How Good Are Canadian Macroeconomic Forecasts?

Richard Deaves*
Michael G. DeGroote School of Business, McMaster University

Introduction

How good is the performance of professional Canadian macroeconomic forecasters? There are several issues of interest. First, taken as a group, are they able to generate useful predictions, in the sense that they are at least able to improve on simple mechanical alternatives. In the latter category, one could, most obviously, predict no change in the relevant variable. For example, if unemployment is currently 10 per cent, would one be just as well off on average by predicting that unemployment would remain at this level rather than making use of a professional forecast? As a second issue, are some macroeconomic variables more predictable than others? For instance, is inflation more susceptible to prognostication than real economic activity? Third, are there substantial divergences in the quality of different forecasting organizations? If not, demanders of professional macroeconomic forecasts need not conduct a careful search for the best product.

These questions are here addressed by an examination of a survey of macroeconomic forecasts conducted by the *Financial Post*. Since 1974, this business newspaper has conducted every February a telephone survey of major financial institutions and forecasting companies for eight macroeconomic variables (real growth, CPI inflation, unemployment, housing starts, the trade balance, corporate profits, the U.S./Canada exchange rate, and the prime rate), and published the results.[1] Except in the first couple of years, there have generally been over 20 respondents. In most years, the Canadian Imperial Bank of Commerce (CIBC), Dominion Securities (now RBC Dominion), DuPont Canada, Informetrica, McLeod Young Weir (now Scotia McLeod), the National Bank, (the former) Royal Trust and Wood Gundy (now CIBC Wood Gundy) have replied. Other respondents have been more occasional.

Here attention will be reserved for arguably four of the most critical macroeconomic magnitudes: real growth, CPI inflation, the unemployment rate and the prime rate.[2] To cite Canadian studies using mostly earlier samples, Daub (1981, 1987, 1993) has investigated the performance of Canadian forecasters for real growth and inflation (using the implicit price deflator for the latter); as have Pesando (1981) and Pesando and Plourde (1988) for short-term interest rates (T-bills and finance company paper).[3] I revisit the issue, looking at data extending into the early 1990s and asking a richer set of questions than previously posed. Both survey means (in section two) and, for a restricted sample, selected individual forecasts (in section three) are assessed.

Performance of Survey Means

Table 1 presents for all four of the variables under consideration various sample statistics which provide information on the performance of the mean survey responses for the entire 20-year sample. Most of these sample statistics are directly based on forecast errors, namely the difference between the actual values of a variable and its forecast value.[4]

The first two rows provide the mean error (ME) and standard deviation (SD) of the forecast errors for each of the variables. Indeed, the latter indicate that these macroeconomic variables are difficult to predict. For example, assuming normality, one would expect that

Table 1 Sample statistics of mean survey responses

	Real growth	CPI inflation	Unemployment rate	Prime rate
Mean Error	-0.293	0.168	-0.045	0.682
Standard Deviation	1.447	1.019	0.743	2.118
Absolute t-statistic	0.907	0.736	0.271	1.441
Bias p-value	0.302	0.265	0.848	0.379
Mean Absolute Error	1.108	0.733	0.495	1.682
Root Mean Squared Error	1.440	1.008	0.725	2.175
Root Mean Squared Error/Standard Deviation	0.605	0.310	0.392	0.690
Theil U Statistic	0.438	0.496	0.615	0.978

one third of the time the prime rate will not be successfully predicted within 2.12 percentage points.

Mean error provides information on bias. Since an unbiased forecaster should be as likely to err on the high side as on the low side, he/she should produce MEs close to zero for reasonably large samples. This is indeed so, as evidenced by the (absolute) t-statistics in the third row which are all insignificantly different from zero. Results from a more rigorous test for bias appear in the fourth row. Regressions are run of actual values on forecast values and a constant. Unbiasedness implies that the intercept should be zero, and the slope unity. An F-test is employed to test this joint hypothesis. The p-values all indicate a safe inability to reject the null of unbiasedness.[5]

Though the survey mean responses do well with respect to bias, they may still be weak in terms of variability about the mean. Mean absolute error (MAE) and root mean squared error (RMSE) capture both bias and variance. RMSE is similar to MAE in that positive and negative errors are equally penalized because forecast errors are squared before being averaged and having the square root taken.[6] It is different in the sense that high errors are penalized relatively more than low errors through the act of squaring. The implicit assumption is that large errors are proportionately more serious than small errors.

To focus on MAEs, the typical absolute error in predicting real economic growth was a little over 1 percentage point; a half of 1 percentage point in predicting the unemployment rate; a bit less than three quarters of 1 percentage point in predicting inflation; and well above 1.5 percentage points for the prime rate. On the surface, it seems that the prime rate was the most difficult to forecast. As for RMSEs, they yield different values but the same ordering.[7]

While a ranking may be justifiable given that all variables are measured in the same units (i.e., percentages), it should be realized, however, that a volatile series will be more difficult to forecast than a stable one. With this in mind, it is more meaningful to focus on the next row of the table, which provides RMSEs relative to SDs for the underlying series.[8] Notice that from this perspective as well the prime rate remains the most difficult to forecast with real growth continuing to come in third. There is however a change in the ranking between inflation and unemployment, with inflation now appearing most predictable. One reason for the lower unadjusted RMSE for unemployment is simply that it is a more stable series.[9]

Table 2 Mean survey response statistics for subperiods

	Real growth	CPI inflation	Unemployment rate	Prime rate
	First subperiod: 1974-83			
Mean Error	-0.787	0.598	-0.036	1.478
Standard Deviation	1.696	1.266	0.983	2.479
Mean Absolute Error	1.324	1.084	0.627	2.317
Root Mean Squared Error	1.792	1.341	0.933	2.777
Root Mean Squared Error/Standard Deviation	0.657	0.680	0.483	0.819
Theil U Statistic	0.421	0.558	0.724	1.096
	Second subperiod: 1984-93			
Mean Error	0.201	-0.262	-0.054	-0.113
Standard Deviation	0.999	0.425	0.445	1.388
Mean Absolute Error	0.839	0.380	0.363	1.048
Root Mean Squared Error	0.969	0.481	0.426	1.321
Root Mean Squared Error/Standard Deviation	0.483	0.365	0.286	0.530
Theil U Statistic	0.594	0.318	0.408	0.727

To investigate the stability of these findings, the full 20-year sample was partitioned into two equal subperiods, with the relevant statistics shown in Table 2.[10] In general, the first subperiod can be characterized as one of high and rising inflation, while the second enjoyed lower and more stable price changes. The fact that the first subperiod was more turbulent than the second is evidenced by the fact that the standard deviations of the actual values in the second period relative to the first range from 67 per cent (for inflation) to 77 per cent (for unemployment). Notice the tendency of forecasters, during the first (second) subperiod, to underestimate (overestimate) inflation and interest rates. Clearly to a great extent the upward/downward movements in these two variables were unanticipated.[11] As for relative forecasting performance, due to the recent more stable environment, all forecast SDs, MAEs, RMSEs and RMSEs relative to actual SDs were lower during 1984-93 than during 1974-83, and sometimes substantially so.

One problem suggests itself. What if the "best" forecast is itself not very useful? In particular, what if one could have done just as well by naively predicting no change in the relevant variable?[12] The way to get around this problem is to calculate the Theil U statistic, which is simply the ratio of the RMSE of the relevant forecasting procedure to the RMSE of the naive no-change prediction. Obviously a Theil U of one or more implies that the forecasting methodology could not even outperform the naive no-change "straw man." These Theil Us are shown for the full

Table 3 Theil Us for selected survey participants, 1979-90

	Real growth	CPI Inflation	Unemployment rate	Prime rate
Survey mean	0.479	0.299	0.720	1.097
BC Credit Union	0.462	0.401	0.773	1.218
CIBC	0.560	0.313	0.655	N.A.
Dominion Securities	0.527	0.293	0.738	1.066
Dupont Canada	0.444	0.347	0.655	1.082
National Bank	0.592	0.395	0.835	1.217
McLeod Young Weir	N.A.	0.351	0.632	1.426
Informetrica	N.A.	0.550	0.671	1.078

sample in the last row of Table 1, and for the two subperiods in the last row of the relevant panels of Table 2.

Focussing on the full-sample results, the news is for the most part favorable for the professional forecasting community. Real economic growth, inflation and the unemployment rate (in this order) can be predicted with some success incremental to the naive approach.[13] Notice that the rate of unemployment, though it had the lowest forecast ME, SD, MAE and RMSE, is actually the least susceptible to improvement of the three.

The prime rate stands out as the noticeable exception. Despite a clear informational advantage,[14] the Theil U is only slightly below unity, suggesting no value added over simply predicting the same value for the coming year. It should be noted that the inability of professionals to improve on the naive no-change prediction in the case of short-term interest rates is consistent with the results of Belongia (1987), Pesando and Plourde (1988) and McNees (1992).[15]

As for the subperiod results, notice that forecasters exhibited enhanced ability to better the naive no-change prediction for all variables other than real economic activity during 1984-93. Whether this is attributable to improvements in forecasting techniques or to the more stable environment is open to debate. In addition, notice that some predictable component, albeit a small one relative to the other variables, does seem to have been uncovered by forecasters of late for the prime rate.

Comparing Forecasters

Those demanding economic forecasts might naturally wonder whether some institutions do better than others. Suppliers will of course also be interested as long as their performance is strong. Of course it is possible that different institutions have different strengths: one might be adept at forecasting prime while another's comparative advantage lies in predicting real economic activity.

Some might question the validity of focussing on institutions rather than individuals. Given the mobility of individuals between institutions, not to mention changes in overall staffing and techniques employed, can one meaningfully examine institutions? To the extent that there is in an organization a consistent approach to prediction, a forecasting "production function" as it were, then the answer is yes. Since this issue is not easily resolved, one should of course take any rankings with a grain of salt.[16]

This caveat notwithstanding, Table 3 provides a little evidence. Since participation in the survey is inconsistent for many respondents, I focus on the longest subsample where there was a reasonable number of respondents in each category participating every year. This is the 12-year period 1979-90. For brevity, I focus exclusively on the Theil U statistics. As a point of reference, I also show the

Theil U for the overall survey mean for this 12-year period.

A few salient observations can be made. Focussing on the overall survey, the same variables that were somewhat predictable in the overall sample were also somewhat predictable in this subsample. Nevertheless, there is one change in ranking. During 1979-90, inflation showed the most scope for predictability versus the naive approach, with real economic growth coming next, followed by unemployment. Once again prime remained elusive, since one would have been better off simply predicting prime would remain at its current level. Not only did the latter naive forecast dominate the survey mean, *but, remarkably, it also outperformed all six individual forecasters during this subperiod.*

Did one or two institutions rise above the crowd on an overall basis? If one ranks the Theil Us from best (i.e. lowest) to worst (i.e. highest), one institution, Dominion Securities, is at the top for two of the four variables, with no bottom performances. Another institution, the National Bank, is at the bottom on two occasions, with no points in the win column.

Another approach is to restrict ourselves to institutions that appear in all four categories (namely, the BC Credit Union, Dominion Securities, Dupont Canada and the National Bank), rank all competitors in each category, add up the individual ranks to arrive at total ranks, and then compare them. The survey mean is used as the fifth "competitor." The results are as follows: Dupont (7), Dominion Securities and the survey mean (10), BC Credit Union (17) and the National Bank (18). It is interesting that by this measure an industrial company comes out on top, surpassing the three financial institutions. Nevertheless, given the lack of observations, both time-series and cross-sectional, these comparative results should not be taken too seriously.[17]

Conclusions

In closing, it is appropriate to summarize the answers that we found to the questions posed at the outset. Taken as a group, how useful were professional predictions? I am able to conclude that they are useful in being able to outperform the naive no-change prediction. There are two caveats. First, though this was true for real growth, inflation and unemployment (to varying degrees), the prime rate was an exception. For the full sample period, any predictable short-term changes in short-term interest rates appear to have been swamped by noise, though during 1984-93 there does seem to have been some predictable component that was captured by forecasting organizations. Second, it should be pointed out that there do exist other fairly obvious mechanical models (for example, ARIMA models) that may do as well as the forecasters. This issue was not explored here.

Finally, were there substantial divergences in the quality of different forecasting organizations? To be fair, one cannot really say based on our scant evidence, although some suggestive evidence was uncovered that certain forecasters tended to do better than others.

Notes

* The author would like to acknowledge the excellent research assistance of Carlos Leite, as well as the helpful comments of two anonymous referees. The usual disclaimer regarding errors holds. I thank Catherine Harris of the *Financial Post* for providing me with the data.

1. Thus we have 20 sets of observations. Note that the US exchange rate did not appear until 1979.

2. More specifically, for real growth, GNP was used initially and GDP later (first in 1990); the unemployment rate and the CPI levels for the year (from which inflation is calculated) are averages of monthly observations; and the prime rate is that in effect at the end of the year.

3. Daub's 1987 sample ends in 1984 and Pesando's in 1985, while my sample runs a further eight years.

4. The actual values were taken from the following year's newspaper article (except for 1993 where various issues of the *Bank of Canada Review* and the *Canadian Economic Observer* were used), and thus might differ slightly from the final numbers due to revisions.

5. One interesting result of the regressions is that (except in the case of the prime rate) the slope coefficient is over one (though never significantly so). This is reminiscent of the conservatism of forecasters found by Spiro (1989).

6. More formally, RMSE is calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T} (y_t - \hat{y}_t)^2}{T}}$$

where y_t is the variable at time t and $\hat{y}_t$ is the prediction for t made at some point before t.

7. Daub (1987) found substantially higher RMSEs for real growth (1.91 per cent) and GDP deflator inflation (2.06 per cent). There are two likely reasons for this. First, it seems that on balance his earlier period was somewhat more exposed to macroeconomic shocks. Second, he investigated individual forecasters rather than survey averages, and it is well known (Zarnowitz, 1984) that RMSEs for individual forecasts tend to be higher than those based on pooled data.

8. A byproduct of this division is to "deunitize" the RMSEs: both the latter and the SDs are measured in terms of the original units. One would expect that the standard deviations of the forecasts are less than those of the actual series (Samuelson, 1965). Indeed the ratios of the latter to the former range from 1.15 to 1.58.

9. The SDs for the underlying series are 2.38 per cent, 1.85 per cent, 3.25 per cent and 3.15 per cent for real growth, unemployment, inflation and prime respectively.

10. Given the limited number of observations, the t-statistics and p-values have been omitted.

11. None of the t-statistics are statistically significant (at 5 per cent), though inflation during the second subperiod comes the closest.

12. When the macroeconomic variables refer to growth rates (as in real activity and inflation), one can also say that the naive prediction is for the same change in the underlying magnitude.

13. Are the lower RMSEs for three of the four variables reflective of statistically significant superiority relative to the straw man? Using an indirect and stringent test of significantly different RMSEs, as suggested by Ashley, Granger, and Schmalensee (1980), one cannot unequivocally conclude that this is so. See Deaves (1993) for details on the test. Implicitly, one requires both significantly less bias and significantly less variability.

14. The survey takes place in February, while the naive "forecast" is based on the December prime rate. Informational advantage also exists to some extent in the case of the other variables, but is less of an issue because of the time lag that exists (for the other three variables) in the collection and release of the relevant macroeconomic information .

15. Earlier studies by Pesando (1981) and Throop (1981) did find some incremental success for professional forecasts.

16. One can also question the tracking of individuals over time. To the extent that a forecast is really a team effort, can one meaningfully compare an individual's forecast, while working at Institution X at one point in his/her career, to the same individual's forecast, now at Institution Y at another point in his/her career? Of course, the present data source has really forced the issue. Though the *Financial Post* does identify the (lead) individual responsible for a forecast, there are too many staffing changes to arrive at a reasonable sample for individuals.

17. Note that Daub (1987) concludes "there appears to be no difference among forecasters."

References

Ashley, R., C.W.J. Granger, and R. Schmalensee (1980) "Advertising and Aggregate Consumption: An Analysis of Causality," *Econometrica*, Vol. 48, No. 5, July, pp. 1149-67.

Belongia, M.T. (1987) "Predicting Interest Rates: A Comparison of Professional and Market-based Forecasts," *Federal Reserve Bank of St. Louis Review*, Vol. 69, No. 3, March, pp. 9-15.

Daub, M. (1981) "The Accuracy of Canadian Short-term Economic Forecasts Revisited," *Canadian Journal of Economics*, Vol. 14, No. 3, August, pp. 499-507.

Daub, M. (1987) *Canadian Economic Forecasting: In a World Where All's Unsure* (Kingston: McGill-Queen's University Press).

Daub, M. (1993) "A Comparison of Canadian and American Economic Forecasts," *Canadian Business Economics*, Vol. 1, No. 2, pp. 61-68.

Deaves, R. (1993) "Forecasting Canadian Short-term Interest Rates," McMaster University Working Paper.

McNees, S.K. (1992) "How Large Are Economic Forecast Errors?" *New England Economic Review*, July/August, pp. 25-42.

Pesando, J.E. (1981) "On Forecasting Interest Rates: An Efficient Markets Perspective," *Journal of Monetary Economics*, Vol. 8, No. 4, November, pp. 305-18.

Pesando, J.E., and A. Plourde. (1988) "The October 1979 Change in the U.S. Monetary Regime: Its Impact on the Forecastability of Canadian Interest Rates," *Journal of Finance*, Vol. 43, No. 1, pp. 217-39.

Samuelson, P. (1965) "Proof that Properly Anticipated Prices Fluctuate Randomly,"

Industrial Management Review, Spring, pp. 41-49.

Spiro, P.S. (1989) "Improving a Group Forecast by Removing the Conservative Bias in its Components," *International Journal of Forecasting*, Vol. 5, No. 1, pp. 127-31.

Throop, A.W. (1981) "Interest Rate Forecasts and Market Efficiency," *Federal Reserve Bank of San Francisco Economic Review*, Spring, pp. 29-43.

Zarnowitz, V. (1984) "The Accuracy of Individual and Group Forecasts from Business Outlook Surveys," *Journal of Forecasting*, Vol. 3, pp. 11-23.