

STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
659 Degree project in economics
Spring 2020

The Use of Judgment Versus Models

Forecasting Accuracy in Sveriges Riksbank 2010-2014

Emma Hamre (24033) and Sandra Lanros (24234)

Abstract: In this study, we ask, how did the use of judgments and models affect the accuracy of the Riksbank's forecasts during the years 2010-2014? We examine the two components, judgment and models, that determine the Sveriges Riksbank's published forecasts for three macro indicators: GDP growth, CPIF inflation and the repo rate. The analysis compares the relative forecasting accuracy of the central bank's primary models, DSGE and BVAR, and the final published forecasts containing the judgmental adjustments of the Riksbank's Executive Board. The study draws on both graphical and econometric analyses. Our findings show that in the nowcast (one-two quarters), the published forecast outperformed the two model forecasts. The forecasting accuracy of the published forecasts, however, progressively worsened at four to eight quarters for CPIF inflation and the repo rate, whereas the BVAR model, compared to the DSGE model, systematically improved in accuracy. By incorporating judgment in the Riksbank's nowcasts, we find an improvement in accuracy – more so than what would have been the case if the models were solely used. Judgmental adjustments are thus important for accuracy in the short term.

Keywords: Central Banks, Forecasting, Judgment, Monetary Policy

JEL: E47, E52, E58, E65

Supervisor: Anna Dreber Almenberg
Date submitted: May 14, 2020
Date examined: May 25, 2020
Discussants: Carl Edvin Steinvall and Milton Manngård
Examiner: Johanna Wallenius

Acknowledgments

We would like to express our sincerest gratitude for the guidance of our supervisor Anna Dreber Almenberg and our co-supervisor Lars E.O. Svensson. We would also like to extend our gratitude toward Jesper Lindé for providing us with the data essential to our empirical analysis. Further, we are also grateful for the valuable discussions surrounding the analysis of our data and the feedback we received from Cristina Clerici. Lastly, we are thankful for the enlightening discussions with former and current Riksbank officials – helping us better understand the Riksbank’s forecasting process and certainly providing greater perspective on the period under review.

Table of Contents

1. Introduction	4
2. Background	6
2.1 Sveriges Riksbank	6
2.2 Describing the Riksbank's Monetary Policy	7
2.3 Defining the Riksbank's Main Macro Models	7
2.4 Judgment and Models in Forecasting	9
3. Literature Review	9
3.1 Goodfriend and King (2016)	10
3.2 Lindé and Reslow (2017)	10
3.3 Iversen et al. (2016)	10
4. Purpose of Study	11
5. Theoretical Framework	12
5.1 Bounded Rationality	12
5.2 Ecological Rationality	13
5.3 When is more preferred to less?	14
6. Data and Method	14
6.1 Data	14
6.1.1 Model Forecasts	15
6.1.2 Published Forecasts	15
6.1.3 Data Transformation	15
6.2 Complementary Qualitative Research	16
6.3 Defining the Period of Wishful Thinking	16
6.4 Multicollinearity, Serial Correlation and Sample Size Concerns	16
7. Empirical Strategy	17
7.1 Visual Inspection of Time Period and Real-Time Forecasts	17
7.2 Forecast Deviations and Bias	20
7.3 Testing Forecast Errors for Statistical Significance	23
7.3.1 GDP Growth Mean Squared Errors Comparisons	24
7.3.2 CPIF Inflation Mean Squared Errors Comparisons	25
7.3.3 Repo Rate Mean Squared Errors Comparisons	25
7.4 Optimal Weighting Ratios	26
7.4.1 Optimal GDP Growth Model Forecasts	27
7.4.2 Optimal CPIF Inflation Model Forecasts	28
7.4.3 Optimal Repo Rate Model Forecasts	30
8. Discussion	31
8.1 Visual Inspection and Forecast Deviation Sections	31
8.2 Testing Forecast Errors for Statistical Significance	34
8.3 Optimal Weighting Ratios	35
8.4 Limitations and Future Studies	37

9. Conclusion	38
10. References	40
Appendices.....	44
Appendix A.....	44
Appendix B.....	50
Appendix C.....	54
Appendix D.....	55
Appendix E	58
Appendix F	59

1. Introduction

The financial crisis of 2008 brought an abrupt end to the golden age for Central Bankers (Goodhart and Rochet 2011). It left policy makers around the world caring for weak economies many years following the height of the turmoil. Sweden was one of many countries strongly affected, experiencing a 6% drop in GDP from 2008/Q3 to 2009/Q3 (Goodfriend and King 2016). The country's sharp fall in GDP experienced in these two years stemmed not primarily from the housing collapse, but rather from the global decline in international trade (Goodhart and Rochet 2011). Sweden experienced a remarkable rebound in 2010, more than what can be said for similar industrialized countries. The evident rise in exports led many monetary policy makers in Sweden's central bank, Sveriges Riksbank, to believe raising interest rates was a reasonable response (Goodfriend and King 2016). But the years that followed proved this to be "wishful thinking".¹ In hindsight, the assumptions made about the Swedish economy during 2010-2014 have been described as overly optimistic. Criticism over the Riksbank's reasoning soon followed. It became clear the state of the economy was not quite at the level they had predicted – systematically overestimating inflation, among other things (Konjunkturinstitutet 2013).

The period also displayed a notable degree of polarization within the central bank's Executive Board of Governors (Goodfriend and King 2016). Differences in judgment expressed by the Deputy Governors during this time largely concerned the growing dissent over the desired repo rate path. The fundamental objectives of the Riksbank's monetary policy were also under question (Svensson 2010a). With a majority favoring a more aggressive response to accumulating household debt, the Riksbank dramatically hiked the repo rate in June 2010, even more so than the Bank of England and the US Federal Reserve (Lindé and Reslow 2017). Meanwhile in October 2010, the minority advocated keeping the repo rate unchanged and following a steadily increasing path to a level of 2.7 percent. Effective communication and debate appeared to be lacking. "The two sides talked past each other", describe Goodfriend and King (2016) in a review of the central bank's performance, explaining, "it was not helpful that minutes and interviews by Board members displayed a degree of brusqueness uncharacteristic of normal public debate in Sweden." (9) In other words, the disagreement within the Executive Board was not a secret. In a recent interview with a former Deputy Governor, they recall the conflicts as being abnormally public (Ekholm 2020). And the economic situation that unfolded provides substance to the argument that the majority was in fact too optimistic (Goodfriend and King 2016).

The question then is, what made the Riksbank's predictions so inaccurate? Ex ante forecasts are the backbone of the Riksbank's decision-making, and this paper furthermore aims to examine the two key components of the Riksbank's published forecasts – model forecasts and Executive Board judgment. The model forecasts referenced are produced by the Monetary Policy Department as preparatory material for each monetary policy decision (Sveriges Riksbank 2019a). Previous studies have reviewed the forecast performance of the Riksbank in relation to their use of judgment versus models in the final published forecasts, however, with differing conclusions.

¹ The period of "wishful thinking" was confirmed by former Deputy Governor Lars E.O. Svensson, and also affirmed in interviews by other Riksbank officials (Svensson 2010; Ekholm 2020; Hansson 2020).

One study finds the central bank guilty of an overreliance on models, while another responds with a counterargument that this is a myth (Goodfriend and King 2016; Lindé and Reslow 2017). Neither of these studies take a position as to which component provides the most accurate forecast. This has been done by a study that compares the forecasts produced by the models with the judgmental forecasts published by the Riksbank. The findings indicate that the model-based forecasts have been more accurate than the published forecasts that include judgmental adjustments by the Executive Board (Iversen et al. 2016).²

The aim of this paper is to examine whether this is the case for the specific period of wishful thinking. We will also perform additional elements of data analysis that have yet been done by these three main studies we will be referencing – that is, testing mean squared errors (MSE) for statistical significance and finding the optimal model weighting ratios. This paper will, in other words, investigate whether the element of judgment or model forecasts ultimately brought the Riksbank further or closer to the actual outcome during 2010-2014.

We will begin by providing an overview of the Riksbank's role as a central bank, followed by descriptions of the published forecasts' main macro models and the indicators our study will examine. The forecasting process, from model-making to voting, will also be covered in order to clarify how the two components of models and judgment come into play. We will also further describe the key studies upon which we compare and build our own framework of analysis, followed by a section dedicated to the theoretical body of work that aims to shed light on how we as humans process information. Is more information preferred to less, or vice versa?

The main form of analysis conducted is done through empirical study. The data analysis can be shortly summarized as being both graphical and econometric. In addition to the visual inspection of multiple graphs for all forecast horizons, we further analyze the performance of the period's nowcasts (one to two quarters). We do so by testing differences in mean squared errors for statistical significance and by comparing the derived optimal model forecasts with the published forecasts. Qualitative material – in the form of meetings minutes and interviews – and comparisons to previous research, allow us to further support our own findings in section 8.

We find that published forecasts outperformed both model forecasts in the nowcasts (one-two quarters). The forecasting accuracy of the published forecasts, however, progressively worsened for CPIF inflation and the repo rate for four-eight quarters, whereas the BVAR model compared to the DSGE model, systematically improved in accuracy for these two indicators. By incorporating judgment in the Riksbank's nowcasts, we find an improvement in accuracy – more so than what would have been the case if the models were solely used. Judgmental adjustments are thus important for accuracy in the short term. Whether judgmental adjustments could have been even more accurate in the nowcasts is another question and relates to a further study on the source of the criticized optimism connected to the period reviewed.

More specifically, this thesis is comprised of eleven sections, the second of which delves into relevant background information and key concepts this study will be handling. In section 3, we

² Goodfriend and King (2016) review the period of 2010-2015. Lindé and Reslow (2017) study the years 2006-2016. Finally, Iversen et al. (2016) evaluate the period of 2007-2013. See section 3 for further descriptions.

list previous studies that also uncover the forecasting components – judgments and models – and their respective performance and influence. Section 4 emphasizes the value of our research question and the contribution we aim to make. Section 5 explores relevant theory and the subsequent theoretical framework of this study. Section 6 describes the data collection and method of analysis, followed by section 7, presenting our analysis of our empirical data and the four main methods of analysis: visual inspection and real-time forecasts, forecast deviations and bias, testing forecast errors for statistical significance, and lastly, optimal weighting ratios. These results are then discussed in section 8, succeeded by concluding remarks in section 9. References and additional figures and tables are presented in section 10 and 11 respectively.

2. Background

2.1 Sveriges Riksbank

Central banks are fundamental institutions that influence economies nationally, and often, economies across their borders. Sveriges Riksbank, established in 1668, is the oldest of such institutions. It is a public authority under Sweden's Parliament, the Riksdag. At the core of all its monetary policy decisions lies the promise to maintain price stability and safeguard the secureness of the payment system (Wetterberg 2009). In entirety, it is entrusted with the following tasks: to issue and retain the value of money, decide upon monetary policy that aims to produce a stable, low rate of inflation, and to ensure safe and efficient payments in the economy (Sveriges Riksbank 2018).

The 1990s were transformative years for the Riksbank to say the very least. An extensive reassessment of monetary policy and the Riksbank's capabilities began following grave international recessions. The interest rate was raised to 500 percent with hope to defend the Swedish krona, but failed to do so (Sveriges Riksbank n.d.). Price stability became the new central objective following the unsuccessful defense of the Swedish krona in 1992 and the shift to a floating exchange rate. This new objective was anchored to the introduction of an inflation target at two percent in 1993 – a numerical inflation target typical of advanced countries (Svensson 2010b). The Riksbank first aimed to maintain the Swedish krona's domestic purchasing power after leaving the gold standard in 1931, becoming the first central bank to standardize price stability in monetary policy. This was however lost during the price regulations of WWII. The central bank was also given a more independent role in 1999 following changes to the Sveriges Riksbank Act. Law confirmed the Riksbank's present task of maintaining price stability on January 1, 1999 (Sveriges Riksbank n.d.). Several European central banks were also given greater independence in the 1990s following the Maastricht Treaty in 1992. Ever since 1999, six members of the Executive Board have led the Riksbank. These positions are currently appointed by the Riksdag's assigned General Council (Wetterberg 2009). Following the crisis of 2008, the Riksbank transitioned from the CPIX to the CPIF index – a consumer price index with a fixed mortgage rate (Iversen et al. 2016). All these developments were pivotal in shaping the Riksbank's current position and monetary policy objectives.

2.2 Describing the Riksbank's Monetary Policy

Its relative independence from Swedish Parliament has created an additional reason for transparency. This has primarily been achieved through the publication of a summarizing report of each monetary policy meeting. These meetings are held six times per annum and concern the future monetary policy decided by the Executive Board and their understanding of current economic development. In 2007, the central bank began publishing its own forecast for the repo rate path in effort to further increase transparency. All forecasts are made on a quarterly basis. As a result, the Deputy Governors of the Executive Board must participate in the forecasting process in a more definitive way – assessing what repo rate path they believe will provide for well-balanced monetary policy (Hallsten and Tägtström 2009). The repo rate path informs the general public and everyone in the financial markets of the Riksbank's expectations for the direction of monetary policy. However, as it disclaims on its website, "the interest rate path is a forecast, not a promise" (Sveriges Riksbank 2019c).

Models have supported the Riksbank monetary policy decisions since the inflation target of two percent was implemented. Stabilization, in many ways, is the ultimate goal. The Executive Board not only aims to stabilize inflation around the two percent target, but also to stabilize other macro indicators, such as employment and production, around their respective long-term paths (Sveriges Riksbank 2019b). To do so, the Riksbank conducts flexible inflation targeting. In a speech at the Netherlands Bank in 2009, former Deputy Governor Lars E.O. Svensson describes flexible inflation targeting as the following:

Flexible inflation targeting means that monetary policy aims at stabilizing *both* inflation around the inflation target and the real economy ... By stabilizing the real economy I mean stabilizing the resource utilization around a normal level, keeping in mind that monetary policy cannot affect the long-term level of resource utilization. (Svensson 2010b, 1)

Forecasts are hence critical in this approach to monetary policy. Monetary policy decisions are delayed in time lags and thus their effect on the various macro indicators – on both inflation and the real economy – is delayed (Svensson 2010b).

The forecasts that this study will be evaluating are those for GDP growth, CPIF inflation and the repo rate. Since 1994, the repo rate has been the Riksbank's so-called "policy rate" and main instrument for monetary policy. As described by the Riksbank, it "is the rate of interest at which banks can borrow or deposit funds at the Riksbank for a period of seven days" (Sveriges Riksbank 2020). The Riksbank is able to affect the interest rates and demand in the Swedish economy – and in doing so influence inflation – by raising or lowering the repo rate (Sveriges Riksbank 2020).

2.3 Defining the Riksbank's Main Macro Models

Models are central in the forecasting process. This study will examine the output produced by the dynamic stochastic general equilibrium (DSGE) model and the Bayesian vector autoregressive (BVAR) model. These are the main models of interests as they are central to the forecasting process and are both present in the key studies referenced in this paper, which allows for better comparison. The two main general equilibrium models are used for the medium-term forecast,

but also as material for the nowcasts (Iversen et al. 2016; Lindé and Reslow 2017). There is foreground to show that models have been bearing an increasingly important weight in forecasting activities at the Riksbank (Hallsten and Tågström 2009; Goodfriend and King 2016). Whether this is reflected in the final published forecasts, will be covered in the literature review. Regardless, it is reasonable to state that the models are of great value. Further, the empirical properties and forecasting performance of large structural macro models have seen large improvement. Advancements have also been made in the role of time series models that utilize historical economic data, as opposed to the sole leverage of economic theory (Hallsten and Tågström 2009). A high-ranking member of the Monetary Policy Department (henceforth MPD) staff also spoke in an interview to the systematic evaluations and refinements made to the models, essentially by every quarter. They referred to the many stages of correction and adjustments made before the MPD's analysis is presented to the Executive Board.³

The model-based forecasts can be categorized into three groups based on horizon. First, nowcasts, or short-term forecasts, concern the current or following quarter (Iversen et al. 2016). This varies between different forecasting occasions (Lindé and Reslow 2017). The second category is the set of forecasts covering the medium-term for a number of macro indicators such as CPIF inflation, GDP growth and the repo rate. As described by Iversen et al. (2016), "these forecasts are conditioned on the international forecast and the short-term forecast." (5) The third category is the long-term, recognized as the forecasts that surpass the medium-term horizon of two years. In relation to the models, the shortest and longest forecasts are computed differently. While the shortest forecasts are produced using empirical links in volumes of available data, the longest forecasts are based on theoretical economic links (Andersson and Reijer 2015).

The first model we will evaluate is the DSGE model, "used to produce macroeconomic forecasts, alternative scenarios, and for monetary policy analysis" (Adolfson et al. 2013, 4). The model, internally referred to as Ramses, assumes Sweden to be a small open economy and is informative in the sense of showing possible outcomes in various manipulatable simulations (Iversen et al. 2016). The final good outputs are produced "by combining the domestic homogenous good with specific imported inputs for each type of final good," (Ibid., 4). According to officials from the MPD, Ramses will soon be replaced by another DSGE model called Maja.⁴ Adolfson et al. (2007) describe the DSGE model's applications as follows:

An important part of the Riksbank's analysis of future economic developments is the analysis of scenarios concerning possible future paths for wages, interest rates or a different economic development abroad, for example. Ramses can be used to analyze the combined consequences of a proposed alternative scenario. (19)

The DSGE model proposes alternative scenarios by simulating exogenous shocks with different probability distributions and is thus quite flexible in its utility. By studying these exogenous shocks, the Riksbank is able to understand why the economy is not progressing with historical patterns, as well as understand the importance of each explanatory factor (Andersson et al. 2015).

³ Interview with member #2 of Monetary Policy Department Staff, February 27, 2020

⁴ Interview with member #1 of Monetary Policy Department Staff, February 18, 2020; Interview with member #3 of Monetary Policy Department Staff, February 27, 2020

The BVAR model is a time series model first introduced in 2003. It is useful in capturing a number of important macro variables that have a tendency to show different correlations over time.⁵ A high-ranking member at the MPD describes BVAR as a purely empirical model that builds on several variables, producing better forecasts as a model with less theoretical restrictions. The model's forecasts for CPIF inflation and the repo rate supposedly perform well "in absolute terms and relative to the DSGE model and the Riksbank's published forecasts" (Iversen et al. 2016, 17). The Riksbank's current version of BVAR includes three foreign variables – trade-weighted GDP growth, CPI inflation and the repo rate – and weighs-in domestic variables such as hours worked, the trade-weighted nominal exchange rate, GDP growth, CPIF inflation and the repo rate (Iversen et al. 2016). Further, the model uses steady-state priors to produce information about the long-run values of different variables (Iversen et al. 2016; Adolfson et al. 2007).

2.4 Judgment and Models in Forecasting

When analyzing the use of judgment in the forecasting process we refer to the “information, knowledge, and views outside the scope of a particular model” (Svensson 2005, 2). This interpretational account of judgment is central to our study as it allows us to define model-based forecasts *a contrario* as all information, knowledge and views produced *inside* the scope of a particular model. By construction, this forms a mutually exclusive definition that is also collectively exhaustive. Further, we are convinced that this distinction is consistent both with the Riksbank interviewees’ recital of judgment and model-based forecasting. Although we recognize that model forecasts are never entirely devoid of judgment, personal assessments performed by the MPD prior to and during the Main Forecasting Meeting will still count as model product against this definition. In other words, any judgment applied *a posteriori* will exclusively refer to that made by the Executive Board. Finally, this study does not delve into the grounds of this judgment.

3. Literature Review

Large intellectual credit is often given to econometrician Henri Theil who propelled the field of forecast evaluation foreword (Zarnowitz and Mincer 1969). This study will examine two forecasting components, judgment and models. Jansson and Vredin (2003) were among the first to investigate and identify two forms of discretionary monetary policy in Sweden’s inflation-targeting central bank: ‘policy shocks’ (deviations from systematic policy) and ‘judgments’ in forecasting. Previous work that has delved into the use of judgments and models in the Riksbank’s forecasting process during the period of interest mainly consists of three papers. To understand our contribution to this body of literature, we begin by evaluating Goodfriend and King’s (2016) review of the Riksbank’s Monetary Policy in the years 2010 to 2015, followed by the rebuttal by Lindé and Reslow (2017). We will then build on the study by Iversen et al. (2016) that evaluates which component yields the most accurate forecast.

⁵ Interview with member #1 of Monetary Policy Department Staff, February 18, 2020.

3.1 Goodfriend and King (2016)

Every four years, the Committee on Finance of the Riksdag commissions an external and independent review of Swedish monetary policy. The most recent review was made by Goodfriend and King (2016). The review is an independent evaluation of the Riksbank's monetary policy during the years 2010-2015. Centered around the events post the financial crisis of 2008, the locus of the evaluation is the controversial tightening of monetary policy beginning in 2010. The authors employ a strictly qualitative method in their rigorous inspection of the Executive Board's minutes of monetary policy meetings. They source complementary material from interviews with Riksbank staff and other economic agents of importance. Goodfriend and King conclude that the Riksbank over relied on model forecasts during the evaluation period and that this could explain some of the inaccuracy of the Riksbank's published forecasts. Goodfriend and King (2016) write, "there was heavy reliance, among both the majority of the Board and the dissenters alike, on forecasts produced by models developed by Riksbank staff" (9) and that "a greater degree of humility about those models would have been appropriate." (76) They proceed to advise future caution in regard to the use of model forecasts in monetary policy decision-making. The final assessment being that, "the Riksbank should re-examine its methods for producing forecasts for both the world economy and overseas interest rates to help the Executive Board focus on the big issues surrounding the outlook." (10) Goodfriend and King's review is one account of how alleged model preference in the forecasting process has worked to the detriment of the Riksbank.

3.2 Lindé and Reslow (2017)

As a rebuttal to the review, Lindé and Reslow – both employees of the Riksbank at the time – conducted their own study of the Riksbank's reliance on models versus judgment in the years 2010-2015. Using the Riksbank's published forecasts in the medium term from 2006-2016 for GDP growth, CPIF inflation and the repo rate, the authors were able to extrapolate the direct impact that the models had had on the published forecasts. The authors then studied the influence that the models had on both the level and the revision of forecasts. In both cases, the authors found that judgment had a primary role in the Riksbank's forecasts during the period of review. Lindé and Reslow thus refute the allegation, or myth, that the Riksbank's forecasting process is governed by models. They write, "our results show in all clarity that the view disseminated by Goodfriend and King – that the Riksbank blindly relies on and follows its models – is entirely misleading and is merely a myth." (28)

To build on the ideas and material entered in the discussion about forecasting and monetary policy by Goodfriend and King (2016) and Lindé and Reslow (2017), respectively, we ask how, or if at all, these findings have benefitted the Riksbank. If, indeed, the Riksbank makes predominant use of judgments relative to models, has this improved their forecasting accuracy and, particularly so, in the reviewed period of 2010-2014? This is where our study will attempt to build and add to the findings of Iversen et al. (2016).

3.3 Iversen et al. (2016)

Unlike the two previously mentioned studies, Iversen et al. (2016) delve into the question of which component yields better forecasts. The study evaluates the forecasts made by the Riksbank

in real time from 2007-2013, comparing the forecasts produced by the DSGE model and the BVAR model with the judgmental forecasts published by the Riksbank. The authors take a stance on the usefulness of founding the published forecasts on the two components, judgment and models, respectively. Like Goodfriend and King (2016) and Lindé and Reslow (2017), the authors begin by investigating the question of which component holds more power – that is, how have the model forecasts informed the published forecasts. The second dimension refers to the relative accuracy of judgments versus models – a question that we will uncover further. Their conclusions are described as the following:

We have shown that model-based forecasts have provided important input into the Riksbank's published forecasts, but the model-based forecasts have often been more accurate than the published forecasts. In particular, the BVAR model forecasts for inflation and the repo rate have performed very well, both in absolute terms and relative to the DSGE model forecasts and the Riksbank's published forecasts (Iversen et al. 2016, 17).

As the scope of their study partly overlaps with the period of wishful thinking, the findings may indicate support for the argument that the Executive Board's own judgment was systematically overestimating the economic situation. The findings show that the model-based forecasts are more consistently accurate than the published forecasts. We will attempt to shed light on whether this is the case in the following contents of this paper. As recommended by Iversen et al. (2016), we will also use additional formal methods, as mentioned previously, to further evaluate the accuracy of the forecasting components.

4. Purpose of Study

Criticism and questioning of the central bank's decision-making arise given the nature of its role as a public establishment. The fulfillment of its motto *Hinc robur et securitas* (henceforth strength and security) is as relevant to examine now as it was a century ago. As Swedish economist Gustav Cassel writes in *Riksbanken under krisen 1907–1908*, “a central bank gives guidelines for the country's entire economic life through the choice of its interest rates” (Cassel 1908, 90). Evaluating the foundational reasoning that lies behind its monetary policy is therefore a timeless essentiality and will remain true for the target-oriented and independent Riksbank for many reasons. First, it is important for the Riksbank to maintain legitimacy and public understanding of the relevance of the monetary policy they conduct (Bernanke et al. 1999). Jansson and Vredin (2003) also emphasize the importance of external evaluation in regard to analyzing whether the Riksbank succeeds with achieving its goals. As Goodfriend and King (2016) point out, lively debate on monetary policy provides a valuable contribution to the future of Swedish monetary policy.

From a policy analysis perspective, Jonung (1999) describes Sweden as an exceptionally well-suited object of study. The primary reason being the “traditional openness of Swedish society” (277) with the central bank's transparency allowing for a great volume of material to be evaluated. As transparency is at the heart of the institution's maintenance of legitimacy, the Sveriges Riksbank is a particularly interesting case to evaluate.

With the luxury of hindsight, we evaluate the two components of the Riksbank's forecasts during the period of 2010-2014. Our aim is three-fold. First, we aim to build on and compare the results with previous studies (as detailed in the section 3). Second, we investigate the grounds for the heavy criticism – specifically regarding the judgmental adjustments made by the Executive Board – surrounding the period of review. Third, we aim to reboot the discussion of the appropriate use of judgment versus models, especially in the setting of such a central policy-making institution. The question thus holds global relevance. In light of this, we aim to answer our research question: how did the use of judgments and models affect the accuracy of the Riksbank's forecasts during the years 2010-2014? We define accuracy as 'capable of or successful in reaching the intended target'. The question is, thus, not what a reasonable margin of difference between forecast and outcome is, but whether models or judgment brought the Riksbank's forecasts closer or farther from the actual outcome.

5. Theoretical Framework

Given the fundamental differences in how models and humans register information, our theoretical framework will surround ideas of how we as humans process information. Differences about the marginal utility of information has led to a plenitude of studies. While some scholars believe that more information necessarily leads to better informed decisions, others have pointed out mankind's limited capacity to process large quantities of information as reason for the decreasing marginal utility of information. Hertwig and Todd (2003) inspect both accounts. By noting the general perception that more information is always better than less, and that ignorance impedes good decision-making, it would appear that marginal utility of information is increasing. Hertwig and Todd also investigate the limits to our processing of information, going so far as to question whether the concern is even necessary. They point out that skepticism about our information-processing capacity is usually coupled with the assumption that these limitations are impairing (Hertwig and Todd 2003). Based on this latter remark, we ask if informed decisions are made not despite the limits to the processing mind but because of them? We explore the extents of the two differing accounts – whether more information is preferred to less – and their academic foundation, through the lens of theories of bounded and ecological rationality. Finally, we relate it to our quest to determine the potential value of the Executive Board's judgment in addition to the model forecast.

5.1 Bounded Rationality

The first mention of bounded rationality appears in *Models of Man* (Simon 1957, 198; Barros 2010). Herbert Simon, who has been recognized as the founding father of bounded rationality, believed the economic agent to be cognitively limited and, by extension, rationally bounded. Simon argued that cognitive limitations in the form of time constraints, imperfect and incomplete information impose boundaries on the agent's rational mind. As Kleinmuntz (1990) notes, "Simon meant that people do not think rationally because doing so, he posited, requires excessive cognitive effort." (298) Simon, thus, contradicted neoclassical economic theory in arguing against global rationality and believing that temporal and informational concerns hinder rational thinking.

It is of particular interest to this study if bounded rationality can be evoked in scenarios tainted by uncertainty, like that in the period of wishful thinking. In the setting of the Riksbank,

uncertainty can be argued to naturally overshadow the forecasting process. According to a high-ranking member of MPD staff, some amount of uncertainty is ever-present as monetary policy is conducted with temporal lags and models are always mere simplifications of reality.⁶ In the five years that the period of wishful thinking lasted, specific concern regarding household debt and difficulties estimating foreign exchange rates evoked additional uncertainty. As stated by Head of General Secretariat Anders Vredin, the Executive Board could not afford to neglect the impending risks of financial instability (Vredin 2020). When the foreseeable future becomes increasingly ambiguous, the assumption of rational choice is too simplistic. If underlying uncertainty cannot be reduced this would imply that there are too many interpretational barriers to act rationally – that is, economic agents, such as the Riksbank, fail to rationalize all information available. In light of this, the unprecedented events and uncertainty following the financial crisis could have posed extraordinary challenges for the Riksbank and other economic agents to understand and accurately predict economic developments.

If mankind is not equipped to process large quantities of information, what are then the preconditions for smart decision-making in uncertain times? In the following section, we seek answers by exploring the notion of ecological rationality.

5.2 Ecological Rationality

Simon also inspired others to ponder the effects of uncertainty on decision-making and rationality. Gigerenzer et al. (2000) introduce a related component of bounded rationality that stresses the importance of the environment in which decisions are made. The new component, ecological rationality, is rationality defined by its fit with reality. At the core of ecological rationality lies the presupposition that decision-makers often must consider direct threats to their future survival and reproduction. Faced with these environmental challenges, decision-makers need fast, frugal and accurate methods of making inference in order to behave adaptively.

I use the term uncertainty for situations where the exhaustive and mutually exclusive set of future states and their consequences is not known or knowable. In these frequent real-world situations where the assumptions of rational choice theory do not hold, individuals and institutions can nevertheless make smart decisions using psychological tools such as fast-and-frugal heuristics (Gigerenzer 2018, 304).

Gigerenzer et al. (2000) also hypothesize whether fast-and-frugal heuristics can be an alternative to optimization strategies in human decision-making. What are then the strengths of heuristics? Gigerenzer and his colleagues argue two things: simplicity and domain-specificity. Adaptive decision-making comes about when both of these two elements are matched with the information of the specific environment. With fast-and-frugal heuristics at hand, human biases such as the cognitive limitations introduced by Simon need not adversely affect smart decision-making. (Gigerenzer 1991) believes that rational thinking is not so much conditional on cognitive limitations as it is on the environment in which the agent navigates. Critics of Gigerenzer, most notably Kahneman and Tversky, object to this concept, arguing that some human biases should be characterized as errors or fallacies (Kahneman and Tversky 1974; Vranas 1999).

⁶ Interview with member #2 of Monetary Policy Department Staff, February 27, 2020.

In practice, the Riksbank finds a fast-and-frugal heuristic in the Taylor rule. A simple instrument rule, the policy rule lets the instrument (here, the repo rate) float with the inflation and the output gap⁷ (Svensson 2003). Although the Taylor rule and its performance is not the object of our study, it is one example of the Riksbank's use of a simple and domain-specific rule to circumvent uncertainty and other cognitive barriers. One could also argue that the inflation target is yet another example of a fast-and-frugal-heuristic. However, given the development of CPIF inflation during 2010-2014, "a risk began to build up that the role of the inflation target as a benchmark for price-setting and wage information would begin to weaken – that the nominal anchor that has been an important part of the favorable developments in Sweden since the early-1990s crisis, would begin to slip up" (Ingves 2019, 5). It was therefore of outmost importance for the Riksbank to maintain public confidence in the inflation target, even when inflation expectations were dwindling (Ingves 2019). This highlights a possible weakness of fast-and-frugal-heuristics, that is, they lose substance when confidence in them is lost. This further confirms the importance of evaluating forecasting accuracy.

5.3 When is more preferred to less?

The notion of ecological rationality is an alternative view to the prevailing perception that more information and more computation leads to better informed decisions. In short, that more is more. Different to bounded rationality, it not only presupposes cognitive capacity issues, but inherent environmental structures that limit rational choice. It theorizes that there exists an equivocal point where too much information impairs inference. In short, when less is preferable to more. Interestingly, and perhaps most relevant for this study, is that it provides an explanation as to how decision-makers can make better decisions when both time and knowledge is scarce. Taken at face value, it could explain why the judgment of the Executive Board – whatever its foundational basis – have the potential to outperform the most intricate algorithm of the Riksbank's macro models. Whether this is the case will be elaborated in our findings.

6. Data and Method

6.1 Data

The data used in this study was granted to us by economist Jesper Lindé – staff member of the Research Division at the Riksbank's Monetary Policy Department. This is the data used in the paper co-written by André Reslow (2017). Model forecasts from 2013 and onwards can be accessed through the Riksbank's current data management system, while forecasts prior to 2013 are made available on Databiblioteket. The published forecasts are publicly available in the Monetary Policy Reports and Updates on the Riksbank's website (Lindé and Reslow 2017). The original time series data stretches from 2006-2016 and contains model forecasts and published forecasts for the three macro indicators: GDP growth, CPIF inflation and the repo rate. As previously mentioned, prior to 2020, the Executive Board met six times per annum (Riksbank 2020) which corresponds to three Monetary Policy Reports (MPR) and three Monetary Policy Updates (MPU) in our data.

⁷ Output gap is defined as the gap between actual output and potential output (Svensson 2003).

6.1.1 Model Forecasts

Model forecasts used in this study are produced by the Riksbank's two primary macro models, DSGE and BVAR. For the purpose of this study, we make use of model forecasts conducted and saved at the point of the main forecast meeting. The model-based forecasts used in this study assume current and international events to be exogenous. Therefore, model forecasts have been conditioned *a priori* on a nowcast and a forecast for international developments.

6.1.2 Published Forecasts

Published forecasts comprise of some weighted combination of the BVAR and the DSGE model forecasts, the published forecast from the previous forecasting round (a practice known as forecast smoothing) and some element of judgment (Lindé and Reslow 2017). Both Iversen et al. (2016) and Lindé and Reslow (2017) estimate the relationship as follows:

$$F_{j,t+h}^{New} = \omega_R F_{j,t+h}^R + \omega_B F_{j,t+h}^B + (1 - \omega_R - \omega_B) F_{j,t+h}^{Old} + e_{j,t+h}, \quad (1)$$

Where $F_{j,t+h}^{New}$ is the published forecast, established in time t , for some variable j in time $t + h$. Further, the published forecast is the sum of some weight, ω_R (ω_B), attributed to the Ramses (BVAR) forecast, $F_{j,t+h}^R$ ($F_{j,t+h}^B$), and some weight, $(1 - \omega_R - \omega_B)$ attributed to the previously published forecast, $F_{j,t+h}^{Old}$. Ramses is the name of the Riksbank's DSGE model. We use DSGE for convenience. In their regression, the element of judgment is contained in the error term, $e_{j,t+h}$. However, we recognize that this could just as easily capture "noise" in the model.

6.1.3 Data Transformation

In our efforts to reconstruct the data, we begin by excluding all observations outside of the period under review. We then transform the data vertically to categorize observations by meetings, by which Monetary Policy Reports (MPRs) and Monetary Policy Updates (MPUs) are given the same status, instead of by macro indicator. This allows us to sort dates by the most recent forecast, assuming the most accurate forecast was produced at the most recent meeting. After sorting the data and omitting missing values, the dataset ultimately consists of the variables GDP growth, CPIF inflation and the repo rate, with 19, 18 and 17 observations respectively. The number of observations differ because of missing values for certain dates.

In the analysis of our data, we begin by visually inspecting the forecasts for one to two (nowcast), four and eight quarters from the meeting occasion. For these three horizons, we also examine their respective forecast bias performances across the period of wishful thinking. We proceed by exclusively studying the nowcast for our t-tests and optimal model weighting ratios. Despite the fact that monetary policy action typically cannot affect much of the real economy in the short-term because of time lags, we recognize that accurate medium-term forecasts are conditional on accurate short-term forecasts, as described by Iversen et al. (2016). In addition, nowcasts reflect the Executive Board's understanding of the economy in real time and their subsequent repo rate decisions in the moment, based their well-balanced monetary policy considerations.

6.2 Complementary Qualitative Research

In addition to our data collection, interviews were conducted with seven individuals including current and former staff of the Riksbank. Among them were two former Deputy Governors: Lars E.O. Svensson (2007-2013) and Karolina Ekholm (2009-2014). Interviews were held in the months of January and February of this year and lasted approximately 30-90 minutes. The purpose of the interviews was two-fold: to add to our understanding of the Riksbank's forecasting process and to obtain a nuanced perception of the period of wishful thinking.

6.3 Defining the Period of Wishful Thinking

Several individuals were asked to define the period of wishful thinking. Most notably, Svensson estimated the period to have stretched between June 2010 and December 2013 (Svensson 2020). Ekholm, on the other hand, approximated the period to have lasted somewhat longer; from mid 2010 until sometime during spring or summer 2014 (Ekholm 2020). In deciding what time period to study, we also reviewed Lindé and Reslow (2017) who motivated cut-off points by stating the following: “The third period, 2011/Q1-2014/Q2, refers to a time when the Riksbank hiked the interest rate, before subsequently starting to cut it again.” (41) By visual inspection of the minutes of the meeting, and in particular by votes and reservations, the first account of a majority vote to hike the interest rate occurred during the Executive Board's meeting on June 30, 2010. The session corresponds to the second quarter of that year in our dataset, which is why we have chosen 2010/Q2 as our early cut-off point. As for the closing cut-off point, we stand in full agreement with Lindé and Reslow (2017) and have thus used the monetary policy meeting on July 2, 2014. Hence, the period of wishful thinking stretches from 2010/Q2 to 2014/Q3.

6.4 Multicollinearity, Serial Correlation and Sample Size Concerns

Model forecasts are rarely uncorrelated. In fact, model forecasts are typically highly correlated which raises the concern of potential multicollinearity. Moreover, because the Riksbank practices linear forecast smoothing, that is, uses the forecast from the previous forecasting round when producing the new forecast, errors risk being serially correlated. In addition, because of the Riksbank's internal debate concerning a normalized repo rate path, previous repo rate forecasts are actively considered in each subsequent forecast occasion. The prevalence of multicollinearity and serial correlation have great implications for our t-tests, which is why it is necessary to consider all findings against this fact. Our optimal model weighting method also suffers from the same kinds of limitations as Equation 1, presented in section 6, used by Iversen et al. (2016) and Lindé and Reslow (2017). This includes potential risks for multicollinearity. Given our focus on this specific period, this data, although limited, is exhaustive. Finally, a small sample introduces statistical power weaknesses which is why it is important to consider effect sizes, as we will discuss in succeeding sections.

7. Empirical Strategy

7.1 Visual Inspection of Time Period and Real-Time Forecasts

Before reviewing the performance of the so-called nowcasts in real time, let us begin by looking at the data from a wider lens. We recognize the importance of establishing a general understanding of forecast accuracy across all horizons. What can we discern from a graphical display of *all* MPR forecasts – short-term to medium-term – compared to the actual observed values for each indicator? Figure 1-3 provide some intriguing insights.⁸ The charts show how the forecasts perform compared to the last known outcome for each indicator. They also illuminate oscillations in the economic environment, providing greater contextual understanding of this period we review. With the last known outcomes distinguished by the black line, we can see the relative performance of each MPR (published forecasts) and MPU (revised forecasts) in the collection of shorter lines.

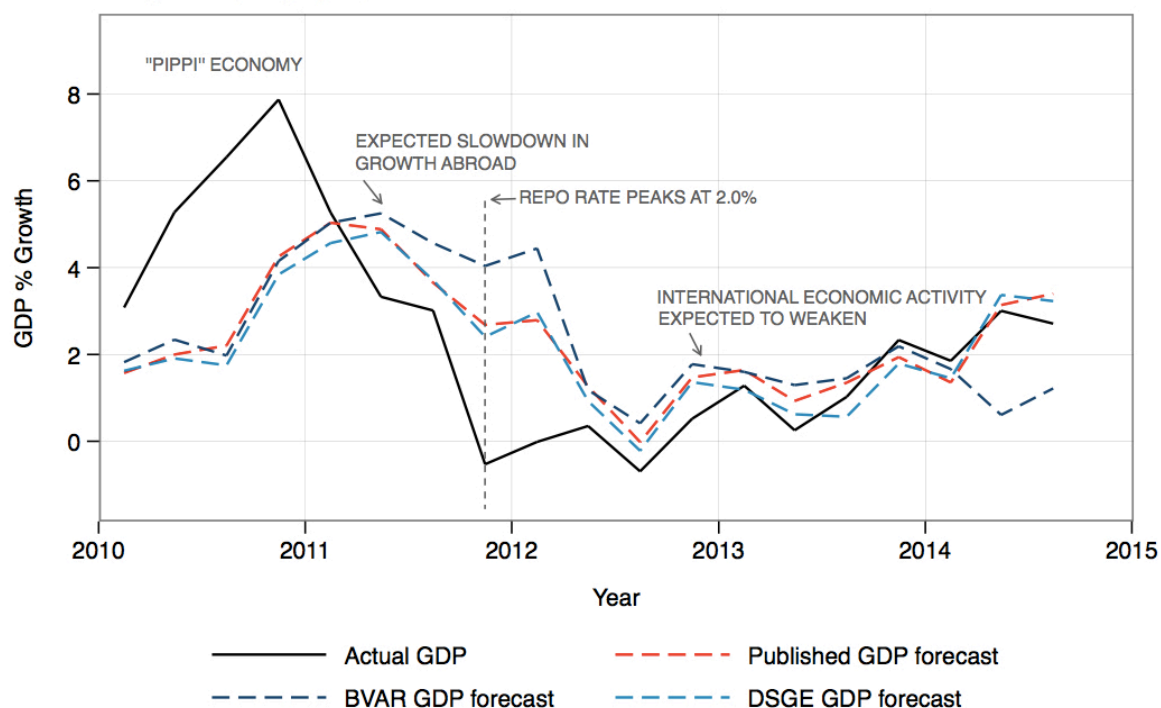
Firstly, it is difficult to discern any considerable differences in the GDP growth forecasts as seen in Figure 1. On the other hand, the DSGE model forecast from 2011 to early 2012, showed a relatively more accurate prediction of the steep fall in GDP growth than the other two forecasting components. That being said, the same can be seen for the BVAR model forecasts around the peak of GDP growth between 2016 and 2017. Secondly, Figure 2 provides foreground for some more visible comparisons. Chart A in this figure, indicates a persistent tendency for the Executive Board to publish overestimations of the basal inflationary pressure throughout this entire period. This partly explains the excessively high expectations of how quick the repo rate can be normalized. It may also indicate a tendency to expect a quicker return to the desired inflation target of two percent. Nonetheless, the same can be said for the DSGE model forecasts. It is important to note, however, that the nowcasts for CPIF inflation – shown in the beginning of each line – appear to be the most accurate to the outcome, and thus suggests that forecasting accuracy is greatest in the short-term. This will be constructive to consider in following analyses of the real-time forecasts. In other words, accuracy varies across horizons and thus the findings for nowcasts will not be representative for all forecasting horizons. Cross-sectionally, the DSGE model and published forecasts show quite similar forecasting paths for CPIF inflation and the repo rate.⁹ The BVAR model displays more accurate, systematically lower forecasts for both CPIF inflation and the repo rate.

How do the nowcasts compare to the actual observed values for each indicator? Figure 4 displays the DSGE, BVAR and published forecasts in dotted lines together with the outcome (the black line) for our macro indicators. The first row in the figure shows the charts for GDP growth and CPIF inflation, while the second row presents the chart for the repo rate.

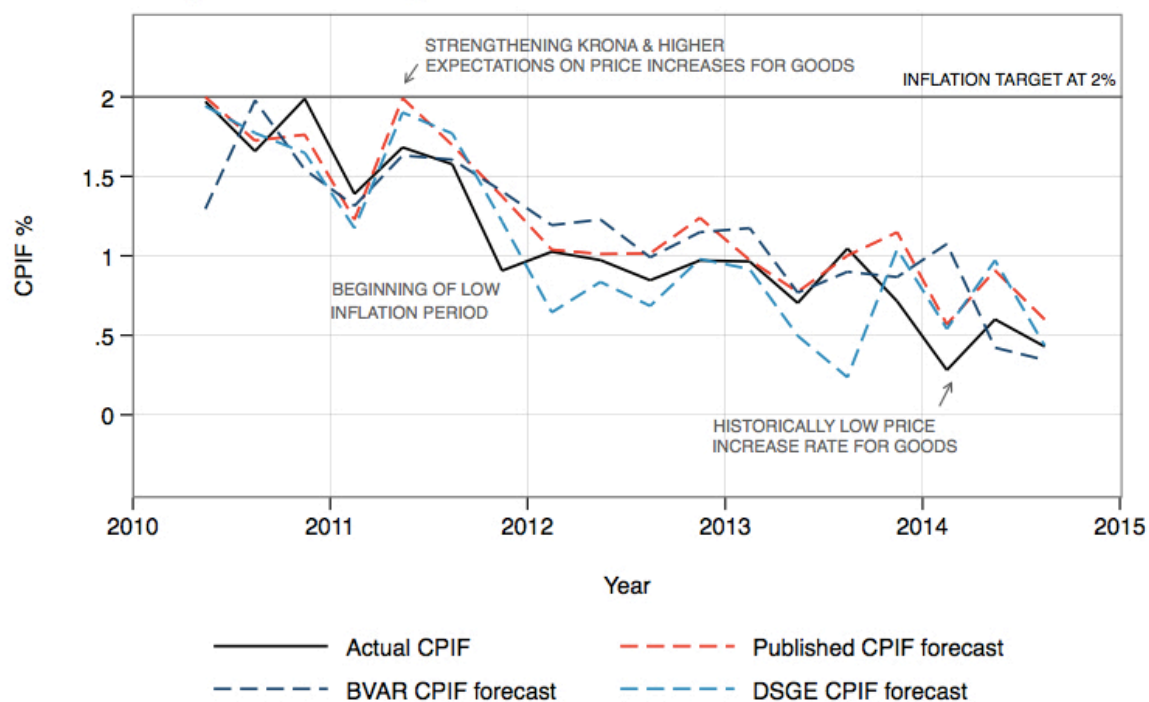
⁸ Consult Appendix A for Figure 1-3 illustrating forecast performance across the macroeconomic indicators.

⁹ See Appendix A, Figure 2-3.

A. GDP % Growth



B. CPIF INFLATION



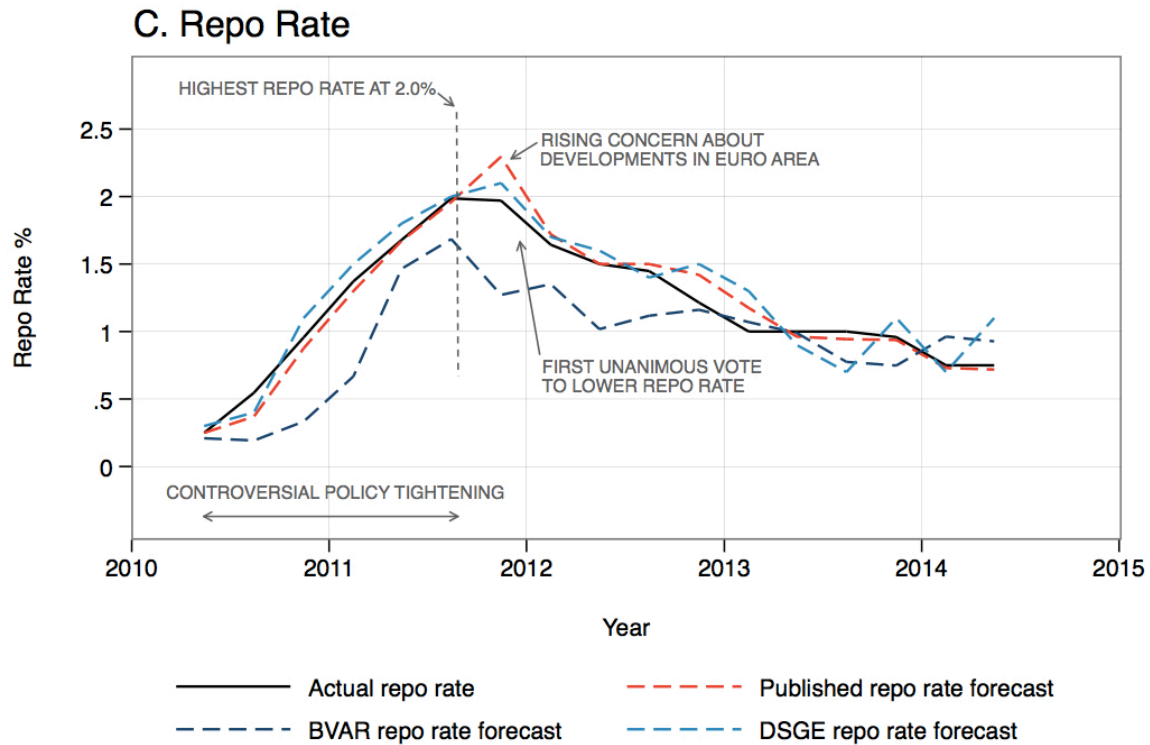


FIGURE 4. ACTUALS AND REAL-TIME FORECASTS. Note the figure charts the development of the nowcasts in real time – the most recent short-term forecast for each date – along with the line for actual outcome from 2010-2014. Chart A shows the actuals and real-time forecasts for GDP growth, while Chart B and C show the actuals and real-time forecasts for CPIF inflation and the repo rate respectively. Each forecasting component is displayed in the dashed lines. See legends for specific reference. Data source: Lindé and Reslow (2020).

From Figure 4, we discern a significant lag and underestimation in both the published and the models' forecast of the peak and succeeding drop in GDP growth by the turn of year 2010. In January 2011, Secretary-General of the Organization for Economic Co-operation and Development (OECD), José Angel Gurria, compared the Swedish economy's strength to that of the fictional character Pippi Longstocking, known for her abnormal strength (Svensson 2011). Head of MPD, Jesper Hansson, also referred to the 2010 Swedish Economy as "the Pippi Longstocking economy", emphasizing the International Monetary Fund's optimism about Swedish growth post financial crisis (Hansson 2020). Both the DSGE model and the Riksbank's published forecast appear to have outperformed the BVAR model in forecasting GDP growth under the period of review. As for CPIF inflation, substantial model forecast differences are hard to discern with both models tailing the outcome relatively closely. We can also note that the published and DSGE model forecasts follow similar paths throughout the entire period. We recognize, however, that as of 2014, all forecasts have predicted an imminent return to the two percent inflation target, faster than the actual development. Implications of this can be observed in the repo rate chart, with both model and published forecasts motivating a higher than justified repo rate by the end of the period.

7.2 Forecast Deviations and Bias

In order to compare the predictive accuracy of the BVAR, DSGE and published forecasts, respectively, we define three new variables.

Actual deviation: the absolute difference between the published forecast and the outcome. (2)

DSGE deviation: the absolute difference between the DSGE forecast and the outcome. (3)

BVAR deviation: the absolute difference between the BVAR forecast and the outcome. (4)

The deviations measure what is commonly known as the forecast error – the residual of an observation that is unpredicted by the forecast. We denote this as:

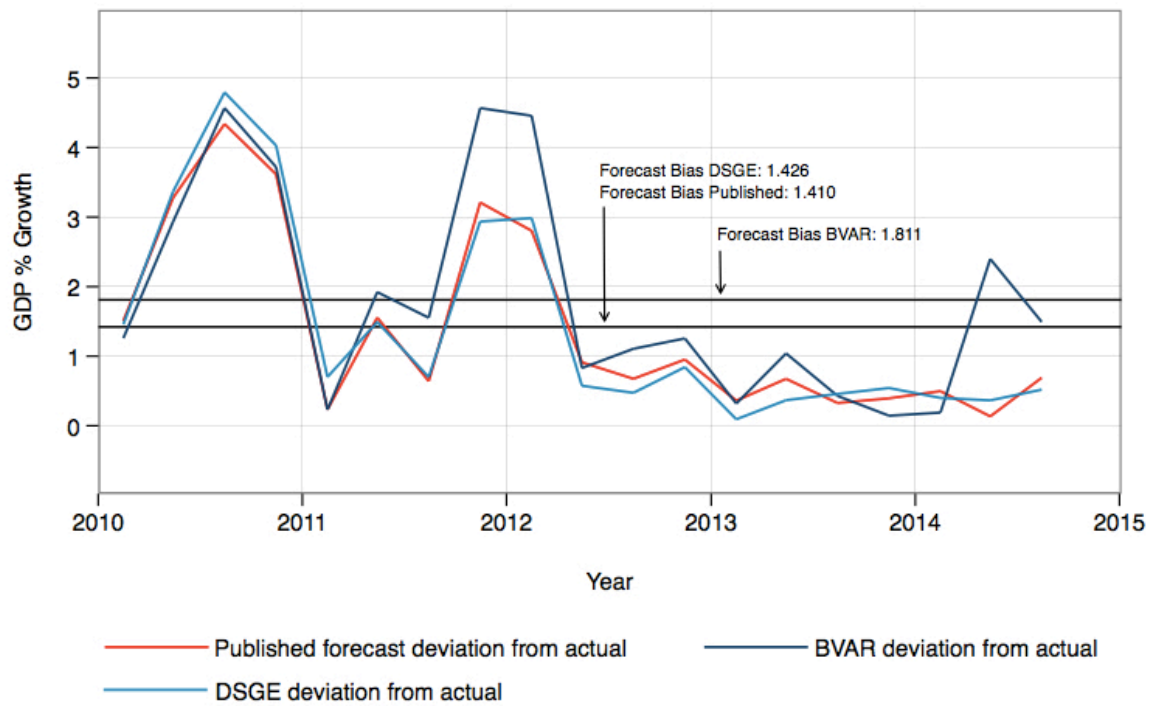
$$\hat{e}_{t+h} = \hat{y}_{t+h} - y_{t+h} \quad (5)$$

where the size of the forecast error, $\hat{e}_{t+h}$, is measured as the difference between the value predicted by the forecast, $\hat{y}$, and the observed value, y . The current time period is noted by t and the forecast horizon by h . A negative $\hat{e}_{t+h}$ can be interpreted as an underestimation of the true value and vice versa, but in this study, we measure all deviations in absolute terms. In studying the nowcast, we set h equal to a maximum of 2. By plotting forecast errors for all three of our macro indicators, we are able to graphically observe any important differences in patterns.

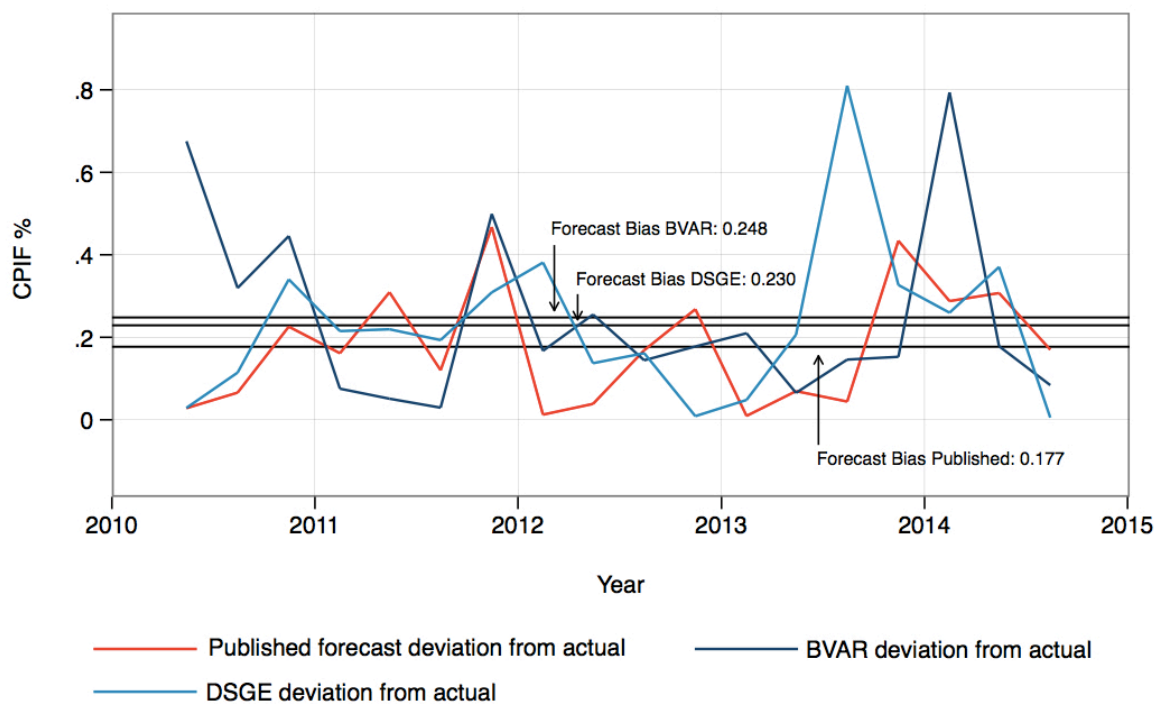
Figure 5 shows the real-time forecasts' deviation from the actual outcome during the period of interest, where the macro indicator unit is displayed on the y-axis for each respective chart and the deviation is plotted across time by the x-axis. The charts isolate differences in forecasting accuracy from 2010-2014. It is important to note that each forecast component performs relatively better than its competitors at some point in time, which indicates that the potential for a component to outperform another in accuracy is not improbable. Another important disclaimer relates to the inability to ascribe deviations any rankings in terms of what is considered standard. The systematic differences, however, can still be identified. The key insight from Figure 5 is the noticeable greater deviations of the BVAR model forecasts for all macro indicators. This is further supported by the reference lines for average forecast bias, showing the BVAR model performing worst across the board in this measure.¹⁰ The opposite is shown to be the case for published forecasts, for which the average forecast bias is lowest in all three charts. Looking at the overall forecasting performance, the GDP growth forecasts show notable similar predictions compared with the other two variables, while the largest deviations are visible in the repo rate. The repo rate chart also shows a relatively flatter line for published forecasts, which does bring up the question of whether the Riksbank demonstrates a tendency to prefer a normalized repo rate path.

¹⁰ All forecast bias results are available in Appendix C.

A. GDP Growth



B. CPIF Inflation



C. Repo Rate

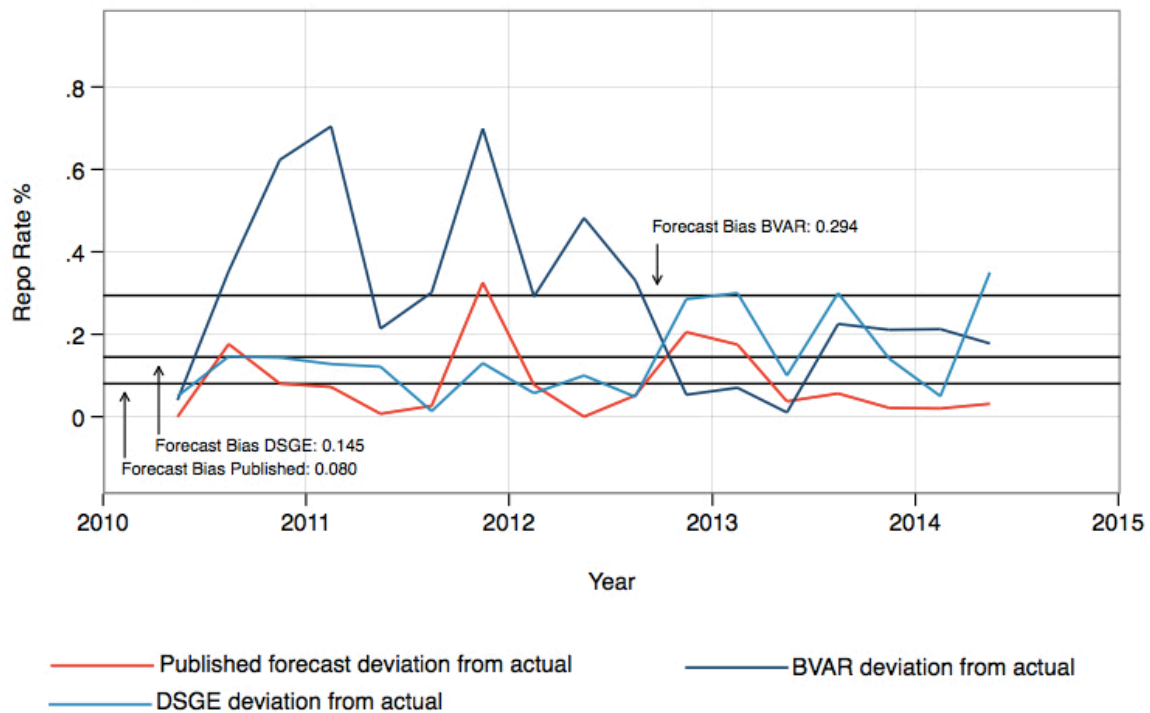


FIGURE 5. NOWCAST DEVIATIONS FROM ACTUAL. Note the figure charts the deviation of the nowcasts in real time – the most recent short-term forecast for each date – from the actual outcome for 2010-2014. Chart A shows the deviations of the nowcasts for GDP growth from the actual outcome, while Chart B and C show the deviations of the nowcasts for CPIF inflation and the repo rate from their respective actual outcome. Each forecasting component is displayed in the solid lines. See legends for specific reference. Data source: Lindé and Reslow (2020).

In Figure 6, we compare the deviations of the forecasts four quarters from the meeting occasion. We note that the average forecast bias for GDP growth has decreased across all forecasting components compared to the bias reported in the nowcast.¹¹ This result is somewhat unexpected as forecast errors typically increase the more uncertain the future is (Lindé and Reslow 2017). For CPIF inflation, the results appear more intuitive, with higher average forecast biases reported for all three forecasting components. Unsurprisingly, greater diffidence in future CPIF inflation taints the ability for the Executive Board to determine the repo rate – with both models and the published forecast reporting a higher forecast bias compared to the nowcast. As previously established, the inflationary pressure was consistently overestimated during this period (Ingves 2019), which is more evident in this forecast horizon.

Lastly, given the time lag in monetary policy, we extend the horizon even further to consider eight quarters ahead. This is shown in Figure 7. We expect forecast biases for all macro indicators to have risen since our analysis of the real-time forecasts in the fourth quarter and even more so since our nowcast analysis. For GDP growth, the pattern is somewhat hard to distinguish. As for the published forecast, average forecast bias has risen relative to the four-quarter horizon, but not

¹¹ Consult Appendix B for Figure 6 and 7.

enough to surpass the bias reported in the nowcast. For the BVAR and the DSGE model forecast, average GDP forecast biases have decreased even beyond the estimates reported in the fourth quarter. This implies that model forecasts become more accurate over a longer time horizon. As for CPIF inflation, the bias has increased in the published and the BVAR model forecast, but decreased ever so slightly in the DSGE model forecast compared to the fourth-quarter estimates. Finally, the repo rate forecast biases are the highest in the eighth-quarter analysis. Henceforth, as we have previously motivated, we shall center our analysis around the nowcasts.

Comparisons are continued through the visual distribution of deviations in histograms.¹² The histograms in Figure 8-10 bring our focus back to forecasting accuracy across all horizons with the purpose of evaluating the aggregated (i.e. all-horizon encompassing) deviation performances. Shown as the percentage of all observations, deviations from the actual outcome look relatively similar for GDP growth, just like we established while looking at Figure 1. Again, a key takeaway from these charts is that forecasting accuracy varies for all forecasting components and the variables of interest. The deviation distributions for all forecasting components in GDP growth are quite similar. For the repo rate however, it appears that the DSGE and BVAR models show a similar flatter distribution with a single peak in the left tail.

7.3 Testing Forecast Errors for Statistical Significance

In the second part of the data analysis, we aim to compare levels of "noise" produced by the real-time nowcasts. In other words, we examine that which is left unexplained by the forecasts. We do this by testing the difference in the mean squared errors (MSE) for statistical significance, obtained by regressing the outcome on each forecast. We hypothesize that the difference in means is equal to zero against the alternative that it is different from zero using appropriate two-sided Student's t-tests. Failure to reject the null hypothesis implies we cannot rule out the possibility that forecasts produce equal amounts of noise, or conversely, that they have equal forecasting abilities. We thus hypothesize the following:

$$H_0: \bar{e}_{PUB}^2 - \bar{e}_{DSGE}^2 = 0. \quad (6.1) \quad H_1: \bar{e}_{PUB}^2 - \bar{e}_{DSGE}^2 \neq 0. \quad (6.2)$$

$$H_0: \bar{e}_{PUB}^2 - \bar{e}_{BVAR}^2 = 0. \quad (7.1) \quad H_1: \bar{e}_{PUB}^2 - \bar{e}_{BVAR}^2 \neq 0. \quad (7.2)$$

$$H_0: \bar{e}_{BVAR}^2 - \bar{e}_{DSGE}^2 = 0. \quad (8.1) \quad H_1: \bar{e}_{BVAR}^2 - \bar{e}_{DSGE}^2 \neq 0. \quad (8.2)$$

where *PUB* is published forecast abbreviated. We calculate MSE using the following formula, where *f* is an abbreviation for forecast – as in the chosen forecast component for analysis:

$$\bar{e}_f^2 = MSE = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}, \quad i = 1, 2 \dots n, \quad (9)$$

where MSE is denoted by $\bar{e}_f^2$, the predicted value by $\hat{y}$, and the observed value by y . Individual forecasts are denoted by i whereas n marks the sample size.

¹² Consult Appendix D for histograms.

In the following three sub-sections, we present the results from the t-tests with concluding comments on what can be discerned. This method of analysis allows us to draw connections to the observations noted in our visual inspection, and thus strengthening our findings by considering different angles of examination.

7.3.1 GDP Growth Mean Squared Errors Comparisons

TABLE 4. GDP GROWTH NOWCASTS T-TEST OF MEAN SQUARED ERRORS

Forecast	Published	BVAR	DSGE
Published	.	-2.212* (.581)	-1.145 (.225)
BVAR	-2.212* (.581)	.	2.285* (.450)
DSGE	-1.145 (.225)	2.285* (.450)	.

Notes: Table 4 reports t-statistics from a test of the null hypothesis that the difference in mean squared errors (MSE) is null across nowcasts with 18 degrees of freedom. Standard mean errors are reported in parentheses. Stars denote statistical significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: Lindé and Reslow 2020.

From Table 4, we note that the difference in the MSEs of the BVAR model and the published forecast is statistically different from zero at the 5% significance level (p-value of .0401). We, however, fail to reject the null at any conventional significance level when comparing the DSGE model with the published forecast (noting a p-value of .2671). Finally, the difference between the BVAR model and the DSGE model is statistically different from zero at the 5% significance level (p-value of .0347). We advise caution in the interpretation of statistically significant results considering the size of our sample. Upon calculations of effect size, we note small effects (i.e. less than 0.2) in all mean comparisons of the forecasts. This bears implications for the statistical power of our t-tests and also increases the risk of type II errors, i.e. when we erroneously fail to reject the null. For the effect sizes, consult Table 7 in Appendix E.

Failure to reject the null in the single instance of the comparison between the DSGE model and the published forecast suggests we cannot exclude the possibility that both forecast produce equal amounts of noise – that is, perform equally well in predicting the outcome. Is this consistent with previous graphical observations? Given our findings in Figure 1-3, the suggestion that the DSGE model forecasts and published forecasts present similar findings is further supported. This is also the case in the forecast bias findings, for which the two components produce remarkably similar average biases. In other words, the Executive Board produces similar forecasts to those of the DSGE model and thus arguably relies more on the DSGE model than the BVAR model – or at least, this seems to be the case for GDP growth. Whether this trend is present in the other macro indicators will be further developed in section 7.4.

7.3.2 CPIF Inflation Mean Squared Errors Comparisons

In Table 5, we test the difference in MSEs for our CPIF inflation nowcasts. We note the absence of any statistically significant result. Failure to reject the null in all three instances means we cannot exclude that the mean squared errors from our forecasts are equal. That is, model and published forecasts produce equal amount of noise. For the effect sizes, consult Table 8 in Appendix E.

TABLE 5. CPIF INFLATION NOWCAST T-TEST OF MEAN SQUARED ERRORS

Forecast	Published	BVAR	DSGE
Published	.	-1.609 (.043)	-1.835 (.022)
BVAR	-1.609 (.043)	.	.671 (.043)
DSGE	-1.835 (.022)	.671 (.043)	.

Notes: Table 5 reports t-statistics from a test of the null hypothesis that the difference in mean squared errors (MSE) is null across nowcasts with 17 degrees of freedom. Standard mean errors are reported in parentheses. Stars denote statistical significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: Lindé and Reslow 2020.

Is this consistent with our previous findings? As noted from the CPIF inflation chart in Figure 4, substantial differences are difficult to discern, which supports the findings from the t-test in Table 5. The forecasting components during this period show a relatively similar path along the actual CPIF inflation line, either above or below, which may explain the similar absolute indications of noise.

7.3.3 Repo Rate Mean Squared Errors Comparisons

In Table 6, we test the difference in MSEs for the repo rate nowcasts. In the case of the published forecast and the BVAR model, we note a statistically significant result at the 1% significance level (p-value of .0029). We also reject the null when comparing the DSGE model and the BVAR model at the 5% significance level (p-value of .0476). We fail, however, to reject the null when testing the DSGE model forecast and the published forecast (p-value of .0941). For the effect sizes, consult Table 9 in Appendix E.

As in the case of GDP growth, failure to reject the null implies we cannot exclude the equality of published and DSGE model MSEs. Once again, let us consult our previous findings. The repo rate chart in Figure 4 validates this finding. The relatively close forecast biases for these two components can also provide additional support. As established in section 7.3.1, it seems the published and the DSGE model forecasts show similar forecasting accuracy, which may be an indication of how the model serves the Executive Board.

TABLE 6. REPO RATE NOWCAST T-TEST OF MEAN SQUARED ERRORS

Forecast	Published	BVAR	DSGE
Published	.	-3.478** (.022)	-1.773 (.015)
BVAR	-3.478** (.022)	.	2.135* (.023)
DSGE	-1.773 (.015)	2.135* (.023)	.

Notes: Table 6 reports t-statistics from a test of the null hypothesis that the difference in mean squared errors (MSE) is null across nowcasts with 16 degrees of freedom. Standard mean errors are reported in parentheses. Stars denote statistical significance levels *p<0.05, **p<0.01, ***p<0.001. Data source: Lindé and Reslow 2020.

Acting on the proposed research extensions of Iversen et al. (2016), we continue the analysis by exploring an additional formal method of analyzing the predictive accuracy of the different forecasting components.

7.4 Optimal Weighting Ratios

In this final section of the data analysis, we examine if a weighted average of the two model forecasts achieves a better fit to the actual than the published forecast. In a hypothetical world where the Riksbank could perfectly optimize model weights, we need a measure to evaluate this optimal model's performance against the published forecast. This is where RMSE comes into use. By comparing the two forecasts, we can evaluate the value of judgmental adjustments. We find the optimal weighting ratio by minimizing the root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}, i = 1, 2 \dots n, \quad (10)$$

where the predicted value is denoted by $\hat{y}$, and the observed value by y . Individual forecasts are denoted by i whereas n marks the sample size.

We use RMSE, the sample standard deviation of the forecast error, to measure how well forecasts predict the outcome, y . Although there are many ways to measure predictive accuracy, we prefer RMSE because of its simple interpretation and its comparison relevance with Iversen et al. (2016). We also compare the RMSE of the optimally weighted model forecast with that of the published forecast. The smaller the RMSE, the better the forecast. It is important to note that our optimization excludes the previously published forecast, neglecting any explanatory power that it might hold. It is conceivable that by practicing forecast smoothing, a smaller RMSE could be obtained if the previously published forecast bettered the predictive accuracy. It is weakly supported by the fact that the Riksbank often does not make considerable forecast revisions to its own forecasts but relies on the previous forecast as a predictor of the subsequent forecast. However, Lindé and Reslow (2017) argue that this is much less common of a practice for model forecasts. Finally, we disregard any component that is the combination of both model and

judgment as would be the case with forecast smoothing. In this instance, we are solely interested in the predictive accuracy of model forecasting. As per Equation 1, the relationship between the DSGE and the BVAR model is linear. We also impose three additional constraints on our regression:

$$\text{Model weights must sum to 1: } \beta_{BVAR} + \beta_{DSGE} = 1 \quad (11)$$

$$\text{Model weights must be greater or equal to 0: } \beta_{BVAR} \geq 0 \text{ and } \beta_{DSGE} \geq 0 \quad (12)$$

$$\text{We do not allow the optimal model to have an intercept.} \quad (13)$$

We impose the first constraint because of our aim to make the models the only explanatory variables. Further, we want non-negative weights for interpretation reasons and intuitive results. Either the Executive Board relies fully on one model or a combination of both. A negative weight would thus not be realistic. Lastly, we do not include an intercept as this is not accounted for in Equation 1. In section 7.3.3, we present the results from our optimal weighting calculations. Consult Tables 10, 11 and 12 in Appendix F for coefficient estimates and RMSE scores.

7.4.1 Optimal GDP Growth Model Forecasts

When optimizing the GDP growth model forecast, Stata reports a ratio of 1:0. That is, the DSGE model forecast is preferred in its entirety to the BVAR model forecast. Therefore, our optimal GDP forecast corresponds to the DSGE forecast, with the exception of an intercept. As previously detailed, both the DSGE model forecast and the published forecast failed to accurately anticipate the GDP growth hike in November 2010 and its succeeding fall. Post-2011, the forecasts develop in much the same fashion without any discernable deviations. Upon comparison of RMSE, the optimal forecast produces a lower score. This would indicate that the optimal forecast is a better fit to the actual than the published forecast.

The dominance of the DSGE model, in this case, goes in line with the observable lower forecast bias for this model compared to its model counterpart, the BVAR model. See Table 1 in Appendix C for the exact forecast bias averages. As for the relatively minuscule lower forecast bias shown by the published forecast compared to that of the DSGE model, this can partly be explained by the nature of taking the arithmetic average of the deviations – that is, it provides a general overview rather than a point by point evaluation. In any case, an analysis that combines the mean and some measure of variance yields ground for better understanding. The optimal weighting ratio provides us with this needed evaluation.

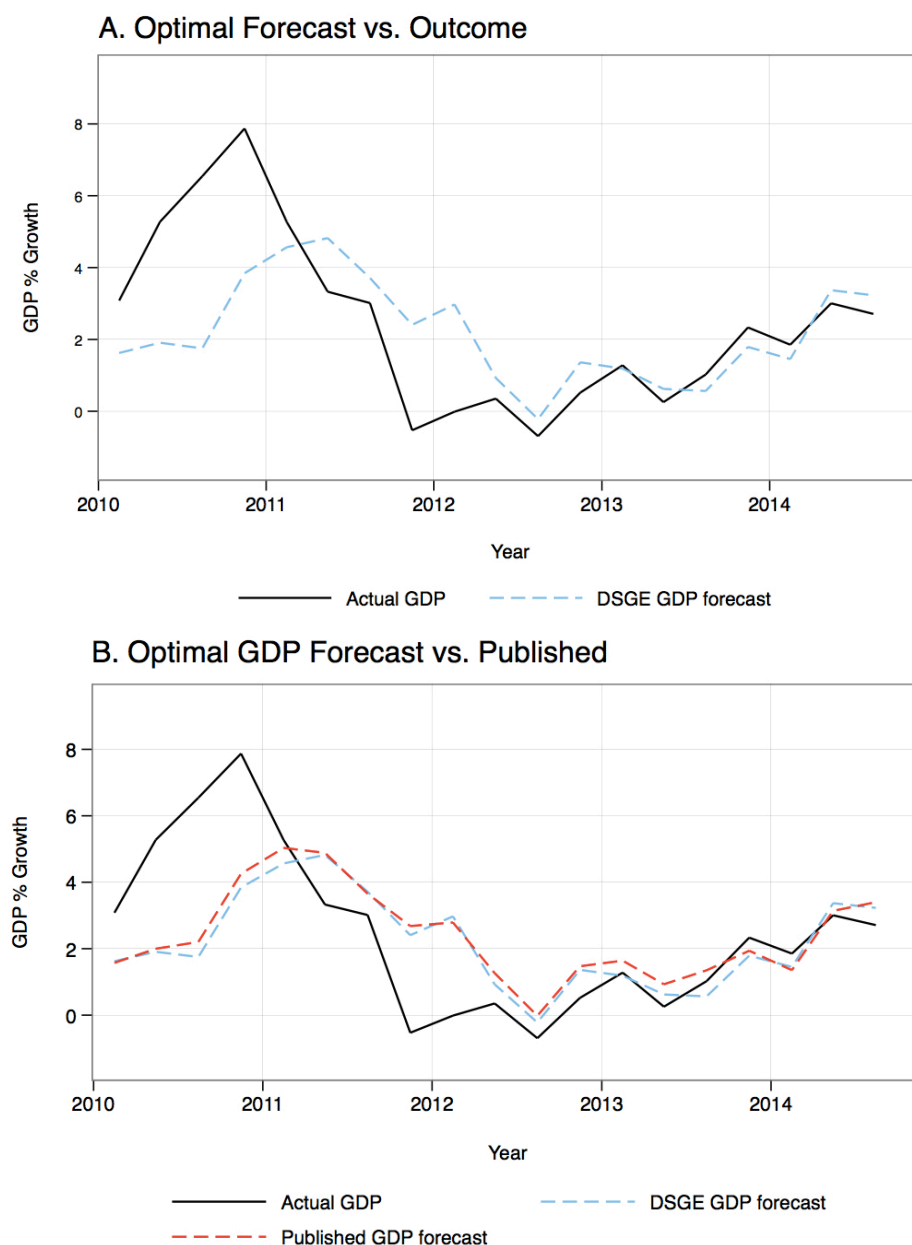


FIGURE 10. OPTIMAL FORECASTS FOR GDP GROWTH. Note the figure charts the development of the optimal model forecasts, along with the line for actual outcome and the competing published forecast from 2010-2014. Chart A shows the optimal model forecast for GDP growth, while Chart B shows the optimal model forecast in comparison with the published forecast for GDP growth. The forecasting components are displayed in the dashed lines. See legends for specific reference. Data source: Lindé and Reslow (2020).

7.4.2 Optimal CPIF Inflation Model Forecasts

We find the optimal model forecast for this indicator assuming a 58:42 ratio, DSGE: BVAR.¹³ As detailed in the visual inspection of the forecasts, both the BVAR and the DSGE model forecasts tail the outcome without any substantial differences in performance, which is why this ratio is unsurprising. Against the published forecast, the optimal model forecast still underperforms,

¹³ Consult Appendix F, Table 11 for estimates.

reporting a higher RMSE score. Comparing the accuracy performance of the two models in section 7.3.2, this ratio makes sense given the relatively close average forecast biases. As for the underperformance of the optimal model forecast, this is not entirely unexpected given the lower average forecast bias of the published forecasts and the noticeable 2% larger deviation, by both models, in mid-2014 to mid-2015.

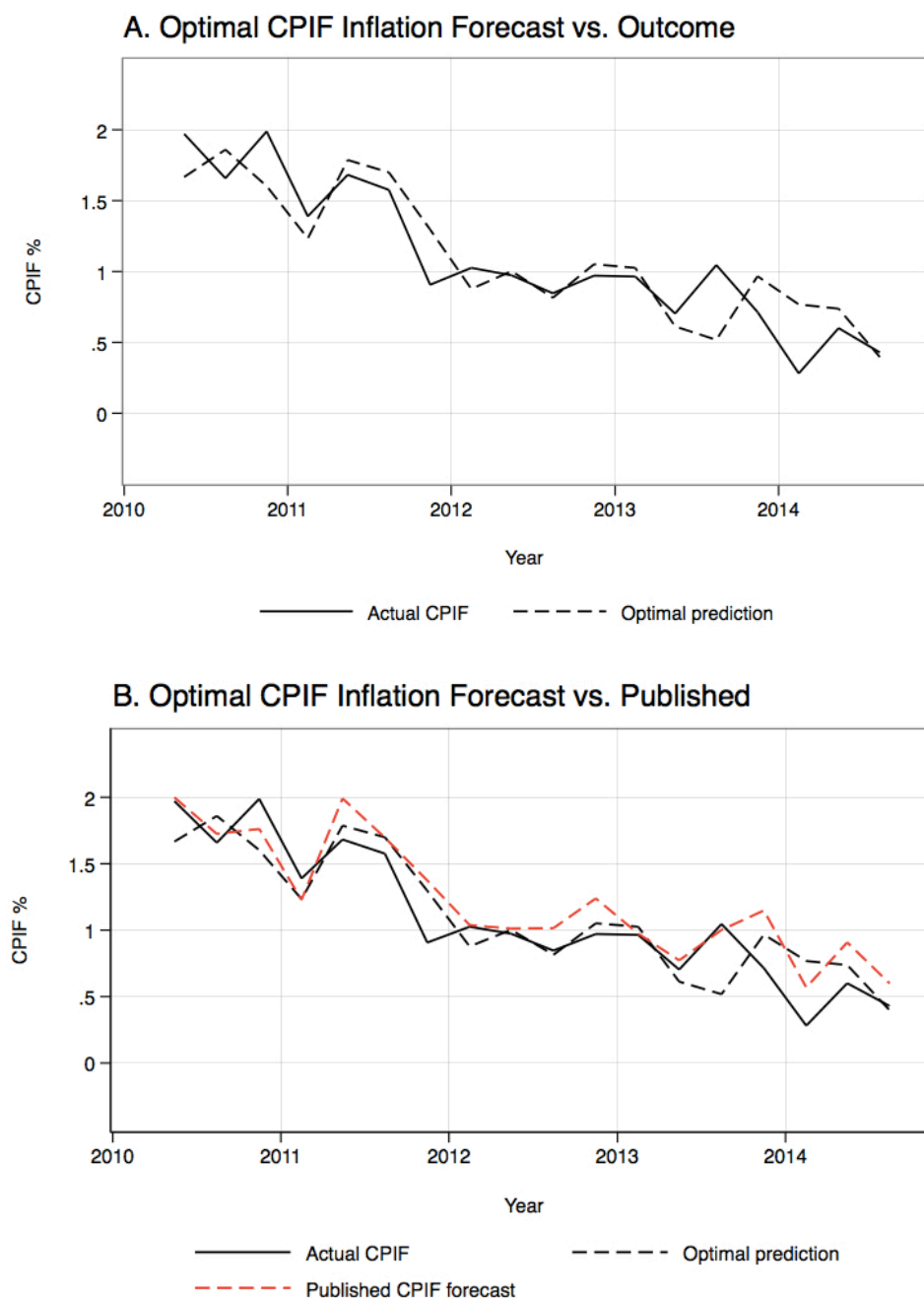


FIGURE 11. OPTIMAL FORECASTS FOR CPIF INFLATION. Note the figure charts the development of the optimal model forecasts, along with the line for actual outcome and the competing published forecast from 2010-2014. Chart A shows the optimal model forecast for CPIF inflation, while Chart B shows the optimal model forecast in comparison with the published forecast for CPIF inflation. The forecasting components are displayed in the dashed lines. See legends for specific reference. Data source: Lindé and Reslow (2020).

7.4.3 Optimal Repo Rate Model Forecasts

The optimal forecast assumes a model weight ratio of 71:29, DSGE: BVAR.¹⁴ Again, the result is unsurprising given what has been detailed in the sections about forecast biases and t-test.

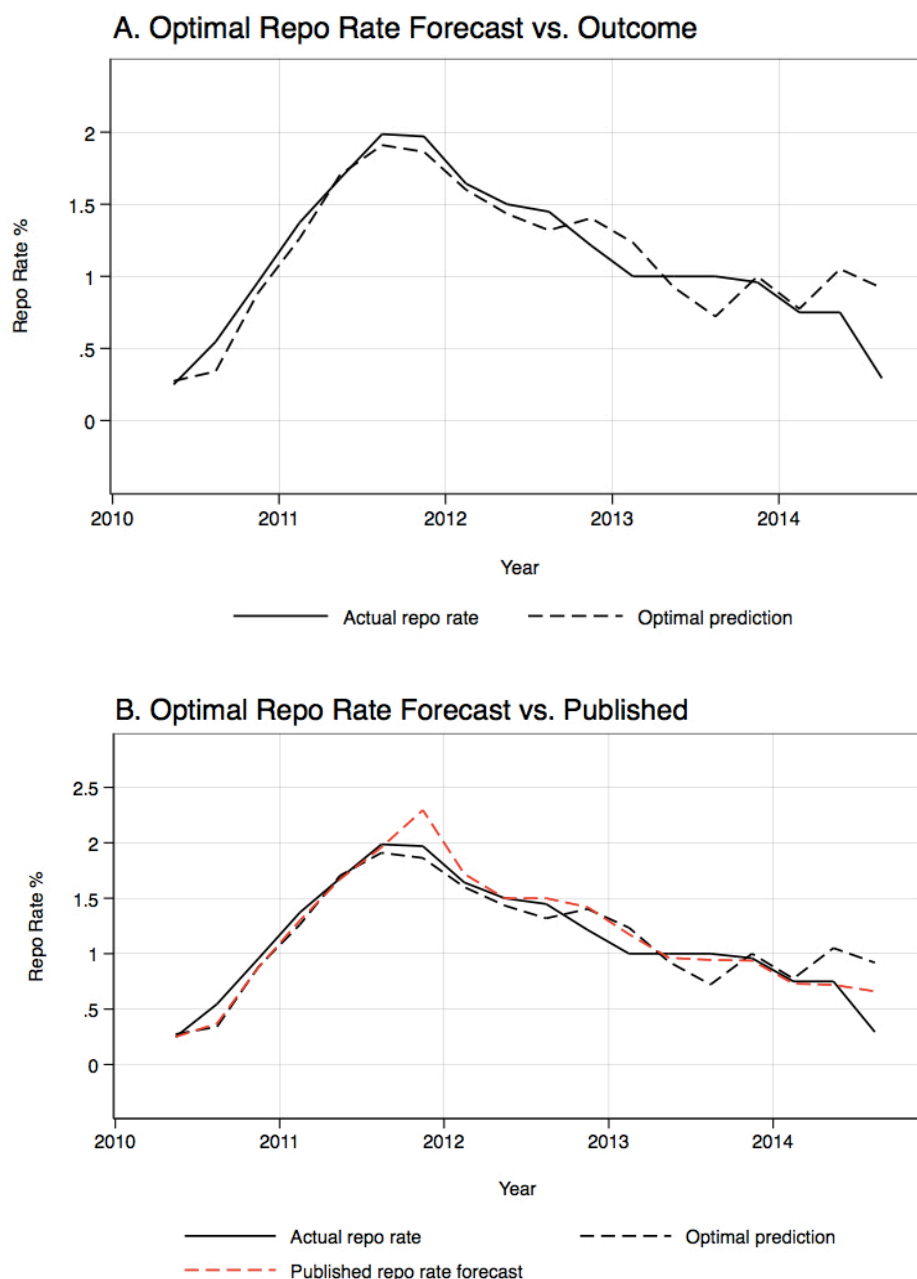


FIGURE 12. OPTIMAL FORECASTS FOR REPO RATE. Note the figure charts the development of the optimal model forecasts, along with the line for actual outcome and the competing published forecast from 2010-2014. Chart A shows the optimal model forecast for the repo rate, while Chart B shows the optimal model forecast in comparison with the published forecast for the repo rate. The forecasting components are displayed in the dashed lines. See legends for specific reference. Data source: Lindé and Reslow (2020).

¹⁴ Consult Appendix F, Table 12 for estimates.

The fact that the Riksbank uses priors for its models, which return the inflation to target, is suggestive in the overshoot in estimation by the end of the period. In terms of fit to the actual, the published forecast outperforms the optimized model forecast. Once again, this performance order is not entirely unreasonable given the larger forecast bias, of nearly 1.5 percent, by the BVAR model compared to the DSGE model. As for the better accuracy performance of the published forecast, its lower forecast bias shown in Figure 2 provides further support. It is also important to note that the relative forecast performances for the repo rate need to be put into perspective. The repo rate is fundamentally not the same as the other macro indicators. While serving as a macroeconomic indicator, it is also an instrument at the hand of the Riksbank. We will further elaborate on this point and all the other results from our data analysis, and their ultimate consistency with the existing body of research in section 6.

8. Discussion

As we have established, the years 2010-2014 have been characterized as a period where the Riksbank's Executive Board displayed supposedly unreasonable amounts of wishful thinking. Over the years, the belief in the Riksbank's overreliance on models grew, particularly following Goodfriend and King's (2016) review. As Lindé and Reslow (2017) explain, "the conclusions of Goodfriend and King appear to have spread both in the mass media and in academia." (28) In rebuttal, they conclude that models do not have a central role in explaining the Executive Board's published forecasts, and that in fact, judgments play an active role in the forecast revisions as well (Lindé and Reslow 2017). If this were the case, it would suggest that the forecasts containing the judgment of the Executive Board – the published forecasts – should have performed relatively worse than the material produced by the model forecasts they were provided. Hence, we ask, which forecasting component was then in fact most accurate? Iversen et al.'s (2016) findings suggest the Riksbank would benefit more from relying on models, and as per their recommendation, we investigate this further with additional formal method. How did the use of judgments and models affect the accuracy of the Riksbank's forecasts during the year 2010-2014? The data reveals part of the story. The qualitative documentation – meeting minutes and interviews – reveals another layer, and the previous research helps solidify the narrative.

8.1 Visual Inspection and Forecast Deviation Sections

As we first noted upon the visual inspection of the data, no forecasting component is an oracle. Comparing the different forecasts with the observed outcomes may reveal systematic performance differences, but it also reveals no model or additional element of judgment to be immune to over or under estimation from quarter to quarter. The similar comment is made, concerning models specifically, by Lindé and Reslow (2017), writing, "It should be remembered, however, that it is not uncommon for different models to be better or worse in different periods." (30). As for the patterns we can deduce, it is clear that the models do not show the same overall level of performance – the BVAR model being the last place contender on the real-time forecast level. This is where the discussion of the performance differences across horizons presents itself. As identified upon the visual inspection of all forecast horizons in Figure 1, the BVAR model shows a notably tighter development along the actual outcome for the two macro indicators, CPIF inflation and the repo rate. Thus, conversely, the BVAR model is the best

performing on this scale. Lindé and Reslow (2017) state, “It is interesting to see how the BVAR model deviates with systematically lower inflation and repo rate forecasts that are much closer to the actual outcome during the period.” (31). The findings by Iversen et al. (2016) further support this observation, writing, “For inflation and the repo rate, the BVAR forecasts outperform the DSGE and the published forecasts.” (15) The Riksbank’s forecast of the imminent return to two percent inflation, criticized in the published forecasts during this period, would therefore have benefited from a greater reliance on the material produced by the BVAR model. According to Jesper Hansson, this imminent return was a problem in the DSGE model at the time, causing the Riksbank to retrospectively recalibrate the model. Consequently, Hansson shared they now lean more towards the BVAR model (Hansson 2020). As for the dominance of published forecasts on the nowcast level for CPIF inflation, this is also observed by Iversen et al. (2016), concluding that the published forecasts perform “fairly well” in the very short horizons, but less so in the long run. (10)

We thus see it important to emphasize the functional differences of these models. Our previous results are supported by the existing body of research, but also confirmed by those who know the models best. When asked about the differences in forecasting accuracy and usability between the DSGE and BVAR models, and their respective strengths and weaknesses, the high-ranking member of MPD staff first explained that the DSGE model is completely structural. They continued by sharing: “That is, it is based on micro foundations, assuming preferences of different actors in the economy – assuming something about basic behavior.”¹⁵

Another MPD staff member further explains that the DSGE model’s greatest contributions are its insights on shocks, giving the example, it is “such that goes directly to how careful people are to save or consume.”¹⁶ The BVAR model, however, can be made structural and according to the previously named high-ranking member, “models with less theoretical restrictions such as BVAR, make better forecasts”.¹⁷ This would support the model’s systematic better performance of accuracy in a longer horizon for both CPIF inflation and the repo rate, but not its evident weaknesses in the nowcasts, as seems to be the case in our data analysis. According to CPIF inflation findings by Iversen et al. (2016), “conditional DSGE model forecasts are competitive with the BVAR model forecasts at horizons up to six quarters” (11), which would suggest the BVAR model’s relatively weaker performance in the short run.

As for the similarities in forecasting accuracy between the DSGE model and published forecasts visible across all macro indicators, this is also noted by Iversen et al. (2016), who write, “the published forecasts for inflation and the repo rate have been more closely related to the DSGE model forecasts than the BVAR forecasts.” (17) The extremely close forecasts for the GDP growth we observe in the forecast bias, not noted by Iversen et al. (2016), can be attributed to the entirety of horizons they consider and the slightly different period of focus. They continue by arguing that, “this is because the DSGE model’s advantages when it comes to interpreting the forecasts and building a story around the forecasts.” (3) According to an MPD staff member, the

¹⁵ Interview with member #2 of Monetary Policy Department Staff, February 27, 2020.

¹⁶ Interview with member #1 of Monetary Policy Department Staff, February 18, 2020.

¹⁷ Interview with member #2 of Monetary Policy Department Staff, February 27, 2020.

BVAR model does not deliver explanations like those of the DSGE model, stating that the DSGE model provides explanations based on various disruptions that can occur.¹⁸

Iversen et al. (2016) touch upon an observation that relates to suspected tendency of the Executive Board to favor one model over another. That is that, the DSGE model differs from the BVAR model in anticipating a faster return to the target inflation rate of two percent. This could provide part of an explanation as to why the DSGE and published forecasts show similar developments across all horizons in Figure 1. Given the Riksbank's view of an appropriate monetary policy path, the inclination to expect a faster return to the formal inflation target should be expected and "natural for the judgmental forecast," according to Iversen et al. (2016, 12). This could, however, be problematic if the models themselves produce unrealistic returns to the inflation target and provide misleading substance for the bias in the judgmental adjustments. We will elaborate on this further in 8.2 and 8.3.

As for longer forecast horizons, the published forecast continually worsens for CPIF inflation and the repo rate compared to the nowcast. We see the inverse relationship for GDP growth. This finding is interesting given the increased reliance on judgmental adjustments for longer horizons, as already established. We see much of the same pattern for the BVAR model with the exception that average GDP growth biases are systematically higher across all horizons. Compared to DSGE, the BVAR model is better at predicting CPIF inflation and the repo rate looking both one and two years ahead. This confirms previous research of which we have already discussed at length. Sweeping conclusions are, however, hard to draw considering that both models outperform the published forecast in predicting the repo rate when the horizon is eight quarters but are, in turn, outperformed by the published forecast at four. It is not surprising that this is the case given the repo rate's duality as an indicator and instrument. As for CPIF inflation, the only horizon at which the published forecast outperforms the model forecasts is at the nowcast. Again, this contradicts suggestions that model influence is decreasing in importance over time.

The unique usefulness and superiority of each model also highlights the importance of being aware of each model's limitations. As the high-ranking MPD member shares, the BVAR model may give a lot of valuable input in the long run, but less so in the immediate future.¹⁹ This leads to the discussion on the marginal utility of additional information introduced in our theoretical framework. A model may be able to produce a forecast for a certain horizon or macro indicator, but this does not imply it will be entirely useful. The additional material, or in other words, the additional information, may show that less information is preferred to more, especially if a model systematically produces worse forecasts for a particular macro indicator or horizon. The models' functions and consequent predictive capabilities are critical. To ensure the maximum accuracy of the Riksbank's forecasts, the reliance on a particular forecasting component should thus vary upon context. Discussions in 8.2 and 8.3 will support this further.

¹⁸ Interview with member #2 of Monetary Policy Department Staff, February 27, 2020.

¹⁹ Interview with member #2 of Monetary Policy Department Staff, February 27, 2020.

8.2 Testing Forecast Errors for Statistical Significance

The overarching takeaway from our testing of forecast errors is the failure to reject the null in the DSGE and published forecast comparison and across all macro indicators. In other words, we cannot rule out the possibility of equal forecasting ability. This further confirms the findings presented in section 8.1 above. Particularly in the case of GDP growth, the Riksbank's preference for the DSGE model is coherent with the optimal model GDP growth forecast, which attributes zero weight to the BVAR model.

Further, failure to reject the null indicates a reliance on the DSGE model as a base for reasoning. The DSGE model's dominance over the Riksbank's forecast in the nowcast technically contradicts some of the existing research. Yet if we refer to the forecast biases for CPIF inflation and the repo rate for eight quarters ahead, the BVAR model outperforms the rest. This is the pattern we should expect to see according to Iversen et al. (2016). That is, we should expect the BVAR model to dominate inflation and repo rate forecasts – something we cannot observe for nowcasts. In other words, the BVAR model's functional capabilities are catered towards a longer horizon. Thus, it would be wise for all future studies to be mindful of the fact that model accuracy varies with forecast horizon.

If there is a preferred model component, it is important, to say the least, that we fully evaluate the strengths but also the weaknesses of this component. Since the alleged period of wishful thinking, the Riksbank has developed a successor, Maja, to the current DSGE model.²⁰ According to Hansson, post-2014, the DSGE model was recalibrated and more weight was shifted towards BVAR to avoid overoptimistic forecasts of how fast inflation could return to the target (Hansson 2020). This bears implications for the Riksbank's main instrument, the repo rate, which is discussed in further detail in section 8.3.

As for the new DSGE model, Maja, it has been rewired to more accurately estimate the correlation between the Swedish economy and international developments – something of which was later realized to be insufficient during our period of review.²¹ For instance, the Riksbank underestimated the effect of the strengthening krona, as well the developments in the euro area, leading to a delayed response in early 2012 (Goodfriend and King 2016). Consequently, with the macro models unable to detect this correlation, the only possibility for this information to be included would be through the judgmental adjustments made by the Executive Board. In practice, the value of any adjustment was made contingent on the Executive Board's understanding of the economy and its access to outside-model information at the time.

From this we gather two things. Firstly, it supports the idea that additional inside-model information could improve forecasting and by extension monetary policy decision-making. This would support the idea that more is, indeed, more and that despite any cognitive limitations and bounded rationality among the Executive Board, this was, in fact, the only way the Riksbank could better the accuracy of their forecasts. In this light, dissent among the Deputy Governors

²⁰ Interview with member #2 of Monetary Policy Department Staff, February 27, 2020.

²¹ Interview with member #2 of Monetary Policy Department Staff, February 27, 2020.

could be interpreted as a positive sign that the model-based forecasts were given merited criticism. Secondly, it shows the importance of great flexibility in the forecasting process and the ability to adapt what information should be considered over time. For the Governors of the Executive Board, this arises in the ongoing development of understanding the economy in real-time. For the models, this is factored into systematic improvements and model revisions.

We see a further connection to the discussion on ecological rationality and how decisions-makers often must consider direct threats to their future survival and reproduction. For the Riksbank, it was the danger of losing the public's confidence in the inflation target and with it the legitimacy of their forecasts. With CPIF inflation undershooting target since 2010, Governor of Sveriges Riksbank, Stefan Ingves, explains that it became "urgent to bring inflation back up towards two percent fairly rapidly... It was therefore clear that the Riksbank needed to focus entirely on the task of maintaining confidence in the inflation target" (Ingves 2019, 5). In times of uncertainty, the Riksbank must act on the most alarming information.

8.3 Optimal Weighting Ratios

The debate about model-based forecasting versus judgmental adjustments has aroused studies in support of either one of the two components. Once again, we call upon the two intuitively opposing findings to propel our discussion forward. While Iversen et al. (2016) argue that model-based forecasts often outperformed the Riksbank's published forecasts in predictive accuracy during the period reviewed, Lindé and Reslow's (2017) calculations suggest judgmental adjustments had a dominating influence on the very same published forecasts. To determine the value of judgmental adjustments, and with them the value of additional outside-model information, we compared the performance of the published forecast to that of a hypothetical, optimized model-based forecast. These model-based forecasts are computed only using nowcasts, for reasons explained in section 6.1.3. On two out of three accounts the published forecast outperformed our optimal model forecast.

The optimal model's slight preference for the DSGE model over the BVAR model, in the case of CPIF inflation, is consistent with the reported forecast biases and RMSEs. It also matches the relative advantages of the DSGE model, as identified by Iversen et al. (2016). In spite of the fact that BVAR model forecasts for inflation and the repo rate performed well during the period reviewed, as seen in Figure 2-3, the published forecasts for these indicators have put greater faith in the DSGE model. One plausible explanation as to why this is, is provided by Iversen et al. (2016) who argue that "this is because of the DSGE model's advantages when it comes to interpreting the forecasts and building a story around the forecast." (17) However, as the published forecast outperformed the optimal forecast, judgmental adjustments should be considered value-adding. In other words, they yielded more accurate Riksbank forecasts in this horizon. Not only does this finding contradict Iversen et al. (2016), but it suggests the Riksbank does well to trust its own judgmental adjustments. This is yet but another account of how forecasting accuracy differs across horizons.

An additional factor, which has already been alluded to in section 8.1, is the tendency of the DSGE model to approach the inflation target faster than the BVAR model. According to Iversen et al. (2016), the tendency has led to “a systematic overprediction of inflation since 2011.” (11) They view the tendency of biased inflation forecasts as prevalent at all central banks that engage in inflation-targeting. This then pulls into question the advantages and disadvantages of the practice. Svensson puts the discussion into a historical perspective, sharing, “I believe it fair to say that never before in monetary policy history has an incentive system been set up with such strong incentives for optimal policy decisions” (Svensson 1999, 663). Thus, we cannot say whether the Riksbank would have fared better without inflation-targeting incentives. A fast-and-frugal-heuristic in this uncertain time may have been better than none at all.

Related to the discussion of bias, several sources in the Riksbank have suggested that the Monetary Policy Department (MPD) actively considers the positioning and likely sentiments of the Executive Board while constructing the model-based forecasting reports. It has been confirmed by a number of sources within the Riksbank that the MPD may present a weaker inflation forecast upon knowledge about the Executive Board’s pessimism. According to the high-ranking MPD member, “in order for the Executive Board to feel comfortable with what is later published, we need to consider their [the Governors’ of the Executive Board] position. It can be about certain preferences or spirit, or about attributing different weights to different types of forecasts.”²² Albeit, the Executive Board remains outermost responsible for the Riksbank’s monetary policy, often enough, the MPD has the last say about model forecasts. This relates to the fundamental distinction between judgment and models, and how models are never entirely devoid of the former. If judgmental adjustments by the MPD are made in anticipation of the Executive Board’s reservations, underlying bias could be rooted in the model-based forecasts. To revert to the question about wishful thinking, it could be that the supporting model material was overoptimistic in the first place and that succeeding adjustments inflated the bias even more. This was also suggested by Svensson (2016), who criticized the Executive Board for intentionally exploiting overoptimistic inflation forecasts to justify a higher repo rate. Thus, multiple model biases could be at play, including biased inflation expectations, underestimated correlations and those caused by anticipation of the Executive Board’s reservations. The degree of consciousness and intentionality, however, differs. In any case, biased estimates and underlying unrealistic assumptions in the model-based forecasts could mislead judgment and cause even greater inaccuracies.

Interestingly, our calculations found the optimal GDP growth forecast to outperform the published forecast. We attribute this to three factors. Firstly, as shown by Lindé and Reslow (2017), the influence of judgment is somewhat lower for GDP growth compared to the other two macro variables. In essence, the published GDP growth forecast relies relatively more on model components compared to the other macro indicators. Secondly, the DSGE model has certain advantages over the BVAR model in predicting GDP growth.²³ As already mentioned, the latter model is mainly preferred for the repo rate and inflation forecasts in the longer horizon. In other words, we expect the DSGE to be relatively more insightful for GDP growth predictions.

²² Interview with member #2 of Monetary Policy Department Staff, February 27, 2020.

²³ Interview with member #1 of Monetary Policy Department Staff, February 18, 2020.

Third, and lastly, the models' influence decreases the longer the forecast horizon (Lindé and Reslow 2017). Because our calculations are made with nowcasts, this suggests that the model influence should be relatively high to judgmental adjustments. Thus, we expect an optimal model forecast that mainly draws on the DSGE model. What we however failed to anticipate, was the optimal model's complete disregard for the BVAR model. A plausible explanation is the BVAR's usefulness for longer forecast horizons, as previously established. Although sweeping conclusions like the BVAR model's uselessness in predicting GDP growth in the short run are tempting, perhaps it would be more useful to put it through a theoretical lens. In this instance, the BVAR model had a limited accuracy contribution for the Riksbank. In other words, less information, particularly from this model, generated more.

As for the repo rate, one must consider the fundamental interdependence between the instrument and the CPIF inflation forecast. A slight dependence on the DSGE model in the inflation forecast complicates the projected repo rate path. Given that the DSGE model prognosticates a faster return to the inflation target in the short horizon, it implies a higher than justified repo rate. This is what we were weakly able to discern following the visual inspection of the forecasts and in particular by the end of the period reviewed. This was, in turn, magnified in the longer horizons. As seen in Chart B of Figure 4, by the end of 2013 and beginning of 2014, actual CPIF inflation was still declining, yet all forecasts were predicting a rise. Thus, underlying bias in the inflation forecast transcends to the repo rate forecast. If there ever was any wishful thinking, it could be attributed to upwardly biased inflation forecasts. There are also two cases where published forecasts outperformed the optimal model forecasts for CPIF inflation and the repo rate. It could demonstrate a predictive capability that models lack in this short-term horizon and given that this period succeeded a financial crisis. That is, the Executive Board is able to consider information that the models do not, suggesting that more information is preferred. It also speaks to the general forecasting contributions of judgments. Svensson (2005) explains it as, "the crucial and beneficial role of judgment – information, knowledge, and views outside the scope of a particular model – in modern monetary policy." (2)

8.4 Limitations and Future Studies

As any person well versed in statistics, it is our obligation to communicate the statistical power weaknesses of our forecast analysis. Despite being exhaustive for the nowcasts of this period, a sample size of as many as 19 observations – as is the case for GDP growth nowcasts – cannot compete with the power a larger sample offers. Thus, any statistically significant result and any failure to reject the null that we have reported should be evaluated against this fact. We also refer to other power-limiting considerations in 6.4. As our data analysis is primarily focused on the nowcasts, this presents some limitations and challenges. One being the extent of comparison we can make to previous studies and all their findings. Forecasts display different predictive accuracy over varying horizons. Thus, more can be said about the relative accuracy performance of the forecasting components across a longer horizon. In addition, because our study is largely quantitative, we cannot, as a result, take a stand on what exact information was available to the Executive Board at the time and the foundations to the over optimistic characterization of this period. This would require, for instance, further qualitative analysis of minutes, which would be interesting in future work.

This leads us to our recommendations for future studies. First, future research should continue to expand to include more models other than the DSGE and BVAR models, as well as more formal methods of analysis. The aggregate macroeconomic model called ‘Moses’ is an example. There are even further possibilities to additional formal methods, such as those presented by Mincer and Zarnowitz (1969), as recommended by Iversen et al. (2016). Second, it is relevant that research should formally evaluate whether there is a tendency to choose models that persistently favor the Riksbank’s inflation target. After discussions with the MPD, this could apparently be evaluated further.²⁴ This is also supported by the observations made in our data analysis. Third, and lastly, further forecast accuracy comparisons can be made with external actors. Accuracy comparisons of Riksbank published forecasts with forecasts produced by external sources is an example. For instance, whether the Riksbank and external forecasters persistently overestimated inflation during the period reviewed would be of relevance. According to an economic commentary by MPD staff member, Mårten Löf, “both the Riksbank and other forecasters have overestimated inflation on several occasions during the period 2013-2015” (Andersson et al. 2015, 1). It could also be extended to the relationship between forecast accuracy and market expectations.²⁵ Consequently, we see many extensions for future study.

9. Conclusion

In this study, we ask how did the use of judgments and models affect the accuracy of the Riksbank’s forecasts during the years 2010-2014? We examine two components – judgment and models – that determine the Sveriges Riksbank’s published forecasts for the three macro indicators: GDP growth, CPIF inflation and the repo rate. The analysis compares the relative forecasting accuracy of the central bank’s primary models, DSGE and BVAR, and the final published forecasts containing the judgmental adjustments of the Riksbank’s Executive Board. The study draws on both graphical and econometric analyses to answer the research question.

Initial visual inspection allows for analysis across all forecast horizons during the period of review. We also examine the average forecast bias for the nowcasts, as well as for four and eight quarters ahead. We continue the analysis of the nowcasts by testing the difference in mean squared errors for significance to deduce whether the forecasting components display equal forecasting abilities. Finally, we examine if a weighted average of the two model forecasts achieves a better fit to the actual outcome than the published forecast using comparisons of RMSEs.

To summarize our findings, published forecasts outperformed both model forecasts in the nowcasts, although only slightly better than the DSGE model’s forecasts for GDP growth. The forecasting accuracy of the published forecasts across four to eight horizons, however, progressively worsened for CPIF inflation and the repo rate, whereas the BVAR model,

²⁴ Interview with member #1 of Monetary Policy Department Staff, February 18, 2020.

²⁵ See related comparisons made by Svensson (2010). As Lars E.O. Svensson states in a speech on evaluating monetary policy, “As monetary policy is mainly about managing expectations, particularly expectations concerning future policy rates, it is of great interest to see what extent a published repo rate path has been predicted by the market and other analysts.” (Svensson 2010b, 14)

compared to DSGE, systematically improved. In longer horizons, a substantial difference in forecasting performance is difficult to discern for GDP growth. By incorporating judgment in the Riksbank's nowcasts, we found an improvement in accuracy – more so than what would have been the case if the models were solely used. Judgmental adjustments are thus important for accuracy in the short term. As the development across longer horizons show, judgments become steadily worse at forecasting for CPIF inflation and the repo rate. Whether judgmental adjustments could have been even more accurate in the nowcasts is another question and relates to a further study on the source of the criticized optimism. In the nowcast horizon, our analysis suggests that the judgmental adjustments made in this period in fact bettered the accuracy. However, because models can source and process a substantially larger pool of information, more so than the Executive Board, the relative contribution of the model forecast is still significant and increases as the horizon expands. In the short term, however, models have proved to be an important determinant in the forecasting process. Thus, the better accuracy of the published forecast cannot entirely ignore the influence of the models.

As Swedish economist Gustav Cassel wrote, “a central bank gives guidelines for the country's entire economic life through the choice of its interest rates” (Cassel 1908, 90). Given the legitimacy and the public expectations surrounding an institution of the Sveriges Riksbank's caliber, independent evaluation of its forecasting process is of outmost importance. While the Riksbank has had a tradition of regularly evaluating its published forecasts to external forecasts, model forecasts have earned significantly less review. In this study, we therefore regard both components in evaluating accuracy in poor performing times when the Riksbank failed to achieve its goals. Finally, we ask what lessons this tumultuous period could provide for understanding future monetary policy. We especially emphasize the importance of unbiased forecasts, the significance of forecasting horizon and the role of additional information.

10. References

- Adolfson, M., Laséen, S., Christiano, L., Trabandt, M. and Walentin, K., 2013. "Ramses II: Model Description," *Occasional Paper Series*, **12**.
- Adolfson, M., Laséen, S., Lindé, J. and Villani, M., 2007. "RAMSES: A New General Equilibrium Model for Monetary Policy Analysis," *Sveriges Riksbank Economic Review*, **2**: 5-40.
- Andersson, B., Corbo, V. and Löf, M., 2015. "Why Has Inflation Been So Low?" *Sveriges Riksbank Economic Review*, **3**: 5-47.
- Andersson, M.K. and Den Reijer, Ard H. J., 2015. "Nowcasting," *Sveriges Riksbank Economic Review*, **1**: 75-89.
- Barros, G., 2010. "Herbert A. Simon and the Concept of Rationality: Boundaries and Procedures," *Brazilian Journal of Political Economy*, **30**(3): 455-72.
- Bernanke, B.S., Laubach, T., Mishkin, F.S. and Posen, A.S., 1999. *Inflation Targeting: Lessons from the International Experience*. Woodstock: Princeton University Press.
- Cassel, G., 1908. *Riksbanken Under Krisen 1907-1908*. Stockholm: Hugo Gebers förlag.
- Ekholm, K. Former Deputy Governor 2009-2014, Sveriges Riksbank. Interview February 20, 2020, Stockholm University, Stockholm.
- Gigerenzer, G., 2018. "The Bias Bias in Behavioral Economics," *Review of Behavioral Economics*, **5**(3-4): 303-36.
- Gigerenzer, G., 1991. How to Make Cognitive Illusions Disappear: Beyond 'Heuristics and Biases'," *European Review of Social Psychology*, **2**: 83-115.
- Gigerenzer, G., Todd, P.M. and Group, A.R., 2000. *Simple Heuristics That Make Us Smart*. Oxford, New York: Oxford University Press.
- Goodfriend, M., King, M., 2015. *Review of the Riksbank's Monetary Policy 2010-2015*. 2015/16: RFR7. Stockholm: Riksdagstryckeriet.
- Goodhart, C. and Rochet, J., 2011. *Evaluation of the Riksbank's Monetary Policy and Work with Financial Stability 2005-2010*. Stockholm: Riksdagstryckeriet.
- Hallsten, K. and Tägtström, S., 2009. "The Decision-Making Process: How the Executive Board of the Riksbank Decides on the Repo Rate," *Sveriges Riksbank Economic Review*, **1**: 69-86.

Hansson, J. Head of Monetary Policy Department, Sveriges Riksbank. Interview February 17, 2020, Sveriges Riksbank, Stockholm.

Hertwig, R. and Todd, P.M., 2003. More Is Not Always Better: The Benefits of Cognitive Limits. *Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making*, 213-31. New Jersey: John Wiley and Sons, Ltd.

Ingves, S., "Swedish Monetary Policy Experiences after the Global Financial Crisis: What Lessons are There for Other Countries?" October 15, 2019, Speech, Money Macro and Finance Research Group, Stockholm.

Interview with member #1 of Monetary Policy Department Staff, February 18, 2020, Sveriges Riksbank, Stockholm.

Interview with member #2 of Monetary Policy Department Staff, February 27, 2020, Sveriges Riksbank. Stockholm.

Iversen, J., Laséen, S., Lundvall, H. and Söderström, U., 2016. *Real-Time Forecasting for Monetary Policy Analysis: The Case of Sveriges Riksbank*. Stockholm: Sveriges Riksbank.

Jansson, P. and Vredin, A., 2003. *Forecast-Based Monetary Policy: The Case of Sweden*. UK: Blackwell Publishing Ltd.

Jonung, L., 1999. *Looking Ahead Through Rear-View Mirror. Swedish Stabilisation Policy as a Learning Process, 1975-1995. A Summary*.

Kahneman, D. and Tversky, A., 1974. *Judgment under Uncertainty: Heuristics and Biases*. *Science* 185, **4157**: 1124-31.

Kleinmuntz, B., 1990. "Why We Still Use Our Heads Instead of Formulas: Toward an Integrative Approach," *Psychological Bulletin*: 296–310.

Konjunkturinstitutet, 2013. *Konjunkturläget Augusti 2013*, Stockholm: 105-13.

Lindé, J. and Reslow, A., 2020. *Dataunderlag 2006 till 2016*, Distributed by Jesper Lindé.

Lindé, J. and Reslow, A., 2017. "It's a Myth that the Riksbank's Forecasts Have Been Governed by Models," *Sveriges Riksbank Economic Review*, **1**.

Simon, Herbert A., 1957. *Models of Man: Social and Rational: Mathematical Essays on Rational Human Behavior in Society Setting*. New York: John Wiley & Sons, Ltd.

Svensson, L.E.O. Former Deputy Governor 2007-2013, Sveriges Riksbank. Interview February 24, 2020, Stockholm School of Economics, Stockholm.

Svensson, L.E.O., 2005. "Monetary Policy with Judgment: Forecast Targeting," *International Journal of Central Banking*, **1**(1): 1-54.

Svensson, L.E.O., 2003. "What Is Wrong with Taylor Rules? Using Judgment in Monetary Policy Through Targeting Rules," *Journal of Economic Literature*, **41**(2): 426-477.

Svensson, L.E.O. "Why a Low Repo Rate for an Extended Period?" Speech, Stockholm, May 4, 2010. BIS Review. Available: https://www.bis.org/review/r100505c.pdf?fbclid=IwAR0z_-QNkjisAyyWnCFxYmINrMpw8uGe21z9glgU9wnxEZhXG-_Bon1dtIo. [Feb 24, 2020].

Svensson, L.E.O., 2010. Inflation Targeting. National Bureau of Economic Research Working Paper no. 16654.

Svensson, Lars E. O., Randall S. Kroszner, and David Wessel. "Practical Monetary Policy: Examples from Sweden and the United States [with Comments and Discussion]." *Brookings Papers on Economic Activity*, 2011, 289-352. Available: www.jstor.org/stable/41473601. [May 14, 2020].

Sveriges Riksbank 2020, "Repo Rate, Deposit and Lending Rate," Sveriges Riksbank, Stockholm. Available: <https://www.riksbank.se/en-gb/statistics/search-interest--exchange-rates/repo-rate-deposit-and-lending-rate/> [May 1, 2020].

Sveriges Riksbank 2019a, "The Path to a Monetary Policy Decision," Sveriges Riksbank, Stockholm. Available: <https://www.riksbank.se/en-gb/monetary-policy/the-path-to-a-monetary-policy-decision/> [May 1, 2020].

Sveriges Riksbank 2019b, "Monetary Policy Report September 2019," September, Sveriges Riksbank, Stockholm. Available: <https://www.riksbank.se/globalassets/media/rapporter/ppr/engelska/2019/190905/monetary-policy-report-september-2019.pdf> [March 28, 2020].

Sveriges Riksbank 2019c, "What is the Interest Rate Path?" Sveriges Riksbank, Stockholm. Available: <https://www.riksbank.se/en-gb/monetary-policy/what-is-monetary-policy/what-is-the-interest-rate-path/> [May 1, 2020].

Sveriges Riksbank 2018, "What is Monetary Policy?" January, Sveriges Riksbank, Stockholm. Available: <https://www.riksbank.se/en-gb/monetary-policy/what-is-monetary-policy/> [Feb 21, 2020].

Sveriges Riksbank n.d., "1999: New Law Gives the Riksbank More Independence," Available: <https://www.riksbank.se/en-gb/about-the-riksbank/history/historical-timeline/1900-1999/new-law-gives-the-riksbank-more-independence/> [Feb 23, 2020].

Vranas, B.M.P., 1999. "Gigerenzer's Normative Critique of Kahneman and Tversky," *Cognition*, **76**(3): 179-93.

Vredin, A. Head of General Secretariat, Sveriges Riksbank. Interview February 18, 2020, Sveriges Riksbank, Stockholm.

Wetterberg, G., 2009. *Pengarna & Makten*. Stockholm: Sveriges Riksbank i Samarbete med Atlantis.

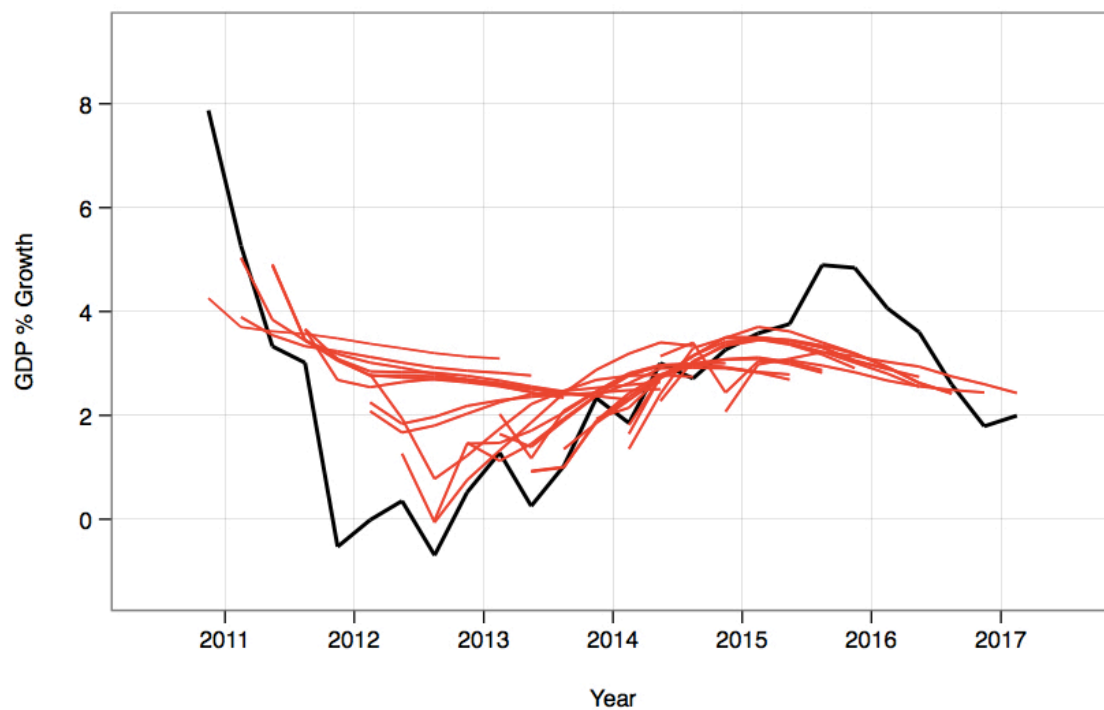
Zarnowitz, V. and Mincer, J.A., "The Evaluation of Economic Forecasts," In: J.A. Mincer, ed, *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*. NBER: 3-46.

Appendices

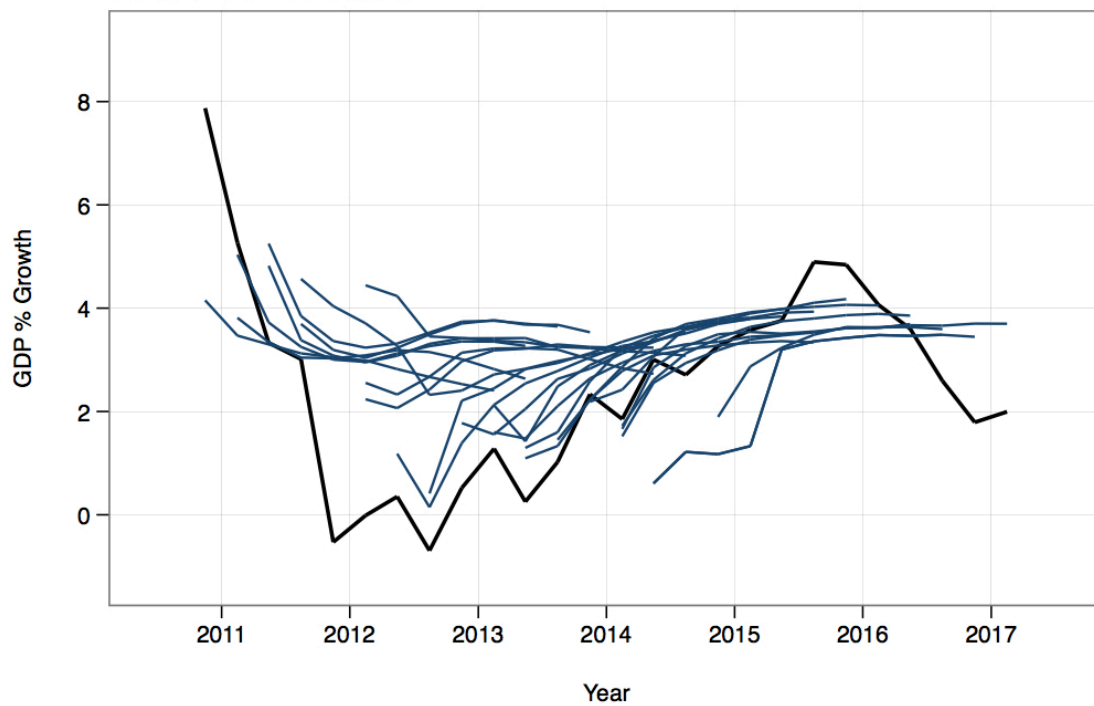
All charts and tables in this study are produced using Stata: Software for Statistics and Data Science.

Appendix A

A. Published Forecasts



B. BVAR Forecasts



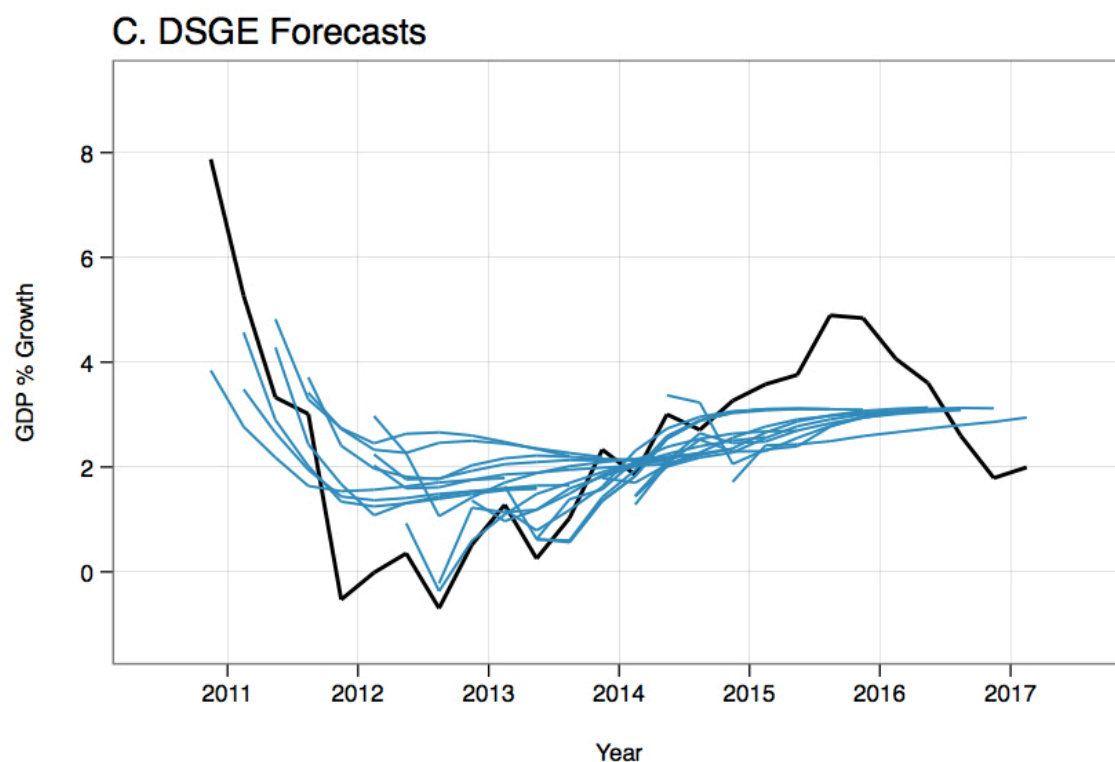
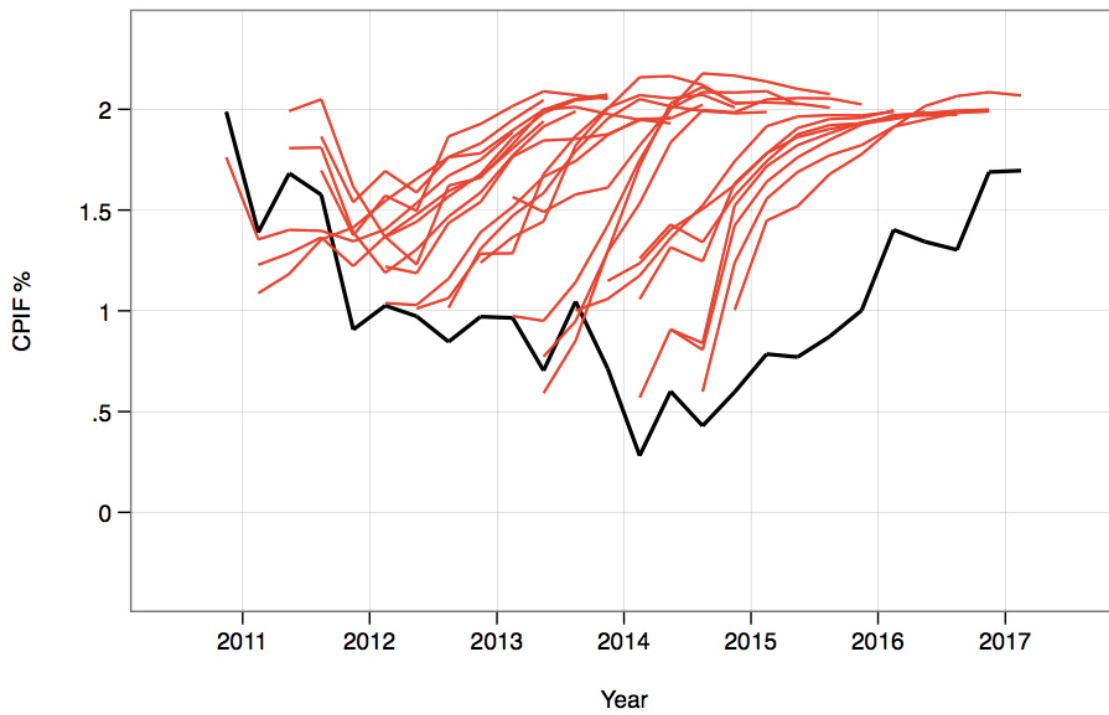
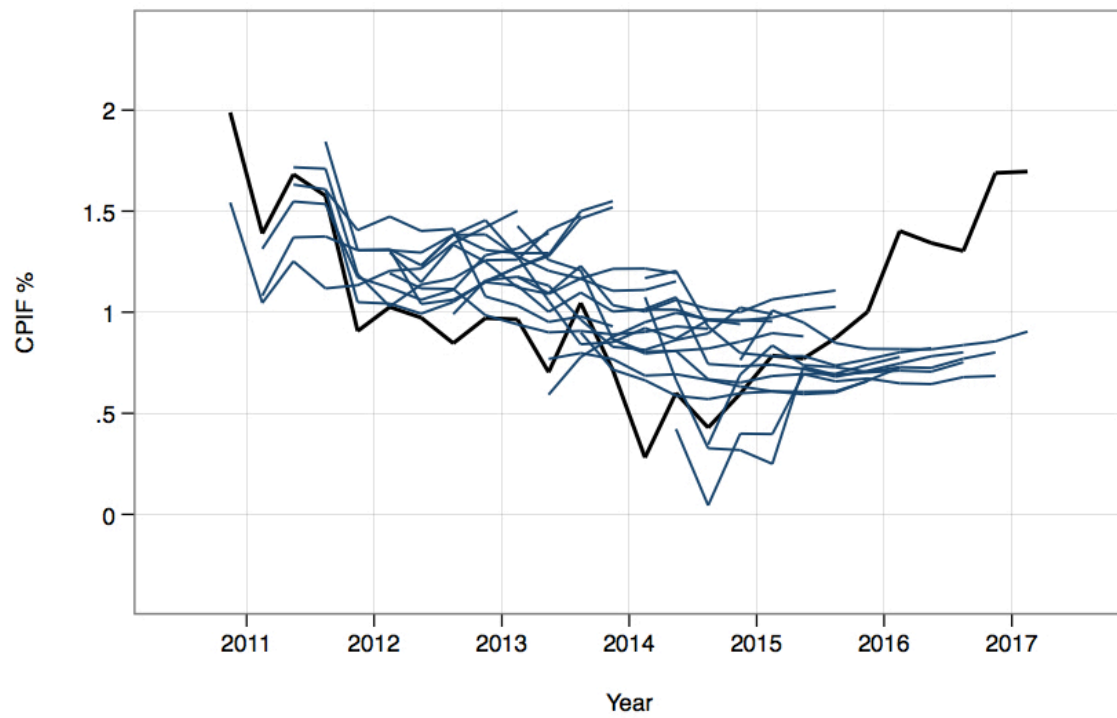


FIGURE 1. GDP GROWTH AND FORECASTS. Note the figure charts the development of the forecasts for GDP growth across all horizons – for each forecasting component – material produced at every meeting, along with the line for actual outcome – the black solid line – from 2010-2014. Chart A displays the published forecasts, while Charts B and C display the BVAR and DSGE forecasts respectively. The forecasting components are displayed in colored solid lines. Data source: Lindé and Reslow (2020).

A. Published Forecasts



B. BVAR Forecasts



C. DSGE Forecasts

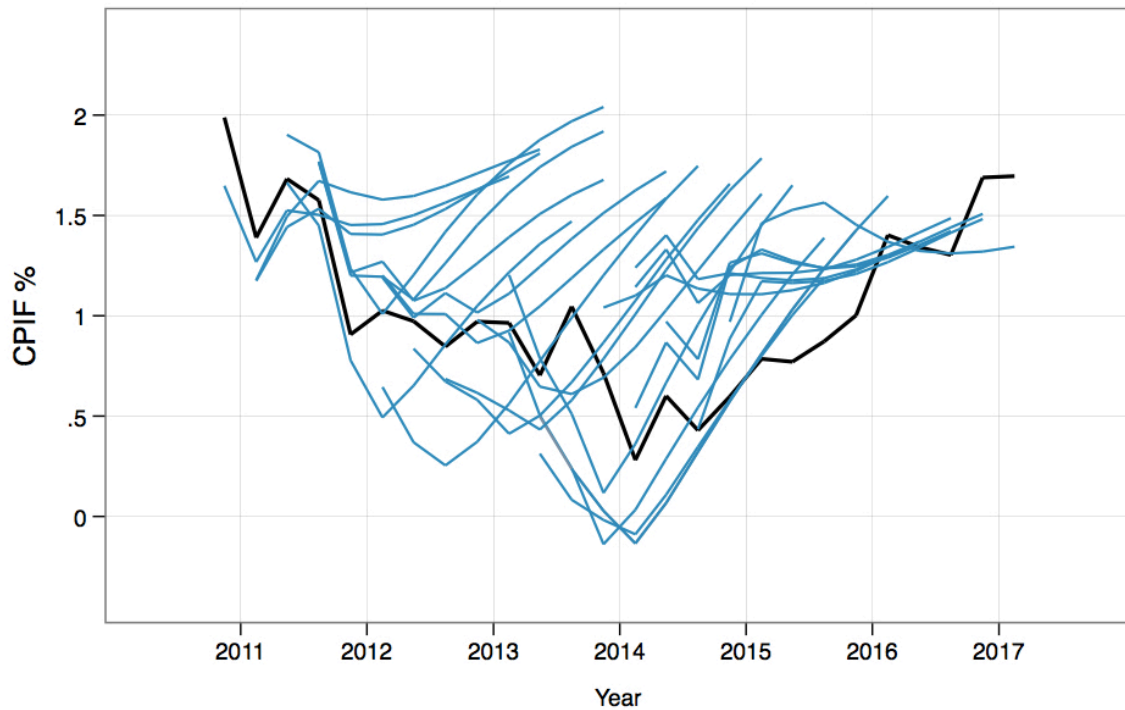
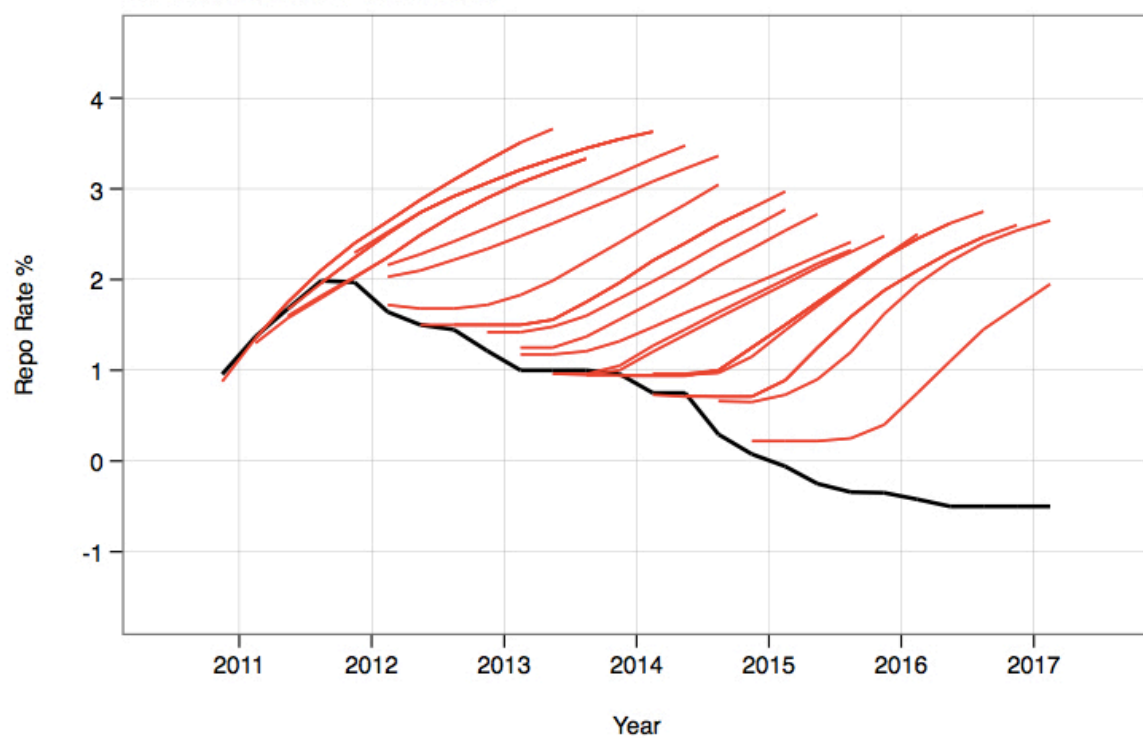
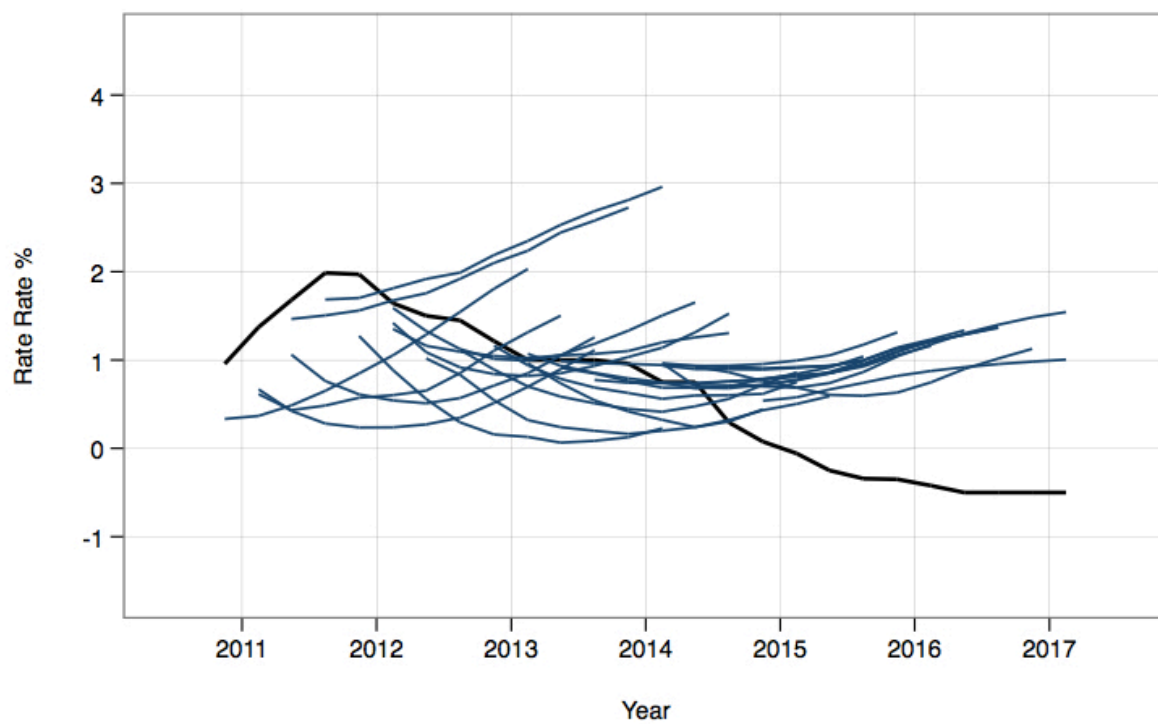


FIGURE 2. CPIF INFLATION AND FORECASTS. Note the figure charts the development of the forecasts for CPIF inflation, across all horizons – for each forecasting component – material produced at every meeting, along with the line for actual outcome – the black solid line – from 2010-2014. Chart A displays the published forecasts, while Charts B and C display the BVAR and DSGE forecasts respectively. The forecasting components are displayed in colored solid lines. Data source: Lindé and Reslow (2020).

A. Published Forecasts



B. BVAR Forecasts



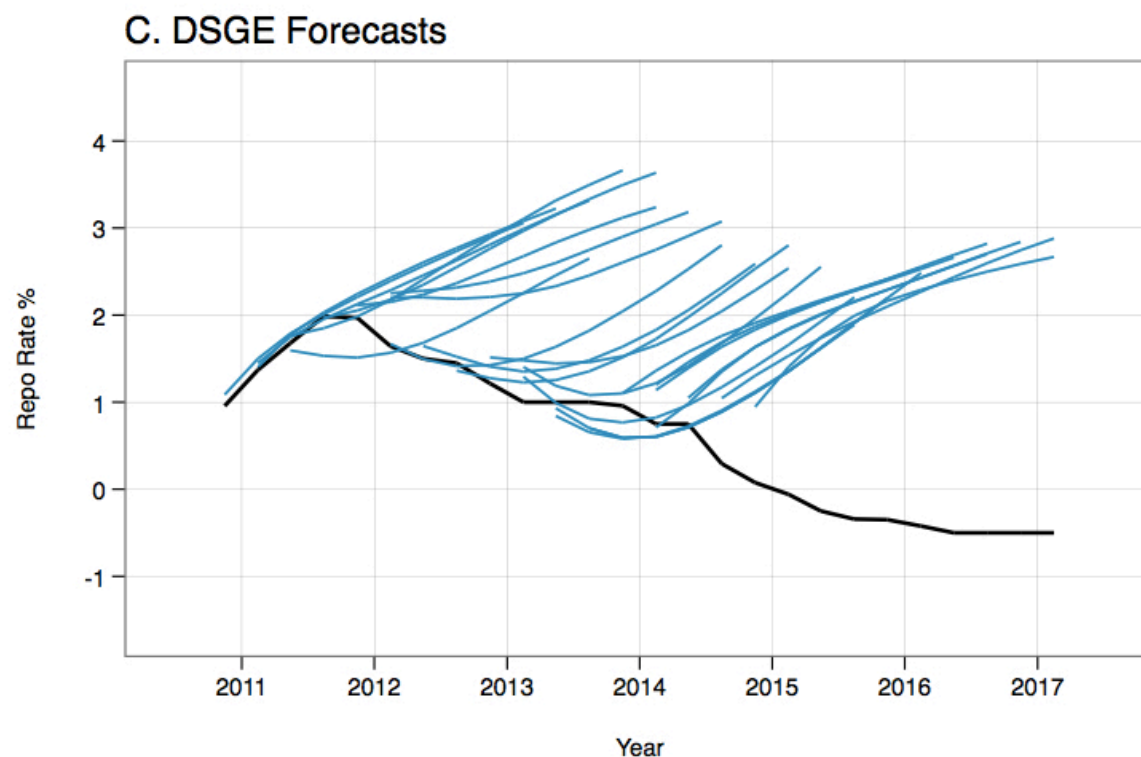
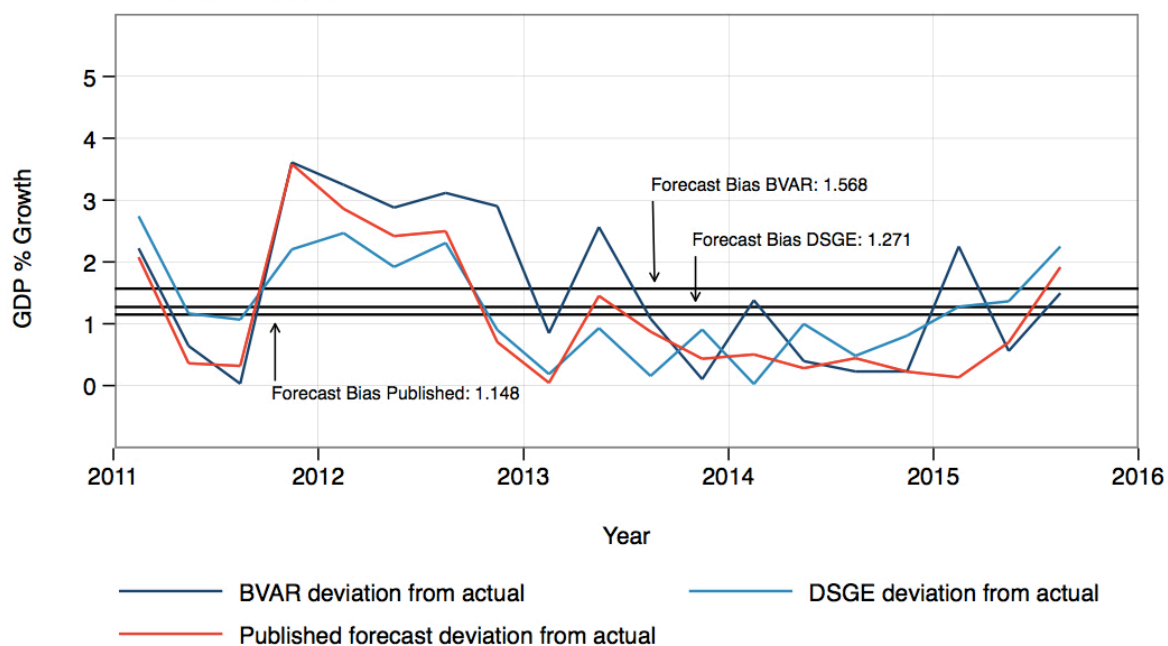


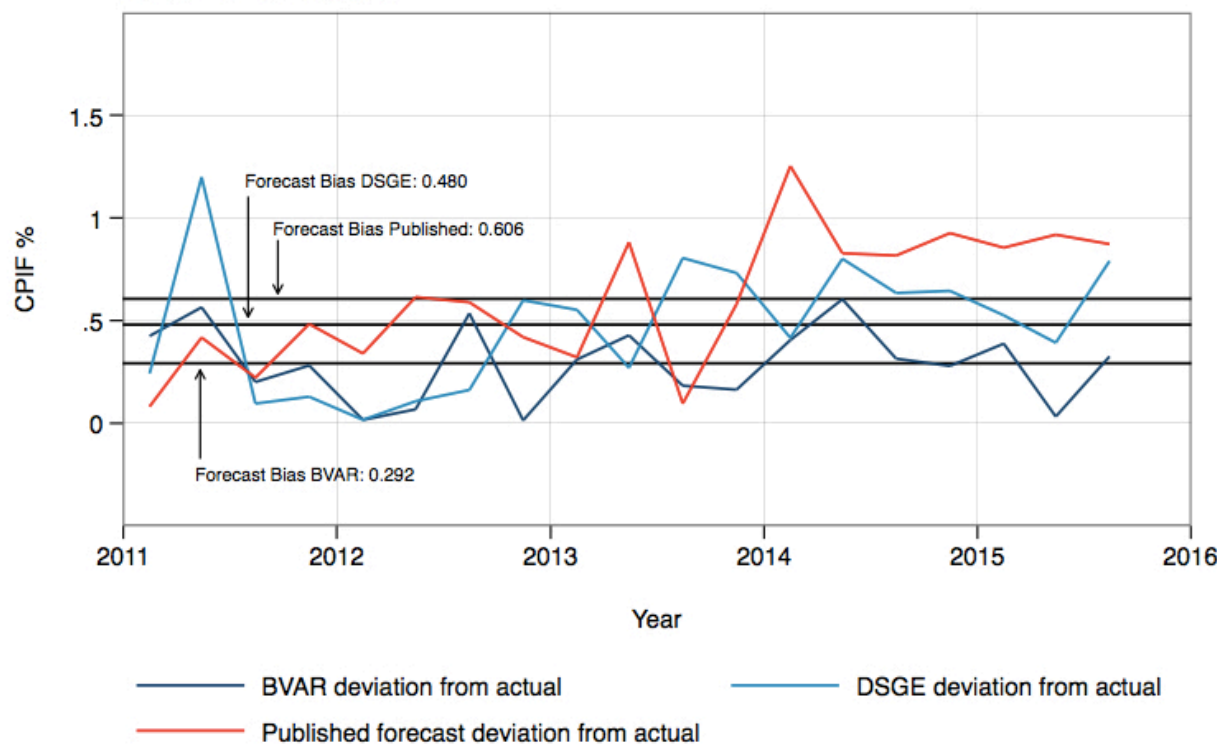
FIGURE 3. REPO RATE AND FORECASTS. Note the figure charts the development of the forecasts for the repo rate, across all horizons – for each forecasting component – material produced at every meeting, along with the line for actual outcome – the black solid line – from 2010-2014. Chart A displays the published forecasts, while Charts B and C display the BVAR and DSGE forecasts respectively. The forecasting components are displayed in colored solid lines. Data source: Lindé and Reslow (2020).

Appendix B

A. GDP Growth



B. CPIF Inflation



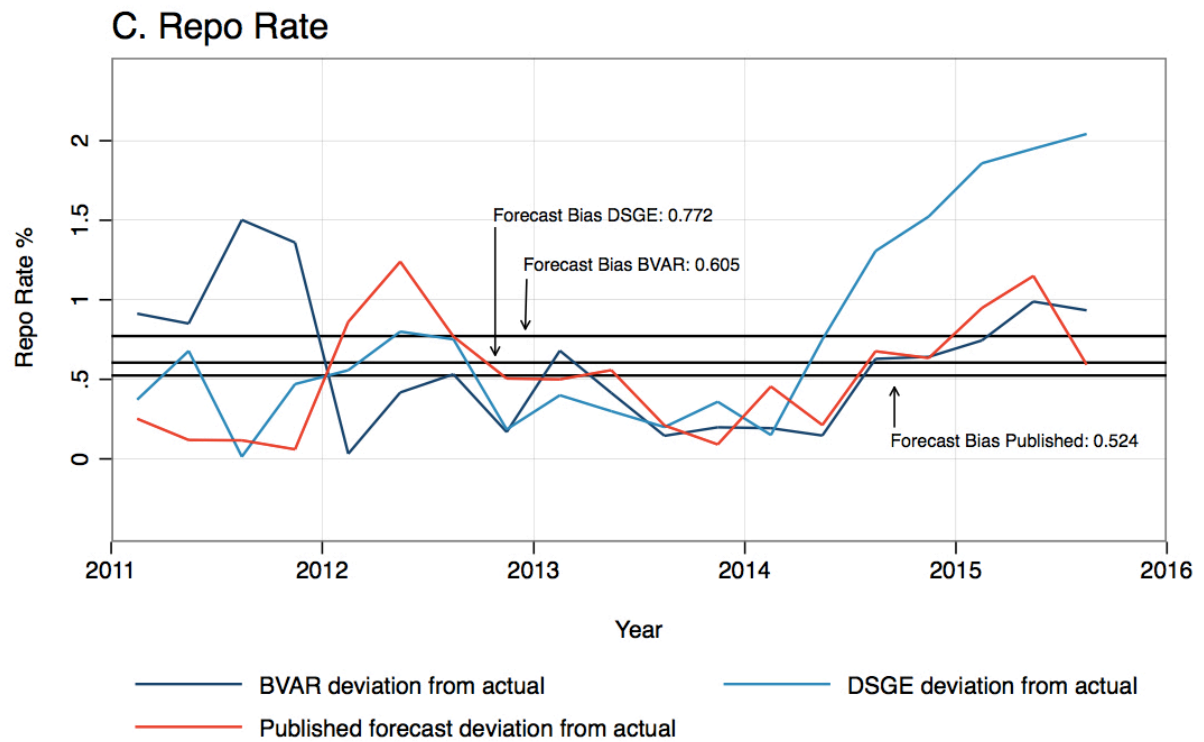
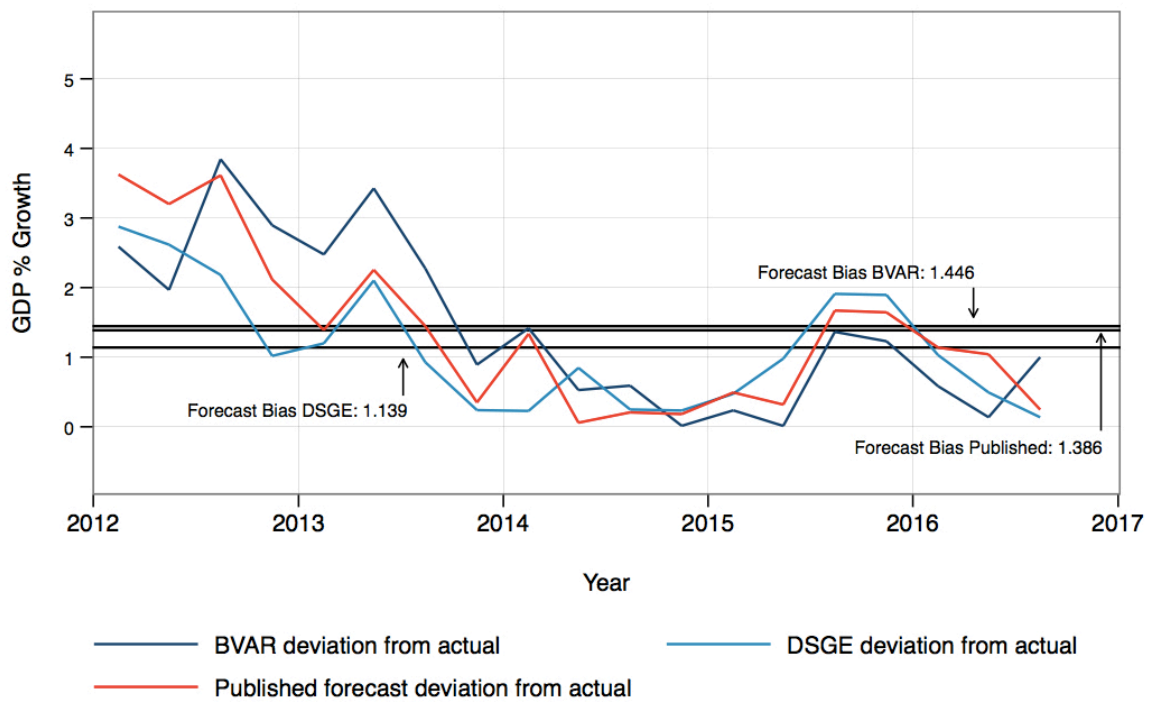
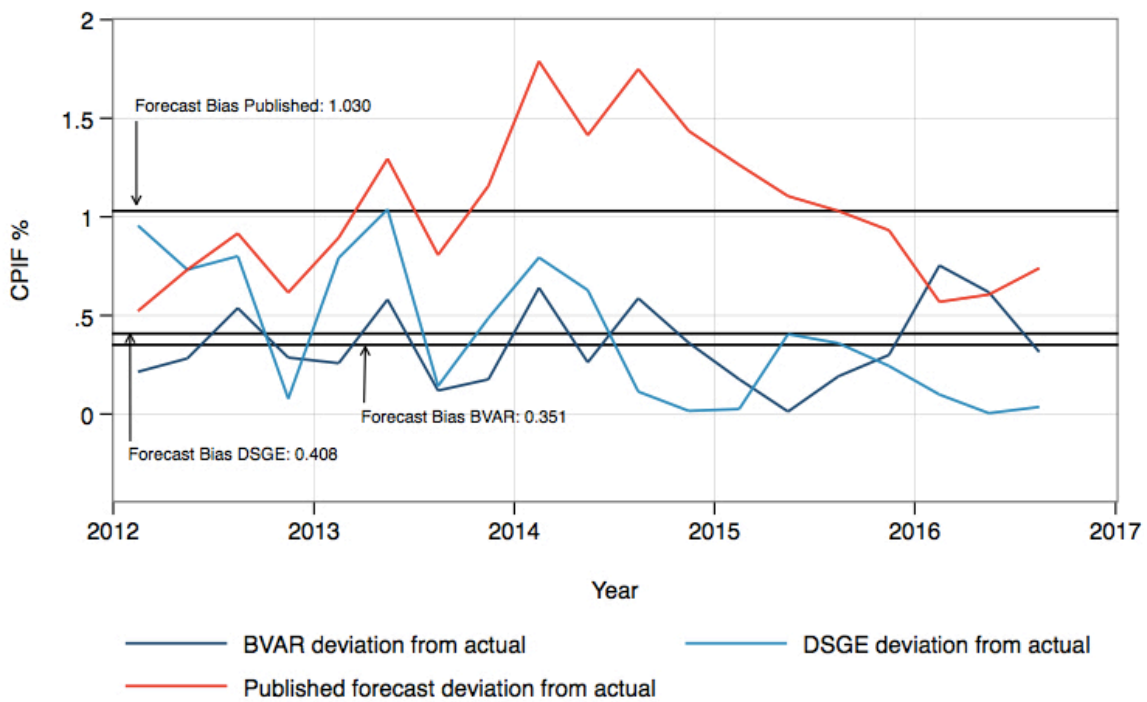


FIGURE 6. QUARTER FOUR FORECAST DEVIATIONS FROM ACTUAL. Note the figure charts the deviation of the forecasts in quarter four from the actual outcome for 2010-2014. Chart A shows the deviations of the forecasts for GDP growth from the actual outcome, while Chart B and C show the deviations of the forecasts for CPIF inflation and the repo rate from their respective actual outcome. Each forecasting component is displayed in the solid lines. See legends for specific reference. Data source: Lindé and Reslow (2020).

A. GDP Growth



B. CPIF Inflation



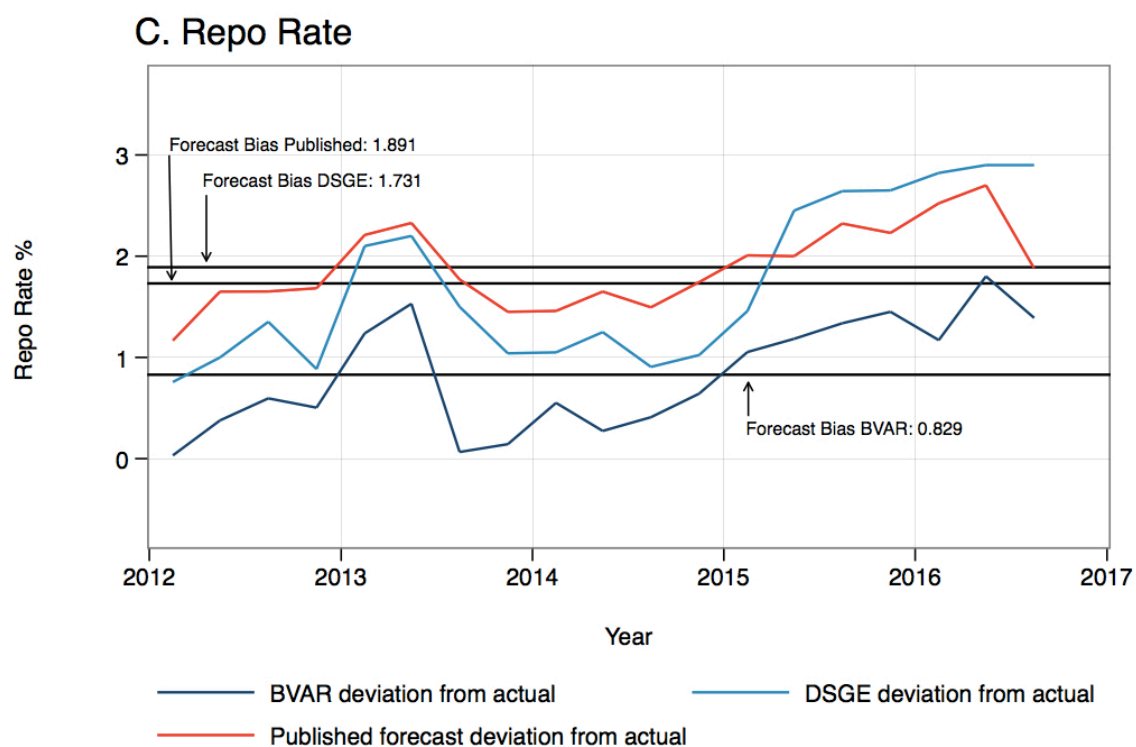


FIGURE 7. QUARTER EIGHT FORECAST DEVIATIONS FROM ACTUAL. Note the figure charts the deviation of the forecasts in quarter eight from the actual outcome for 2010-2014. Chart A shows the deviations of the forecasts for GDP growth from the actual outcome, while Chart B and C show the deviations of the forecasts for CPIF inflation and the repo rate from their respective actual outcome. Each forecasting component is displayed in the solid lines. See legends for specific reference. Data source: Lindé and Reslow (2020).

Appendix C

TABLE 1. AVERAGE FORECAST BIAS OF DEVIATIONS FROM ACTUAL

Macro Indicator	Published	BVAR	DSGE
GDP Growth	1.410 (.306)	1.811 (.352)	1.426 (.330)
CPIF Inflation	.177 (.034)	.248 (.052)	.230 (.045)
Repo Rate	.080* (.022)	.294 (.053)	.145 (.025)

Notes: Table 1 reports the average forecast bias for each macroeconomic indicator, computed by taking the average forecast deviation for GDP growth, CPIF inflation and the repo rate. The degrees of freedom are 18, 17 and 16 respectively. Stars denote statistical significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: Lindé and Reslow (2020).

TABLE 2. QUARTER FOUR AVERAGE FORECAST BIAS OF DEVIATIONS FROM ACTUAL

Macro Indicator	Published	BVAR	DSGE
GDP Growth	1.148 (.247)	1.568 (.281)	1.271 (.189)
CPIF Inflation	.606 (.073)	.292 (.042)	.480 (.072)
Repo Rate	.524 (.082)	.605 (.096)	.772 (.148)

Notes: Table 2 reports the average forecast bias for each macroeconomic indicator, computed by taking the average forecast deviation for GDP growth, CPIF inflation and the repo rate. Df. 18. Stars denote statistical significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: Lindé and Reslow (2020).

TABLE 3. QUARTER EIGHT AVERAGE FORECAST BIAS OF DEVIATIONS FROM ACTUAL

Macro Indicator	Published	BVAR	DSGE
GDP Growth	1.386 (.264)	1.446 (.272)	1.139 (.200)
CPIF Inflation	1.030 (.088)	.351 (.047)	.408 (.082)
Repo Rate	1.891 (.093)	.829 (.127)	1.731 (.181)

Notes: Table 3 reports the average forecast bias for each macroeconomic indicator, computed by taking the average forecast deviation for GDP growth, CPIF inflation and the repo rate. Df. 18. Stars denote statistical significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: Lindé and Reslow (2020).

Appendix D

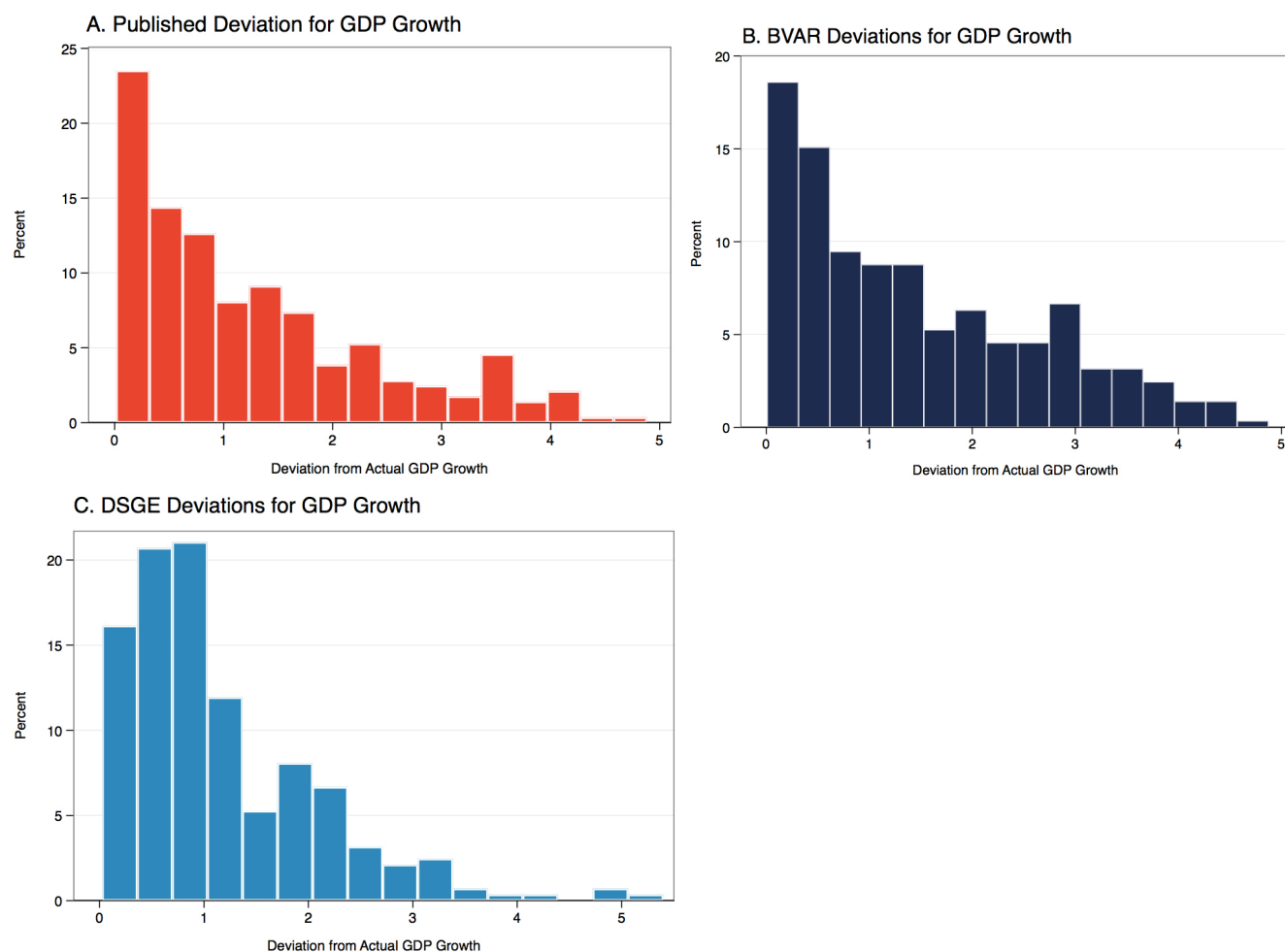


FIGURE 8. DISTRIBUTION OF FORECAST DEVIATIONS FROM ACTUAL GDP GROWTH. Note the figure charts the distributions of the forecast deviations from the actual outcome across all horizons for each forecasting component for GDP growth. The A Charts display the published forecasts, while Charts B and C display the BVAR and DSGE forecasts respectively for the years 2010-2014. The forecasting components are differentiated by color. Data source: Lindé and Reslow (2020).

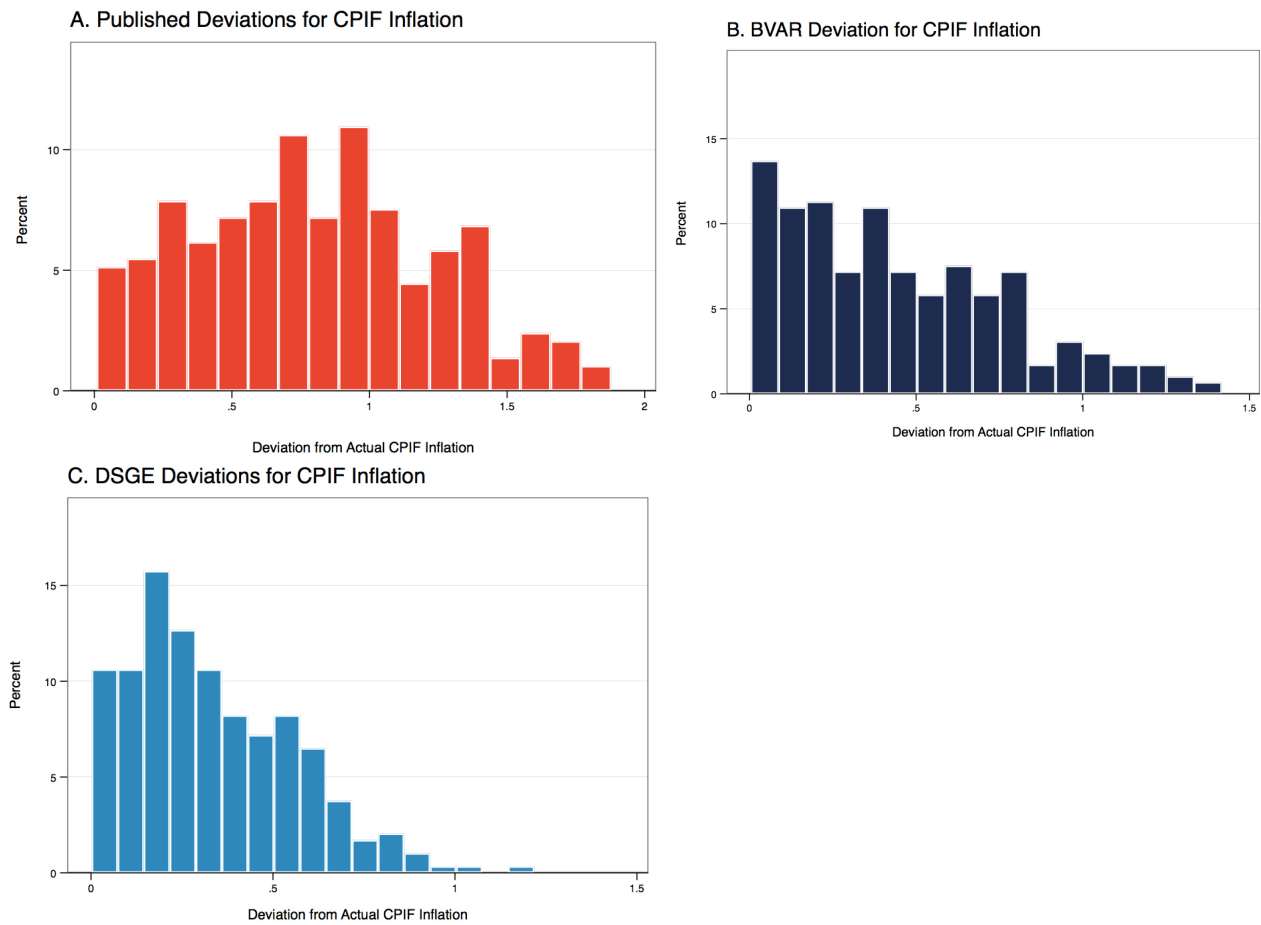


FIGURE 9. DISTRIBUTION OF FORECAST DEVIATIONS FROM ACTUAL CPIF INFLATION.

Note the figure charts the distributions of the forecast deviations from the actual outcome across all horizons for each forecasting component for CPIF inflation. The A Charts display the published forecasts, while Charts B and C display the BVAR and DSGE forecasts respectively for the years 2010-2014. The forecasting components are differentiated by color. Data source: Lindé and Reslow (2020).

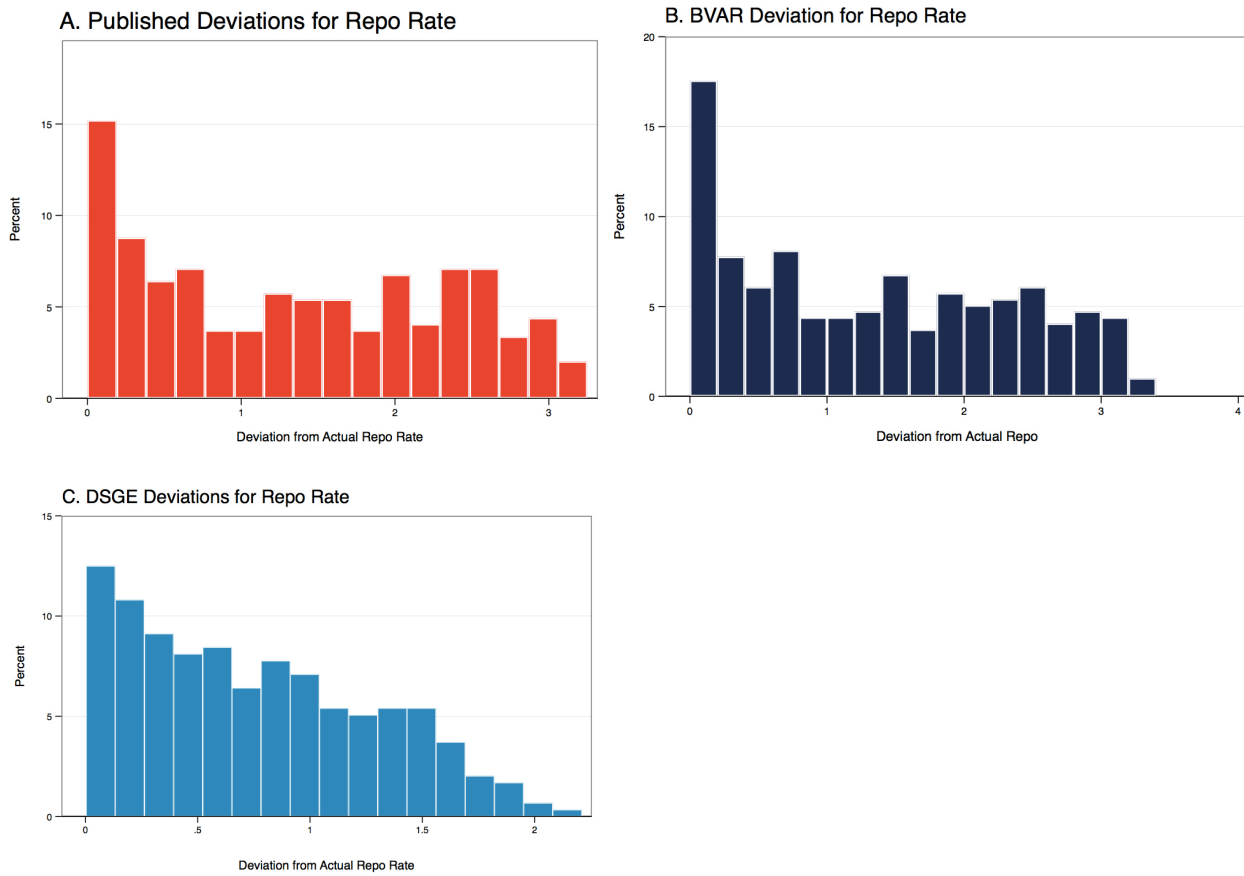


FIGURE 10. DISTRIBUTION OF FORECAST DEVIATIONS FROM ACTUAL REPO RATE.

Note the figure charts the distributions of the forecast deviations from the actual outcome across all horizons for each forecasting component for the repo rate. The A Charts display the published forecasts, while Charts B and C display the BVAR and DSGE forecasts respectively for the years 2010-2014. The forecasting components are differentiated by color. Data source: Lindé and Reslow (2020).

Appendix E

TABLE 7. EFFECT SIZE NOWCASTS GDP GROWTH: HEDGES' G

Forecast	Published	BVAR	DSGE
Published	.	.051	.135
BVAR	.051	.	.177
DSGE	.135	.177	.

Notes: Table 7 reports Hedges' g, an estimate of the effect size for two independent samples using variables. Data source: Lindé and Reslow (2020).

TABLE 8. EFFECT SIZE NOWCAST CPIF INFLATION: HEDGES' G

Forecast	Published	BVAR	DSGE
Published	.	.148	.313
BVAR	.148	.	.190
DSGE	.313	.190	.

Notes: Table 8 reports Hedges' g, an estimate of the effect size for two independent samples using variables. Data source: Lindé and Reslow (2020).

TABLE 9. EFFECT SIZE NOWCAST REPO RATE: HEDGES' G

Forecast	Published	BVAR	DSGE
Published	.	.485	.123
BVAR	.485	.	.645
DSGE	.123	.645	.

Notes: Table 9 reports Hedges' g, an estimate of the effect size for two independent samples using variables. Data source: Lindé and Reslow (2020).

Appendix F

TABLE 10. GDP OPTIMAL WEIGHTING

	(1) Optimal Forecast b	(2) Published Forecast b
DSGE GDP Forecast	1.000	.
Published GDP Forecast	.	1.027**
Constant	.	-.017
RMSE	1.948	2.024

Notes: Table 10 reports coefficients from (1) regression of actual GDP on the fitted values obtained by optimizing model weights. (2) reports the coefficient from the regression of actual GDP on the published forecast. The optimal forecast has been constrained so as not to report a constant. We report root mean squared errors (RMSE). Stars denote statistical significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, 19 df. Data source: Lindé and Reslow (2020).

TABLE 11. CPIF INFLATION OPTIMAL WEIGHTING

	(1) Optimal Forecast b	(2) Published Forecast b
BVAR CPIF Forecast	.425*	.
DSGE CPIF Forecast	.575**	.
Published CPIF Forecast	.	1.059***
Constant	.	-.201
RMSE	.257	.194

Notes: Table 11 reports coefficients from (1) regression of actual CPIF inflation on the fitted values obtained by optimizing model weights. (2) reports the coefficient from the regression of actual CPIF inflation on the published forecast. The optimal forecast has been constrained so as not to report a constant. We report root mean squared errors (RMSE). Stars denote statistical significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, 18 df. Data source: Lindé and Reslow (2020).

TABLE 12. REPO RATE OPTIMAL WEIGHTING

	(1) Optimal Forecast b	(2) Published Forecast b
BVAR Repo Rate Forecast	.285*	.
DSGE Repo Rate Forecast	.715***	.
Published Repo Rate Forecast	.	.908***
Constant	.	.070
RMSE	.211	.138

Notes: Table 12 reports coefficients from (1) regression of actual repo rate on the fitted values obtained by optimizing model weights. (2) reports the coefficient from the regression of actual repo rate on the published forecast. The optimal forecast has been constrained so as not to report a constant. We report root mean squared errors (RMSE). Stars denote statistical significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, 17 df. Data source: Lindé and Reslow (2020).