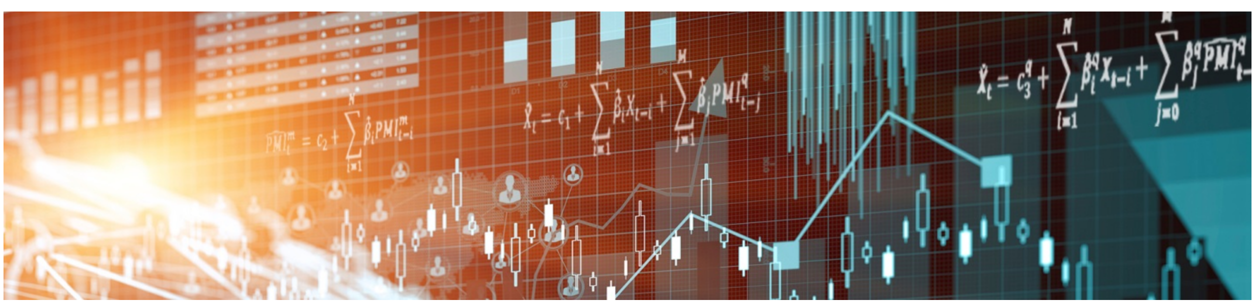
BANK OF CANADA
BANQUE DU CANADA

Staff Analytical Note/Note analytique du personnel 2017-21

Evaluating Real GDP Growth Forecasts in the Bank of Canada *Monetary Policy Report*



by André Binette and Dmitri Tchebotarev

Canadian Economic Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
abinette@bankofcanada.ca
dtchebotarev@bankofcanada.ca

www.bank-banque-canada.ca

Acknowledgements

The authors are grateful to Russell Barnett, Karyne Charbonneau, Harriet Jackson and Eric Santor for helpful comments and suggestions. We thank Alexa Evans for her excellent research assistance and help assembling data related to inflation forecasts. We also thank Meredith Fraser-Ohman for valuable editorial suggestions. All remaining errors are our own.

ISSN 2369-9639

Abstract

This paper examines the quality of projections of real GDP growth taken from the Bank of Canada Monetary Policy Report (MPR) since they were first published in 1997. Over the last decade, it has become common practice among the central banking community to discuss forecast performance publicly. The assessment we undertake is on annual forecasts as well as the average prediction over the policy horizon. We find that the MPR is more accurate than a naïve forecast model and marginally superior to a consensus of professional forecasters. The accuracy of the MPR annual predictions, measured by the root-mean-square prediction error (RMSPE), improves from 1.6 to 0.6 percentage points as more data become available. On a two-year average basis, the RMSPEs are about 1.0 percentage point for forecasts made in April and October. Our results also suggest that the bias present in MPR forecasts is often not statistically significant for both annual and two-year projections. Nonetheless, we found a tendency to overpredict growth at the beginning of the forecast cycle. Finally, at the beginning of the forecast cycle, the MPR correctly predicts the sign of the change in annual real GDP growth roughly 50 per cent of the time, improving to about 75 per cent at the end of the cycle. The sign of change is correctly predicted roughly 60 per cent of the time for the two-year average prediction.

Bank topics: Business fluctuations and cycles; Econometric and statistical methods; Monetary policy
JEL codes: C, C5, C52, E, E3, E32, E37, E5, E52

Résumé

La qualité des projections de la croissance du PIB réel depuis leur première parution dans le Rapport sur la politique monétaire de la Banque du Canada en 1997 est examinée dans cette étude. Au cours des dix dernières années, il est devenu de plus en plus courant pour les banques centrales de discuter publiquement de la performance de leurs projections. Notre évaluation porte sur les projections annuelles et sur le taux moyen des prévisions de croissance pour la période de deux ans qui forme l’horizon de la politique monétaire. Nous constatons que les projections du Rapport sont plus précises que celles d’un modèle naïf, et d’une qualité légèrement supérieure à celle des projections du consensus de prévisionnistes professionnels. La précision des prévisions annuelles du Rapport, qui est mesurée par la racine carrée de l’erreur quadratique moyenne (RCEQM), augmente, passant de 1,6 à 0,6 point de pourcentage, dès lors qu’une plus grande quantité de données est disponible. Pour les taux moyens des prévisions sur deux ans, la RCEQM est d’environ 1,0 point de pourcentage si l’on considère les projections faites en avril et en octobre. Nos résultats montrent que le biais présent dans les projections du Rapport n’est en général pas statistiquement significatif, tant pour les prévisions annuelles que pour les prévisions sur deux ans. Nous découvrons néanmoins une tendance à surestimer les taux de croissance au début du cycle de projection. Enfin, au début du cycle de projection, le Rapport prévoit correctement l’orientation de la croissance annuelle du PIB réel pratiquement 5 fois sur 10; en fin de cycle, cette proportion s’élève à 75 %. Quant au taux moyen des prévisions sur deux ans, le Rapport prévoit correctement l’orientation de la croissance environ 6 fois sur 10.

Sujets : Cycles et fluctuations économiques; Méthodes économétriques et statistiques; Politique monétaire
Codes JEL : C, C5, C52, E, E3, E32, E37, E5, E52

1 Introduction

The Bank of Canada began publishing the *Monetary Policy Report* (MPR) in 1995, and since 1997, it has included numerical predictions[1] of real gross domestic product (GDP) growth for the current and following year. The year 2017 marks the 20th anniversary of these projections, and as such, provides enough data to reflect on their performance. Over the last decade, it has become common practice among the central banking community to discuss forecast performance publicly, including assessments by the Riksbank, the Federal Reserve, the Bank of England and the Reserve Bank of New Zealand.[2,3] A clear assessment of forecast accuracy is desirable since the forecasts represent a key input to the monetary policy decision. Forecasts are also a useful communications tool, since they convey considerable information about the central bank's views on the economy. The forecast is, indeed, a focal point of many observers (e.g., markets and firms), and poor forecast performance could affect the credibility of a central bank. Worse, poor forecasting could lead to policy mistakes that would have ramifications for economic outcomes and impair the central bank's ability to achieve its monetary policy objectives. That said, most central bankers would agree that any policy decision relies not only on these forecasts but also on a broader assessment of the risks surrounding the outlook. At the Bank of Canada, this is referred to as the risk management framework.[4]

The economic outlook presented in the MPR, which includes the forecast and narrative, reflects many weeks of work and dialogue.[5] Four times a year, starting with the publication by Statistics Canada of the latest quarterly National Economic Accounts data, staff begin a thorough process to conduct the best possible analysis of all developments since the previous MPR. This process engages many departments in the institution and culminates with the presentation of the staff view to Governing Council. The staff view becomes the foundation of the economic outlook later published in the MPR. The staff view is based on state-of-the-art models, substantial analyses of key issues, and consultations with businesses and associations across Canada. In this regard, the *Business Outlook Survey* provides crucial insights into current economic conditions and expected developments. With the staff view in mind, Governing Council deliberates and discusses the risks to the outlook for inflation. Following these deliberations, Governing Council's consensus view of the outlook for growth and inflation and the main risks are published in the MPR. It is worth noting that there is no single "MPR model," like ToTEM or LENS.[6] The MPR outlook is the culmination of a complex process where models and informed judgment play a vital role in determining the final projection.

In this paper, we look back on 20 years of real GDP forecasts published in the MPR to assess how well these forecasts match the realized data and how their performance compares with that of a naïve forecast model and a survey of external forecasters. When compared against these two benchmarks, we find that the MPR is more accurate than a naïve model that uses a 10-year rolling mean as its forecast and marginally superior to a consensus of professional forecasters. This result holds for annual predictions, two-year average projections and most forecast horizons. The accuracy of the MPR annual predictions, measured by the root-mean-square prediction error (RMSPE), improves from 1.6 to 0.6 percentage points (pp) as more data become available. On a two-year average basis, the RMSPEs are about 1.0 pp for both forecast horizons. Our results also suggest that the bias present in MPR forecasts is often not statistically significant for either annual or two-year projections. Indeed, according to the Mincer-Zarnowitz test, MPR predictions are deemed to be optimal. Nonetheless, we found a significant tendency to overpredict growth at the beginning of the forecast cycle. Finally, we have looked at the ability of the MPR to correctly predict the sign of change in

1. In this note, we use prediction, forecast and projection interchangeably.
2. See, for example, Vredin et al. (2017), Chang and Hanson (2015), Habgood (2015) and Lees (2016).
3. Appendix A presents a list of some forecast evaluations at central banks and other institutions.
4. For a discussion of the Bank of Canada risk management framework see Poloz (2014).
5. For an extensive discussion about monetary policy decision making at the Bank of Canada see Macklem (2002) and Murray (2013).
6. ToTEM and LENS are currently two of the most important macroeconomic models at the Bank of Canada. For a description of each model see Dorich et al. (2013) and Gervais and Gosselin (2014).

real GDP growth. At the beginning of the forecast cycle, the MPR correctly predicts a pickup or slowdown in annual real GDP growth roughly 50 per cent of the time, improving to about 75 per cent at the end of the cycle. The sign of change is correctly predicted roughly 60 per cent of the time for the two-year average prediction.

The rest of this note is organized as follows: Section 2 discusses the data we use for the different MPR projections followed by a discussion of the relevant GDP series from Statistics Canada. The section also includes a brief presentation of the Consensus Economics data used to evaluate the MPR forecast performance. Section 3 presents a number of performance metrics for several forecast horizons. In this section, we evaluate specific annual predictions as well as the average prediction over the projection horizon. Section 4 compares the forecast accuracy of the MPR with two benchmarks: a naïve forecast (the ex-ante unconditional mean) and private sector consensus. Finally, Section 5 summarizes the main results.

2 Data

2.1 MPR predictions

The Bank of Canada has published the MPR two to four times a year since 1995, to inform the Canadian public about its view of the economy and the implications for inflation. As of 2009, the Bank of Canada publishes four complete MPRs per year in January, April, July and October.[7]

From these reports, we have extracted the real GDP predictions. From 2004 onward this is straightforward, as the predictions are explicitly presented in a table. However, for MPRs published before 2004, the numerical figures must be recovered from the text of the MPR. For example, in May of 2000, the MPR read, "Taking into account the upward revision to growth in 1999 as well as recent data, the Bank projects output growth in Canada in 2000 of about 4 to 4.5 per cent." We have entered this in our data set as a growth rate of 4.25 per cent. Before 1997, even this information is not readily available, so our analysis starts in 1997.

We evaluate only real GDP predictions from the MPR publications because inferring the inflation forecast is impractical before 2004.[8] In the same May 2000 MPR as above, the inflation projection is described as follows: "On balance, with some demand pressures emerging this year, the Bank sees core inflation rising gradually towards the midpoint of the inflation-control range." From this, it is not clear what the projected inflation numbers were for 2000 and 2001.

Because of the changes in timing of the publication of the MPR, and limitations on which information was reported consistently through time, we restrict our data set in several ways. In addition to exclusively focusing on real GDP predictions, we assess only predictions made in April and October, because they are the most consistent MPR publication dates.[9] Similarly, we solely evaluate annual forecasts since comprehensive quarterly predictions became available only in October 2005. Moreover, in recent years, the quarterly predictions were limited to the first two quarters of the projection and not the full forecast horizon. Lastly, we assess forecasts two years ahead (current year (t) and next (t+1)) because the three-year-ahead projections were not published before 2005. In this way, we ensure the longest possible sample for analysis.

We end up with four predictions for each year. Table 1 shows the timing of the MPR predictions (forecast cycle). For example, the first real GDP growth forecast for 2016 is published in the April 2015 MPR. We

7. In 1995, the MPR was released in May and November of each year. In 2000, the Bank added short updates in February and August. Two years later, the publication schedule changed and detailed MPRs were issued in April and October while updates were released in January and July.

8. Although there is not enough data to conduct a complete assessment of the MPR inflation forecasts, we present the annual RMSPEs in Appendix B. We use the same data as for the forecast confidence intervals published in the MPR, but on an annual basis.

9. Before 2002, the Bank of Canada published MPRs in May and November instead of April and October, but we still label them as if they were circulated in April and October respectively for ease of discussion/analysis.

refer to it as the first/longest forecast horizon (FH1). In October 2015, the second prediction for 2016 growth is made, referred to as FH2. The last two annual forecasts for 2016 growth are made in April and October of the same year (2016) and are called FH3 and FH4, respectively. FH4 represents the last/shortest forecast horizon for a given target year. FH1 is the forecast made with no quarterly information about the target year while FH4 exploits much more data, including quarterly real GDP figures for five of the seven quarters relevant for calculating the annual growth rate of the target year.[10] At horizon FH4, we have more than 80 per cent of the necessary information to compute the annual figure of the target year.

We undertake the forecast assessment for two-year average growth rates as well as annual. The reasoning is twofold. First, monetary policy takes about two years to fully propagate through the economy. Second, we believe that quarterly or annual volatility in the growth rate of real GDP is less relevant for the conduct of monetary policy. In our view, a central bank should focus more on the expected level of real GDP at the end of the relevant policy horizon and less on the volatility of the intermediate point forecasts. We construct the two-year series as the geometric average of the current and following year. The two-year growth rate g of g_1 (current year growth) and g_2 (next year growth) is computed as $g = \sqrt{(1+g_1)(1+g_2)} - 1$. Since we need the data from the second year to know the realized value, we label the two-year average growth rate using the second year, e.g., 2016 represents the average real GDP growth for 2015 and 2016. The projections are made in April and October of the first year and labelled APRIL and OCTOBER, respectively. For the April two-year prediction, less than 20 per cent of the quarterly data is available, and about 44 per cent is known for the October forecast. In all, we have around 20 data points per forecast horizon in our data set.

Table 1: Timing of the MPR predictions—forecast horizon/cycle

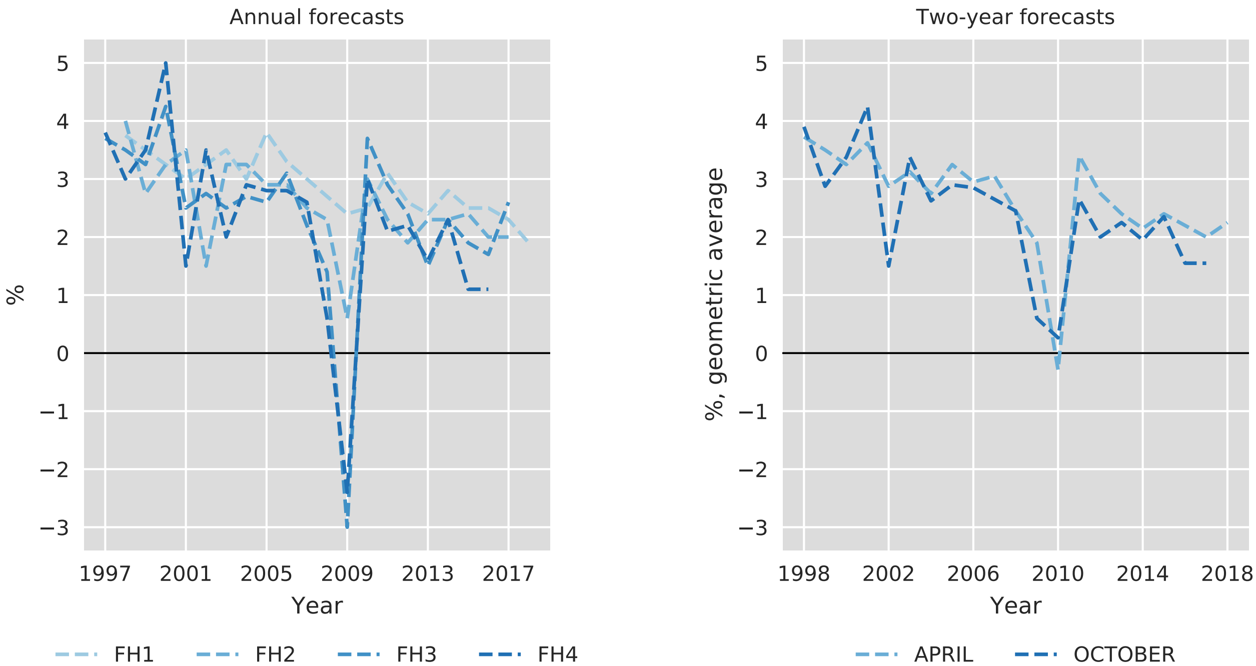
MPR	Target Year: 1997	1998		…	2015		2016	
Apr '97	FH3	FH1	APRIL					
Oct '97	FH4	FH2	OCTOBER					
Apr '98		FH3		…				
Oct '98		FH4						
		⋮					⋮	
Apr '14					FH1	APRIL		
Oct '14					FH2	OCTOBER		
Apr '15					FH3		FH1	APRIL
Oct '15				…	FH4		FH2	OCTOBER
Apr '16							FH3	
Oct '16							FH4	

Note: FH1, FH2, FH3 and FH4 represent annual forecasts while APRIL and OCTOBER indicate two-year predictions.

Figure 1 below presents the real GDP forecast (annual and two-year average) as published in the April and October MPR since 1997. We observe that the MPR outlook for real GDP appears to be downgraded often, as more information becomes available—roughly two-thirds of the time. It echoes the results of a recent analysis by Guénette et al. (2016) about serial disappointment in the Bank of Canada forecasts. This phenomenon is observed throughout the sample.

10. See Cross and Wyman (2011) for a detailed explanation of the relationship between quarterly and annual growth rates.

Figure 1: What do the MPR forecasts for real GDP growth look like?



2.2 Real GDP series

To calculate the prediction errors used to assess the MPR performance, we need an annual series for realized GDP. We use the final real GDP series as opposed to the real-time estimate because the former makes use of the most recent information, and represents a complete measure of real economic activity that should be more relevant to the conduct of monetary policy.[11] Thus we use the latest release of GDP at market prices in chained 2007 dollars.[12] The quarterly series is transformed into an annual series by averaging the level.

The conceptual definition of the real GDP series has changed over the period of our analysis. Following two papers by the Federal Reserve,[13] we have adjusted the real GDP series for three main methodological changes made by Statistics Canada. First, in 2001, Statistics Canada modified the treatment of software.[14] Second, they switched from a Laspeyres to a chain Fisher volume index.[15] Finally, in 2012, Statistics Canada broadened the capital boundary in the Canadian System of National Accounts (CSNA) to include (R&D) and military weapons.[16] It is important to adjust the GDP series for these conceptual changes since forecasters, like the Bank of Canada, were predicting output growth based on existing methodology. Adjusting the GDP series eliminates distortions in the forecast errors.

Regarding the change in the treatment of software, Statistics Canada estimates the impact on the growth rate of real GDP to be quite small—less than 0.1 pp over 1997 to 2000.[17] Unfortunately, Statistic Canada's

11. Research at the Board of Governors of the Federal Reserve System also used the latest vintage in their forecast assessment, as the resulting errors should provide a better measure of uncertainty—Nalewaik and Scherling (2014).

12. Series v62305752 published by Statistics Canada.

13. See Reifschneider and Tulip (2007) and Nalewaik and Scherling (2014).

14. Software is now treated like any other capital input. Previously it was assumed to be fully depreciated during the production period like intermediate inputs.

15. There are two reasons mentioned by Statistics Canada for this: first, it provides users with a more rigorous measure of real GDP growth between two consecutive periods and second, to make the Canadian measure comparable with the United States. For more detail see Chevalier (2003).

16. See Tebrake et al. (2011).

17. For a detailed discussion of the treatment of software spending in the CSNA see Jackson (2002).

study does not provide an estimate for 2001 and 2002. To proxy the effect for these two years, we assume the average impact over 1997 to 2000. Overall, the impact on growth is small, as much of the software spending was already included in GDP under hardware and government expenditures.[18] For the switch to the chain Fisher volume index, the average difference between the growth rates of the two indices is about 0.1 pp before 2002—Laspeyres growth has been, on average, higher. That said, this change has by far the largest impact on real GDP growth with a maximum effect of 0.4 pp in 1999. Finally, we did not find any formal analyses concerning the repercussions of the capitalization of research and development (R&D) and military weapons on the growth rate of real GDP in Canada for each year between 1997 and 2012. An analysis from the Federal Reserve Board of Governors indicated that accounting for the capitalization of R&D in the evaluation of US GDP forecasts resulted in only very minor gains. The Federal Reserve Board of Governors has proposed to simply remove the contribution to growth of R&D from real GDP. We concur with this methodology and remove the contribution of both R&D and military weapons from GDP growth in our analysis. The average impact, over the period considered, is close to zero and the maximum effect on growth is 0.2 pp in 2000 due to a surge in R&D expenditures of more than 50 per cent.

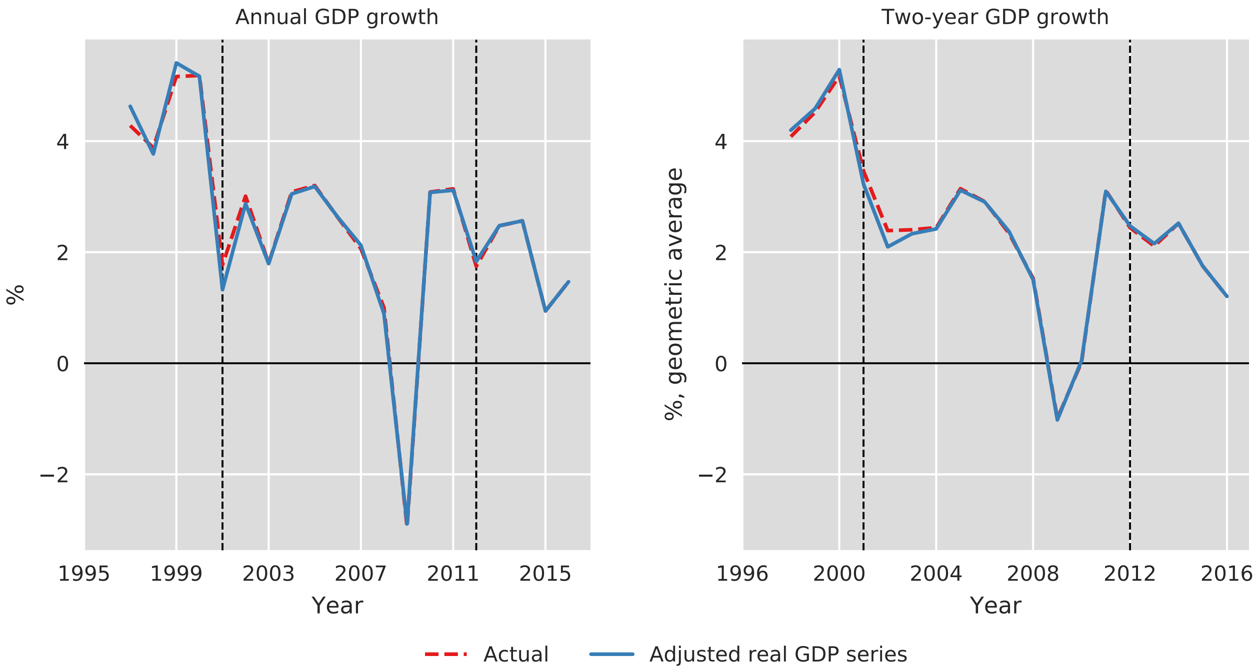
Figure 2 shows the latest estimate of real GDP growth (annual and two-year average) along with an adjusted version to correct for the conceptual changes through time. While these changes have impacted the average annual growth of real GDP, the overall effect appears relatively small and concentrated in the early part of the sample. As discussed above, the switch to the chain Fisher volume index seems to explain most of the divergence observed before 2002. Appendix C presents the estimated annual impact on real GDP growth for each change. In the end, we have created slightly different adjusted series (one for each forecast horizon—FH1 to FH4 and APRIL and OCTOBER) to match the timing of the methodological changes to that of the projections. For example, in April 2001 when the bank first predicted annual real GDP growth for 2002 (FH1), it was not aware of the upcoming changes in the measurement of GDP. However, by October 2001 (FH2), the first two changes (Laspeyres and Software) were fully accounted for in the Bank of Canada forecast. Thus, to assess the forecast performance for 2002, we adjust the GDP series when evaluating FH1, but not FH2, FH3 or FH4.

2.3 Consensus forecast

To assess the MPR projections of real GDP growth, we require benchmarks against which to do so. Typically, the performance is first compared with a naïve forecast model that predicts only the unconditional mean of the series. However, a more convincing test is to compare the MPR with a survey of professional forecasters, which we obtain from Consensus Economics (CE). CE has been surveying between 12 and 19 forecasters on the Canadian outlook each month since 1989. The fact that they collect the information in the first half of the month ensures the independence of the predictions—there is no risk of the professional forecasters reacting to the publication of the MPR. To match our forecast series, we extract the predictions from the CE publications and label the May and November forecasts before 2001 as April and October.

18. The price differential also helps explain the limited influence on real GDP growth—hardware prices have been decreasing much faster than prices for software.

Figure 2: Conceptual changes have little impact on real GDP growth after 2001



Note: Dotted vertical lines mark the introduction of the different conceptual changes to real GDP. The first dotted line is at 2001, corresponding to the change from a Laspeyres to a chain Fisher volume index and the modifications to the treatment of software. The second dotted line is at 2012, marking the shift in the treatment of military weapons and R&D spending.

3 MPR Forecast Performance

There is a large set of metrics available to evaluate point estimates.[19] In this paper, as with many other forecast assessments, we focus on unbiasedness and accuracy.[20] Additionally, we assess the directional forecast: can the MPR outlook predict an increase or a decrease in the growth rate of real GDP? That said, on their own, these metrics provide insufficient information about the overall performance of the MPR projections. We therefore examine the results of the widely used Mincer-Zarnowitz regression, which gives a sense of the quality of the MPR predictions by jointly evaluating bias and efficiency.

3.1 Unbiasedness: pessimism, optimism or realism?

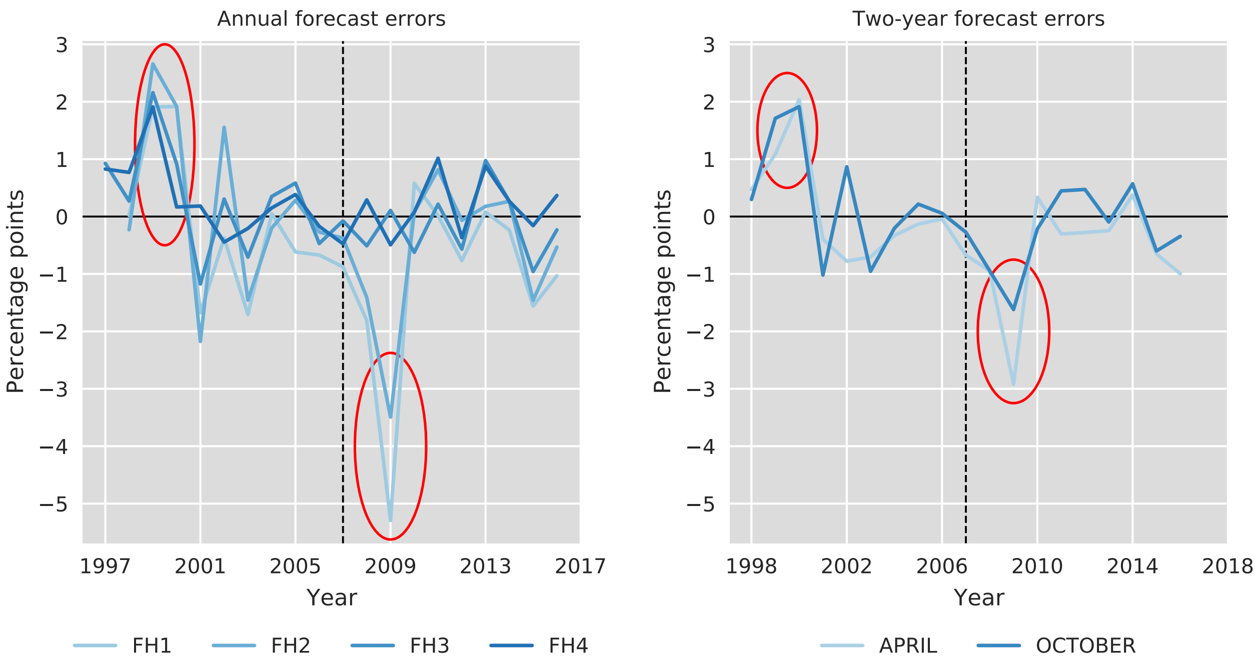
A desirable characteristic of any forecast is to be unbiased, meaning that the forecaster makes positive errors as often as negative errors and the magnitudes of these errors are comparable in size. A cursory glance at Figure 1 shows that the outlook for real GDP growth at all horizons tends to be revised downward in the MPR. Moreover, Figure 3 shows that the MPR errors appear biased.[21] For example, it seems that the Bank tends to overpredict growth early on for both the annual and the two-year average forecasts (FH1, FH2 and

19. See for example Mincer and Zarnowitz (1969) and Clark and McCracken (2013) for an overview.
20. See Appendix A for examples.
21. In this paper, forecast errors represent the difference between actual real GDP growth and the forecasted figure: $e_t^{FH_i} = Y_t - Y_t^{FH_i}$.

APRIL) while underpredicting growth late in the forecast cycle on an annual basis (FH3 and FH4). These observations could be driven by the Asian crisis and the Great Recession, circled in red in Figure 3.

A regression of the forecast errors on a constant (Table 2) suggests that the bias is not statistically significant for most forecast horizons and subsamples, with a few exceptions.[22] For FH1 and APRIL, the Bank tends to overpredict growth. Splitting the sample in two indicates that this may be caused by the Great Recession in 2009. For FH4, the bank tends to underpredict growth. Again, splitting the sample indicates that this is mostly due to the increase in growth after the Asian crisis in 1999–2000. Abstracting from these events, however, suggests that the MPR forecasts of real GDP growth are unbiased for most forecast horizons and subsamples, except for the initial predictions (FH1 and APRIL).

Figure 3: MPR prediction errors for all forecast horizons



Note: Red circles indicate the Asian crisis and Great Recession. The dotted vertical line indicates the split of the sample in half (2007).

To summarize, while the MPR predictions of real GDP real growth tend to be downgraded through time, they are statistically unbiased for most forecast horizons and subsamples. Nonetheless, we found a significant tendency to overpredict growth initially, especially after correcting for the Asian crisis and Great Recession. The initial optimism bias over the full sample amounts to 0.7 pp on an annual basis (FH1). For the two-year average, the initial overprediction remains important at 0.3 pp (APRIL).

3.2 Accuracy: How big are the MPR prediction errors?

The choice of the accuracy measure is theoretically associated with a particular loss function, which is in turn linked to a measure of central tendency. Typically, the policy-maker assumes a quadratic loss function,

22. To test for unbiasedness, we used the following equation: $e_t^{FH_i} = \beta_0 + \mu_t$. Only if β_0 is statistically different from zero is the forecast considered biased. This equation is estimated by ordinary least squares (OLS), and the covariance matrix is heteroskedasticity and autocorrelation consistent (HAC) based on West and Newey (1987).

Table 2: In general we cannot reject the hypothesis that MPR forecasts are unbiased, except for the initial predictions

Forecast Horizon	FH1	FH2	FH3	FH4	APRIL	OCTOBER
Full Sample	-0.63^{*}	-0.21	0.09	0.25^{*}	-0.27	0.01
	$(-0.66)^{***}$	$(-0.31)^{*}$	(-0.07)	(0.16)	$(-0.34)^{***}$	(-0.12)
1997–2006	-0.13	0.23	0.32	0.36	0.13	0.30
	$(-0.71)^{***}$	(-0.36)	(0.01)	(0.19)	(-0.29)	(-0.13)
2007–2016	-1.09^{**}	-0.60	-0.14	0.14	-0.63^{**}	-0.26
	$(-0.62)^{***}$	(-0.28)	(-0.14)	(0.14)	$(-0.38)^{***}$	(-0.11)

Note: Numbers in parentheses indicate the bias when the following errors are removed: 1999 and 2000 (Asian crisis) for all forecast horizons, and 2009 (Great Recession) for FH1, FH2, APRIL and OCTOBER.

*, **, *** indicate forecast errors that are statistically different from zero at 10%, 5%, and 1% significance levels, respectively. The HAC adjustment is based on West and Newey (1987).

since it is more costly to make one large error than many small mistakes—significant deviations from the expected path of key macroeconomic variables are more likely to lead to policy mistakes. Consequently, the optimal point forecast is the mean.[23] In this note, we avoid any discussion about the definition of the Bank of Canada's point estimates, whether it is the mean, median or mode. We choose to focus on the quadratic loss function—the RMSPE as defined by Equation 1, even if it could be suboptimal if the MPR point forecast is in reality the mode or the median.

$$RMSPE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^{actual} - Y_t^{predicted})^2} \tag{1}$$

The RMSPE is mathematically related to another desirable characteristic of any forecast: unbiasedness. Equation 2 below presents the relationship between the two.

$$RMSPE = \sqrt{\text{bias}^2 + \text{prediction error variance}} \tag{2}$$

In the case of unbiasedness, the RMSPE is equivalent to the standard deviation of the forecast errors. Thus, avoiding systemic prediction error has two benefits—lowering the bias directly, and indirectly reducing the RMSPE. Table 3 below presents the accuracy of the MPR outlook for the annual and two-year projections at different forecast horizons, along with the impact of the bias.

A few observations are worth highlighting. First, as expected, the performance of the MPR in predicting real GDP growth improves as more information becomes available. As mentioned earlier, this could largely be explained by the fact that we know more than 80 per cent of the necessary information to compute the annual figure at horizon FH4. Of course, this assumes no revision to the quarterly figures.[24] In contrast, the performance of the two-year projection is relatively consistent across the MPRs over the last 20 years with

23. For more detail on this see Wallis (1999) or Gneiting (2011).

24. Quarterly real GDP revisions negatively impact the performance for FH4. Without revisions, the RMSPE would be lower by about 0.3 pp over 1997–2016. This estimate compares the performance of a naïve model using actual and real-time real GDP data. The naïve prediction is simply the average quarterly growth rate over the past 10 years.

Table 3: Size of the MPR forecast errors and impact of the bias

	Annual				Two-year	
	FH1	FH2	FH3	FH4	APRIL	OCTOBER
Full sample	1.63	1.41	0.78	0.64	1.00	0.86
	(1.50)	(1.39)	(0.77)	(0.59)	(0.96)	(0.87)
1997–2006	1.25	1.50	0.95	0.74	0.89	1.01
	(1.24)	(1.48)	(0.90)	(0.64)	(0.87)	(0.97)
2007–2016	1.91	1.32	0.55	0.52	1.08	0.70
	(1.57)	(1.18)	(0.53)	(0.50)	(0.88)	(0.65)

Note: Table shows RMSPEs along with the standard deviation of the MPR forecast errors in parentheses. The relation between RMSPE and the standard deviation is given in Equation 2.

only a marginal improvement between April and October. That said, we observe a pronounced improvement in forecast performance over the last decade.

Second, the performance of the MPR seems to have improved in the last decade for both the annual and the two-year average forecasts, except at the longest horizon (FH1 and APRIL) where the impact of the Great Recession is probably the biggest. In April 2008, the Bank of Canada, and many other forecasters, did not anticipate the meltdown of real economic activity in 2009.

Third, to add to our discussion of the bias (largely insignificant at most horizons), its impact on the accuracy of the predictions is small, except at horizon FH1 and APRIL over the last 10 years where the gap amounts to 0.2 and 0.3 pp for the two-year and the annual predictions, respectively. Otherwise, the difference between the RMSPE and the standard deviation of the forecast errors is less than 0.1 pp.

Finally, the calculated RMSPE suggests that a simple confidence interval (one standard deviation or 68 per cent) indicates that for any given year, the actual outcome of real GDP growth could be anything in a range of ±1.6 pp at the beginning of the forecast cycle, dropping to ±0.6 pp at the end of the forecast cycle. This suggests that uncertainty around the annual point forecast is non-negligible. For comparison, another frequently used measure of uncertainty, the dispersion of private sector forecasts,[25] suggests intervals that are much smaller—0.9 pp and 0.3 pp for FH1 and FH4, respectively.[26] On a two-year average basis, the calculated RMSPEs suggest a still–considerable confidence interval with a range of about ±1.0 pp. Again, the dispersion of private sector forecasts suggests an interval that is much smaller at about 0.6 pp.

3.3 How good are the MPR projections?

So far, we have discussed the bias and the accuracy of the MPR projections, but it is difficult to say whether the MPR produces good predictions overall. A typical test for good predictions is to estimate Equation 3.

$$Y_t = \beta_0 + \beta_1 Y_t^{FH_i} + \epsilon_t \tag{3}$$

This equation is referred to as the Mincer-Zarnowitz regression. It regresses realized values on a constant and the predictions (Y^{FH_i}). If the intercept is far from zero, the forecast is, on average, biased. If the slope

25. For a discussion on the use of private sector forecast dispersion as a measure of uncertainty see D'Amico and Orphanides (2008) and Lahiri and Sheng (2010).

26. The dispersion of private sector forecasts—not to be confused with the RMSPE—is calculated by eliminating the most extreme value (high and low) and taking the difference between the remaining highest and lowest forecast (see Bank of Canada 2013).

(β_1) is different from one, the forecast has consistently underpredicted or overpredicted deviations from the mean (the predictions are deemed to be inefficient). Finally, if the R^2 is low, then little of the variation of the variable is captured by the forecast. These three elements together represent a more complete test of the overall forecast performance.

A simple way to visualize the Mincer-Zarnowitz equation is to produce a scatter plot of the realized versus the predicted value (prediction-realization diagram).[27] A good forecast will lie on the 45-degree line. Figure 4 below shows the scatter plots for all horizons. The scatter plots suggest that MPR forecasts at the longest horizon (FH1, FH2 and APRIL) tend to be biased and consistently underpredict deviation from the mean. At the shortest horizons, the overall performance of the MPR annual predictions improves significantly, reflecting increased availability of data. For the two-year average, the performance is similar between APRIL and OCTOBER.

After estimating Equation 3, a few results stand out for both the annual and the two-year average forecasts (Table 4). First, the bias gets closer to zero as we progress in the forecast cycle.[28] Similarly, the slope coefficient gets closer to one with more information. Furthermore, the R^2 of the regression increases significantly across forecast horizons. For the annual projection, the R^2 moves from a low of 0.34 to 0.90 at horizon FH4. For the two-year forecast, the R^2 increases from 0.56 in APRIL to 0.65 in OCTOBER. These findings confirm the visual inspection of Figure 4—the MPR performance is good and improves with more information. Finally, we can jointly test the null hypothesis of a zero constant and a slope coefficient of one. This is often referred to as an optimality test—do the forecasts use all available information? Statistically, we cannot reject (at the 10 per cent level) the hypothesis that jointly $\beta_0 = 0$ and $\beta_1 = 1$. The optimality result is surprising for FH1 but could be explained by the large imprecision in the estimates. Overall, the MPR predictions are deemed to be optimal according to this test.

Table 4: Results of the Mincer-Zarnowitz regression (1997–2016)

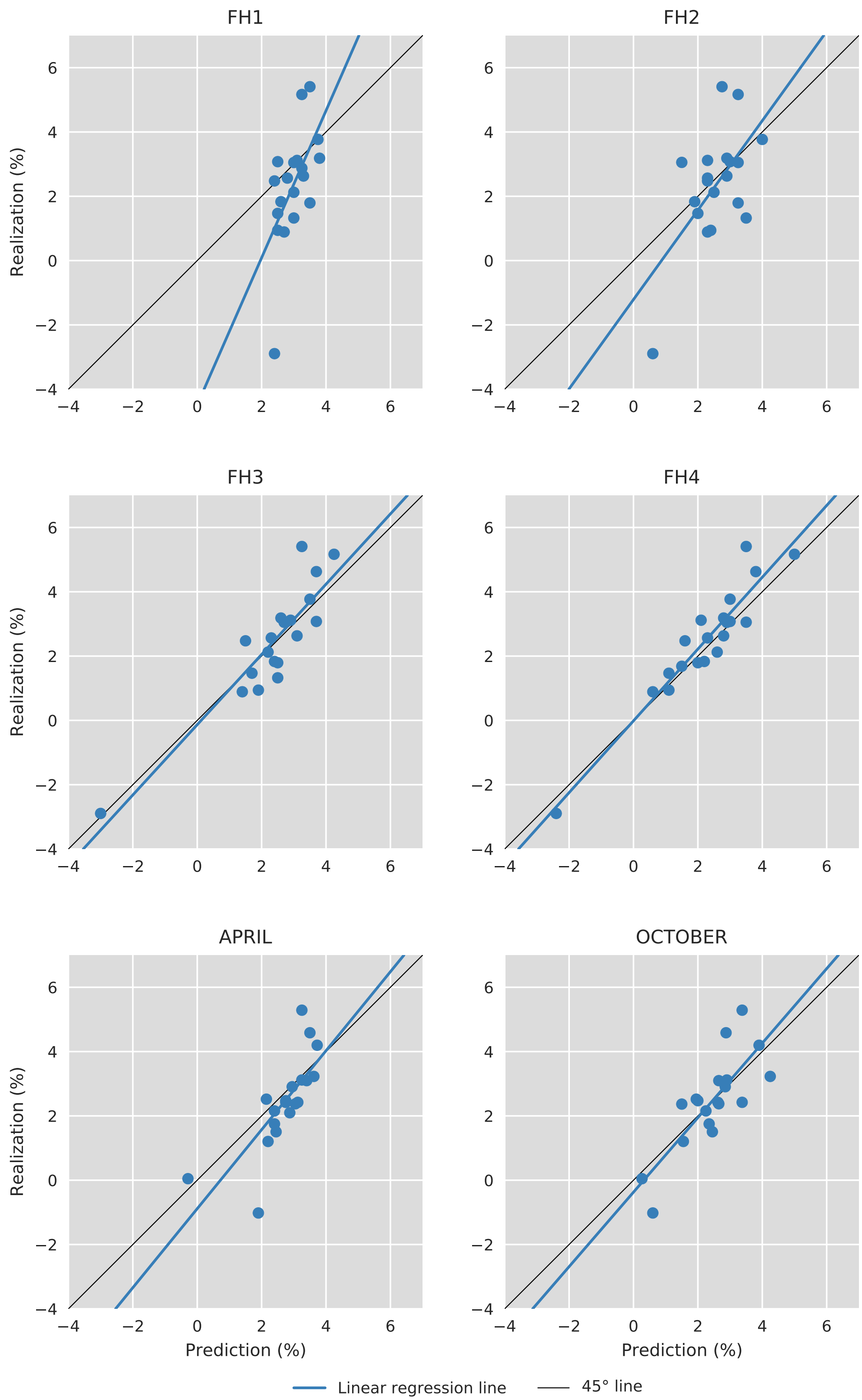
	Annual				Two-year	
Forecast Horizon	FH1	FH2	FH3	FH4	APRIL	OCTOBER
Constant	−4.49	−1.21	−0.13	−0.01	−0.88	−0.40
	(2.62)	(1.82)	(0.23)	(0.08)	(0.70)	(0.59)
Forecast	2.29	1.39	1.09	1.12	1.23	1.17
	(0.85)	(0.66)	(0.12)	(0.06)	(0.27)	(0.23)
R-Squared	0.34	0.37	0.81	0.90	0.56	0.65
Joint F-Test	2.61	0.32	0.32	1.84	1.40	0.27

Note: F-Test of the joint hypothesis of $\beta_{Constant} = 0$ and $\beta_{Forecast} = 1$ is not rejected at the 10% significance level. Estimated by OLS, and the covariance matrix is adjusted for the presence of heteroscedasticity and serial correlation (West and Newey 1987). Values in parentheses indicated HAC adjusted standard errors.

27. This diagram was first introduced by Theil (1961).

28. The constant in these regressions differs from the bias estimate presented above since the slope coefficient is not constrained to one.

Figure 4: MPR prediction-realization diagram for all forecast horizons



3.4 Predicting downturn and upturn—"hit ratio"

While forecast precision is necessary, the direction of a prediction is also crucial. Predicting turning points is one of the most arduous tasks in economic forecasting, but if done successfully the benefits could be important since monetary policy actions only gradually impact real economic activity. The "hit ratio" indicates how often a forecaster correctly predicts an increase or a decrease in the growth rate of a series.[29] Table 5 shows the ability of the MPR to predict shifts in the rate of change of real GDP growth for the different forecast horizons. The results illustrate how challenging it is to correctly predict the change in the dynamics of real GDP growth. Early on, the MPR barely meets the 50 per cent bar for the annual predictions and improves to about 75 per cent at the shortest forecast horizon (FH4). Similarly, predicting downturns and upturns in the two-year projection is difficult, with a success rate of roughly 60 per cent.

Table 5: "Hit ratio" for all forecast horizons

Forecast Horizon	FH1	FH2	FH3	FH4	APRIL	OCTOBER
Hit Ratio	47%	58%	74%	74%	59%	65%

4 Forecast Comparison

An assessment of the MPR performance would be incomplete without a comparison with common benchmarks. As a first test, we compare the MPR projection with a naïve model—the ex-ante unconditional mean of the series. The unconditional mean is computed using a 10-year rolling window based on real-time data.[30] This comparison tells us if the MPR outlook for real GDP is informative. When the ratio of $\frac{RMSPE_{MPR}}{RMSPE_{NAIVE}}$ is below one, then the MPR projection of real GDP growth provides additional information beyond the rolling mean of the series. When the ratio is above one, the opposite is true. For the annual forecast, we can also determine whether the MPR projection is informative not only for the first year but also for the second (referred to as forecast memory).[31] The forecast memory indicates the horizon at which a forecaster provides useful signals.

The second benchmark is much more difficult to meet—are the MPR predictions more accurate than the average view of more than 12 forecasters ($\frac{RMSPE_{MPR}}{RMSPE_{CE}}$)? The tendency for group forecasts to surpass any individual forecast has been observed for over a century.[32] In empirical studies, forecast combinations have frequently been found to generate better predictions on average than approaches based on the ex-ante best individual forecasting model.[33] This phenomenon is known as the wisdom of crowds.

Figures 5 and 6 show the results for the naïve benchmark and the comparison with consensus. Initially, the information content of the MPR projections is about 20 per cent and increases over the first three forecast horizons to reach about 50 per cent. By the end of the forecast horizon, the information content is more limited, as most of the relevant data are now known. That said, the MPR outlook retains a slight edge with an RMSPE of 0.6 compared with 0.7 for the naïve model. The results also indicate the value of the MPR outlook after the first year of the projection (forecast memory of at least two years—the RMSPE ratios are below one for horizon FH1 and FH2). For the two-year projection, the information content amounts to roughly one-third with only a small difference between APRIL and OCTOBER. For the comparison with

29. See Chang and Hanson (2015) or Timmermann (2006a) for further discussion about directional accuracy.
30. The real-time data reflect the initial release of real GDP as recorded by the Bank.
31. The Riksbank discussed this concept in its forecasting assessment of 2007—Andersson, Karlsson and Svensson (2007).
32. See Galton (1907).
33. See Stock and Watson (2004) and Timmermann (2006b) for a thorough discussion. Granziera, Luu and St-Amant (2013) found similar results in the context of Canadian real GDP.

consensus, the results illustrate a small gain on average for the MPR, which is quite good considering the literature on forecast combination. Only at the shortest horizon (FH4) does consensus seem to marginally outperform the MPR outlook.

Figure 5: MPR outlook outperforms a naïve model

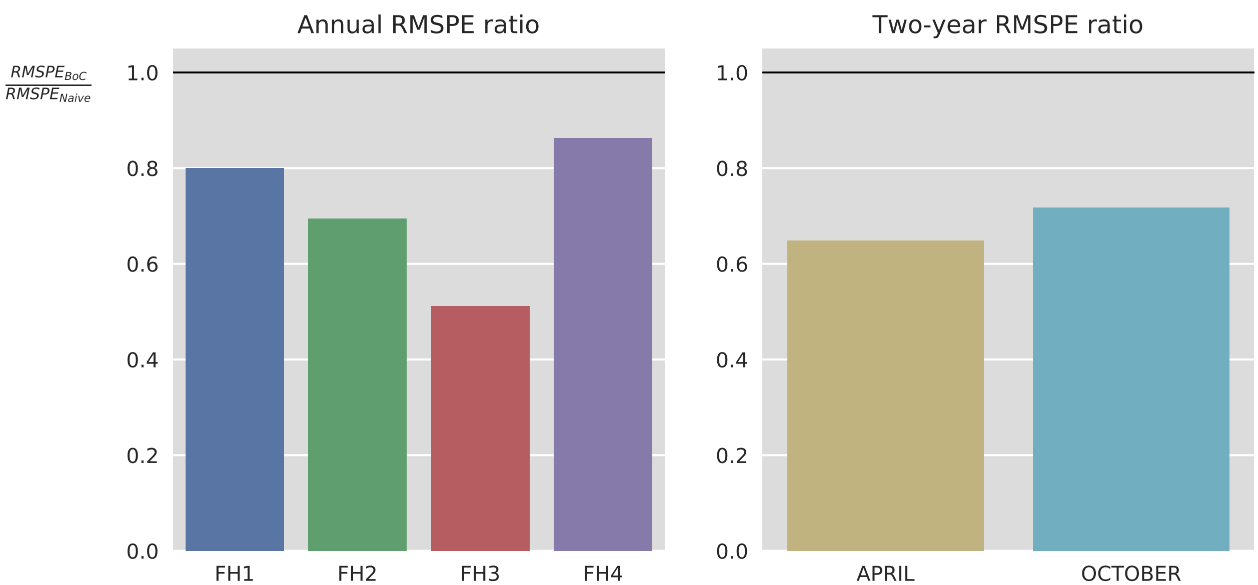
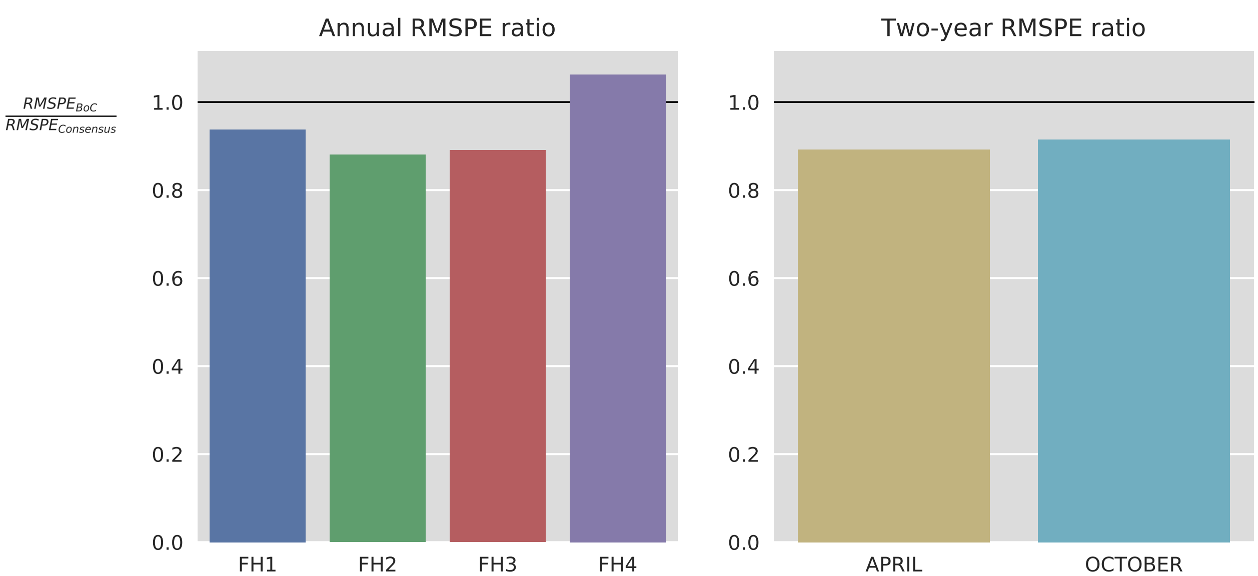


Figure 6: MPR outlook versus the wisdom of crowds—slight edge to the MPR



5 Conclusion

Based on 20 years of data, this paper examines the quality of MPR projections for real GDP growth. The analysis not only looks at the performance of the annual predictions but also assesses the two-year average prediction given the importance of correctly determining the expected level of real GDP at the end of the policy horizon. The predictions are compared with the final estimate of real GDP growth adjusted for three important conceptual changes through time. Adjusting for conceptual changes eliminates distortions in the forecast errors by taking into account the existing methodology at the time of producing the forecast.

To judge overall performance, we examine a number of metrics and several forecast horizons. When compared against two benchmarks, we find that the MPR is more accurate than a naïve forecast model and marginally superior to consensus. This result holds for annual predictions, two-year average projections and most forecast horizons. The accuracy of the MPR annual predictions, measured by the RMSPE, improves from 1.6 to 0.6 pp as more data become available. On a two-year average basis, the RMSPE is about 1.0 pp for both forecast horizons. Our results also suggest that the bias present in MPR forecasts is often not statistically significant for both annual and two-year projections. Indeed, according to the Mincer-Zarnowitz test, MPR predictions are deemed to be optimal. Nonetheless, we found a significant tendency to overpredict growth at the beginning of the forecast cycle and underpredict at the end of the cycle. The optimism bias over the full sample amounts to 0.7 pp on an annual basis and 0.3 pp for the two-year average. Finally, we have assessed the ability of the MPR to predict the sign of changes in real GDP growth correctly. At the beginning of the forecast cycle, the MPR correctly predicts a change in the pace of annual real GDP growth roughly 50 per cent of the time, improving to about 75 per cent at the end of the cycle. The sign of change is correctly predicted roughly 60 per cent of the time for the two-year average prediction.

In sum, the MPR predictions appear relatively good according to the tests used in this note. Looking at the performance of the MPR in different ways has helped identify scope for improvement. We see three areas for improvement: 1) increasing point forecast accuracy early on in the forecast cycle; 2) reducing bias at the longest forecast horizon; and 3) improving the ability to predict a change in the dynamic of real GDP growth. Improvements along these margins will help mitigate the risk of policy mistakes and support the already high level of credibility the Bank has achieved over the last 25 years of inflation targeting.

References

Andersson, M. K., G. Karlsson and J. Svensson. 2007. "The Riksbank's Forecasting Performance." Sveriges Riksbank Working Paper Series No. 218.

Bank of Canada. 2013. *Monetary Policy Report* (October).

Chang, A. C. and T. J. Hanson. 2015. "The Accuracy of Forecasts Prepared for the Federal Open Market Committee." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series No. 2015-062.

Chevalier, M. 2003. "Chain Fisher Volume Index Methodology." Statistics Canada Income and Expenditure Accounts Technical Series.

Clark, T. and M. McCracken. 2013. "Advances in Forecast Evaluation." Chap. 20, 2:1107–1201. Elsevier.

Cross, P. A. and D. Wyman. 2011. "The relationship between monthly, quarterly, and annual growth rates." *Canadian Economic Observer.*

D'Amico, S. and A. Orphanides. 2008. "Uncertainty and Disagreement in Economic Forecasting." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series No. 2008-56.

Dorich, J., M. K. Johnston, R. R. Mendes, S. Murchison and Y. Zhang. 2013. "ToTEM II: An Updated Version of the Bank of Canada's Quarterly Projection Model." Bank of Canada Technical Report No. 100.

Galton, F. 1907. "Vox populi (The wisdom of crowds)." *Nature* 75 (7): 450–451.

Gervais, O. and M.-A. Gosselin. 2014. "Analyzing and Forecasting the Canadian Economy through the LENS Model." Bank of Canada Technical Report No. 102.

Gneiting, T. 2011. "Making and Evaluating Point Forecasts." *Journal of the American Statistical Association* 106 (494): 746–762.

Granziera, E., C. Luu and P. St-Amant. 2013. "The Accuracy of Short-Term Forecast Combinations." *Bank of Canada Review* (Summer): 13–21.

Guénette, J.-D., N. Labelle, M. Leduc and L. Rennison. 2016. "The Case of Serial Disappointment." Bank of Canada Staff Analytical Note No. 2016-10.

Habgood, A. 2015. "Evaluating forecast performance." Independent Evaluation Office - Bank of England.

Jackson, C. 2002. "Capitalization of Software in the National Accounts." Statistics Canada.

Lahiri, K. and X. Sheng. 2010. "Measuring forecast uncertainty by disagreement: The missing link." *Journal of Applied Econometrics* 25 (4): 514–538.

Lees, K. 2016. "Assessing forecast performance." *Reserve Bank of New Zealand Bulletin* 79 (June): 1–19.

Macklem, T. 2002. "Information and Analysis for Monetary Policy: Coming to a Decision." *Bank of Canada Review* (Summer): 11–18.

Mincer, J. and V. Zarnowitz. 1969. "The Evaluation of Economic Forecasts." In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance,* 3–46. National Bureau of Economic Research, Inc.

Murray, J. 2013. "Monetary Policy Decision Making at the Bank of Canada." *Bank of Canada Review* (Autumn): 1–9.

Nalewaik, J. and A. Scherling. 2014. "Updated Historical Forecast Errors (4/9/2014)." Board of Governors of the Federal Reserve System.

Poloz, S. S. 2014. "Integrating Uncertainty and Monetary Policy-Making: A Practitioner's Perspective." Bank of Canada Staff Discussion Paper No. 2014-6.

Reifschneider, D. and P. Tulip. 2007. "Gauging the Uncertainty of the Economic Outlook from Historical Forecasting Errors." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series No. 2007-60.

Stock, J. H. and M. W. Watson. 2004. "Combination forecasts of output growth in a seven-country data set." *Journal of Forecasting* 23 (6): 405–430.

Tebrake, J. et al. 2011. "A preview of the historical revision of the Canadian System of National Accounts." Statistics Canada.

Theil, H. 1961. "Economic forecasts and policy." Contributions to economic analysis. North-Holland Pub. Co.

Timmermann, A. 2006a. "An Evaluation of the World Economic Outlook Forecasts." International Monetary Fund Working Paper No. 06/59.

———. 2006b. "Forecast combinations." *Handbook of Economic Forecasting* 1:135–196.

Vredin, A. et al. 2017. "Evaluation of the Riksbank's forecasts." *Riskbank Studies* (May).

Wallis, K. F. 1999. "Asymmetric density forecasts of inflation and the Bank of England's fan chart." *National Institute Economic Review* 167 (1): 106–112.

West, K. D. and W. K. Newey. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3): 703–708.

Appendix A

Table A1: Examples of central bank and international organization forecast assessments

Bank/Institution Name	Publication Name	Date of Last Forecast Assessed	Date of Paper
Central Bank of Brazil	Evaluation of the Central Bank of Brazil Structural Model's Inflation Forecasts in an Inflation Targeting Framework	2000Q3	July 2001
Federal Reserve Bank of St. Louis	Evaluating FOMC Forecasts	2001Q4	August 2002
Sveriges Riksbank	An evaluation of the Riksbank's forecasting performance	2006	March 2007
European Commission	The track record of the Commission's forecasts - an update	2005	October 2007
Reserve Bank of Australia	Estimates of Uncertainty around the RBA's Forecasts	2011Q4	July 2012
European Central Bank	An assessment of Eurosystem staff macroeconomic projections	2012	May 2013
Bank of England	Speech: Forecast errors	2013	May 2013
Danmarks Nationalbank	Danmarks Nationalbank's Projections for the Danish Economy 2008-12	2012	January 2014
OECD	OECD Forecasts During and After the Financial Crisis: A Post Mortem	2012	February 2014
International Monetary Fund, Independent Evaluation Office	On the Accuracy and Efficiency of IMF Forecasts: A Survey and Some Extensions	2011	February 2014
International Monetary Fund, Independent Evaluation Office	The Quality of IMF Forecasts	2011	May 2014
Federal Reserve Board of San Francisco	Persistent Overoptimism about Economic Growth	2015	February 2015
OECD	Lessons from OECD forecasts during and after the financial crisis	2012	March 2015
Federal Reserve Board	The Accuracy of Forecasts Prepared for the Federal Open Market Committee	2008	July 2015
Bank of England	Evaluating forecast performance	2014	November 2015
European Commission	European Commission's Forecasts Accuracy Revisited: Statistical Properties and Possible Causes of Forecast Errors	2014	March 2016
Reserve Bank of New Zealand	Assessing forecast performance	2015	June 2016
Bank of Japan	The effects of a central bank's inflation forecasts on private sector forecasts: Recent evidence from Japan	2015	August 2016
Sveriges Riksbank	Evaluation of the Riksbank's forecasts	2016	May 2017

Appendix B

Figure B1: MPR annual inflation forecast errors

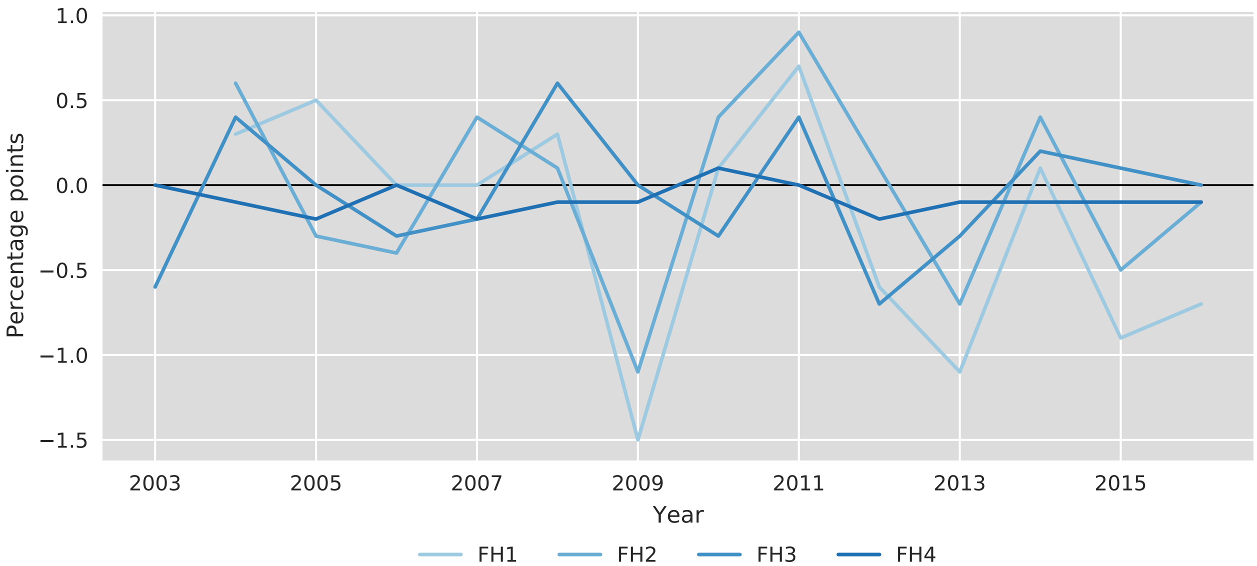
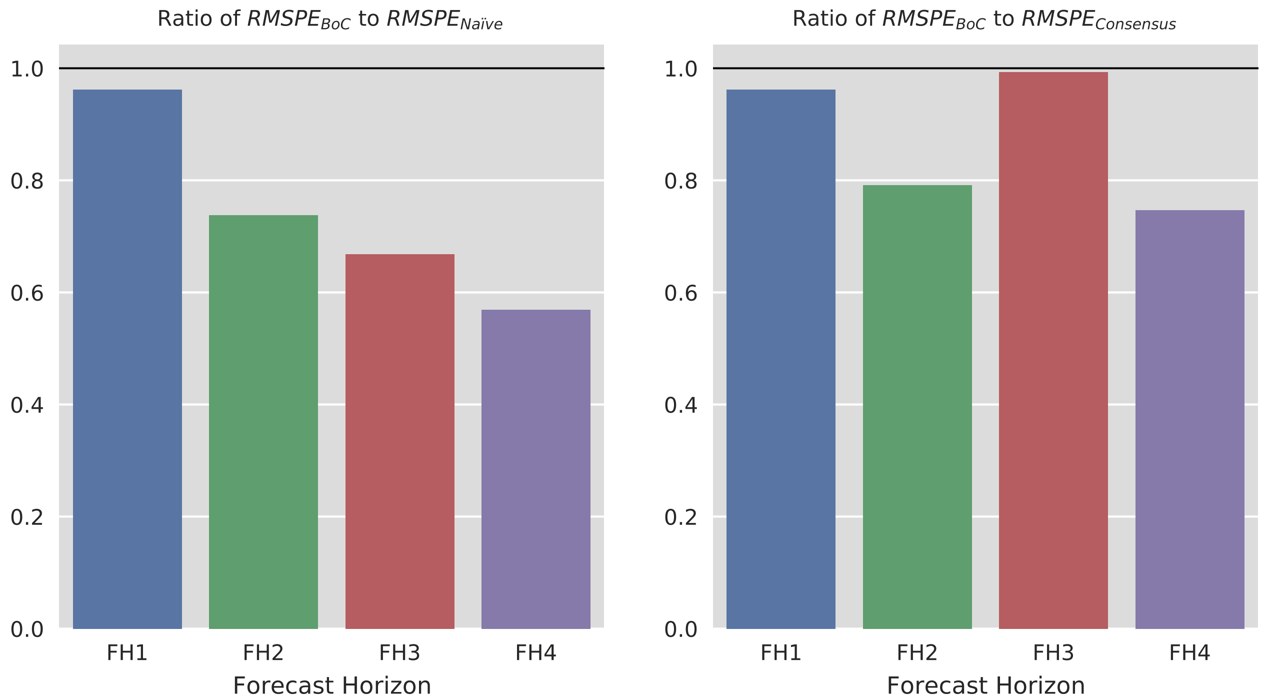


Table B1: RMSPEs of MPR inflation forecast errors trend down as the forecast horizon shortens

	FH1	FH2	FH3	FH4
RMSPE	0.68	0.55	0.37	0.12

Figure B2: MPR inflation forecast increasingly outperforms a naïve forecast model as the forecast horizon shortens, and outperforms the CE forecast, particularly in FH2 and FH4[34]



34. The naïve forecast is the ex-ante unconditional mean of the quarterly growth in the total CPI series. Because this series is not seasonally adjusted, the 10-year rolling average is computed separately for each quarter of the year. Blindly using a two per cent forecast, while not ideal, yields an RMSPE of 0.72.

Appendix C

Table C1: Impact of conceptual changes on the growth rate of real GDP

Date	Chain Fisher (Current Methodology)	Laspeyres Index	Software	R&D and Weapons	Adjusted GDP
1997	4.28	0.29	0.06	−0.12	4.63
1998	3.88	−0.04	0.05	0.02	3.77
1999	5.16	0.36	0.08	0.03	5.41
2000	5.18	0.29	0.08	0.23	5.17
2001	1.77	−0.29	0.07	0.09	1.32
2002	3.01	−0.11	0.07	−0.04	2.87
2003	1.80			0.01	1.79
2004	3.09			0.04	3.05
2005	3.20			0.02	3.18
2006	2.62			−0.01	2.63
2007	2.06			−0.06	2.12
2008	1.00			0.11	0.89
2009	−2.95			−0.06	−2.89
2010	3.08			0.01	3.08
2011	3.14			0.03	3.12
2012	1.75			−0.09	1.83
2013	2.48			−0.00	2.48
2014	2.57				2.57
2015	0.94				0.94
2016	1.47				1.47

Note: First four columns decompose the methodological changes to GDP. Adjusted GDP = Chain Fisher + Laspeyres - Software - R&D and Weapons.