations

$$x_t = \gamma_0 + \gamma_1 F_{ith+1} + \gamma_2 Z_{ith+1} + \epsilon_{ith} \tag{2}$$

$$F_{ith} = \alpha_0 + \alpha_1 F_{ith+1} + \alpha_2 Z_{ith+1} + \xi_{ith}, \tag{3}$$

where equation (2) tells us the ex-post optimal behavior that would have minimized the error and

[7] This has a direct link to the rational expectations literature, Muth (1961), and the Efficient Markets Hypothesis literature, Fama (1970).

[8] I calculate the mean of others, of those that are available in each survey month, due to the unbalanced nature of the data. n is the number of forecasters, including forecaster i.

[9] $\beta_0 = \beta_1 = \beta_2 = 0$ is necessary for strong and weak efficiency but not sufficient for claiming strong efficiency.

equation (3) tells us the actual behavior. The difference between the coefficients in equation (2) and the coefficients in (3) is the coefficients found in equation (1) such that $\beta = \gamma - \alpha$.

As mentioned before, the exercise described above resembles some existing work in the literature. For example, Batchelor and Dua (1992), Lahiri and Sheng (2008), Deschamps and Ioannidis (2013), Coibion and Gorodnichenko (2015), Dovern et al. (2015), Bordalo et al. (2018) and Broer and Kohlhas (2018) are all papers that follow the same spirit of evaluating forecasters' rationality and optimal use and reaction to information. The empirical part of this paper differs from the existing literature since the focus is on estimating the *relative* use of own versus others' information and not on general estimation of over- and underreaction to information or over- and underuse of the own lagged forecast. Also, the rich dataset used in this paper allows for a large number of observations that provide valuable precision in the estimates. It also allows for the high-precision study of heterogeneities such as variable, country, year and horizon.[10]

4 Results

Table 3 reports the results from estimating equations (1), (2) and (3). In column (1), the dependent variable is the forecast error, and we expect zero coefficients to not rule out efficient use of information. We interpret a positive coefficient on *Lagged Others* as underuse of the competitors' forecasts, and a negative coefficient is interpreted as overuse of information. Hence, from the table, we interpret the -0.762 as overuse of the lagged own forecast while we interpret the 0.721 as underuse of the lagged forecast from others. Due to the information structure of annual averages and the assumption that forecasters have access to each other's estimates, the forecasts are assumed to potentially be serially correlated and correlated across firms within the same survey month. To combat this, the standard errors are robust to two-way clustering (Cameron et al. (2011) and Cameron and Miller (2015)) at the forecaster (firm) level and the survey month level.

Columns (2) and (3) report the estimated coefficients from equations (2) and (3), respectively. The coefficients in column (2) can be interpreted as the suggested optimal behavior; i.e. what they should have done. The estimates suggest that it would be optimal to put very high attention (0.923) on the *Lagged Others* while putting very low attention on the own *Lagged Forecast* (0.054). The estimates suggest that others' forecasts should get 17 times a higher weight. Column (3) reports the estimated

[10] In the Appendix, Section A.2, I estimate some alternative specification to equations (1), (2) and (3) as robustness checks to compare and relate to the existing literature. Note that some papers use a different type of forecast data—data with a fixed horizon instead of a fixed target as in this paper.

actual behavior where we see that the own *Lagged Forecast* gets a very high weight (0.816), while *Lagged Others* gets a relatively small weight (0.202). The difference between columns (2) and (3) give us the error found in column (1). For example, for the *Lagged Others* we have that $0.923 - 0.202 = 0.721$, and hence indicates underuse of forecasts from competitors. Hence, the forecasts appear to be too persistent, and forecasters do not seem to realize the importance of pooling information from others.

To ensure that the estimation results in Table 3 are robust, I perform a sensitivity analysis regarding the actual behavior estimates. Column (1) of Table 4 re-reports column (3) from Table 3. Column (2) then adds forecaster (firm) fixed effects, while column (3) adds country fixed effects. Since forecasters are defined at the country level, the country fixed effects will be nested in the firm fixed effects. Hence, there is no reason to include both at the same time. Column (4) has firm and variable (GDP and inflation) fixed effects, while column (5) adds time effects, represented by target year and survey month. Column (6) adds horizon fixed effects. Lastly, one concern is that we are missing controls for some common information received in period t. There might be some information (for example actual raw data) that is released in period t and observed by everyone. To control for this, I also include country times survey month fixed effects. These effects absorb all common information in each period. From the estimates in columns (1) to (7), it is clear that the estimated actual behavior is very robust. It would make little sense to replicate the same exercise for optimal behavior since we, in that case, are interested in the counterfactual optimal behavior from using lagged forecasts. See also Appendix Section A.2 for some additional alternative specifications that capture a more general reaction to new information.

Table 3: Decomposing the Efficiency Error

Dep. Variable:	(1) Forecast Error	(2) Outcome	(3) Forecast
Lagged Forecast	-0.762***	0.054**	0.816***
	(0.024)	(0.027)	(0.008)
Lagged Others	0.721***	0.923***	0.202***
	(0.027)	(0.031)	(0.010)
Observations	179,310	179,310	179,310
R^2	0.066	0.543	0.966

Notes: Standard errors robust to two-way clustering at the forecaster (firm) level and the survey month level. Estimated equations are (1), (2) and (3). *,**,*** represent the 10%, 5%, 1% significance levels.

Table 4: Robustness of Actual Behavior

Dep. Variable:	(1) Forecast	(2) Forecast	(3) Forecast	(4) Forecast	(5) Forecast	(6) Forecast	(7) Forecast
Lagged Forecast	0.816***	0.804***	0.816***	0.804***	0.804***	0.804***	0.803***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Lagged Others	0.202***	0.218***	0.205***	0.219***	0.216***	0.216***	0.221***
	(0.010)	(0.011)	(0.010)	(0.011)	(0.009)	(0.009)	(0.009)
Observations	179,310	179,310	179,310	179,310	179,310	179,310	179,310
R^2	0.966	0.967	0.966	0.967	0.970	0.970	0.973
Firm		✓		✓	✓	✓	✓
Country			✓				
Variable				✓	✓	✓	✓
Target Year					✓	✓	✓
Survey Month					✓	✓	✓
Horizon						✓	✓
Country × Survey Month							✓

Notes: Standard errors robust to two-way clustering at the forecaster (firm) level and the survey month level. Estimated equation is (3) adding the different fixed effects. *,**,*** represent the 10%, 5%, 1% significance levels.

5 Heterogeneity in Actual Behavior?

The results in the previous section suggest that forecasters are inefficient in their use of competitors' forecasts. The forecasters pay too little attention to their competitors, resulting in forecasts that are too persistent. Specifically, forecasters allocate an 80 to 20 relative weight between their own and others' forecasts in the forecasting process. The general result of underreaction to new information and too much attraction to the own lagged forecast is hence in line with previous literature. The results presented in Tables 3 and 4 are obtained using pooled data over country, variable, time and horizon. To further ensure the robustness of the results, this section investigates potential heterogeneity in the actual behavior. The rich dataset at my disposal allows for a high-precision heterogeneity analysis.

5.1 Country and Year

The first heterogeneities investigated are country and time. Forecasters might behave differently in different countries. Forecasters might also have changed their behavior over time. The graphs in Figure 1 plot the estimated coefficients for the different heterogeneities. The estimation is done by restricting the sample selection to each country and year respectively. The solid lines correspond to the pooled estimates from Table 3 and act as a reference point. From the graphs, we do not observe any big heterogeneities regarding either country or year. The forecasters seem to behave the same regardless of country, and the behavior is stable over time.

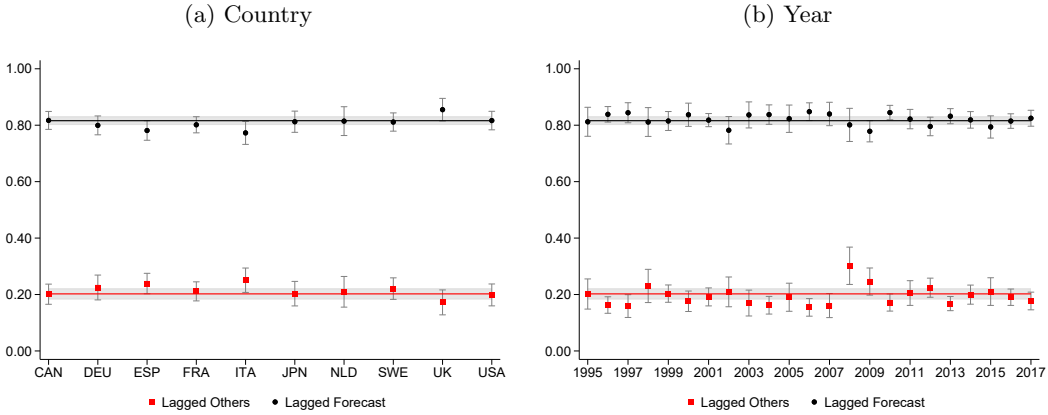

Figure 1: Country and Year Analysis

Notes: Solid lines refer to the estimates from the pooled regression in column (3) of Table 3. The dots refer to estimates from equation (3) using the different subsamples. The 95% confidence intervals are constructed using standard errors robust to two-way clustering at the forecaster (firm) level and the survey month level.

5.2 Variable

To ensure that the results are stable between different variables, I also estimate equation (3) using each variable separately. Column (1) in Table 5 shows the pooled baseline results for reference. Columns (2) and (3) show the results from estimating the main specification in equation (3) using variable specific subsamples. Column (2) shows the GDP results, while column (3) shows the inflation results. From the table, we see that the coefficients are very similar.

The obtained dataset from Consensus Economics contains many more variables then GDP growth and inflation. This allows for an easy extension to also consider the actual behavior regarding several more variables. Therefore, I also estimate equation (3) using subsamples based on several other variables.[11] Tables 6 and 7 present the results for a large number of variables. We can first note that the results are extremely consistent across variables. The *Lagged Forecast* has a coefficient of around 0.8, while *Lagged Others* receives a coefficient of around 0.2. We can also note that many variables have a much smaller sample than GDP and inflation have. This is due to several factors. First, a smaller number of forecasters in each country do a forecast for many of these variables. It is also the case that some of these variables are only observed in a small number of countries. Some variables are also only present during a small period of the sample, and many of them change definition over time. However, the results confirm that the behavior is very consistent—forecasters put a high weight on their own lagged forecast and a low weight on their competitors' lagged forecasts.

[11] Some variables refer to levels instead of annual growth rates.

Table 5: Variable Analysis

	Pooled	GDP growth	Inflation (CPI)
	(1)	(2)	(3)
Dep. Variable:	Forecast	Forecast	Forecast
Lagged Forecast	0.816***	0.831***	0.792***
	(0.008)	(0.010)	(0.007)
Lagged Others	0.202***	0.192***	0.219***
	(0.010)	(0.013)	(0.008)
Observations	179,310	90,920	88,390
R^2	0.966	0.968	0.963

Notes: Standard errors robust to two-way clustering at the forecaster (firm) level and the survey month level. Estimated equations are (3). *,**,*** represent the 10%, 5%, 1% significance levels.

Table 6: Variable Heterogeneity, Part 1

	Consumpt.	Wholesale	Invest.	Equip. Invest.	Indust. Prod.	Manuf. Prod.	Mining Prod.	Corp. Profits
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable:	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast
Lagged Forecast	0.82***	0.79***	0.82***	0.79***	0.76***	0.82***	0.85***	0.81***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Lagged Others	0.19***	0.24***	0.21***	0.25***	0.27***	0.20***	0.20***	0.23***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	90,368	8,020	66,795	20,123	54,559	13,971	1,970	21,957
R^2	0.96	0.92	0.95	0.93	0.95	0.94	0.87	0.90

Notes: Standard errors robust to two-way clustering at the forecaster (firm) level and the survey month level. Estimated equations are (3). *,**,*** represent the 10%, 5%, 1% significance levels.

Table 7: Variable Heterogeneity, Part 2

	Empl.	Unempl.	Wages	New Car	Car Sales	Housing Starts	Current Account	Producer Prices
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable:	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast
Lagged Forecast	0.81***	0.79***	0.85***	0.93***	0.79***	0.82***	0.67***	0.81***
	(0.02)	(0.01)	(0.01)	(0.03)	(0.02)	(0.03)	(0.07)	(0.01)
Lagged Others	0.19***	0.21***	0.15***	0.02	0.22***	0.18***	0.33***	0.23***
	(0.02)	(0.01)	(0.01)	(0.03)	(0.02)	(0.03)	(0.06)	(0.01)
Observations	6,320	69,425	48,653	1,595	10,541	24,100	75,068	36,438
R^2	0.94	1.00	0.97	0.91	0.98	1.00	0.95	0.92

Notes: Standard errors robust to two-way clustering at the forecaster (firm) level and the survey month level. Estimated equations are (3). *,**,*** represent the 10%, 5%, 1% significance levels.

5.3 Horizon

The results in Table 3 also pool over the forecast horizon. The main specification estimates an average behavior over horizons 1 to 23 months. However, results from Lahiri and Sheng (2008) suggest that there is heterogeneity in forecast behavior over horizons. As before, I estimate the main specification using a sample selection for each horizon.

The results are presented as dots in Figure 2. Again, the solid lines correspond to the pooled results. We see that on longer horizons the weight on *Lagged Others* is small and the weight on the own *Lagged Forecast* is high. As the horizon decreases, the weights converge and end up being equal at horizon one. Hence, as the horizon decreases, forecasters behave more optimally, but still far from optimal, since optimal would be a high weight on *Lagged Others*. The estimated optimal weight on the own *Lagged Forecast* is around 0.05 and hence far from the actual behavior on both long horizons and short horizons, even though the error is smaller on shorter horizons.

The fact the forecasters behave differently at different horizons might suggest some strategic decisions at play. It might be that alternative motives are stronger when we are far away from the realization of the outcome, and as forecasters approach the realization they switch to concerns regarding the size of the forecast error.

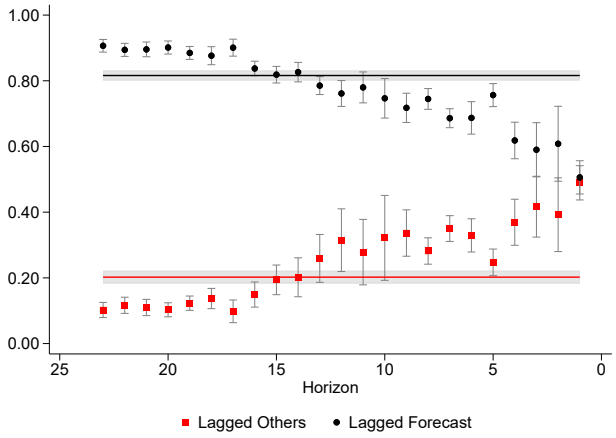

Figure 2: Horizon Analysis

Notes: Solid lines refer to the estimates from the pooled regression in column (3) of Table 3. The dots refer to estimates from equation (3) using the different horizon subsamples. The 95% confidence intervals are constructed using standard errors robust to two-way clustering at the forecaster (firm) level and the survey month level.

6 Baseline Model to Understand Optimal Behavior

The empirical results suggest inefficient behavior among forecasters. Specifically, the results suggest that it is optimal to put a relatively high weight on the lagged forecast of others (the mean of the competitors) compared to the forecasters' own lagged forecast. The decomposition performed in this paper tells us that, on average, the optimal ratio between others' forecasts and own forecasts should be around 17.

This section presents a baseline model to help us understand these empirical predictions of the suggested optimal behavior. Hence, the model is used to ensure that the empirical prediction of optimal behavior is reasonable. I will consider two versions of the model: one model that has fully informed forecasters that perfectly filters out all the signals and one model with forecasters that follow a simplified "rule of thumb." The rule of thumb will be such that forecasters, instead of filtering out signals, will utilize the lagged actual forecasts as "reduced form" aggregates. Hence, the rule of thumb model is more in line with the empirical approach.

6.1 A Model with Fully Informed Forecasters

Consider a three-period model where nature draws x at the dawn of time. Assume that we have n forecasters, indexed by i, where all try to forecast x. Each forecaster will produce a forecast in each of the three periods such that each forecaster has three attempts at predicting x.[12] At the end of time, the true value is revealed and each forecaster faces the loss

$$L_i = (x - F_{i3})^2 + (x - F_{i2})^2 + (x - F_{i1})^2, \tag{4}$$

where F_{ih} is forecaster i's forecast of x at Horizon h (h periods before the realization of x). Each forecaster minimizes the loss with respect to F_{i3}, F_{i2} and F_{i1}, where we from the first-order conditions have that $F_{ih} = \mathbb{E}_{ih}[x]$. The forecasters have a common prior about x that follows the true distribution $x \sim N(\mu, \sigma_x^2)$. In each period, each forecaster receives a private noisy signal $S_{ih} = x + \epsilon_{ih}$, where $\epsilon_{ih} \sim N(0, \sigma_{\epsilon_h}^2)$. The level of noise $\sigma_{\epsilon_h}^2$ is assumed to be the same for everybody.

In Period One, Horizon 3, each forecaster receives the private signal and is called upon to declare a forecast. The best each forecaster can do is to forecast the expected value given the information set

[12] Note that the model considers just one realization of x such that the t indexation from the empirical exercise is omitted.

(containing the prior and the signal), $\Omega_{i3} = \{\mu, S_{i3}\}$. Hence, the forecast is

$$F_{i3} = \mathbb{E}_{i3}[x|\Omega_{i3}] = \frac{p_\mu}{p_\mu + p_{s_3}}\mu + \frac{p_{s_3}}{p_\mu + p_{s_3}}S_{i3}, \tag{5}$$

where we have used the definition $p_v \equiv \frac{1}{\mathbb{V}[v]}$, such that p is a precision parameter defined as one over the variance. Hence, $p_\mu = \frac{1}{\sigma_x^2}$ and $p_{s_3} = \frac{1}{\sigma_{\epsilon_3}^2}$.[13] To simplify the notation, rewrite the forecast as

$$F_{i3} = w_{\mu|3}\mu + w_{s_{3|3}}S_{i3}, \tag{6}$$

such that w is a relative weight where the weights sum to one.[14] At Horizon 2, each forecaster receives a new private signal and observes the Horizon 3 forecasts of their competitors. The forecaster filters out the signals from the competitors and produces a new forecast:[15]

$$F_{i2} = w_{\mu|2}\mu + w_{s_{3|2}}S_{i3} + \sum_{j\neq i} w_{s_{3|2}}S_{j3} + w_{s_{2|2}}S_{i2}. \tag{7}$$

Following the Horizon 2 structure we will have the Horizon 1 forecast

$$F_{i1} = w_{\mu|1}\mu + \sum_{q=2}^{3} w_{s_{q|1}}S_{iq} + \sum_{j\neq i}\sum_{q=2}^{3} w_{s_{q|1}}S_{jq} + w_{s_{1|1}}S_{i1}, \tag{8}$$

where we see that a forecast is a weighted average of the prior, the own lagged signals, the lagged signals of others, and the own new private signal.[16]

If we allow for an arbitrary number of horizons (H) and outcomes (t), we can rewrite any forecast in a compressed form as

$$F_{ith} = w_{\mu|h}\mu + \sum_{q=h+1}^{H}\sum_{j=1}^{n} w_{s_{q|h}}S_{jtq} + w_{s_{h|h}}S_{ith}. \tag{9}$$

Given the information structure under fixed target forecasts, the information set will grow over horizons such that as the forecast horizon becomes shorter, the error is expected to be smaller.[17] From (9) we see that a forecast in the model is a weighted average of all signals and the prior. The model does not

[13] Equation (5) follows from Bayesian updating.
[14] Hence, $w_{\mu|3}$ is the prior weight at Horizon 3 and $w_{s_{3|3}}$ is the weight on the Horizon 3 signal at Horizon 3.
[15] The forecaster is assumed to have full knowledge about the prior structure and the precision in all signals. Hence, fully informed forecasters have the ability to filter out the signal from the observed forecasts.
[16] The weights follow the structure presented in Equation (5).
[17] See Andersson et al. (2017) for an analysis of the expected size of forecast errors at different horizons.

provide an expression of the form that is used in the empirical exercises,

$$F_{ith} = \beta_0 + \beta_1 F_{ith+1} + \beta_2 Z_{ith+1} + \varepsilon_{ith}, \tag{10}$$

where the forecast directly depends on the lagged forecasts and not the signals. However, the model does provide a prediction regarding the relationship between β_1 and β_2 in the reduced form estimation.

Consider that we only have three forecasters (a, b and c) such that $n = 3$. If we then write the Horizon 2 forecast from equation (7) for forecaster a in terms of lagged forecast instead of lagged signals, we get

$$F_{a2} = \omega_\mu \mu + \omega_{f3} F_{a3} + \omega_{z3} Z_{a3} + \omega_{s2} S_{i2}, \tag{11}$$

where $Z_{a3} = \frac{1}{2}\big(F_{b3} + F_{c3}\big)$. To replicate the individual signal weights, we then need that $\omega_{f3} = \frac{p_\mu + p_{s_3}}{P_2}$, $\omega_{z3} = 2\frac{p_\mu + p_{s_3}}{P_2}$, $\omega_{s2} = \frac{p_{s_2}}{P_2}$ and $\omega_\mu = -2\frac{p_\mu}{P_2}$, where $P_2 = p_\mu + 3p_{s_3} + p_{s_2}$. Hence, we have that $\omega_{z3} = (n-1)\omega_{f3}$ such that the relationship between the own lagged forecast and the lagged forecast of the others is proportional in the number of forecasters.[18]

Hence, we should expect $\beta_2 = (n-1)\beta_1$ in the estimation such that the relative weight between the own forecast and that of others should be dependent on n (the number of forecasters). In the empirical estimation, I found that it would be optimal to have a ratio of 17 between others' forecasts and the own forecasts, implying in the model that we have around 18 forecasters. This lines up very well with the firm number statistics in Table 1, where we can read that we have on average 19.8 forecasters.

6.2 Forecasters with "Rule of Thumb"

The baseline model above assumed highly advanced forecasters with full information about the signal structure. In this section, I present an alternative model that imposes some form of "bounded ability" to ensure that we do not ask too much from our forecasters. I follow the setup in the fully informed model, but assume now that forecasters cannot, or at least do not, filter out the individual signals. Hence, I assume now that forecasters just use some simplifying rule of thumb where they make a weighted average of the lagged forecasts (own and others') and the new private signal.

Assume the following structure. At Horizon 3, each forecaster receive the private signal and declares

[18] Expanding this to more horizons would require that we include all lagged forecasts and not only the latest. The per-horizon relationship between the own forecasts and those of others, however, do stay the same. The relationship always follows $\omega_{zh} = (n-1)\omega_{fh}$.

a forecast just as before:

$$F_{i3} = \mathbb{E}_{i3}[x|\Omega_{i3}] = \frac{p_\mu}{p_\mu + p_{s_3}}\mu + \frac{p_{s_3}}{p_\mu + p_{s_3}}S_{i3}. \tag{12}$$

In Period Two, each forecaster receives a new private signal and observes the Horizon 3 consensus forecast of the competitors (Z_{i3}). The information set is now the own lagged forecast, the lagged consensus forecast of others, and the new private signal, $\Omega_{i2} = \{F_{i3}, Z_{i3}, S_{i2}\}$. Hence, this setup treats Z_{i3} as a single signal such that the Horizon 2 forecast will be

$$F_{i2} = \frac{p_{f_3}}{p_{f_3} + p_{z_3} + p_{s_2}}F_{i3} + \frac{p_{z_3}}{p_{f_3} + p_{z_3} + p_{s_2}}Z_{i3} + \frac{p_{s_2}}{p_{f_3} + p_{z_3} + p_{s_2}}S_{i2}, \tag{13}$$

where we again will have that we can simplify the notation to

$$F_{i2} = w_{f2}F_{i3} + w_{z2}Z_{i3} + w_{s2}S_{i2}. \tag{14}$$

The Horizon 1 forecast will then be

$$F_{i1} = w_{f1}F_{i2} + w_{z1}Z_{i2} + w_{s1}S_{i1}. \tag{15}$$

The optimal relationship between the weights, again, will depend on the number of forecasters such that $w_{zh} = (n-1)w_{fh}$.[19] This is the same prediction as in the fully informed model.

However, if a forecaster follows this process by not filtering out the actual signals, the forecast will be less efficient compared to the fully informed forecast. This is because a forecast using aggregate information such as lagged forecasts will have nested information where old information is weighted too much compared to new private information.[20] Hence, a fully efficient forecaster needs to filter out all the signals.

[19] As before, I assume that everyone has the same signal precision in private information. I also assume that each forecaster assumes that everyone is equally good.

[20] A simple example to consider is the usage of a common prior. If all forecasters have the same prior and then put a weight on each other's forecasts, we will have that the prior is weighted too much relative to private information. This prior problem is noted by Kim et al. (2001), where the authors show analytically that consensus forecasts are inefficient since too much weight is attributed to analysts' common information relative to their private information. This is further highlighted in Crowe (2010), where the author argues that consensus forecasts are inefficient due to over-weighting older information at the expense of new private information.

7 Extending the Model to Understand the Actual Behavior

From the empirical exercise, I found that forecasters appear to be inefficient in their use of competitors' forecasts. Two facts are of interest. First, the forecasts are too persistent. Second, the forecasters change their behavior over horizons. The results show that the actual behavior of forecasters is such that they put a high weight on their own lagged forecast and this weight is different at different horizons. The empirical exercise also provided the relative use between others' versus own lagged forecasts. This is of good use since the model framework provides a prediction for that exact relationship. The estimated actual behavior is a ratio of 0.25 while the optimal was estimated to be 17. This section aims to extend the model to understand why forecasters seem to behave inefficiently compared to the baseline model. It could be that we have incorrect assumptions about the loss function of forecasters that make us believe they are inefficient when they are not.[21]

The fact that forecasters overuse their own signal might suggest that they are overconfident in their own ability. This is reasonable to believe since this behavior bias often is called "the mother of all biases." Overconfidence is well studied in psychology, and De Bondt and Thaler (1995) state that: "perhaps the most robust finding in the psychology of judgment is that people are overconfident."[22] However, I find it unlikely that overconfidence is the only explanation since this would require unreasonable levels of overconfidence. A simple example can provide some intuition about the level.

Assume a forecaster is overconfident about their own signal. The beliefs are such that their own precision is perceived to be κp_i, where $\kappa > 1$.[23] Continue to assume that forecasters have correct beliefs about others and that they correctly can filter out the true signals from their competitors. This will result in a higher relative weight on their own signals.

To understand the level of κ, assume that we have 20 forecasters. A non-overconfident forecaster will put $1/20 = 0.05$ relative weight on the own signal and $19/20 = 0.95$ on the others' combined. If the forecaster has a κ of 2, this reflects that the own ability is twice that of competitors, and the weight structure changes to 0.095 on the own forecast and 0.905 on the others' forecasts. With $\kappa = 10$, we get 0.34 on the own and 0.66 on the others', and with $\kappa = 100$ we have 0.84 on the own forecast and 0.16 on the others'. Hence, with many forecasters, we need a very high κ to replicate the relationships found in the data. Overconfidence is then most likely not a reasonable (sole) explanation since the κ levels needed seem to be a bit extreme. It is also the case that overconfidence does not provide any

[21] However, they are still going to be inefficient in minimizing forecast errors, but this might not be their only objective.
[22] See also Kahneman and Tversky (1972) and Tversky and Kahneman (1974) for more on this type of cognitive bias.
[23] See, for example Daniel et al. (1998) for similar modeling of overconfidence in the securities markets.

strategic trade-off that produces horizon dynamics as those found in the data.

To explain both the general underuse of competitors' forecasts, as well as the horizon dynamics in this behavior, I instead consider the introduction of a revision cost. Assume that a forecaster releases a first estimate of the GDP growth in 2018 of 1 percent. After one month, the forecaster observes the forecasts of the competitors and realizes that the consensus view of the others is a GDP growth of 3 percent. When the forecaster releases the next estimate, it would be optimal to revise up to the 3 percent mark, but this would imply a forecast revision of 2 percentage points, subject to the new private signal. A revision that large would need substantial explanation since no big news has been uncovered. Telling the public that the forecast last month was wrong after seeing the forecasts of others would probably be a big hit to the forecaster's credibility.[24] What the forecaster could do instead is to slowly revise toward the others to preserve credibility and hide the big revision in many small steps.

It is important to note that the revision cost can be a product of forecasters' responses to the demand side of the forecasting business. Hence, the end user of the forecasts might demand, or reward, consistency and thus punish inconsistent forecasters.

Introducing a cost from revisions gives rise to a trade-off between loss from accuracy and loss from revisions. Together with horizon discounting, this trade-off introduces a horizon trade-off that results in a higher degree of "forecast smoothing." The next section outlines the technical details of the revision cost extension.

7.1 Revision Cost to Preserve Credibility

In the baseline model, it was assumed that forecasters' sole purpose was to minimize the forecast error. I now add an explicit revision cost to the loss function. The baseline model consisted of three periods (horizons). I now extend the model to 23 horizons to understand the horizon dynamics.[25] Assume that the forecaster faces the following minimization problem:

$$\min_{F_{ih}} \quad \sum_{h=1}^{23} g(h)(\mathbb{E}_{ih}[x] - F_{ih})^2 + r(h)(F_{ih} - F_{ih+1})^2. \tag{16}$$

[24] This follows some of the reasoning in Ehrbeck and Waldmann (1996), who present a similar mechanism in a repeated game where there is a trade-off between minimizing forecast errors and looking good before the outcome is realized.
[25] It is important to note that the predictions from the baseline model do not depend on the number of horizons.

In each horizon, the forecaster faces a trade-off between the minimization of the forecast error and the revision cost.[26] The forecast error is horizon discounted by the function $g(h)$ such that $g(h) > g(h+1)$ to account for the fact that it is harder to forecast at longer horizons.[27] The revision cost is also horizon discounted such that $r(h) > r(h+1)$. The relation between $g(h)$ and $r(h)$ then determines the relative cost between forecast errors and revisions. From the first-order condition of the minimization problem, we have that a forecast at Horizon h is given by

$$F_{ih} = \frac{g(h)}{g(h)+r(h)+r(h-1)}\,\mathbb{E}_{ih}[x] + \frac{r(h)}{g(h)+r(h)+r(h-1)}F_{ih+1} + \frac{r(h-1)}{g(h)+r(h)+r(h-1)}\,\mathbb{E}_{ih}[F_{ih-1}]. \tag{17}$$

Note that $\mathbb{E}_{ih}[x]$ is formed as in the baseline model. It is a weighted average of the prior and the signals as presented in the baseline model. Solving backward, we have that the Horizon 1 forecast is[28]

$$F_{i1} = \frac{g(1)}{g(1)+r(1)}\,\mathbb{E}_{i1}[x] + \frac{r(1)}{g(1)+r(1)}F_{i2}. \tag{18}$$

Solving the Horizon 2 forecasts then yields

$$F_{i2} = \frac{g(2)}{g(2)+r(2)+r(1)}\,\mathbb{E}_{i2}[x] + \frac{r(2)}{g(2)+r(2)+r(1)}F_{i3} + \frac{r(1)}{g(2)+r(2)+r(1)}\,\mathbb{E}_{i2}[F_{i1}], \tag{19}$$

substitute (18) into (19) and rearrange to get[29]

$$F_{i2} = \frac{g(2)+r(1)m_1}{g(2)+r(1)m_1+r(2)}\,\mathbb{E}_{i2}[x] + \frac{r(2)}{g(2)+r(1)m_1+r(2)}F_{i3}, \tag{20}$$

where $m_1 = \frac{g(1)}{g(1)+r(1)}$. Simplify the notation to

$$F_{i2} = m_2\,\mathbb{E}_{i2}[x] + (1-m_2)F_{i3}. \tag{21}$$

[26] In the first period, there is no lagged forecast. Hence, the solution in the first period is trivial and follows $F_{i24} = \mathbb{E}_{i24}[x]$.
[27] See Andersson et al. (2017) for an analysis of the expected size of forecast errors at different horizons. In the baseline model, the discounting does not have any implications at all since it would drop out from the first-order conditions.
[28] Note that by assumption $F_{i0} = x$ and $r(0) = 0$, since the outcome is known at horizon zero.
[29] Note that $\mathbb{E}_{i2}[x] \neq \mathbb{E}_{i1}[x]$, $\mathbb{E}_{i2}[\mathbb{E}_{i1}[x]] = \mathbb{E}_{i2}[x]$ and that $\mathbb{E}_{i2}[F_{i2}] = F_{i2}$.

Then, for a general Horizon h forecast, we have that

$$F_{ih} = \frac{g(h) + r(h-1)m_{h-1}}{g(h) + r(h-1)m_{h-1} + r(h)} \mathbb{E}_{ih}[x] + \frac{r(h)}{g(h) + r(h-1)m_{h-1} + r(h)} F_{ih+1}, \tag{22}$$

and simplify any expression of this form as

$$F_{ih} = m_h \mathbb{E}_{ih}[x] + (1 - m_h) F_{ih+1}, \tag{23}$$

where $(1 - m_h)$ can be seen as a smoothing parameter. In the empirical results, we found that m_h was decreasing in h such that $m_1 > m_2 > ... > m_{23}$. Hence, the two functions $g(h)$ and $r(h)$ need to be decreasing in h at different rates. Specifically, we need $g(h)$ to decrease faster than $r(h)$ as the horizon increases. The intuition for this is that the revision cost needs to be relatively lower compared to forecast error cost at long horizons compared to short horizons. It then becomes apparent that it is the relative size, and not the actual size, of $g(h)$ and $r(h)$ that is important, such that we could simplify and define $g(h) = 1 - r(h)$. Different functions of $g(h)$ (and/or $r(h)$) could then provide the dynamics needed to explain the empirical results.

8 Concluding Remarks

From the empirical exercise, I found that forecasters appeared to behave inefficiently since they underuse information from their competitors. This behavior results in too persistent forecasts given an assumption of forecast error minimization. The empirical results are robust for several heterogeneities, such as country, year and variable. However, a horizon decomposition revealed that forecasters behave differently at different forecast horizons. Specifically, they pay more attention to their competitors as the forecast horizon decreases, and hence behave more efficiently as they approach the realization of the outcome—but they are still far from an optimal level.

A simple model of noisy and private information was used as a framework to analyze the behavior of forecasters. Under the assumption of forecast error minimization, the model and the empirical exercise predicted that it is optimal for the individual forecaster to pay high attention to what the other forecasters do. These results originate from the fact that we assume that each forecaster has private information such that there is a benefit from pooling information from many forecasters. These results regarding the importance of pooling are no big surprise and are in line with standard findings

in the forecasting literature.

Previous studies have found that forecasters appear to put high weight on own lagged information such that they are too slow in incorporating new public information. In the empirical part of this paper, I focused on the estimation of the relative use of own lagged forecasts and the lagged forecasts of competitors. The results show that forecasters put four times a relative higher weight on the own lagged forecast compared to the forecasts of competitors, while the optimal would be a 17 times higher weight on the competitors' forecasts compared to the own lagged forecast.

To explain the observed inefficient behavior, the baseline model was extended. First, I find it unlikely that the cognitive bias of overconfidence is the source of the observed behavior since the level of overconfidence needed is very large. For overconfidence to explain the observed behavior, we need forecasters who think they are around 100 times better than their competitors. I instead proposed that forecasters face a loss from revisions. Rewriting the forecasters' loss function to include a cost from making revisions, together with horizon discounting, that is relatively different for the cost from forecast errors and revisions can explain the actual behavior. The revision cost introduces a horizon-varying, forecast-smoothing parameter that generates the persistence and horizon dynamics that we observed in the behavior. It is, of course, also possible (and likely) that some degree of overconfidence is present together with the proposed revision cost.

The idea behind the revision cost is that users of forecasts value consistency, and the reputation of the forecasters will be damaged if the forecaster justifies revisions based on forecasts from competitors. This results in a trade-off between accuracy and consistency over horizons for the forecasters—a trade-off that depends on the relative cost between revisions and forecast errors.

References

Ager, P., Kappler, M., and Osterloh, S. (2009). The Accuracy and Efficiency of the Consensus Forecasts: A Further Application and Extension of the Pooled Approach. *International Journal of Forecasting*, 25(1):167–181.

Andersson, M. K., Aranki, T., and Reslow, A. (2017). Adjusting for information content when comparing forecast performance. *Journal of Forecasting*, 36(7):784–794.

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

Batchelor, R. and Dua, P. (1992). Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting*, 11(2):169–181.

Bernhardt, D., Campello, M., and Kutsoati, E. (2006). Who Herds? *Journal of Financial Economics*, 80(3):657–675.

Bordalo, P., Gennaioli, N., Ma, Y., and Shleifer, A. (2018). Overreaction in Macroeconomic Expectations. Working Paper.

Broer, T. and Kohlhas, A. (2018). Forecaster (Mis-)Behavior. CEPR Discussion Papers 12898, C.E.P.R. Discussion Papers.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust Inference With Multiway Clustering. *Journal of Business & Economic Statistics*, 29(2):238–249.

Cameron, A. C. and Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372.

Chen, Q. and Jiang, W. (2006). Analysts' Weighting of Private and Public Information. *Review of Financial Studies*, 19(1):319–355.

Cipullo, D. and Reslow, A. (2019). Biased Forecast to Affect Voting Decisions? The "Brexit" Case. Working Paper 364, Sveriges Riksbank.

Coibion, O. and Gorodnichenko, Y. (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *American Economic Review*, 105(8):2644–78.

Crowe, C. W. (2010). Consensus Forecasts and Inefficient Information Aggregation. Technical report, International Monetary Fund Working Paper 10/178.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor Psychology and Security Market Under- and Overreactions. *The Journal of Finance*, 53(6):1839–1885.

De Bondt, W. F. and Thaler, R. H. (1995). Chapter 13 Financial Decision-making in Markets and Firms: A Behavioral Perspective. In *Finance*, volume 9 of *Handbooks in Operations Research and Management Science*, pages 385–410. Elsevier.

Deschamps, B. and Ioannidis, C. (2013). Can Rational Stubbornness Explain Forecast Biases? *Journal of Economic Behavior and Organization*, 92:141–151.

Dovern, J., Fritsche, U., Loungani, P., and Tamirisa, N. (2015). Information rigidities: Comparing average and individual forecasts for a large international panel. *International Journal of Forecasting*, 31(1):144–154.

Ehrbeck, T. and Waldmann, R. (1996). Why Are Professional Forecasters Biased? Agency versus Behavioral Explanations. *The Quarterly Journal of Economics*, 111(1):21–40.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417.

Gallo, G. M., Granger, C. W. J., and Jeon, Y. (2002). Copycats and Common Swings: The Impact of the Use of Forecasts in Information Sets. *IMF Staff Papers*, 49(1):4–21.

Isiklar, G., Lahiri, K., and Loungani, P. (2006). How quickly do forecasters incorporate news? Evidence from cross-country surveys. *Journal of Applied Econometrics*, 21(6):703–725.

Kahneman, D. and Tversky, A. (1972). Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology*, 3:430–454.

Kim, O., Lim, S. C., and Shaw, K. W. (2001). The Inefficiency of the Mean Analyst Forecast as a Summary Forecast of Earnings. *Journal of Accounting Research*, 39(2):329–336.

Lahiri, K. and Sheng, X. (2008). Evolution of Forecast Disagreement in a Bayesian Learning Model. *Journal of Econometrics*, 144(2):325–340.

Lamont, O. A. (2002). Macroeconomic Forecasts and Microeconomic Forecasters. *Journal of Economic Behavior & Organization*, 48(3):265–280.

Laster, D., Bennett, P., and Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. *The Quarterly Journal of Economics*, 114(1):293–318.

Lucas, R. (1973). Some International Evidence on Output-Inflation Tradeoffs. *American Economic Review*, 63(3):326–34.

Marinovic, I., Ottaviani, M., and Sørensen, P. N. (2013). Chapter 12 – Forecasters' Objectives and Strategies. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*, pages 690–720. Elsevier.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: Journal of the Econometric Society*, pages 315–335.

Nordhaus, W. D. (1987). Forecasting Efficiency: Concepts and Applications. *The Review of Economics and Statistics*, 69(4):667–674.

Ottaviani, M. and Sørensen, P. N. (2006). The Strategy of Professional Forecasting. *Journal of Financial Economics*, 81(2):441–466.

Rülke, J.-C., Silgoner, M., and Wörz, J. (2016). Herding behavior of business cycle forecasters. *International Journal of Forecasting*, 32(1):23–33.

Silver, N. (2012). *The Signal and the Noise: Why So Many Predictions Fail-but Some Don't*. Penguin Publishing Group.

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

Tanaka, M., Bloom, N., David, J. M., and Koga, M. (2018). Firm Performance and Macro Forecast Accuracy. Working Paper 24776, National Bureau of Economic Research.

Tversky, A. and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157):1124–1131.

Woodford, M. (2001). Imperfect Common Knowledge and the Effects of Monetary Policy. Technical report, National Bureau of Economic Research.

Zarnowitz, V. (1984). The Accuracy of Individual and Group Forecasts from Business Outlook Surveys. *Journal of Forecasting*, 3(1):11–26.

A Appendix

A.1 Tables

Table A1: Outcome Statistics

		GDP growth			**Inflation**		
Country	Obs.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Canada	14,872	7,457	2.2	1.4	7,415	1.8	0.6
France	18,600	9,331	1.3	1.1	9,269	1.3	0.8
Germany	27,481	13,739	1.4	1.7	13,742	1.4	0.6
Italy	14,336	7,228	0.5	1.6	7,108	1.9	1.1
Japan	18,450	9,409	1.0	2.1	9,041	0.2	0.9
Netherlands	9,586	4,959	1.6	1.8	4,627	1.8	0.8
Spain	13,927	7,011	2.0	2.0	6,916	2.2	1.4
Sweden	12,885	6,469	2.3	2.0	6,416	1.1	1.1
UK	23,663	12,539	1.8	1.6	11,124	2.8	0.9
USA	25,510	12,778	2.5	1.4	12,732	2.2	1.0
Total	179,310	90,920	1.7	1.8	88,390	1.7	1.2

Notes: Actual outcome data for annual average GDP growth and inflation between 1995 and 2018. Each outcome is repeated for each corresponding forecast observation in the monthly survey by Consensus Economics between 1995 and 2017.

Table A2: Error Statistics

		GDP growth			**Inflation**		
Country	Obs.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Canada	14,872	7,457	-0.2	1.1	7,415	-0.1	0.6
France	18,600	9,331	-0.3	0.8	9,269	-0.1	0.6
Germany	27,481	13,739	-0.2	1.3	13,742	-0.2	0.6
Italy	14,336	7,228	-0.6	1.1	7,108	-0.0	0.6
Japan	18,450	9,409	-0.2	1.8	9,041	-0.1	0.6
Netherlands	9,586	4,959	-0.1	1.3	4,627	-0.0	0.5
Spain	13,927	7,011	0.0	1.0	6,916	0.0	0.9
Sweden	12,885	6,469	-0.1	1.7	6,416	-0.4	0.9
UK	23,663	12,539	-0.2	1.1	11,124	0.1	0.8
USA	25,510	12,778	-0.0	1.0	12,732	-0.0	0.8
Total	179,310	90,920	-0.2	1.3	88,390	-0.1	0.7

Notes: Forecasts from the monthly survey by Consensus Economics between 1995 and 2017. Data refer to ex-post forecast errors where errors are defined as outcome minus forecast. In each month, forecasters are asked to forecast the current and next year's annual average growth rate. This results in 24 potential forecast origins (and horizons) for each year's outcome.

A.2 Robustness – Alternative Specifications

As a robustness check, and as a comparison to the literature, I estimate, using individual forecast data

$$R_{ith} = \delta_0 + \delta_1 R_{ith+1} + \varepsilon_{ith} \tag{A1}$$

$$x_t - F_{ith} = \delta_2 + \delta_3 R_{ith} + \varepsilon_{ith}, \tag{A2}$$

where $R_{ith} = F_{ith} - F_{ith+1}$ is the forecast revision. We expect to observe zero estimates on all δ coefficients. Positive estimates on δ_1 or δ_3 tell us that the forecasters underreact to new information. Positive δ_1 implies that an upward revision is followed by an additional upward revision. Hence, the initial revision was not sufficiently large. A positive δ_3 tells us that the upward revision is insufficient since it predicts a positive forecast error and hence an underreaction. I also perform these regressions using the average forecast (consensus mean) to see how the average forecast responds to new information,

$$\bar{R}_{th} = \delta_0 + \delta_1 \bar{R}_{th+1} + \varepsilon_{th} \tag{A3}$$

$$x_t - \bar{F}_{th} = \delta_2 + \delta_3 \bar{R}_{th} + \varepsilon_{th}, \tag{A4}$$

where the average forecast is the mean of all forecasts: $\bar{F}_{th} = \frac{1}{n}\sum_{i=1}^{n} F_{ith}$. The average revision is then $\bar{R}_{th} = \bar{F}_{th} - \bar{F}_{th+1}$. The interpretation of the δ coefficients stays the same but now refers to the response of the average forecast instead of the average response of forecasters. Table A3 reports the estimation of equations (A1), (A2), (A3) and (A4). The first column presents the estimates from equation (A1), where we observe a positive estimate of 0.079. This means that an upward revision in the last period predicts an upward revision also in the current period. The second column presents the results from estimating equation (A2). We again observe a positive estimate, a coefficient of 0.358. The upward revision today is insufficient since it predicts a positive forecast error, and hence we interpret this as underreaction by the forecasters. In columns (4) and (5), we see that these results also hold for the average forecast. Hence, forecasters underreact to new information. These results are in line with Coibion and Gorodnichenko (2015) and Dovern et al. (2015) but go in the opposite direction to some of the findings in Broer and Kohlhas (2018) and Bordalo et al. (2018).

Table A3: Reaction to New Information

	Individual Forecast		Average Forecast	
	(1)	(2)	(3)	(4)
Dep. Variable:	Forecast Revision	Forecast Error	Forecast Revision	Forecast Error
Forecast Revision		0.358***		1.605***
		(0.078)		(0.277)
Lagged Forecast Revision	0.079***		0.502***	
	(0.018)		(0.038)	
Observations	158,181	179,310	10,046	10,536
R^2	0.006	0.007	0.249	0.055

Notes: Standard errors robust to two-way clustering at the forecaster (firm) level and the survey month level for columns (1) and (2). Columns (3) and (4) have standard errors robust to two-way clustering at country level and survey month level. Estimated equations are (A1), (A2), (A3) and (A4). *,**,*** represent the 10%, 5%, 1% significance levels.