Economic Studies 193

André Reslow

Electoral Incentives and Information Content in Macroeconomic Forecasts

Department of Economics, Uppsala University

Visiting address: Kyrkogårdsgatan 10, Uppsala, Sweden
Postal address: Box 513, SE-751 20 Uppsala, Sweden
Telephone: +46 18 471 00 00
Telefax: +46 18 471 14 78
Internet: http://www.nek.uu.se/

ECONOMICS AT UPPSALA UNIVERSITY

The Department of Economics at Uppsala University has a long history. The first chair in Economics in the Nordic countries was instituted at Uppsala University in 1741.

The main focus of research at the department has varied over the years but has typically been oriented towards policy-relevant applied economics, including both theoretical and empirical studies. The currently most active areas of research can be grouped into six categories:

* Labour economics
* Public economics
* Macroeconomics
* Microeconometrics
* Environmental economics
* Housing and urban economics

Additional information about research in progress and published reports is given in our project catalogue. The catalogue can be ordered directly from the Department of Economics.

André Reslow

Electoral Incentives and Information Content in Macroeconomic Forecasts



UPPSALA
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in Hörsal 2, Ekonomikum, Kyrkogårdsgatan 10, Uppsala, Friday, 19 March 2021 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Gisle James Natvik (BI Norwegian Business School).

Abstract

Reslow, A. 2021. Electoral Incentives and Information Content in Macroeconomic Forecasts. *Economic studies* 193. 184 pp. Uppsala: Department of Economics, Uppsala University.
ISBN 978-91-506-2862-3.

Essay I (with Davide Cipullo): This essay introduces macroeconomic forecasters as new political agents and suggests that they use their forecasts to influence voting outcomes. The essay develops a probabilistic voting model in which voters do not have complete information about the future economy and rely on professional forecasters when forming beliefs. The model predicts that optimal forecasters with economic interests (stakes) and influence publish biased forecasts before a referendum. The theory is tested using data surrounding the Brexit referendum. The results show that forecasters with stakes and influence released more pessimistic and incorrect estimates for GDP growth subject to the leave outcome than other forecasters.

Essay II (with Davide Cipullo): This essay documents the existence of Political Forecast Cycles. A theoretical model of political selection shows that governments release overly optimistic GDP growth forecasts ahead of elections to increase the reelection probability. The theory is tested using forecast data from the United States, the United Kingdom, and Sweden. The results confirm key model predictions and show that governments overestimate short-term GDP growth by 10 to 13 percent during campaign periods. Moreover, the bias is larger when the incumbent is not term-limited or constrained by a parliament led by the opposition. Furthermore, election timing determines the size of the bias at different forecast horizons.

Essay III: This essay assesses to what extent forecasters use competitors' forecasts efficiently. Empirical results using a large panel of forecasters suggest that forecasters underuse information from their competitors when forecasting GDP growth and inflation. The results also show that forecasters pay more attention to competitors when releasing short-term forecasts than medium-term forecasts. A belief updating model with noisy and private information supports the underuse interpretation and predicts that it is optimal to pay sizable attention to competitors' work. Furthermore, the essay shows that a revision cost model can only match the observed behavior if asymmetric horizon discounting between cost from revisions and loss from forecast errors is assumed.

Essay IV (with Michael K. Andersson and Ted Aranki): This essay proposes a method to account for differences in release dates when assessing an unbalanced panel of forecasters. Cross-institutional forecast evaluations may be severely distorted because forecasts are made at different points in time and thus with different amounts of information. The proposed method computes the timing effect and the forecaster's ability (performance) simultaneously. Simulations demonstrate that evaluations that do not adjust for the differences in information may be misleading. The method is also applied to a real-world data set of 10 Swedish forecasters, and the results show that the forecasters' ability ranking is affected by the proposed adjustment.

Keywords: Brexit, Electoral Cycles, Forecast Evaluation, Forecasting, Imperfect Information, Interest Groups, Political Selection, Voting

André Reslow, Department of Economics, Box 513, Uppsala University, SE-75120 Uppsala, Sweden.

© André Reslow 2021

ISSN 0283-7668
ISBN 978-91-506-2862-3
urn:nbn:se:uu:diva-433165 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-433165)

Dedicated to all users of forecasts

Acknowledgments

There are many people to whom I owe gratitude. First and foremost I am extremely grateful to my supervisors, Mikael Carlsson and Jesper Lindé, for their invaluable advice and continuous support during my Ph.D. studies. Your knowledge and guidance have greatly contributed to my academic research and the completion of this thesis.

My gratitude further extends to all members of the Department of Economics at Uppsala University. You are all part of making the department a fantastic research environment and your feedback have greatly improved this thesis. I would especially like to thank Mikael Elinder, Eva Mörk, Luca Repetto and Erik Öberg for our many talks and your extensive feedback. Furthermore, I wish to acknowledge the help provided by the administrative staff at the department throughout my studies.

To my cohort: Daniel Bougt, Davide Cipullo, Sofia Hernnäs, Sebastian Jävervall, Lillit Ottosson, Anna Thoresson, and our SLU friends Julian Lozano and Jonathan Stråle. The friendship and immense support we have for each other are fantastic, and you have all made these years truly amazing. I would like to give a special acknowledgment to Davide with whom I have co-authored two essays. I greatly enjoy working with you. Our long sessions in front of the whiteboard have been inspiring and essential in fostering my passion for the research process. A special thanks also to Sebastian for being a fantastic friend, office-mate, and commuter-buddy.

I am also indebted to several people outside of Uppsala University. I would like to thank Michael K. Andersson and Ted Aranki with whom I co-authored one essay. Thanks to all my colleagues at Sveriges Riksbank for your feedback and support. I am also truly grateful for the continued support throughout all of my studies from Sirus Dehdari with whom I did both my bachelor's and master's studies. Many thanks to Janne Tukianien for your valuable feedback at my final seminar. I would also like to express immense gratitude to David Cesarini for inviting me to spend the 2018/19 academic year at New York University. I greatly enjoyed our many talks and my year at NYU was truly inspiring. Financial support from Handelsbankens Forskningsstiftelser is gratefully acknowledged. Finally, it goes without saying, this thesis would not have been completed without the encouragement and support from family and friends.

Uppsala, January 2021
André Reslow

Contents

Introduction 9
The Essays 11
A Primer on Forecast data 16
Ten Commandments 17
References 19

I Biased Forecasts to Affect Voting Decisions? The Brexit Case 21
1 Introduction 22
2 Theoretical Framework 25
3 Taking the Model to Data 35
4 Results 44
5 Concluding Remarks 54
References 56
Appendix 59

II Electoral Cycles in Macroeconomic Forecasts 87
1 Introduction 88
2 Theoretical Framework 93
3 Institutional Background and Data 102
4 Empirical Strategy 105
5 Results 108
6 Robustness Checks 115
7 Concluding Remarks 117
References 119
Appendix 124

III Inefficient Use of Competitors' Forecasts? 137
1 Introduction 138
2 Data 141
3 Empirical Strategy 143
4 Results 146
5 Heterogeneity in Actual Behavior? 148
6 Baseline Model to Understand Optimal Behavior 151
7 Extending the Model to Understand Actual Behavior 154
8 Concluding Remarks 158
References 159
Appendix 162

IV Adjusting for Information Content when Comparing Forecast Performance 165
1 Introduction 166
2 Empirical Framework 167
3 Monte Carlo Simulation 171
4 Empirical Application 177
5 Conclusions 182
References 183

Introduction

"Difficult to see. Always in motion is the future."
— Yoda, *The Empire Strikes Back*

The future is uncertain. When we think about buying a home, changing job, voting for a politician, or applying to college, our decision is based on our expectations about the future. Such expectations are forecasts. We can often do the forecasting ourselves. Before leaving our homes, we form expectations about the weather to decide whether we should bring an umbrella. Based on our knowledge of the season and a glance through the window, we can roughly assess the probability of rainfall. Other times we need help from experts. When planning an outdoor party, days in advance, we might tune in for the news' weather segment for a multi-day forecast from the TV meteorologist. Likewise, when deciding whether to subscribe to a mortgage with fixed or variable interest rates, we might read the recommendation from the newspaper experts and consult our local bank office.

Forecasting is important. Firms, households, and policymakers constantly produce forecasts. For example, firms routinely forecast sales to guide their strategic planning; forecasting plays a key role in marketing decisions; businesses, households, and governments use many different forecasts in their budgeting process. All of us continuously form expectations about the future. Hence, we are all forecasters. In fact, forecasting is so important that some have made it their career and livelihood. We call them professional forecasters. We turn to these professionals for expert guidance when making important decisions. The question is: can we trust them?

In the episode "The Weatherman," of the TV-series "Curb Your Enthusiasm," the main character, Larry David, accuses the weatherman of having alternative motives. In the episode, Larry's friend cancels their next day round of golf due to a rain forecast. When Larry wakes up the following day, he notices it is sunny and a clear blue sky. Larry then starts to outline his hypothesis: "Predicts rain. Doesn't rain. It is not the first time this has happened." Larry then questions whether the weatherman merely forecasts rain to keep everyone off the golf course so that he can have it for himself. To investigate his theory, Larry heads down to the golf course, and lo and behold, he finds the weatherman on the ninth hole on an otherwise wide-open course. In fact, there is something to Larry's observation: "Predicts rain. Doesn't rain. It is not the first time this has happened." In weather forecasting, *wet bias* is an established term. The term refers to the often observed phenomenon that weather forecasts

consistently predict a higher probability of rainfall than occurring. Despite my sympathy for Larry's theory, golf incentives are unlikely to be the cause of the wet bias. However, one potential explanation is, according to Silver (2012), alternative objectives. The TV meteorologists might expect a greater backlash from the viewers if it rains when it is not supposed to, compared to predicting rain when it turns out to be a nice day. Larry's reaction does, however, suggest the opposite.

Also economic forecasts can be biased. For example, Charemza and Ladley (2016) show that central banks' inflation forecasts have a bias towards the inflation targets. Hence, central banks tend to predict a return to the target that is too rapid. A bias can be both intentional and unintentional. In the case of central banks' inflation forecasts, the bias can arise due to both. It could be that the central banks are overconfident in their ability to conduct monetary policy. In such a case, they will unintentionally introduce a bias towards their target. The bias can also arise intentionally from an attempt to create a self-fulfilling prophecy. Suppose firms and households believe the forecast of a rapid return to the inflation target. In that case, they may act in such a way that it becomes self-fulfilling to some degree.

As with anything else, there are good and bad forecasts. It is also the case that some things are easier than others to predict. While meteorologists struggle to predict rainfall days in advance, they can accurately predict the sunrise and sunset time. They predict the sunrise/sunset with accuracy within a minute, and they can do so years in advance. Economic forecasting is not straightforward. However, as with weather forecasting, it is generally easier to predict a realization in the next week than a realization years into the future. One challenge with economic outcomes is that they are a result of human decisions. Hence, an economic forecaster needs to understand and predict human behavior. This is undoubtedly challenging. In addition, economic outcomes are affected by nature. For example, a drought in South America lowers the crop yield, triggering a global rise in food prices. Predicting food prices is then easier in the short-term than in the long-term since we have a better understanding of the current crop yield than the crop yields years ahead.

The current ongoing COVID pandemic also demonstrates the challenge of doing economic forecasts. In January 2020, the consensus 2020 GDP growth among Swedish forecasters was 1.1%. By April 2020, this figure had been revised down to -5.6%. Some might think that the January estimates merely reflect that economists do not understand epidemiology. However, also the experts failed to see the coronavirus impact. On January 14, 2020, the World Health Organization tweeted: "Preliminary investigations conducted by the Chinese authorities have found no clear evidence of human-to-human transmission of the novel #coronavirus (2019-nCoV) identified in #Wuhan, #China." It is then fair to say that it is was challenging to predict the 2020 economic downturn.

Given that economic forecasting is challenging, most of us rely on professionals to do the job for us. The professionals can put more time and resources into their forecasts, and they are likely to outperform most of us even if also they struggle from time to time. The economic outlook is often presented in mass media and rests on the work of the professionals. Given that we form expectations and make important decisions based on forecasts from professionals, we must clearly understand these forecasters' objectives, motives, and performances. The aim of this thesis is to contribute to this understanding. Hence, the thesis is about the behavior and performance of macroeconomic forecasters. The thesis consists of four independent essays. The first two essays introduce political incentives in macroeconomic forecasts. The essays show that forecasters who prefer a particular outcome in an election or referendum will try to influence voters' beliefs by releasing biased forecasts. The other two essays study the efficiency and performance of forecasters with a focus on the role of information content. Specifically, how forecasters use information from competitors and how crucial it is to consider the timing and information when assessing forecast performance.

The Essays

Essay I: Biased Forecasts to Affect Voting Decisions? The Brexit Case
The first essay is co-authored with Davide Cipullo and studies the relationship between macroeconomic forecasters and voters in times of a referendum. In the essay, we introduce macroeconomic forecasters as new political agents in a probabilistic voting model, and we suggest that they use their forecasts to influence voting outcomes. We model a setting similar to the Brexit referendum, where voters have to decide between remaining in status quo and leaving for an alternative. We assume that voters have incomplete information when assessing the two alternatives. When forming expectations, voters update their beliefs relying on macroeconomic forecasts from professionals. Voters who care about the economy and have individual ideological preferences vote for the alternative that yields them the highest utility.

Macroeconomic forecasters are assumed to have heterogeneous economic interests in the outcome of the referendum. Some forecasters prefer one of the alternatives, and they do so to different degrees. A forecaster who prefers a specific outcome in the referendum will, due to the information asymmetry, have incentives to publish biased information in an attempt to swing the voters to the preferred outcome. One important aspect is that forecasters are not equally influential: some of them can influence voters more than others.

The decision problem of forecasters becomes rather complicated as they have a multidimensional objective. They are assumed to minimize forecast errors. They are also assumed to care about consistency over time. In addition, they have to balance the attempt to influence the referendum result. The de-

cision is further complicated by the presence of other forecasters, which generates a common pool problem. In other words, each forecaster must respond optimally to the best decision made by all other forecasters in the economy. It is also the case that the decision problem becomes highly nonlinear since the probability of victory for an outcome in the referendum is endogenous and dependent on the released forecasts.

The theoretical model predicts that it is optimal for forecasters with economic interest (stakes) and influence to publish biased forecasts prior to a referendum. We test this theory using high-frequency data at the forecaster level surrounding the Brexit referendum.

In the empirical analysis, the goal is to compare the forecasts for GDP growth published by forecasters with stakes and influence to those released by other institutions. To do this, we use three different measures of stakes: being a financial institution, being located in the City of London, and the stock market drops the days after the referendum results. In order to determine the influence, we use several Google-based measures. The empirical strategy then involves estimating a dynamic difference-in-differences model surrounding the Brexit referendum. The results show that forecasters with stakes and influence released much more pessimistic and incorrect estimates for GDP growth subject to the leave outcome than other forecasters.

The essay extends two strands of literature. First, the information bias of political agents and second, the strategic behavior and objective function of macroeconomic forecasters. The results are of high importance. The propaganda bias reduces the welfare of voters, who, in equilibrium, may not cast a vote for the preferred choice. The results also have broader implications, as the paper highlights that we need to consider electoral incentives in macroeconomic forecasts more generally.

Essay II: Electoral Cycles in Macroeconomic Forecasts

The second essay builds on the implications from the first essay. Also coauthored with Davide Cipullo, the second essay documents the existence of *Political Forecast Cycles*. There is a large literature on Political Business Cycles and Political Budget Cycles. The early literature by Nordhaus (1975) on political business cycles provided models where politicians exploit the so-called Phillips curve by inflating the economy during election years to reduce unemployment, to gain an electoral advantage. Rogoff and Sibert (1988), Rogoff (1990), and others, expanded the literature to variables such as taxes, government spending, and deficits, resulting in the vast literature on political budget cycles. This essay extends the concept to macroeconomic forecasts. In the essay, we ask whether governments release overly optimistic GDP growth forecasts ahead of elections to gain an electoral advantage. The essay provides a theoretical rationale and empirical evidence of such behavior.

We consider a model of political selection consisting of voters and one incumbent who faces a random opponent during an election. In the model, the

ability of politicians impacts fiscal policy that contributes to GDP growth. We task the government (the incumbent politician) with supplying forecasts about the economy to the public. We assume that voters observe these forecasts via mass media and use these observations to form beliefs about the politician's ability. However, given that the government is a prime source for forecasts about the economy, there will be an incentive for the incumbent government to manipulate the forecasts to influence the voters. The key trade-off is in the hands of the incumbent, who needs to balance reelection incentives and the reputation cost of lying.

We use forecaster-level panel data from the United States, the United Kingdom, and Sweden to empirically test key model predictions. We employ a difference-in-differences approach for each country that allows us to compare government with non-government forecast errors in and out of campaign periods. This allows us to identify whether the government behaves differently during campaign periods than other forecasters compared to itself in non-campaign periods. The empirical results confirm key model predictions. Specifically, we document that governments overestimate short-term GDP growth by 10 to 13 percent during campaign periods. Furthermore, we find that the bias is larger when the incumbent government is not term-limited or constrained by a parliament led by the opposition. Consistent with the model, we also find that the election timing and amount of available information determine the size of the bias at different forecast horizons.

The theoretical model highlights an important equilibrium condition. If voters are rational and discount the bias, the incumbent does not gain any advantage in equilibrium. Yet, the incumbent faces a utility (reputation) loss due to her estimates' low accuracy. If voters instead are naive, then the bias will be effective and foster the reelection probability, resulting in an incumbency advantage. Hence, the model clarifies the rationale behind an election cycle bias both in the case of naive and in the case of sophisticated voters. The bias arises from lack of commitment if voters are rational and from manipulation of voters' beliefs if they are naive and do not expect the incumbent to be biased. From the politician's view, the problem is similar to a prisoner's dilemma situation, while the voters solve a problem similar to a matching pennies game.

In addition to a loss in voter welfare, biased forecasts could also be damaging to the economy in a broader perspective. Coibion et al. (2018) show that firms care about macroeconomic variables such as GDP, inflation, and unemployment. They also show that firms update their beliefs (forecasts) when presented with forecasts from professionals. Furthermore, Tanaka et al. (2019) show that firms' GDP forecasts are associated with their investment and employment choices. Hence, biased forecasts could result in biased and inefficient firm and household choices if they fail to account for the bias.

Essay III: Inefficient Use of Competitors' Forecasts?
The third essay assesses the extent to which macroeconomic forecasters make efficient use of forecasts from their competitors. The premise of the essay is a world with incomplete and private information. In such a world, an important source of information for a macroeconomic forecaster is the forecasts of the competitors. It is well established in the forecasting literature that it is optimal to pool information from many forecasters (see, e.g., Zarnowitz, 1984 and Hendry and Clements, 2004). The typical empirical finding is that consensus (average) forecasts are more accurate over time than individual forecasts. This essay evaluates whether forecasters adopt this well-established result in their work or not.

The essay takes an empirical starting point using a large panel of forecasters. The results suggest that forecasters do not make efficient use of information from their competitors. More specifically, forecasters underuse the available estimates from their competitors in their estimates for current and next years' GDP growth and inflation. The essay then studies several heterogeneities and finds that forecasters appear to behave differently at different horizons. When they are far away from the realization of the outcome, the attention to competitors is at its minimum. As the realization of the outcome emerges (i.e., the forecast horizons decreases), forecasters start to increase the attention to others and start to behave more optimally.

In order to understand these empirical findings, a theoretical model of belief updating is adopted. A baseline model with noisy and private information supports the underuse interpretation and predicts that it is optimal to pay considerable attention to competitors' work. However, these interpretations are based on the baseline assumption that forecasters only try to minimize forecast errors. Two different extensions to the baseline model are made to consider alternatives to pure inefficiency.

First, the findings of this essay arguably do not reflect pure cognitive bias. In particular, overconfidence—which leads to the overweighting of private information—is unlikely to be the sole explanation for the observed behavior. A back-of-the-envelope calculation shows that for overconfidence to explain the observed behavior, forecasters need to think that the precision in the private signal that they receive is 100 times more precise than the private signal received by any competitor. This is deemed to be highly unlikely.

Second, the model is extended so that forecasters have to pay a cost when revising forecasts. This cost captures a preference for consistency. The extension introduces a trade-off between minimizing the forecast errors and minimizing the revisions. However, the essay shows that the revision cost model can only match the observed horizon heterogeneity in the data if it is assumed that the cost from revisions and the cost from forecast errors have asymmetric horizon discounting.

These results highlight that the objective function of forecasters seems to differ from the often assumed objective of pure forecast error minimization.

Specifically, that the forecast horizon matters for the strategic behavior and efficient use of competitors' information. In order to optimally use and evaluate forecasts, users of forecasts need to understand the objective and nature of the forecasters. The results of this essay contribute to this understanding and argue that forecasters are either inefficient or minimizing more than forecast errors.

Essay IV: Adjusting for Information Content when Comparing Forecast Performance

The fourth essay is co-authored with Michael K. Andersson and Ted Aranki, and is published in the *Journal of Forecasting*.[1] The essay studies the role of timing and information when assessing forecast performance.

Since many important economic and political decisions are based on economic forecasts it is crucial that these predictions are as accurate as possible. Therefore, evaluations are regularly undertaken to assess the performance of forecasts and the forecasters. This essay points out that cross-institutional forecast evaluations may be severely distorted by the fact that forecasts are made at different points in time, and thus with different amounts of information. In the essay, we propose a statistical method to account for these differences when analyzing an unbalanced panel of forecasters.

The situation that we have in mind is the following. Suppose we have a set of forecasters that are forecasting each years GDP growth. In addition, suppose that forecasters are releasing their forecasts during different months during the year. Comparing the performance of the forecasts released in, for example, May to the ones released in August would be misleading since the August forecast had access to a couple of months of more information.

The aim of this essay is to compute and control for differences in information content that arise due to timing differences. The method estimates a timing effect and the forecaster's ability simultaneously. We perform a Monte Carlo simulation to demonstrate that evaluations that do not adjust for the differences in information content may be misleading. We find the simulation helpful since it allows us to test the method in a controlled environment. However, the method is also applied to a real-world data set of 10 Swedish forecasters for the period 1999–2015. The panel data is unbalanced and the results show that the performance ranking of the forecasters is affected by the proposed adjustment. Consequently, when adjusting for timing and the information content we change our assessment of the relative ability of the forecasters. Hence, the proposed method in this essay extends the toolbox available to evaluate forecast performance.

[1] Published as Andersson, M. K., Aranki, T., and Reslow, A. (2017), Adjusting for Information Content when Comparing Forecast Performance, *Journal of Forecasting*, 36(7):784–794.

A Primer on Forecast data

In the context of this thesis, a forecaster should in general not be seen as an individual person. You should rather think of a forecaster as an institution or firm. Private banks, research institutions, consultant agencies, government agencies, and labour unions, are some examples of forecasters occurring in the data used in this thesis. We call them *macroeconomic* forecasters since they provide forecasts for macro variables such as GDP growth and inflation.

All the forecast data used in this thesis follow the same structure. The forecasts refer to *current* and *next* years' annual growth rates. For example, if we observe the release of a forecast in April 2021, that forecast would contain an estimate for the 2021 (current) year's GDP growth rate, and an estimate for the 2022 (next) year's GDP growth rate. Hence, at each forecast release we observe forecasts with different horizons (the current and next year). During a year, we can observe multiple forecasts released from a forecaster. Hence, we observe for each targeted outcome multiple forecast at different horizons.

Figure 1 illustrates the structure of the data. In the left panel we see that a forecast released during the second quarter of 2021 targets two different outcomes: the 2021 and the 2022 years' annual growth rates. Suppose now a forecaster that releases a new forecast each quarter. In that case, there exists 8 different forecasts for each year's outcome. This is illustrated in the right panel of Figure 1. Note that we would expect the forecast for the 2022 outcome released in Q4 of 2022 to be more accurate than the forecast released in Q1 of 2021. This expectation rests on the fact that a forecaster has more information at the shorter horizon. The horizon is generally thought of as the distance in time between the release of the forecast and the realization of the outcome.

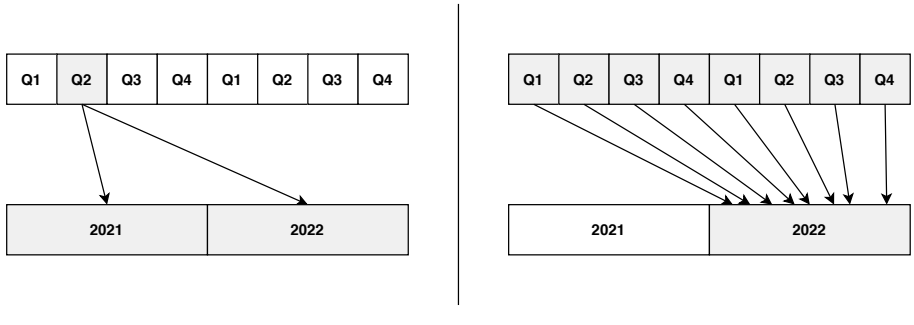


Figure 1. Structure of Data

Notes: The top-row boxes represent different forecast release dates, and the bottom-row boxes represent the different outcomes that is the target of the forecasts.

Ten Commandments

Phil Tetlock ends his book *Superforecasting* with Ten Commandments for aspiring superforecasters (Tetlock & Gardner, 2015). Tetlock's commandments provide a useful takeaway of the key themes in his book for how we all can become better forecasters. The main theme of my thesis is the behavior (or misbehavior) of macroeconomic forecasters. Hence, the main takeaways of this thesis can best be summarized in guidelines for the users of forecasts.

Ten Commandments for Users of Forecasts

(1) Use the consensus rather than a individual forecast
Is is generally the case that the consensus (average) forecast is better than the forecast from an individual forecaster. This is well established in the forecasting literature and further supported by the results in this thesis. Hence, users of forecasts are advised to use information from several forecasters.

(2) Be aware that forecasters have different abilities
Some forecasters are better than others. Many factors determine the quality of the forecasts. Some forecasters have more resources than others, and some have more or better information than others. Hence, when we use forecasts, we need to assess the quality of our source. For example, it is ex-ante likely that a Swedish bank makes a better forecast of the future house prices in Sweden than a foreign bank.

(3) Use a weighted average if possible
Given that forecasters have different abilities (as stated by the second commandment), you can improve on the first commandment. You can outperform the consensus (unweighted average) forecast by creating a weighted average based on the different abilities. However, this relies on you knowing (or accurately predicting) the true abilities of the forecasters.

(4) Remember that the forecast horizon matters
The forecast horizon (i.e., the distance between the release of the forecast and the realization of the outcome) matters. Forecasts with shorter horizons are generally better than forecasts with longer horizons. Remember that the horizon might impact the strategic behavior of the forecasters.

(5) Gather as much information as possible
It is usually the case that we make better decisions the more information that we have. Gathering more data ourselves is beneficial, as is gathering more forecasts. Having more and better information about the forecasters' objectives is beneficial, as is having a good understanding of their performance and ability.

(6) Do not be overconfident, but remember that most of us are
Overconfidence is "the mother of all biases." Do not overestimate your own ability. Seek information from others, but be aware that also your source of information might be overconfident. Even professional forecasters are likely to be overconfident and overestimate their forecast ability.

(7) Think of the underlying objectives
Think of the underlying objectives when you observe a forecast. Try to figure out if the forecasters' objectives are in line with your own. When thinking of their objectives, the natural starting point is to assume that they try to minimize forecast errors. However, there can be more to it. They might, for example, have consistency concerns (minimize revisions).

(8) Do not be naive
The forecaster might be lying! Do not be naive and think that the forecaster can not have alternative motives. The forecaster might be trying to influence your beliefs to their benefit. The weatherman might be manipulating your golf plans; the central bank might try to manipulate your inflation expectations; and the government might try to influence your voting decision.

(9) Be aware of the perils of commitment solutions
Try to find a way for the forecaster to not act on alternative motives, but be aware of the perils of different solutions. Outsourcing is a potential but not infallible commitment device. Consider, for example, the problem when politicians outsource the forecasting task to another agency. If the managers or workers of that agency are appointed, paid, and otherwise dependent on the politicians, they might act to please their superiors.

(10) Strike a balance between following commandments and living life
Following commandments might be exhausting. You need to balance the cost of following commandments with the benefit of forming better expectations. Some forecasts are more important than others. Your weather expectations for your normal workday is likely to be less important than those for your wedding day. Do the hard work and invest the resources where they matter the most and yield the highest return.

References

Charemza, W., & Ladley, D. (2016). Central banks' forecasts and their bias: Evidence, effects and explanation. *International Journal of Forecasting*, *32*(3), 804–817.

Coibion, O., Gorodnichenko, Y., & Kumar, S. (2018). How Do Firms Form Their Expectations? New Survey Evidence. *American Economic Review*, *108*(9), 2671–2713.

Hendry, D. F., & Clements, M. P. (2004). Pooling of forecasts. *The Econometrics Journal*, *7*(1), 1–31.

Nordhaus, W. D. (1975). The Political Business Cycle. *The Review of Economic Studies*, *42*(2), 169–190.

Rogoff, K. (1990). Equilibrium Political Budget Cycles. *American Economic Review*, *80*(1), 21–36.

Rogoff, K., & Sibert, A. (1988). Elections and Macroeconomic Policy Cycles. *The Review of Economic Studies*, *55*(1), 1–16.

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

Tanaka, M., Bloom, N., David, J. M., & Koga, M. (2019). Firm Performance and Macro Forecast Accuracy. *Journal of Monetary Economics*.

Tetlock, P., & Gardner, D. (2015). *Superforecasting: The art and science of prediction*. Random House.

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

Essay I

Biased Forecasts to Affect Voting Decisions? The Brexit Case

with Davide Cipullo

Acknowledgments: We would like to thank Eva Mörk for extensive feedback, and Alberto Alesina, Fabio Canova, Davide Cantoni, Mikael Carlsson, David Cesarini, Sylvain Chassang, Sirus Dehdari, Mikael Elinder, Christopher Flinn, Lucie Gadenne, Georg Graetz, Oliver Hart, Isaiah Hull, Ethan Kaplan, Andreas Kotsadam, Horacio Larreguy, Barton E. Lee, Jesper Lindé, Andreas Madestam, Jonathan Meer, Torsten Persson, Fabien Postel-Vinay, Luca Repetto, Martin Rotemberg, Petr Sedláček, Daniel Spiro, Francisco José Veiga, Karl Walentin, Ekaterina Zhuravskaya and the participants in the seminars held at Uppsala University, Uppsala Centre for Fiscal Studies (UCFS), Sveriges Riksbank, Harvard University, New York University, the 75^{th} Annual Congress of the International Institute of Public Finance, the 34^{th} Annual Congress of the European Economic Association, and the 2020 ASSA Meetings in San Diego for their dialogue and comments. Thanks to Amanda Kay from the HM Treasury for endowing us with data in digital format. Financial support from *Handelsbankens Forskningsstiftelser* is gratefully acknowledged.

1 Introduction

Several issues of great economic relevance have recently been addressed using referenda: the referendum held in the United Kingdom to leave the European Union, the referendum held in Greece on the agreements with the EU institutions to solve the debt crisis, the referendum held in Italy on a major change to the national Constitution, and the referendum held in Catalonia on the independence from Spain. Many of the debates leading up to those referenda focused on the potential effects on economic growth, using estimates published by professional macroeconomic forecasters. Economic forecasts can be easily communicated to and understood by voters, even if advanced competence, modeling, and equipment are required to produce a forecast. Voters can use the forecasts to obtain information about economic variables, such as GDP growth, before turning to the ballot.[1] In many public debates, economic forecasts are taken as a given, without considering that the institutions publishing the forecasts may be promoting their interests.

In this paper, we introduce macroeconomic forecasters as political agents and argue that they may exploit their information advantage to influence the voting process. Our approach combines a simple theoretical framework, which shows how forecasting institutions can profit from the asymmetry of information with respect to voters, and an empirical analysis, which uses a panel of forecasters surveyed every month before and after the Brexit referendum. Different forecasters face different incentives. First, a forecaster has incentives to favor one of the outcomes at the expense of the other if they have an economic interest to defend, and the referendum outcome threatens this interest. Second, a forecaster can have an impact on the outcome of the decision-making process only if she is influential enough. The model predicts, and the empirical results confirm that forecasters with stakes in and influence over the referendum decision released more pessimistic and more incorrect estimates of GDP growth rate than the other institutions.

We set up a probabilistic voting model in which voters do not have complete information about the potential states of the economy after a referendum and therefore rely on professional forecasters to form beliefs. In the model, the voters' decision rule is to support the outcome that yields them the highest utility (Lindbeck & Weibull, 1987). Still, voters' beliefs on the potential states of the economy depend on published forecasts rather than on the states themselves. Forecasters' economic interests (stakes) in the outcome are heterogeneous, and some can influence voters' beliefs more than others. Forecasters with stakes and influence face a trade-off between the accuracy of their forecast and the attempt to influence the referendum result. Accuracy is mea-

[1] The relationship between voters and macroeconomic forecasters can be understood in the light of Downs (1957). Rational agents lack incentives to invest in collecting costly information before voting because the probability of casting the decisive vote to swing the election outcome is negligible.

sured by the forecast error, whereas the information asymmetry provides the opportunity to influence the voters. In equilibrium, forecasters with stakes and influence release intentionally biased forecasts to make swing voters change their voting decision. The model predicts the presence of an extensive as well as an intensive margin of propaganda bias. Forecasters with positive stakes and influence are more likely to release incorrect forecasts than other forecasters, and the size of the propaganda bias is increasing in both parameters. After the vote, forecasters face a trade-off between accuracy and consistency of their estimates, so that the propaganda bias is observable in the subsequent releases after the vote.

We test our theory using high-frequency data at the forecaster level collected in connection with the EU membership referendum (also known as the Brexit referendum) held in 2016 in the United Kingdom. In the empirical analysis, we compare the forecasts for GDP growth published by forecasters with stakes and influence to those released by other institutions. We use three different measures of stakes: being a financial institution, being located in the City of London, and the stock market drops the days after the referendum results. In order to determine the influence of each forecaster, we use several Google Trends and Google News measures.

The Brexit referendum is ideal for testing our theory for at least two reasons. First, the economic consequences are difficult to predict for voters since no country had previously experienced a retreat from the EU. Second, several forecasters have economic interests threatened by Brexit. Third, we document that voters were actively searching for information about the economic effects of the referendum approaching the vote.

We document that forecasters with stakes and influence released short-run GDP growth rate estimates subject to the Leave outcome that were between 0.6 and 0.8 percentage points lower than the estimates released by other institutions. The actual outcome for GDP growth in 2017 shows that these forecasters were more incorrect than other institutions and that the propaganda bias explains up to 50 percent of the short-term forecast errors. We also find that the difference between the groups of forecasters comes primarily from pessimistic forecasts on investments and trade exposure. In addition, we test the implications of our model at the intensive margin. The results confirm the theoretical prediction of increasingly more pessimistic forecasts when either stakes or influence increases. The empirical setup at hand does not allow to estimate the impact of the propaganda bias on the voters' decision, which would require the comparison with a counterfactual world in which forecasters release unbiased estimates. However, daily-level survey data from the *British Election Study* show that the release of pessimistic pre-referendum reports affected voters' beliefs about the issues at stake, and increased the interest in the referendum and the willingness to vote.

The propaganda bias is estimated in proximity to the referendum, while forecasts released by different institutions converge within a few months after

the vote. The convergence is consistent with a two-fold interpretation. First, after the result was realized, there was still scope for forecasters to influence the implementation of a *hard* or *soft* Brexit. Second, they might have decided to adjust their forecasts slowly to preserve their credibility, in line with the trade-off between accuracy and consistency after the vote postulated in the theoretical model.

Mass media are a key channel through which voters receive relevant information, including forecasts. However, we address and rule out the possibility that our results are driven by voters' selective exposure to information consistent with their prior beliefs and by forecasters' biased beliefs on the economic effects of the referendum on the economy. First, we exploit the variation across newspapers in the number of estimates published by each forecaster in the sample. The results exclude that forecasters whose forecasts are more likely to be read by a pro-Remain or progressive audience are the drivers of our findings. Second, we compare our main results with comparable periods after the vote and with other major shocks to the UK economy to assess that forecasters with stakes and influence do not tend to underestimate GDP growth in connection with negative economic shocks.

This paper extends two strands of literature. First, earlier literature has shown that special interest groups (see, e.g., Baron, 1994; Besley & Coate, 2001; Grossman & Helpman, 1996) and media (see, e.g., DellaVigna et al., 2014; Enikolopov et al., 2011; Qin et al., 2018) are active players in the political economy and may release biased pieces of information in order to affect individuals' beliefs and, in turn, voting behavior (Durante et al., 2019; Martin & Yurukoglu, 2017). Our theoretical model and empirical results suggest that also macroeconomic forecasters exploit their information oligopoly to influence voters' beliefs. Second, the strategic literature on the behavior of forecasters, Laster et al. (1999) develop a theoretical model in which forecasters' payoffs are based on two criteria: their accuracy and their ability to generate publicity. There is a trade-off between the two as efforts to attract publicity compromise accuracy (see also Croushore, 1997; Marinovic et al., 2013; Ottaviani & Sørensen, 2006). Our theoretical model proposes an alternative trade-off and shows that the strategic behavior of macroeconomic forecasters can also be generated by a propaganda bias coming from the attempt to influence voters.

The propaganda bias may reduce the welfare of voters, who, in equilibrium, may not necessarily cast a vote for the preferred choice compared to a world with unbiased forecasters. Naive voters are predicted to make systematic voting errors in line with the outcome preferred by macroeconomic forecasters, while sophisticated voters make the correct choice in expectations but are incorrect for particular realizations of stakes and influence. If voters are rational, the propaganda bias generates an Pareto inefficient equilibrium in this information market since forecasters in expectations pay an accuracy cost without systematically influencing the referendum result.

2 Theoretical Framework

We consider two types of agents: voters and forecasters. Voters have to choose between two alternatives, $A \in \{L,R\}$, each of which is associated with an economic outcome y^A. L represents the decision to leave the status quo, and R the decision to remain. We start by presenting a simplified version of our theoretical model, in which we assume one forecaster and that both voters and the forecaster correctly observe y^R while they do not have complete information on y^L. While the forecaster receives unbiased, noisy information about y^L, voters have to rely on a noisy media report on the released forecast. This framework represents a standard model of asymmetric information: voters are prospective and care about the economy in the future but need professional forecasters to form beliefs before voting. Forecasters have the opportunity to exploit the information asymmetry by strategically releasing biased information if they have an economic interest to defend that is at stake during the referendum.

This simple version is helpful to follow the intuition behind the model, familiarize with the notation, and get an intuitive solution. We later extend the model in several dimensions and solve it with the help of numerical methods: first, we assume that voters and forecasters do not have perfect information on y^R either; second, we consider several forecasters, that are heterogeneous in stakes and influence; third, we extend the timing of the model allowing forecasters to release multiple updates before and after the vote to follow the dynamic evolution of the propaganda bias; fourth, we extend the objective function of macroeconomic forecasters. In the basic model, the forecaster trades-off accuracy with the attempt of influencing the referendum results. In the extended version, forecasters also care about the consistency over time in the forecasts that they release.

The model presented in this paper considers a referendum in which voters decide whether to remain in a status quo or leave for an alternative. However, an analogous intuition would apply to a setting of general elections in which the incumbent politician faces a random opponent.

2.1 Basic Model

Voters

Consider a continuum of voters with total mass 1, with linear preferences over policy outcomes represented by $W(y) = y$. Following the well-established probabilistic voting model (Lindbeck & Weibull, 1987), we assume that voters make their decision based on the state of the economy and their ideological preferences. Individual i prefers alternative L over alternative R if and only if

$$y^L \geq y^R + \sigma_i, \tag{1}$$

where the ideology parameter σ_i captures all preferences at the individual level in support of R that are orthogonal to $W(\cdot)$. We assume that σ_i is uniformly distributed over the interval $\left[-\frac{1}{2\phi}, \frac{1}{2\phi}\right]$ with density $\phi > 0$.

We assume that the agents in the economy do not observe y^L, although they know that $y^L \sim \mathcal{N}\left(\mu^L, \tau_L^{-1}\right)$. The forecaster receives a noisy, private signal $\widetilde{y}^L = y^L + \varepsilon$, where $\varepsilon \sim \mathcal{N}\left(0, \tau_\varepsilon^{-1}\right)$ is a random error, and releases a forecast $F^L = \mathbb{E}\left(y^L|\widetilde{y}^L\right) + b$, where b is an additive bias that the forecasters can attached to its estimates released to the public.[2] Voters observe, via mass media, a noisy measure of F^L from the forecaster:

$$\widetilde{F}^L = F^L + e \tag{2}$$

where $e \sim \mathcal{N}\left(0, \tau_e^{-1}\right)$ is a media transition error capturing the friction between the information released by the forecaster and the signal received by voters. Therefore, voters' decision rule in (1) can be expressed as

$$\mathbb{E}\left(y^L|\widetilde{F}^L\right) \geq y^R + \sigma_i, \tag{3}$$

where $\mathbb{E}\left(y^L|\widetilde{F}^L\right)$ is the posterior belief about y^L consistent with Bayesian updating. Formally,

$$\mathbb{E}\left(y^L|\widetilde{F}^L\right) = (1-\gamma)\mu^L + \gamma\left[\widetilde{F}^L - \mathbb{E}\left(b|\widetilde{F}^L\right)\right], \tag{4}$$

where $\mathbb{E}\left(b|\widetilde{F}^L\right)$ is the voters' expectations on the propaganda bias attached by the forecaster. The parameter $\gamma = \frac{\omega\tau_L^{-1}}{\omega^2(\tau_L^{-1}+\tau_\varepsilon^{-1})+\tau_e^{-1}}$, captures the relative influence of the forecaster on the voters' posterior belief about y^L. Hence, the forecasters influence is a function of the precision in the prior, τ_L, the precision in the signal, τ_ε, and the precision in the media report, τ_e.[3]

Voters that are indifferent between the two alternatives are denoted swing voters. According to (3), they are defined by the relationship

$$\widetilde{\sigma} = \mathbb{E}\left(y^L|\widetilde{F}^L\right) - y^R.$$

By ranking voters according to their ideological parameter, all individuals with $\sigma_i \in \left[-\frac{1}{2\phi}, \widetilde{\sigma}\right]$ will then vote in favor of alternative L. We define π^L to be the share of votes in society in support of L, and $p^L = P\left(\pi^L > \frac{1}{2}\right)$ is, by extension, the probability that L wins in a binary competition. The share of votes that L

[2] The forecaster's expectations about y^L is $\mathbb{E}(y^L|\widetilde{y}^L) = (1-\omega)\mu^L + \omega A$, where $\omega = \frac{\tau_\varepsilon}{\tau_L+\tau_\varepsilon}$.

[3] $\omega = \frac{\tau_\varepsilon}{\tau_L+\tau_\varepsilon}$ comes from the forecaster's Bayesian updating over y^L.

receives in the population is

$$\pi^L = \int_{-\frac{1}{2\phi}}^{\widetilde{\sigma}} \phi \, di = \phi\Big(\widetilde{\sigma} + \frac{1}{2\phi}\Big)$$
$$= \frac{1}{2} + \phi\Big\{\gamma\Big[F^L + \varepsilon - \mathbb{E}\Big(b|\widetilde{F}^L\Big)\Big] + \Big(1-\gamma\Big)\mu^L - y^R\Big\},$$

while the probability that L wins is given by

$$p^L = P\Big(\gamma\varepsilon - y^R + \Big(1-\gamma\Big)\mu^L > \gamma\Big[\mathbb{E}\Big(b|\widetilde{F}^L\Big) - F^L\Big]\Big)$$
$$= 1 - G\Big(\gamma\Big[\mathbb{E}\Big(b|\widetilde{F}^L\Big) - F^L\Big]\Big), \tag{5}$$

where $G\big(\cdot\big)$ is the cumulative distribution function of the gaussian random variable $\gamma\varepsilon - y^R + (1-\gamma)\mu^L \sim \mathscr{N}\big(-y^R + (1-\gamma)\mu^L, \gamma^2\tau_e^{-1}\big)$. Therefore, the probability that L wins depends on F^L instead of the true y^L, providing the incentive for the forecaster to influence the voting outcome if she has economic interests at stake during the vote.

Forecaster

Assume a forecaster can either face an economic loss under L or be indifferent between the two alternatives. The forecaster minimizes the following loss function with respect to F^L, given μ^L, the observed signal $\widetilde{y}^L$, and voters' strategies:

$$\min_{F^L} \; p^L\Big\{\eta C + \frac{1+\gamma}{2}\Big[F^L - \mathbb{E}\Big(y^L|\widetilde{y}^L\Big)\Big]^2\Big\}, \tag{6}$$

where $F^L \in \big[\underline{F^L}, \overline{F^L}\big]$ is the forecast subject to alternative L, $C > 0$ represents a cost associated with alternative L, p^L is the probability of leaving the status quo and the parameter $\eta \geq 0$ captures the stakes of the forecaster on the referendum outcome.[4] We model the loss function of forecasters in the spirit of Laster et al. (1999), modifying their trade-off between accuracy and publicity into a trade-off between the accuracy of the released estimate and the will of favoring the preferred outcome in the referendum. If the forecaster faces a loss in the event that L wins, then it has a direct economic interest in the referendum result and hence has stakes, while if the forecaster is indifferent between the two alternatives, then it has no stakes in the referendum result, and only aims to minimize the forecast error.

[4]In the occasion of a referendum, forecasters, may either have stakes in support of one of the alternatives or being indifferent between the two alternatives. We assume for simplicity that $\eta \geq 0$ so that forecasters can either have preferences in support of R or being indifferent between the two alternatives. The sign of η determines the sign of the propaganda bias, but not its presence.

This simple loss function, which entails a linear incentive for bias and a quadratic incentive for accuracy, shows that the problem at hand is not as straightforward as it seems. First, the forecaster can influence both the probability of paying the cost attached to L, ηC, and the probability of paying the accuracy cost, $\frac{1+\gamma}{2}\left[F^L - \mathbb{E}\left(y^L|\widetilde{y}^L\right)\right]^2$, which is endogenous to the referendum outcome. Second, more influence increases the marginal impact of F^L on p^L, but also the cost that the forecaster pays if she releases incorrect estimates.

Assuming an interior solution, the first-order condition with respect to F^L takes the form

$$\begin{aligned}&\gamma g(\cdot)\left[\eta C + \frac{1+\gamma}{2}\left[F^L - \mathbb{E}\left(y^L|\widetilde{y}^L\right)\right]^2\right] \\ &+ \left(1+\gamma\right)\left[F^L - \mathbb{E}\left(y^L|\widetilde{y}^L\right)\right]\left[1 - G(\cdot)\right] = 0,\end{aligned} \tag{7}$$

where $g(\cdot)$ is the probability density function of the random variable having the cumulative distribution function equal to $G(\cdot)$. The model is closed by taking voters' expectation over (7) to derive the expected propaganda bias $\mathbb{E}\left(b|\widetilde{F}^L\right)$. From (7), we derive the following proposition.

Proposition 1. ***Existence of political equilibrium with unbiased forecasts*** *Under the assumptions of the model, $F^L = \mathbb{E}\left(y^L|\widetilde{y}^L\right)$ is part of a political equilibrium if and only if $\eta = 0$ or $\gamma = 0$.*

Proposition 1 predicts that the forecaster releases its best estimate for y^L approaching the referendum only if she has no stakes or no influence. This result is not surprising since a forecaster who does not prefer one alternative over the other or cannot influence voting behavior does not face a trade-off between accuracy and the referendum outcome, and only aims to minimize the forecast error. The optimal bias, for any $\eta > 0$, $\gamma > 0$ is implicitly determined by (7) and takes the form

$$F^L - \mathbb{E}\left(y^L|\widetilde{y}^L\right) = -f\left(\eta, \gamma, C, \mu^L, y^R\right) < 0. \tag{8}$$

The optimal bias derived in (8) implies the following proposition.[5]

Proposition 2. ***Existence of political equilibrium with a propaganda bias*** *Under the assumptions of the model, necessary and sufficient conditions for $F^L \in \left[\underline{F^L}, \mathbb{E}\left(y^L|\widetilde{y}^L\right)\right)$ to be part of a political equilibrium are $\eta > 0$ and $\gamma > 0$.*

Proposition 2 predicts that in a political equilibrium a forecaster with stakes ($\eta > 0$) and influence ($\gamma > 0$) will publish biased estimates for alternative

[5]The forecasters' objective function is cubic in F^L and hence is convex only in a subset of its domain. Among the two solutions of (7), we select the one satisfying the second-order conditions with respect to F^L.

L approaching the referendum. The bias appears in the form of pessimistic forecasts for alternative L as forecasters with stakes are assumed to prefer alternative R.

Intuition and Mechanisms

This basic theoretical framework is tractable and straightforward. Nevertheless, it provides sufficient insights about the incentives that the asymmetry of information gives to the forecaster in this strategic game. A forecaster who releases biased estimates solves the trade-off between accuracy and the attempt to influence the outcome of the voting process by taking double advantage of its strategy. The optimal choice of F^L takes into account that if R prevails, the propaganda bias is costless in terms of ex-post accuracy. Indeed, the bias reduces the probability of paying the economic cost C in the event alternative L wins, but it also reduces the probability of paying the accuracy cost $\frac{1+\gamma}{2}\left[F^L - \mathbb{E}\left(y^L|\tilde{y}^L\right)^2\right]$. The strategic manipulation of the forecast is very appealing for the forecaster, who can take the opportunity to influence the voters at no cost. Voters, instead, face a utility loss compared to a world with unbiased forecasts if the propaganda bias is decisive to swing the referendum result.

The forecaster biases its estimates independently on voters' information on its stakes and influence and regardless of how voters form their expectations about the bias. In the case of rational voters, the propaganda bias reduces the welfare of the forecaster since it reduces its accuracy without influencing the expected referendum result. If voters are naive (i.e., if they expect the forecaster to not bias), the propaganda bias reduces the welfare of voters since they are induced to change the voting strategy systematically towards R.

This model imposes that $\eta \geq 0$ so that the forecaster either prefers R or is indifferent between the two alternatives. If, instead, $\eta \leq 0$, (i.e., that the forecaster is in favor of alternative L), then the propaganda bias would result in an overestimation of y^L. For instance, the latter would be consistent with the findings in Giles (2019), who documents that the HM Treasury, lead by a pro-Remain minister, released very pessimistic forecasts approaching the vote and that, conversely, the pro-Leave *Economists for Brexit* released very optimistic forecasts before the referendum.

2.2 Extended Model

In this version, we model the announcement periods, the referendum campaign, and the post-referendum periods. At the beginning of the game, the political economy is subject to a status-quo with an associated economic outcome, and J forecasters and voters are informed that a referendum will take place between the status quo and an alternative. During the campaign period k, the two alternatives are discussed, and voters have to choose between

$A \in \{L, R\}$, each of them associated with an economic outcome y^A. As in the basic version, L represents the decision to leave the status quo for the alternative, while R is the decision to remain.

We assume that voters do not observe y^A, although they know that $y^A \sim \mathcal{N}(\mu^A, \tau_A^{-1})$, such that $\tau_L = \rho\tau_R$ with $\rho \in (0,1)$, and use forecasts to update their beliefs over y^L and y^R before casting the vote. The parameter ρ indicates that voters have better knowledge about the current state R than about the alternative L. During periods 1 to $k-1$, each forecaster (indexed by j) releases a forecast for y^R, denoted $F^R_{j,t}$. During the campaign period k, each forecaster releases both $F^R_{j,k}$ and $F^L_{j,k}$. After the vote, forecasters will release only the forecast subject to the alternative winning the referendum. To simplify the exposition we now assume that forecasters perfectly observe y^L and y^R, and that voters are naive and expect forecasters to release unbiased forecasts.

Forecasters strategically release their forecasts to the public based on three competing incentives. First, they have incentives in minimizing the forecast error (i.e., the difference between the forecast and the ex-post realization of the economic outcome subject to the same state). Second, forecasters aim to influence voters' beliefs over the referendum by manipulating the released forecasts. Third, forecasts subject to the same state need to be consistent over time for the forecaster to preserve its credibility.

Voters

The voters' decision rule follows (1). Differently from the basic model, we assume now that in each period, they observe a noisy measure of $F^A_{j,t}$ from multiple forecasters

$$\widetilde{F}^A_{j,t} = F^A_{j,t} + e_{j,t}, \tag{9}$$

where the media transition error $e_{j,t} \sim \mathcal{N}(0, \tau_{j,t}^{-1})$, such that $\tau_{j,t}^{-1} = \beta\tau_{j,t-1}^{-1}$ with $\beta \in (0,1)$, captures the heterogeneity over time in the precision of the signal that forecasters send to voters. The assumption that the variance of the transition error is reduced every period reflects the idea that forecasts become more precise, as the forecast horizon decrease, approaching the realization of the outcome.[6] We define $\tau_j := \tau_{j,0}$ such that $\tau_{j,t}^{-1} = \beta^t\tau_j^{-1}$.

Voters aggregate information from different forecasters using Bayesian updating, and form posterior beliefs about the alternative states of the economy

[6] See, for example, Andersson et al. (2017) for an analysis of the expected size of forecast errors at different horizons.

before casting a vote. Formally,

$$\mathbb{E}\Big(y^L|\widetilde{F}^L_{1,k},...,\widetilde{F}^L_{J,k}\Big) = \frac{\tau_L}{\tau_L+\Big(\frac{1}{\beta}\Big)^k\sum_{j=1}^J\tau_j}\mu^L + \Big(\frac{1}{\beta}\Big)^k\sum_{j=1}^J\frac{\tau_j}{\tau_L+\Big(\frac{1}{\beta}\Big)^k\sum_{j=1}^J\tau_j}\widetilde{F}^L_{j,k}$$

$$= m\mu^L + \Big(\frac{1}{\beta}\Big)^k\sum_{j=1}^J\gamma_j\widetilde{F}^L_{j,k}, \tag{10}$$

and

$$\mathbb{E}\Big(y^R|\widetilde{F}^R_{1,1},...,\widetilde{F}^R_{J,k}\Big) = \frac{\tau_R}{\tau_R+\sum_{j=1}^J\sum_{t=1}^k\Big(\frac{1}{\beta}\Big)^t\tau_j}\mu^R + \sum_{j=1}^J\sum_{t=1}^k\frac{\Big(\frac{1}{\beta}\Big)^k\tau_j}{\tau_R+\sum_{j=1}^J\sum_{t=1}^k\Big(\frac{1}{\beta}\Big)^t\tau_j}\widetilde{F}^R_{j,t}$$

$$= \nu\Big[\frac{1}{\rho}m\mu^R + \sum_{j=1}^J\gamma_j\sum_{t=1}^k\Big(\frac{1}{\beta}\Big)^t\widetilde{F}^R_{j,t}\Big], \tag{11}$$

where $m := \frac{\tau_L}{\tau_L+\Big(\frac{1}{\beta}\Big)^k\sum_{j=1}^J\tau_j}$, $\gamma_j := \frac{\tau_j}{\tau_L+\Big(\frac{1}{\beta}\Big)^k\sum_{j=1}^J\tau_j}$ and $\nu := \frac{\tau_L+\Big(\frac{1}{\beta}\Big)^k\sum_{j=1}^J\tau_j}{\frac{1}{\rho}\tau_L+\frac{1-\Big(\frac{1}{\beta}\Big)^k}{\beta-1}\sum_{j=1}^J\tau_j}$, so that $\Big(\frac{1}{\beta}\Big)^k\sum_{j=1}^J\gamma_j = 1-m$.[7] Consistently with the basic version of the theoretical framework, the parameter γ_j is a sufficient statistics to determine the influence of each forecaster on the voters' posterior belief.

Naive voters do not expect forecasters to release biased $F^A_{j,t}$, and hence vote based on

$$\mathbb{E}\Big(y^L|\widetilde{F}^L_{1,k},...,\widetilde{F}^L_{J,k}\Big) \geq \mathbb{E}\Big(y^R|\widetilde{F}^R_{1,1},...,\widetilde{F}^R_{J,k}\Big) + \sigma_i. \tag{12}$$

Following the same steps as in the basic model, the probability that alternative L wins is

$$p^L = 1 - G\Big(m\Big(\frac{\nu}{\rho}\mu^R - \mu^L\Big) - \sum_{j=1}^J\gamma_j\Big[\Big(\frac{1}{\beta}\Big)^kF^L_{j,k} - \nu\sum_{t=1}^k\Big(\frac{1}{\beta}\Big)^tF^R_{j,t}\Big]\Big), \tag{13}$$

where $G\big(\cdot\big)$ is the cumulative distribution function of the gaussian random variable $\sum_{j=1}^J\gamma_j\Big[\Big(\frac{1}{\beta}\Big)^ke_{j,k} - \nu\sum_{t=1}^k\Big(\frac{1}{\beta}\Big)^te_{j,t}\Big]$. Analogously to the basic model,

[7] Notice that the final step of (11) comes from the application of the formula for the geometric series $\sum_{x=1}^n x^n = \frac{x-x^{n+1}}{1-x}$.

forecasters have the leverage to influence the voting result and to favor the most preferred outcome. The allocation of the propaganda bias between $F^L_{j,t}$ and $F^R_{j,t}$ depends on the number of pre-referendum time periods during which forecasters released a forecast for y^R, the relative precision of voters' prior ρ, and on the variation over time in the precision of forecaster's signal determined by β.

Forecasters

We assume a discrete number of J forecasters who have perfect information on y^L and y^R and can face an economic loss under L or be indifferent between the two alternatives. Each forecaster minimizes the following expected loss function with respect to $F^L_{j,t}$ and $F^R_{j,t}$, given y^A as well as other forecasters' and voters' strategies:

$$\begin{aligned}\min_{F^L_{j,t},F^R_{j,t}} \quad & \left(1-p^L\right)\sum_{t=1}^{k}\left(\frac{1}{\beta}\right)^t\left(1+\gamma^R_j\right)\left[\frac{\alpha}{2}\left(F^R_{j,t}-y^R\right)^2+\frac{1-\alpha}{2}\left(F^R_{j,t}-F^R_{j,t-1}\right)^2\right] \\ & +p^L\left[\left(\frac{1}{\beta}\right)^k\left(1+\gamma^L_j\right)\frac{\alpha}{2}\left(F^L_{j,k}-y^L\right)^2+\eta_j C\right] \\ & +\sum_{A\in\{L,R\}}p^A\sum_{t=k+1}^{T}\left(\frac{1}{\beta}\right)^t\left(1+\gamma^A_j\right)\left[\frac{\alpha}{2}\left(F^A_{j,t}-y^A\right)^2+\frac{1-\alpha}{2}\left(F^A_{j,t}-F^A_{j,t-1}\right)^2\right],\end{aligned} \tag{14}$$

where $F^A_{j,t}\in[\underline{F^A_j},\overline{F^A_j}]$ represents the forecast released by institution j in period t subject to alternative A. $C>0$ represents a cost associated with alternative L, p^L is the probability of leaving the status quo and the parameter $\eta_j\geq 0$ captures the stakes of each forecaster.[8] Moreover, $\alpha\in[0,1]$ indicates the relative weight of the incentive for accuracy compared to the incentive for consistency. In this version of the model, forecasters trade-off accuracy with the will of favoring the preferred outcome in the referendum, as well as the consistency of released estimates over time.

Predictions

The numerical solution of the extended model, whose details are presented in Section A.1 in the Appendix, provides additional insights beyond the prediction of a propaganda bias in the estimates released by macroeconomic forecasters with influence over voters and economic interest at stake during the referendum.[9] First, the model indicates how forecasters allocate the bias across

[8] $\eta_j\geq 0$ implies that we assume forecasters do not have a strict preference in support of L. The sign of η_j determines the sign of the propaganda bias at the individual level, but not its presence.

[9] In Figure A3 in the Appendix, we show how the main predictions of the theoretical model are affected by the calibration of parameters. These comparative dynamics exercises, on the

different alternatives when voters do not have complete information about any of the two. Second, the model predicts whether there also exists an intensive margin of propaganda bias. Third, the model shows the dynamic evolution of forecasts after the vote, when forecasters no longer have electoral incentives.

In Figure 1, we report how the bias in $F^L_{j,k}$ and $F^R_{j,k}$ vary as a function of η_j (stakes) and γ_j (influence). In both heatmaps, the red areas represent the cases of largest bias, whereas the combination of parameters for which the model does not predict any bias are reported in dark blue. The figure shows that, in equilibrium, forecasters with stakes and influence do not allocate equally the propaganda bias across alternatives. In the left panel, the estimates released by forecasters with highest values of stakes and influence are substantially different from y^L, while the right panel shows that the propaganda bias is, albeit present, very small to what concerns forecasts subject to alternative R.

(a) Alternative L *(b)* Alternative R

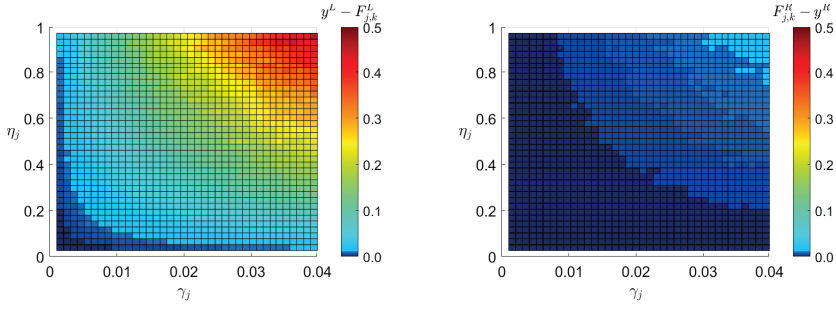


Figure 1. Propaganda Bias in period k

Notes: The figure reports the bias in $F^L_{j,k}$ and $F^R_{j,k}$ as a function of the individual parameters η_j (y-axis) and γ_j (x-axis). The parameters are reported on the axes of the graph, whereas different values of the bias are reported with different colors, as summarized in the legend. Dark blue represent $F^A_{j,k} = y^A$, whereas red represent the relatively most biased forecasts. Note that we have adjusted the sign of the bias to have the same color scale in both graphs. All other parameters have been calibrated as described in Section A.1 in the Appendix.

The intuition behind this result comes from the trade-off that forecasters with stakes and influence face at time k. Forecasters choose the propaganda bias optimally at the level that equalizes the marginal benefit (i.e., the impact on p^L scaled by the level of stakes) and the marginal cost. In terms of marginal benefit, (13) implies that a propaganda bias on $F^L_{j,k}$ is more profitable because voters weight more the signals subject to L observed in period k than the ones subject to alternative R. The influence of $F^R_{j,k}$ on the posterior belief is mitigated by all signals sent in the previous periods by forecasters, as well as by

one hand, show the robustness of the theoretical predictions, and on the other hand, provide additional insights useful to understand in depth the incentives that macroeconomic forecasters with stakes and influence face in connection with an important referendum.

the much more precise prior that voters have about on R than they do on L. Marginal cost of propaganda bias, according to (14), is instead higher under alternative R because forecasters would have to pay an additional cost for a large revision from the previous release, making it relatively more profitable to release a biased forecast subject to the less likely alternative.

Figure 1 clarifies the existence of an intensive margin of propaganda bias, in terms of both stakes and influence. Among institutions with $\eta_j > 0$ and $\gamma_j > 0$, indeed, there is a monotonic relationship between each of the two and the size of the bias. Forecasters with relatively higher stakes (influence) released more biased estimates than their colleagues with positive but lower levels of η_j or γ_j, holding the other individual parameter constant. While it is not surprising that higher stakes lead to a more pronounced propaganda bias, influence increases the marginal benefit, as well as the marginal cost of a biased forecast, and hence its impact on the size of the propaganda bias was ambiguous at first sight.

(a) Constant γ_j *(b)* Constant η_j

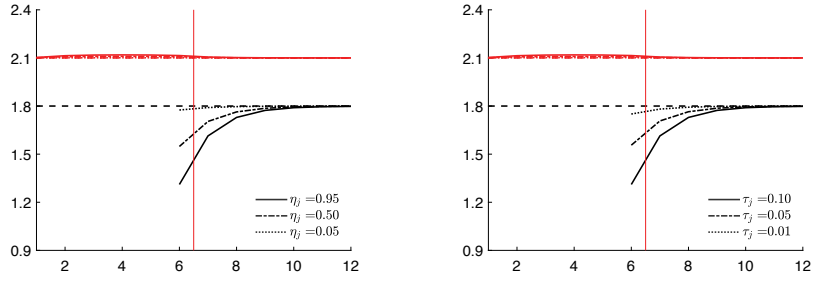


Figure 2. Evolution of propaganda bias over time

Notes: The figure reports the evolution of $F^L_{j,t}$ and $F^R_{j,t}$ over time. Dashed lines represent y^L and y^R, while solid lines as well as dotted and dashed-dotted represent forecasts. Black lines represent variables referring to alternative L, while red lines represent variables referring to alternative R. In the left graph, τ_j is held constant at 0.1, so that γ_j is on average equal to 0.04. In the right graph, η_j is held constant at 0.95. All other parameters have been calibrated as described in Section A.1 in the Appendix.

Figure 2 shows how $F^L_{j,t}$ and $F^R_{j,t}$ evolve over time around the referendum for different levels of η_j and γ_j. In the left graph, we hold constant γ_j and compare forecasters with different values of η_j. In the right panel, we hold constant η_j and compare forecasters with different values of γ_j. In both graphs, the solid lines, as well as the dotted and dashed-dotted lines, represent forecasts, while the dashed lines represent respectively y^R and y^L according to our baseline calibration. All variables referring to alternative R are reported in red, while variables referring to alternative L are shown in black. Figure 2 confirms the predictions presented in Figure 1 and highlights the convergence of $F^L_{j,t}$ after

the referendum. Even if forecasters with stakes and influence do not have any electoral incentives after the vote, the presence of an incentive for consistency over time implies that different forecasters release different estimates for some periods after the vote.

3 Taking the Model to Data

We test the predictions of our theoretical model using the EU membership referendum, known as Brexit, held in June 2016 in the United Kingdom. Several reasons make the Brexit referendum ideal for empirically testing the model. First, a withdrawal from the European Union would have exposed several forecasters to substantial losses. Second, it was difficult for voters to anticipate the effects of their choice on the economy since no country had previously withdrawn from the European Union. Third, the probability of leaving the European Union was considered low before the vote, making it relatively more convenient to bias the forecasts conditional on a leave outcome. In this section, we start by providing relevant details about the Brexit referendum. Section 3.2 then describes the forecast data, Section 3.3 describes the identifying assumptions, Section 3.4 describes our measures of stakes and influence, and Section 3.5 present our estimation strategy.

3.1 The Brexit Referendum

In January 2016, the UK Prime Minister David Cameron announced a referendum on the EU membership that would take place on June 23 of the same year. The referendum was formally non-binding since the Parliament maintained the right to make the final decision on the issue. Still, the government clarified before the vote its willingness to commit to the voters' preference.

During the campaign, which started in mid-April, the economic effects of the eventual withdrawal from the European Union, and, potentially, from the Single European Market (Dhingra et al., 2015; Kierzenkowski et al., 2016), were a major argument against Brexit. Governmental agencies, forecasters, media, and European and international public institutions warned the British citizens about a large economic downturn, especially due to a drop in investments (Dhingra et al., 2016a) and exports (Dhingra et al., 2016b) if the UK withdrew from the EU. The voters themselves seemed to be concerned about the future state of the economy. According to Google Trends summary reports, the number of online searches for "Brexit GDP," "Brexit pound," and "Brexit economy" increased between 10 and 100 times on a relative scale in the weeks approaching the referendum date (see Figure A1 in the Appendix). On May 23, the UK government's ministry for economics and finance (HM Treasury) released a report (HM Treasury, 2016) containing the estimated immediate economic impact of a Leave vote. According to the report, a Leave outcome

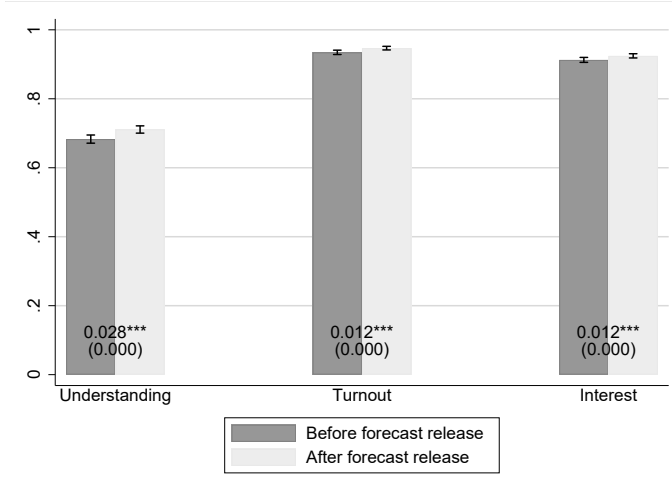


Figure 3. Changes in voters' beliefs after forecast release

Notes: The graph reports the share of respondents in the British Election Study panel (Wave 8) which a) agree or strongly agree with the following statements: *I have a pretty good understanding of the important issues at stake in the EU referendum*; b) announce that they would very likely or likely vote in the EU membership referendum; c) respond that they are very interested or somewhat interested in the EU referendum. Dark gray bars represent respondents surveyed before the release negative forecasts conditional on Leave by the HM Treasury on May 23rd, while light gray bars represent respondents surveyed after the release. The sample has been constrained to include only individuals surveyed in a window of ten calendar days around the HM Treasury release. The difference between the two sample average is reported with standard errors in parentheses. Standard errors and confidence intervals are robust to heteroskedasticity. *,**,*** represent the 10%, 5%, 1% significance levels.

would have generated a reduction in GDP between 3.6 and 6 percent within two years, with higher inflation and unemployment, as well as a depreciation of house prices and the sterling. We access the *British Election Study* panel data to correlate the release of the report with voters' beliefs on the relevance of the referendum and voters' understanding of the issues at stake. In the wave 8 of the *British Election Study* survey, a subsample of respondents was invited to complete the survey on each day approaching the referendum between May 6 and June 22. Figure 3 documents that respondents surveyed within ten days after the publication of the report were more likely to understand the relevant issues at stake on the referendum, more likely to participate in the election, and more interested in the debate approaching it than respondent surveyed in the days before the publication of the report. The evidence reported in Figure 3, and the Google Trends data reported in Figure A1, suggests that voters use, and search for, macroeconomic forecasts approaching the vote as a rele-

vant source of information and that they update their prior beliefs after being exposed to them.

In the weeks approaching the referendum, the Remain side was leading according to 66 percent of the opinion polls, and often with a winning margin of at least 5 percentage points. Macroeconomic forecasters, as well as bookmakers, were predicting the victory of the Remain side. According to Consensus Economics (2016a), forecasters were assigning a probability of 63 percent to Remain, whereas the bookmakers assigned Remain a probability around 85 percent in the final days before the vote (see Figure A2 in the Appendix). However, the referendum results reversed all predictions. On June 23, a majority (51.9%) of the voters decided to leave the European Union. Prime Minister David Cameron, who had campaigned to remain in the EU despite the opposition of several ministers and party colleagues, announced his resignation the day after the referendum.

3.2 GDP Forecasts

In Figure 4, we present a few examples of forecasts for the long-term impact of Brexit on GDP available to voters before the referendum. The figure shows that forecasts were largely heterogeneous, even if most of them agreed that Brexit would damage the economy. Consistently with the evidence reported in Giles (2019), we document that the think tank of pro-Leave economists *Economists for Brexit* (Minford, 2016) was one of the most optimistic about the potential effects of Brexit on the economy, while the HM Treasury, led by a pro-Remain minister, has been among the most pessimistic and incorrect forecaster.

However, a formal comparison across forecasters of these estimates would hardly be feasible since these forecasts differ in multiple dimensions. First, many forecasters provide bounds for the effect of the withdrawal on economic growth where intervals do not rely on the same set of assumptions about future agreements between the UK and the European Union. Second, many forecasters report estimates for the net effect of Brexit (i.e., they compare GDP if Brexit with baseline GDP) without providing additional information on their baseline forecast. Third, forecasts released at different points in time are heterogeneous in terms of available information.[10] Lastly, the forecast horizon considered by each forecaster may very, and long-term estimates do not allow to address which forecasters published more accurate estimates than the others.

[10]For example, Mansfield (2014), Bertelsmann Stiftung (2015) and Open Europe (2015) was released before 2016, while the others in Figure 4 was released in 2016. Also, those released within 2016 was conditional on different information sets, since, for example, Oxford Economics (2016) was published in March, OECD (2016) in April and Ebell and Warren (2016) in May.

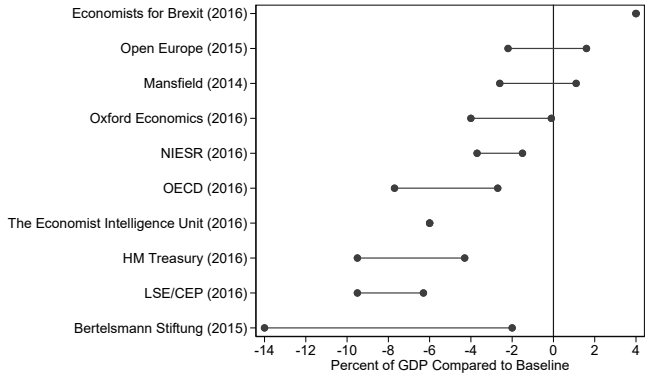


Figure 4. Long Run Effects on GDP

Notes: See cited papers for respective source. The Economists for Brexit is cited as Minford (2016), NIESR as Ebell and Warren (2016), and LSE/CEP as Dhingra et al. (2016a).

A standard approach in the literature to exploit comparable estimates across forecasters relies on surveys of professional forecasters. In surveys, all forecasters are asked their best estimates for the same target year at the same point in time, given a common set of assumptions. We use the data collection *Forecasts for the UK Economy* from the HM Treasury (n.d.). The dataset is a monthly survey of independent forecasters collected by the Treasury and is publicly available. The collection covers 44 forecasters from January 2012 to April 2018. The sample mainly contains financial institution and research companies, which are also the forecasters updating their estimates most frequently.[11]

The dataset neither includes the forecasts produced by the HM Treasury itself, nor those released by, for example, the *Economists for Brexit* pro-leave think tank as they are not producing forecasts regularly. At the beginning of each month, each forecaster in the sample is surveyed, and the results are quickly released online.[12] The data contain short-term forecasts for GDP growth and its components: private and government Consumption, Investments, Imports, and Exports. Our main focus is on the forecasts for the GDP growth rate in the year $t+1$, such that forecasts released in 2016 refer to GDP growth in 2017.[13] The use of short-term forecasts allows us to evaluate the ex-

[11] See Figure A5 and Table A3 in the Appendix for the distribution of forecasters by industry, as well for a comparison of the forecasts released before the announcement of the EU membership referendum.

[12] Not all forecasters release new predictions every month, but we observe when the latest available prediction was released so that we exclude the ones that were not updated in the occasion of a new survey from the empirical analysis.

[13] All data refer to the changes in annual figures expressed in percent. For the January collection, the forecasts refer to the year t.

post accuracy of the released forecast on the basis of the actual GDP growth rates. Nevertheless, short-term forecasts are also very informative about the potential effects of Brexit in the subsequent years, as short and long-term forecasts are positively correlated.[14] Table A4 in the Appendix provides descriptive statistics of the relevant forecasts.

A limitation of the HM Treasury (n.d.) survey is that forecasts collected each month are contingent on the most likely outcome based on available information. Hence, we observe conditional forecasts subject to the victory of Remain before the referendum, while we observe forecasts subject to the victory of Leave only after the referendum.[15] It is important to note that the forecasts subject to the victory of the Leave alternative should reflect that the actual withdrawal from the EU would have taken at least two years, according to the Treaty of the European Union (Art. 50). Therefore, the forecasts are subject to the referendum outcome and not the actual withdrawal.

3.3 Identifying assumptions

The theoretical model introduced in Section 2 predicts that institutions with stakes (i.e., $\eta_j > 0$) and influence (i.e., $\gamma_j > 0$) release biased forecasts subject to Leave before the referendum. However, the *Forecasts for the UK economy* data only measures the forecasts subject to Leave after the referendum. Therefore, a difference in forecasts subject to the Leave outcome would only indicate a propaganda bias if the forecasts released just after the referendum reflect those released just before.

We take advantage of a special release of a broadly equivalent survey of professional forecasters to assess the validity of the assumption that forecasts just after the vote reflect those just before. In the April, May, and June releases of the Consensus Economics report, surveyed forecasters were asked—besides the usual request for their central forecast analogous to the HM Treasury (n.d.) data—to report their conditional forecast subject to Leave anonymously. Consensus Economics and HM Treasury (n.d.) survey approximately the same set of forecasters. More specifically, only six institutions surveyed by the HM Treasury are not included in the Consensus Economics report, while Consensus Economics surveys three institutions not included in the HM Treasury (n.d.). Figure 5 shows that the average forecast subject to Leave as computed according to the anonymous survey conducted before the referendum in June

[14] Figure A6 in the Appendix shows that forecasts for GDP growth in the year $t+1$ and forecasts for GDP growth in the year $t+3$ are highly correlated in the subsample of institutions releasing medium-term forecasts on a quarterly basis. A regression controlling for time and individual fixed effects yields an R^2 of 0.55.

[15] Table A5 in the Appendix shows by comparing the June and July 2016 surveys how the sample averages changed substantially at the time of the referendum. More specifically, forecasts for GDP growth decreased from 2 percent to less than 1 percent.

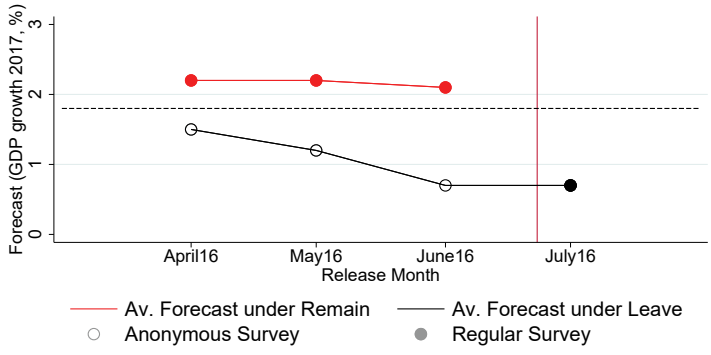


Figure 5. Consensus Forecasts for Leave and Remain around the Referendum

Notes: The graph reports the average forecast for the 2017 GDP growth rate conditional on Remain (red) and Leave (black), published by Consensus Economics in June and July 2016. The vertical line represents the referendum date, while the dashed horizontal line represents the actual GDP growth rate in 2017. Source: Authors' elaboration on data from Consensus Economics, 2016a and Consensus Economics, 2016b.

is equal to the average forecast subject to Leave collected in the first regular survey after the vote.[16]

The validity of our identifying assumption is further reassured by the fact that forecasters had only seven calendar days between the referendum and the day in which the HM Treasury began the survey for the July release of the collection.[17] In this limited window of time, it is unlikely that they updated the estimates or got new information about the economy other than the referendum result. Moreover, it is costly for forecasters to revise their estimates subject to the same alternative in the very short run. A large revision from a forecast under the Remain alternative to a forecast under Leave is justified and does not affect credibility, whereas the publication of a significantly different estimate from the one previously released that is subject to the same alternative would reduce credibility substantially. This argument is consistent with the forecasters' loss function in the extended model, as well as with the empirical results in Nordhaus (1987), who motivates that forecasters move away slowly from the last period's forecast to an emerging reality and the concept of consistency developed in Deb et al. (2018).

On the contrary, our identifying assumption would be violated if forecasters with stakes have biased beliefs about negative shocks to the economy that

[16]We can confirm the validity of our identifying assumption at the individual level for one institution included in the HM Treasury (n.d.) sample. The Economist Intelligence Unit (2016) shows that the forecast for GDP growth in 2017 released by the Economist Intelligence Unit in June 2016 before the vote is analogous to the forecast included in the July 2016 release of the *Forecasts for the UK Economy*.

[17]The July 2016 edition of *Forecasts for the UK Economy* was published on July 20 and contained information from forecasters surveyed between July 1 and July 13.

affect their profits. In Section 4.4, we address and rule out this possibility by comparing the results of our empirical analysis to their counterparts estimated before and after the 2008 financial crisis and the 2001 attacks to the World Trade Center in New York.

3.4 Measures of Stakes and Influence

In our model, stakes represent the economic loss that a forecaster faces in the event the United Kingdom leaves the European Union. We argue that Brexit is likely to damage financial institutions and the institutions located in the City of London financial district, more than other forecasters. Hence, we measure stakes with an indicator equal to 1 if the forecast is a financial institution and 0 otherwise, and alternatively with an indicator capturing whether the institution is located in the City of London's financial district. Ramiah et al. (2017) estimates that the victory of Leave has reduced the stock market prices of the banking sector by 15.37 percent in the very short run compared to baseline. Our data show that the financial institutions in our sample faced, on average, a reduction in stock market prices of 16.37 percent in the two days after the referendum. In our sample, we observe 22 banks and 25 forecasters located in the city of London's financial district, respectively.

Among financial institutions, we use the percentage decline in stock market prices in the two banking days after the referendum to obtain a variation in stakes at the intensive margin (see Section A.2 in the Appendix for details). Forecasters have been very differently exposed to the immediate effects of Brexit, as reported in Figure 6, which shows stock market losses ranging between 1.8 percent and 31 percent. Nevertheless, Figure 6 reassures that none of the institutions listed on financial markets increased its market capitalization just after the referendum.

It is not obvious how to measure influence. In the model, influence represents the weight that each individual forecaster has in the formation of voters' posterior beliefs. We propose proxies that aim at capturing how known each institution is and whether it is established in the UK public debate. The first approach measures influence from the point of view of the public. We use Google Trends to measure how often the users search for an individual forecaster on the web.[18] The second approach aims at capturing media coverage. We retrieve the number of times in which each institution is mentioned in a Google News search during 2015. In both cases, we create an indicator equal to 1 if the institution scores above the median forecaster and 0 otherwise to investigate the extensive margin. We then use the full support of the Google

[18]Google Trends releases a normalized score on a weekly basis, such that the value 100 is assigned to the most-visited forecaster in the week of the largest number of visits. We then aggregate all summary reports for the year 2015, before the announcement of the referendum.

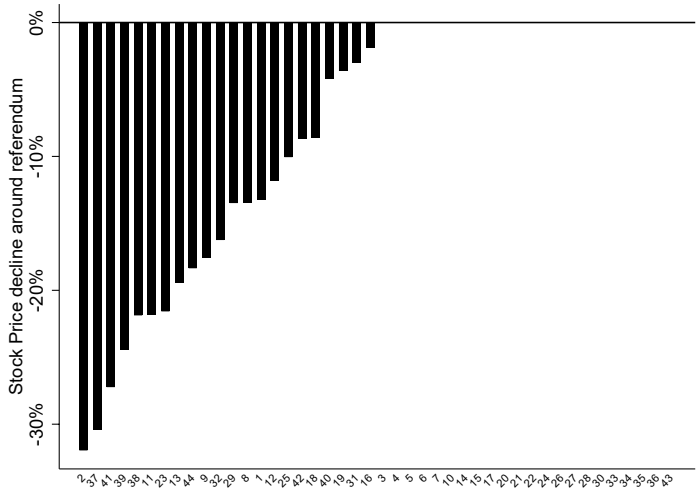


Figure 6. Stock Price Decline

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. The horizontal axis shows a forecaster ID. The figure reports for each financial institution the percentage variation in stock market prices between the referendum date and the second market day after the vote.

Trends and Google News measures (in logs) to proxy for influence at the intensive margin.[19]

In Section 4.1, we exploit heterogeneity in influence towards media outlets targeting different groups of voters. For each media outlets, we retrieve the number of articles that mention GDP forecasts released by an institution during 2015, and create an indicator equal to 1 if forecasts are reported more than the forecasts from the median forecaster, and zero otherwise.

3.5 Estimation

We investigate the existence of a propaganda bias by estimating the following baseline regression model

$$F_{j,m} = \theta_j + \delta_m + \mathbb{1}(\eta_j \gamma_j > 0) \sum_{k=-5}^{4} \beta_k \mathbb{1}(m = k) + \varepsilon_{j,m}, \tag{15}$$

where θ_j represents the forecaster fixed effects, δ_m represents the survey time effects and $k = -5, ..., 4$ measures the distance in months from the first sur-

[19] See Section A.2 in the Appendix for details on the group assignment and how the different definitions are correlated (see Table A6 in the Appendix). Depending on the definition, we identify between 15 and 17 forecasters to simultaneously have stakes and influence according to our measures.

vey after the vote. The indicator function $\mathbb{1}(\eta_j \gamma_j > 0)$ allows us to compare forecasters with stakes (η_j) and influence (γ_j) to the other institutions in the sample.

The dependent variable is the forecasts for GDP growth rate in the next year, where $F_{j,m}$ is the central forecast released by institute j in survey month m. β_0 estimates the propaganda bias around the date of the referendum, while $\beta_1, \ldots, \beta_4$ capture the eventual persistence of the effect after the vote and $\beta_{-1}, \ldots, \beta_{-5}$ reflect different judgments across groups between the announcement of the referendum and the vote. A negative β_0 would be consistent with the theoretical prediction that forecasters with stakes and influence have intentionally released pessimistic forecasts to influence voter behavior.

In the model, η_j and γ_j are treated as exogenous parameters, but empirically they are potentially correlated with omitted variables that also affect the published forecasts. For instance, influential forecasters might have become such because they have had better accuracy in the past or forecasters with stakes might be more pessimistic than others at any period. The panel structure of our data allows us to control for all time-unvarying characteristics that are determinant of published forecasts and potentially correlated with stakes or influence in a dynamic difference-in-differences setup under the parallel trends assumption, where the treatment (i.e., the referendum result) is at least partially unexpected by the forecasters and only changes the central forecast from alternative R to alternative L.[20] In order to corroborate that our estimates are not due to omitted confounders or selection, we also show the results using a specification excluding the forecaster-specific fixed effects as well as several robustness checks (see Section 4.4).

Economic forecasts are serially correlated due to persistence and the structure of annual outcomes, and they are potentially correlated across different institutions within the same survey date since institutions share information and models at least partially (see, e.g., Andersson et al., 2017; Davies & Lahiri, 1995). For this reason, we use standard errors robust to two-way clustering (Cameron et al., 2011; Cameron & Miller, 2015) at the forecaster and the survey levels.

Our measures of stakes and influence defined in Section 3.4 identify which forecasters have higher stakes and greater influence in the sample, but they do

[20]The literature investigating correlations and plausible causal relationships between socioeconomic, historical and demographic characteristics of UK districts and the referendum results have constantly been increasing in the past few months. For instance, Viskanic (2020) finds that areas with a higher concentration of Polish migrants are associated with a larger vote share of the Leave. On the contrary, Becker et al. (2017) do not find any correlation between migration, trade exposure, and the variation across-districts in support of the Leave side, while individual characteristics such as per-capita income in the district and education have a much larger explanatory power and a negative effect. Alabrese et al. (2019) find using a large individual-level survey that the support for Leave is associated with personal characteristics like age, ethnicity, education, use of smartphones and the internet, and life satisfaction.

not guarantee that the remaining institutions have no stakes or no influence. If some forecasters with positive stakes and influence turned out to be in the *control* group, then our estimates would suffer an attenuation bias. First, all the forecasters that the HM Treasury (n.d.) reports in the survey might be influential. In that case, it should be assumed that $\gamma_j > 0$ for all institutions. Second, if all forecasters have stakes, then it should be assumed that $\eta_j > 0$ for all. We limit this potential concern by proposing two additional specifications in which we compare institutions with and without stakes, and institutions with and without influence. We expect to detect a larger coefficient in absolute terms in the event of an attenuation bias or, conversely, an attenuated coefficient.

4 Results

We report the estimation results for the extensive margin of propaganda bias in Table 1. In column (1), we suppress the forecaster-specific fixed effects, and columns (2)–(4) report the results from estimating the model in equation (15) using different measures of stakes and influence. In column (5), we compare institutions with stakes and forecasters without, while in column (6), we compare influential and non-influential forecasters. The difference in forecasts released by the two groups of forecasters in the first survey after the referendum is reported in the first row of the table, while coefficients labeled with $(+1)$, ..., $(+4)$ estimate the eventual persistence of the difference in the subsequent months, and coefficients labeled with (-1), ..., (-5) capture pre-referendum differences in forecasts conditional on a Remain outcome.

In column (2) of Table 1, we estimate that forecasters with stakes and influence published a GDP growth rate forecast that was 0.638 percentage points lower than the other institutions. The result is larger in magnitude than the coefficient in column (1), suggesting that the potential selection bias at work without accounting for the unobserved heterogeneity would have underestimated the propaganda bias of forecasters with stakes and influence. In columns (3) and (4), we estimate coefficients of -0.745 and -0.601, respectively, showing that results are robust to changes in the measures of stakes and influence. In columns (5) and (6), we estimate a coefficient of -0.755 percentage points for the forecasters with stakes compared to their competitors and of -0.807 percentage points for the institutions with influence.

All specifications strongly confirm the predictions of our theoretical model about the presence of a propaganda bias, namely that forecasters with stakes and influence released more pessimistic forecasts for GDP growth subject to the victory of the Leave around the Brexit referendum. In opposite, we document that anticipated coefficients, subject to the victory of the Remain, are never distinguishable from zero. This evidence is in line with the extended model, which predicted a negligible propaganda bias subject to Remain.

Table 1. *Estimation of Propaganda Bias in GDP Growth Forecasts*

	Stakes × Influence				Stakes	Influence
	(1)	(2)	(3)	(4)	(5)	(6)
Group × Ref.	-0.526***	-0.638***	-0.745***	-0.601***	-0.755***	-0.807***
	(0.181)	(0.168)	(0.185)	(0.171)	(0.204)	(0.164)
Gr. × Ref. (+1)	-0.613***	-0.598***	-0.529***	-0.597***	-0.743***	-0.451***
	(0.146)	(0.177)	(0.177)	(0.176)	(0.146)	(0.171)
Gr. × Ref. (+2)	-0.460***	-0.405***	-0.471***	-0.450***	-0.536***	-0.491***
	(0.154)	(0.150)	(0.148)	(0.149)	(0.155)	(0.152)
Gr. × Ref. (+3)	-0.420***	-0.483***	-0.473***	-0.450***	-0.479***	-0.443***
	(0.159)	(0.149)	(0.154)	(0.150)	(0.151)	(0.153)
Gr. × Ref. (+4)	-0.121	-0.127	-0.157	-0.125	0.001	-0.370***
	(0.145)	(0.120)	(0.120)	(0.120)	(0.149)	(0.127)
Gr. × Ref. (−1)	0.015	0.001	0.005	0.002	-0.033	0.073
	(0.114)	(0.092)	(0.093)	(0.092)	(0.111)	(0.090)
Gr. × Ref. (−2)	-0.050	-0.076	-0.026	-0.074	-0.077	-0.000
	(0.114)	(0.097)	(0.098)	(0.095)	(0.113)	(0.095)
Gr. × Ref. (−3)	-0.044	-0.115	-0.067	-0.114	-0.092	-0.017
	(0.118)	(0.087)	(0.090)	(0.087)	(0.097)	(0.092)
Gr. × Ref. (−4)	0.085	0.053	0.081	0.055	-0.004	0.135
	(0.144)	(0.099)	(0.097)	(0.098)	(0.127)	(0.102)
Gr. × Ref. (−5)	-0.169	-0.157	-0.168	-0.128	-0.116	-0.079
	(0.131)	(0.108)	(0.110)	(0.107)	(0.113)	(0.117)
Observations	1,643	1,643	1,643	1,643	1,643	1,643
R^2	0.679	0.775	0.776	0.775	0.778	0.777
Fixed Effects		✓	✓	✓	✓	✓
Survey Effects	✓	✓	✓	✓	✓	✓
Stakes	Banks	Banks	Banks	City	Banks	
Influence	GTrends	GTrends	GNews	GTrends		GTrends

Notes: The dependent variable is GDP growth rate in period $t+1$. For each column, the column title defines the group assignment. The estimated equation is (15). Standard errors robust to two-way clustering at the forecaster and the survey levels are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

The estimated propaganda bias is very large, statistically significant, and precisely estimated. It explains, depending on the specification, up to 50 percent of the forecast error at the time of the referendum.[21] In all specifications, we find that in the subsequent surveys, the two groups weakly converge in their forecasts as point estimates approach zero after four months in five of the specifications. This additional evidence is consistent with a two-fold interpretation: first, the decision-making process leading to the withdrawal of the UK from the European Union did not end with the realization of the referendum result, potentially inducing forecasters to continue biasing the estimates. In-

[21] Figure A8 reports the distribution of released forecasts just before and just after the referendum. It shows that in the first survey after the referendum there was a clear cluster of forecasters with stakes and/or influence in the bottom of the distribution of published scenarios, whereas this evidence was not in place in the June survey.

deed, the victory of the Leave side opened new discussions among policymakers about the terms of the negotiation with the EU partners, the opportunity of remaining or not in the Single European Market (*soft* or *hard* Brexit) and the choice of the new prime minister after the immediate resignation of PM David Cameron. Second, even when the forecasters who were trying to influence the policy-making process no longer have had incentives to pursue their objectives, it might have taken time to converge back to their competitors' forecast to not lose credibility compared to their competitors, as predicted by the theoretical model presented in Section 2.

Evidence in support of the parallel trends assumption is presented in Figure 7, in which we plot the average GDP growth rate forecast for the following year released by the two groups of institutions. The figure shows that for many time periods prior to the referendum, forecasters with stakes and influence and those without released on average the same GDP growth rate forecasts (see Figure A9 in the Appendix for the other group specifications).

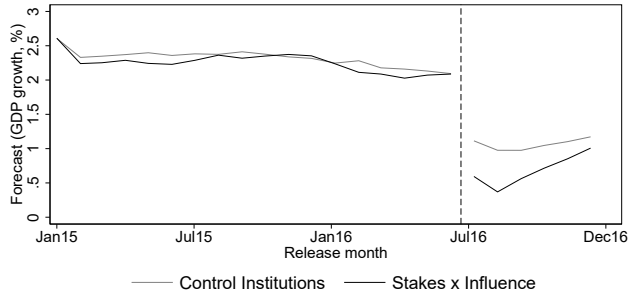


Figure 7. Pre-Referendum Trends

Notes: The measure of stakes is financial institution (Banks). The measure of influence is Google Trends. The graph plots the average GDP growth rate forecast for the next year among forecasters belonging to different levels of stakes and influence. The black line represents the group with stakes and influence, while the gray line represents the remaining control institutions.

4.1 Alternative mechanisms

In principle, the empirical results presented in Table 1 can be interpreted in light of two alternative mechanisms to the one shown in the theoretical model, which focuses on electoral incentives. On the one hand, if voters are more likely to believe in information confirming their priors and those of them willing to vote for Leave select into different media outlets compared to those willing to vote for Remain, then market incentives may induce forecasters reaching different groups of voters to produce alternative estimates. On the other hand, if forecasters' beliefs about the future states of the economy are

affected by their exposure, then our result would reflect differences in beliefs rather than strategic manipulation to affect voters. In what follows, we elaborate on each of the two alternative mechanisms and provide empirical evidence against the interpretation of our findings in that light.

Selective exposure. Suppose voters with different idiosyncratic preferences choose to gather information from alternative media outlets because they are more willing to believe information confirming their priors. Market incentives induce mass media to publish information aligned with the opinions of their customers, and in turn, to select macroeconomic forecasts aligned with those views. In turn, market incentives would also induce forecasters to please their final readers and media outlets in which their estimates are published more often.

We consider four major UK general-interest daily newspapers different in their market shares and their readers' political views, as well as a daily newspaper, specialized in the financial sector. Figure 8 reports the share of respondents in the *British Election Study* panel data considering each media outlet as the newspaper they read more often by political preferences (Panel (a)) and by support for Leave and Remain in the occasion of the 2016 EU Membership referendum. Figure 8 shows evidence of selective exposure based on voters' political positions and support for Leave or Remain in the EU Membership referendum (Panel (b)). More specifically, conservative and UKIP voters, as well as supporters of Leave are more likely to read *The Daily Mail* and *The Telegraph* than Labour and Liberal Democrats (lib-dem) voters as well as supporters of Remain, who mostly read *The Guardian*. The *Financial Times*, whose customers have a special interest in the financial sector, is not popular across different voters' preferences.

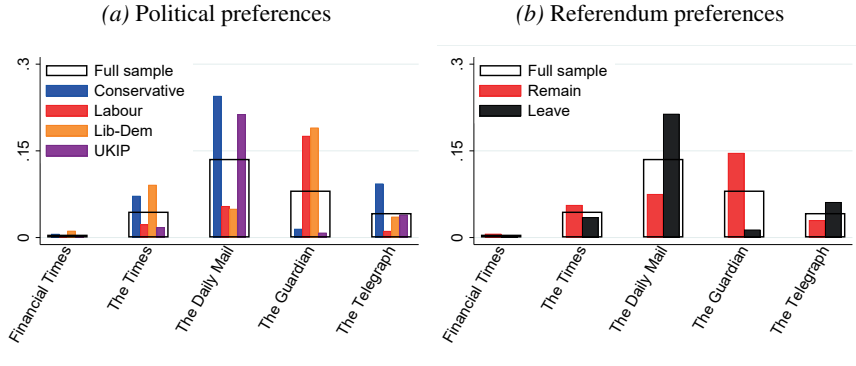


Figure 8. Newspaper Exposure by Political and Referendum Preferences

Notes: The graph reports the share of respondents in the British Election Study panel (Wave 10), which mentions each newspaper as the one they read most often. In Panel (a), the sample is divided based on the party that each respondent is most likely to vote for in the following general election. In Panel (b), the sample is divided based on the self-reported voting decision in the EU Membership referendum.

If the evidence of biased forecasts released by institutions with stakes and influence were driven by voters' selective exposure and the related market incentives to media and forecasters, then we should expect that forecasters whose estimates are reported more often on *The Guardian* should release the most biased forecasts. Conversely, forecasters often mentioned on *The Daily Mail* and *The Telegraph* should not bias their releases. Table 2 disproves this conjecture. In each column of Table 2, we replace our baseline indicators of influence with indicators of influence varying at the newspaper level.[22] More specifically, we compute the number of articles in which the estimates for GDP growth released by each forecaster were mentioned in a specific newspaper during 2015 and divide the sample at the median value.

Table 2. *Estimation of Propaganda Bias: Influence by mass media*

	Stakes × Influence				
	(1)	(2)	(3)	(4)	(5)
Group × Referendum	-0.363**	-0.812***	-0.604***	-0.359**	-0.773***
	(0.176)	(0.181)	(0.207)	(0.183)	(0.167)
Group × Ref. (+1)	-0.909***	-0.566***	-0.318*	-0.746***	-0.915***
	(0.166)	(0.177)	(0.192)	(0.177)	(0.163)
Group × Ref. (+2)	-0.517***	-0.473***	-0.335**	-0.475***	-0.452***
	(0.136)	(0.146)	(0.153)	(0.144)	(0.141)
Group × Ref. (+3)	-0.454***	-0.463***	-0.353**	-0.483***	-0.510***
	(0.153)	(0.158)	(0.164)	(0.154)	(0.153)
Group × Ref. (+4)	-0.096	-0.125	0.134	-0.047	-0.030
	(0.123)	(0.123)	(0.130)	(0.129)	(0.130)
Observations	1,643	1,643	1,643	1,643	1,643
R^2	0.775	0.777	0.773	0.774	0.778
Fixed Effects	✓	✓	✓	✓	✓
Survey Effects	✓	✓	✓	✓	✓
Measure of Stakes	Banks	Banks	Banks	Banks	Banks
Measure of Influence	Financial Times	The Times	The Daily Mail	The Guardian	The Telegraph

Notes: The dependent variable is GDP growth rate in period $t+1$. The relevant group assignment is Stakes x Influence. In each column, the measure of influence is an indicator equal to 1 if GDP growth forecasts released by the forecaster were mentioned during 2015 at least as much as the median forecaster in the respective news outlet. The estimated equation is (15). Standard errors robust to two-way clustering at the forecaster and the survey levels are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

The estimated coefficients reported in Table 2 show that the most pessimistic forecasts are published by forecasters often mentioned in newspapers read by a heterogeneous pool of voters—as *The Times*—or by voters more likely to have conservative preferences and to vote in support of the Leave. The table also excludes another channel of selective exposure. If individuals in the financial sector prefer to read negative news about the potential effects

[22] See Section A.2 in the Appendix for additional information about data sources and details about the definition of the measures.

of Brexit on the economy, then forecasters more likely to have their estimates reported on the *Financial Times* should release very pessimistic releases. In contrast, in columns (1) and (4), the estimated coefficients are substantially smaller than in the other specifications. The results in Table 2 are corroborated by Figure A10 and Table A7 in the Appendix, which show that the relative influence of each individual forecaster is homogeneous across media outlets and that the correlation between the measures used in Table 2 is not negligible.

Biased beliefs. Suppose forecasters with stakes in the referendum have biased beliefs about the potential effects of leaving from the EU, while forecasters not exposed to Brexit rely on more objective estimates. Then, our results would reflect differences in beliefs rather than in the attempt to influence the referendum outcome. To show that our results cannot be explained in light of forecasters' biased beliefs in connection with negative shocks to the economy, we compare the evolution of released forecasts after the referendum as well as on occasions of other major shocks to the UK economy, unrelated with electoral incentives.

First, the convergence after the referendum of the forecasts released by different institutions is not consistent with the possibility that forecasters with stakes have biased beliefs, which would have required the two groups to behave differently from each other in the subsequent months after the vote as well.

Second, forecasts released during the same months of 2017, after the referendum but before the realization of the Leave outcome, and subject to the same horizons as the forecasts exploited in the main analysis rule out the possibility that the convergence is due to the approaching horizons. More specifically, in Figure A11 in the Appendix, we estimate the same regression model as in equation (15) at different points in time to assess whether there is evidence of similar estimates in other periods. The figures show a large jump in the estimates at the referendum and the months just following, while the pre-referendum estimates are centered at zero. Moreover, forecasters both with and without stakes and influence publish very similar estimates throughout the year 2017 and especially in the releases subject to the same forecast horizon, confirming that our results are not consistent with forecasters' biased beliefs about the economic effects of Brexit.

Third, we estimate a version of equation (15) in which we compare financial institutions and other forecasters before and after the unexpected beginning of the 2008 financial crisis and the 2001 attack on the World Trade Center in New York.[23,24] The results in Table A8 in the Appendix, summarized in

[23] We identify the unexpected beginning of the financial crisis with the bankruptcy of the Lehman Brothers Holdings Inc. on September 15, 2008.

[24] In the time span of our main data (from January 2012 to April 2018), we do not identify any negative event that can be comparable to the withdrawal from the European Union. Therefore, we digitize the older publications of the *Forecasts for the UK Economy* collection from The

Figure 9, show that in the first survey after each event there is no evidence of different behavior of institutions with stakes compared to their competitors. Coefficients are never distinguishable from zero in the first and second survey after the event, and they are much smaller in magnitude compared to the ones estimated in proximity to the EU membership referendum.[25] Moreover, we do not observe a significant revision of the forecasts in the first survey after the events, confirming that forecasters are unlikely to adjust their forecasts during a very limited window of time after an unexpected shock.

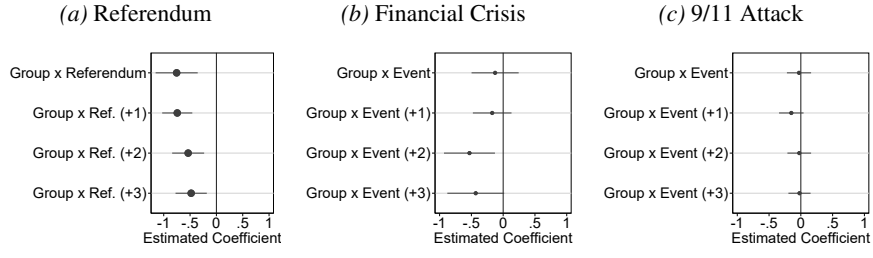


Figure 9. Effect during the EU Referendum and in the Occasion of Other Events

Notes: The measure of stakes is financial institution (Banks). The graphs report estimated coefficients and 95% confidence intervals from estimating (15), assuming that everyone is influential, in the occasions of the EU membership referendum; of the bankruptcy of the Lehman Brothers; and of the attacks to the World Trade Center in New York. The dependent variable is the GDP growth rate in the period $t+1$. Standard errors are robust to twoway clustering at the forecaster and the survey levels. Confidence intervals represent the 5% significance level.

4.2 Intensive Margin

The theoretical model also predicts that among forecasters with stakes in and influence over the referendum outcome, an increase in either characteristic is associated with a more pronounced bias. This section tests that prediction.

As described in Section 3.4, we measure stakes using the short-run percentage decline in the stock market prices, and proxy for influence using a continuous version of the Google Trends and Google News variables described earlier (in logs). In Table 3, we estimate Difference-in-Differences models of

National Archives (n.d.), enlarging the sample back to the year 1998. We investigate forecasters' behavior around the time of the 2001 terrorist attack by estimating equation (15) in a sample between the year 1998 and the year 2003, while we explore the reaction to the financial crisis restricting the sample to observations between the year 2004 and the year 2010. Compared to the sample used in the main empirical analysis and described in Section 3.4, a slightly different group of forecasters has been surveyed in the more recent publications.

[25] In the event of the financial crisis forecasters of both groups released more optimistic forecasts than the observed realization of the outcome. An "overreaction" of institutions with stakes should be interpreted as a better forecast, rather than a low-quality one driven by panic.

the form

$$F_{j,m} = \theta_j + \delta_m + \beta_1 \eta_j \mathbb{1}(\eta_j \gamma_j > 0)\mathbb{1}(m = 0) + \beta_2 \gamma_j \mathbb{1}(\eta_j \gamma_j > 0)\mathbb{1}(m = 0) + \varepsilon_{j,m}, \tag{16}$$

where the terms $\eta_j \mathbb{1}(\eta_j \gamma_j > 0)\mathbb{1}(m = 0)$ and $\gamma_j \mathbb{1}(\eta_j \gamma_j > 0)\mathbb{1}(m = 0)$ represent the interaction between the group indicator with the intensive margin variables γ_j and η_j at the time of the referendum.

The results in Table 3 confirm the predictions about the existence of an intensive margin of propaganda bias. In columns (1) and (2), where we estimate the coefficients β_1 and β_2 in two separate regressions, we detect a large and negative correlation between the continuous measures of stakes and influence and the released forecast at the time of the referendum. Specifically, a one-standard-deviation increase in the stock price loss after the referendum is associated with more pessimistic forecasts of 0.361 percentage points, while a one standard deviation increase in influence is associated with a lower $F_{j,k}$ of 0.252 percentage points.

Table 3. *Estimation of Propaganda Bias at the Intensive Margin*

	Stakes × Influence			Stakes	Influence
	(1)	(2)	(3)	(4)	(5)
Group × Ref. × Stock Price	-0.361***		-0.316***	-0.330***	
	(0.094)		(0.102)	(0.098)	
Group × Ref. × log(Trend)		-0.252***	-0.067		-0.297***
		(0.093)	(0.084)		(0.092)
Observations	1,643	1,643	1,643	1,643	1,643
R^2	0.770	0.769	0.770	0.770	0.769
Fixed Effects	✓	✓	✓	✓	✓
Survey Month Effects	✓	✓	✓	✓	✓
Measure of Stakes	Banks	Banks	Banks	Banks	
Measure of Influence	GTrends	GTrends	GTrends		GTrends

Notes: The dependent variable is GDP growth rate in period $t+1$. For each column, the column title defines the relevant group assignment. Stock Price and log(Trend) have been standardized to have mean 0 and variance 1. In all specifications, continuous measures of stakes and influence have been standardized to have zero mean and unit variance. The estimated equation is (16). Standard errors robust to two-way clustering at the forecaster and the survey levels are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

In column (3), we estimate the parameters β_1 and β_2 in the same regression, as stated in equation (16), and confirm that both variables are negatively correlated with the forecast for GDP growth rate in the next year, although the coefficient attached to the continuous measure of influence is not statistically significant.

In columns (4) and (5), we repeat the exercise interacting the continuous measures of stakes and influence with the groups defined in columns (5) and (6) of Table 1, detecting large and statistically significant negative coefficients.

Additional evidence that the bias is determined also at the intensive margin is presented in the Appendix. In Table A9 we repeat the exercise done in Table 3 using Google News instead of Google Trends as the measure of influence, and in Figure A12 we show that there is a negative relationship between the distance $F^L_{j,k} - F^R_{j,k}$ and the continuous measures of stakes and influence.

4.3 GDP Decomposition

The empirical results show that the forecasters with stakes and influence predicted a larger downturn in the economy than their competitors. We proceed by decomposing the effect on the GDP growth rate in its components. This contributes to the interpretation of the forecasters' behavior around the time of the referendum, as it highlights whether biased forecasts were published based on a precise rationale consistent with the general beliefs about the potential economic effects of Brexit. If forecasters conducted the propaganda bias in a rational manner, we expect to detect heterogeneous effects in line with the supposed economic effects of a Leave outcome and consistent with predictions from standard macroeconomic models. Investments and trade are volatile and pro-cyclical, while consumption does not react as much, and government expenditure usually increases as a response to economic crises.

According to the expenditure approach, the GDP can be decomposed as follows:

$$Y = C + I + G + (X - M), \tag{17}$$

so that the GDP growth rate can be expressed as

$$g_Y = g_C \varepsilon_C + g_I \varepsilon_I + g_G \varepsilon_G + (g_X \varepsilon_X - g_M \varepsilon_M), \tag{18}$$

where C is household consumption, I is investments, G is government consumption, X is exports and M is imports, while g and ε represent respectively the growth rate of each component and its share of GDP. In Figure 10 we plot results of estimating equation (15) for each component. The symbols report the estimated propaganda bias at the time of the referendum (β_0 in equation (15)) for each of the specifications used in Table 1.

The results show that the propaganda bias in investments is very pronounced, as estimated coefficients are around -2 percentage points, significant at least at the 10 percent level for most specifications. Consistent with the stylized evidence that investments are usually much more volatile than GDP, and that they are supposed to be among the major driving channels of the economic effects of Brexit (Dhingra et al., 2016b), the estimated coefficients are much larger than the ones estimated for GDP growth. In addition, our data forecast the short-term effects of the referendum result, and investments usually react immediately to changes in the political or economic environment.

Trade is expected to be another major channel of the effect of Brexit on economic growth (Dhingra et al., 2016a). Looking at the trade components of

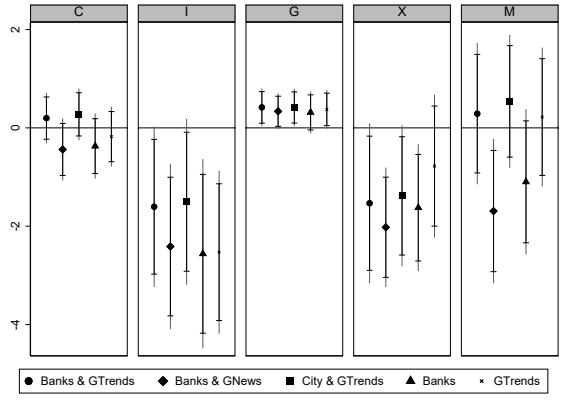


Figure 10. Estimation of Propaganda Bias in GDP Growth Components

Notes: The dependent variables are Household consumption, Investments, Government consumption, Exports and Imports in period $t+1$. All specifications include individual fixed effects and survey fixed effects. The estimated equation is (15), and the reported coefficient is β_0. Standard errors robust to two-way clustering at the forecaster and the survey levels. Gray lines represent 5% significance level while black lines represent 10% significance level.

GDP, we find large and negative coefficients on export, of similar magnitude as the growth of investments, while estimated coefficients on imports are not distinguishable from zero. Overall, the institutions which released biased forecasts for GDP growth also predicted lower trade activity than their competitors due to a pessimistic forecast on export growth.

As expected, the coefficients on household consumption are never statistically different from zero at conventional significance levels. We detect small positive coefficients on government consumption, which translate into more pessimistic forecasts released by institutions with stakes and influence, even if generating the opposite effect on GDP growth according to equation (18). This exercise shows that forecasters with stakes and influence conducted their propaganda bias by reporting much more negative views on investments and export growth, together with an excessive increase in government consumption to counteract part of the downturn.

4.4 Robustness Checks

We perform a number of robustness checks to validate our empirical strategy and exclude that our estimates are driven by chance or large variability in a relatively small sample.

First, we reduce the number of surveys included in the sample to the months much closer to the referendum. Figure A13 in the Appendix reports the estimated coefficients and confidence intervals for β_0 estimated with the support

of several different windows of time. Estimated coefficients are stable for all specifications and are not sensitive to the time span of the data.

Second, we show that our results do not depend on the arbitrary thresholds chosen to determine the most influential institutions in the sample. Specifically, we move the thresholds used to separate influential and non-influential forecasters in order to alter the composition of the two groups. The results of this exercise, as presented in Figure A14, show that a large propaganda bias is estimated using other thresholds as well and that the coefficients reported in Table 1 do not represent extreme estimates.

Third, we replicate the main results using alternative measures from Google News that disentangle the general knowledge of each institution from the influence of its forecasting activity. The measure used in Table 1 captures the overall media coverage of each forecaster, thus pooling cases in which each institution was mentioned in the UK news for several possible reasons.[26] In Table A10, we show that the results hold when we refine the Google News measure by requiring the words "GDP" and "forecast" to be part of the same newspaper article as the name of the forecasting institutions.

Fourth, we perform a Monte Carlo simulation with 10,000 draws, in each of which we randomly assign half of the institutions to a placebo treatment group, and estimate equation (15). Figure A15 shows the empirical density of the coefficient estimated at every draw, as well as where in the distribution the coefficients reported in Table 1 lie. Our results, as expected, always lie in the lower parts of the distribution, which is symmetric and centered in zero.

5 Concluding Remarks

Voters are seldom completely aware of different political platforms and the economic consequences of their choices before casting a vote since they lack incentives to invest in gathering costly information. Traditionally, we think of special interest groups and media as having some monopoly power and releasing biased pieces of information in order to affect individuals' beliefs and, in turn, their voting behavior.

In this paper, we have introduced macroeconomic forecasters as political agents and suggested that they exploit their information advantage over the future states of the economy to influence the policy-making process. First, we have theoretically analyzed a framework of asymmetric information between forecasters and voters approaching a referendum. Forecasters are better informed on the future state of the economy under each of the potential outcomes of a referendum. Voters care about the economy in the future, but they have to rely on scenarios published by professional forecasters to form beliefs before the referendum. Under the assumptions of our model, it is optimal for

[26] For instance, consider that *Barclays* has been the title sponsor of the soccer's Premier League between 2011 and 2016.

forecasters with stakes and influence to publish biased scenarios instead of their best estimate. Second, we have tested the predictions of the model using the EU membership referendum, also known as the Brexit referendum, held in the UK in 2016.

The results show that forecasters with stakes and influence released GDP growth forecasts in the case of a Leave outcome that was more pessimistic than the forecasts released by other institutions. These results confirm the theoretical predictions about a propaganda bias in macroeconomic forecasts released by institutions with stakes and influence. We also find that the propaganda bias is present at the intensive margin, which is consistent with the predictions from the model.

The predictions of biased forecasts in equilibrium differ from the lobbying models of campaign expenditures in electoral competitions (see, e.g., Baron, 1994), despite the very similar setup, because of the institutional nature of referenda. In the models of electoral competition, policy convergence implies that, in equilibrium, organized groups face no incentives to favor one candidate over another, as the two will implement the same platform after the vote. The policy outcome is affected by the presence of special interest groups, but the voting is not. Instead, in a referendum, policy outcomes are given ex-ante and are divergent. In equilibrium, forecasters may have a preference for one over the other.

The propaganda bias might impact the welfare of both voters and forecasters. In the case of the Brexit referendum, in which the Leave side won, the realization of shocks to preferences was determinant in generating the outcome that forecasters did not prefer. In that case, voters did not face any welfare loss compared to a world of unbiased forecasters, although the race was more uncertain because of the bias. On the contrary, forecasters with stakes and influence ended up paying a high accuracy cost due to the bias and facing the economic loss attached to Brexit. In addition to those presented in the model, the propaganda bias might generate additional welfare reductions because of general equilibrium effects if consumers and investors make consumption and investment decisions based on the forecasts. If forecasts are biased, then economic agents may make incorrect decisions that could, in turn, reduce GDP.

Our results contribute to the political economy literature by proposing economic forecasters as an additional player, and to the forecasting literature, by highlighting an additional strategic behavior underlying forecasts errors. According to our theoretical predictions and empirical results, macroeconomic forecasters may use their information advantage to influence the decision-making process and favor their most preferred outcome. We recommend that voters and policymakers take this into account when forming their beliefs to avoid systematic mistakes.

References

Alabrese, E., Becker, S. O., Fetzer, T., & Novy, D. (2019). Who voted for Brexit? Individual and regional data combined. *European Journal of Political Economy*, *56*, 132–150.

Andersson, M. K., Aranki, T., & Reslow, A. (2017). Adjusting for Information Content when Comparing Forecast Performance. *Journal of Forecasting*, *36*(7), 784–794.

Baron, D. P. (1994). Electoral Competition with Informed and Uninformed Voters. *American Political Science Review*, *88*(1), 33–47.

Becker, S. O., Fetzer, T., & Novy, D. (2017). Who Voted for Brexit? A Comprehensive District-level Analysis. *Economic Policy*, *32*(92), 601–650.

Bertelsmann Stiftung. (2015). Costs and benefits of a United Kingdom exit from the European Union.

Besley, T., & Coate, S. (2001). Lobbying and Welfare in a Representative Democracy. *The Review of Economic Studies*, *68*(1), 67–82.

BetData. (n.d.). UK EU Referendum 2016 [Data retrieved November 30, 2017 from https://betdata.io/historical-odds/uk-eu-referendum-2016].

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

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

Consensus Economics. (2016a). Consensus Forecasts: G7 & Western Europe 2016/06.

Consensus Economics. (2016b). Consensus Forecasts: G7 & Western Europe 2016/07.

Croushore, D. D. (1997). The Livingston Survey: Still Useful after all These Years. *Business Review-Federal Reserve Bank of Philadelphia*, *2*, 1–12.

Davies, A., & Lahiri, K. (1995). A new Framework for Analyzing Survey Forecasts using Three-dimensional Panel Data. *Journal of Econometrics*, *68*(1), 205–228.

Deb, R., Pai, M. M., & Said, M. (2018). Evaluating Strategic Forecasters. *American Economic Review*, *108*(10), 3057–3103.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., & Zhuravskaya, E. (2014). Cross-border Media and Nationalism: Evidence from Serbian Radio in Croatia. *American Economic Journal: Applied Economics*, *6*(3), 103–32.

Dhingra, S., Ottaviano, G., & Sampson, T. (2015). Should we Stay or Should we Go? The Economic Consequences of Leaving the EU. *London School of Economics and Political Science, CEP*.

Dhingra, S., Ottaviano, G., Sampson, T., & Van Reenen, J. (2016a). The Consequences of Brexit for UK Trade and Living Standards. *London School of Economics and Political Science, CEP*.

Dhingra, S., Ottaviano, G., Sampson, T., & Van Reenen, J. (2016b). The Impact of Brexit on Foreign Investment in the UK. *London School of Economics and Political Science, CEP*.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, *65*(2), 135–150.

Durante, R., Pinotti, P., & Tesei, A. (2019). The Political Legacy of Entertainment TV. *American Economic Review*, *109*(7), 2497–2530.

Ebell, M., & Warren, J. (2016). The Long-Term Economic Impact of Leaving the EU. *National Institute Economic Review*, *236*(1), 121–138.

FT Research. (n.d.). UK's EU referendum, Brexit Poll tracker [Data retrieved November 30, 2017 from https://ig.ft.com/sites/brexit-polling/].

Giles, C. (2019). Politics is Falling on Brexit but Economics Has Been on the Money. *FinancialTimes.com, March 14, 2019*.

Grossman, G. M., & Helpman, E. (1996). Electoral Competition and Special Interest Politics. *The Review of Economic Studies*, *63*(2), 265–286.

HM Treasury. (2016). HM Treasury Analysis: The Immediate Economic Impact of Leaving the EU.

HM Treasury. (n.d.). Forecasts for the UK Economy [Data retrieved from https://www.gov.uk/government/collections/data-forecasts].

Kierzenkowski, R., Pain, N., Rusticelli, E., & Zwart, S. (2016). The Economic Consequences of Brexit. *OECD Economic Policy Papers*, *16*.

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal on optimization*, *9*(1), 112–147.

Laster, D., Bennett, P., & Geoum, I. S. (1999). Rational Bias in Macroeconomic Forecasts. *The Quarterly Journal of Economics*, *114*(1), 293–318.

Lindbeck, A., & Weibull, J. W. (1987). Balanced-Budget Redistribution as the Outcome of Political Competition. *Public choice*, *52*(3), 273–297.

Mansfield, I. (2014). A Blueprint for Britain: Openness not Isolation.

Marinovic, I., Ottaviani, M., & Sørensen, P. N. (2013). Chapter 12 – Forecasters' Objectives and Strategies. In G. Elliott & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (pp. 690–72). Elsevier.

Martin, G. J., & Yurukoglu, A. (2017). Bias in Cable News: Persuasion and Polarization. *American Economic Review*, *107*(9), 2565–99.

Minford, P. (2016). Brexit and Trade: What are the options?

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

OECD. (2016). *The consequences of Brexit: a taxing decision* (OECD Economic Policy Paper No. 16).

Open Europe. (2015). What if...? The Consequences, challenges and opportunities facing Britain outside EU.

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

Oxford Economics. (2016). Assessing the Economic Implications of Brexit.

Qin, B., Strömberg, D., & Wu, Y. (2018). Media Bias in China. *American Economic Review*, *108*(9), 2442–76.

Ramiah, V., Pham, H. N., & Moosa, I. (2017). The Sectoral Effects of Brexit on the British Economy: Early Evidence from the Reaction of the Stock Market. *Applied Economics*, *49*(26), 2508–2514.

The Economist Intelligence Unit. (2016). "Brexit" will cost 6% of GDP.

The National Archives. (n.d.). Forecasts for the UK Economy [Data retrieved May 10, 2018 from https://www.nationalarchives.gov.uk/webarchive/].

Thomson Reuters Eikon. (n.d.). Daily Price History [Data retrieved June 21, 2018 from https://eikon.thomsonreuters.com/index.html].

Viskanic, M. (2020). *Fear and Loathing on the Campaign Trail: Did Immigration Cause Brexit?* (Sciences Po publications No. 100).

Appendix

A.1 Details on the Numerical Solution of the Extended Model

The extended version of the theoretical framework is solved numerically to take into account the recursivity in the forecasters' problem generated by the dynamics and by their strategic interaction, which also introduces a common pool problem among forecasters in the economy.

We calibrate the parameters of the extended model with the aim of being consistent with the main intuitions behind the theoretical framework and the application of the model predictions to the Brexit referendum. In Figure A3, we perform a number of comparative dynamics exercises to show how the predictions of the model vary as a function of exogenous parameters.

As shown in Table A1, the number of forecasters in the economy is set to 44 to reproduce the data collection *Forecasts for the UK economy*, collected and released by the HM Treasury on a monthly basis. Consequently, we calibrate the number of time periods T to 12, hence imposing that the forecasters in the model release a new estimate each month for a full year. We set $k = 6$, so that between January and May forecasters only release $F^R_{j,t}$, while they start publishing forecasts subject to alternative L from June onwards. The measure of stakes η_j is drawn at the individual forecaster level from a uniform distribution between 0 and 1, while τ_j, from which the measure of influence γ_j is generated, is drawn from a uniform distribution between 0 and 0.1.

y^L and y^R are respectively calibrated to 1.8 and 2.1. The former figure replicates the actual GDP growth observed in the UK in the year 2017, after the victory of Leave in the occasion of the Brexit referendum, while the latter replicates the average forecast for GDP growth in 2017 reported in the latest release before the referendum of *Forecasts for the UK economy*. We make this choice since we cannot observe ex-post what the GDP growth in 2017 would have been had the UK voters decided not to leave the European Union. We also calibrate voters' prior μ^A to be correct, so that $\mu^A = y^A \; \forall \, A \in \{L, R\}$.

The parameter α, which represents the relative weight of accuracy and consistency in the cost structure of forecasters, is calibrated to 0.5, while β is imposed equal to 0.99 to represent that the quality of the signal that forecasters send to voters does not change substantially in precision over time in the short run.

τ_L and ρ are set respectively to 0.1 and to 10^{-3} in order to represent that voters have a very precise prior subject to R, while their prior information about alternative L is of much lower precision.

Lastly, C is calibrated to 40 with the aim of calculating an expected probability of Brexit consistent with Consensus Economics (2016a) and the bookmakers' odds (see Figure A2 in the Appendix). It is worth to notice that the calibration of other parameters defines acceptable bounds for C.

Given the parameters, we solve the model by supplying an initial guess of forecasts for all forecasters and allowing all forecasters to update their fore-

casts to minimize the individual loss given the guess on the forecast of others.[27] We then update our guess using the resulting forecasts from the optimization problem performed by all individuals and allow all forecasters to re-optimize given the updated guess. We repeat this process until the forecasters no longer wish to change their forecasts, given the forecasts of others. Due to the randomness of the draws of η_j and τ_j, we always simulate multiple economies in order to find the model solution in expectations.

We assume an unbiased forecast as the initial guess in the baseline calibration. Nevertheless, we performed a robustness check in which we allowed for different initial guesses. More specifically, we perturbated the initial guess by adding a random number drawn from a normal distribution. The model was solved 100 times, and the obtained equilibrium solution is the same in all cases.

The graphs in Figure 1 are based on 500 simulated economies. In the baseline calibration, we obtain a p^L of 0.22 with voter posteriors that are $\widetilde{F}^L = 1.6$ and $\widetilde{F}^R = 2.1$. When we calibrate $C = 0$, such that all forecasters provide unbiased forecasts, p^L is equal to 0.31 while voters' posteriors are $\widetilde{F}^L = 1.8$ and $\widetilde{F}^R = 2.1$. Hence, the bias in the baseline calibration reduced p^L by 9 percentage points.

In the graphs in Figures 2 and A3, we report the mean of 100 simulated economies for each parameter case. In all graphs we keep the base calibration reported in Table A1 and only vary the parameter of interest. We fix the η_j and τ_j values for one forecaster to be $\eta_j = 0.95$ and $\tau_j = 0.1$. Hence, we show the forecasts for a forecaster with high levels of stake and influence. A τ_j of 0.1 corresponds to an expected γ_j of 0.04.

Figure A4 reports the forecasts of a forecaster with $\eta_j = 0.95$ and $\tau_j = 0.1$ in the baseline economy. The solid line reports the median forecast based on 100 simulated economies while the dotted lines report the 1^{st} and 99^{th} percentiles. This illustrates how the bias varies depending on the random draw of the other forecasters' individual parameters η_j and γ_j in the economy.

Table A1. *Calibration of parameters*

T	k	J	y^L	y^R	μ^L	μ^R	C	β	α	ρ	τ_L	η_j	τ_j
12	6	44	1.8	2.1	1.8	2.1	40	0.99	0.5	0.001	0.1	$\mathscr{U}[0,1]$	$\mathscr{U}[0,0.1]$

Notes: Calibration of the exogenous parameters introduced in the extended model presented in Section 2.2. See Section A.1 in the Appendix for details on how we solve the model numerically and a discussion of the parameters.

[27] We use the Nelder-Mead simplex algorithm as described in Lagarias et al. (1998) in order to minimize the loss function.

A.2 Data

Google News, Google Trends, and Influence by Newspaper

Google Trends allows us to retrieve, for each institution, a measure of the number of Google searches from the public, relative to the most searched forecaster. We restrict to the UK and only consider searches in the 2015 calendar year (prior to the announcement of the referendum). After downloading weekly data in which the most searched institution scores 100, we aggregate on a yearly level and assign the binary measure of influence based on a threshold set at the median value, so that half of the forecasters take the value 1 and the other half take the value 0 (see Figure A7a). The same variable is also used for the analysis of the intensive margin.

The number of search results on Google News gives an indication of how influential an institution is according to the media. If frequently mentioned in the news, then the institution is more influential than if very rarely mentioned. We record the number of mentions in a Google News search in the United Kingdom during 2015. Then, the binary measure of influence is constructed based on a threshold of 7,000 citations, so that half of the forecasters are above and half are below (see Figure A7b).

We compute measures of influence by different media outlets, which vary in terms of the political and referendum preferences of their readers by using the Dow Jones Factiva newspaper search engine. For each forecaster, we search for the name associated with the words "GDP" and "forecast," and export the number of articles reporting this information during 2015. We then construct an indicator equal to 1 if the forecaster has been mentioned by each specific media more than the median forecaster, and zero otherwise.

In Table A10, we consider alternative measures from Google News that corroborate the robustness of our findings. First, we span several time periods using the Archive settings of the search engine in order not to restrict the research to only the calendar year 2015.[28] Second, we restrict the news search such that we only consider media releases in which the name of each forecaster is accompanied either by the word "GDP" or by the words "GDP" and "forecast." This allows us to separate the cases in which an institution is reported in the media because of its forecasting activity from the cases in which the institution is mentioned for other reasons. In all cases, the binary measure of influence is constructed based on thresholds dividing the sample into two groups of equal size.

[28]The Google News search for Archive settings enlarges the research from the oldest article that the search engine is able to retrieve up to the date of the current research. The last access to this research was on December 7, 2017.

Banks, City and Stock Price

We determine whether a forecaster is a financial institution by referring to each forecaster's official web page and relying on how the institution describes herself. We label those which best can be described as a financial institution as Banks. We also assert that all the institutions labeled as Banks are quoted on international financial markets. We also propose an alternative measure of stakes based on the geographical location of each forecaster. Specifically, we make use of the group assignment to City or Non-City made by HM Treasury in its data collection *Forecast for the UK Economy* under the assumption that forecasters located in the City of London's financial district have higher stakes than the others.

For the investigation of the intensive margin, we have computed for each institution the percentage decline in the stock market price after the referendum. Specifically, between the referendum date (since both the London and the New York stock markets closed before the announcement of the referendum results) and the second banking day after the referendum results (see Figure 6). We make this choice based on the stylized fact that the decline in market prices has been continuous not only on the very first day after the vote (a Friday) but also on the subsequent Monday. The data source for this analysis is Thomson Reuters Eikon (n.d.).

A.3 Figures and Tables

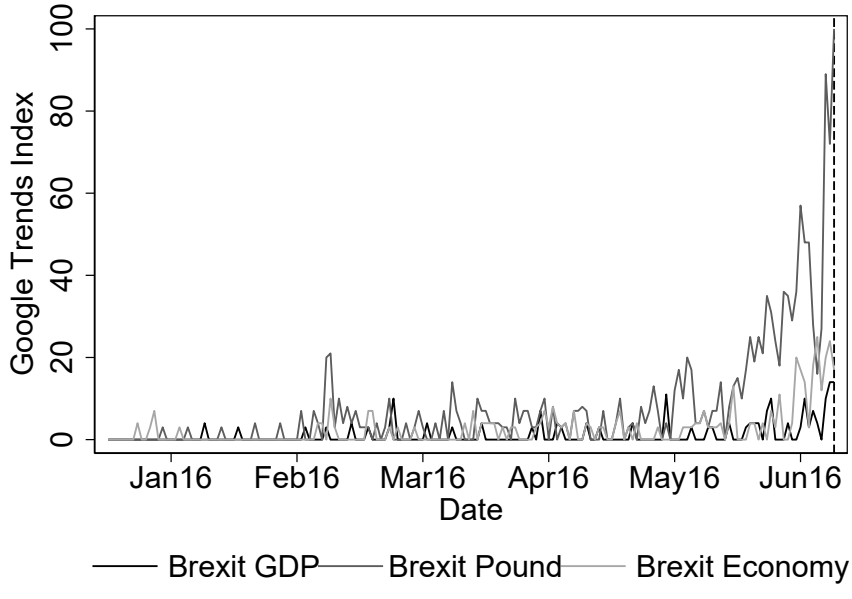
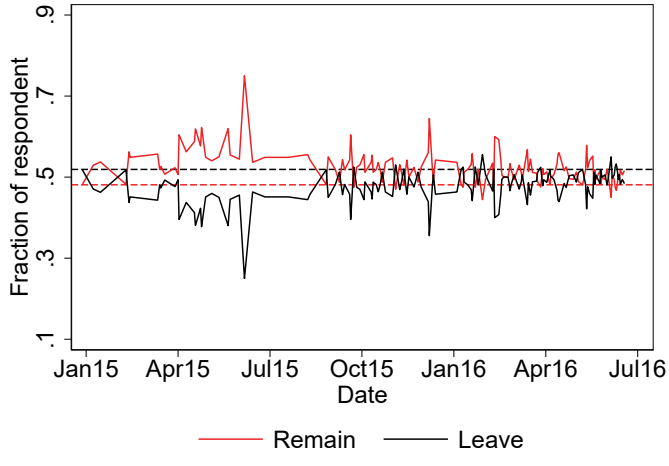


Figure A1. Brexit and the Economy Approaching the Referendum

Notes: The figure shows the Google Trends summary reports for the search entries "brexit GDP," "brexit pound" and "brexit economy" on a daily basis before the referendum. Source: Authors' elaboration on data from Google Trends.

(a) Opinion Polls



(b) Bookmakers' Odds

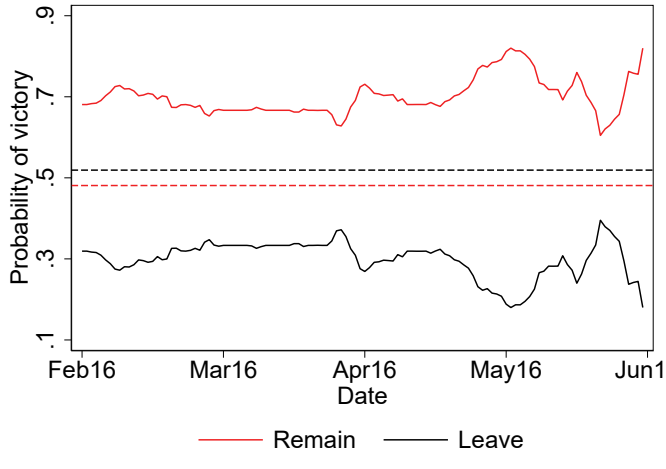


Figure A2. Opinion Polls and Bookmakers' Odds Approaching the Referendum

Notes: Panel (a) reports the daily averages of all opinion polls recorded by the Financial Times between January 2015, before the official announcement of the referendum, and June 22, 2016. Source: Authors' elaboration on data from the FT Research, n.d. Panel (b) reports the daily average of the odds released by all bookmakers recorded by the portal Betdata.io from the announcement of the referendum date until June 22, 2016. Source: Authors' elaboration on data from BetData, n.d. In both panels, dashed lines represent the referendum result.

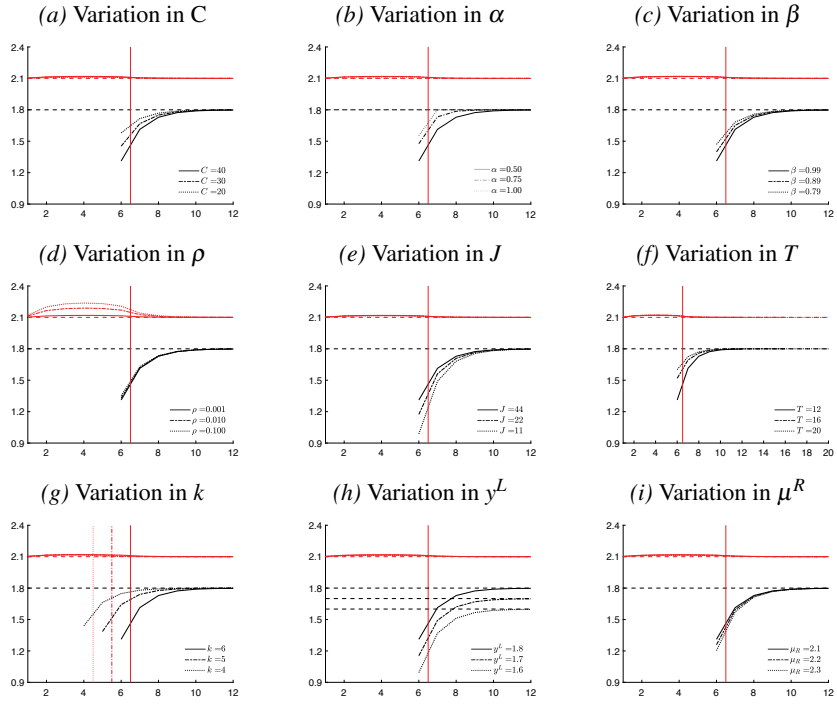


Figure A3. Comparative Dynamics with respect to the Exogenous Parameters

Notes: The figure reports the evolution of $F^L_{j,t}$ and $F^R_{j,t}$ over time. Dashed lines represent y^L and y^R, while solid lines as well as dotted and dashed-dotted represent forecasts. Black lines represent variables referring to state L, while red lines represent variables referring to state R. In each graph, one parameter at the time is allowed to vary according to the values reported in the legends. τ_j is held constant at 0.1, so that γ_j is on average equal to 0.04, and η_j is held constant at 0.95. All other parameters have been calibrated as described in Section A.1.

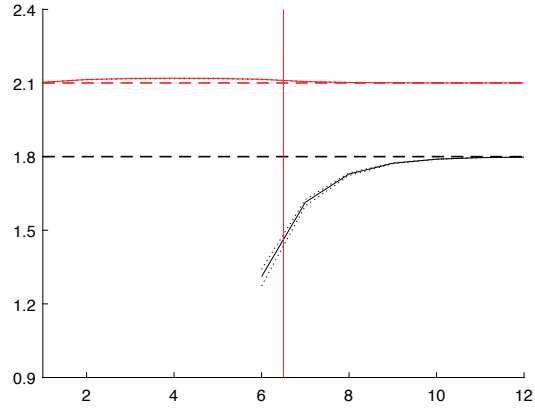


Figure A4. Baseline Calibration

Notes: The figure reports the evolution of $F^L_{j,t}$ and $F^R_{j,t}$ over time. Dashed lines represent y^L and y^R, while solid line report the median forecast based on 100 simulated economies while the dotted lines report the 1^{st} and 99^{th} percentiles. Black lines represent variables referring to state L, while red lines represent variables referring to state R. τ_j is held constant at 0.1, so that γ_j is on average equal to 0.04, and η_j is held constant at 0.95. All other parameters have been calibrated as described in Section A.1.

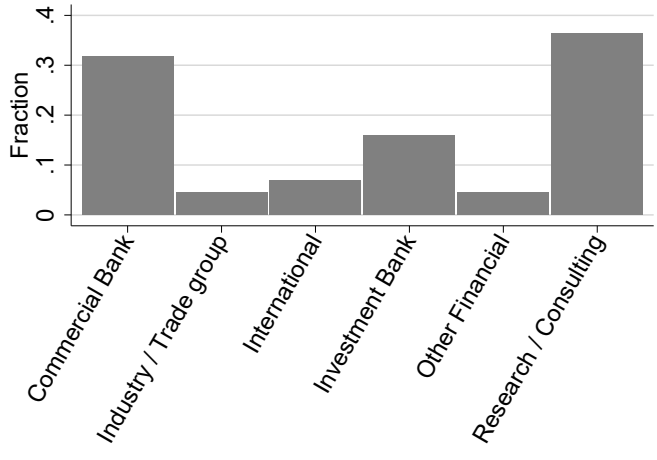


Figure A5. Distribution of Sampled Forecasters by Industry

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. The figure reports the proportion of forecasters in the sample belonging to each of the industries specified in the horizontal axis.

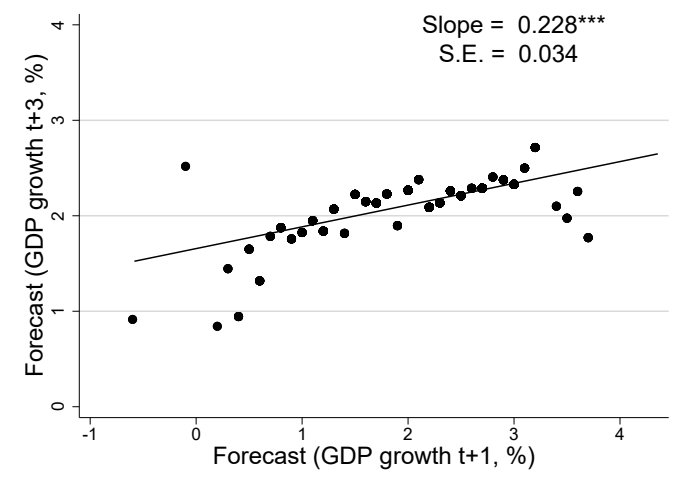
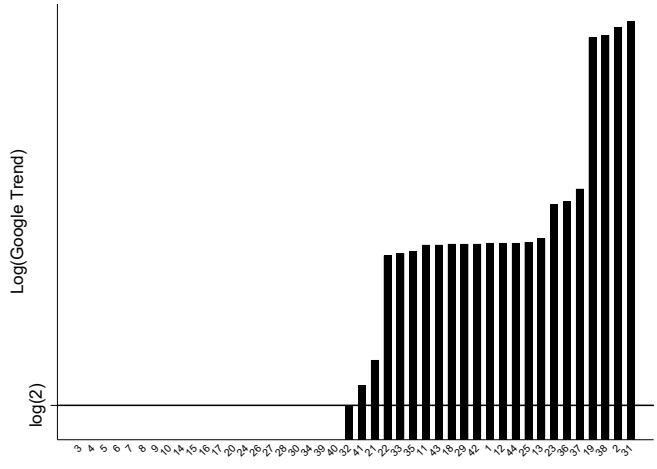


Figure A6. Correlation between Short and Medium-term Forecasts

Notes: All forecasters surveyed by HM Treasury between January 2012 and June 2016 which reported within the same release both short-term and medium-term forecasts. The figure reports the bivariate regression $F_{j,t}^{t+3} = \alpha + F_{j,t}^{t+1} + \varepsilon_{j,t}$. Heteroskedasticity-robust standard errors are reported. *,**,*** represent the 10%, 5%, 1% significance levels. Markers represent the sample average of $F_{j,t}^{t+3}$ within bins of 0.1 percentage points of $F_{j,t}^{t+1}$.

(a) Google Trends



(b) Google News

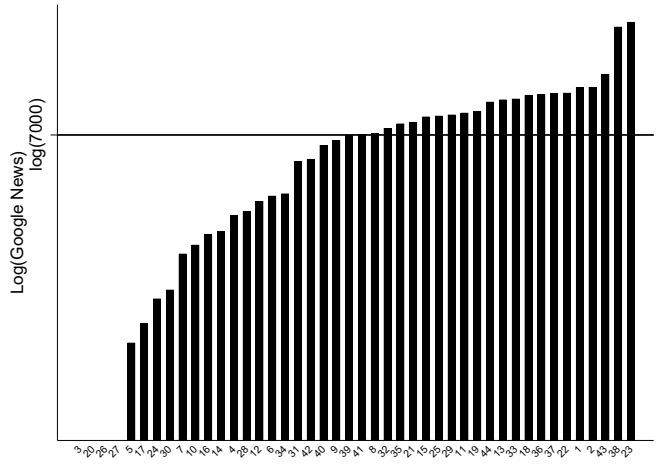
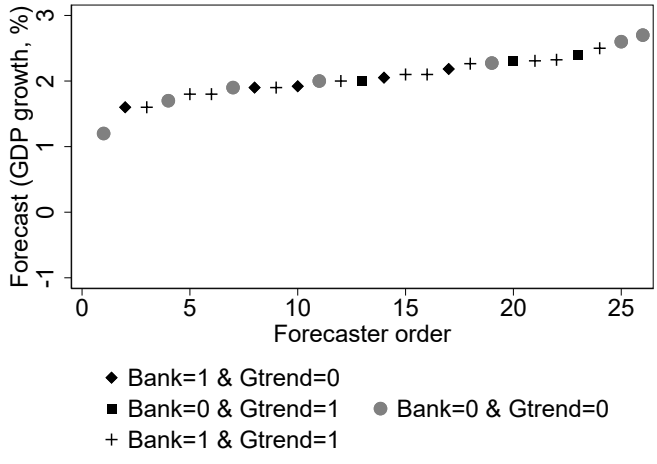


Figure A7. Google Measures

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. The horizontal axis shows a forecaster ID. Panel (a) plots the number of searches (in logarithms) that the general public has done for each institution according to Google Trends. Panel (b) plots the number of citations (in logarithms) that each institution has reported in Google News. In both panels, the horizontal line represents the threshold used to assign binary measures of influence used in the extensive margin analysis.

(a) June 2016



(b) July 2016

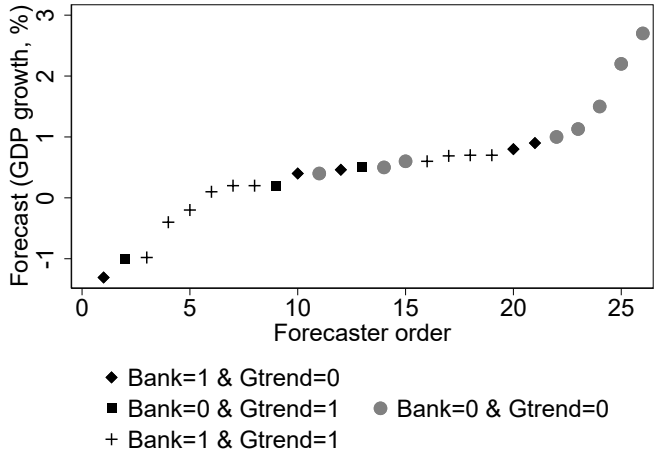


Figure A8. Forecast for the 2017 GDP Growth Rate Before and After the Referendum

Notes: All forecasters surveyed by HM Treasury in June and July 2016. The relevant measure of stakes is to be a financial institution (Banks). The relevant measure of influence is Google Trends (see Section 3.4 for details). Each marker represents an individual GDP growth forecast for period $t+1$. Black non-circle markers represent forecasters with stakes or influence, while gray circle markers represent the control institutions.

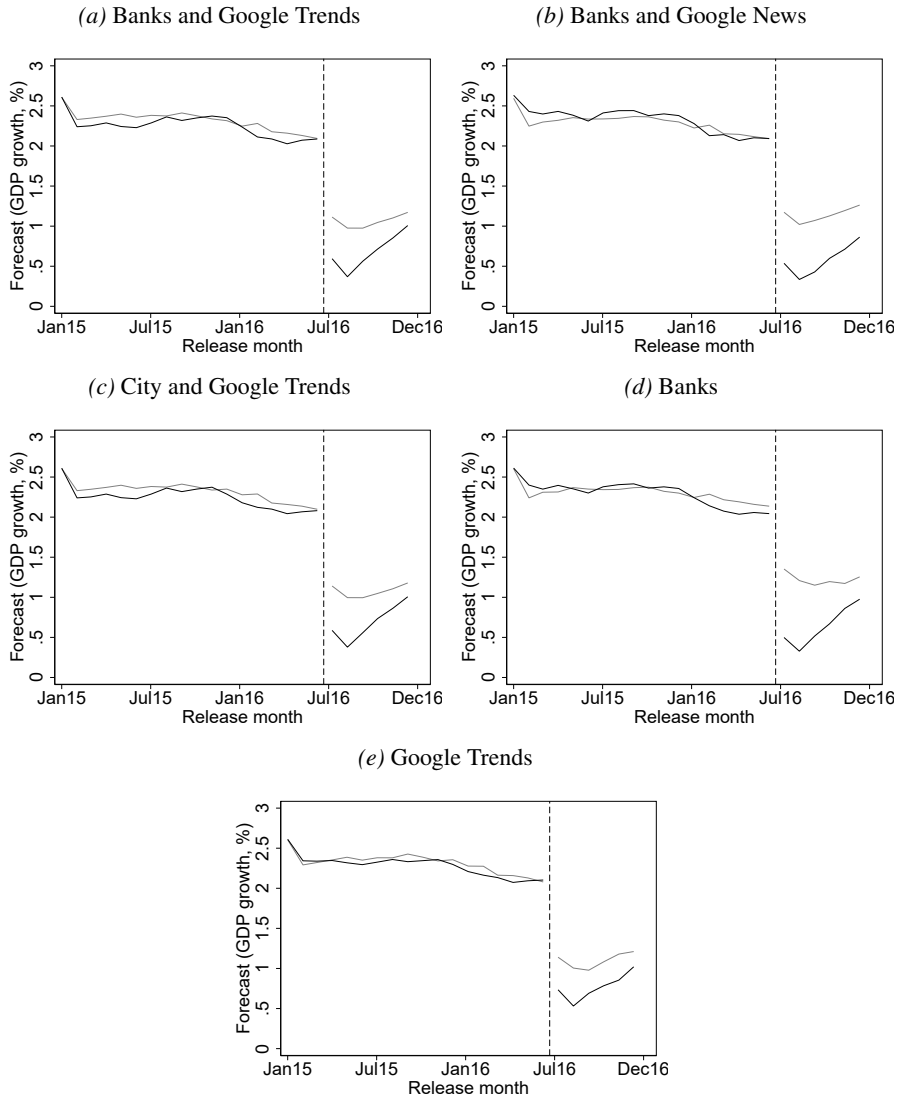


Figure A9. Pre-Referendum Trends for all Measures of Stakes and Influence

Notes: Each graph plots the average GDP growth forecast for period $t+1$ among institutions belonging to each of the different groups under investigation between January 2015 and December 2016. Specifically, black lines represent the group under investigation as defined is Section 3.5, while gray lines represent the remaining control institutions.

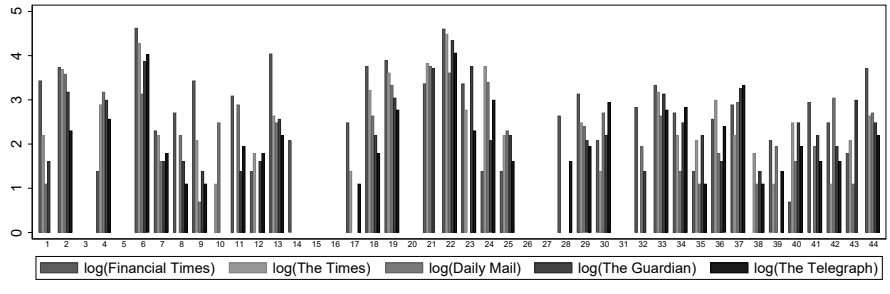


Figure A10. Influence by Newspaper

Notes: The graph reports the natural logarithm of the number of articles mentioning the forecasts for GDP growth released by each forecaster in different daily newspapers during 2015.

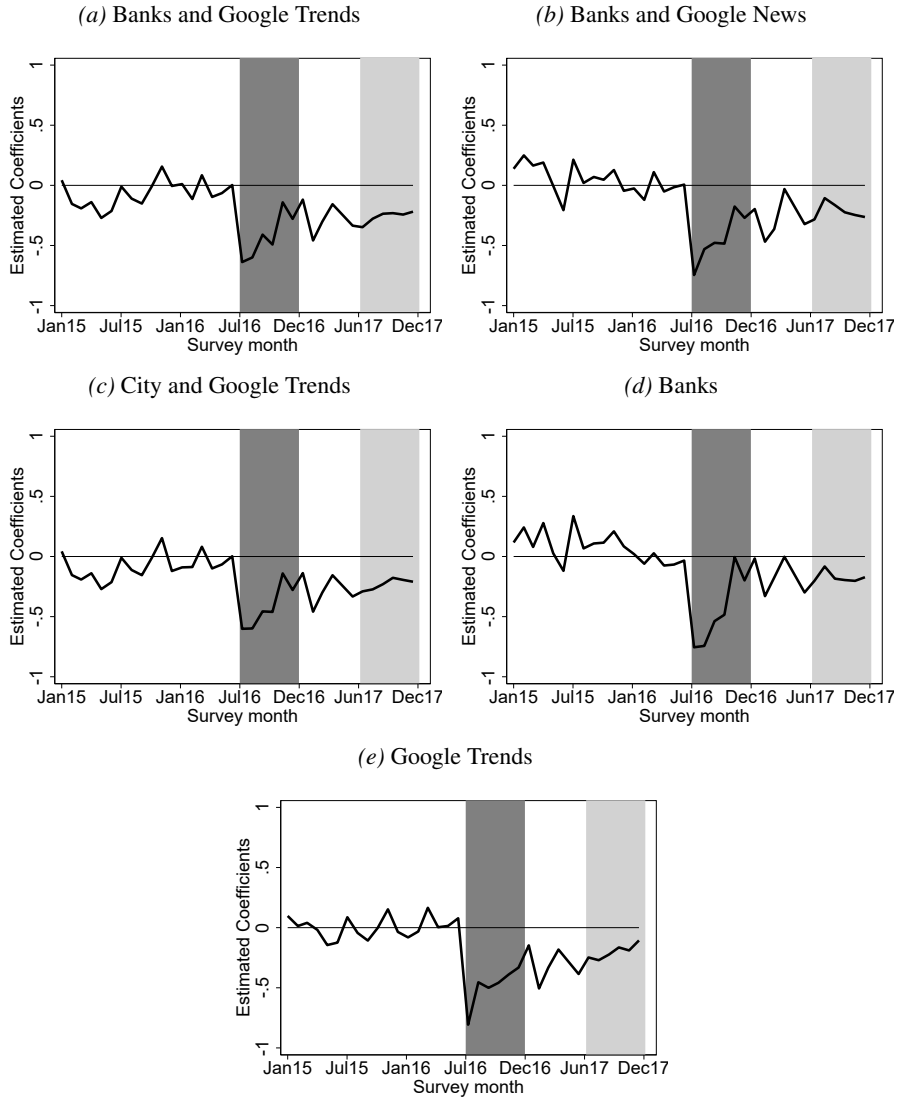
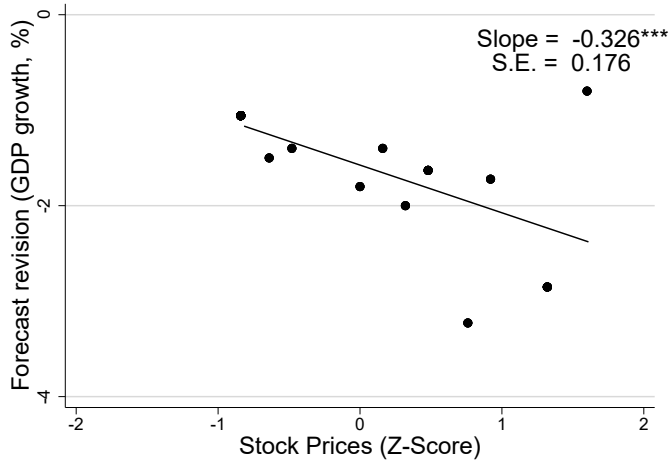


Figure A11. Estimated Propaganda Bias at Different Points in Time

Notes: In each graph, we replicate the results in Table 1, columns (2)–(6) by assuming a placebo referendum at every month between January 2015 and April 2018. All specifications include forecasters' fixed effects and survey fixed effects. The dependent variable is GDP growth rate in period $t+1$. The line represents the estimated coefficient for β_0. The dark gray area represents the five surveys included in the results presented in Table 1, while the light gray area represents result estimated for surveys released during the same month of the following year.

(a) Stock Prices



(b) Google News

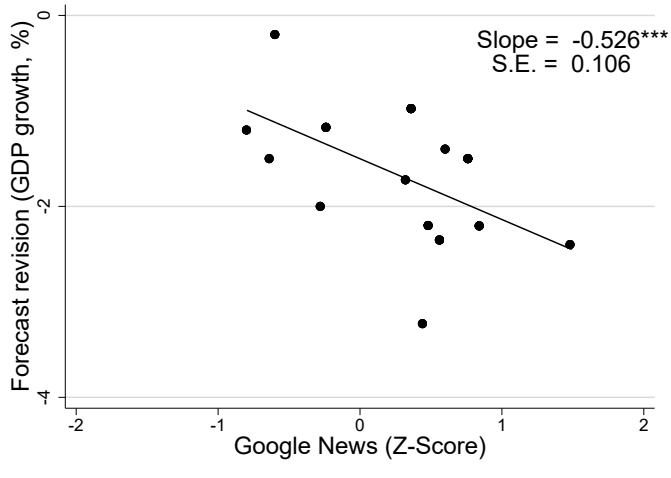
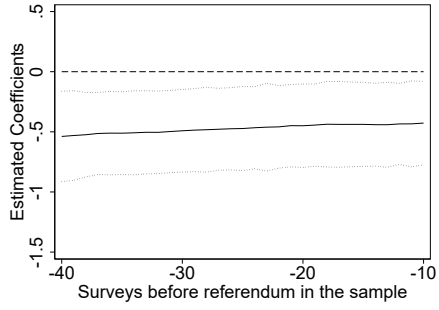


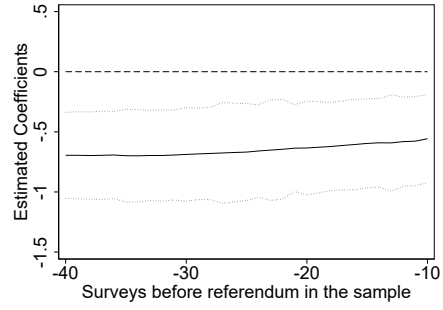
Figure A12. $F^L_{j,k} - F^R_{j,k}$ as a Function of Stakes and Influence

Notes: Panel (a) reports the bivariate regression $F^L_{j,k} - F^R_{j,k} = \alpha + \beta StockPrice_j + u_{j,k}$. Panel (b) reports the bivariate regression $F^L_{j,k} - F^R_{j,k} = \gamma + \delta GoogleNews_j + v_{j,k}$. In both panels, $F^L_{j,k}$ represents the forecast published in July 2016 by each institution and $F^R_{j,k}$ the forecast published in June 2016. Heteroskedasticity-robust standard errors are reported. *,**,*** represent the 10%, 5%, 1% significance levels. Markers represent the sample average of $F^L_{j,k} - F^R_{j,k}$ within bins of 0.04 standard deviation units of *StockPrice* and *GoogleNews*.

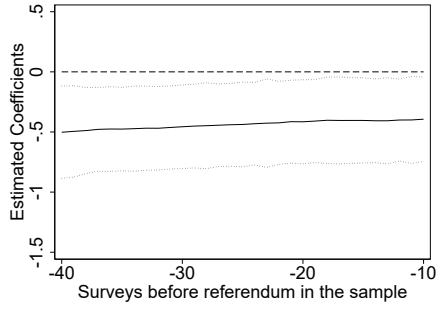
(a) Banks and Google Trends



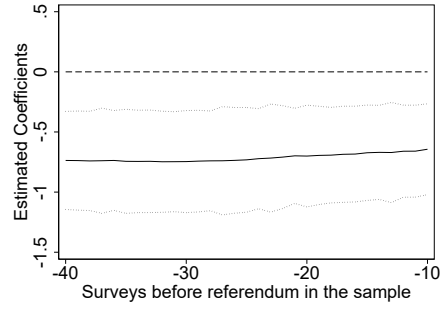
(b) Banks and Google News



(c) City and Google Trends



(d) Banks



(e) Google Trends

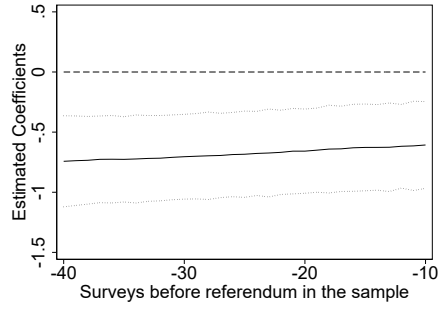
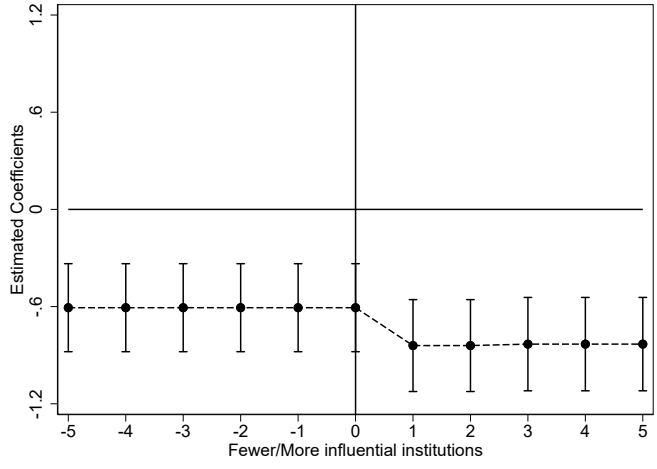


Figure A13. Sensitivity to Changes in the Time Span

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. The dependent variable is GDP growth rate in period $t+1$. In each graph, we replicate the results in Table 1, columns (2)–(6) by estimating with the support of sample that spans a different number of months. The black solid line represents estimated coefficients, while dotted lines represent the 95% confidence intervals. All specifications include forecasters' fixed effects and survey fixed effects.

(a) Google Trend



(b) Google News

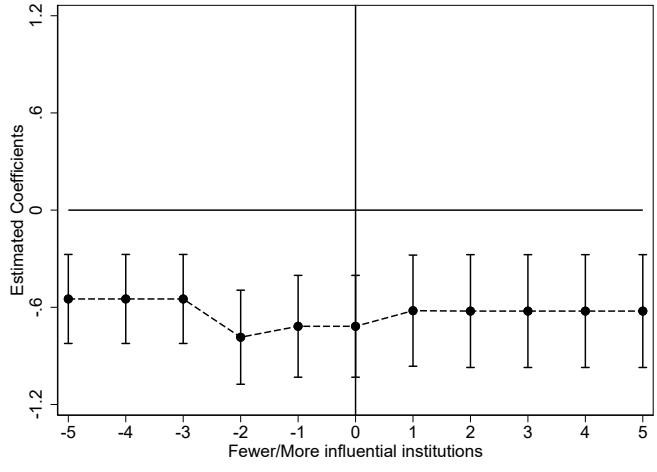


Figure A14. Sensitivity to the Definitions of Influence

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. The dependent variable is GDP growth rate in period $t+1$. In each panel, we estimate (15) by including in or excluding from the group of institutions with influence up to five forecasters before interacting with the stakes measure (Banks). All specifications include forecasters' fixed effects and survey fixed effects. Standard errors are robust to two-way clustering at the forecaster and the survey levels. For all regressions, graphs report estimated coefficients and 95% confidence intervals.

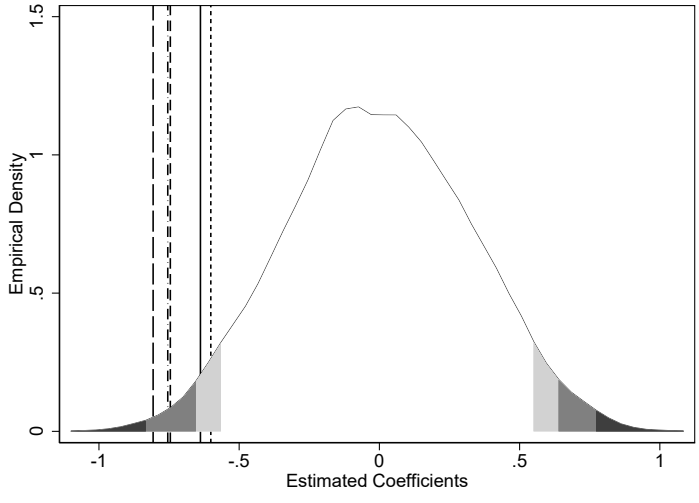


Figure A15. Monte Carlo Simulation: Random Assignment to Groups

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. The dependent variable is GDP growth rate in period $t+1$. We perform a Monte-Carlo simulation with 10,000 draws, in each of which we randomly assign half of the institutions to one group and the other half to another group. At each draw, we then estimate equation (15). All specifications include forecasters' fixed effects and survey fixed effects. We then plot the empirical density of estimated coefficients. Vertical lines represent the estimated coefficient obtained in columns (2)–(6) of Table 1. The areas shaded in gray show the 1%, 5% and 10% tails of the distribution.

Table A2. *Timeline of the United Kingdom European Union Membership Referendum*

Jan. 22, 2013	Prime Minister David Cameron announced that a referendum on EU membership would be held before the end of 2017, on a renegotiated package, if elected in 2015
May 22, 2014	The UK Independence Party (UKIP) gets 26 percent of the vote in European elections and becomes the largest UK party in the European Parliament
May 7, 2015	The Conservative Party won the majority in 2015 general elections
May 27, 2015	The European Union Referendum Act 2015 (c. 36) was unveiled in the Queen's Speech
Dec. 17, 2015	The Act is given Royal Assent
Jan. 5, 2016	PM Cameron says ministers are free to campaign on either side
Feb. 20, 2016	PM Cameron announced the referendum date (23 June 2016)
Apr. 15, 2016	Start of the official campaign period
May 6, 2016	Start of British Election Study panel (Wave 8)
May 23, 2016	Release of HM Treasury analysis: the immediate economic impact of leaving the EU
June 22, 2016	End of British Election Study panel (Wave 8)
June 23, 2016	The United Kingdom European Union membership referendum
June 24, 2016	PM Cameron announces resignation after vote for Brexit
July 9, 2016	A petition calling for a second referendum was rejected by the Government
July 11, 2016	Theresa May formally declared leader of the Conservative Party
July 13, 2016	Theresa May appointed Prime Minister by Queen Elizabeth II
Jan. 24, 2017	The Supreme Court: the Government needs parliamentary approval to trigger Article 50
Mar. 29, 2017	Prime Minister Theresa May triggers Article 50, which starts the clock on the process of the UK leaving the EU
Jan. 31, 2020	The United Kingdom left the European Union

Notes: This table reports the key dates of the UK membership referendum, before and after the vote. Source: Authors' elaboration on information from https://www.bbc.com/news/politics.

Table A3. *Forecasts Before Referendum Announcement, by Forecaster Type*

(a) *Commercial Bank*

Variable	Mean	St. Dev.	Obs.
GDP	2.038	0.730	350
Private consumption	1.921	0.819	350
Fixed investment	4.874	3.227	350
Government consumption	-0.368	0.966	348
Total exports	4.229	1.739	339
Total imports	3.727	1.905	337

(b) *Industry / Trade group*

GDP	2.242	0.543	42
Private consumption	2.176	0.577	42
Fixed investment	4.992	1.522	42
Government consumption	0.123	0.885	42
Total exports	3.978	1.230	42
Total imports	3.421	1.073	42

(c) *International*

GDP	2.037	0.569	14
Private consumption	1.882	0.519	7
Fixed investment	9.350	5.445	12
Government consumption	-0.970	1.109	7
Total exports	3.488	0.902	14
Total imports	2.831	0.856	14

(d) *Investment Bank*

GDP	1.862	0.756	163
Private consumption	1.833	0.895	162
Fixed investment	4.005	2.227	162
Government consumption	0.043	1.079	162
Total exports	3.351	1.364	161
Total imports	3.232	1.403	161

(e) *Other Financial*

GDP	2.001	0.678	93
Private consumption	2.080	0.715	93
Fixed investment	5.072	1.833	93
Government consumption	-0.667	0.832	93
Total exports	4.105	1.631	93
Total imports	3.819	1.493	93

(f) *Research / Consulting*

GDP	1.945	0.685	364
Private consumption	1.850	0.882	364
Fixed investment	4.760	2.334	363
Government consumption	-0.055	1.140	362
Total exports	3.264	2.427	320
Total imports	3.014	2.591	320

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. All variables represent yearly growth rates (%) and refer to year $t+1$.

Table A4. *Descriptive Statistics*

Variable	Mean	St. Dev.	Obs.
GDP	1.792	0.721	1,643
Private consumption	1.724	0.853	1,620
Fixed investment	3.429	3.225	1,626
Government consumption	0.208	1.087	1,618
Total exports	3.556	1.785	1,520
Total imports	2.984	2.044	1,518

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. All variables represent yearly growth rates (%) and refer to year $t+1$.

Table A5. *Aggregate Views around the Referendum*

	June 2016	July 2016
Variable	Mean (St. Dev.)	Mean (St. Dev.)
GDP	2.092 (0.339)	0.926 (1.041)
Private consumption	2.187 (0.399)	1.004 (1.370)
Fixed investment	4.234 (1.396)	-1.288 (4.608)
Government consumption	0.705 (0.706)	0.830 (0.758)
Total exports	3.436 (1.619)	2.808 (1.685)
Total imports	3.269 (1.503)	1.278 (2.610)

Notes: All forecasters surveyed by HM Treasury in June and July 2016. All variables represent yearly growth rates (%) and refer to year $t+1$.

Table A6. *Correlation Matrix of the Assignment to Groups*

	Banks	City	GTrends	GNews	Log(Trend)	Log(News)	Stock Price
Banks	1.00	0.82	0.45	0.36	0.50	0.55	0.79
City	0.82	1.00	0.46	0.37	0.49	0.58	0.71
GTrends	0.45	0.46	1.00	0.73	0.86	0.68	0.43
GNews	0.36	0.37	0.73	1.00	0.59	0.74	0.46
Log(Trend)	0.50	0.49	0.86	0.59	1.00	0.62	0.43
Log(News)	0.55	0.58	0.68	0.74	0.62	1.00	0.53
Stock Price	0.79	0.71	0.43	0.46	0.43	0.53	1.00

Notes: Correlation between the groups described in Section 3.4. All forecasters surveyed by HM Treasury between January 2012 and April 2018. Banks is an indicator taking the value 1 if the institution self-reports itself as a financial institution on the official website, and 0 otherwise. City is an indicator taking the value 1 if the institution is located in the City of London according to HM Treasury information, and 0 otherwise. Google Trend is an indicator taking the value 1 if the institution has a score above the threshold value reported in Figure A7a and 0 otherwise. GNews is an indicator taking the value 1 if the institution has a score above the threshold value reported in Figure A7b, and 0 otherwise. Log(Trend) and Log(News) are the continuous measures of influence associated with GTrends and GNews. Stock Prices is a continuous variable representing the drop in capitalization of each company between the referendum day and two days after.

Table A7. *Correlation Matrix: Influence by Newspaper*

	Financial Times	The Times	The Daily Mail	The Guardian	The Telegraph
Financial Times	1.00	0.48	0.49	0.44	0.39
The Times	0.48	1.00	0.35	0.67	0.72
The Daily Mail	0.49	0.35	1.00	0.49	0.35
The Guardian	0.44	0.67	0.49	1.00	0.58
The Telegraph	0.39	0.72	0.35	0.58	1.00

Notes: Correlation between number of times in which the forecast released by each forecaster has been reported in each of the considered newspapers in 2015. All forecasters surveyed by HM Treasury between January 2012 and April 2018.

Table A8. *Reaction to Negative Economic Events*

	Stakes			
	Financial Crisis		Terrorist Attacks	
	(1)	(2)	(3)	(4)
Group × Event	-0.127	-0.092	-0.029	-0.009
	(0.189)	(0.194)	(0.097)	(0.106)
Group × Event (+1)	-0.173	-0.179	-0.150	-0.105
	(0.155)	(0.154)	(0.098)	(0.098)
Group × Event (+2)	-0.530***	-0.456**	-0.025	-0.031
	(0.204)	(0.213)	(0.095)	(0.100)
Group × Event (+3)	-0.431*	-0.320	-0.021	-0.060
	(0.228)	(0.237)	(0.089)	(0.095)
Group × Event (-1)	-0.104	-0.031	0.042	0.054
	(0.191)	(0.201)	(0.114)	(0.124)
Group × Event (-2)	-0.375*	-0.390*	0.224**	0.270**
	(0.198)	(0.204)	(0.111)	(0.109)
Group × Event (-3)	-0.204	-0.308	-0.024	-0.082
	(0.191)	(0.193)	(0.077)	(0.081)
Group × Event (-4)	-0.108	-0.114	0.122	-0.006
	(0.175)	(0.178)	(0.090)	(0.088)
Group × Event (-5)	-0.034	-0.096	0.026	0.025
	(0.163)	(0.174)	(0.101)	(0.103)
Observations	1,954	1,954	1,632	1,632
R^2	0.885	0.885	0.717	0.717
Fixed Effects	✓	✓	✓	✓
Survey Effects	✓	✓	✓	✓
Measure of Stakes	Banks	City	Banks	City
Measure of Influence				

Notes: Columns (1) and (2): All forecasters surveyed by HM Treasury between January 2004 and December 2010. Columns (3) and (4): All forecasters surveyed by HM Treasury between January 1998 and December 2003. The dependent variables is the GDP growth rate in period $t+1$. For each column, the column title defines the relevant group assignment. All specifications include forecaster fixed effects and survey fixed effects. The estimated equation is (15), assuming that everyone is influential. Columns (1) and (2): $k = 0$ on the occasion of the first survey after the bankruptcy of the Lehman Brothers Holdings Inc. on September 15, 2008. Columns (3) and (4): $k = 0$ on the occasion of the first survey after the attack to the World Trade Center in New York on September 11, 2001. Standard errors robust to two-way clustering at the forecaster and the survey levels are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table A9. *Propaganda Bias at the Intensive Margin: Influence by Google News*

	Stakes × Influence			Stakes	Influence
	(1)	(2)	(3)	(4)	(5)
Group × Ref. × Stock Price	-0.431***		-0.319***	-0.330***	
	(0.096)		(0.107)	(0.098)	
Group × Ref. × log(News)		-0.602***	-0.253		-0.833***
		(0.152)	(0.155)		(0.146)
Observations	1,643	1,643	1,643	1,643	1,643
R^2	0.771	0.770	0.771	0.770	0.772
Fixed Effects	✓	✓	✓	✓	✓
Survey Effects	✓	✓	✓	✓	✓
Measure of Stakes	Banks	Banks	Banks	Banks	
Measure of Influence	GNews	GNews	GNews		GNews

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. The dependent variable is GDP growth rate in period $t+1$. For each column, the column title defines the relevant group assignment. Stock Price and log(News) have been standardized to have mean 0 and variance 1. In all specifications, continuous measures of stakes and influence have been standardized to have zero mean and unit variance. All specifications include forecaster fixed effects and survey fixed effects. The estimated equation is (16). Standard errors robust to two-way clustering at the forecaster and the survey levels are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table A10. *Estimation of Propaganda Bias: Alternative Google News Measures*

	Stakes × Influence			
	(1)	(2)	(3)	(4)
Group × Referendum	-0.745***	-0.413**	-0.745***	-0.753***
	(0.185)	(0.193)	(0.188)	(0.215)
Group × Ref. (+1)	-0.529***	-0.654***	-0.656***	-0.209
	(0.177)	(0.174)	(0.177)	(0.206)
Group × Ref. (+2)	-0.471***	-0.511***	-0.514***	-0.401***
	(0.148)	(0.143)	(0.144)	(0.151)
Group × Ref. (+3)	-0.473***	-0.484***	-0.473***	-0.388**
	(0.154)	(0.149)	(0.155)	(0.156)
Group × Ref. (+4)	-0.157	-0.064	-0.157	-0.409***
	(0.120)	(0.129)	(0.122)	(0.129)
Group × Ref. (-1)	0.005	-0.010	0.044	0.071
	(0.093)	(0.098)	(0.096)	(0.093)
Group × Ref. (-2)	-0.026	-0.137	-0.026	0.090
	(0.098)	(0.098)	(0.098)	(0.097)
Group × Ref. (-3)	-0.067	-0.117	-0.017	0.074
	(0.090)	(0.090)	(0.092)	(0.088)
Group × Ref. (-4)	0.081	-0.008	0.081	0.175*
	(0.097)	(0.112)	(0.098)	(0.099)
Group × Ref. (-5)	-0.168	-0.112	-0.114	-0.080
	(0.110)	(0.114)	(0.116)	(0.111)
Observations	1,643	1,643	1,643	1,643
R^2	0.776	0.774	0.777	0.775
Fixed Effects	✓	✓	✓	✓
Survey Effects	✓	✓	✓	✓
Measure of Stakes	Banks	Banks	Banks	Banks
Measure of Influence	GNews	Gnews Archive	Gnews + GDP	GNews + GDP Forecast

Notes: All forecasters surveyed by HM Treasury between January 2012 and April 2018. The dependent variable is GDP growth rate in period $t+1$. The relevant group assignment is Stakes x Influence. In Column (1), the measure of influence is the Google News measure as described in Section 3.4. In Column (2), we extend the search to Archive settings. In Column (3), we require the media release to contain the word "GDP", while in column (4), we require the media release to contain both the words "GDP" and "forecast". All specifications include forecaster fixed effects and survey fixed effects. The estimated equation is (15). Standard errors robust to two-way clustering at the forecaster and the survey levels are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Essay II

Electoral Cycles in Macroeconomic Forecasts

with Davide Cipullo

Acknowledgments: We would like to thank Eva Mörk and Mikael Carlsson for extensive feedback and Enghin Atalay, Alessandra Bonfiglioli, David Cesarini, Sirus Dehdari, Mikael Elinder, Peter Fredriksson, Nils Gottfries, Georg Graetz, Christoph Hedtrich, Sebastian Jävervall, Jesper Lindé, Torben Mideksa, Luca Repetto, Olli Ropponen, Uta Schönberg, Anna Thoresson, Janne Tukiainen, Erik Öberg, and the participants at seminars held at Uppsala University, Uppsala Centre for Fiscal Studies (UCFS), the 76^{th} Annual Congress of the International Institute of Public Finance (IIPF), and the 35^{th} Annual Congress of the European Economic Association (EEA) for their dialogues and comments. Financial support from *Handelsbankens Forskningsstiftelser* is gratefully acknowledged.

1 Introduction

Elections empower voters to select the most preferred politician and ensure the accountability of elected officials. However, the empowerment is limited when voters lack sufficient knowledge regarding the traits of competing candidates. Often, voters use observable signals to infer crucial information about candidates before the election. For example, voters can use fiscal-policy outcomes and economic-performance indicators to learn about the incumbent's ability to handle the economy. Since voters lack sufficient information, the incumbent candidate can attempt to raise the reelection probability by adjusting the tax-collection and public-spending composition before an election. This is the well-known *Political Budget Cycle*.[1] When a political budget cycle is rampant and voters fail to consider its repercussions, elections fail to deliver an effective selection and good accountability of politicians.

This paper documents the existence of *Political Forecast Cycles*. Policy outcomes and economic realizations help voters to learn about the incumbent government. However, elections often occur before the complete realization of policy decisions' effects on public goods or economic growth. Therefore, to thoroughly evaluate the incumbent candidate, voters need to form expectations about the future effects of policies. Forming such expectations is challenging, and voters need to seek readily available information.[2] Macroeconomic forecasts published by governments are publicly available and reported in mass media. Voters may use these forecasts when forming beliefs about the economy and the incumbent politician's ability. When voters interpret economic performance as a signal of a leader's ability, the incumbent leader can gain an electoral advantage by intentionally manipulating the forecasts released to the public before an election.[3] Hence, on top of strategically using fiscal policy, the incumbent government may also strategically use the policy-outcome forecasts to gain an electoral advantage.[4]

We provide a theoretical rationale and empirical evidence of electoral incentives in macroeconomic forecasts released by the government. We show, in a theoretical model of political selection, that governments release overly

[1] The early literature by Nordhaus (1975) and MacRae (1977) on political business cycles provided models where politicians exploit the so-called Phillips curve by inflating the economy during election years to reduce unemployment. Rogoff and Sibert (1988) and Rogoff (1990) expanded the literature to variables such as taxes, government spending, and deficits, resulting in the vast literature on political budget cycles.

[2] Voters do not invest in costly information since the probability of casting the decisive vote is negligible (Downs, 1957).

[3] See, for example, Markus (1988, 1992) and Bagues and Esteve-Volart (2016) for empirical evidence that general economic conditions shape voters' support for the incumbent.

[4] The notion of politically compromised forecasts can be further motivated by anecdotal evidence. After the 2010 general election in the United Kingdom, the newly appointed government created the Office for Budget Responsibility for the primary purpose of providing unbiased forecasts and ending government interference with economic and fiscal forecasting (Giugliano, 2015, in the *Financial Times*).

optimistic GDP growth forecasts just before general elections. Using high-frequency forecaster-level data from the United States, the United Kingdom, and Sweden, we confirm key model predictions. We document that governments overestimate short-term GDP growth by 10 to 13 percent during campaign periods. Furthermore, we find that the bias is larger when the incumbent government is not term-limited or constrained by a parliament led by the opposition. Consistent with the model, we also find that the election timing and amount of available information determine the size of the bias at different forecast horizons.

Our theoretical framework adapts the Persson and Tabellini (2002) model of political selection by allowing the government to release forecasts for GDP growth to the public. In the model, voters use the forecasts to update their beliefs about the incumbent's ability before casting a vote. The incumbent politician faces a trade-off between the accuracy of the released forecast (due to reputational concerns) and the incentive to bias the estimates to increase reelection probability.

The first prediction of the model is that the government releases overly optimistic GDP growth forecasts before an election. Hence, the model predicts the existence of political forecast cycles. Second, we show that the incumbent government biases its estimates only in the presence of contingent electoral incentives. For example, incumbents who can run for reelection are predicted to bias their forecasts more than term-limited incumbents. Third, we show that the bias depends on the political strength in parliament of the incumbent government. For instance, a government politically aligned with the parliament releases more biased estimates than a government facing a parliament controlled by the opposition. We can also interpret this prediction through the lens of Jones and Olken (2005). They show that there is a stronger link between individual leaders and economic growth when there are fewer constraints and checks and balances on a leader's power. In our model, the bias is larger when the link between the politician's ability and the economic outcome is stronger. Fourth, the month of the year in which elections occur predicts whether the government biases the forecasts targeting the election-year outcome or the following year's growth. In our model, biased forecasts for the election year's GDP growth are more effective in shaping the voters' beliefs. However, they are also more costly for the government than biased forecasts for the following year's outcome. When elections are held later in the year, the marginal cost of bias dominates the expected marginal benefit and the incumbent bias its forecasts for future growth rather than for growth in the election year.

To test the model's theoretical predictions, we propose an empirical strategy that allows us to identify cycles using country-level data even in the absence of sharp quasi-experiments. We exploit the multiplicity of agencies that release forecasts for a country's GDP growth on a high-frequency basis to compare forecasters targeting the same outcome. More specifically, our data allow us

to identify electoral cycles by combining three different sources of variation. We compare forecasts released by the government with forecasts released by other institutions in the same period, forecasts released before the election and those released after it within the same year, and forecasts released in election years and those released in off-election years. The countries in our sample differ in terms of institutions since we include presidential and parliamentary democracies as well as majoritarian and proportional electoral systems. Also, the frequency and the timing of elections differ since we include countries voting every second year, countries voting more seldom, and countries voting both earlier and later in the year. These institutional differences strengthen our empirical results' external validity and allow us to test the theoretical predictions.

For all three countries, we detect large electoral cycles in the forecast bias for short-term GDP growth. In other words, we detect the existence of political forecast cycles. In our estimations, we find that the coefficient of interest—which captures the additional impact of the pre-election months on the bias in government forecasts—ranges between 0.16 and 0.31 percentage points. As compared to average GDP growth in our samples (2.5% in the U.S., 1.6% in the U.K., and 2.3% in Sweden), our results suggest that governments overestimate economic performances by 10 to 13 percent ahead of elections.

Consistent with the theoretical predictions, we find that the election timing matters. Specifically, we find that, in the United States, where elections occur late in the year, the government releases biased forecasts targeting the following year instead of the election year. We find for the United Kingdom, where elections in our sample take place during the spring, a bias in the forecasts targeting the election year but not the following year. We find for Sweden, where elections take place in September, that the government releases biased forecasts for both the election year's and the following year's outcome during campaign periods.

We take advantage of the presence of term-limits and cases of divided government in the United States to test the other theoretical predictions. Consistent with the model, the results show that the bias is more pronounced when the president has contingent individual reelection incentives and the party of the president has the majority in both branches of Congress. Thus, this highlights electoral incentives as the primary channel behind the evidence of the overestimation of GDP growth approaching elections.

The theoretical model also shows that it is inefficient to release biased forecasts if voters are rational—i.e., expect the government to overestimate economic growth—because it is costly to bias while the probability of being reelected does not increase in equilibrium. In contrast, if voters do not expect the government to release biased forecasts, the incumbent politician will gain electoral advantage from the bias since there would be an impact on posterior beliefs about the politicians' ability. Hence, the bias in equilibrium leads

to an increase in reelection probability only in the case of naive voters, and, subsequently, the bias results in an incumbency advantage.

The inefficient outcome in the case of rational voters arises since the politicians cannot credibly commit not to bias and are forced to bias since voters expect them to. The data availability and institutional settings in Sweden and the United Kingdom allow us to evaluate two possible tools available to the government to commit not to bias. First, we exploit the reform after the 2010 general election that outsourced the government's primary forecasting function in the United Kingdom from the H.M. Treasury to the Office for Budget Responsibility (OBR). This reform was motivated by the notion of politically compromised forecasts and aimed to end government interference with economic and fiscal forecasting (Giugliano, 2015, in the *Financial Times*). Our results show that outsourcing the forecasting function reduced the overall forecast error compared to the period before the reform. However, it failed to reduce or eliminate the cyclical bias estimated in proximity to elections—suggesting that outsourcing does not necessarily represent a credible commitment device. Second, we exploit the heterogeneity across forecasters in the public sector in Sweden. Several public agencies provide forecasts that are independent of those released by the Ministry of Finance. We find that only the Ministry of Finance releases biased forecasts approaching elections. Other administrative agencies, which either are independent from the central government or have mild connections to it, release estimates that do not follow the election cycle. These results indicate that governments can limit the damage to voters generated by biased forecasts if other agencies release unbiased information to the public.

While our model only considers the potential loss to voter welfare, biased forecasts could also damage the economy from a broader perspective. Coibion et al. (2018) show that firms care about macroeconomic variables such as GDP, inflation, and unemployment. They also show that firms update their beliefs (forecasts) when presented with forecasts from professionals. Furthermore, Tanaka et al. (2019) show that firms' GDP forecasts are associated with their investment and employment choices. Hence, biased forecasts could result in inefficient firm and household decisions if they fail to account for electoral incentives.

This paper extends several strands of literature. First, it contributes to the vast literature that studies electoral cycles.[5] Even though the literature predominantly covers political budget cycles and the underlying conditions, it is

[5]Empirical evidence of electoral cycles in fiscal policy is well established and can be found in, for example, Akhmedov and Zhuravskaya (2004), Alesina et al. (1992), Alesina and Paradisi (2017), Bartolini and Santolini (2009), Brender and Drazen (2013), Drazen and Eslava (2010), and Repetto (2018); and Shi and Svensson (2006). The literature has also expanded to study the underlying conditions for political budget cycles: country development (Schuknecht, 1996), political fragmentation (Perotti & Kontopoulos, 2002), transparency and political polarization (Alt & Lassen, 2006), media freedom (Veiga et al., 2017), budget process and checks and

not limited to electoral cycles in fiscal policy. For example, Brown and Dinc (2005) and Dam and Koetter (2012) provide evidence for electoral cycles in bank bail-outs, and Müller (2020) identifies electoral cycles in macroprudential regulations. We contribute by establishing that government forecasts follow the election cycle and directly impact the re-election of incumbent politicians. More specifically, we show that releasing biased forecasts is sufficient for generating an incumbency advantage if voters do not expect the government to release biased estimates. Thus, we also extend the literature on the mechanisms behind incumbency advantages, often observed in the data.[6]

Second, we highlight a theoretical and empirical connection between electoral incentives and bias in forecasts released by the government. Previous work has focused on bias in revenue forecasts as a tool to not comply with balance-budget requirements both in election and off-election years (e.g., Bohn & Veiga, 2020; Picchio & Santolini, 2020). Others have drawn attention to cognitive biases worsening the quality of the forecasts in abnormal macroeconomic periods (Krause, 2006) and depending on party ideology (Frendreis & Tatalovich, 2000).[7] More broadly, we contribute to research on strategic incentives of macroeconomic forecaster (e.g., Laster et al., 1999; Marinovic et al., 2013) by establishing the different trade-offs faced by the government as compared to private forecasters. This paper generalizes the work in Cipullo and Reslow (2021), which introduces the concept of electoral incentives of macroeconomic forecasters theoretically and empirically. The analysis in Cipullo and Reslow (2021) is limited to cases of referenda with substantial potential consequences for the economy. This paper shows that governments systematically released biased estimates approaching each election.

The results in this paper also highlight a need for watchfulness in the macroeconomic literature. Macroeconomic forecasts are often used in research studying, for example, information rigidity (e.g., Coibion & Gorodnichenko, 2015) or economic uncertainty (e.g., Altig et al., 2020; Bomberger, 1996). Researchers need to be careful regarding the incentives of forecasters, and consider that some of them, during certain periods, may face electoral incentives. For instance, greater disagreement between forecasters before an election might depend on both increased economic uncertainty and divergent electoral incentives.

---

balances (Saporiti & Streb, 2008), and politician characteristics (Hayo & Neumeier, 2012), to mention a few.

[6]The existence of an incumbency advantage has empirically been confirmed in numerous papers (e.g., Erikson, 1971; Freier, 2015; Gelman & King, 1990; Lee, 2008) There are several commonly invoked sources of incumbency advantage: access to resources (such as staff) attached to legislative office, press coverage, name recognition, and pork-barrel spending.

[7]See also Krause et al. (2006), who isolates the relationship between political incentives and government revenue forecasts by comparing the U.S. states in which the forecasting function is controlled by a combination of politically appointed and merit-selected subordinates, with states in which the forecasting function is completely controlled by politically appointed staff.

2 Theoretical Framework

This section proposes a simple model to analyze the incentives that incumbent politicians face when releasing GDP growth forecasts before an election. For the purpose at hand, our model builds on the two-period career concern model from Persson and Tabellini (2002) by allowing for purely office-motivated politicians and exogenous fiscal policy as well as the forecasting function of the government.

2.1 Setup

Consider a two-period model of voters and one incumbent politician who will face a random opponent in an election held during the first period. In the first period, the incumbent politician implements a fiscal policy and is tasked with supplying forecasts about the economy to the public. When the incumbent politician releases the forecasts, she can decide whether to release her best (unbiased) prediction or release manipulated (biased) information. After the election, the winner implements a fiscal policy during the second period.

Neither voters nor the incumbent politician perfectly observe the current or future state of the economy. However, they do know the structure of the economy and prior distributions of relevant variables. In addition, the incumbent politician receives noisy unbiased information about the economy. Before the election, the incumbent politician releases the forecasts, and voters use noisy media reports on the released forecasts when deciding on whom to support.

The politician in office is office-motivated and receives an exogenous rent from holding office. This rent is zero otherwise. When the incumbent releases the forecasts, she faces a trade-off between the accuracy of the forecasts and the incentive to overestimate the state of the economy to increase the reelection probability. Voters aim to elect the politicians that yield the highest future economic outcome, net of ideological preferences, which are orthogonal to the economy.

In this political economy, the economic outcome comes as a combination of fiscal policy and an idiosyncratic shock to productivity. We think of GDP growth as the economic outcome. In each period, the incumbent politician implements a fiscal policy that comes exogenously as a function of her innate ability. Formally, the economic outcome is given by

$$y_T = \lambda \eta^P + \nu_T, \tag{1}$$

where $\eta^P \sim \mathcal{N}(\bar{\eta}, \tau_\eta^{-1})$ is the innate ability of the politician in office, and $\nu_T \sim \mathcal{N}(0, \tau_\nu^{-1})$ is a productivity shock, orthogonal to η^P. The politician's true ability is unobserved by all individuals in the economy, including the politician herself. The parameter $\lambda \in (0,1)$, known to all individuals in the economy, represents the fiscal-policy transformation from the ability to the

performance of the economy.[8] According to (1), the politician in this model is not endowed with a tool to manipulate fiscal policy to increase the probability of reelection. This assumption keeps the model more tractable and allows us to show that bias in macroeconomic forecasts directly impacts electoral success. The model consists of two periods, $T = \{1,2\}$. We think of a period as a calendar year. The election is held during period one at time $t \in (0,1)$. We define t such that $t \to 0$ represents the beginning of period one, and $t \to 1$ represents the end of period one. Figure 1 illustrates the timing of the model.

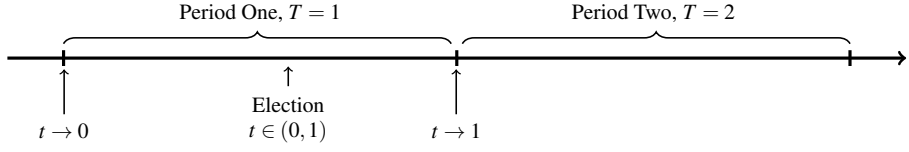


Figure 1. Timing Illustration

2.2 Information

The true state of the economy, y_T, is unobserved by voters and the incumbent politician. However, just before the election takes place, at point t of the first period, the incumbent receives two private signals about the future realizations of the economy.[9] Formally, the politician observes

$$\widetilde{y}_T = y_T + \varepsilon_T, \tag{2}$$

where $\varepsilon_T \sim \mathcal{N}(0, \tau_{\varepsilon_{T,t}}^{-1})$ is a random error. Hence, the government observes signals for both periods' outcomes.[10] To model the fact that new information about the economy increases in precision over time, we assume that $\tau_{\varepsilon_{T,t}} = \tau^{\frac{1}{h}}$, where $\tau > 1$ is a scaling parameter and $h = T - t$ is defined to be the forecast horizon. The forecast horizon measures the distance between the realization of the outcome and the release of the forecast. The true state of the economy in a given period T is revealed to both voters and the politician as the period ends.

[8] See, for example, Meriläinen (2020) for a study of the links between politician quality and fiscal policy, Besley et al. (2011) for the link between leader education and economic growth, and Jones and Olken (2005) who show that changes in political leadership lead to changes in economic growth.

[9] To reduce the number of timing parameters in the model, we assume that the signals arrive and the forecasts are released at the same time as the election is held.

[10] Note that during period one, the economic policy in period two is not yet realized. Hence, the signal represents a reduced form expectation regarding the outcome—conditional on unchanged economic policies, such that it implicitly assumes reelection of the incumbent.

The shape of the functional form of the signal precision $\tau_{\varepsilon_{T,t}}$ accounts for the fact that the precision in the signals that the incumbent politician receives increases non-linearly in the forecast horizon when forecasting annual growth rates. When the election is scheduled at the beginning of the period, then $t \to 0$ and $\tau_{\varepsilon_{1,t}} \to \tau$ while $\tau_{\varepsilon_{2,t}} \to \tau^{\frac{1}{2}}$. Conversely, if the election takes place at the end of the period, then $t \to 1$ and $\tau_{\varepsilon_{1,t}} \to \infty$ while $\tau_{\varepsilon_{2,t}} \to \tau$. Hence, at $t \to 1$, the signal for the current year is received without noise and the government observes the true state.[11]

As mentioned before, the incumbent politician is tasked with supplying forecasts about the economy to the public. The politician first forms expectations about the economy using prior knowledge and the two signals

$$\mathbb{E}(y_T|\widetilde{y}_1,\widetilde{y}_2) = m_{0,T}\lambda\bar{\eta} + m_{1,T}\widetilde{y}_1 + m_{2,T}\widetilde{y}_2, \tag{3}$$

where $m_{0,T}$, $m_{1,T}$ and $m_{2,T}$ are optimal weights according to Bayesian updating.[12] The incumbent politician then releases separate forecasts for targets $T = \{1,2\}$, which are potentially biased by an additive factor b_T:

$$F_T = \mathbb{E}(y_T|\widetilde{y}_1,\widetilde{y}_2) + b_T. \tag{4}$$

We assume that voters receive forecast information via mass media.[13] The media observes the forecasts and generates a noisy report

$$\widetilde{F}_T = F_T + e_T, \tag{5}$$

such that voters observe $\widetilde{F}_T$, where $e_T \sim \mathcal{N}(0,\tau_e^{-1})$ is a random media error. Noise in the media report reflects the observation that the media releases news articles in which the official figures are mitigated in longer pieces of information. The media noise also plays the technical role of keeping the model probabilistic from the politician's perspective.

2.3 Politician

The incumbent politician chooses an optimal bias in the forecasts by trading off accuracy of the estimates and the probability of reelection. Formally, politician I chooses b_1 and b_2 to maximize her expected utility

$$\max_{b_1,b_2} \ \mathbb{E}\Big(p^I(\cdot)|\widetilde{y}_1,\widetilde{y}_2\Big)R - \sum_{T\in\{1,2\}} \tau_{\varepsilon_{T,t}}\big(b_T\big)^2, \tag{6}$$

[11] See Figure A1a in the Appendix for an illustration of the signal precision as a function of t.

[12] Due to the common component $\lambda\eta^I$, the signal for the period one outcome, $\widetilde{y}_1$, is informative not only for the first period but also for the second period, and vice versa. See Section A.1 in the Appendix for the weights.

[13] See, for example, Durante and Knight (2012), Durante et al. (2019), Gentzkow et al. (2011), and Knight and Tribin (2018) for the media's role in shaping beliefs as an information channel.

where $p^I(\cdot)$ is the probability that politician I wins the election held in period 1. $R > 0$ is the exogenous rent for holding office, and b_T is the bias that the politician has the opportunity to add to her forecast for target period T.

We assume that the cost of biasing a forecast for a given target is quadratic in the bias.[14] We also assume that the marginal cost of a forecast error is decreasing in the precision of the available information. Hence, it is not as costly to bias the forecast for the second period as it is for the first period since the available information is more precise for the first period. Therefore, we add to the cost component of (6) the precision in the signals, $\tau_{\varepsilon_{T,t}} = \tau^{\frac{1}{h}}$, as a scaling term.[15] Figure A1b in the Appendix illustrates the evolution of the cost structure in (6) as t moves from 0 to 1. The figure shows that the cost is relatively similar across the two target periods when the horizon is long and explodes for period one when t approaches 1 since the forecast horizon then approaches 0.

It is worth noting that the cost function is a reduced-form representation of a more complex process. One way to think about the cost function is as a reputation cost that negatively impacts future reelection probabilities and outside options of the politician who releases biased information. The willingness to maintain a reputation is often thought of as the mechanism that keeps politicians in check (see, e.g., Besley & Case, 1995a).

2.4 Voters

Consider a continuum of voters indexed by i with total mass 1. We assume that voters have period-by-period linear preferences over policy outcomes represented by $W(y_T) = y_T$ and ideological preferences against the incumbent $\sigma_i \sim \mathscr{U}[-\frac{1}{2\phi}; \frac{1}{2\phi}]$ with density $\phi > 0$.[16] Hence, the utility of voters can be represented as

$$U_{i,T} = y_T - \sigma_i. \tag{7}$$

From (7) we see that voters care about the economy as well as having individual ideology preferences. The ideology parameter σ_i captures all policy preferences that are orthogonal to the economic outcome y_T.

[14] Note that, according to (4), the bias is analogous with the expected forecast error. Quadratic loss of the forecast error is a standard assumption and the most common in the forecasting literature (see, e.g., Granger, 1999; Granger & Pesaran, 2000; Laster et al., 1999).

[15] The intuition for this assumption is as follows: When $t \to 1$, the politician has perfect information about y_1, but not about y_2. Hence, the cost of releasing a biased forecast should be approaching infinity for period one, but not for period two. See, for example, Andersson et al. (2017) for an analysis of the relationship between horizon and forecast error when forecasting annual growth rates.

[16] These assumptions about the voters' utility allow for an electorate that is not too polarized, so that at least some voters can be persuaded by the forecasts.

At the time of election, prospective voters decide whether to support the incumbent from period one or the random opponent based on rational expectations over future economic performance.[17] Formally, voter i votes for the incumbent if and only if

$$\mathbb{E}\Big(y_2^I|\widetilde{F}_1,\widetilde{F}_2\Big)-\sigma_i\geq\mathbb{E}\Big(y_2^O|\widetilde{F}_1,\widetilde{F}_2\Big). \tag{8}$$

Given that the productivity shock in (1) is orthogonal to either politician's abilities, the unique component of y_2 that voters value for their decision is η^P. Therefore, voters use the forecasts on economic performance released by the incumbent government in period one as a signal of her innate ability η^I. The voter then decides whether to support the incumbent or vote for the random opponent with expected innate ability $\mathbb{E}(\eta^O)=\bar{\eta}$. Hence, we can rewrite the decision problem as

$$\mathbb{E}\Big(\lambda\eta^I|\widetilde{F}_1,\widetilde{F}_2\Big)-\sigma_i\geq\lambda\bar{\eta}. \tag{9}$$

The decision rule in (9) shows that voters cast their vote based on the comparison between the expected innate ability of the incumbent politician and the expected ability of the challenger. The expected ability of the incumbent is formed from observable information. In contrast, the expected ability of an opponent, drawn at random, is equal to the average ability $\bar{\eta}$.

Voters perform Bayesian updating using the prior distribution of abilities as well as the observed forecasts $\widetilde{F}_1$ and $\widetilde{F}_2$ to infer the ability of the incumbent politician. Their posterior belief about $\lambda\eta^I$ is given by

$$\mathbb{E}\left(\lambda\eta^I|\widetilde{F}_1,\widetilde{F}_2\right)=\gamma_0\left(\lambda\bar{\eta}\right)+\gamma_1\left(\widetilde{F}_1-\mathbb{E}(b_1|\widetilde{F}_1,\widetilde{F}_2)\right)+\gamma_2\left(\widetilde{F}_2-\mathbb{E}(b_2|\widetilde{F}_1,\widetilde{F}_2)\right), \tag{10}$$

where γ_0, γ_1 and γ_2 represent the optimal weighting according to Bayes' rule.[18] In equation (10), $\mathbb{E}(b_T|\widetilde{F}_1,\widetilde{F}_2)$ represents the voters' posterior belief about the biases, consistent with all observables and Bayesian beliefs about unobservables. Hence, to form rational expectations about $\lambda\eta^I$, voters must take into account that they cannot perfectly identify the $\lambda\eta^I$ component in y_1 and y_2 due to the productivity shocks ν_T, and the errors in the signals ε_T. They also need to account for the media errors, e_T, and the potential bias, b_T, in the forecasts.

[17] Voters are prospective in the sense that the relevant utility outcome for the election decision regards future policies and economic outcomes. However, they are retrospective in their beliefs formation. Voters form beliefs conditional on already implemented policies and a static incumbent ability.

[18] See Section A.1 in the Appendix for the weights and Figure A1c for the dynamics of γ_T with respect to the timing of the election. In the model, we assume that the voters only use forecasts published by the government. If voters were to have access to additional information or other forecasts, for example, from private forecasters, they would weight all the available information. Hence, the voters' weight on the government forecasts would be lower, and, in turn, the politicians' impact on voters' beliefs would be lower, but still present. See Cipullo and Reslow (2021) for a model with multiple forecasters.

2.5 Equilibrium

Voters that are indifferent about the incumbent and the random opponent are denoted swing voters, and according to (9) they are defined by

$$\widetilde{\sigma} = \mathbb{E}\left(\lambda \eta^I | \widetilde{F}_1, \widetilde{F}_2\right) - \lambda \bar{\eta}. \tag{11}$$

All individuals who prefer the incumbent, and this preference is stronger or equal to the preference of the swing voters, will support the incumbent. Therefore, the equilibrium vote share for the incumbent politician is given by

$$\pi^I = \int_{-\frac{1}{2\phi}}^{\widetilde{\sigma}} \phi \; di = \frac{1}{2} + \phi\left[\mathbb{E}\left(\lambda \eta^I | \widetilde{F}_1, \widetilde{F}_2\right) - \lambda \bar{\eta}\right], \tag{12}$$

while the probability that the incumbent politician wins the electoral competition is equal to

$$p^I = P\left(\pi^I > \frac{1}{2}\right) = G\left(\mathbb{E}(\lambda \eta^I | \widetilde{F}_1, \widetilde{F}_2)\right), \tag{13}$$

where $G(\cdot)$ is the cumulative distribution function of the random variable $\lambda\bar{\eta} + \gamma_1 e_1 + \gamma_2 e_2 \sim \mathcal{N}\left(\lambda\bar{\eta}, \tau_e^{-1}(\gamma_1^2 + \gamma_2^2)\right)$.[19] Substituting (13) in (6), and taking politician's expectations over (10), the politician maximizes

$$\max_{b_1, b_2} \; G\Big(\mathbb{E}\left(\mathbb{E}(\lambda \eta^I | \widetilde{F}_1, \widetilde{F}_2) | \widetilde{y}_1, \widetilde{y}_2\right)\Big) R - \sum_{T \in \{1,2\}} \tau_{\varepsilon_{T,t}} \left(b_T\right)^2. \tag{14}$$

The first-order conditions with respect to b_T for an interior solution imply that

$$b_T = \frac{R\gamma_T}{2\tau_{\varepsilon_{T,t}}} g\Big(\mathbb{E}\left(\mathbb{E}(\lambda \eta^I | \widetilde{F}_1, \widetilde{F}_2) | \widetilde{y}_1, \widetilde{y}_2\right)\Big) \; > 0, \tag{15}$$

$\forall \, T \in \{1,2\}$, so that it is optimal for the government to release biased forecasts for both periods when approaching an election. In equation (15), $g(\cdot)$ is the probability density function associated with the cumulative distribution function $G(\cdot)$. In a Perfect Bayesian Equilibrium, voters' beliefs about the bias in the forecasts are consistent with optimal strategies, and strategies are optimal given consistent beliefs, such that

$$\mathbb{E}(b_T | \widetilde{F}_1, \widetilde{F}_2) = \frac{R\gamma_T}{2\tau_{\varepsilon_{T,t}}} g\Big(\mathbb{E}(\lambda \eta^I | \widetilde{F}_1, \widetilde{F}_2)\Big). \tag{16}$$

Finally, the politician's expectations over (16), which is needed to solve for (15), closes the model implicitly.

[19] Hence, the random variable captures the uncertainty that arises due to the media error, e_T.

2.6 Model Predictions and Intuition

This section presents several testable model predictions and further intuition for how to interpret the model parameters. While some predictions can be derived directly from equation (15), we also solve the model using numerical methods to derive additional predictions.[20] The model predictions of interest are as follows.

Prediction 1: existence of bias. The first and main prediction of the model is the existence of bias. The model predicts that, during campaign periods, the government will overestimate the forecast for both the election year (period one) and the following year (period two) to influence voters' beliefs. This overestimation generates an electoral cycle in the forecast bias of the government. The prediction follows directly from equation (15) and establishes a theoretical rationale behind the existence of political forecast cycles.

Prediction 2: electoral incentives. Equation (15) shows that the bias is increasing in R. This prediction is not surprising since R captures the strength of the electoral incentives of the incumbent politician. Since term limits, which prevent incumbents from running for reelection, represent a case of reduction in R (see, e.g., Besley & Case, 1995a, 1995b), the model predicts that term-limited incumbents will release more accurate forecasts than incumbents who can run for reelection.[21]

Prediction 3: divided government. λ determines the correlation between ability and economic growth. Hence, λ captures how strong the link between ability and economic growth is and can be interpreted to capture the incumbent government's political strength. For instance, λ is higher when the president is politically aligned with the parliament compared to when the government is divided. Figure 2a shows that the bias is negligible for low values of λ (i.e., when the incumbent politician's ability plays a limited role in shaping the economy) and the bias increases when λ increases. Therefore, the model predicts that the bias is reduced if the government is not aligned with the parliament. We can also interpret this prediction through the lens of Jones and Olken (2005). They show that there is a stronger link between individual leaders and economic growth when there are fewer constraints and checks and balances on a leader's power.

Prediction 4: election timing. The month of the year in which elections occur is a crucial ingredient of the model. Both the impact of a forecast on the voters' posterior belief about the incumbent's ability and the marginal cost of bias depend on the forecast horizon. In the model, the election timing is

[20] See Table A1 in the Appendix for calibration of parameter values.

[21] Harrington (1992) shows in an OLG model that infinitely lived parties that care for survival into office also after the end of the incumbent politician's career can induce the incumbent to trade individual incentives for partisan incentives. In such cases, R is lower—not zero—for term-limited incumbents. If parties do not affect the incumbent politician's incentives, term limits move R to zero. Our model, under both assumptions, predicts that term limits reduce the bias.

determined by the parameter t. More specifically, $t \to 0$ represents an election held in the beginning of the year. $t \to 1$, instead, indicates elections held at the end of the year. Figure 2b shows that the model predicts that politicians will strategically bias differently across the two available target periods, depending on t. For small values of t, the bias is comparable between the two targets. Although the period one forecast has a higher weight on voters' posterior beliefs about the incumbent's ability, the cost is also always higher for the first period than the second period. When t approaches one, the cost of biasing the forecast for the first period rapidly approaches infinity while the influence on voters' beliefs is bounded. Therefore, for large values of t, it becomes relatively more profitable for politicians to bias the forecast subject to only the longer horizon, even if this has a lower impact on voters' beliefs.[22]

(a) Divided Government *(b)* Election Timing

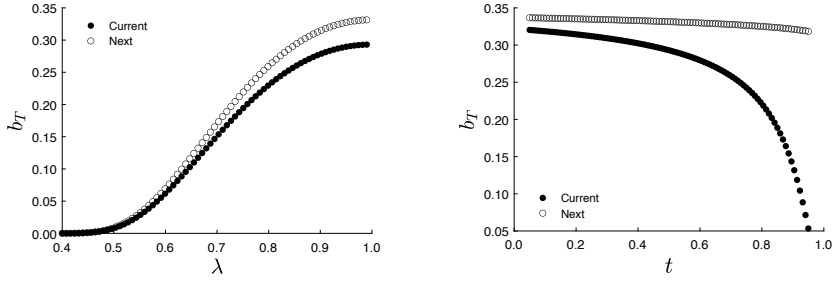


Figure 2. Model Predictions

Notes: Model predictions based on the calibration presented in Table A1. The filled dots represent the bias in the forecasts targeting the election period ($T = 1$) while the non-filled dots represent the bias in the forecasts targeting the following year ($T = 2$).

2.7 Commitment vs. Manipulation

Suppose voters are rational and expect the government to release biased estimates. In this case, they can account for it in expectations, and the signaling tool used by the politician will represent a Pareto inefficiency. The bias is costly and neither increases the equilibrium probability of success nor helps voters sort out able incumbents in equilibrium. However, politicians who cannot commit credibly not to bias are forced to do so as long as voters expect them to overestimate economic growth.

Suppose voters are naive in the sense that they do not expect the government to overestimate the economy intentionally. In this case, the politician will gain

[22] See Figure A1 in the Appendix for the signal precision, cost of bias, and impact on voter beliefs as a function of t.

electoral advantage from the bias since posterior beliefs about the politicians' ability would be affected by its magnitude. Compared to a counterfactual world with no bias, voters then would be worse off because of systematic voting mistakes, and incumbent politicians would be better off thanks to an increased reelection probability.

We solve our model for the case of naive voters by imposing the condition $\mathbb{E}(b_T|\widetilde{F}_1,\widetilde{F}_2)=0$ instead of equation (16). Voters perform Bayesian updating according to equation (10) but do not account for the bias, while the politician's expectations about voters' expectations are correct. Figure 3 documents that the probability of reelection is higher for any level of innate ability when voters are naive compared to when voters are rational. For instance, politicians with average innate ability have a probability of reelection equal to 50 percent when voters account for the bias. In contrast, the probability of reelection is higher under the assumption of naive voters.[23]

Hence, the overestimation of economic growth generates an incumbency advantage only if voters do not expect the government to release biased estimates. The bias is present both in the case of rational and naive voters, yet the mechanisms behind it and the consequences for the electoral competition change substantially. While the bias arises from lack of commitment under rational voters, it comes from manipulation of voters' beliefs if voters are naive.

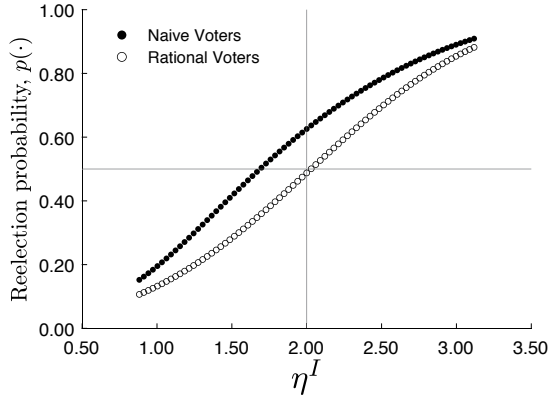


Figure 3. Reelection Probability under Rational and Naive Voters

Notes: Model prediction based on the calibration presented in Table A1. The filled dots represent the bias assuming naive voters while the non-filled dots represent the bias assuming rational voters. We illustrate the reelection probability for the ability levels (η^I) that lie within one-half standard deviation from the mean of 2.

[23]For ability levels further away from the average, the gain in the probability of reelection is attenuated.

3 Institutional Background and Data

We test the predictions of the theoretical model by exploiting high-frequency panel data at the forecaster level from the United States, Sweden, and the United Kingdom. We have data that contains the short-term (current and next year) GDP growth forecasts from multiple forecasters for all three countries. The countries included in the sample are primarily selected because they satisfy our data availability needs. We require at least bi-annual releases from both the government and a pool of non-government forecasters—of which one release is before and one is after the election day. Multiple countries strengthen the external validity of our empirical results and allow us to test the theoretical predictions by including countries that are heterogeneous in terms of both electoral rules and institutions and the seasonality of elections and length of electoral cycles. We combine the forecast data with real-time data on the actual GDP growth realization in each year to measure the ex-post forecast error.[24] In the remainder of this section, we summarize the relevant information about the form of government and national elections in the countries included in the sample as well as the sources and main qualities of the data.

3.1 United States

Background. The United States is a presidential democracy. The union is composed of fifty states, which elect their representatives to the national parliament (*Congress*) and participate in the election of the president. Elections at any level are held on the first Tuesday after the first Monday in November in even-numbered years.

Congress holds the legislative power and is formed by the *House of Representatives* and the *Senate*. The members of the House of Representatives serve two-year terms representing the people of a single constituency. The Senate members serve six-year terms, with staggered elections, so that every second year, approximately one-third of the Senate is up for election.

Presidential elections take place every fourth year.[25] The popular vote for the president is indirect. Each state is assigned as many delegates to the *Electoral College* as its number of Congress members, and each state has the authority to determine whether to assign all its delegates to the candidate with the most votes or proportionally.[26] Incumbent presidents are term-limited after

[24] GDP growth releases are subject to revision, and definition changes over time. Real-time data make it possible to compare the observed economic growth with the released forecasts as they represent the first release of the outcome available in each period.

[25] If the president resigns at some point during the term, the vice president steps in until the natural end of the term. Since the first contested election held in 1796, the country has never experienced early presidential elections.

[26] As of 2019, only Maine and Nebraska assign their delegates based on a proportional system. Notice that the District of Columbia, which does not elect any voting member of Congress,

two consecutive terms. The variation in the timing of elections has generated relatively many cases of *Divided Government*, in which the party of the president does not have the support of the majority of both branches of Congress.

Data. We build our panel of forecasts based on the *Livingston Survey* collection. The Livingston Survey is a survey of banking, governmental, academic, and other forecasters. It was started in 1946 by the columnist Joseph Livingston and is currently published in June and December every year. Our data report GDP growth forecasts for the current and the next year released between 1972 and 2018.[27] We create an indicator called *government* that takes the value 1 for all forecasters labeled as Government in the Livingston Survey and 0 otherwise. We combine this data with real-time actual GDP growth rate data from the Federal Reserve Bank of Philadelphia and information we collected ourselves about relevant political variables. Panel A of Table A2 in the Appendix presents descriptive statistics regarding our data for the United States, and Figure A4a provides the forecast coverage by institution type.

3.2 Sweden

Background. Sweden is a parliamentary democracy. The national parliament (*Riksdagen*) is composed of one house, elected every fourth year on the second Sunday in September.[28] Members of the parliament are elected based on a proportional representation system in small constituencies and an entry threshold of 4% of the votes.

The head of political power is the prime minister. After each general election, the parliament votes on the incumbent prime minister and determines whether she can remain in power or she has to step down. Every time the prime minister resigns, the speaker holds consultations with the parties represented in parliament and appoints a new prime minister, who needs to receive approval from the majority of the parliament.[29]

Data. Our analysis builds on real-time data collected by the National Institute of Economic Research (*Konjunkturinstitutet*), reporting forecasts released between 1994 and 2018 for the current and next year's growth by the most prominent forecasters for the Swedish economy. The collection includes financial institutions as well as public and international agencies, trade and

is awarded the same number of delegates as the lowest-represented states. Currently, D.C. is guaranteed three members so that the members of the electoral college amount to 538.

[27] The collection of the survey has been the responsibility of the Federal Reserve Bank of Philadelphia since 1990. The relevant forecasts for our study have been collected since 1972.

[28] Until 2010, general elections took place on the third Sunday of September.

[29] According to the Constitution, the parliament has four chances to approve the proposed prime minister before an early election becomes mandatory. The most recent early election took place in 1958.

industry unions, and labor unions.[30] We combine the forecast data with real-time actual GDP growth data from Statistics Sweden (1994–1998) and the OECD (1999–2018).

In Sweden, official forecasts are released independently by multiple governmental divisions and agencies, which differ in terms of distance from the electoral incentives of policymakers. We define the government forecasts to be those released by the Ministry of Finance and are under the finance minister's direct control. The presence of several other government forecasters, such as the National Institute of Economic Research, the Swedish Public Employment Service (*Arbetsförmedlingen*), the Swedish National Debt Office (*Riksgäldskontoret*), and the Swedish National Financial Management Authority (*Ekonomistyrningsverket*), allows us to disentangle those who may have stakes in the election results from those that are without incentives to affect voters' beliefs.[31] Panel B of Table A2 in the Appendix presents descriptive statistics regarding our data for Sweden, and Figure A4b reports the forecast coverage by institution type.

3.3 United Kingdom

Background. The United Kingdom is a parliamentary democracy. The national parliament is composed of the *House of Commons* and the *House of Lords*. The head of the executive power is the prime minister, who is appointed by the King or Queen. Conventionally, the monarch appoints the person who is most likely to gather the confidence of the majority of the House of Commons.[32]

Historically, the length of terms in the United Kingdom was not predetermined by law, even if, in most cases, elections took place every fourth or fifth year. Since 2011, the House of Commons is elected every fifth year while membership to the House of Lords is granted by appointment, based on heredity or official function. The House of Commons is the only house in the parliament assigned the right to vote in favor of or against the government.

Early elections are called either when it is impossible to form a government with the confidence of the House of Commons or following an explicit decision by the incumbent government, which has the formal power to choose the election date. In this century, early elections were called in 2017 and 2019 since it has been difficult to approve the set of bills necessary to effect the

[30] The data collection is updated immediately when a forecaster releases a new update, so we can observe the exact timing instead of a screen-shot determined by a survey date.

[31] The National Institute of Economic Research is an agency under the Ministry of Finance tasked to perform *independent* analysis and forecasts for the Swedish and international economy as a basis for economic policy.

[32] In the event of the prime minister's resignation, or a loss of the confidence from the House of Commons, the monarch has the opportunity of appointing a new prime minister, whose government needs the support from the House of Commons.

withdrawal from the European Union after the result of the 2016 Brexit referendum. In both cases, early elections have been called following a decision by the incumbent government.

Data. Government forecasts for GDP growth and other macroeconomic indicators had been released by the H.M. Treasury (the U.K. Ministry of Finance) until the 2010 elections. Subsequently, the newly appointed conservative government outsourced the task to a newly formed agency (Office for Budget Responsibility), motivated by the presumption that the Labour party had previously used the Treasury's forecasts to boost their reelection probability.

Our empirical analysis builds on the monthly survey *Forecasts for the U.K. economy*, conducted and released by the H.M. Treasury, observed between January 1998 and April 2018. The survey publishes forecasts for the current and next year released by financial institutions as well as research companies, industrial and public forecasters that are either located in the City of London's financial district or elsewhere. We merge this data with the government forecasts released by the H.M. Treasury itself, observed between 1998 and May 2010, and forecasts from the Office for Budget Responsibilities between June 2010 to the end of 2017.[33] Real-time data on actual GDP growth is collected from the OECD. Descriptive statistics for the United Kingdom are presented in Panels C and D of Table A2 in the Appendix, while Figure A4c provides the forecast coverage by institution type.

4 Empirical Strategy

The theoretical model presented in Section 2 predicts that the incumbent government manipulates the economic growth forecasts just before elections to increase the probability of being reelected. The main empirical challenge is that a significant forecast error might depend both on the attempt to influence voters before the election and on confounders that make it more difficult to develop forecasts approaching a vote. For instance, a comparison between the ex-post forecast error that the government makes before and after the ballot date would be affected by the different forecast horizons, as all pre-election forecasts would systematically be released when relatively less information is available. Similarly, a comparison between the forecasts released in election years and the estimates published in off-election years would be affected by the additional uncertainty that the election outcome generates (see, e.g., Bloom, 2014; Bloom et al., 2007), as well as by the presence of electoral cycles in real macroeconomic variables (see, e.g., Alesina et al., 1992). Lastly, a

[33] In the main analysis, we restrict the sample to the observations collected before the decision to outsource the forecasting competence to the independent OBR. In Section 5.2, we investigate whether the replacement of forecasts released by the H.M. Treasury itself with the newly formed independent agency had an impact on forecast errors and their potential election cyclicality.

cross-sectional comparison between forecasts released by the government and the ones released by other forecasters would be threatened by the potential differences in available information between the government and other forecasters and alternative incentives motivating the forecasting activity of national governments and private firms.

Our data's panel structure and high frequency allow us to alleviate the potential endogeneity concerns by simultaneously combining three sources of variation. In particular, we compare i) variation within a forecaster, across different periods; ii) variation within each forecasting horizon, across forecasts released in election and off-election years; iii) variation within each year, across predetermined election dates. The dependent variables of interest are the ex-post forecast errors at different forecast horizons. We define the forecast errors as the difference between the forecast and the real-time realization of the outcome. The use of the forecast error as the dependent variable is beneficial in two ways. First, it allows us to address whether the government releases systematically different estimates than private forecasters during the pre-election periods compared to differences in off-election periods. Second, it provides information regarding which institutions were releasing on average overly optimistic or pessimistic forecasts compared to the ex-post realization of GDP growth.

We explore whether the government releases biased GDP growth forecasts during campaign periods with the difference-in-differences model

$$E_{i,h,T} = \theta_i + \delta_h + \mu_T + \alpha\ campaign_{h,T} + \beta\ government_i \times campaign_{h,T} + \epsilon_{i,h,T}, \tag{17}$$

where $E_{i,h,T}$ is the forecast error by institution i for the GDP growth outcome in year T at horizon h. The forecast error is defined as the difference between the forecast and the realization, and the horizon is defined as the difference (in months) between the end of year T and the forecast's release month. The indicator $campaign_{h,T}$ takes the value one if the forecast is released in the same year of a national election and before the election day and zero otherwise. Likewise, $government_i$ is an indicator taking the value one if the forecaster is defined to be the government and zero otherwise. θ_i is the forecaster fixed effect, while δ_h and μ_T are respectively the horizon and the year fixed effects.[34] For each country, we estimate equation (17) using the current-year forecasts and the next-year forecasts, separately. Current-year forecasts are those released within the same year to which the outcome refers, while next-year forecasts are those released during the year preceding the target year of the forecast.

The coefficient of interest, β, captures the government forecasts' additional forecast error during election campaigns compared to other institutions and

[34]For the United States, many forecasters are included in the sample only for a few years. To not lose an important source of variation, we estimate a version of (17) for the United States in which we replace the forecaster fixed effects with industry (type) fixed effects.

the government itself in non-campaign periods. A positive and significant β would imply that the government systematically overestimates GDP growth during campaign periods, given the information available at the time of the release. We also present the results using the following specification, in which we replace the individual forecaster fixed effects with a common constant and a government indicator:

$$E_{i,h,T} = \theta + \delta_h + \mu_T + \alpha\ campaign_{h,T} + \beta\ government_i \times campaign_{h,T} + \psi\ government_i + \epsilon_{i,h,T}. \quad (18)$$

The estimation of (18) is informative to show the overall robustness of our findings and endow the reader with additional information on how the government's forecasts compare with the estimates published by other forecasters in non-campaign periods.

The validity of the empirical strategy rests on two main identifying assumptions. First, absent the campaign, the difference between the forecasts released by the government and those released by other forecasters would reflect the difference observed by the two groups of forecasters in periods without elections. This assumption reflects the standard parallel trends assumption of difference-in-differences models. Second, the forecasts released by institutions in the control group should not be affected by political campaign incentives at the time of the release and should reflect the efficient use of the available information. In Section 6, we perform several robustness checks to alleviate potential concerns regarding our identifying assumptions.

The literature on opportunistic election timing often assumes that early elections are timed for partisan advantage by incumbent governments to coincide with favorable circumstances, such as peaks in economic performance (see, e.g., Balke, 1990; Kayser, 2005). Therefore, the empirical literature on political budget cycles strongly relies on the exogeneity of the election schedule with respect to the fiscal policy implemented by politicians. In the context of this paper, this is of no concern. Favorable circumstances, such as high economic growth, do not contest our results. Our analysis's validity only requires that governments do not call early elections as a consequence of overly optimistic forecasts, which seems very unlikely. Moreover, as described in Section 3, two of the three countries included in our sample have not called an early election during the sample period.

We estimate the model separately for the United States, Sweden, and the United Kingdom, taking a threefold advantage from showing evidence from multiple countries. First, these countries vote in different months of the year. Suppose the predictions of the model find support in the data. In this case, the governments in each of the countries should bias differently across forecast targets based on whether elections are held at the beginning, in the middle, or at the end of the year. Second, this allows us to address external-validity concerns and establish the presence of electoral cycles in the government's macroeconomic forecasts as consistent evidence across institutional

rules, election frequency, and election timing. Third, data availability differs across countries in terms of the number of elections, frequency of forecast updates, and the number of observations. Restricting the attention to either one of the countries would have generated a trade-off between the characteristics.[35]

In all specifications, the inference is based on two-way cluster robust standard errors (Cameron et al., 2011; Cameron & Miller, 2015) at the forecaster and the horizon-by-target year levels. In this way, we account for the potential autocorrelation in the forecast error of each forecaster as well as for the cross-sectional correlation of forecasts subject to the same target and available information.[36]

5 Results

In this section, we present the results of the empirical analysis. We start by providing evidence that, in all countries in our sample, the government overestimates GDP growth in the months approaching an election compared to the other forecasters in the economy and the government itself in off-election years and the months following the vote. We also show that the forecasts for the current-year and the next-year GDP growth are differently biased depending on the season in which each country casts its vote.

Table 1 reports the results using the current-year forecasts in columns (1) and (2), while columns (3) and (4) refer to the estimations using the next-year forecasts. In columns (1) and (3) we estimate equation (18), while in columns (2) and (4) we estimate equation (17). The coefficient of interest β—which captures the additional impact of the pre-election months on the bias in government forecasts—is reported as the interaction term *Government* $\times$ *Campaign*. In the estimations without forecaster fixed effects, the coefficient ψ—which captures the average difference in the forecast error between private and government forecasters—is reported with the label *Government*. Even if our empirical strategy is not designed to identify ψ consistently, its size and sign can still be informative about the government's behavior, independently of elections. For instance, the government may in general be overconfident about its policy's effects on the economy (e.g., Krause, 2006) or bias the forecasts to deviate from balanced-budget requirements (e.g., Bohn & Veiga, 2020; Picchio & Santolini, 2020).

[35] For instance, our data for the United States goes back to the 1972 election but contains only biannual observations, while our U.K. data are monthly but observed only between 1998 and 2010 given our main specification.

[36] For Sweden, we only have 20 forecasters in the sample, so cluster-robust inference is not feasible (Donald & Lang, 2007). We account for this data limitation by calculating two-way cluster robust standard error at the forecaster and the horizon-by-target year levels based on 999 wild bootstrap repetitions at the individual level (Cameron et al., 2008). The two-way wild bootstrap standard errors turn out, as expected, to be more conservative than the standard two-way clustered standard errors.

Table 1. *Estimated Election Cycle Bias*

	Current		Next	
	(1)	(2)	(3)	(4)
Panel A. United States				
Government × *Campaign*	0.049	0.043	0.274***	0.265***
	(0.069)	(0.070)	(0.095)	(0.094)
Government	-0.030*		0.017	
	(0.018)		(0.045)	
Observations	3,082	3,082	3,057	3,057
R^2	0.454	0.460	0.748	0.750
Panel B. Sweden				
Government × *Campaign*	0.113***	0.113***	0.315***	0.307***
	(0.035)	(0.037)	(0.055)	(0.052)
Government	0.017		0.062	
	(0.035)		(0.053)	
Observations	1,028	1,028	1,034	1,034
R^2	0.663	0.683	0.907	0.915
Panel C. United Kingdom				
Government × *Campaign*	0.172***	0.156**	0.039	0.015
	(0.036)	(0.062)	(0.148)	(0.170)
Government	0.101***		0.453***	
	(0.026)		(0.055)	
Observations	3,471	3,471	3,271	3,271
R^2	0.590	0.649	0.855	0.901
Fixed Effects		✓		✓
Year Effects	✓	✓	✓	✓
Horizon Effects	✓	✓	✓	✓

Notes: The dependent variable is the GDP growth rate forecast error, where the error is defined as the difference between the forecast and the outcome. In columns (1) and (3) the estimated equation is (18), while in columns (2) and (4) the estimated equation is (17). In Panel A. United States, the fixed effects are replaced with type effects. Standard errors robust to two-way clustering at the forecaster and the horizon-by-year levels are in parentheses. In Panel B. Sweden, the standard errors are based on 999 wild bootstrap repetitions at the individual level. *,**,*** represent the 10%, 5%, 1% significance levels.

Panel A of Table 1 presents the results of the empirical analysis for the United States, where elections are held every second year in November. We detect that the government, during campaign periods, releases overly optimistic forecasts for GDP growth in the next year of 0.265–0.274 percentage points. At the same time, we estimate very small—and statistically indistinguishable from zero—coefficients in columns (1) and (2) where we use the current-year forecasts. We can interpret these findings through the lens of the model. When elections are held late in the year, it is too costly for the incum-

bent government to bias the current-year forecasts. Subsequently, electoral cycles are found only in the forecasts with the longer horizon.

In Panel B, we report the main results for Sweden, in which elections are held every fourth year in September. When elections occur in the middle of the year, the theoretical model predicts that the government manipulates both the current-year and next-year forecasts. In line with that prediction, we find that the government, during campaign periods, releases overly optimistic forecasts of 0.113 percentage points for the current-year GDP growth and 0.307–0.315 percentage points for the next-year growth.

In Panel C, we present the empirical results for the United Kingdom, in which elections in our sample take place during the spring. Prediction 4 of the theoretical model predicts that, in the case of elections held early during the year, the government has a large incentive to bias the forecast for the current-year outcome. This prediction finds support in the data. We document that the government, during campaign periods, overestimates GDP growth for the current year by 0.156–0.172 percentage points. However, we do not find any evidence of significant biased releases targeting the next year. For the United Kingdom, the manipulation of released forecasts approaching the elections is, however, accompanied by evidence that the government, in general, is more optimistic compared to the other forecasters. Specifically, we detect a general overestimation of approximately 0.10 percentage points in the current-year forecasts and 0.45 percentage points in the next-year forecasts, as reported by the coefficients attached to the *Government* indicator in columns (1) and (3).

The results presented in Table 1 confirm predictions 1 and 4 presented in Section 2.6. Governments in all countries in our sample release overly optimistic GDP growth forecasts during campaign periods. Moreover, they decide whether to overestimate GDP growth for the current year or next year based on the month in which the election occurs.

5.1 Reelection Incentives and Divided Government

We take advantage of our data from the United States, which spans several years and elections, to test additional predictions from the theoretical model derived in Section 2.6. In Table 2, we interact the variable of interest from (17) *Government* $\times$ *Campaign* with institutional indicators capturing the parameters of our model. First, in columns (1) and (4), we investigate if the bias depends on whether voters are called upon to select only the members of Congress or also the president. The results show that forecasts are equally biased approaching both types of elections since the estimated coefficients are neither statistically nor economically distinguishable from zero. This result does not come as a surprise since presidents are substantially affected by the outcome of a mid-term election, even if they are not unseated due to a defeat.

Table 2. *Reelection Incentives and Divided Government*

	Current			Next		
	(1)	(2)	(3)	(4)	(5)	(6)
Gov. × *Cam.* × *Presidential*	-0.026			0.021		
	(0.102)			(0.228)		
Gov. × *Cam.* × *Term-Limited*		-0.158			-0.416*	
		(0.130)			(0.219)	
Gov. × *Cam.* × *Divided Gov.*			-0.035			-0.238
			(0.103)			(0.256)
Government × *Campaign*	0.057	0.112	0.068	0.251*	0.431***	0.424*
	(0.084)	(0.089)	(0.107)	(0.133)	(0.105)	(0.235)
p-value: sum of coefficients = 0	0.729	0.649	0.652	0.091	0.936	0.086
Observations	3,082	3,082	3,082	3,057	3,057	3,057
R^2	0.460	0.460	0.461	0.751	0.751	0.751
Type Effects	✓	✓	✓	✓	✓	✓
Year Effects	✓	✓	✓	✓	✓	✓
Horizon Effects	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the GDP growth rate forecast error, where the error is defined as the difference between the forecast and the outcome. The estimated equations are versions of equation (17), in which we add an interaction between $Government_i \times Campaign_{h,T}$ and each of the indicator variables specified in the table. Standard errors robust to two-way clustering at the forecaster and the horizon-by-year levels are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Second, the theoretical model predicts that the electoral bias in government forecasts depends on the incumbent politician's electoral incentives. When R decreases, the incumbent is predicted to reduce the bias in the estimates released to the public. In columns (2) and (5), we test this prediction by exploiting term limits that prevent the incumbent president from competing for a third consecutive term. The results show that the government overestimate next-year GDP growth by 0.4 percentage points when the president is not term-limited, while the government does not release biased forecasts when the president is serving for a second consecutive term.

Prediction 3 indicates that the bias is smaller when the politician's relevant traits have a limited impact on economic outcomes. In the United States, the institutional setting generates non-unusual cases in which the president's party does not have the majority of seats in one or both branches of Congress. In such cases, the president and Congress limit each other's autonomy and can either compromise to agree on a policy or stop each other's decision.[37] Despite a loss in precision, the results in column (6) of Table 2 suggest, in line with the prediction, that government forecasts are relatively less biased when the president is not politically aligned with both branches of Congress.

[37] See, for example, Persson et al. (1997), Persson et al. (1998), and Persson et al. (2000) for more on the separation of powers in a presidential-congressional regime and how this can improve accountability of elected officials through checks and balances.

5.2 Outsourcing and Multiple Forecasters

As the theoretical model shows, it is inefficient to bias if voters are rational. The inefficient outcome arises since the politicians cannot credibly commit not to bias and are forced to do so since voters expect them to. This section evaluates different approaches and potential commitment devices to limit the negative economic effects of electoral cycles in forecasts. Specifically, we exploit the reform that outsourced the U.K. government's primary forecasting function in 2010 and the presence of multiple forecasters in the public sector of Sweden.

Forecast Outsourcing in the United Kingdom

During the 2010 election campaign, the opposition leader David Cameron openly criticized the incumbent Labour party for manipulating the Treasury's forecasts and advocated for the creation of an independent government budget office. After winning the election, the new government announced the creation of the Office for Budget Responsibilities (OBR), to which the government outsourced its primary forecasting function. Chancellor George Osborne was quoted as saying at the creation of the OBR that it would "rebuild confidence" in economic forecasts from the government. Later, Prime Minister David Cameron recalled "fiddled forecasts and fake figures" before the OBR was set up—blaming the Labour Party for manipulating the forecasts.

We exploit the introduction of the OBR to evaluate whether outsourcing has been effective to limit the government's opportunity to bias its releases using a triple difference specification. We estimate

$$\begin{aligned} E_{i,h,T} = {} & \theta_i + \delta_h + \mu_T + \alpha_0\, campaign_{h,T} + \alpha_1\, campaign_{h,T} \times post_{h,T} \\ & + \beta_0\, government_i \times campaign_{h,T} \\ & + \beta_1\, government_i \times campaign_{h,T} \times post_{h,T} \\ & + \psi_1\, government_i \times post_{h,T} + \varepsilon_{i,h,T}, \end{aligned} \tag{19}$$

where $post_{h,T}$ is an indicator variable taking the value one if the forecast was released after June 2010 and zero otherwise. Hence, we extend the dataset from May 2010 to April 2018, and $post_{h,T}$ indicates the behavior of the OBR relative to the H.M. Treasury. The coefficients of interest are β_1, capturing the cyclical component of the forecast error after the 2010 reform, and ψ_1, which is informative of whether the introduction of the reform was able to correct the average over-optimism in forecasts released by the government in general. We also estimate a version where we suppress the forecaster fixed effect in (19) and instead include a government indicator to obtain the specification

presented in (20):

$$\begin{aligned} E_{i,h,T} = \theta + \delta_h + \mu_T + \alpha_0\, campaign_{h,T} + \alpha_1\, campaign_{h,T} \times post_{h,T} \\ + \beta_0\, government_i \times campaign_{h,T} \\ + \beta_1\, government_i \times campaign_{h,T} \times post_{h,T} \\ + \psi_0\, government_i + \psi_1\, government_i \times post_{h,T} + \varepsilon_{i,h,T}. \end{aligned} \tag{20}$$

This natural experiment addresses one of the key points developed in the theoretical model. Suppose voters expect governments to bias their forecasts at the time of an election. In this case, the bias survives in equilibrium even if it does not increase the likelihood that the incumbent is reelected. This Pareto inefficiency arises because politicians cannot commit not to bias their releases. Outsourcing of the government's forecasting function to an external agency can represent a credible commitment if the agency does not have connections to the government's electoral incentives and does not aim to please the incumbent.

The results of this policy evaluation exercise are presented in Table 3. The estimates of β_1, presented in the first row in Table 3, show that the outsourcing did not reduce the additional forecast error during election years. However, the estimates for ψ_1 show that the introduction of the OBR had a significant impact on correcting the general over-optimism in the government forecasts (see the *Government* $\times$ *Post Reform* coefficients). Since the reform, the average error in OBR releases has been more in line with the average forecast error of private forecasters.

Table 3. *Effect of the 2010 U.K. Forecast Reform*

	Current		**Next**	
	(1)	(2)	(3)	(4)
Government × *Campaign* × *Post Reform*	0.001	0.000	-0.128	-0.142
	(0.059)	(0.076)	(0.111)	(0.116)
Government × *Campaign*	0.170***	0.160***	0.073	0.053
	(0.045)	(0.056)	(0.120)	(0.124)
Government × *Post Reform*	-0.113***	-0.156**	-0.285***	-0.348***
	(0.040)	(0.061)	(0.067)	(0.099)
Government	0.085***		0.418***	
	(0.025)		(0.055)	
Observations	5,551	5,551	5,140	5,140
R^2	0.587	0.626	0.830	0.867
Fixed Effects		✓		✓
Year Effects	✓	✓	✓	✓
Horizon Effects	✓	✓	✓	✓

Notes: The dependent variable is the GDP growth rate forecast error, where the error is defined as the difference between the forecast and the outcome. The estimated equations are (19) and (20). Standard errors robust to two-way clustering at the forecaster and the horizon-by-year levels are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

The results presented in Table 3 indicate that outsourcing does not imply a commitment to not bias the forecasts before an election but does improve the average quality of the releases. Moreover, the outsourcing might have affected voters' beliefs about the bias in forecasts. Hence, the outsourcing might have manipulated voters' preferences toward the incumbent if the reform changed voters' beliefs about government bias. There are several reasons why outsourcing may not work as a credible commitment device. The OBR is highly connected to the government, and executives are appointed by politicians. Executives may bias to please the government to ensure reappointment. Furthermore, the actual independence of the OBR is not undisputed, and the Treasury may have sought to meddle with the OBR forecasts (Ralph, 2015, in *The Times*).

Heterogeneous Government Forecasting in Sweden

In Sweden, several additional institutions and agencies in the public sector, whose interests might not coincide with the central government's electoral incentives, release forecasts regularly. The main results presented in Panel B of Table 1 refer to the bias of the Ministry of Finance compared to all other forecasters in the sample. This section repeats the exercise, focusing on the other forecasters in the public sector (albeit still sorted under the government). The economic consequences of electoral cycles in the government's forecasts will be attenuated if agencies with little or no electoral incentives release unbiased estimates to the public.

In Table 4, we report the results from estimating equation (17) among government agencies other than the Ministry of Finance.[38] In columns (1) and (3), we compare the National Institute of Economic Research (NIER)—whose core objective is to produce *independent* forecasts—to all other forecasters in the sample. In columns (2) and (4), we compare forecasts released by the Swedish Public Employment Service (PES), the Swedish National Debt Office (NDO), and the Swedish National Financial Management Authority (NFMA) with the estimates released by other forecasters.[39] The results presented in columns (1) and (3) show no evidence of electoral cycles in the forecasts released by the independent forecast agency NIER. In column (2), the estimated coefficient for the current year is just slightly attenuated compared to the results in Table 1. However, it is not statistically different from zero. The results presented in Table 4 suggest that lower electoral incentives are associated with less pronounced electoral cycles in the different forecasting functions of the public sector. Hence, the negative consequences of biased forecasts may be mitigated if the public sector can also provide unbiased information.

[38] In this analysis, we drop the Ministry of Finance (MoF) from the sample.

[39] We analyze PES, NDO, and NFMA jointly due to the limited amount of data.

Table 4. *Heterogeneous Government Forecasting in Sweden*

	Current		Next	
	(1)	(2)	(3)	(4)
NIER × *Campaign*	-0.063		0.043	
	(0.046)		(0.054)	
PES, NFMA, NDO × *Campaign*		0.108		0.006
		(0.216)		(0.083)
Observations	945	945	950	950
R^2	0.680	0.680	0.915	0.915
Fixed Effects	✓	✓	✓	✓
Year Effects	✓	✓	✓	✓
Horizon Effects	✓	✓	✓	✓

Notes: The dependent variable is the GDP growth rate forecast error, where the error is defined as the difference between the forecast and the outcome. The estimated equation is (17), where the $Government_i$ indicator has been replaced with the agencies specified in the table. Standard errors robust to two-way clustering at the individual and the horizon-by-year levels based on 999 wild bootstrap repetitions at the individual level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

6 Robustness Checks

This section elaborates on some concerns related to our empirical strategy and performs a battery of robustness checks to validate our results. We start by validating our use of private forecasters as a suitable control group for the government's forecasts. Cipullo and Reslow (2021) show that private forecasters with stakes in and influence over the outcome of a referendum release biased forecasts to affect voters' beliefs. In principle, in the case of general elections, forecasters may have either partisan incentives—for instance, forecasters in the financial sector may prefer a conservative government for corporate tax motives—or they may lean in support of the incumbent in light of valuable connections that have been implemented during the term. In what follows, we show that, in a setup of general elections, private forecasts do not have strong electoral incentives and hence represent a suitable control group. First, to exclude the possibility that forecasters systematically support the incumbent government or the challenger candidate, we estimate (17) for the different types of forecasters in our data sets. The results, presented in Figure A5 in the Appendix, show that, for all countries in our sample and forecasters of all industries, campaign periods are not associated with a change in forecasting behavior.

Second, we exclude the possibility of partisan bias by estimating how the forecasts released by different types of institutions evolve during campaigns depending on the party affiliation of the incumbent government.[40] We present

[40]For the United States, the label *Left* refers to a Democratic president. For Sweden, the *Left* indicator is equal to 1 if the center-left Social Democratic Party (*Socialdemokraterna*) is in the governing coalition. We cannot perform this analysis for the U.K. since the observations

the results from this exercise in Figure A6 in the Appendix. The results exclude the possibility that our control group systematically supports or penalizes incumbents approaching an election because of party affiliation.

Third, we replicate our main results for each country by excluding one forecaster at a time from the institutions in the control group. This exercise guarantees that our findings do not depend on the presence of potential outliers among private forecasters. We present the estimated coefficients in Figure A7 in the Appendix. Some individual forecasters are more pessimistic while others are more optimistic, so removing them one at a time moves the estimated coefficients, but not substantially. Therefore, we conclude that our estimated coefficients from Table 1 are not driven by any outliers since they always lie at the center of the distribution.

As described in Section 4, one of our identifying assumptions reflects the standard parallel trends assumption of difference-in-differences models. However, a formal test is not feasible in our context due to the fuzzy definition of campaign periods within the election year and the repetition of the campaign treatment over time. Notwithstanding, we propose three tests to alleviate the concerns about the fulfillment of the parallel trends assumption.

First, we provide in Figure A8 an estimation where we show the government behavior over the entire election cycle.[41] More specifically, we interact the government indicators with a set of dummies capturing all years of the election cycle as well pre-election and post-election indicators for the election year. From the figure, we confirm our main results for the U.S. since we find no significant bias during campaign periods in the current-year forecasts. Furthermore, the bias is the same as for non-government forecasters in off-election years. In the next-year forecasts, we again confirm our main results of a campaign period bias and no significant different behavior in other periods. For Sweden, we gain additional insights compared to our main results. In both the estimation using the current-year forecasts and the estimation using the next-year estimates, we see clear evidence of a phase-out of the electoral cyclicality (see, e.g., Cipullo & Reslow, 2021). The bias is also present in the months after the election, suggesting that the government slowly revises its forecasts to the unbiased reality. In the estimation using the current-year forecasts, we find that the off-election year behavior is the same as non-government forecasters. In the estimation using next-year forecasts, the figure suggests that the government might be slightly more optimistic during the two years following the election and somewhat more pessimistic in the third year.

In light of the insights from Figure A8, we provide a second test. We show in Table A3 that the results are very similar by estimating equation (17) with an indicator equal to one for all forecasts released during the election year instead

included in our sample before the OBR outsourcing refers to years in which the center-left Labour Party was in government.

[41] We are unable to perform this estimation for the U.K. since it does not have a well-defined election cycle.

of the observations released approaching the election. Compared to equation (17), the identifying variation in this test suppresses the comparison between forecasts released just before and just after an election within the same year. The rationale behind this test is that it is unclear ex-ante whether observations just after an election are good counterfactuals since they can (as indicated in Figure A8) reflect a phase-out of the electoral cyclicality. The results presented in Table A3 show that the estimated coefficients using this specification are consistent with a phase-out. For the United States and Sweden, which have elections late in the year, the coefficients are in line with the ones presented in Table 1. The results for the United Kingdom, as expected, are more attenuated with this approach, given that the elections are held much earlier in the year.

In our main specification, the campaign is defined as starting in January during the election year. This choice of starting point is somewhat arbitrary. Therefore, we perform a sensitivity analysis in which we manipulate the starting point of the campaign. In Figure A9 in the Appendix, we plot the estimated coefficients for different campaign-starting months for Sweden and the United Kingdom, ranging from January of the election year to the month before the election date.[42] We do not find any difference between the estimated coefficients in our main specification and the ones reported in Figure A9 for the U.K. and the forecasts for GDP growth in the next year in Sweden. However, for Sweden, the estimated coefficients relative to the forecasts for the current year are slightly more fragile to this test. More specifically, β is significantly different from zero only if the definition of campaign includes either the entire calendar year or only the final month before the election. Nevertheless, the evidence that the largest coefficient is estimated when the focus is restricted to the final month approaching the election is firmly in line with our suggested mechanism.

7 Concluding Remarks

Voters are perfectly informed about neither the current and future states of the economy nor the ability of the incumbent government. Macroeconomic forecasts are useful to update voters' beliefs about the economy and the ability of the incumbent politician. In this paper, we document the existence of political forecast cycles. Our theoretical framework allows us to study the relationship between voters and the forecasting function of the government. The incumbent politician aims to gain electoral advantages by releasing biased GDP growth forecasts to the public ahead of elections. Rational voters discount the bias and form expectations about the ability of the incumbent. In equilibrium, the incumbent politician does not gain any advantage from the bias yet faces a utility loss due to her estimates' low accuracy. This Pareto

[42] We cannot perform this check for the U.S. since our data is biannual and we only observe one forecast before the election date.

inefficiency rests crucially on the assumption of rational voters who expect politicians to bias their releases. Suppose voters instead are naive, as it may be more conservative to think in the absence of previous rigorous evidence about this government behavior. In this case, the bias will be effective and foster the incumbent's equilibrium reelection probability, resulting in an incumbency advantage. The empirical results confirm key model predictions and disclose electoral cycles in government forecasts. Specifically, governments overestimate short-term GDP growth by 10 to 13 percent during campaign periods. Furthermore, the bias is larger when the incumbent government is not term-limited or constrained by a parliament led by the opposition. We also find that the election timing and amount of available information determine the size of the bias at different forecast horizons.

Biased forecasts pose a problem if voters fail to account for the electoral incentives. In addition to a potential loss in voter welfare, biased forecasts could also damage the economy in a broader perspective. Firms update their beliefs when presented with information about forecasts from professionals. In addition, firms' forecasts are associated with their investment and employment decisions. Hence, the bias can result in inefficient firm and household decisions.

The direct implications of our results are twofold. On the one hand, voters should consider that, when releasing forecasts prior to an election, the government may try to influence its survival probability. On the other hand, outsourcing the government's forecasting function may represent a tool to constrain future governments' behavior. However, according to our empirical findings, the practical implementation of such outsourcing is still far from a perfect commitment device.

The policy implication echoes those from the central bank independence literature. One of the main points of Rogoff's theory of the conservative central banker (Rogoff, 1985) is that independence reduces "political induced" variability and bias. Separating monetary policy from electoral incentives may serve to shield the economy from political business cycles by removing pre-election manipulation of monetary policy as in the Nordhaus (1975) model or reducing partisan shocks to policy following elections as in the Alesina (1988) model. Likewise, in the case of government forecasts, the public would benefit from receiving forecasts also from independent agencies with the sole purpose of providing unbiased estimates.

References

Akhmedov, A., & Zhuravskaya, E. (2004). Opportunistic Political Cycles: Test in a Young Democracy Setting. *The Quarterly Journal of Economics*, *119*(4), 1301–1338.

Alesina, A. (1988). Macroeconomics and politics. *NBER macroeconomics annual*, *3*, 13–52.

Alesina, A., Cohen, G. D., & Roubini, N. (1992). Macroeconomic Policy and Elections in OECD Democracies. *Economics & Politics*, *4*(1), 1–30.

Alesina, A., & Paradisi, M. (2017). Political Budget Cycles: Evidence from Italian Cities. *Economics & Politics*, *29*(2), 157–177.

Alt, J. E., & Lassen, D. D. (2006). Transparency, Political Polarization, and Political Budget Cycles in OECD Countries. *American Journal of Political Science*, *50*(3), 530–550.

Altig, D., Baker, S. R., Barrero, J. M., Bloom, N., Bunn, P., Chen, S., Davis, S. J., Leather, J., Meyer, B. H., Mihaylov, E., Mizen, P., Parker, N. B., Renault, T., Smietanka, P., & Thwaites, G. (2020). *Economic Uncertainty Before and During the COVID-19 Pandemic* (NBER Working Paper Series No. 27418).

Andersson, M. K., Aranki, T., & Reslow, A. (2017). Adjusting for Information Content when Comparing Forecast Performance. *Journal of Forecasting*, *36*(7), 784–794.

Bagues, M., & Esteve-Volart, B. (2016). Politicians' Luck of the Draw: Evidence from the Spanish Christmas Lottery. *Journal of Political Economy*, *124*(5), 1269–1294.

Balke, N. S. (1990). The Rational Timing of Parliamentary Elections. *Public Choice*, *65*(3), 201–216.

Bartolini, D., & Santolini, R. (2009). *Fiscal rules and the opportunistic behaviour of the incumbent politician: evidence from Italian municipalities* (CESifo Working Paper Series No. 2605).

Besley, T., & Case, A. (1995a). Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits. *The Quarterly Journal of Economics*, *110*(3), 769–798.

Besley, T., & Case, A. (1995b). Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition. *American Economic Review*, *85*(1), 22–45.

Besley, T., Montalvo, J. G., & Reynal-Querol, M. (2011). Do Educated Leaders Matter? *The Economic Journal*, *121*(554), F205–227.

Bloom, N. (2014). Fluctuations in Uncertainty. *Journal of Economic Perspectives*, *28*(2), 153–76.

Bloom, N., Bond, S., & Van Reenen, J. (2007). Uncertainty and Investment Dynamics. *The Review of Economic Studies*, *74*(2), 391–415.

Bohn, F., & Veiga, F. J. (2020). Political forecast cycles. *European Journal of Political Economy*.

Bomberger, W. A. (1996). Disagreement as a Measure of Uncertainty. *Journal of Money, Credit and Banking*, *28*(3), 381–392.

Brender, A., & Drazen, A. (2013). Elections, Leaders, and the Composition of Government Spending. *Journal of Public Economics*, *97*, 18–31.

Brown, C. O., & Dinc, I. S. (2005). The Politics of Bank Failures: Evidence from Emerging Markets. *The Quarterly Journal of Economics*, *120*(4), 1413–1444.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, *90*(3), 414–427.

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

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

Cipullo, D., & Reslow, A. (2021). Biased Forecast to Affect Voting Decisions? The "Brexit" Case. In A. Reslow (Ed.), *Electoral Incentives and Information Content in Macroeconomic Forecasts*. Department of Economics, Uppsala University.

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

Coibion, O., Gorodnichenko, Y., & Kumar, S. (2018). How Do Firms Form Their Expectations? New Survey Evidence. *American Economic Review*, *108*(9), 2671–2713.

Dam, L., & Koetter, M. (2012). Bank Bailouts and Moral Hazard: Evidence from Germany. *The Review of Financial Studies*, *25*(8), 2343–2380.

Donald, S. G., & Lang, K. (2007). Inference with Difference-in-Differences and Other Panel Data. *The Review of Economics and Statistics*, *89*(2), 221–233.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, *65*(2), 135–150.

Drazen, A., & Eslava, M. (2010). Electoral Manipulation via Voter-Friendly Spending: Theory and Evidence. *Journal of Development Economics*, *92*(1), 39–52.

Durante, R., & Knight, B. (2012). Partisan Control, Media Bias, and Viewer Responses: Evidence from Berlusconi's Italy. *Journal of the European Economic Association*, *10*(3), 451–481.

Durante, R., Pinotti, P., & Tesei, A. (2019). The Political Legacy of Entertainment TV. *American Economic Review*, *109*(7), 2497–2530.

Erikson, R. S. (1971). The Advantage of Incumbency in Congressional Elections. *Polity*, *3*(3), 395–405.

Freier, R. (2015). The Mayor's Advantage: Causal Evidence on Incumbency Effects in German Mayoral Elections. *European Journal of Political Economy*, *40*, 16–30.

Frendreis, J., & Tatalovich, R. (2000). Accuracy and Bias in Macroeconomic Forecasting by the Administration, the CBO, and the Federal Reserve Board. *Polity*, *32*(4), 623–632.

Gelman, A., & King, G. (1990). Estimating Incumbency Advantage without Bias. *American Journal of Political Science*, 1142–1164.

Gentzkow, M., Shapiro, J. M., & Sinkinson, M. (2011). The effect of newspaper entry and exit on electoral politics. *American Economic Review*, *101*(7), 2980–3018.

Giugliano, F. (2015). Robert Chote reappointed as chairman of OBR. *Financial Times, September 16, 2015*.

Granger, C. W. (1999). Outline of forecast theory using generalized cost functions. *Spanish Economic Review*, *1*(2), 161–173.

Granger, C. W., & Pesaran, M. H. (2000). Economic and statistical measures of forecast accuracy. *Journal of Forecasting*, *19*(7), 537–560.

Harrington, J. E. J. (1992). The Role of Party Reputation in the Formation of Policy. *Journal of Public Economics*, *49*(1), 107–121.

Hayo, B., & Neumeier, F. (2012). Leaders' Impact on Public Spending Priorities: The case of the German Laender. *Kyklos*, *65*(4), 480–511.

Jones, B. F., & Olken, B. A. (2005). Do Leaders Matter? National Leadership and Growth Since World War II. *The Quarterly Journal of Economics*, *120*(3), 835–864.

Kayser, M. A. (2005). Who Surfs, Who Manipulates? The Determinants of Opportunistic Election Timing and Electorally Motivated Economic Intervention. *The American Political Science Review*, *99*(1), 17–27.

Knight, B., & Tribin, A. (2018). The Limits of Propaganda: Evidence from Chavez's Venezuela. *Journal of the European Economic Association*, *17*(2), 567–605.

Krause, G. A. (2006). Beyond the Norm: Cognitive Biases and the Behavioral Underpinnings of US federal Agency Macroeconomic Forecasts. *Rationality and Society*, *18*(2), 157–191.

Krause, G. A., Lewis, D. E., & Douglas, J. W. (2006). Political Appointments, Civil Service Systems, and Bureaucratic Competence: Organizational Balancing and Executive Branch Revenue Forecasts in the American States. *American Journal of Political Science*, *50*(3), 770–787.

Laster, D., Bennett, P., & Geoum, I. S. (1999). Rational Bias in Macroeconomic Forecasts. *The Quarterly Journal of Economics*, *114*(1), 293–318.

Lee, D. S. (2008). Randomized experiments from non-random selection in US House elections. *Journal of Econometrics*, *142*(2), 675–697.

MacRae, C. D. (1977). A Political Model of the Business Cycle. *Journal of Political Economy*, *85*(2), 239–263.

Marinovic, I., Ottaviani, M., & Sørensen, P. N. (2013). Chapter 12 – Forecasters' Objectives and Strategies. In G. Elliott & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (pp. 690–72). Elsevier.

Markus, G. B. (1988). The Impact of Personal and National Economic Conditions on the Presidential Vote: A Pooled Cross-Sectional Analysis. *American Journal of Political Science*, 137–154.

Markus, G. B. (1992). The Impact of Personal and National Economic Conditions on Presidential Voting, 1956–1988. *American Journal of Political Science*, 829–834.

Meriläinen, J. (2020). *Politician Quality, Ideology, and Fiscal Policy*.

Müller, K. (2020). *Electoral Cycles in Macroprudential Regulation*.

Nordhaus, W. D. (1975). The Political Business Cycle. *The Review of Economic Studies*, *42*(2), 169–190.

Perotti, R., & Kontopoulos, Y. (2002). Fragmented Fiscal Policy. *Journal of Public Economics*, *86*(2), 191–222.

Persson, T., Roland, G., & Tabellini, G. (1997). Separation of Powers and Political Accountability. *The Quarterly Journal of Economics*, *112*(4), 1163–1202.

Persson, T., Roland, G., & Tabellini, G. (1998). Towards micropolitical foundations of public finance. *European Economic Review*, *42*(3), 685–694.

Persson, T., Roland, G., & Tabellini, G. (2000). Comparative Politics and Public Finance. *Journal of Political Economy*, *108*(6), 1121–1161.

Persson, T., & Tabellini, G. E. (2002). *Political Economics: Explaining Economic Policy*. MIT press.

Picchio, M., & Santolini, R. (2020). Fiscal rules and budget forecast errors of Italian municipalities. *European Journal of Political Economy*, *64*, 101921.

Ralph, A. (2015). Treasury has sought to meddle with OBR forecasts. *The Times, September 14, 2015*.

Repetto, L. (2018). Political Budget Cycles with Informed Voters: Evidence from Italy. *The Economic Journal*, *128*(616), 3320–3353.

Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. *The quarterly journal of economics*, *100*(4), 1169–1189.

Rogoff, K. (1990). Equilibrium Political Budget Cycles. *American Economic Review*, *80*(1), 21–36.

Rogoff, K., & Sibert, A. (1988). Elections and Macroeconomic Policy Cycles. *The Review of Economic Studies*, *55*(1), 1–16.

Saporiti, A., & Streb, J. M. (2008). Separation of Powers and Political Budget Cycles. *Public Choice*, *137*(1-2), 329–345.

Schuknecht, L. (1996). Political Business Cycles and Fiscal Policies in Developing Countries. *Kyklos*, *49*(2), 155–170.

Shi, M., & Svensson, J. (2006). Political Budget Cycles: Do they differ across countries and why? *Journal of Public Economics*, *90*(8-9), 1367–1389.

Tanaka, M., Bloom, N., David, J. M., & Koga, M. (2019). Firm Performance and Macro Forecast Accuracy. *Journal of Monetary Economics*.

Veiga, F. J., Veiga, L. G., & Morozumi, A. (2017). Political Budget Cycles and Media Freedom. *Electoral Studies*, *45*, 88–99.

Appendix

A.1 Model

Politician's economy expectations: weights in equation (4),

$$m_{1,T} = \frac{var(\widetilde{y}_2)cov(y_T,\widetilde{y}_1) - cov(\widetilde{y}_1,\widetilde{y}_2)cov(y_T,\widetilde{y}_2)}{var(\widetilde{y}_1)var(\widetilde{y}_2) - cov(\widetilde{y}_1,\widetilde{y}_2)^2} \tag{A1}$$

$$m_{2,T} = \frac{var(\widetilde{y}_1)cov(y_T,\widetilde{y}_2) - cov(\widetilde{y}_1,\widetilde{y}_2)cov(y_T,\widetilde{y}_1)}{var(\widetilde{y}_1)var(\widetilde{y}_2) - cov(\widetilde{y}_1,\widetilde{y}_2)^2} \tag{A2}$$

$$m_{0,T} = 1 - m_{1,T} - m_{2,T}, \tag{A3}$$

where

$$var(\widetilde{y}_T) = var(y_T) + var(\varepsilon_T) \tag{A4}$$

$$= \lambda^2\tau_\eta^{-1} + \tau_\nu^{-1} + \tau_{\varepsilon_{T,t}}^{-1} \tag{A5}$$

$$cov(\widetilde{y}_1,\widetilde{y}_2) = \lambda^2\tau_\eta^{-1} \tag{A6}$$

$$cov(y_T,\widetilde{y}_T) = \lambda^2\tau_\eta^{-1} + \tau_\nu^{-1} \tag{A7}$$

$$cov(y_1,\widetilde{y}_2) = \lambda^2\tau_\eta^{-1} \tag{A8}$$

$$cov(y_2,\widetilde{y}_1) = \lambda^2\tau_\eta^{-1}. \tag{A9}$$

Voters' ability expectations: weights in equation (10),

$$\gamma_1 = \frac{var(\widetilde{F}_2)cov(\widetilde{F}_1,\lambda\eta^I) - cov(\widetilde{F}_1,\widetilde{F}_2)cov(\widetilde{F}_2,\lambda\eta^I)}{var(\widetilde{F}_1)var(\widetilde{F}_2) - cov(\widetilde{F}_1,\widetilde{F}_2)^2} \tag{A10}$$

$$\gamma_2 = \frac{var(\widetilde{F}_1)cov(\widetilde{F}_2,\lambda\eta^I) - cov(\widetilde{F}_1,\widetilde{F}_2)cov(\widetilde{F}_1,\lambda\eta^I)}{var(\widetilde{F}_1)var(\widetilde{F}_2) - cov(\widetilde{F}_1,\widetilde{F}_2)^2} \tag{A11}$$

$$\gamma_0 = 1 - \gamma_1 - \gamma_2, \tag{A12}$$

where

$$var(\widetilde{F}_T) = var(m_{1,T}\widetilde{y}_1 + m_{2,T}\widetilde{y}_2 + e_T) \tag{A13}$$

$$= m_{1,T}^2 var(\widetilde{y}_1) + m_{2,T}^2 var(\widetilde{y}_2) + 2m_{1,T}m_{2,T}cov(\widetilde{y}_1,\widetilde{y}_2) + \tau_e^{-1} \tag{A14}$$

$$cov(\widetilde{F}_1,\widetilde{F}_2) = cov(m_{1,1}\widetilde{y}_1 + m_{2,1}\widetilde{y}_2\ ,\ m_{1,2}\widetilde{y}_1 + m_{2,2}\widetilde{y}_2) \tag{A15}$$

$$= m_{1,1}m_{1,2}var(\widetilde{y}_1) + m_{2,1}m_{2,2}var(\widetilde{y}_2)$$

$$+ (m_{1,1}m_{2,2} + m_{2,1}m_{1,2})cov(\widetilde{y}_1,\widetilde{y}_2) \tag{A16}$$

$$cov(\widetilde{F}_T,\lambda\eta^I) = cov(m_{1,T}\widetilde{y}_1 + m_{2,T}\widetilde{y}_2,\lambda\eta^I) \tag{A17}$$

$$= m_{1,T}cov(\widetilde{y}_1,\lambda\eta^I) + m_{2,T}cov(\widetilde{y}_2,\lambda\eta^I) \tag{A18}$$

$$= (m_{1,T} + m_{2,T})var(\lambda\eta^I) \tag{A19}$$

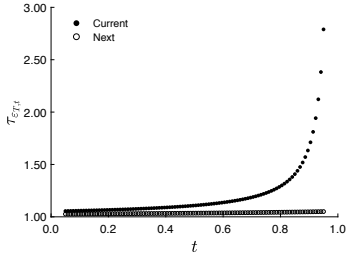
$$= (m_{1,T} + m_{2,T})\lambda^2\tau_\eta^{-1}. \tag{A20}$$

Table A1. *Calibration of Model Parameters*

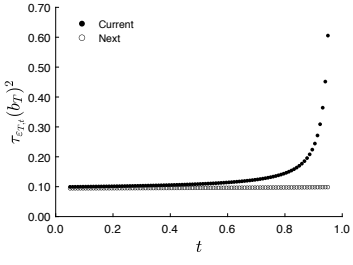
t	λ	R	$\bar{\eta}$	τ	τ_η	τ_v	τ_e
0.50	0.99	3.00	2.00	1.10	0.20	0.70	0.70

Notes: Calibration of the exogenous parameters in the theoretical model. We set $\bar{\eta} = 2$ and $\lambda = 0.99$ to match a benchmark average annual GDP growth of around 2 percent. Based on this calibration we provide simulated example data in Figures A2 and A3.

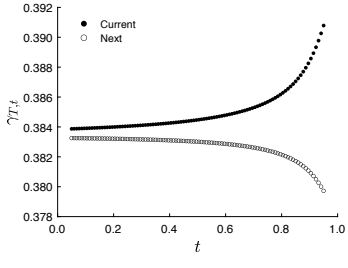
(a) Signal Precision



(b) Cost of Bias



(c) Impact on Voter Beliefs



(d) Bias Ratio

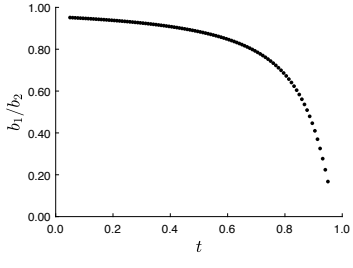


Figure A1. Election Timing and Model Dynamics

Notes: Model predictions based on the calibration presented in Table A1. In graph A1b we assume an arbitrary bias level of $b_1 = b_2 = 0.3$. In graph A1d we show $\frac{b_1}{b_2} = \frac{\gamma_1}{\gamma_2}\frac{\tau_{\varepsilon_{2,t}}}{\tau_{\varepsilon_{1,t}}}$.

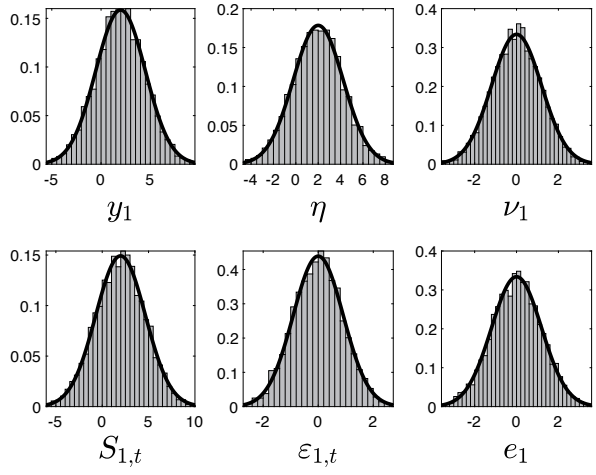


Figure A2. Simulated Data

Notes: The graph is based on the calibration presented in Table A1. Each histogram represent simulated data based on 5,000 repetitions, while each curves represent the underlying probability density function. Bars have been normalized to match the PDF.

(a) Ability Expectations *(b)* Reelection Probabilities

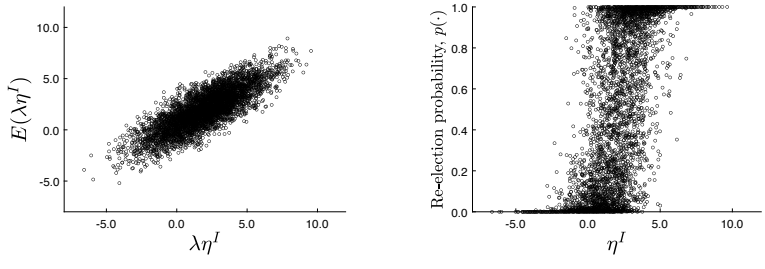


Figure A3. Ability Expectations and Reelection Probabilities

Notes: The graph is based on 5,000 simulations using the calibration presented in Table A1. Graph (a) shows, on the y-axis, voters' expectations about the ability of the incumbent politician following (10), and, on the x-axis, the true innate ability of the incumbent. Graph (b) shows, on the y-axis, the incumbent politician's reelection probability, and, on the x-axis, the politician's true ability.

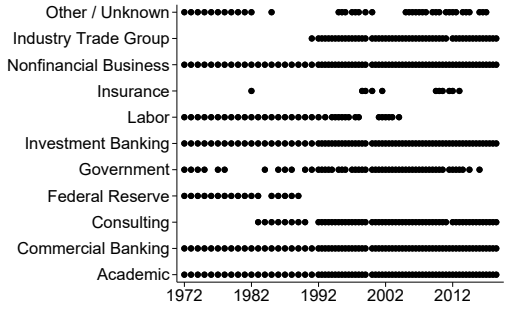
A.2 Data

Table A2. *Descriptive Statistics*

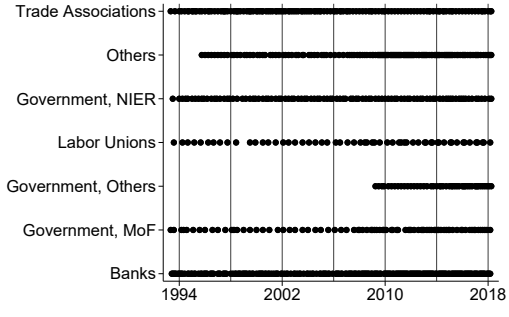
	Obs.	Mean	SD	Min	Max
Panel A. United States					
Forecast Error (Current Year)	3,082	0.05	0.88	-12.21	10.73
Forecast Error (Next Year)	3,057	0.34	1.90	-6.80	11.46
Government	3,082	0.03	0.17	0	1
Election Year	3,082	0.51	0.50	0	1
Campaign	3,082	0.36	0.48	0	1
Year	3,082	1995.45	12.75	1972	2018
Quarter	3,082	2.61	0.92	2	4
Presidential Election	3,082	0.48	0.50	0	1
Term-Limited Incumbent	3,082	0.34	0.48	0	1
President Not Aligned with Congress	3,082	0.70	0.46	0	1
Panel B. Sweden					
Forecast Error (Current Year)	1,028	-0.02	0.87	-3.26	4.04
Forecast Error (Next Year)	1,034	0.18	1.78	-5.46	7.54
Government	1,042	0.08	0.27	0	1
Election Year	1,042	0.28	0.45	0	1
Campaign	1,042	0.17	0.37	0	1
Year	1,042	2008.05	7.05	1994	2018
Month	1,042	6.97	3.47	1	12
Panel C. United Kingdom (Pre reform)					
Forecast Error (Current Year)	3,471	-0.00	0.67	-2.75	4.35
Forecast Error (Next Year)	3,271	0.41	1.77	-3.69	7.55
Government	3,477	0.01	0.08	0	1
Election Year	3,477	0.18	0.39	0	1
Campaign	3,477	0.09	0.28	0	1
Year	3,477	2003.73	3.66	1998	2010
Month	3,477	6.32	3.57	1	12
Panel D. United Kingdom (Post reform)					
Forecast Error (Current Year)	2,080	0.05	0.52	-1.65	2.22
Forecast Error (Next Year)	1,869	0.20	0.92	-3.11	3.02
Government	2,080	0.01	0.09	0	1
Election Year	2,080	0.31	0.46	0	1
Campaign	2,080	0.08	0.27	0	1
Year	2,080	2013.84	2.34	2010	2018
Month	2,080	6.49	3.56	1	12

Notes: See Section 3 for an explanation of the data. For Sweden, the Government indicator refers to the Ministry of Finance, and, for the U.K., the Government indicator refers to H.M. Treasury in the pre-reform periods and the Office for Budget Responsibilities (OBR) in the post-reform periods.

(a) United States



(b) Sweden



(c) United Kingdom

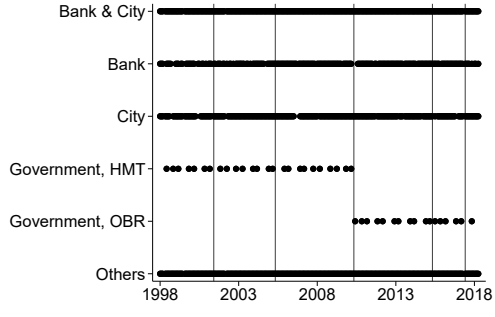


Figure A4. Forecast Frequency by Institution Type

Notes: Each dot shows the existence of at least one forecast at a given survey release date for a specific forecaster or type. For Sweden, our main government definition is the Ministry of Finance (MoF), while Government Others refer to the PES, NDO and NFMA. For Sweden and the U.K., the vertical lines mark the election dates.

A.3 Robustness

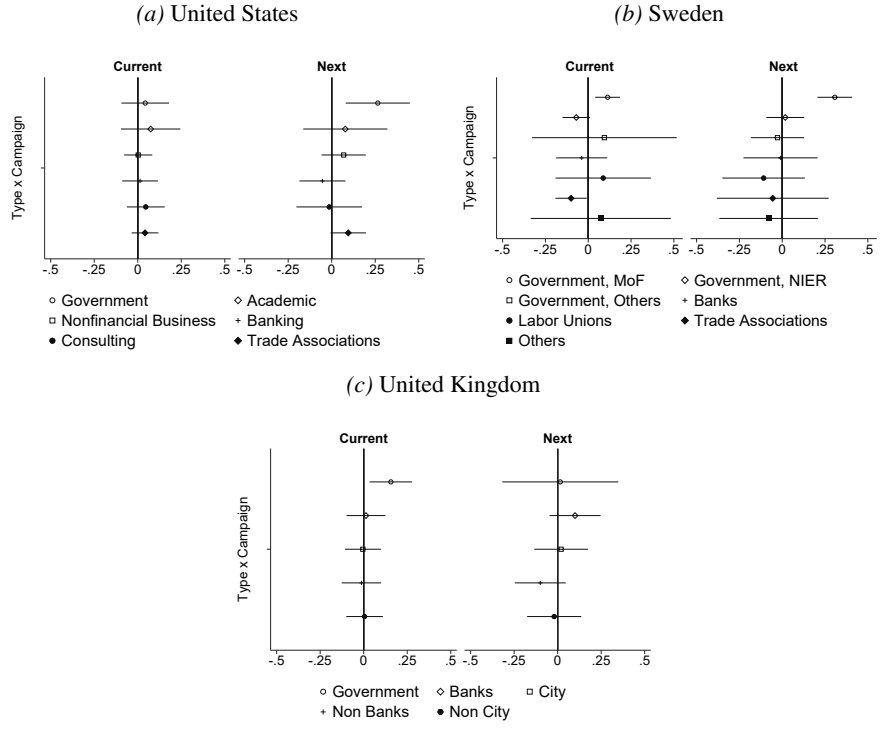
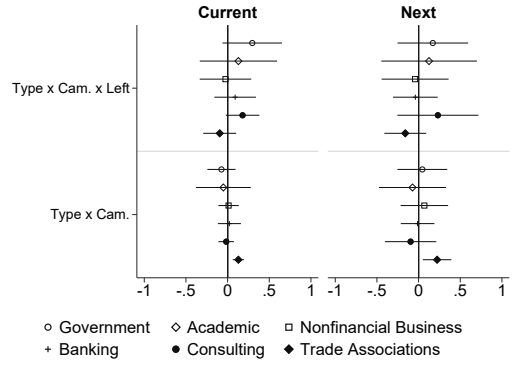


Figure A5. Estimated Coefficients for the *Control* Forecasters

Notes: The dependent variable is the GDP growth rate forecast error, where the error is defined as the difference between the forecast and the outcome. The markers show the estimated coefficients for β by estimating (17), where the $Government_i$ indicator has been replaced with each of the industry-specific markers reported in the label. The government markers correspond to the results presented in Section 5. 95% confidence intervals are based on standard errors robust to two-way clustering at the forecaster and the survey levels. For Sweden, 95% confidence intervals are based on standard errors robust to two-way clustering at the individual and the horizon-by-year levels based on 999 wild bootstrap repetitions at the individual level.

(a) United States



(b) Sweden

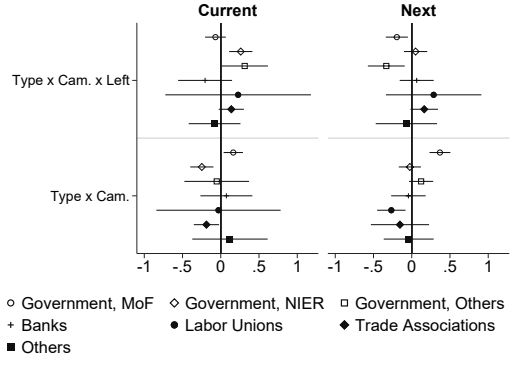


Figure A6. Estimated Coefficients for the Control Forecasters—By Identity of the Ruling Party

Notes: The dependent variable is the GDP growth rate forecast error, where the error is defined as the difference between the forecast and the outcome. The estimated equations are versions of equation (17), in which the $Government_i$ indicator has been replaced with each of the industry-specific markers reported in the label, and the $Government_i \times Campaign_{h,T}$ has been interacted with an indicator for the political affiliation of the government. For the U.S., the label *Left* refers to a Democratic president and the $Government_i$ indicator. For Sweden, the *Left* indicator is equal to 1 if the center-left Social Democratic (*Socialdemokraterna*) party is in the governing coalition. Markers show the estimated coefficients from an interaction between the type of forecaster (defined by the label) with the campaign variable and the Left variable. 95% confidence intervals are based on standard errors robust to two-way clustering at the forecaster and the survey levels. For Sweden, 95% confidence intervals are based on standard errors robust to two-way clustering at the individual and the horizon-by-year levels based on 999 wild bootstrap repetitions at the individual level.

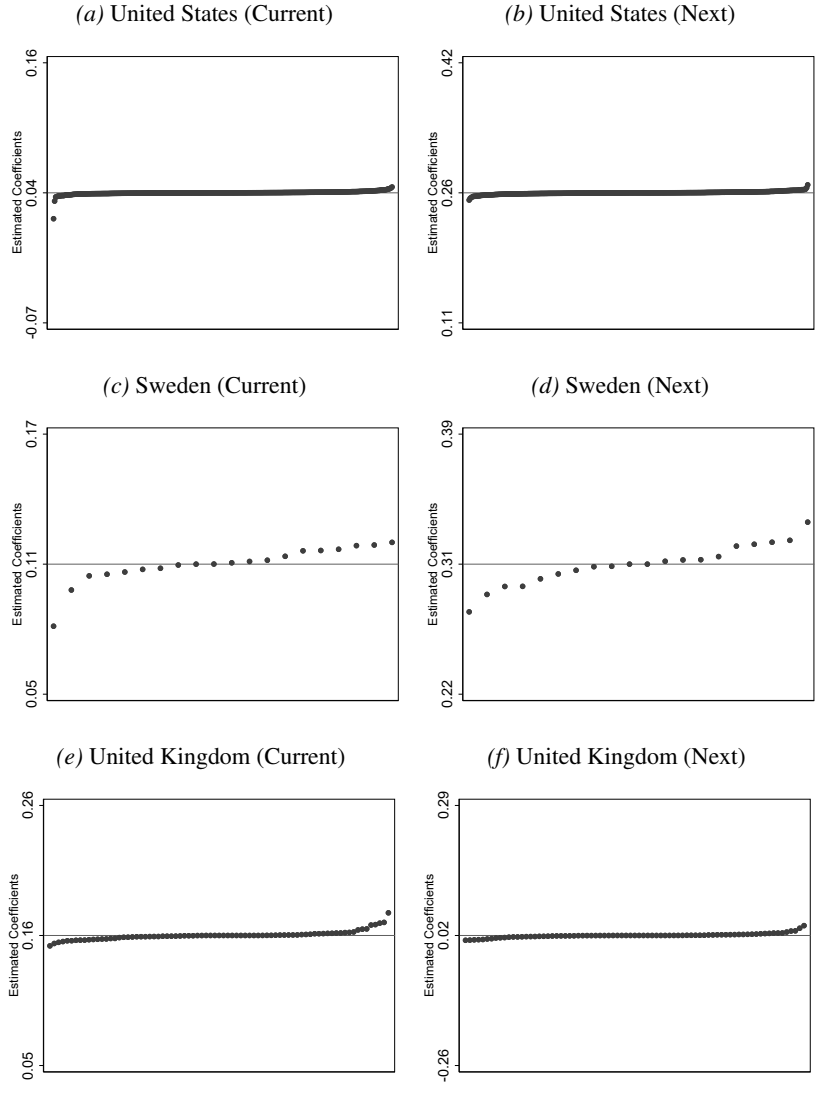


Figure A7. Removing One Forecaster at a Time from the Control Group

Notes: Each dot represents an estimation of equation (17) by removing one forecaster at a time from the group of control forecasters. The estimations have been sorted from lowest to highest. Solid lines correspond to the main results presented in columns (2) and (4) of Table 1. The upper limit on the y-axis of each graph represents the estimated coefficient from Table 1 plus the estimated standard error. Likewise, the lower limit represents the estimated coefficient from Table 1 minus the estimated standard error.

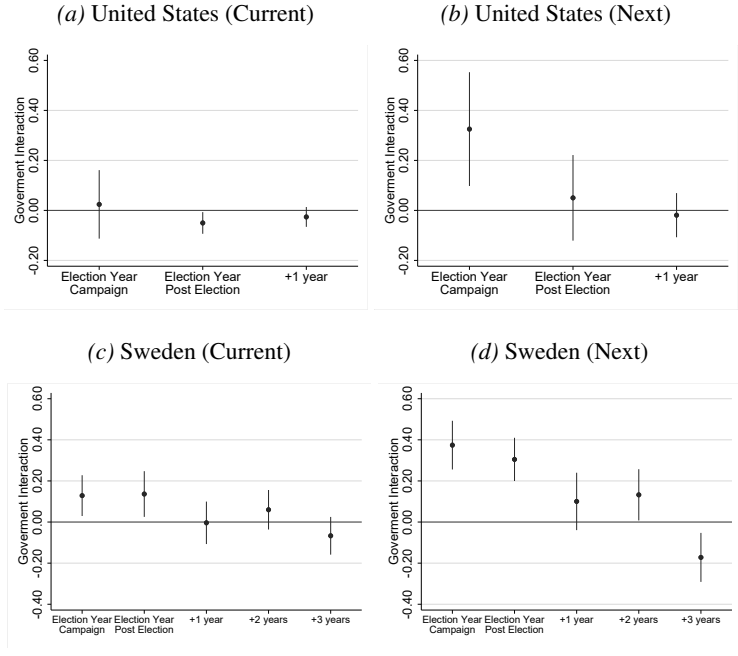


Figure A8. Estimation of Election Cycle

Notes: The horizontal axis refers to the election cycle where the election year has been split into pre-(campaign) and post-election. The "+1 year" then refers to the year after the election year. The horizontal axis refers to the origin for when the forecast was released, while the graph titles provide the forecast target (current or next year). The estimated equation is $E_{i,h,T} = \theta + \delta_h + \mu_T + \alpha_c cam_{h,T} + \alpha_p post_{h,T} + \beta_c gov_i \times cam_{h,T} + \beta_p gov_i \times post_{h,T} + \sum_{j=1}^{J} \beta_j gov_i \times \mathbb{1}(election\ year + j)_{h,T} + \epsilon_{i,h,T}$, where $J = 1$ for the U.S. and $J = 3$ for Sweden. $post_{h,T}$ is an indicator taking the value 1 when the forecast is released within an election year but after the election, and $cam_{h,T}$ is an indicator taking the value 1 when the forecast is released within an election year but before the election. We are unable to perform the same estimation for the U.K. since it does not have a well-defined election cycle. The coefficients displayed in the graphs correspond to the β coefficients in the estimated equation. Standard errors robust to two-way clustering at the forecaster and the survey levels are reported at the 5% significance level. In the case for Sweden, the standard errors are based on 999 wild bootstrap repetitions at the individual level.

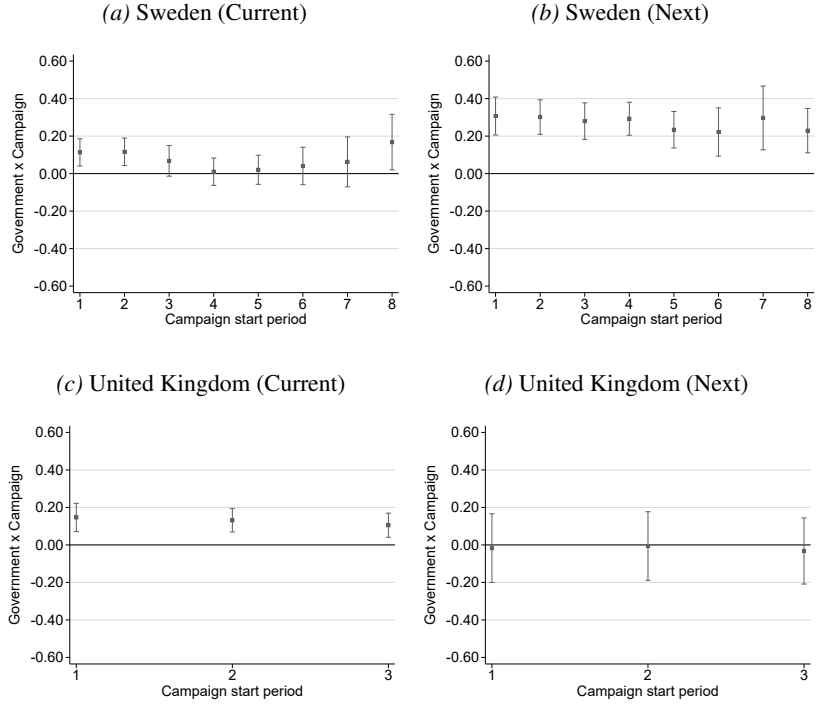


Figure A9. Sensitivity of the Estimates to the Definition of the Campaign Period

Notes: The graphs plot the estimated coefficients and 95% confidence intervals from estimating equation (17) by alternating the definition of the $Campaign_{h,T}$ indicator to include only the months between each marker on the x-axis and the election date. Coefficients at period 1 replicate the results presented in Columns (2) and (4) of Table 1. 95% confidence intervals are based on standard errors robust to two-way clustering at the forecaster and the survey levels. For Sweden, 95% confidence intervals are based on standard errors robust to two-way clustering at the individual and the horizon-by-year levels based on 999 wild bootstrap repetitions at the individual level.

Table A3. *Estimated Election Cycle Bias: Election Year Definition*

	Current		Next	
	(1)	(2)	(3)	(4)
Panel A. United States				
Government × *Election Year*	0.026	0.024	0.255***	0.254***
	(0.046)	(0.047)	(0.062)	(0.063)
Government	-0.026		-0.019	
	(0.022)		(0.051)	
Observations	3,082	3,082	3,057	3,057
R^2	0.454	0.460	0.748	0.750
Panel B. Sweden				
Government × *Election Year*	0.125***	0.134***	0.337***	0.346***
	(0.026)	(0.027)	(0.049)	(0.043)
Government	-0.002		0.019	
	(0.038)		(0.057)	
Observations	1,028	1,028	1,034	1,034
R^2	0.661	0.680	0.907	0.915
Panel C. United Kingdom				
Government × *Election Year*	0.039	0.033	-0.073	-0.076
	(0.050)	(0.061)	(0.090)	(0.115)
Government	0.113***		0.472***	
	(0.028)		(0.051)	
Observations	3,471	3,471	3,271	3,271
R^2	0.582	0.641	0.855	0.901
Fixed Effects		✓		✓
Year Effects	✓	✓	✓	✓
Horizon Effects	✓	✓	✓	✓

Notes: The dependent variable is the GDP growth rate forecast error, where the error is defined as the difference between the forecast and the outcome. In columns (1) and (3) the estimated equation is (18), while in columns (2) and (4) the estimated equation is (17), with, in both cases, the *Campaign* indicator replaced with an *Election Year* indicator equal to 1 for each forecast released during an election year in all estimations. In Panel A. United States, the fixed effects are replaced with type effects. Standard errors robust to two-way clustering at the forecaster and the horizon-by-year levels are in parentheses. In Panel B. Sweden, the standard errors are based on 999 wild bootstrap repetitions at the individual level. *,**,*** represent the 10%, 5%, 1% significance levels.

Essay III

Inefficient Use of Competitors' Forecasts?

Acknowledgments: I would like to thank Mikael Carlsson, Jesper Lindé and Erik Öberg for the extensive feedback, and Daniel Bougt, Tobias Broer, Fabio Canova, David Cesarini, Davide Cipullo, Nils Gottfries, Sofia Hernnäs, Isaiah Hull, Sebastian Jävervall, Luca Repetto, Janne Tukiainen, 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 data used in this paper is copyright protected against redistribution but can be purchased from the provider. Some of the work on this essay was conducted at Sveriges Riksbank. 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.

1 Introduction

Uncertainty poses a problem for economic agents, who all need to make important decisions based on expectations about the future. Forming expectations is not trivial. While the average consumer, investor, voter, and even policymaker is likely to lack the competence to derive optimal predictions, they can obtain helpful information from those who have made it their livelihood to predict the future: *professional forecasters*. Firms update their expectations when exposed to forecasts from professionals (see, e.g., Coibion et al., 2018), and firms' forecasts are positively and significantly associated with their input choices (see, e.g., Tanaka et al., 2019). When economic agents use forecasts in their decision process, they would like those forecasts to be as accurate and efficient as possible.

Even so, professional forecasters also operate in a world with incomplete information. Even though they are highly skilled, they have different skillsets as well as different information sets. Therefore, an essential source of information for a macroeconomic forecaster is the forecasts from competitors. The starting point of this paper is to assess the extent to which macroeconomic forecasters make efficient use of forecasts from their competitors. If forecasts are inefficient, the resulting outcome can be inefficient firm, household, and policy decisions. Nordhaus (1987) presents a useful concept of weak and strong forecast efficiency when evaluating efficiency in macroeconomic forecasts. This paper augments the weak efficiency concept in Nordhaus (1987) to include competitor information in addition to the own lagged information. The paper also deploys a theoretical beliefs updating model with noisy and private information to better understand the forecasters' optimal and actual behavior.

The empirical analysis makes use of high-frequency forecaster firm-level panel data from Consensus Economics. The data used cover ten countries—Canada, France, Germany, Italy, Japan, the Netherlands, Spain, Sweden, the United Kingdom, and the United States—and the data span forecasts made between 1995 and 2017. The dataset is a monthly survey of forecasters who are asked to provide forecasts for the current and next year's annual growth rates. Hence, the forecasts are best characterized as a set of fixed target forecasts where multiple predictions are made for the same outcome at a decreasing forecast horizon. The primary analysis focuses on forecasts for GDP growth and inflation (consumer prices). The data offers a dynamic panel of 591 forecasters. Using this data, it is found that forecasters *underuse* competitor information in their forecasts for current and next year's annual GDP growth and inflation. The results suggest that forecasters put a sub-optimal high weight on their own lagged forecast and too low a weight on their competitors' work. The empirical results also show that the forecasters behave differently at different forecast horizons. As the forecast horizon decreases, when approaching the realization of the outcome, the forecasters increase the attention to their competitors. However, they are still far from an optimal level.

In line with the empirical analysis, a baseline model with noisy and private information predicts that, for a forecaster to minimize forecast errors, it is optimal to pay considerable attention to competitors' work and, consequently, put a high weight on the forecasts from others. The suggestion that it is optimal to pool information from many forecasters has been raised numerous times and is well-known in the forecasting literature (see, e.g., Elliott et al., 2008; Hendry & Clements, 2004; Zarnowitz, 1984). The typical empirical finding is that consensus (average) forecasts are more accurate over time than individual forecasts. The exercise in this paper should not be seen as a test of whether or not it is optimal to pool forecasts. Instead, it is a test of whether or not forecasters adopt this well-known result in their work.[1] It is important to note that, even if forecasters are efficient and optimally use competitor information, it is still optimal for end-users, such as firms, households, and policymakers, to pool from multiple forecasters due to private information that each forecaster holds. Hence, the well-known pooling result should hold regardless.

Two versions of the model are considered: one with fully informed forecasters and one with "rule of thumb" forecasters. In the fully informed model, the forecaster perfectly filters out signals from competitors to produce the (theoretically) most efficient forecast. The rule of thumb model is used to impose limitations on the forecasters' ability. The forecasters' rule of thumb is to use lagged aggregate forecasts instead of filtering out individual signals. This model, hence, is more in line with the empirical strategy.

The two models differ in 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.[2] The models' information structure matches the specific structure of fixed target forecasts used in the empirical analysis. The baseline models do not have any form of strategic trade-off and assume that forecasters' objective is pure forecast error minimization.

In order to explain the observed actual (seemingly inefficient) behavior, the baseline model is extended. The extensions are made to consider alternatives to pure inefficiency. First, the paper rules out one pure cognitive bias by considering forecasters to be overconfident in their ability (signal precision), leading to a behavior response of overweighting private information. It is unlikely that overconfidence can be the sole explanation for the observed behavior. A back-of-the-envelope calculation shows that for overconfidence to explain the observed behavior, forecasters need to think that the precision in their private signal is 100 times more precise than the private signal received by any competitor.

[1] An ongoing example of incorporating the well-known pooling 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).

[2] In this type of model (see, e.g., Lucas, 1973; Sims, 2003; Woodford, 2001), agents continuously update their information sets and their beliefs via signal extraction.

As an alternative, an explicit cost from forecast revisions to capture a credibility loss from making substantial changes is added to the model (see, e.g., Ehrbeck & Waldmann, 1996). Given that forecast smoothing is observed in the data, it is intuitive to consider a backward-looking extension to the model. The extension introduces a trade-off between minimizing the forecast error and minimizing the revisions. However, this paper shows that the revision cost model can only match the observed horizon heterogeneity if it is assumed that the revision cost and the cost from forecast errors have asymmetric horizon discounting.

The main contribution of this paper relates to the literature on the behavior of macroeconomic forecasters. First, the documentation of the *relative* use between forecasters' own lagged forecasts and those of others and how this relationship differs between horizons. The empirical analysis 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 past forecast and more attached to new information when making forecast revisions. The analysis is also related to, for example, Lahiri and Sheng (2008), who estimate a Bayesian learning model with heterogeneity aimed at explaining forecast disagreement and its evolution over horizons. More recent related work includes, among others, Coibion and Gorodnichenko (2015), Dovern et al. (2015), Broer and Kohlhas (2018) and Bordalo et al. (2020), who all estimate how forecasters react to new information in general.[3] It is important to keep in mind that forecasters can be efficient (or inefficient) in using competitor information and still under- or overreact to new private information, or different forms of new public information, as is found in the papers cited above. Second, this paper shows in a noisy information framework with private information that this behavior cannot be explained by overconfidence alone. It also shows that the introduction of asymmetric horizon discounting is crucial for revision costs to explain the observed behavior.

The notion of strategic behavior, in addition to forecast-error minimization, among macroeconomic forecasters is not novel. For example, Laster et al. (1999) introduces a trade-off problem for the forecasters, where they have to balance accuracy 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. Cipullo and Reslow (2021) propose an alternative trade-off where forecasters trade accuracy with electoral incentives. Although strategic behavior is well established in the lit-

[3] 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 than efficient weights on their private information when they forecast corporate earnings.

erature, this paper highlights that, given the observed horizon heterogeneity, strategic behavior is likely to be horizon dependent.

The paper is also related to the work on herd behavior by, for example, Banerjee (1992), but with some differences.[4] In the Banerjee (1992) model, herding is doing what everyone else is doing, even if private information suggests doing something different. In this paper's framework, forecasters still put positive weight on private information and do not explicitly aim to herd. Note that herd behavior, as in the Banerjee (1992) model, would result in the opposite of what is found in this paper. Herd behavior would display inefficient *overuse* of information from others and *underuse* of own information.

2 Data

This paper utilizes high-frequency forecaster firm-level forecast data from a monthly survey by Consensus Economics. The data used in this paper covers ten countries—Canada, France, Germany, Italy, Japan, the Netherlands, Spain, Sweden, the United Kingdom, and the United States—and span forecasts made between 1995 and 2017. The forecasts refer to the current and next year's annual growth rates. For example, the forecasters were asked 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).[5] In order to calculate forecast errors, the actual realizations are needed. The outcome data is collected from the OECD database. For inflation, the latest available numbers are used. For GDP growth, the outcome is given by the first available release (often referred to as real-time data).[6]

Table 1 reports firm number descriptive statistics. The sample consists of 591 forecasting firms.[7] 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 of the forecast. Nevertheless, the correlation between the target country and the origin country is very high. 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

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

[5] For the U.K., the inflation forecasts refer to both RPIX and HICP (CPI). The RPIX covers the entire period, but toward the end of the sample many forecasters produce a forecast for both measures while others change entirely to HICP. All available price forecasts are used.

[6] GDP is subject to large revisions and real-time data is often used when evaluating forecast performance for variables subject to revisions. Real-time GDP growth data is constructed using the OECD revision database.

[7] Some firm name entries in the raw data have been corrected due to spelling and naming errors.

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.

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; most observations are from Germany and the United States, and the fewest are from the Netherlands. The number of observations for GDP growth and inflation is similar in total and for each country. The average GDP growth forecast is 1.9 percent, while the average inflation forecast is 1.8 percent. Tables A1 and A2 in the appendix reveal a slight overestimation of both 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: N refers to the total number of firms in the entire sample, while Mean, Median, Min, and Max refer to firms by survey month. *Source*: Consensus Economics.

Table 2. *Forecast Statistics*

		GDP growth			Inflation		
Country	Obs.	Obs.	Mean	SD	Obs.	Mean	SD
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: Forecasters are asked each month to forecast the current and next year's annual growth rate, resulting in 24 potential forecast origins (horizons) for each year's outcome between 1995 and 2017.

3 Empirical Strategy

3.1 Theoretical Underpinnings

It is well established in the forecasting literature that pooling information from multiple forecasters is beneficial (see, e.g., Elliott et al., 2008; Hendry & Clements, 2004; Zarnowitz, 1984). The typical finding is that consensus (average) forecasts are more accurate over time than individual forecasts. The takeaway from this result is that economic agents are advised to rely on forecasts from multiple sources instead of a single forecaster.

To set ideas, let x_t be the target outcome and F_{ith} the forecast for x_t made by forecaster i, at a horizon of h periods before the realization of x_t. Given this notation, the forecast F_{ith+1} is released one period *before* the release of the forecast F_{ith}. Further assume that the forecasters in each period have access to a private signal $S_{ith} = x_t + \epsilon_{ith}$, where $\epsilon_{ith} \sim N(0, \sigma_\epsilon^2)$ is a random error, as well as some prior knowledge μ_{ith}. The forecasters release to the public a forecast given by $F_{ith} = \mathbb{E}(x_t|\mu_{ith}, S_{ith}) = (1-w)\mu_{ith} + wS_{ith}$, where w is given by optimal Bayesian updating.[8] In a repeated forecast setting where a forecaster releases multiple forecasts for each x_t, the prior can be expressed as $\mu_{ith} = F_{ith+1}$. Hence, the prior is the forecast released in the previous period. The rationale behind this assumption is that the previous forecast should summarize all previously known information in an optimal way. A forecast at horizon h is then given by the forecast at horizon $h+1$ updated with the new private signal. In the case of infinite many forecasters, the consensus (average) forecast can be expressed as $\overline{F}_{th} = \int_0^1 F_{ith} di$. In this simple model, $\overline{F}_{th} = \int_0^1 F_{ith} di = x_t$ represents a perfect forecast since the noise in the signal, ϵ_{ith}, cancels out when aggregating over the infinite amount of forecasters.

Even though the assumption of an infinite amount of forecasters is unrealistic, it illustrates an understanding of the well-known pooling result. Over time, consensus forecasts are more accurate than individual forecasts since the error in the noisy information is mitigated. Hence, $\overline{F}_{th}$ is expected to be a better forecast than F_{ith}. This well-known result prompts third-party users of forecasts, such as firms, voters, and policymakers, to collect information from multiple sources. The same advice should also hold for the forecasters themselves. Going back to the expression of the prior knowledge in the forecaster's updating function, it was assumed above that $\mu_{ith} = F_{ith+1}$. However, given that F_{ith+1} at horizon h is public knowledge, and that we established that $\overline{F}_{th}$ is expected to be a better forecast than F_{ith}, the optimal prior should be $\mu^*_{ith} = \overline{F}_{th+1}$. Hence, the optimal prior should then be a weighted average of the own lagged forecast as well as the lagged forecast of the competitors.

[8] In order to simplify the model, it is assumed that there is no public signal. This assumption does not impact the main takeaway from the example.

3.2 Estimation

Nordhaus (1987) presents a useful concept of weak and strong forecast efficiency to assess the efficiency in forecasts. 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.[9] Nordhaus (1987) then argues that a weaker concept, where instead of the full information set I_{ith}, only the forecaster's own past forecast should be included. 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 forecast produced in the previous period. 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.

This paper continues to assume, as in Nordhaus (1987), that forecasters know their own past forecasts and further add the assumption that they know the past (average) forecasts of their competitors, $Z_{ith+1} = \frac{1}{n-1} \sum_{j \neq i} F_{jth+1}$.[10] Hence, Z_{ith+1} is the leave-out mean for each forecaster i. This results in an augmented weak efficiency definition of minimizing $\mathbb{E}_{ith}[(x_t - F_{ith})^2 | F_{ith+1}, Z_{ith+1}]$, which implies that $\mathbb{E}_{ith}[x_t - F_{ith} | F_{ith+1}, Z_{ith+1}] = 0$. 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.[11] $\beta_2 > 0$ 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 forecast error and 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 $\beta = \gamma - \alpha$.

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

[10] Due to the unbalanced nature of the data, the mean of others—those that are available in each survey month—is used instead of each separate forecast. n is the number of forecasters, including forecaster i.

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

The estimation rests on a couple of timing assumptions. Consider especially the estimation of equation (3). First of all, there can be no reverse causality since a forecast released in a certain month (the dependent variable) can neither impact the own released forecast in the previous month nor the forecast released in the previous month by competitors. Second, the timing of the estimation also rules out, by assumption, the existence of simultaneity. The assumption is that there is no information in period h that simultaneously impacts the dependent variable (the current month's forecast) and the regressors (the previous month's forecasts). Hence, the error term in period h is assumed to be uncorrelated with any of the regressors (dating to the last month) as it, by assumption, captures new information received in the current month (horizon h). In contrast, the regressors capture all previously known information.

It is also assumed that a forecast released today cannot impact the forecast released today by another forecaster. The rationale behind this assumption is that forecasters are competitors and are assumed to not share information before releasing it to stakeholders and the public. This assumption rules out the possibility that F_{ith+1} directly impacts Z_{ith+1} and vice versa. The impact rather goes via one lag, such that F_{ith+1} impacts Z_{ith}, and Z_{ith+1} impacts F_{ith}.[12] Furthermore, the lagged forecast represents all the previously known information since the past forecasts should summarize all past information. F_{ith+1} and Z_{ith+1} are hence going to be collinear since they are determined, partially, by the same information set. The information set that determines F_{ith+1} is $\{I_{th+1}, S_{ith+1}\}$ while the information set that determines Z_{ith+1} is $\{I_{th+1}, \sum_{j\neq i} S_{jth+1}\}$, where I_{th+1} is all publicly available information at the time of forecast release—including all previously released forecasts—and S_{ith+1} is the private signal of forecaster i.[13] Hence, the information sets differ by the private information in time $th+1$, resulting in multicollinearity, with estimations likely suffering from inflated standard errors.

Due to the information structure of annual average growth rates and the assumption that forecasters have access to each other's lagged estimates, the forecasts are likely to be serially correlated and correlated across firms within the same survey month. To tackle these issues, inference is based on two-way cluster robust standard errors (Cameron et al., 2011; Cameron & Miller, 2015) at the forecaster (firm) level and the country-by-survey month level.

[12] Hence, there should not be any concern regarding reflection problems (Manski, 1993).

[13] Note that a regression of the form $F_{ith} = \alpha + \beta Z_{ith} + e_{ith}$ would likely yield $\beta = 1$ due to the fact that all forecasts today are determined to a large extent by the same information set. However, as Angrist (2014) points out, most often, this cannot be interpreted as a causal relationship. Individual outcomes are often highly correlated with group average outcomes without a causal peer-effect. The same would hold in this paper, since, by assumption, forecaster i at horizon h does not have access to Z_{ith}. They do, however, have access to Z_{ith+1}.

4 Results

Table 3 reports the results from estimating equations (1), (2) and (3). In columns (1) and (4), the dependent variable is the forecast error, and under the hypothesis that forecasters use competitor's forecasts efficiently, the expectation is for these coefficients to be zero. A positive coefficient on *Lagged Others* is interpreted as underuse of the competitors' forecasts, while a negative coefficient is interpreted as overuse of competitors' forecasts. Hence, the coefficient -0.777 in column (1) is interpreted as overuse of the lagged own forecast. In contrast, the coefficient 0.735 is interpreted as underuse of the lagged forecast from others.

Columns (2) and (3) report the estimated coefficients from equations (2) and (3) using GDP growth forecasts while columns (5) and (6) correspond to the inflation forecasts. The coefficients in columns (2) and (5) can be interpreted as the suggested optimal behavior—that is, what they should have done. The estimates suggest that it would be optimal to put very high attention (0.926) on *Lagged Others* while putting very low attention on the own *Lagged Forecast* (≈ 0.05). The estimates suggest that the mean of others' forecasts should get around 17 times higher weight than the own lagged forecast.

Columns (3) and (6) report the estimated actual behavior where it can be seen that the own *Lagged Forecast* gets a very high weight (≈ 0.8), while *Lagged Others* gets a relatively small weight (≈ 0.2). The difference between columns (2) and (3) and between (5) and (6) provide the error found in columns (1) and (4), respectively. For example, for *Lagged Others*, for inflation, $0.926 - 0.219 = 0.707$, which indicates underuse of forecasts from competitors. Given the results in Table 3, the forecasts appear to be too persistent, and forecasters do not seem to internalize the importance of pooling information from others.

Table 3. *Decomposing the Efficiency Error*

	GDP growth			Inflation		
Dep. Variable:	(1) Forecast Error	(2) Outcome	(3) Forecast	(4) Forecast Error	(5) Outcome	(6) Forecast
Lagged Forecast	-0.777***	0.054	0.831***	-0.738***	0.055**	0.792***
	(0.033)	(0.037)	(0.010)	(0.026)	(0.027)	(0.007)
Lagged Others	0.735***	0.926***	0.192***	0.707***	0.926***	0.219***
	(0.036)	(0.039)	(0.011)	(0.028)	(0.029)	(0.007)
Observations	90,920	90,920	90,920	88,390	88,390	88,390
R^2	0.055	0.499	0.968	0.100	0.655	0.963

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

Table 4 presents a sensitivity analysis regarding the actual behavior estimates presented in Table 3. The sensitivity analysis is performed to ensure that the estimation results are robust and sound.

First, columns (1) and (3) in Table 4 present results for when forecaster (firm) fixed effects are added to equation (3). The firm fixed effects control for any time-invariant differences in general optimism or pessimism among the forecasters. It is important to note that the estimation refers to a dynamic panel when firm fixed effects are included. One concern using a dynamic panel is the so-called Nickell bias (Nickell, 1981). However, given the reasonably large number of time periods, the estimations' bias should be neglectable. The stability of the coefficients presented in Table 4 mitigate the concerns regarding potential bias in the estimates.

Second, there might be some information that is released in a particular month and observed by everyone. Examples could be actual data published by the official statistics authorities or policy decision from the government. In order to control for and capture these public signals, country-by-survey month fixed effects are included in the estimations presented in columns (2) and (4) of Table 4. These time effects absorb all information in each period that is common to all forecasters.

From the estimates in Table 4, it is clear that the estimated actual behavior is very robust. Nevertheless, the results presented in Table 3 are obtained using data pooled over countries, years, and forecast horizons. Consequently, Section 5 investigates potential heterogeneity in the actual behavior.

Table 4. *Robustness of Actual Behavior*

	GDP growth		Inflation	
	(1)	(2)	(3)	(4)
Dep. Variable:	Forecast	Forecast	Forecast	Forecast
Lagged Forecast	0.808***	0.805***	0.772***	0.771***
	(0.011)	(0.009)	(0.007)	(0.007)
Lagged Others	0.217***	0.228***	0.242***	0.258***
	(0.012)	(0.010)	(0.008)	(0.009)
Observations	90,920	90,920	88,390	88,390
R^2	0.969	0.978	0.964	0.972
Firm	✓	✓	✓	✓
Country × Survey Month		✓		✓

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

5 Heterogeneity in Actual Behavior?

5.1 Country and Year

Forecasters might behave differently for different countries, and forecasters might have changed their behavior over time. The graphs in Figures 1 and 2 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 reference points. From the graphs, there is no evidence of a significant difference in the behavior between countries, and the behavior is stable over time.

(a) GDP growth *(b)* Inflation

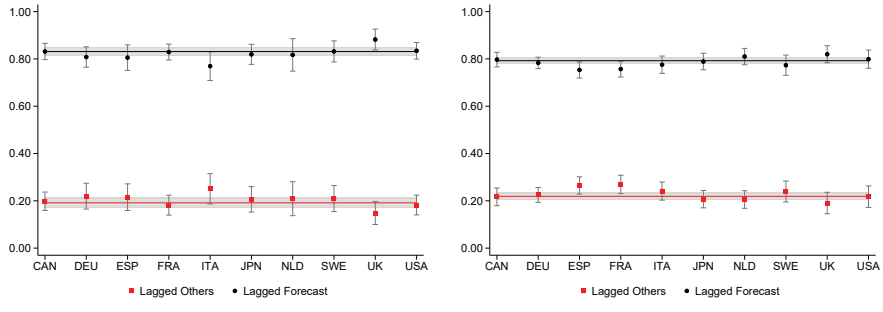


Figure 1. Country Analysis

Notes: Solid lines refer to the estimates in Table 3. The 95% confidence intervals are constructed using two-way clustering at the forecaster and the country-by-survey month level.

(a) GDP growth *(b)* Inflation

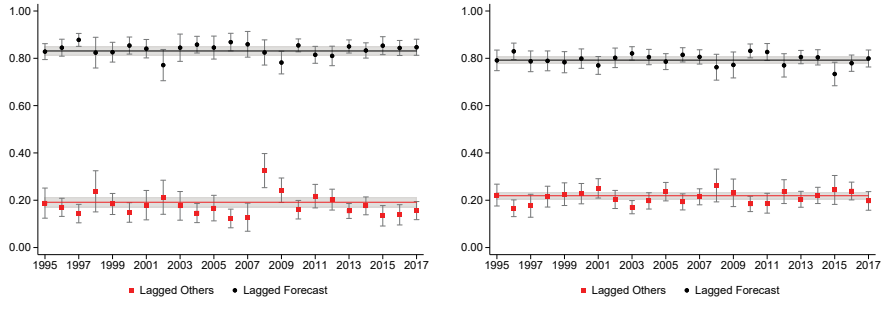


Figure 2. Year Analysis

Notes: Solid lines refer to the estimates in Table 3. The 95% confidence intervals are constructed using two-way clustering at the forecaster and the country-by-survey month level.

5.2 Variable

From Table 3, the behavior of forecasters is similar using either GDP growth or inflation forecasts. The obtained dataset from Consensus Economics contains many more variables than GDP growth and inflation, allowing for an easy extension to also consider the actual behavior regarding additional variables. Equation (3) is estimated using subsamples based on several other variables. Tables 5 and 6 present the results and confirm that the behavior is very consistent. For all variables, the *Lagged Forecast* has a coefficient of around 0.8, while *Lagged Others* receives a coefficient of around 0.2.[14]

Table 5. *Variable Heterogeneity, Part 1*

	Consump.	Wholesale	Inv.	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)	(0.02)	(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)	(0.01)	(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: Estimated equations are (3). Standard errors robust to two-way clustering at the forecaster (firm) level and the country-by-survey month level. *,**,*** represent the 10%, 5%, 1% significance levels.

Table 6. *Variable Heterogeneity, Part 2*

	Empl.	Unempl.	Wages	New Car	Car Sales	Housing Starts	Current Acc.	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: Estimated equations are (3). Standard errors robust to two-way clustering at the forecaster (firm) level and the country-by-survey month level. *,**,*** represent the 10%, 5%, 1% significance levels.

[14]The variables differ in terms of sample size. The difference is due to several factors. First, fewer forecasters in each country do forecasts for many of these variables. Some of these variables are also only observed in a subset of the countries, and other are only present during a short period of the sample. Many of them change definitions over time.

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, the main specification of the actual behavior is estimated using a sample selection for each horizon.

The results are presented as dots in Figure 3. Again, the solid lines correspond to the pooled results from Table 3. From Figure 3, it is evident 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 that forecasters behave differently at different horizons might suggest some strategic decisions at play. It might be that alternative motives are more pronounced far away from the realization of the outcome. As forecasters then approach the realization of the outcome, they switch to concerns regarding the size of the forecast error. This reasoning is extended further in Section 7.

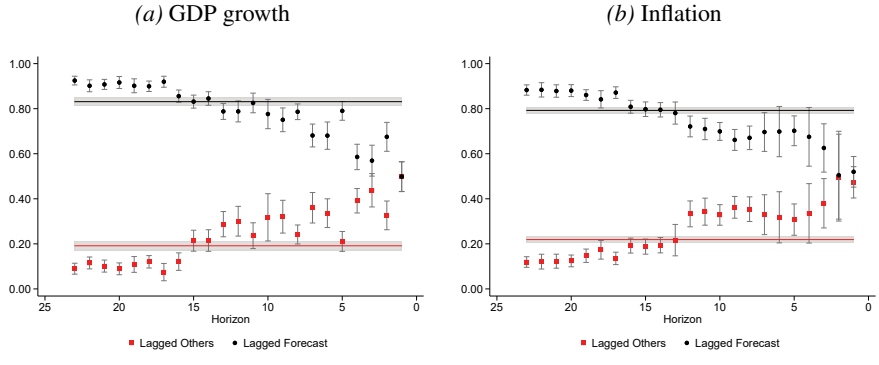


Figure 3. Horizon Analysis

Notes: Solid lines refer to the estimates in Table 3. The 95% confidence intervals are constructed using standard errors robust to two-way clustering at the forecaster and the country-by-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 further 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. Two versions of the model will be considered: one that has fully informed forecasters who perfectly filter out all the signals and one with forecasters who follow a simplified rule of thumb. The rule of thumb will be such that forecasters will utilize the lagged actual forecasts as reduced form aggregates, instead of filtering out signals. Hence, the rule of thumb model is more in line with the empirical approach and the simple framework presented in Section 3.1.

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 they have three attempts at predicting x.[15] 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 from the first-order conditions we 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 $\epsilon_{ih} \sim N(0, \sigma_{\epsilon_h}^2)$ is a random error. The level of noise $\sigma_{\epsilon_h}^2$ is assumed to be the same for everybody.

In period one, i.e., at 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 (containing the prior and the signal). Hence, the forecast is

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

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

where we have used the definition $p_v \equiv \frac{1}{\mathbb{V}[v]}$, such that p is a precision parameter defined as the inverse of the variance. As a result, $p_\mu = \frac{1}{\sigma_x^2}$ and $p_{s_3} = \frac{1}{\sigma_{\epsilon_3}^2}$.[16] 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.[17] 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:[18]

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

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.[19]

Allowing 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.[20] From (9), a forecast in the model is a weighted average of all signals and the prior. The model does not 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.

[16] Equation (5) follows from Bayesian updating.

[17] 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.

[18] 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.

[19] The weights follow the structure presented in Equation (5).

[20] See, for example, Andersson et al. (2017) for an analysis of the expected size of forecast errors at different horizons.

Consider a world with only three forecasters (a, b and c) such that $n = 3$. Then write the horizon 2 forecast from equation (7) for forecaster a in terms of lagged forecast instead of lagged signals,

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

where $Z_{a3} = \frac{1}{2}\left(F_{b3} + F_{c3}\right)$. To replicate the individual signal weights, it is then needed 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, $\omega_{z3} = (n-1)\omega_{f3}$ is such that the relationship between the own lagged forecast and the lagged forecast of the others is proportional in the number of forecasters.[21]

As a result of this, 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, the optimal behavior was estimated 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 of an average of 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. This section presents an alternative model that imposes some form of bounded ability to ensure that we do not ask too much from our forecasters. Follow the same setup as in the fully informed model, but assume now that forecasters cannot, or at least do not, filter out the individual signals. Hence, forecasters use a 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 receives the private signal and declares a forecast just as before:

$$F_{i3} = \frac{p_\mu}{p_\mu + p_{s_3}}\mu + \frac{p_{s_3}}{p_\mu + p_{s_3}} S_{i3}. \tag{12}$$

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. Hence, this setup treats Z_{i3} as an aggregate signal such that the Horizon 2 forecast will be

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

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

where, again, we 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 will, again, depend on the number of forecasters such that $w_{zh} = (n-1)w_{fh}$.[22] 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.[23]

7 Extending the Model to Understand Actual Behavior

From the empirical exercise, two facts are of interest. First, the forecasters pay little attention to their competitors. Second, the forecasters change their behavior over forecast horizons. The results show that the actual behavior of forecasters is such that they put high weight on their own lagged forecast, and this weight is different at different horizons. The empirical exercise also provides relative use between others' versus their own lagged forecasts. This is of good use since the model framework provides a prediction for that relationship. The estimated actual behavior is a ratio (others' over own) 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 the assumed loss function of forecasters is incorrect, resulting in incorrect inefficiency judgment.[24]

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 cognitive bias often is called "the mother of all biases." Overconfidence is

[22] As before, it is assumed that everyone has the same signal precision in the private information. It is also assumed that each forecaster assumes that everyone is equally good.

[23] 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, the prior will be 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.

[24] However, foreasters are still going to be inefficient in minimizing forecast errors, but this might not be their only objective.

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."[25] However, it is unlikely that overconfidence is the only explanation since this would require unreasonable levels of overconfidence. A simple back-of-the-envelope 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 $\kappa > 1$.[26] Continue to assume that forecasters have correct beliefs about others and that they correctly can filter out the signals from their competitors. This will result in a higher relative weight on their own signals (and forecasts).

To understand the level of κ, assume 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, reflecting that the own ability is twice that of competitors, the weight structure changes to 0.095 on the own forecast and 0.905 on the others' forecasts. With $\kappa = 10$, the resulting weights are 0.34 on the own and 0.66 on the others', and $\kappa = 100$ yield 0.84 on the own forecast and 0.16 on the others'. Hence, with many forecasters, κ needs to be very high to replicate the relationships found in the data. Judging this number as high or not is, of course, not straightforward. However, consider the following example; assume the true value of 2 to be predicted. Assume a true signal that offers a 95% confidence interval equal to 2 ± 1. A "hundred times" overconfident forecaster would judge the own signal to be in the 95% confidence interval of 2 ± 0.01. In contrast, each of the competitors is judged to have a signal in the 2 ± 1 interval. 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 strategic trade-off that produces horizon dynamics as those found in the data.

To explain both the general underuse of competitors' forecasts and the horizon dynamics in this behavior, the introduction of a revision cost is considered (see, e.g., Coibion & Gorodnichenko, 2015; Ehrbeck & Waldmann, 1996). The idea behind the revision cost is as follows. 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 its 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, subject to the new private signal. This would imply a forecast revision of 2 percentage points. A revision that large would need substantial explanation since no big news has been uncovered. Telling the public that the

[25] See, for example, Kahneman and Tversky (1972) and Tversky and Kahneman (1974) for more on this cognitive bias.

[26] See Daniel et al. (1998) for similar modeling of overconfidence in the securities markets.

forecast last month was wrong after seeing the forecasts of others would probably be a big hit to the forecaster's credibility.[27] 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.[28] 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 (Nordhaus, 1987).

Introducing a revision cost gives rise to a trade-off between loss from forecast errors and loss from revisions, resulting in a higher degree of forecast smoothing. However, in order to explain the horizon heterogeneity, the cost function needs to be horizon discounted, so that the trade-off also introduces a horizon trade-off. The following section outlines the technical details of a 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. This section adds an explicit revision cost to the loss function. The baseline model consisted of three periods (horizons). This section extends the model to 23 horizons to understand the horizon dynamics.[29] Assume that the forecaster faces the following minimization problem:[30]

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

At each horizon, the forecaster faces a trade-off between the minimization of the forecast error and the forecast revision.[31] 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.[32] 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

[27] 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.

[28] See Nordhaus (1987) for similar reasoning.

[29] It is important to note that the predictions from the baseline model do not depend on the number of horizons.

[30] This is similar to a loss function presented in Coibion and Gorodnichenko (2015), with the difference of having horizon dependent trade-off weights.

[31] In the first period, there is no lagged forecast. Hence, the solution in the first period is trivial and follows $F_{i24} = \mathbb{E}_{i24}[x]$.

[32] See, for example, 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.

forecast at horizon h is given by

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

Note that $\mathbb{E}_{ih}[x]$ is formed as in the baseline model, where it is a weighted average of the prior and the signals. Solving backwards, the horizon 1 forecast is[33]

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

Solving the horizon 2 forecasts then yields

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

Substitute (18) into (19) and rearrange to get[34]

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

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

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

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

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

and simplify any expression of this form as

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

where $(1-m_h)$ can be seen as a smoothing parameter. In the empirical results, it was 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, $g(h)$ needs to decrease faster than $r(h)$ as the horizon increases. The intuition is that forecast errors need to be more costly in relation to revisions on short relative to long 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 one 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.

[33] Note that by assumption $F_{i0} = x$ and $r(0) = 0$ since the outcome is known at horizon zero.

[34] Note that $\mathbb{E}_{i2}[x] \neq \mathbb{E}_{i1}[x]$, $\mathbb{E}_{i2}[\mathbb{E}_{i1}[x]] = \mathbb{E}_{i2}[x]$ and that $\mathbb{E}_{i2}[F_{i2}] = F_{i2}$.

8 Concluding Remarks

From the empirical exercise, forecasters are deemed inefficient since they underuse information from their competitors given an assumption of pure forecast error minimization. The empirical results are robust for several heterogeneities, such as country, year, and variable. Moreover, 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 still far from an optimal level.

A simple model of noisy and private information provides 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 close attention to what the other forecasters do. This result originates from the fact that it is assumed that each forecaster has private information such that there is a benefit from pooling information from many forecasters. The result regarding the importance of pooling is no big surprise and is in line with standard findings in the forecasting literature. Perhaps more surprising is the finding that forecasters do not incorporate this in their revision process.

Previous studies have found that forecasters appear to put high weight on their own lagged information such that they are too slow to incorporate new information. In the empirical part of this paper, the focus is 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 higher relative weight on their own lagged forecast compared to the forecasts of competitors. In comparison, optimal is rather 17 times higher weight on the competitors' forecasts compared to the own lagged forecast.

In order to explain the observed (seemingly) inefficient behavior, the baseline model is extended. First, it is unlikely that the cognitive bias of overconfidence is the only source of the observed behavior. For overconfidence to explain the observed behavior, forecasters need to think that they are around 100 times better than their competitors. It is, of course, possible (and likely) that some degree of overconfidence is present as a partial explanation to the findings. Second, introducing a revision cost can only explain the observed horizon dynamics under the assumption that the horizon discounting is relatively different for the cost from forecast errors and the cost from forecast revisions.

It is well known that the average of multiple forecasts often outperforms individual estimates when compared over time. This well-known result prompts third-party users of forecasts, such as consumers, investors, voters, and policymakers, to collect information from multiple sources. However, the results in this paper show that professional forecasters themselves do not sufficiently internalize this well-established result.

References

Ager, P., Kappler, M., & 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., & Reslow, A. (2017). Adjusting for Information Content when Comparing Forecast Performance. *Journal of Forecasting*, *36*(7), 784–794.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, *30*, 98–108.

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

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

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

Bordalo, P., Gennaioli, N., Ma, Y., & Shleifer, A. (2020). Overreaction in macroeconomic expectations. *American Economic Review*, *110*(9), 2748–2782.

Broer, T., & Kohlhas, A. (2018). *Forecaster (Mis-)Behavior* (CEPR Discussion Papers No. 12898).

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

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

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

Cipullo, D., & Reslow, A. (2021). Biased Forecast to Affect Voting Decisions? The "Brexit" Case. In A. Reslow (Ed.), *Electoral Incentives and Information Content in Macroeconomic Forecasts*. Department of Economics, Uppsala University.

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

Coibion, O., Gorodnichenko, Y., & Kumar, S. (2018). How Do Firms Form Their Expectations? New Survey Evidence. *American Economic Review*, *108*(9), 2671–2713.

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

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

De Bondt, W. F., & Thaler, R. H. (1995). Chapter 13 Financial Decision-making in Markets and Firms: A Behavioral Perspective. *Finance* (pp. 385–410). Elsevier.

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

Dovern, J., Fritsche, U., Loungani, P., & 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., & Waldmann, R. (1996). Why Are Professional Forecasters Biased? Agency versus Behavioral Explanations. *The Quarterly Journal of Economics*, *111*(1), 21–40.

Elliott, G., Komunjer, I., & Timmermann, A. (2008). Biases in macroeconomic forecasts: Irrationality or asymmetric loss? *Journal of the European Economic Association*, *6*(1), 122–157.

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., & Jeon, Y. (2002). Copycats and Common Swings: The Impact of the Use of Forecasts in Information Sets. *IMF Staff Papers*, *49*(1), 4–21.

Gneiting, T. (2011). Making and Evaluating Point Forecasts. *Journal of the American Statistical Association*, *106*(494), 746–762.

Hendry, D. F., & Clements, M. P. (2004). Pooling of forecasts. *The Econometrics Journal*, *7*(1), 1–31.

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

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

Kim, O., Lim, S. C., & 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., & 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., & 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.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, *60*(3), 531–542.

Marinovic, I., Ottaviani, M., & Sørensen, P. N. (2013). Chapter 12 – Forecasters' Objectives and Strategies. In G. Elliott & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (pp. 690–72). Elsevier.

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

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, *49*(6), 1417–1426.

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

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

Rülke, J.-C., Silgoner, M., & 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., & Koga, M. (2019). Firm Performance and Macro Forecast Accuracy. *Journal of Monetary Economics*.

Tversky, A., & 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* (NBER Working paper No. 8673).

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

Appendix

Table A1. *Outcome Statistics*

		GDP growth			Inflation		
Country	Obs.	Obs.	Mean	SD	Obs.	Mean	SD
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	SD	Obs.	Mean	SD
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: Data refer to ex-post forecast errors where errors are defined as the outcome minus the 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.

Essay IV

Adjusting for Information Content when Comparing Forecast Performance*

with Michael K. Andersson and Ted Aranki

Acknowledgments: The authors thank Karl Walentin, Xin Zhang, and the seminar participants at Sveriges Riksbank and the National Institute of Economic Research for their comments and suggestions. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of Sveriges Riksbank or Finansinspektionen. Andersson and Aranki are currently employed by Finansinspektionen, and some of the work on this essay was conducted while the authors worked at Sveriges Riksbank.

*Published as Andersson, M. K., Aranki, T., and Reslow, A. (2017), Adjusting for Information Content when Comparing Forecast Performance, *Journal of Forecasting*, 36(7):784–794.
Copyright John Wiley and Sons; reproduced with permission from the publisher.
Erratum: In the version published in the Journal of Forecasting, the joint hypothesis test in Section 4.2 was incorrectly specified to be $\delta_{i=1} = \delta_{i=2} =, ..., \delta_{i=10} = 0$, while the correct should be $\delta_{i=1} = \delta_{i=2} =, ..., \delta_{i=10}$.

1 Introduction

Many agents in the economy, including economic policymakers, regularly publish forecasts of the development of the economy. Since important economic and political decisions usually are based on forecasts it is crucial that these predictions are as accurate as possible. Therefore, evaluations are regularly undertaken to assess the performance. Usually evaluations compare the forecast at a specific point in time with the outcome, or using a statistical measure based on the two quantities. One such measure is the mean absolute errors (MAE) (see, e.g., Diebold, 2007; Gneiting, 2011). Albeit informative, such evaluations are in general insufficient. A large forecast error, for a particular observation, can be a consequence of a shock that could not have been foreseen. Comparing different agencies is a common way to handle this problem. For example, central banks, international organizations and other institutions regularly publish evaluations of their forecasts and compare them with other forecasters (see, e.g., Bank of England, 2015; European Central Bank, 2013; Sveriges Riksbank, 2016; Timmermann, 2007; Vogel, 2007). More in-depth analyses of forecasts are presented at a less frequent basis. For instance, Blix et al. (2001) compare different forecaster's performance of key Swedish macroeconomic variables. Andersson et al. (2007) estimate the Riksbank's accuracy in relation to that of the National Institute of Economic Research and simple econometric specifications. Davies and Lahiri (1995) and Bauer et al. (2003) compare the agents of the Blue Chip Survey and Boero et al. (2008) assess the Bank of England Survey of External Forecasters. Goh and Lawrence (2006) compares the accuracy of New Zealand forecaster's RMSE and their average relative rank. Cabanillas and Terzi (2012) present an assessment of the European Commission's track record and compares the forecast errors of GDP growth to those of other international institutions.

Comparing forecasters is appealing but not necessarily straightforward. One problem that evaluations face in practice is that forecasts are published at different points in time. In practice, this implies that different forecasters have different amounts of information when they prepare their forecasts. To make a fair comparison, this must be accounted for. One attempt to reduce this problem is to compare forecasts produced at almost the same point in time. This approach is far from flawless since one month of information may be important, for example, if a quarter of the National Account figure is published that particular month. Another problem is that this approach utilizes only a subset of the available forecasts. We argue that a fair comparison between forecasters requires that the differences in information must be adjusted for in a quantitative way.

The contribution of this paper is that we propose a method to account for the fact that forecasts are made at different points in time, and thus with different amounts of information when comparing the accuracy. The typical data set we address is thus an unbalanced panel with forecasts on annual averages. We

represent the amount of information by timing, that is, the distance between the publication date and outcome date. In the model, the timing effect and the respective forecasters underlying performance (ability) are estimated simultaneously. The model rests on the assumption that part of a forecaster's average error can be explained by the timing. A forecaster that publishes its forecasts later than others can also be expected, ceteris paribus, to have a slightly better accuracy. A Monte Carlo simulation shows that this is the case and that the proposed model can correct for this. In an empirical application, the model is used to assess the forecasting performance of Swedish forecasters with respect to GDP growth and CPI inflation for the period 1999–2015. The ranking of the forecasters depends on whether information differences are accounted for or not. Overall, we find the model well-suited for comparing forecast performance.

The paper unfolds as follows: Section 2 presents the empirical framework and the econometric model, and Section 3 holds a Monte Carlo simulation. Section 4 presents the empirical application and Section 5 concludes the paper.

2 Empirical Framework

This section discusses the effect of the information content, measured by time to the outcome, and explains why this is an important determinant of the size of the forecast error. The forecast error is expected to become smaller the nearer (in time) to the outcome the prediction is made. We then present an econometric model that adjusts for differences in timing when estimating the ability of the different forecasters.

2.1 The Forecast Error

Let $\hat{y}_{it,h}$ be the forecast for the growth rate of variable y for year t, made by individual i, h months prior to the end of year t (the period the outcome is known). The error of a particular forecast is then defined by:

$$v_{it,h} = y_t - \hat{y}_{it,h}. \tag{1}$$

This error is expected to decline as the horizon shrinks, because the information set grows over time. For a forecast of a variable expressed as an annual average, the errors decrease at a non-linear rate as quarterly or monthly outcomes of the particular prediction year become available. We exemplify this with quarterly data, by looking at how annual and quarterly growth rates are

related. The annual change of a variable is a function of the quarterly changes during the last four quarters according to:[1]

$$\frac{y_{T,1}-y_{T-1,1}}{y_{T-1,1}}=\frac{y_{T,1}}{y_{T-1,1}}-1=\frac{y_{T,1}}{y_{T-1,4}}*\frac{y_{T-1,4}}{y_{T-1,3}}*\frac{y_{T-1,3}}{y_{T-1,2}}*\frac{y_{T-1,2}}{y_{T-1,1}}-1, \quad (2)$$

where $y_{T,q}$ is the level of variable y in quarter q of year T. Let $g^4_{T,q}$ and $g^1_{T,q}$ be the annual and quarterly change for quarter q of year T. Equation (2) can then be expressed as:

$$1+g^4_{T,1}=(1+g^1_{T,1})*(1+g^1_{T-1,4})*(1+g^1_{T-1,3})*(1+g^1_{T-1,2}). \quad (3)$$

The annual change of quarters 2, 3 and 4 are calculated as equation (3). Thus, the yearly (average) growth for year T, $\Delta_4 y_T$, can be stated as the average of the annual change in the year's quarters:

$$\begin{aligned}\Delta_4 y_T=\frac{1}{4}\Big[&(1+g^1_{T,1})*(1+g^1_{T-1,4})*(1+g^1_{T-1,3})*(1+g^1_{T-1,2})+ \quad (4)\\ &(1+g^1_{T,2})*(1+g^1_{T,1})*(1+g^1_{T-1,4})*(1+g^1_{T-1,3})+\\ &(1+g^1_{T,3})*(1+g^1_{T,2})*(1+g^1_{T,1})*(1+g^1_{T-1,4})+\\ &(1+g^1_{T,4})*(1+g^1_{T,3})*(1+g^1_{T,2})*(1+g^1_{T,1})\Big]-1.\end{aligned}$$

Equation (4) shows that the annual growth rate of year T is a function of all quarterly growth rates from the second quarter of the previous year, $T-1$, until the last quarter of the forecast year, T. The equation also indicates that the quarters weigh differently in terms of yearly growth. Table 1 present the weights.

Table 1. *The relative importance of each quarter in annual average calculation*

	Year: $T-1$				Year: T			
Quarter	1	2	3	4	1	2	3	4
Weight	0/16	1/16	2/16	3/16	4/16	3/16	2/16	1/16
Accumulated weight	0/16	1/16	3/16	6/16	10/16	13/16	15/16	16/16

Notes: The table shows the weights of each quarter in the yearly calculation, together with the proportion of the full year value (accumulated) known at each point in time.

Table 1 reveals that a forecast of the annual average of year T made in the first quarter the year prior to the outcome, $T-1$, do not include any information of the outcome. One quarter later, i.e., the second quarter in year $T-1$, 1/16 of the outcome is known. As the first quarter of the forecast year becomes available, 10/16 of the yearly outcome is known to the forecaster. Note

[1] Similar calculations can easily be generalized to an arbitrary frequency.

that these calculations are stylized and do not consider recurrent revisions of quarterly outcomes, which is a common feature of macroeconomic data.[2] The practical implication of the growing information set is that less of the outcome needs to be forecasted as time proceed. We call this the *outcome effect*.

In addition to the outcome effect, there exist other reasons to why the forecast errors are expected to decrease the closer to the outcome the forecast is made. For example, the degree of persistence in the data. If we assume that the quarterly growth of variable y in period t is a function of the quarterly growth in the previous period we get the following relation:

$$y_t = c + \alpha y_{t-1} + \epsilon_t. \tag{5}$$

For future periods $t+1,...,t+H$, equation (5) is a simplified description of the expected future quarterly growth rates which the forecaster can use at each forecast occasion. The parameter α (persistence of y) determines how far into the future a new outcome will remain important. We call this effect the *persistence effect*.[3]

Given the timing effects (i.e., *outcome effect* and *persistence effect*), we can calculate how the expected forecast errors, in absolute terms, will decrease as the distance to the outcome decreases. In Figure 1 we plot a calculation of the expected forecast error as a function of the timing assuming α to be 0.3 and 0.6 respectively.[4] The size of the forecast error before any information of the outcome is known, i.e., the first quarter of the year prior to the outcome year $T-1$, has been normalized to 1. Figure 1 provides a sketch showing that the functional form of the expected size of the forecast error depends on the outcome effect and the persistence effect. This study uses this functional form to account for the expected decline in the errors.

2.2 The Econometric Model

Given the panel structure of data and following the previous section, the absolute forecast error, denoted by ϵ, can be explained by:[5]

$$\epsilon_{it,h} = \delta M_{it,h} + u_{it,h}, \tag{6}$$

[2]GDP data are usually revised back in time when a new outcome is published. Seasonal adjustments may also lead to a revision of historical GDP observations.

[3]Moreover, indicator series and forecasts given by other institutions may enhance the information set even more.

[4]As a reference, an estimation of α based on quarterly data for Swedish GDP during 1994–2015 gives the value 0.3.

[5]The method may also be adapted to other error measures. The measure should be selected on the base of the loss-function of the forecasters (see, e.g., Patton, 2015). However, comparing different evaluation measures is beyond the scope of this paper.

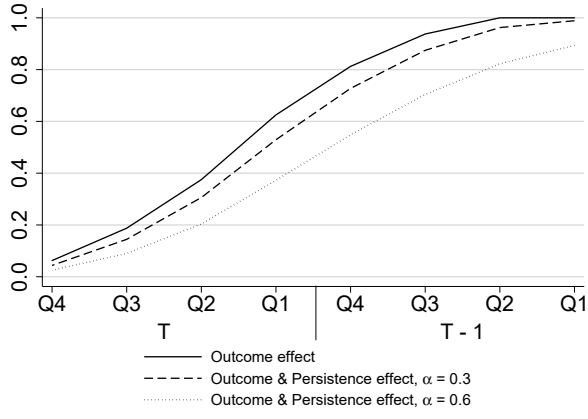


Figure 1. The impact of the timing on the size of the forecast error

Notes: The black solid line indicates the extent to which the expected error depends on the outcome effect, the dashed and dotted lines indicates the extent to which the error depends on the outcome and persistence effects, where the persistence coefficient α is set to 0.3 and 0.6, respectively.

where $M_{it,h}$ is the timing effect, defined as $M_{it,h} = 1 - W_{it,h}$. The accumulated weight, $W_{it,h}$, increases as the parts of the prediction year become known.[6] As $W_{it,h}$ increases $M_{it,h}$ decreases and could be seen as a proxy for missing information at the time of the publication. The coefficient δ represents the marginal effect on the absolute forecast error of having less information. We model the absolute errors in equation (6) without an intercept, hence, at horizon zero the absolute forecast error is also zero since the full outcome is then known. The disturbances $u_{it,h}$ can be divided into two unobserved components and an idiosyncratic random disturbance term, according to:

$$u_{it,h} = \mu_i + \lambda_t + e_{it,h}, \tag{7}$$

where the individual specific effect μ_i is interpreted as the (genuine) ability of forecaster i. This ability is assumed to be constant over time. λ_t is a time specific effect which refers to the year the forecast concerns. This effect is assumed to be common for all forecasters. $e_{it,h}$ is the idiosyncratic disturbance term. Rewriting equation (6) gives the following specification:

$$\epsilon_{it,h} = \delta M_{it,h} + \mu_i + \lambda_t + e_{it,h}. \tag{8}$$

Hence, the method to adjust for information differences when comparing the forecast performance across institutions is based on the assumption that the

[6] In the estimation we calculate the weights based on monthly observations in contrast to the quarterly weights presented in Table 1. Thus, a forecast produced in December will have a different weight compared to a forecast produced in October. We also assume a publication lag of one month, which means that the December outcome becomes available in January.

(absolute) forecast errors can be divided up into different components: a component that is due to the amount of information available at the time of publication ($M_{it,h}$), a component that reflects the different forecasters' general forecasting performance/ability (μ_i) and a component that captures that different years can be more or less difficult to forecast (λ_t).[7]

In equation (8) we assume a common information effect for all years. Nevertheless, we will also control for time varying slope coefficients. One might argue that the fact that different years can be more or less difficult to forecast does not only affect the level of the forecast error but also the functional form of the information effect (i.e., the slope). Hence, a more flexible specification becomes:

$$\epsilon_{it,h} = \delta_t M_{it,h} + \mu_i + \lambda_t + e_{it,h}. \tag{9}$$

We estimate model (9) using the within (or fixed effect) estimator. The within estimator is preferred when the individual-specific factors are constant over time. Moreover, the within estimator allows for correlation between the fixed individual-specific effects and the other explanatory variables, without violating the assumption of no correlation between observable and unobservable determinants of the dependent variable. This is due to the fact that the within transformation wipes out the individual-specific effects (see, e.g., Baltagi, 2001). Since the individual-specific effects in this model, which are interpreted as each forecaster's individual forecast ability, may be correlated with the other explanatory variables, the within estimator is appropriate.

3 Monte Carlo Simulation

3.1 Mode of Procedure and Method of Analysis

As described in previous section, the functional form of how the forecast errors decline over time is given by the *outcome effect* and the *persistence effect*. However, the sizes of these effects are unknown, but can be estimated in an econometric model such as equation (9). This model also estimates individual-specific effects that we argue represent the individual forecasting performances. In order to test if the model actually do capture the ability of the individuals, we conduct a Monte Carlo simulation. In the simulation we evaluate the model's comparison between institutions and contrast the model to a traditional mean absolute error (MAE) statistic. We use the following data generating process for the absolute forecast errors ($\epsilon_{it,h}$):

$$\epsilon_{it,h} = L_{it,h} + \mu_i + \varepsilon_{it,h}, \tag{10}$$

[7] A somewhat similar decomposition for analyzing forecast errors in a panel data setting has been used by Davies and Lahiri (1995), Davies (2006) and Boero et al. (2008) with the difference that they do not adjust for timing.

where,

$$L_{it,h} = 1 - \Big(\sum_{j=h+1}^{24} \omega_j + \sum_{\substack{i=1,\ldots,h \\ j=h,\ldots,1}} \alpha^j \omega_i \Big), \tag{11}$$

$$\omega_h = \begin{cases} \frac{h}{12^2} & \text{if } h \leq 12 \\ \frac{24-h}{12^2} & \text{if } h > 12, \end{cases} \tag{12}$$

$$\varepsilon_{it,h} = |e_{it,h}|, \tag{13}$$

$$e_{it,h} \sim N(0, \sigma^2). \tag{14}$$

$L_{it,h}$ represents the information effect h months before the complete outcome is known. $L_{it,h}$ is calculated as one minus the *outcome effect* and the *persistence effect*. ω_h is the monthly weights in the full outcome. In the simulations we assume that all years are equally hard to predict when generating data.

We model the ability using an individual-specific bias (μ_i) that we assign to each forecaster in the simulation. ε_{it} is an error term that is the absolute value of a normally distributed error with mean zero and variance σ^2.[8]

Furthermore, we create seven fictive forecast institutions (A, B, $\cdots$, G). Each institute produces four forecasts each year according to a fixed publication calendar.[9] This gives eight forecasts from each institute for each outcome since we assume a maximum forecast horizon of 24 months prior to the outcome. In the setup, institute A has the shortest average horizon while institute G has the longest.[10] Thus, A has an advantage, due to most information available, at average when forecasting. Besides being affected by the forecast occasion, the errors are influenced by the individual-specific biases. In the evaluation, one would like the institution with the lowest bias to be ranked by the model as the best (number one). We construct two different assignments (cases) of ability:

1. $\mu_A = \mu_B = \cdots = \mu_G$: Equal ability
2. $\mu_A > \mu_B > \cdots > \mu_G$: Unequal ability, different direction to the timing effect

In the first case, we expect the performance of all institutions to be equal, and in the second case, we expect A to have the poorest forecast ability and G the best. In Case 2, we set the sequential differences between forecasters A–G to

[8]The expected mean of $\varepsilon_{it,h}$ is $\sigma\sqrt{\frac{2}{\pi}}$. In the main simulations we set σ to 0.3, but we also do the simulations using σ equal to 0.15.

[9]Publication months: A = {Mar, Jul., Oct., Dec.}, B = {Mar., Jun., Sep., Dec.}, C = {Mar., May., Sep., Nov.}, D = {Feb., May., Aug., Nov.}, E = {Feb., Apr., Aug., Oct.}, F = {Jan., Apr., Jun., Nov.}, G = {Jan., Mar., Jul., Sep.}.

[10]Average horizon in months: A=11, B=11.5, C=12, D=12.5, E=13, F=13.5, G=14.

0.02. Subsequently, the difference between the best and the worst forecaster is 0.12.[11]

The value of α is varied in the study and we also do two different evaluation samples. The first is a single-year evaluation ($N = 1$) and the second is based on 10 years. However, the evaluations are always computed with $S = 100,000$ Monte Carlo replicates. For each replicate, S_i, we rank the forecast performance according to the estimated ability in equation (9) and the MAE statistic. From the $100,000$ rankings we compute the proportion of each institution in each position.

Figure 2 presents the main findings from the Monte Carlo simulation. Our proposed model is labeled Adj. MAE and is compared to the MAE. The results in the figure are presented as the observed proportion of each forecaster in each ranking position.

3.2 The Equal Ability Case

We start with the case with equal ability and no timing effect (case 1). This case is included because we do not want our procedure to distort the performance of the MAE statistic when no timing effect is present. The simulations are based on a single-year evaluation. The results are given in Figure 2a and 2b for the two statistics. The timing effect is eliminated by equating the persistence coefficient α to one. Since the abilities are set equal for all seven forecasters we expect each forecaster's ranking to be uniform in each position. As given in Figure 2a and 2b both our model and the MAE statistic handle this case well. The rankings are evenly distributed for all forecasters.

In the second equal ability simulation we add a timing effect, consisting only of the outcome effect. The rank proportions are still expected to be uniformly distributed across the institutions. Figure 2c shows that our proposed model achieves the expected. However, the MAE fail to rank the forecasters properly in this case, see Figure 2d. The information advantage of forecaster A is interpreted as A being the best forecaster in about 40 percent of the replicates. This is actually not surprising, but will distort an evaluation.

We have also computed a case of random ability. In this case, each forecaster is given an ability from a uniform distribution between 0 and 0.16 in each replicate.[12] Since the forecasters share the same distribution, the forecasters still have the same expected ranking over the simulation. The result of the random ability case is thus similar to the results shown in case 1 with equal ability.

[11] The values of μ_i and σ are inspired from Swedish GDP and CPI forecast error data.

[12] By setting the upper limit to 0.16 we get an average sequential difference of 0.02 and an average difference between the best and the worst forecaster of 0.12, as in case 2.

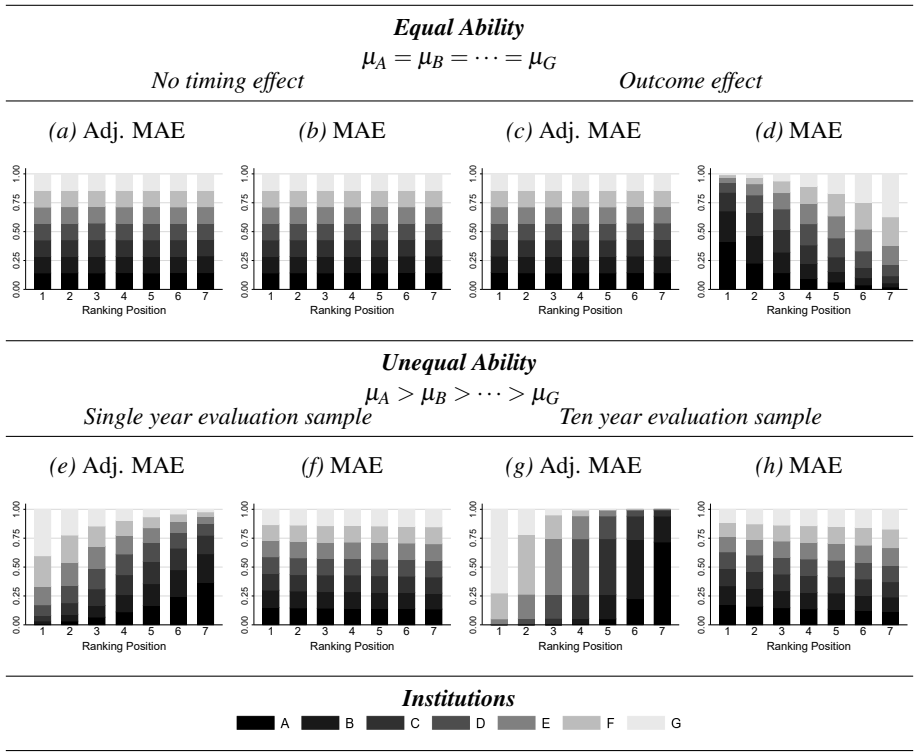


Figure 2. Monte Carlo simulations

Notes: The graphs show the observed proportion of each forecaster in each ranking position according to our proposed model in equation (9) (Adj. MAE) and according to the MAE statistic. In the top panel we have equal ability, $\mu_A = \mu_B = \cdots = \mu_G$ and the sample size, N, is equal to one year. The graphs (a) and (b) show a simulation assuming no timing effect ($\alpha = 1$). Graphs (c) and (d) show a simulation where the timing effect is given by outcome effect only ($\alpha = 0$). The lower panel presents the case of a outcome effect and unequal ability according to $\mu_A > \mu_B > \cdots > \mu_G$, thus, A is expected to be ranked as number seven. The single year evaluation graphs (e) and (f) have N equal to one while the ten year evaluation graphs (g) and (h) have N equal to ten. Results based on S = 100,000 replicates. All simulations have the standard deviation, σ, in the error term set to 0.3.

3.3 An Unequal Ability Case

The lower part of Figure 2 presents results when we do not assume equal ability (case 2). More precisely, we now assume that the generated ability order of the forecasters is such that the forecaster with the shortest average horizon is the poorest, $\mu_A > \mu_B > \cdots > \mu_G$. In other words, the timing effect and the ability now pull in different directions. In this case, we expect forecaster G to be ranked as number one and A as number seven. Figure 2e and 2f show results for a single year, $N = 1$. Since the information effect and the ability works in different directions the MAE statistic concludes that the forecasters are ap-

proximately equally good, see Figure 2f.[13] However, our model does a better job in this case, see Figure 2e. Forecaster G is top ranked in about 40 percent of the cases and forecaster A exhibits the poorest rank in approximately 40 percent of the replicates.

If we instead of evaluating a single year undertake an evaluation based on 10 years, the performance of our model improves further. This is presented in Figure 2g. Forecaster G is now correctly ranked as best in 70 percent of the replicates, forecaster F as second best in 50 percent of the times and forecast A as being worst in over 70 percent of the generated samples. The MAE statistic is still giving wrong rankings, see Figure 2h.

The ability may also work in the "same direction" as the timing effect. Then the unadjusted MAE statistic would work better than in case 2, and slightly better than the adjusted statistic, since it has a bias towards a ranking that follows the direction of the timing effect. The adjusted statistic performance is in line with the results of case 2.

3.4 A Measure of Incorrect Rankings

The overall results from the Monte Carlo are robust also for different designs of the simulations. This is evident in Table 2, which present simulations based on different values of the residual variance, the persistence parameter, as well as the sample size. In the table we present a measure of incorrect rankings (MIR) for a comparison based on observed, $O_{i,r}$, minus expected, $E_{i,r}$, frequency for individual i in ranking r for method j according to the following:

$$MIR_j = \frac{1}{S}\sqrt{\frac{1}{49}\sum_{i=A}^{G}\sum_{r=1}^{7}(O_{i,r}-E_{i,r})^2(1+\rho_i)^2}, \tag{15}$$

where ρ_i is the absolute distance between the expected, $E[r_i]$, and the observed ranking, r_i:

$$\rho_i = |E[r_i] - r_i|. \tag{16}$$

Thus, ρ works as a penalty parameter of having large deviations in the ranking position. In the first ability case we do not have an expected ranking position, $E[r_i]$, meaning that all ρ_i is set to zero, while in the second ability case we set ρ according to equation (16). The column *Adjusted* refers to MIR according to our proposed adjusted MAE statistic, while the column *Unadjusted* refers to MIR based on a traditional MAE statistic.

Overall, for both ability cases, the MIR values are lower for the *Adjusted* method, indicating a smaller ranking error, compared to the *Unadjusted* method. Only when α is one (i.e., no timing effect), the MIR value for both methods

[13] This is of course not always the case. If we vary the size of the information effect and abilities the resulting proportions are also varied.

Table 2. *Monte Carlo simulation: Design and incorrect rankings results*

				Measure of Incorrect Rankings	
μ_i	N	σ	α	Adjusted	Unadjusted
Case 1: $\mu_A = \mu_B = \cdots = \mu_G$	1	0.30	0.0	0.00	0.08
			0.8	0.01	0.08
			1.0	0.00	0.00
		0.15	0.0	0.00	0.15
			0.8	0.01	0.14
			1.0	0.00	0.00
	10	0.30	0.0	0.00	0.20
			0.8	0.02	0.19
			1.0	0.00	0.00
		0.15	0.0	0.00	0.30
			0.8	0.04	0.29
			1.0	0.00	0.00
Case 2: $\mu_A > \mu_B > \cdots > \mu_G$	1	0.30	0.0	0.45	0.62
			0.8	0.46	0.61
			1.0	0.45	0.45
		0.15	0.0	0.36	0.64
			0.8	0.37	0.59
			1.0	0.36	0.36
	10	0.30	0.0	0.28	0.65
			0.8	0.29	0.58
			1.0	0.28	0.28
		0.15	0.0	0.10	0.68
			0.8	0.11	0.56
			1.0	0.10	0.10

Notes: μ_i is the generated individual forecast ability. N is the number of years in the evaluation sample. σ is the standard deviation in the error term and α is the persistence parameter. Note that the value of 0.8 on monthly data approximately corresponds to the line with 0.6 on quarterly data in Figure 1. See equation (10) for the data generating process. The column *Adjusted* refers to the measure of incorrect rankings (MIR) values based on our proposed model in equation (9). The column *Unadjusted* refers to MIR based on a traditional mean absolute error statistic. MIR calculations according to equation (15), where a lower value indicates a smaller ranking error. Results based on 100,000 replications.

are the same. This is, however, in line with what one would expect. All in all, the simulations suggest that the model can handle different situations regarding to the timing effects, which is shown in Figure 2 and Table 2. The Monte Carlo simulations also suggest that our model can adjust for different publication dates and identify the ability of the forecasters. Furthermore, the model does not distort the estimated ability if no information effect is present. The MAE statistic on the other hand is severely affected in the Monte Carlo study by differences in information. We also note that our model performs much better when the sample size increase and that the model performance increases as the variance in the error term is reduced. Hence, we conclude that our model, overall, is preferred compared to an unadjusted statistic.

4 Empirical Application

We now use the model for a real-world application. We pay special attention to two things: The importance of the timing effect by itself and how it affects the ranking. This is done by looking at the data and through estimated quantities in our model.

4.1 Data and Descriptive Statistics

In the application we use data from 10 Swedish forecasters.[14] The forecasts are annual averages for Swedish GDP growth and CPI inflation and are made up to two years prior to the outcome, yielding a maximum forecast horizon of 24 months. The data covers the period 1999–2015. From the forecasts, we calculate the absolute value of the forecast error based on the first available outcome. The absolute errors then give rise to two quantities: the MAE statistic and our timing adjusted ability statistic.

Table 3 presents the data by means of summary statistics for each institution's absolute forecast errors. The number of forecasts differs for the various institutions and so does the average forecast horizon.[15] The average horizon provides a first indication on how the adjustment will affect the evaluation, but is by no means exhaustive. In practice, the horizon for each forecast will matter. The average horizon differs at most by (slightly more than) two months. Two months in this case is definitely enough for potential information disadvantages. The lower part of each panel in Table 3 demonstrates that the forecast errors vary more within each institution's set of errors than between the institutions. This is an indication that forecasters regularly adjust their forecasts, probably due to new outcomes and short-term forecast errors. The small variation between the institutions suggests a prominent "herd behavior" among the forecasters.[16]

4.2 Estimated Timing Effect

This section discusses and estimates the information content effect in the data. Starting with the results reported in Panel A in Table 4. Panel A in Table 4 presents the estimation results of equation (8) for GDP and CPI, respectively.

[14] The evaluated institutions are the same as in Sveriges Riksbank's yearly evaluation (see, e.g., Sveriges Riksbank, 2016): Swedish Ministry of Finance (MoF), HUI Research (HUI), National Institute of Economic Research (NIER), Swedish Trade Union Confederation (LO), Nordea (NDA), the Riksbank (RB), Scandinaviska Enskilda Banken (SEB), Svenska Handelsbanken (SHB), Confederation of Swedish Enterprises (SN), and Swedbank (SWED).

[15] Although there are some regularities regarding when various institutions publish their forecasts the publication calendars are not fixed and can vary between the years.

[16] The occurrence of "herd behavior" among institutions is not particularly remarkable. Forecasters study approximately the same information and have access to each other's forecasts and analyses.

Table 3. *Descriptive statistics for absolute forecast errors, 1999–2015*

			Panel (a) GDP				*Panel (b) CPI*			
	Number of forecasts	Mean horizon	Mean value	Standard deviation	Min	Max	Mean value	Standard deviation	Min	Max
MoF	84	11.7	1.15	1.34	0.0	6.7	0.51	0.60	0.0	3.1
HUI	136	11.3	1.16	1.33	0.0	7.5	0.56	0.62	0.0	3.5
NIER	141	11.7	1.17	1.46	0.0	7.7	0.48	0.63	0.0	3.2
LO	69	12.1	1.25	1.40	0.0	7.2	0.51	0.52	0.0	2.3
NDA	125	12.6	1.13	1.22	0.0	6.9	0.59	0.65	0.0	2.9
RB	164	11.5	1.09	1.30	0.0	6.9	0.55	0.73	0.0	3.8
SEB	136	12.3	1.09	1.24	0.0	6.5	0.51	0.58	0.0	3.5
SHB	115	12.0	1.27	1.36	0.0	7.0	0.56	0.67	0.0	3.3
SN	121	12.1	1.23	1.40	0.0	7.5	0.59	0.68	0.0	3.3
SWED	92	13.5	1.35	1.44	0.0	7.7	0.64	0.70	0.0	3.1
Overall	N=1183	12.0	1.18	1.34	0.0	7.7	0.55	0.65	0.0	3.8
Between	n =10			0.08	1.1	1.6		0.05	0.5	0.6
Within	N/n=118.3			1.34	-0.2	7.7		0.64	-0.1	3.8

Notes: The table describes the absolute errors from the forecasts published by institutions over the period 1999–2015. In total there are 10 forecasters with a total of 1183 forecasts. The overall average of the absolute forecast error is 1.18 for GDP and 0.55 for CPI. *Overall* refers to the entire dataset. The *Between* variation shows the variation of the means of each institution (across time periods), and *Within* refers to the variation about the respective mean of each institution. Note that the between and the within standard deviation does not add up to the overall deviation since the two standard deviations are not calculated around the same mean value.

The models include a full set of fixed forecast year and individual-specific effects. As expected, the contribution from the distance between the forecast and the outcome on the absolute forecast error is positive and significant for both variables. The estimated coefficient for the timing suggests that the marginal effect on the absolute forecast error, of producing the forecast with full information instead of no information of the outcome, is around 1.4 percent for GDP and 1.1 percent for CPI.

We cannot reject that the estimated effect of the information content is common for all forecasters. This is tested by the joint hypothesis $\delta_{i=1} = \delta_{i=2} = , \ldots, \delta_{i=10}$. Moreover, the forecast year effects are strongly significant for both variables. This means that some years have been harder to predict than others. This effect is included in the model to separate the difficulty of different years from the individual effect, which we call ability. We test differences in the ability of the various institutions with a joint F-test. The test finds no significant differences between the institutions' accuracy for the assessed variables. This is because the variation between the forecasters is much smaller than the variation within each forecaster.

Figure 3 plots the observed absolute forecast errors for GDP and CPI for the horizons together with a regression line according to the estimations of timing from Panel A in Table 4. The regression lines represent the best fits of the whole sample of the data, and can be used to adjust the absolute forecast errors with respect to the timing effect. This is done by subtracting the estimated

Table 4. *Estimation results*

	Panel A						**Panel B**					
	GDP			**CPI**			**GDP**			**CPI**		
	Coef.	SE	*p*-val.	Coef.	SE	*p*-val.	Coef.	SE	*p*-val.	Coef.	SE	*p*-val.
Timing	1.425***	0.071	0.000	1.085***	0.036	0.000						
*Timing*Year*												
1999							0.722***	0.189	0.000	1.092***	0.114	0.000
2000							0.262*	0.139	0.060	0.259***	0.067	0.000
2001							2.212***	0.138	0.000	0.616***	0.114	0.000
2002							0.820***	0.101	0.000	0.570***	0.051	0.000
2003							1.208***	0.100	0.000	0.205*	0.118	0.083
2004							1.084***	0.114	0.000	1.179***	0.111	0.000
2005							0.022	0.067	0.738	1.230***	0.077	0.000
2006							1.558***	0.077	0.000	0.211***	0.053	0.000
2007							-0.329***	0.124	0.008	0.270***	0.078	0.001
2008							1.944***	0.119	0.000	1.063***	0.091	0.000
2009							7.563***	0.196	0.000	3.412***	0.154	0.000
2010							3.713***	0.192	0.000	0.316***	0.078	0.000
2011							0.549***	0.119	0.000	1.019***	0.124	0.000
2012							1.429***	0.147	0.000	1.448***	0.093	0.000
2013							0.211*	0.114	0.064	1.468***	0.089	0.000
2014							0.280***	0.081	0.001	1.732***	0.087	0.000
2015							0.194*	0.103	0.061	1.695***	0.113	0.000
Year effects	Yes***			Yes***			Yes***			Yes***		
Individual effects	Yes			Yes			Yes***			Yes**		
Individuals	10			10			10			10		
Observations	1183			1183			1183			1183		
Adj. R-squared	0.642			0.539			0.907			0.745		
RMSE	0.803			0.438			0.410			0.326		
F(16, 1140)							130.07			59.61		
Prob.>F							0.000			0.000		

Notes: *** denotes statistically significant at the 1 percent level, ** significant at the 5 percent level and * significant at the 10 percent level. The table shows selected estimates for 10 Swedish forecasters over the period 1999–2015 together with robust standard errors (SE). *Year* is the the forecast year and not the year of the forecast origin. In Panel A the reported estimates are based on equation (8), while Panel B show estimates from equation (9).

effect from individual absolute errors. The individual effect (μ_i) represent each forecaster's mean deviation from the estimated line (after accounting for year effects). Thus, the estimated individual effect is appropriate to assess and compare the forecasting ability of the different institutions.

It is clear from Figure 3 that the variability in the errors increases with the forecast horizon. Although, the estimated marginal effects presented in Table 4 are still unbiased and consistent, the longer horizons will dominate the short horizons.[17] This is due to the fact that errors are generally further from the expected error (fitted line) in the long horizons and will therefore be adjusted more.

The increasing variability in the errors with the horizon is to a large extent given by the fact that the difficulty to forecast a variable can differ across years.

[17]The presence of heteroskedasticity should not alter the 'central' position of the OLS line. Least square estimates are still unbiased and consistent but not efficient. Thus, the estimated standard errors are biased, so the usual confidence intervals and test statistics are incorrect. The standard errors reported in Table 4 are, however, corrected for heteroskedasticity using White's robust standard errors (White, 1980). Clustering the standard errors on year or institutions does not alter the results.

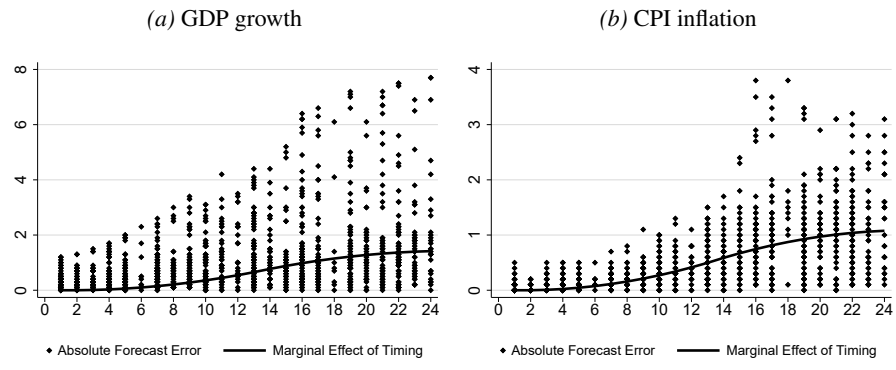


Figure 3. Absolute forecast errors and estimated average effect of timing

This means that both the size of the error and the marginal timing effect can differ across years. To illustrate this we also estimate equation (9), and thus allow for time-varying marginal timing effects (i.e., δ_t). Panel B of Table 4 clearly suggests that there are large and significant variations in the estimated timing effects across different years for both GDP and CPI. The hypothesis of homogeneous timing effects across forecast years is strongly rejected.[18] Subsequently, a model that accounts for heterogeneous timing effects across years is required to capture the behavior of the forecast errors correctly. One caveat with this specification is that it is more sensitive to the number of observations in each year. Our data set has about 60–70 observations per year.

All years, except for GDP in 2007, exhibit the expected timing effect. This is shown by positive estimates in Panel B in Table 4. The GDP timing coefficient for 2007 is negative, which is driven by data.[19] In an application one may think this is appropriate, whereas, it is also possible to restrict the coefficients to be non-negative.

4.3 Estimated Abilities and Rankings

For the flexible specification in Panel B in Table 4 the joint F-test on the individual-specific (ability) effects becomes significant. Thus, a more care-

[18] A joint F-test, presented in Panel B of Table 4, of parameter constancy test of the hypothesis that the coefficients are identical across years is rejected at the 0.0001 level for both GDP and CPI.

[19] During 2007 the forecasters seem to been presented to GDP disinformation as time evolved. In practice, the outcomes during the first half of 2007 were stronger than the forecasts, which led to upward revisions. When the third quarter outcome was published, the first and second quarter outcomes were revised downwards. This had a non-negligible impact on the 2007 annual average, since these quarters together weigh 7/16. See Sveriges Riksbank (2008) for more information. As a consequence this led to the unusual pattern where early forecasts were more accurate than later ones.

ful modeling of the forecast errors pick up systematic differences between different forecasters' GDP and CPI predictions over time.[20]

The estimated individual effects from equation (9), which are interpreted as each institution's forecast ability (adjusted for the differences in timing when forecasting) can be used for a fair comparison of the performance. In Figure 4 we plot the estimated ability for each institution (for GDP and CPI respectively) for the entire period 1999–2015. The figure presents two measures. The black bars show the ability, that is the adjusted mean absolute error, estimated from the individual effects obtained from Panel B in Table 4. This ability is expressed as deviations from the all-institutions average, $\mu_i^* = \hat{\mu}_i - \frac{1}{n}\sum_{i=1}^{n}\hat{\mu}_i$. A positive bar ($\mu_i^* > 0$) corresponds to a larger adjusted forecast error than the average forecaster and thus a lower forecast ability, whereas a negative bar ($\mu_i^* < 0$) implies the opposite. The gray bars represent an MAE statistic, also stated as deviations from the all-institutions average MAE.

Figure 4, ordered by the black bars, shows that there exist differences in the forecasting ability between the assessed institutions. In pair-wise comparisons these differences are sometimes quite small but in other cases non-negligible. Furthermore, it is noted that the two measures give rise to somewhat different rankings between the institutions. Thus, an adjustment of the errors with respect to the differences in timing has implications on the comparison.

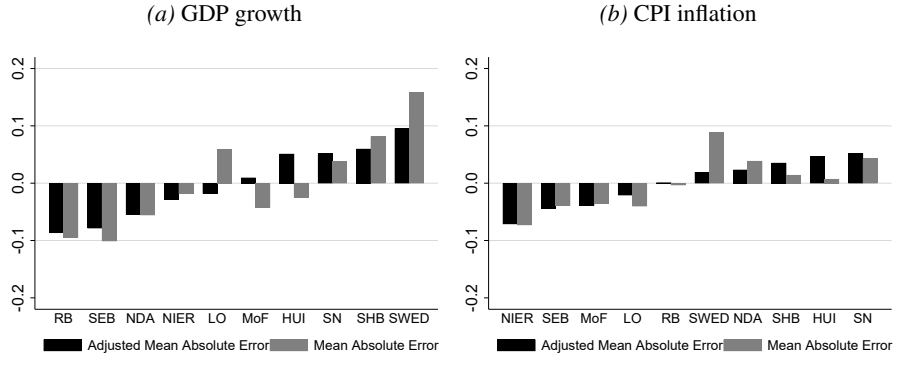


Figure 4. Estimated relative forecast ability

Notes: Adjusted Mean Absolute Error refers to estimated ability according to our proposed model in equation (9). Mean Absolute Error refers to the standard MAE statistic. Both series are demeaned by the overall mean based on all institutions. Thus, a negative value tells us that the individual have a forecast performance above the average.

[20]This joint test investigates the null hypothesis that all forecasters share the same ability against the alternative that at least one of the forecasters ability differs from the others.

5 Conclusions

Different forecasters typically prepare and publish their forecasts at different points in time. This means that the forecasts are based on different amounts of information. When the forecasters' performance is assessed and compared this may severely distort the conclusions. This paper introduces a way to handle these problems by means of a simultaneous estimation of the forecasters' ability and the effect of different information content. We then use the ability measure to compare the forecasters. We argue that a forecast comparison that does not adjust for differences in timing may be misleading due to information advantage or disadvantage. This view is supported by Monte Carlo evidence. The Monte Carlo simulation also concludes that increasing the evaluation sample size is beneficial to the proposed method. The implication of this is that a single year evaluation is not sufficient to retrieve reliable estimates of the forecasters' ability. Finally, the method is applied to a real-world dataset of 10 Swedish forecasters for the sample period 1999–2015. These forecasters are ranked based on their ability to forecast Swedish GDP and CPI inflation. Our first finding is that the timing of the forecasts does matter for the forecast error. Thus, comparisons based on unadjusted mean absolute errors may give unfair results. We conclude that our proposed method is preferred over the traditional statistic when comparing the performance of forecasters.

References

Andersson, M. K., Karlsson, G., & Svensson, J. (2007). An evaluation of the Riksbank's forecasting performance. *Sveriges Riksbank Economic Review*, *3*, 59–75.

Baltagi, B. H. (2001). *Econometric analysis of panel data* (2nd ed.). John Wiley and Sons.

Bank of England. (2015). Evaluating forecast performance.

Bauer, A., Eisenbeis, R. A., Waggoner, D. F., & Zha, T. (2003). Forecast evaluation with cross-sectional data: The blue chip surveys. *Federal Reserve Bank of Atlanta Economic Review*, *88*(2), 17–31.

Blix, M., Wadefjord, J., Wienecke, U., & Ådahl, M. (2001). How good is the forecasting performance of major institutions? *Sveriges Riksbank Economic Review*, *3*, 38–68.

Boero, G., Smith, J., & Wallis, K. F. (2008). Evaluating a three-dimensional panel of point forecasts: The Bank of England Survey of External Forecasters. *International Journal of Forecasting*, (24), 354–367.

Cabanillas, L. G., & Terzi, A. (2012). *The accuracy of the European Commission's forecasts re-examined* (European Economy – Economic Papers No. 476). Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Cipullo, D., & Reslow, A. (2021). Biased Forecast to Affect Voting Decisions? The "Brexit" Case. In A. Reslow (Ed.), *Electoral Incentives and Information Content in Macroeconomic Forecasts*. Department of Economics, Uppsala University.

Davies, A., & Lahiri, K. (1995). A new Framework for Analyzing Survey Forecasts using Three-dimensional Panel Data. *Journal of Econometrics*, *68*(1), 205–228.

Davies, A. (2006). A framework for decomposing shocks and measuring volatilities derived from multi-dimensional panel data of survey forecasts. *International Journal of Forecasting*, (22), 373–393.

Diebold, F. X. (2007). *Elements of Forecasting* (4th ed.). Thomson, South-Western.

European Central Bank. (2013). An assessment of eurosystem staff Macroeconomic projections. *ECB Monthly Bulletin*, 71–83.

Gneiting, T. (2011). Making and Evaluating Point Forecasts. *Journal of the American Statistical Association*, *106*(494), 746–762.

Goh, K. L., & Lawrence, D. (2006). *Treasury's Forecasting Performance: A Head-to-Head Comparison* (New Zealand Treasury Working Paper No. 06/10).

Patton, A. J. (2015). *Evaluating and Comparing Possibly Misspecified Forecasts* (Working paper, March 2015 version).

Sveriges Riksbank. (2008). Material for assessing monetary policy 2005–2007.

Sveriges Riksbank. (2016). Account of Monetary Policy 2015.

Timmermann, A. (2007). *An Evaluation of the World Economic Outlook Forecasts* (IMF Staff Papers No. 1).

Vogel, L. (2007). *How Do the OECD Growth Projections for the G7 Economies Perform? A Post-Mortem* (OECD Working Papers No. 573).

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskadasticity. *Econometrica*, *48*(4), 817–838.

Economic Studies

1987:1 Haraldson, Marty. To Care and To Cure. A linear programming approach to national health planning in developing countries. 98 pp.

1989:1 Chryssanthou, Nikos. The Portfolio Demand for the ECU. A Transaction Cost Approach. 42 pp.

1989:2 Hansson, Bengt. Construction of Swedish Capital Stocks, 1963-87. An Application of the Hulten-Wykoff Studies. 37 pp.

1989:3 Choe, Byung-Tae. Some Notes on Utility Functions Demand and Aggregation. 39 pp.

1989:4 Skedinger, Per. Studies of Wage and Employment Determination in the Swedish Wood Industry. 89 pp.

1990:1 Gustafson, Claes-Håkan. Inventory Investment in Manufacturing Firms. Theory and Evidence. 98 pp.

1990:2 Bantekas, Apostolos. The Demand for Male and Female Workers in Swedish Manufacturing. 56 pp.

1991:1 Lundholm, Michael. Compulsory Social Insurance. A Critical Review. 109 pp.

1992:1 Sundberg, Gun. The Demand for Health and Medical Care in Sweden. 58 pp.

1992:2 Gustavsson, Thomas. No Arbitrage Pricing and the Term Structure of Interest Rates. 47 pp.

1992:3 Elvander, Nils. Labour Market Relations in Sweden and Great Britain. A Comparative Study of Local Wage Formation in the Private Sector during the 1980s. 43 pp.

12 Dillén, Mats. Studies in Optimal Taxation, Stabilization, and Imperfect Competition. 1993. 143 pp.

13 Banks, Ferdinand E.. A Modern Introduction to International Money, Banking and Finance. 1993. 303 pp.

14 Mellander, Erik. Measuring Productivity and Inefficiency Without Quantitative Output Data. 1993. 140 pp.

15 Ackum Agell. Susanne. Essays on Work and Pay. 1993. 116 pp.

16 Eriksson, Claes. Essays on Growth and Distribution. 1994. 129 pp.

17 Banks, Ferdinand E.. A Modern Introduction to International Money, Banking and Finance. 2nd version, 1994. 313 pp.

18 Apel, Mikael. Essays on Taxation and Economic Behavior. 1994. 144 pp.

19 Dillén, Hans. Asset Prices in Open Monetary Economies. A Contingent Claims Approach. 1994. 100 pp.

20 Jansson, Per. Essays on Empirical Macroeconomics. 1994. 146 pp.

21 Banks, Ferdinand E.. A Modern Introduction to International Money, Banking, and Finance. 3rd version, 1995. 313 pp.

22 Dufwenberg, Martin. On Rationality and Belief Formation in Games. 1995. 93 pp.

23 Lindén, Johan. Job Search and Wage Bargaining. 1995. 127 pp.

24 Shahnazarian, Hovick. Three Essays on Corporate Taxation. 1996. 112 pp.

25 Svensson, Roger. Foreign Activities of Swedish Multinational Corporations. 1996. 166 pp.

26 Sundberg, Gun. Essays on Health Economics. 1996. 174 pp.

27 Sacklén, Hans. Essays on Empirical Models of Labor Supply. 1996. 168 pp.

28 Fredriksson, Peter. Education, Migration and Active Labor Market Policy. 1997. 106 pp.

29 Ekman, Erik. Household and Corporate Behaviour under Uncertainty. 1997. 160 pp.

30 Stoltz, Bo. Essays on Portfolio Behavior and Asset Pricing. 1997. 122 pp.

31 Dahlberg, Matz. Essays on Estimation Methods and Local Public Economics. 1997. 179 pp.

32 Kolm, Ann-Sofie. Taxation, Wage Formation, Unemployment and Welfare. 1997. 162 pp.

33 Boije, Robert. Capitalisation, Efficiency and the Demand for Local Public Services. 1997. 148 pp.

34 Hort, Katinka. On Price Formation and Quantity Adjustment in Swedish Housing Markets. 1997. 185 pp.

35 Lindström, Thomas. Studies in Empirical Macroeconomics. 1998. 113 pp.

36 Hemström, Maria. Salary Determination in Professional Labour Markets. 1998. 127 pp.

37 Forsling, Gunnar. Utilization of Tax Allowances and Corporate Borrowing. 1998. 96 pp.

38 Nydahl, Stefan. Essays on Stock Prices and Exchange Rates. 1998. 133 pp.

39 Bergström, Pål. Essays on Labour Economics and Econometrics. 1998. 163 pp.

40 Heiborn, Marie. Essays on Demographic Factors and Housing Markets. 1998. 138 pp.

41 Åsberg, Per. Four Essays in Housing Economics. 1998. 166 pp.

42 Hokkanen, Jyry. Interpreting Budget Deficits and Productivity Fluctuations. 1998. 146 pp.

43 Lunander, Anders. Bids and Values. 1999. 127 pp.

44 Eklöf, Matias. Studies in Empirical Microeconomics. 1999. 213 pp.

45 Johansson, Eva. Essays on Local Public Finance and Intergovernmental Grants. 1999. 156 pp.

46 Lundin, Douglas. Studies in Empirical Public Economics. 1999. 97 pp.

47 Hansen, Sten. Essays on Finance, Taxation and Corporate Investment. 1999. 140 pp.

48 Widmalm, Frida. Studies in Growth and Household Allocation. 2000. 100 pp.

49 Arslanogullari, Sebastian. Household Adjustment to Unemployment. 2000. 153 pp.

50 Lindberg, Sara. Studies in Credit Constraints and Economic Behavior. 2000. 135 pp.

51 Nordblom, Katarina. Essays on Fiscal Policy, Growth, and the Importance of Family Altruism. 2000. 105 pp.

52 Andersson, Björn. Growth, Saving, and Demography. 2000. 99 pp.

53 Åslund, Olof. Health, Immigration, and Settlement Policies. 2000. 224 pp.

54 Bali Swain, Ranjula. Demand, Segmentation and Rationing in the Rural Credit Markets of Puri. 2001. 160 pp.

55 Löfqvist, Richard. Tax Avoidance, Dividend Signaling and Shareholder Taxation in an Open Economy. 2001. 145 pp.

56 Vejsiu, Altin. Essays on Labor Market Dynamics. 2001. 209 pp.

57 Zetterström, Erik. Residential Mobility and Tenure Choice in the Swedish Housing Market. 2001. 125 pp.

58 Grahn, Sofia. Topics in Cooperative Game Theory. 2001. 106 pp.

59 Laséen, Stefan. Macroeconomic Fluctuations and Microeconomic Adjustments. Wages, Capital, and Labor Market Policy. 2001. 142 pp.

60 Arnek, Magnus. Empirical Essays on Procurement and Regulation. 2002. 155 pp.

61 Jordahl, Henrik. Essays on Voting Behavior, Labor Market Policy, and Taxation. 2002. 172 pp.

62 Lindhe, Tobias. Corporate Tax Integration and the Cost of Capital. 2002. 102 pp.

63 Hallberg, Daniel. Essays on Household Behavior and Time-Use. 2002. 170 pp.

64 Larsson, Laura. Evaluating Social Programs: Active Labor Market Policies and Social Insurance. 2002. 126 pp.

65 Bergvall, Anders. Essays on Exchange Rates and Macroeconomic Stability. 2002. 122 pp.

66 Nordström Skans, Oskar. Labour Market Effects of Working Time Reductions and Demographic Changes. 2002. 118 pp.

67 Jansson, Joakim. Empirical Studies in Corporate Finance, Taxation and Investment. 2002. 132 pp.

68 Carlsson, Mikael. Macroeconomic Fluctuations and Firm Dynamics: Technology, Production and Capital Formation. 2002. 149 pp.

69 Eriksson, Stefan. The Persistence of Unemployment: Does Competition between Employed and Unemployed Job Applicants Matter? 2002. 154 pp.

70 Huitfeldt, Henrik. Labour Market Behaviour in a Transition Economy: The Czech Experience. 2003. 110 pp.

71 Johnsson, Richard. Transport Tax Policy Simulations and Satellite Accounting within a CGE Framework. 2003. 84 pp.

72 Öberg, Ann. Essays on Capital Income Taxation in the Corporate and Housing Sectors. 2003. 183 pp.

73 Andersson, Fredrik. Causes and Labor Market Consequences of Producer Heterogeneity. 2003. 197 pp.

74 Engström, Per. Optimal Taxation in Search Equilibrium. 2003. 127 pp.

75 Lundin, Magnus. The Dynamic Behavior of Prices and Investment: Financial Constraints and Customer Markets. 2003. 125 pp.

76 Ekström, Erika. Essays on Inequality and Education. 2003. 166 pp.

77 Barot, Bharat. Empirical Studies in Consumption, House Prices and the Accuracy of European Growth and Inflation Forecasts. 2003. 137 pp.

78 Österholm, Pär. Time Series and Macroeconomics: Studies in Demography and Monetary Policy. 2004. 116 pp.

79 Bruér, Mattias. Empirical Studies in Demography and Macroeconomics. 2004. 113 pp.

80 Gustavsson, Magnus. Empirical Essays on Earnings Inequality. 2004. 154 pp.

81 Toll, Stefan. Studies in Mortgage Pricing and Finance Theory. 2004. 100 pp.

82 Hesselius, Patrik. Sickness Absence and Labour Market Outcomes. 2004. 109 pp.

83 Häkkinen, Iida. Essays on School Resources, Academic Achievement and Student Employment. 2004. 123 pp.

84 Armelius, Hanna. Distributional Side Effects of Tax Policies: An Analysis of Tax Avoidance and Congestion Tolls. 2004. 96 pp.

85 Ahlin, Åsa. Compulsory Schooling in a Decentralized Setting: Studies of the Swedish Case. 2004. 148 pp.

86 Heldt, Tobias. Sustainable Nature Tourism and the Nature of Tourists' Cooperative Behavior: Recreation Conflicts, Conditional Cooperation and the Public Good Problem. 2005. 148 pp.

87 Holmberg, Pär. Modelling Bidding Behaviour in Electricity Auctions: Supply Function Equilibria with Uncertain Demand and Capacity Constraints. 2005. 43 pp.

88 Welz, Peter. Quantitative new Keynesian macroeconomics and monetary policy 2005. 128 pp.

89 Ågren, Hanna. Essays on Political Representation, Electoral Accountability and Strategic Interactions. 2005. 147 pp.

90 Budh, Erika. Essays on environmental economics. 2005. 115 pp.

91 Chen, Jie. Empirical Essays on Housing Allowances, Housing Wealth and Aggregate Consumption. 2005. 192 pp.

92 Angelov, Nikolay. Essays on Unit-Root Testing and on Discrete-Response Modelling of Firm Mergers. 2006. 127 pp.

93 Savvidou, Eleni. Technology, Human Capital and Labor Demand. 2006. 151 pp.

94 Lindvall, Lars. Public Expenditures and Youth Crime. 2006. 112 pp.

95 Söderström, Martin. Evaluating Institutional Changes in Education and Wage Policy. 2006. 131 pp.

96 Lagerström, Jonas. Discrimination, Sickness Absence, and Labor Market Policy. 2006. 105 pp.

97 Johansson, Kerstin. Empirical essays on labor-force participation, matching, and trade. 2006. 168 pp.

98 Ågren, Martin. Essays on Prospect Theory and the Statistical Modeling of Financial Returns. 2006. 105 pp.

99 Nahum, Ruth-Aïda. Studies on the Determinants and Effects of Health, Inequality and Labour Supply: Micro and Macro Evidence. 2006. 153 pp.

100 Žamac, Jovan. Education, Pensions, and Demography. 2007. 105 pp.

101 Post, Erik. Macroeconomic Uncertainty and Exchange Rate Policy. 2007. 129 pp.

102 Nordberg, Mikael. Allies Yet Rivals: Input Joint Ventures and Their Competitive Effects. 2007. 122 pp.

103 Johansson, Fredrik. Essays on Measurement Error and Nonresponse. 2007. 130 pp.

104 Haraldsson, Mattias. Essays on Transport Economics. 2007. 104 pp.

105 Edmark, Karin. Strategic Interactions among Swedish Local Governments. 2007. 141 pp.

106 Oreland, Carl. Family Control in Swedish Public Companies. Implications for Firm Performance, Dividends and CEO Cash Compensation. 2007. 121 pp.

107 Andersson, Christian. Teachers and Student Outcomes: Evidence using Swedish Data. 2007. 154 pp.

108 Kjellberg, David. Expectations, Uncertainty, and Monetary Policy. 2007. 132 pp.

109 Nykvist, Jenny. Self-employment Entry and Survival - Evidence from Sweden. 2008. 94 pp.

110 Selin, Håkan. Four Empirical Essays on Responses to Income Taxation. 2008. 133 pp.

111 Lindahl, Erica. Empirical studies of public policies within the primary school and the sickness insurance. 2008. 143 pp.

112 Liang, Che-Yuan. Essays in Political Economics and Public Finance. 2008. 125 pp.

113 Elinder, Mikael. Essays on Economic Voting, Cognitive Dissonance, and Trust. 2008. 120 pp.

114 Grönqvist, Hans. Essays in Labor and Demographic Economics. 2009. 120 pp.

115 Bengtsson, Niklas. Essays in Development and Labor Economics. 2009. 93 pp.

116 Vikström, Johan. Incentives and Norms in Social Insurance: Applications, Identification and Inference. 2009. 205 pp.

117 Liu, Qian. Essays on Labor Economics: Education, Employment, and Gender. 2009. 133 pp.

118 Glans, Erik. Pension reforms and retirement behaviour. 2009. 126 pp.

119 Douhan, Robin. Development, Education and Entrepreneurship. 2009.

120 Nilsson, Peter. Essays on Social Interactions and the Long-term Effects of Early-life Conditions. 2009. 180 pp.

121 Johansson, Elly-Ann. Essays on schooling, gender, and parental leave. 2010. 131 pp.

122 Hall, Caroline. Empirical Essays on Education and Social Insurance Policies. 2010. 147 pp.

123 Enström-Öst, Cecilia. Housing policy and family formation. 2010. 98 pp.

124 Winstrand, Jakob. Essays on Valuation of Environmental Attributes. 2010. 96 pp.

125 Söderberg, Johan. Price Setting, Inflation Dynamics, and Monetary Policy. 2010. 102 pp.

126 Rickne, Johanna. Essays in Development, Institutions and Gender. 2011. 138 pp.

127 Hensvik, Lena. The effects of markets, managers and peers on worker outcomes. 2011. 179 pp.

128 Lundqvist, Heléne. Empirical Essays in Political and Public. 2011. 157 pp.

129 Bastani, Spencer. Essays on the Economics of Income Taxation. 2012. 257 pp.

130 Corbo, Vesna. Monetary Policy, Trade Dynamics, and Labor Markets in Open Economies. 2012. 262 pp.

131 Nordin, Mattias. Information, Voting Behavior and Electoral Accountability. 2012. 187 pp.

132 Vikman, Ulrika. Benefits or Work? Social Programs and Labor Supply. 2013. 161 pp.

133 Ek, Susanne. Essays on unemployment insurance design. 2013. 136 pp.

134 Österholm, Göran. Essays on Managerial Compensation. 2013. 143 pp.

135 Adermon, Adrian. Essays on the transmission of human capital and the impact of technological change. 2013. 138 pp.

136 Kolsrud, Jonas. Insuring Against Unemployment 2013. 140 pp.

137 Hanspers, Kajsa. Essays on Welfare Dependency and the Privatization of Welfare Services. 2013. 208 pp.

138 Persson, Anna. Activation Programs, Benefit Take-Up, and Labor Market Attachment. 2013. 164 pp.

139 Engdahl, Mattias. International Mobility and the Labor Market. 2013. 216 pp.

140 Krzysztof Karbownik. Essays in education and family economics. 2013. 182 pp.

141 Oscar Erixson. Economic Decisions and Social Norms in Life and Death Situations. 2013. 183 pp.

142 Pia Fromlet. Essays on Inflation Targeting and Export Price Dynamics. 2013. 145 pp.

143 Daniel Avdic. Microeconometric Analyses of Individual Behavior in Public Welfare Systems. Applications in Health and Education Economics. 2014. 176 pp.

144 Arizo Karimi. Impacts of Policies, Peers and Parenthood on Labor Market Outcomes. 2014. 221 pp.

145 Karolina Stadin. Employment Dynamics. 2014. 134 pp.

146 Haishan Yu. Essays on Environmental and Energy Economics. 132 pp.

147 Martin Nilsson. Essays on Health Shocks and Social Insurance. 139 pp.

148 Tove Eliasson. Empirical Essays on Wage Setting and Immigrant Labor Market Opportunities. 2014. 144 pp.

149 Erik Spector. Financial Frictions and Firm Dynamics. 2014. 129 pp.

150 Michihito Ando. Essays on the Evaluation of Public Policies. 2015. 193 pp.

151 Selva Bahar Baziki. Firms, International Competition, and the Labor Market. 2015. 183 pp.

152 Fredrik Sävje. What would have happened? Four essays investigating causality. 2015. 229 pp.

153 Ina Blind. Essays on Urban Economics. 2015. 197 pp.

154 Jonas Poulsen. Essays on Development and Politics in Sub-Saharan Africa. 2015. 240 pp.

155 Lovisa Persson. Essays on Politics, Fiscal Institutions, and Public Finance. 2015. 137 pp.

156 Gabriella Chirico Willstedt. Demand, Competition and Redistribution in Swedish Dental Care. 2015. 119 pp.

157 Yuwei Zhao de Gosson de Varennes. Benefit Design, Retirement Decisions and Welfare Within and Across Generations in Defined Contribution Pension Schemes. 2016. 148 pp.

158 Johannes Hagen. Essays on Pensions, Retirement and Tax Evasion. 2016. 195 pp.

159 Rachatar Nilavongse. Housing, Banking and the Macro Economy. 2016. 156 pp.

160 Linna Martén. Essays on Politics, Law, and Economics. 2016. 150 pp.

161 Olof Rosenqvist. Essays on Determinants of Individual Performance and Labor Market Outcomes. 2016. 151 pp.

162 Linuz Aggeborn. Essays on Politics and Health Economics. 2016. 203 pp.

163 Glenn Mickelsson. DSGE Model Estimation and Labor Market Dynamics. 2016. 166 pp.

164 Sebastian Axbard. Crime, Corruption and Development. 2016. 150 pp.

165 Mattias Öhman. Essays on Cognitive Development and Medical Care. 2016. 181 pp.

166 Jon Frank. Essays on Corporate Finance and Asset Pricing. 2017. 160 pp.

167 Ylva Moberg. Gender, Incentives, and the Division of Labor. 2017. 220 pp.

168 Sebastian Escobar. Essays on inheritance, small businesses and energy consumption. 2017. 194 pp.

169 Evelina Björkegren. Family, Neighborhoods, and Health. 2017. 226 pp.

170 Jenny Jans. Causes and Consequences of Early-life Conditions. Alcohol, Pollution and Parental Leave Policies. 2017. 209 pp.

171 Josefine Andersson. Insurances against job loss and disability. Private and public interventions and their effects on job search and labor supply. 2017. 175 pp.

172 Jacob Lundberg. Essays on Income Taxation and Wealth Inequality. 2017. 173 pp.

173 Anna Norén. Caring, Sharing, and Childbearing. Essays on Labor Supply, Infant Health, and Family Policies. 2017. 206 pp.

174 Irina Andone. Exchange Rates, Exports, Inflation, and International Monetary Cooperation. 2018. 174 pp.

175 Henrik Andersson. Immigration and the Neighborhood. Essays on the Causes and Consequences of International Migration. 2018. 181 pp.

176 Aino-Maija Aalto. Incentives and Inequalities in Family and Working Life. 2018. 131 pp.

177 Gunnar Brandén. Understanding Intergenerational Mobility. Inequality, Student Aid and Nature-Nurture Interactions. 2018. 125 pp.

178 Mohammad H. Sepahvand. Essays on Risk Attitudes in Sub-Saharan Africa. 2019. 215 pp.

179 Mathias von Buxhoeveden. Partial and General Equilibrium Effects of Unemployment Insurance. Identification, Estimation and Inference. 2019. 89 pp.

180 Stefano Lombardi. Essays on Event History Analysis and the Effects of Social Programs on Individuals and Firms. 2019. 150 pp.

181 Arnaldur Stefansson. Essays in Public Finance and Behavioral Economics. 2019. 191 pp.

182 Cristina Bratu. Immigration: Policies, Mobility and Integration. 2019. 173 pp.

183 Tamás Vasi. Banks, Shocks and Monetary Policy. 2020. 148 pp.

184 Jonas Cederlöf. Job Loss: Consequences and Labor Market Policy. 2020. 213 pp.

185 Dmytro Stoyko. Expectations, Financial Markets and Monetary Policy. 2020. 153 pp.

186 Paula Roth. Essays on Inequality, Insolvency and Innovation. 2020. 191 pp.

187 Fredrik Hansson. Consequences of Poor Housing, Essays on Urban and Health Economics. 2020. 143 pp.

188 Maria Olsson. Essays on Macroeconomics: Wage Rigidity and Aggregate Fluctuations. 2020. 130 pp.

189 Dagmar Müller. Social Networks and the School-to-Work Transition. 2020. 146 pp.

190 Maria Sandström. Essays on Savings and Intangible Capital. 2020. 129 pp.

191. Anna Thoresson. Wages and Their Impact on Individuals, Households and Firms. 2020. 220 pp.

192. Jonas Klarin. Empirical Essays in Public and Political Economics. 2020. 129 pp.

193. André Reslow. Electoral Incentives and Information Content in Macroeconomic Forecasts. 2021. 184 pp.