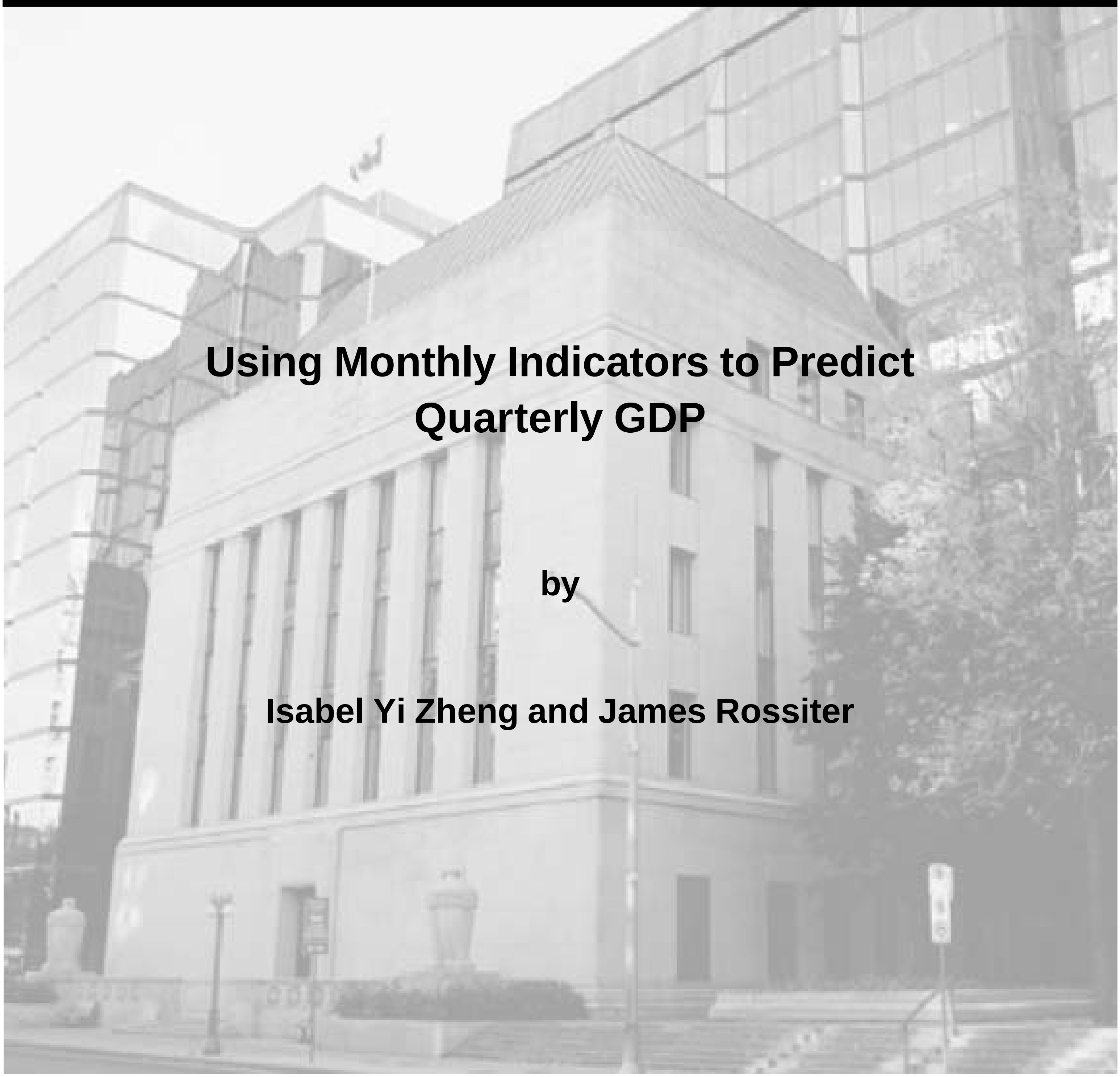
Bank of Canada Banque du Canada

Working Paper 2006-26 / Document de travail 2006-26



Using Monthly Indicators to Predict Quarterly GDP

by

Isabel Yi Zheng and James Rossiter

ISSN 1192-5434

Printed in Canada on recycled paper

Bank of Canada Working Paper 2006-26

August 2006

Using Monthly Indicators to Predict Quarterly GDP

by

Isabel Yi Zheng and James Rossiter

Research Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
izheng@bankofcanada.ca
jrossiter@bankofcanada.ca

Contents

Acknowledgements

The authors wish to thank Gillian Boden, Frédérick Demers, Richard Dion, Bob Fay, Sharon Kozicki, Raphael Solomon, and seminar participants at the Bank of Canada for their helpful comments and suggestions. Philippe Marcil provided excellent technical assistance.

Abstract

The authors build a model for predicting current-quarter real gross domestic product (GDP) growth using anywhere from zero to three months of indicators from that quarter. Their equation links quarterly Canadian GDP growth with monthly data on retail sales, housing starts, consumer confidence, total hours worked, and U.S. industrial production. The authors use time-series methods to forecast missing observations of the monthly indicators; this allows them to assess the performance of the method under various amounts of monthly information.

The authors' model forecasts GDP growth as early as the first month of the reference quarter, and its accuracy generally improves with incremental monthly data releases. The final forecast from the model, available five to six weeks before the release of the National Income and Expenditure Accounts, delivers improved accuracy relative to those of several macroeconomic models used for short-term forecasting of Canadian output. The implications of real-time versus pseudo-real-time forecasting are investigated, and the authors find that the choice between real-time and latest-available data affects the performance ranking among alternative models.

JEL classification: C22, C53
Bank classification: Economic models; Econometric and statistical methods

Résumé

Les auteurs proposent un modèle qui permet de prévoir la croissance du produit intérieur brut (PIB) réel pour le trimestre en cours à partir d'indicateurs mensuels relatifs à une partie ou à la totalité du trimestre concerné. Leur équation lie la croissance trimestrielle du PIB canadien aux statistiques mensuelles des ventes au détail, des mises en chantier de logements, de la confiance des consommateurs, du nombre total d'heures travaillées et de la production industrielle américaine. Lorsque des observations manquent pour certains mois, les auteurs en extrapolent la valeur en ayant recours à des méthodes d'analyse de séries chronologiques. Cette démarche leur permet d'éprouver la validité de la méthode employée en présence d'un volume d'informations mensuelles variable.

Le modèle élaboré peut servir à prédire la croissance du PIB dès le premier mois du trimestre de référence, et sa précision s'améliore chaque fois que paraissent de nouvelles données mensuelles. La dernière prévision du modèle, établie cinq à six semaines avant la publication des chiffres des Comptes nationaux des revenus et dépenses, est de meilleure qualité que les prévisions à court terme produites par plusieurs modèles macroéconomiques au sujet de la production canadienne.

Les auteurs analysent les implications du choix de données en temps réel plutôt que des dernières données disponibles. Ils constatent que ce choix influe sur le classement relatif des modèles du point de vue de la qualité de leurs prévisions.

Classification JEL : C22, C53
Classification de la Banque : Modèles économiques; Méthodes économétriques et statistiques

1. Introduction

Central banks' assessments of the current and future states of the economy play a vital role in the conduct of monetary policy. Providing an accurate and timely assessment of near-term economic growth is challenging, since the delay between the end of a quarter and the publication of its national accounts figures can be up to two months (as in the case of Canada). While more frequent measures of economic activity are published with a shorter lag, these tend to cover only specific sectors of the economy and can be very volatile. Most forecasting models of quarterly economic growth are unable to take these monthly measures into account when they are released, since the models typically consider only (lagged) quarterly data and are not designed to incorporate fewer than three months of a quarter's monthly data.

This paper addresses both issues by building a forecasting model of quarterly economic growth that is able to incorporate anywhere from zero to three months of monthly data in a quarter. Previous research in this field has yielded models for the United States (Trehan 1992; Ingenito and Trehan 1996) and the euro area (Rünstler and Sédillot 2003), among others. This paper develops a single-equation autoregressive distributed lag (ADL) model (the "bridge" equation) for Canadian quarterly gross domestic product (GDP) growth using quarterly averages of monthly indicators as explanatory variables. To forecast current-quarter GDP growth, any missing monthly indicator data are predicted using a variety of univariate and multivariate monthly models. The model produces a forecast at the beginning of the reference quarter, since its performance generally improves with subsequent monthly indicator data releases. We examine the relative performance of this model using both first-release real-time and latest-available data sets. This assessment of real-time performance is unique to this paper: previous research on the topic has generally used revised data. The results suggest that this model would be effective as a tool for the short-term forecasting of output growth in the Canadian economy, and a useful complement to other models used for short-term forecasting of Canadian output. It may be seen as an econometric approximation of the "current analysis" performed by both public and private sector economists.[1]

The paper is organized as follows. Section 2 presents a summary of existing literature. Section 3 describes the general-to-specific strategy used to determine which variables and lag lengths to include and other data issues. The final specification of the bridge equation for GDP comes in section 4, followed by a discussion of the methods to predict missing intraquarter

[1] "Current analysis" (also sometimes referred to as "monitoring") is the term used to describe the short-term forecasting of the economy.

values for monthly indicators. Next, the model's out-of-sample forecasts are compared with those generated by alternative models. Section 7 tackles the issue of model assessment in the context of real-time data, and section 8 concludes with a summary of the findings.

2. Literature Review

A number of studies demonstrate the benefits of incorporating monthly indicators into short-term quarterly GDP forecasting models using various methodologies. Klein and Sojo (1989) use two approaches. The first involves constructing forecast equations for all sub-components of GDP (e.g., predicting quarterly clothing and shoes expenditure with monthly retail sales at apparel and accessory stores) and then aggregating these subcomponents to obtain a forecast for total GDP growth. Unknown monthly data in the quarter are predicted using autoregressive moving average (ARMA) models. Klein and Sojo's second method is that of principal component analysis, where the first principal component of 25 monthly indicators is used to estimate current-quarter GDP. Kitchen and Monaco (2003) estimate 30 equations (each regressing one indicator on GDP with varying months of information) to obtain 30 forecasts for current-quarter GDP growth. These forecasts are then combined using a weighted average based on their $\overline{R}^2$'s. Miller and Chin (1996) employ a different type of forecast aggregation, creating a GDP growth forecast using only quarterly data or only monthly data. The two independent forecasts are then combined using a weighted average that maximizes forecast accuracy. Zadrozny (1990) constructs a multivariate mixed-frequency ARMA model that does not require the interpolation of lower-frequency data.

This paper follows an approach similar to that of both Ingenito and Trehan (1996) and Rünstler and Sédillot (2003). The earliest work to use this approach is Trehan (1992), who finds that real-time forecasts from a univariate equation regressing current-quarter GDP on non-farm payroll employment, industrial production, and retail sales outperforms the Blue Chip consensus forecast for the U.S. economy.[2] His model takes the forecasts of monthly variables from univariate monthly equations to help forecast quarterly GDP growth. In an update to this model, Ingenito and Trehan (1996) examine more than 30 potentially informative monthly indicators, ultimately choosing only two indicators without compromising the performance of the single-equation model.

[2] The Blue Chip consensus is the average of a survey of roughly 50 private sector U.S. forecasters.

Similar approaches have been taken to evaluate the usefulness of high-frequency data in forecasting GDP or other National Accounts components for other economies.[3] Rünstler and Sédillot (2003) review a number of these studies for the euro area, and conclude that while the majority of such models see improvements in near-term forecasts of GDP, "they added little to the timeliness of conjunctural analysis" due to requirements for full quarterly information on indicators that could only be fulfilled two weeks ahead of the first official release of euro area GDP. To assess the forecasting performance of GDP equations with incomplete quarterly information, Rünstler and Sédillot (2003) subsequently propose a method to combine a quarterly univariate bridge equation for GDP with time-series models that forecast missing observations of monthly indicators. For the period 1998Q1 to 2001Q4, they find that GDP growth predictions for the current quarter based on euro area industrial production, retail sales, and car registration are superior to those yielded by autoregressive integrated moving average (ARIMA) forecasts, even when only one additional month of data is used.

Analysis of our model in a real-time environment is necessary to evaluate the model's performance appropriately. This paper examines real-time data (section 7); empirical work on this topic in Canada is fairly sparse, due to a paucity of real-time data.[4] Much real-time research, however, has been conducted on the U.S. and euro area economies.[5] Early research by Denton and Kuiper (1965) and Cole (1969) shows that the variance of forecasting errors increases when preliminary rather than revised data are used. More recently, Diron (2005) studies the implications of data revisions for forecasting GDP growth based on monthly indicators in the context of euro area economic activity. After examining the performance of eight bridge equations relating output growth to various macroeconomic, financial, and survey data, she concludes that the use of revised data does not bias the overall reliability assessment of short-term GDP forecasts and that, in most cases, data revisions contribute less to forecast errors than model misspecification. Koenig, Dolmas, and Piger (2001) suggest that when a model to predict real GDP growth in the United States is estimated using only initial releases of data, its performance is better than if revised data had been used and it compares favourably with the Blue Chip consensus.[6]

---

[3] For example, see Liou and Shen (1996) for Taiwan and Coutiño (2005) for Mexico.
[4] A description of concepts related to data vintage is given in Appendix E.
[5] For a more comprehensive literature review, see Babineau and Braun (2003).
[6] The Koenig, Dolmas, and Piger (2001) (KDP) study differs from the traditional real-time approach where all the observations, initial or not, in a data vintage are used. The KDP approach requires a long history of initial releases, which is very hard to come by in other countries. In addition, their assumption that initial releases are efficient estimates of subsequent releases has not been found to hold generally (see comments in Croushore and Stark 2003, for example).

3. Data and the Indicator Selection Strategy

3.1 The data set

The search for variables to be included in the GDP bridge equation is guided by timeliness, stability, and parsimony. The model's *best* forecast should thus be produced as early as possible and a relatively accurate forecast should be available before the release of the previous quarter's National Accounts. For example, the National Accounts for the third quarter (Q3) are released around 30 November, about 60 days after the quarter ends. By this time, the bridge equation should be able to produce a reliable initial forecast of GDP for the last quarter of the year. The model should yield progressively better estimates of Q4 GDP through to the end of February when the Q4 GDP numbers are released. As such, only monthly indicators with a publication lag shorter than 50 days are considered for inclusion in the model. Additionally, we avoid indicators that undergo frequent and substantial revisions. The model should also contain a limited number of indicators to prevent overspecification and to facilitate updates once the model is operative.

Appendix A lists monthly economic and financial indicators that meet the first two criteria.[7] The indicators are sorted in sequential order of data release; their approximate publication lags are in the last column. Most of the economic indicators directly correspond to components of GDP. Examples include motor vehicle sales (in units), housing starts (in units), and retail sales (constant dollars). The survey-based Index of Consumer Confidence (Conference Board of Canada) and the U.S. Purchasing Managers Index (Institute for Supply Management) are two exceptions. We include these because previous research (such as Koenig 2002) has revealed the potential of these indicators to predict aggregate output growth in Canada and the United States, respectively.

International merchandise trade, industrial production, and manufacturers' shipments, orders, and inventories (MSOI) series are not considered for inclusion in the model. Merchandise trade is dropped due to its high month-to-month volatility and susceptibility to sizable revisions. Despite strong evidence supporting the use of monthly industrial production (IP) to track GDP in several studies for other economies (e.g., Trehan 1992; Rünstler and Sédillot 2003), we do not use Canadian IP in our model, since it is published two months after the reference month. A similar but more comprehensive monthly series available for the Canadian economy is GDP at basic prices by industry. This series is built from various

[7] Monthly financial data are converted from daily frequency using monthly averages.

indicators, including, for instance, manufacturing shipments and retail sales. We exclude GDP at basic prices for several reasons. Much of the data used to construct this series are already included in other variables we use. If we include GDP at basic prices, other domestic variables included in the regression are generally "crowded out" by the basic price series, and the forecasting accuracy does not improve significantly. The only case where the improvement is significant is if GDP at basic price is used alone after two months of data were available. In addition, the series is not available on a real-time basis, and is published with almost a two-month delay. To compensate for the lack of domestic variables that capture monthly external demand or industrial production, we consider U.S. industrial production and other related activity measures, motivated by the economic ties between the two countries. MSOI is not incorporated in the model because, at the time that this exercise was initiated, Statistics Canada did not compile real MSOI data. It has since begun this practice, and future work may consider this series.

The inclusion of financial variables such as short-term and long-term interest rates (both the levels and the spread between them), exchange rates, and commodity prices draws inspiration from empirical evidence of their usefulness for predicting Canadian GDP growth in the literature (for example, Duguay 1994 and Murchison 2001). Finally, the composite leading indicator from Statistics Canada is also considered, given that it is intended to predict cyclical movements in aggregate output. There is some overlap between its components and the other series in Appendix A.[8]

3.2 Selecting indicators for the bridge equation

The bridge equation relates quarterly averages of the monthly indicator variables to quarterly GDP growth. The general specification of the ADL bridge equation is as follows:

$$A(L)y_t = \sum_{i=1}^{k} B_i(L)x_{i,t} + \varepsilon_t, \tag{1}$$

where y_t denotes the log difference of real GDP at market price (i.e., the growth rate of quarterly GDP), $x_{i,t}$ are monthly indicators averaged to quarterly frequency (in first difference of

[8] The leading indicator is a simple unweighted five-month moving average of indexes of the following ten components: housing index (starts and resales), business and personal services employment, S&P/TSX stock price index, money supply (M1), U.S. composite leading indicator, average work week, new orders for durables, shipments/inventories of finished goods, furniture and appliance sales, and other durable goods sales.

logs where appropriate; see Appendix A for specifics), and $A(L)$ and $B_i(L)$ are their respective lag polynomials. The sample period for the bridge equation is 1986Q1 to 2004Q2.

Starting with the indicators identified in the previous section, we adopt a general-to-specific approach to select variables and their lag lengths. In particular, we use a strategy similar to that of Hendry and Mizon (1978) by including as many variables (and lagged terms) as possible in the initial regression, eliminating the most insignificant variables one at a time.[9] Only those with a marginal significance level on the *t*-statistic of below 0.10 are retained. We choose the variables in the initial regression to avoid multicollinearity problems and to accommodate the relatively small sample size.[10]

4. Quarterly Equation for GDP Growth

Our preferred bridge equation is a quarterly ADL model linking quarterly GDP growth to the Canadian index of consumer confidence (*C*), total hours worked (*W*), the number of housing starts (*H*), constant dollar retail sales (*R*), and U.S. industrial production (*IP*):

$$\begin{aligned}\Delta G\hat{D}P_t = 0.002C_t + 0.176\Delta W_t + 0.214\Delta W_{t-1} + 0.025\Delta H_t + 0.146\Delta R_t + \\ 0.271\Delta IP_{t-1} - 0.155\Delta IP_{t-3} + 0.231\Delta GDP_{t-3}\,.\end{aligned} \tag{2}$$

Details of the specification, related statistics, and charts are given in Appendix B. All variables are aggregated to quarterly frequency and enter in first differences, with the exception of consumer confidence, which is a stationary variable. We estimate the equation by ordinary least squares (OLS) with robust standard error calculations over the period from 1986Q3 to 2004Q2 (72 quarters). The equation has an adjusted-R^2 of 0.77 and appears to track turning points fairly well (see Figure B-1). All coefficients have expected signs. Various residual diagnostic tests reveal no discernible specification errors. The root mean squared error of the predictions produced by this model is 1.21 percentage points (quarter-over-quarter annualized rate).[11] This implies a 90 per cent confidence band of ±1.97 percentage points around the model's point forecasts, which can be sizable considering that the mean of the quarterly GDP growth rate since 1986 has been 2.76 per cent (seasonally adjusted annual rates).

---

[9] Eliminating all insignificant variables simultaneously with each iteration does not change the resulting specification.

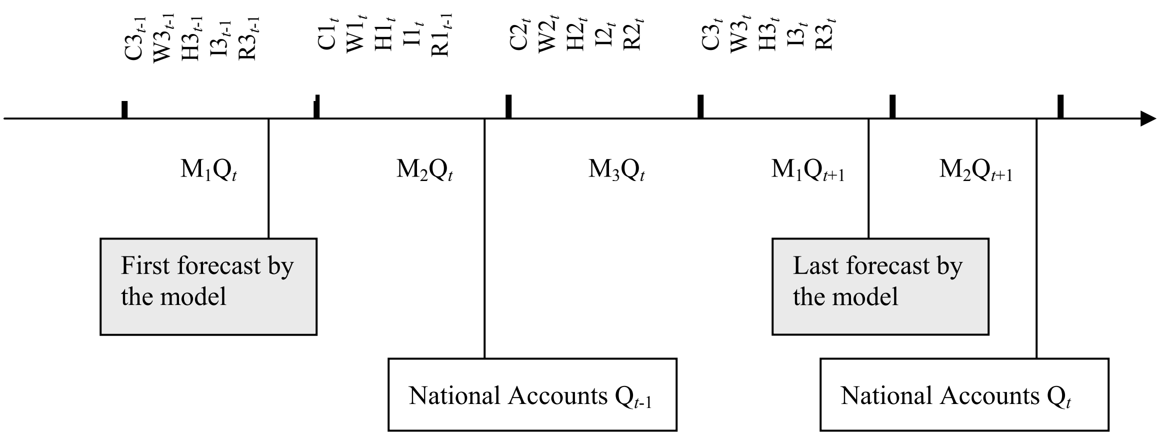
[10] Our sample starts in 1986Q3; retail sales data are not available prior to that date.

[11] This compares with a standard deviation of real GDP growth of 2.67 percentage points over the sample period.

Nevertheless, in this paper, we examine only point forecasts, leaving the issue of forecasting uncertainty to future research.

Overall, the variables selected in the general-to-specific approach signify the importance of consumption indicators, hours worked, and U.S. industrial production data in generating an early estimate of Canadian GDP growth. The publication sequence of the indicators incorporated in the model relative to the release of the National Accounts is depicted in Figure 1. The model's first forecast can be produced around the third week of the first month of the quarter to be forecasted, following the release of all monthly indicators from the preceding quarter.[12] When the full quarterly information on the indicators becomes available, a final forecast based on the model can be made roughly five weeks in advance of the first official release of GDP. While all the series in the bridge equation are potentially subject to some revision, GDP, retail sales, and U.S. industrial production are revised more frequently and thus make the equation more volatile in a real-time environment. Section 7 addresses the impact of data revisions to these series on the model. The analysis in sections 5 and 6 is based on the data as of November 2004.

Figure 1. Schedule of data releases for current quarter



Legend:
C – Consumer confidence W – Total hours worked H – Housing starts
I – U.S. industrial production R – Retail sales
M – Month in the quarter (1, 2, or 3) Q – Quarter
Numbers indicate the month's sequence in a quarter; subscripts indicate the relative quarter. Note that since retail trade is lagged, R2 in fact refers to the release of the first month of the quarter's data.

[12] Quarterly retail sales in this study are calculated as the average of the three previous months at the end of the quarter (e.g., Q4 retail trade is the average of September, October, and November's data), thus eliminating the need to wait for data for the third month of the quarter (December). This measure compares favourably with the conventional quarterly average in the model selection exercise.

5. Predicting the Indicator Variables

The model specified in the previous section uses all three months of the monthly indicator data for the forecasted quarter. In practice, however, we would like to use the model to predict GDP when data for the quarter are only partially available. The worst-case scenario is when no monthly information is available; in this case, the paths of the indicators in the following three months must be predicted. In a real-time application, separate models are required to generate forecasts of the indicators themselves. The quality of the GDP forecasts, therefore, also depends on the performance of the monthly "satellite" models. Each satellite model's forecasts for the remaining months of the quarter are generated at the time of every monthly data release to reflect the most recent information available.

We consider three models used in previous studies to generate monthly forecasts for four indicators (consumer confidence, housing starts, total hours worked, and retail sales).[13] The first is a naïve random-walk-in-growth-rates model (RWG), which predicts growth to be the same rate as that of the last observation. For example, if housing starts grew 2 per cent in the first month of a quarter relative to the previous month, its predicted growth rates in the remaining two months of the quarter would each be set at 2 per cent as well.[14]

The second is a rolling autoregressive model (AR) on monthly growth rates for each indicator, where the lag lengths (up to six) are selected according to the Akaike information criterion. Which lags are included and their estimated coefficients may change over time, as the model is re-estimated each month using a rolling data window of 10 years. We prefer this flexible approach, mainly because it incorporates more recent information and thus is likely better prepared to reflect changing dynamics of a time series. For a general comparison of rolling regressions and fixed-coefficient models, see Stock and Watson (1996).

The final model considered is a vector autoregressive model with parameters estimated using Bayesian procedures (BVAR), as described in Doan, Litterman, and Sims (1984). A well-known problem of unrestricted VARs with many endogenous variables (and thus an excessive number of parameters) is that estimated coefficients are often imprecise and not significantly different from zero, which consequently results in poor forecasting performance. The Bayesian VAR approach tackles this problem by imposing restrictions through prior probability distribution

[13] U.S. industrial production does not enter the quarterly GDP equation contemporaneously, and thus does not need to be forecast.
[14] While an RWG model implies an I(2) process for the variable in the long run, we find that in the short run it fits the data better than a traditional random-walk model.

functions.[15] Such prior distributions (assumed independent normal) reflect the forecaster's belief about the most likely values for the parameters in a VAR.[16] One frequently used prior distribution, the Minnesota prior, reflects the fact that most time series have one single unit root. It does this by assuming that the prior mean for the coefficient on the first lag of the dependent variable is unity while those of all other own- and cross-lags are zero.[17] Prior variances indicate the forecaster's confidence in the specification, with a smaller variance implying higher confidence. Generally speaking, one would assign a smaller variance to coefficients on cross-lag terms than on the dependent variable's own lags, in keeping with the random-walk hypothesis. On the other hand, longer lags are often deemed less important, and thus given a smaller variance around a zero mean.

In this study, to develop a Bayesian VAR for the four monthly indicators, we use priors that would produce a baseline forecast close to a random walk (in levels). The prior variances on the parameters are subsequently adjusted, one at a time, to improve the out-of-sample forecasting performance for each monthly indicator, relative to the baseline model (as in Ingenito and Trehan 1996).[18]

A priori, we expect the rolling AR and the BVAR approach to provide better forecasts for the monthly indicators than the naïve RWG approach. Indeed, this is the case. Table 1 shows root mean squared errors (RMSE) and mean absolute errors (MAE) of quarterly forecasts over the period from 1999Q3 to 2004Q2. Those generated by the RWG method based on up to two months of information are the highest, while those produced by the BVAR approach are generally the lowest. The best method for each set of forecasts is indicated by an asterisk. In many cases, having a second month of data reduces average forecast errors by at least half. Appendix C provides an example, showing the quarterly forecasts versus the actual values/growth rates of housing starts.

[15] An alternative to overcome the overfitting problem is to drop some endogenous variables from the VAR, but this implies a very strong conviction on the forecaster's part that the best coefficients for the excluded variables are zero, no matter what the historical data suggest. Such exclusion restrictions tend to be somewhat extreme and inflexible.

[16] The estimation of a BVAR typically involves maximizing the sample likelihood function weighted by the probability density function of the parameters.

[17] See Todd (1984) for a detailed description of the Minnesota prior.

[18] Details of the prior variance specification are available upon request.

Table 1. Performance of quarterly indicator forecasts: A comparison of three methods

		Consumer confidence[1]			Retail sales		
	Months available:	0	1	2	0	1	2
RMSE	RWG	5.93	3.71	1.04	3.10	2.13	0.62
	AR	3.07*	2.33	1.41	1.01*	1.09*	0.43*
	BVAR	3.64	2.09*	0.85*	1.19	1.19	0.54
MAE	RWG	4.97	2.71	0.80	2.61	1.62	0.51
	AR	2.15*	1.72	0.96	0.79*	0.91*	0.35*
	BVAR	3.03	1.60*	0.71*	0.97	1.00	0.38

		Housing starts			Work hours		
	Months available:	0	1	2	0	1	2
RMSE	RWG	33.75	24.86	9.73	2.28	2.78	0.91
	AR	8.75*	11.90	5.80	0.90*	1.24	0.49
	BVAR	12.19	10.48*	3.31*	1.03	0.84*	0.27*
MAE	RWG	26.32	22.19	7.54	1.83	2.39	0.68
	AR	6.78*	9.98	4.45	0.73*	1.02	0.36
	BVAR	10.02	8.64*	2.66*	0.82	0.69*	0.19*

1. Consumer confidence in levels, all others in growth rates (percentage points).

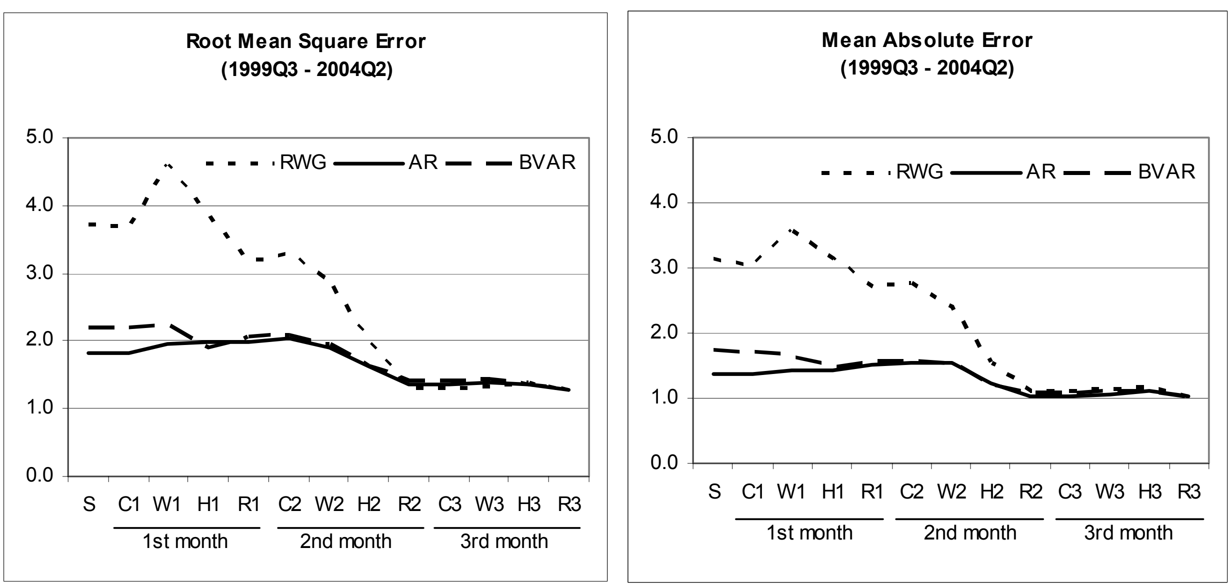
6. Out-of-Sample Forecasts

In this section, we examine the evolution of the forecasting performance of the GDP equation as more data become available during the quarter. The satellite models described in the previous section are used to fill in "missing" observations when data for the monthly indicators are not yet available for the entire quarter. Since we are using a recent vintage of data to mimic the actual ex ante forecasting process, results obtained from that vintage are denoted "pseudo-real time." This exercise helps us answer two questions. First, when can the model start producing useful (when compared with other forecasting models) estimates of GDP for the current quarter? Second, which indicators are relatively more useful in reducing errors in predicting GDP? In order to understand the benefits and limitations of adopting an indicator-based approach to forecasting GDP, we also compare the performance of our model with that of several short-term forecasting models.

6.1 Forecasting GDP as the quarter evolves

Suppose that we want an estimate of Q4 GDP as early as the third week of November. By this time we have data on consumer confidence, total hours worked (from the Labour Force Survey), housing starts, and U.S. industrial production through October, as well as retail sales through September. Thus, we need to forecast November and December values for consumer confidence, total hours worked, and housing starts, as well as October and November values for retail sales. The quarterly averages of these actual and forecast values then feed into the GDP equation to produce an estimate of real output for Q4. A week or two afterwards, when consumer confidence data for November are released, we can obtain a second forecast of GDP. The difference between the two forecasts is that the second requires only a one-month-ahead forecast for consumer confidence, whereas the first requires a two-month-ahead forecast. For all other indicators, two-month-ahead forecasts are necessary for the model to function. As the quarter progresses, the need to fill in “missing months” decreases. The GDP equation produces a new estimate after each indicator release until all the necessary information is received. Repeating this forecasting exercise for every quarter between 1999Q3 and 2004Q2, we obtain 13 sets of GDP forecasts, with each set based on a certain amount of information within the quarter. Figure 2 presents the RMSE (left panel) and the MAE (right panel) for each set. The three lines in the charts represent the three methods of forecasting monthly indicators, namely the RWG approach, the rolling AR model, and the BVAR model. The x-axes are labelled from left to right with the sequential indicator releases.

Figure 2. Prediction errors of GDP over the course of the quarter



Note: S is the starting point when zero months of information are available; C1 is when consumer confidence is released for the 1st month; W3 is the date when hours worked become available for the 3rd month; H is housing starts and R stands for retail sales (again, lagged one month). Errors are expressed in quarter-over-quarter annualized growth rates.

All the lines in Figure 2 exhibit a downward slope starting with the release of the second month of data. Downward slopes indicate a reduction in average prediction errors—quite natural as more data become available and the uncertainty associated with longer forecast horizons diminishes. The improvements with data releases in the AR and BVAR RMSEs and MAEs (around 30 per cent and 40 per cent, respectively) are substantial. The initial GDP forecasts based on the RWG approach result in a root mean squared error of 3.73 percentage points, significantly higher than the RMSE of forecasts given by the rolling AR (1.81 per cent) and the BVAR (2.21 per cent) models (see Figure 2). This gap in performance persists until the release of retail sales data for the second month. In fact, adding the third month of data does not seem to improve the results considerably for any of the models, but this is not surprising. Rünstler and Sédillot (2003) approximate the quarterly growth rate of a variable by decomposing it into a weighted average of the monthly growth rates, showing that the growth rate of the third month of the quarter contributes only 11 per cent to the total quarterly growth rate. Of all three methods for forecasting monthly indicators, the RWG approach fares worst when using the first two months of data releases.

Table 2 shows the contribution by indicator to the reduction in forecast errors where the total reduction is the difference in the error measure (root mean squared error or mean absolute error) between the last and the first GDP forecasts. A negative sign indicates an addition to the forecast error. The most important data releases are retail sales and housing starts in both the first and second months.[19] Interestingly, the retail trade data in the first month do not reduce the RMSE or MAE for either the rolling AR or the BVAR model. In all models, the second-month release of housing starts alone generally accounts for more than one-third of the reduction in RMSE and MAE. While these particular data releases prove essential to the performance of the model's forecast, the following caveats should be kept in mind. First, the results are obtained for the five-year period between 1999Q3 and 2004Q2, which is not long enough to allow one to draw generalized conclusions. Second, a small number of potentially useful indicators are excluded from our study for the sake of a timely forecast. Some of these indicators may model the "underlying dynamics" of the economy better than the ones used here. Finally, a different method for forecasting monthly indicators, particularly one that improves the two-step-ahead forecasts, would likely reduce the benefit of having the actual data for the second month, thus altering the relative rankings among the indicators.

[19] For retail sales, this corresponds to the last calendar month of the previous quarter and the first calendar month of the current quarter.

Table 2. Contribution to forecast error reduction by indicator

Indicators	RWG RMSE	RWG MAE	AR RMSE	AR MAE	BVAR RMSE	BVAR MAE
C1	2.9%	4.6%	-2.5%	3.3%	1.6%	3.5%
W1	-37.6%	-25.0%	-25.0%	-15.2%	-5.8%	6.9%
H1	30.1%	19.6%	-3.5%	0.0%	36.4%	27.4%
R1	27.1%	20.1%	-2.0%	-26.8%	-16.6%	-15.5%
C2	-4.7%	-2.8%	-7.4%	-5.5%	-2.9%	0.2%
W2	16.6%	18.1%	23.4%	3.6%	14.6%	4.4%
H2	35.9%	41.3%	49.6%	82.8%	33.5%	48.4%
R2	28.3%	20.4%	50.4%	52.6%	24.2%	17.2%
C3	-0.1%	-0.3%	2.2%	2.0%	0.9%	0.2%
W3	-1.4%	-2.1%	-7.2%	-7.1%	-4.4%	-6.5%
H3	-1.5%	-0.3%	5.1%	-12.2%	8.5%	1.8%
R3	4.4%	6.5%	16.9%	22.3%	10.1%	12.1%
Total reduction (percentage points)	2.47	2.11	0.54	0.36	0.95	0.72

Notes:
1. C stands for consumer confidence, W for total hours worked, H for housing starts, and R for retail sales. Numbers indicate how many months of the quarter are available.
2. A negative sign indicates an addition to the forecast error.

6.2 Comparison with other models

After examining the performance of the GDP bridge equation at various points in time during the quarter, it is beneficial to compare our results with those of other models. Most studies show comparisons with an autoregressive model. This comparison is convenient but not as meaningful to practitioners, who prefer to compare new models with existing ones. At the Bank of Canada, the staff responsible for current analysis use several models to support their short-term economic forecasting. Since our intention of developing an indicator-based model is to enhance the current pool of models, we compare our approach with existing models, putting emphasis on forecasting performance.

6.2.1 Benchmark models

A total of five sets of forecasts for one-quarter-ahead GDP growth rates from 1999Q3 to 2004Q2 are obtained from five benchmark models. Duguay's (1994) model is an IS curve augmented with real commodity prices. Duguay Adjusted is an updated version of Duguay (1994) that includes a lag of the dependent variable and the change in the consumer confidence index. Murchison (2001) provides an equation for forecasting the Canadian output gap (an estimate of aggregate demand relative to the productive capacity of the Canadian economy). This model is commonly referred to as NAOMI (North American Open-economy Macroeconomic Integrated model). We construct the forecast for GDP by adding the predicted value of the output gap to estimated potential output. We also use a rolling autoregressive model (Quarterly AR) for quarterly GDP, with its lag structure determined by the data using the Akaike information criterion. Its estimation is based on a rolling-data window of ten years. Finally, we consider a threshold autoregressive model (TAR) that captures some non-linearity in the data, also estimated with a ten-year rolling window. The TAR model allows for two different sets of coefficients and lag structures (regimes), depending on whether past growth observations are below or above an econometrically determined threshold level. The benchmark models perform poorly with a sample restricted to the same range as our indicator model (i.e., 1986Q3 to 2004Q2), so we take a conservative approach and present only results in the unrestricted case.[20]

6.2.2 Comparison with benchmarks

One-quarter-ahead forecast errors suggest that the two time-series models (Quarterly AR and Threshold AR) perform quite well, while the macroeconomic ones underperform. With full quarterly information, our model tracks GDP growth and avoids large errors better than Duguay, Duguay Adjusted, and NAOMI. Still, each model experiences periods of idiosyncratic weakness, during which prediction errors are uncharacteristically large.

Table 3 shows the one-quarter-ahead RMSEs for the five benchmark forecasts (shaded), as well as for the monthly indicator model (based on the AR satellite model) with varying months of information. The rows are sequentially ordered based on the release dates of the information required to estimate the model, from earliest to latest. The notation provides the chronological release of the data for the monthly indicator model, as in Table 2 (e.g., the row “C2” shows the RMSE from the bridge equation run at the time that consumer confidence is released for the second month of the quarter). The Duguay Adjusted, NAOMI, Quarterly AR, and TAR models, however, can all be initially run at the same time, since they all rely on lagged GDP, which is released at this time.

[20] There is no presumption that the unrestricted case represents an optimal case in the sense that it yields the lowest forecast errors that can be achieved with the given model specifications.

Table 3. Root mean squared errors by release sequence

Model	RMSE
S	1.81
C1	1.82
W1	1.96
H1	1.98
R1	1.99
Duguay	1.89
C2	2.03
W2	1.90
H2	1.63
R2	1.36
Duguay Adjusted	1.71
NAOMI	1.85
Quarterly AR	0.81
TAR	1.03
C3	1.35
W3	1.39
H3	1.36
R3	1.27

The RMSEs for the macroeconomic models (Duguay, Duguay Adjusted, and NAOMI) are all much higher than those of the statistical forecasts (Quarterly AR, TAR). The monthly indicator model initially produces fairly high RMSEs. However, following the second-month release of housing starts (H2), three weeks before the majority of benchmark models can be run, the RMSE drops sharply and remains below those of the other macroeconomic models (as one would expect, since it contains more data). By the end of the quarter, it has fallen even further, though the accuracy of the indicator model never exceeds that of the statistical models.

While RMSEs help provide a simple way of comparing forecasting models, we deploy more rigorous statistical methods to assess forecast accuracy. We use a modified version of the Diebold-Mariano test to examine the equality of mean squared forecast errors in a small sample, as proposed by Harvey, Leybourne, and Newbold (1997, hereafter HLN). The third column in Table D-1 shows the *p*-value for the HLN test under the null hypothesis that the forecasts from the two models are statistically indistinguishable, where Model 1 denotes the indicator model with various amounts of monthly information and Model 2 denotes a benchmark. We also run forecast encompassing tests in the spirit of Harvey, Leybourne, and Newbold (1998). The forecasts generated by Model 1 are said to "encompass" those given by Model 2 if the latter embody no information absent in the former. In Table D-1, the last two columns indicate whether the null hypothesis that one model encompasses the other is rejected at a 5 per cent level.

According to the HLN test, there is generally no significant difference between the forecasts given by the indicator-based model with incomplete quarterly information and the macroeconomic models.[21] When the quarter is complete, the indicator-based model produces a set of forecasts significantly different than those of NAOMI at the 10 per cent level. However, the difference between the indicator model and the Duguay and Duguay Adjusted remains insignificant. The encompassing tests, however, offer evidence that the indicator-based model (with two or more months of information) is more efficient than the three macroeconomic models.

The HLN tests reject the null hypothesis that forecasts from the monthly indicator model are equivalent to those from the Quarterly AR model, regardless of how many months of data are used. The corresponding encompassing test results indicate that there is useful information in the AR forecasts beyond what is contained in the indicator-based forecasts, and that the converse is not true. Evidence thus far seems to suggest that the Quarterly AR model yields better forecasts than the monthly indicator-based model over the period from 1999Q3 to 2004Q2. The TAR model also appears to outperform the indicator models in a number of cases.

Finally, the monthly indicator model with zero months of information does not compare favourably with the benchmarks. This is hardly surprising, given the poor performance of the satellite models (RWG, rolling AR, and BVAR) for forecasting horizons beyond two months (see Tables 1 and 2). However, as Table 3 shows, the fact that most of the benchmark models cannot be run until late in the quarter implies that there is still value to the less-accurate early-quarter forecasts produced by the monthly indicator model.

We conclude this section with some caveats. First, we use five years of data to evaluate the models. This sample, while short, contains some unusual episodes for the Canadian economy (the 11 September 2001 shock and the 2003 dip caused by SARS, BSE, and the Ontario power outage).[22] It is not clear whether these anomalies could affect the relative performance of the various models. Second, the analysis thus far is based on the vintage of data available as of November 2004. The real-time performance of the models could differ.

[21] The one exception is for the RWG model with zero months of information, when compared with the Duguay Adjusted benchmark. It is only just barely above a 10 per cent significance level, however.

[22] The SARS virus significantly impacted the tourism industry, the discovery of BSE (or mad cow disease) in an Albertan cow crippled cattle trade, and the power outage that occurred in Northeastern North America hindered production in Ontario for a number of days.

During our forecast period, the AR and the TAR models contain only one lag of GDP. Since the prior quarter's estimate of GDP is revised in each release of the National Accounts, real-time forecasts from the Quarterly AR and TAR models may be worse than reported in this section. Similar problems may exist for other benchmark models. While the problem of outliers is hard to avoid, given the limited sample size, we address real-time data issues in the following section.

7. Real-Time Analysis

Significant revisions to variables in the bridge equation suggest a need to conduct a separate real-time analysis.[23] We focus on vintage data for quarterly GDP and two of the five monthly variables—retail sales and U.S. industrial production—both of which are subject to more frequent revision than the rest. Survey-based consumer confidence is never revised, revisions to housing starts occur only for the months within the current reporting quarter and are usually modest, and the labour force data, which contain hours worked, are infrequently revised (usually only the result of re-benchmarking).[24] Since the latter two monthly variables are subject to minor revisions, the results presented in this section provide only a rough approximation of the impact of using real-time data.

The GDP vintages are taken from an archive at the Bank of Canada, the U.S. industrial production vintages are from the "Real-Time Data Set" maintained at the Federal Reserve Bank of Philadelphia, and the retail sales vintages are the authors' calculations based on nominal vintages supplied by Statistics Canada. Since the deflators for retail sales are normally derived from components of the seasonally adjusted consumer price index, and since only seasonal adjustment factors of these components are revised over time, we assume that the revision pattern of real retail sales reflects that of the nominal series.

7.1 Descriptive statistics

Table E-2 (in Appendix E) shows standard summary statistics for the first release, current estimate (as of November 2004), and cumulative revisions between the two estimates for GDP, retail sales, and U.S. industrial production. The table also reports the mean and maximum gaps (i.e., the difference between the largest and smallest values of a particular observation for a given time period, across all vintages of data). Figure E-1 plots the current

[23] Appendix E contains definitions related to vintage data.
[24] Real-time data for the latter two series have proven difficult to find, which is another reason they are not treated in real time in this section. Benchmarking revisions are occasionally performed when the need arises to redefine or revise data series over longer historical periods.

estimates of the three series against their first estimates for the period from 1998Q4 to 2004Q2. Several features of the revisions are notable. For example, first-releases (i.e., Statistics Canada's initial estimate) of GDP growth are on average underestimated by 0.44 percentage points (annualized). This is also evident from the first chart in Figure E-1, where a large cluster of the dots appears above the diagonal. Second, average monthly revisions to retail sales are larger in percentage terms than those made to U.S. industrial production. However, the mean of retail sales revisions is likely exaggerated by a few particularly large revisions, as the distribution of the revisions is positively skewed. Indeed, the median revision is much smaller than the mean. Third, some of the revisions are fairly large. The maximum revisions to retail sales and U.S. industrial production are twice as large as the standard deviations of the respective series. Extremely large revisions to GDP are rare, but some are in the order of magnitude of one standard deviation. For all three series, the gap between the extreme estimates of a particular observation across all vintages is typically greater than the cumulative revision. This implies that later vintages partially reverse previous revisions.

7.2 Procedure for real-time analysis

Since the revisions to three important variables in the bridge equation are often substantial, we implement the following algorithm to mimic the actual ex ante forecasting process.

1. Estimate the coefficients of the bridge equation using only the data available at the time the forecast would have been made.
2. Construct the GDP growth rate forecast for the quarter of interest. Extrapolate missing values of monthly indicators in the quarter following the rolling AR approach, as described in section 5, using only data that were available at the time of the forecast.[25]
3. Compare the forecast with both the first and the current (2004Q2 vintage) estimates of GDP growth for that quarter.
4. Repeat steps 1 through 3 following each indicator release from the beginning of 1999Q3 to 2004Q2.
5. Compute summary measures of forecast accuracy (RMSE and MAE) from the resulting sets of out-of-sample forecasts, and compare them with those obtained from the pseudo-real-time forecast. Significant differences imply that the bridge equation is sensitive to the choice between real-time and revised data.
6. Generate real-time forecasts using the Quarterly AR model—the best model for short-term forecasting of GDP according to the pseudo-real-time analysis. Compare the results with those from the real-time bridge equation to determine whether the use of real-time data alters the relative ranking between the two.

[25] This particular approach is selected for its simplicity and its performance in predicting monthly indicators that is no worse than the other two (RWG and BVAR).

7.3 Main results

Figure E-2 in the appendix shows the progression of the root mean squared errors and mean absolute errors as more information becomes available through the quarter, for three different scenarios:

1. *pseudo*-real-time forecasts from the bridge equation assessed against *current* GDP estimates (i.e., those presented in section 6),
2. real-time forecasts from the bridge equation assessed against *current* GDP estimates, and
3. real-time forecasts from the bridge equation assessed against *initial* GDP estimates.

The two conventional measures of forecast accuracy based on real-time forecasts follow a very similar path to those based on pseudo-real-time forecasts, regardless of the GDP measure against which they are assessed. The RMSEs (and MAEs) of real-time forecasts are slightly higher than those of pseudo-real-time forecasts when both are compared against the current estimates of GDP. Using the first estimates of GDP yields RMSEs (and MAEs) of real-time forecasts lower than the RMSEs of pseudo-real-time forecasts. The HLN test indicates, however, that the differences between real-time forecasts and pseudo-real-time ones are generally not significant (see the first two panels in Table E-3 in the appendix). The assessment of the forecast accuracy of the bridge equation based on revised data does not lead to a materially different conclusion for the sample period we study (1999Q3 to 2004Q2).[26]

The performance of the Quarterly AR model changes dramatically in the real-time context. As shown in the last two panels of Table E-3, not only does the HLN test confirm that in most cases the forecasts made by the Quarterly AR model are significantly different from those of the monthly indicator model, but the former are less accurate as well. This is not surprising, given the tendency for the dynamically specified AR model to retain only one lag and thus expose itself to revisions to the previous quarters' GDP. These results highlight the possibility that relative performance among short-term forecasting models can change with the choice of data vintage. We suspect some deterioration in the performance of the macroeconomic benchmark models as well, but to a lesser extent than the Quarterly AR model. Future research could investigate this hypothesis.

[26] We investigate the possibility that Q1 benchmark revisions (i.e., substantial revisions often made to the previous four years of data due to changes in methodology, definitions, and further information) might affect the real-time performance of the model. The results suggest, somewhat surprisingly, that the RMSEs of Q1 forecasts are relatively low, and that it is the Q4 forecasts that contain the highest RMSEs. One possible explanation is that the benchmark revisions contain revisions that are consistent among different indicators, while for the Q4 forecast certain revisions have been withheld until a final reconciliation with the Q1 National Accounts release in the following year.

The real-time forecasting exercise for the indicator model does not fully meet the criteria set out in Fair and Shiller (1990). In particular, we identify the variables in the bridge equation using the November 2004 vintage of data over the full sample (as in section 4), rather than the data for the period prior to the time of forecast. However, the coefficients on these variables are estimated on a real-time basis. As mentioned above, the short sample of real-time data we have for retail sales (starting in 1997) precludes a more thorough analysis. In addition, the insensitivity of the indicator model to the choice of data vintage may have been a result of the lack of sizable revisions in the recent sample period. Finally, a small sample size, such as the one in our study, means that statistical power may be weak and that the conclusions should not be generalized to longer samples without additional research.

8. Conclusion

This paper presents a model for current-quarter real GDP growth using monthly indicators. Our model produces a GDP estimate as early as the first month of the reference quarter. Its forecasting accuracy generally improves with incremental monthly data releases. The final estimate from the model, available five to six weeks before the release of the National Accounts, compares favourably with those given by several macroeconomic models.

The use of real-time data—reflecting revisions to three variables in the indicator model—does not result in average forecast errors that differ significantly from those produced in the pseudo-real-time exercise, when only the latest available vintage is used. This conclusion holds true whether the real-time forecast errors are assessed against first estimates of GDP growth or the most recent estimates. The Quarterly AR model—the best model in the pseudo-real-time forecasting exercise—loses its advantage when real-time data are used instead, suggesting that it has high sensitivity to data revisions as a result of its lag-length structure.

The model may evolve considerably in the future as the economy and available data (such as new series and earlier release dates) change. Nevertheless, this paper suggests that incorporating monthly data into a quarterly model is worthwhile and that this framework should be used as a tool in the short-term forecasting of output growth in the Canadian economy.

References

Babineau, B. and N. Braun. 2003. "Output Growth Forecasts in Real-Time." Department of Finance Canada Working Paper No. 2003-07.

Cole, R. 1969. "Data Errors and Forecasting Accuracy." In *Economic Forecasts and Expectations: Analyses of Forecasting Behavior and Performance*, edited by J. Mincer, 47–82. New York: National Bureau of Economic Research.

Coutiño, A. 2005. "On the Use of High-Frequency Economic Information to Anticipate the Current Quarter GDP: A Study Case for Mexico." *Journal of Policy Modeling* 27: 327–44.

Croushore, D. and T. Stark. 2003. "A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter?" *Review of Economics and Statistics* 85(3): 605–17.

Denton, F.T. and J. Kuiper. 1965. "The Effect of Measurement Errors on Parameter Estimates and Forecasts: A Case Study Based on the Canadian Preliminary National Accounts." *Review of Economics and Statistics* 47(2): 198–206.

Diebold, F.X. and R.S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics* 13: 253–63.

Diron, M. 2005. "Short-Term Forecasts of Euro Area Real GDP Growth: An Assessment of Real-Time Performance Based on Vintage Data." European Central Bank Working Paper No. 622.

Doan, T., R. Litterman, and C. Sims. 1984. "Forecasting and Conditional Projection Using Realistic Prior Distributions." *Econometric Reviews* 3(1): 1–100.

Duguay, P. 1994. "Empirical Evidence on the Strength of the Monetary Transmission Mechanism in Canada." *Journal of Monetary Economics* 33: 39–61.

Fair, R.C. and R.J. Shiller. 1990. "Comparing Information in Forecasts from Econometric Models." *American Economic Review* 80(3): 375–89.

Harvey, D.I., S.J. Leybourne, and P. Newbold. 1997. "Testing the Equality of Prediction Mean Squared Error." *International Journal of Forecasting* 13: 281–91.

———. 1998. "Tests for Forecast Encompassing." *Journal of Business and Economic Statistics* 16(2): 254–59.

Hendry, D.F. and G.E. Mizon. 1978. "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study by the Bank of England." *Economic Journal* 88: 549–63.

Ingenito, R. and B. Trehan. 1996. "Using Monthly Data to Predict Quarterly Output." Federal Reserve Bank of San Francisco *Economic Review* 3: 3–11.

Kitchen, J. and R. Monaco. 2003. "Real-Time Forecasting in Practice: The U.S. Treasury Staff's Real-Time GDP Forecast System." *Business Economics* (October): 10–19.

Klein, L.R. and E. Sojo. 1989. "Combinations of High and Low Frequency Data in Macroeconometric Models." In *Economics in Theory and Practice: An Eclectic Approach*, edited by L.R. Klein and J. Marquez, 3–16. Norwell, MA: Kluwer Academic Publishers.

Koenig, E.F. 2002. "Using the Purchasing Managers' Index to Assess the Economy's Strength and the Likely Direction of Monetary Policy." Federal Reserve Bank of Dallas *Economic and Financial Policy Review* 1(6): 1–14.

Koenig, E.F., S. Dolmas, and J.M. Piger. 2001. "The Use and Abuse of 'Real-Time' Data in Economic Forecasting." Federal Reserve Bank of St. Louis Working Paper No. 2001-015A.

Liou, R.-W. and C.-H. Shen. 1996. "The Use of High Frequency Data to Improve Macroeconometric Forecast." *International Economic Journal* 10(2): 65–83.

Miller, P.J. and D.M. Chin. 1996. "Using Monthly Data to Improve Quarterly Model Forecasts." Federal Reserve Bank of Minneapolis *Quarterly Review* 20(2): 16–33.

Murchison, S. 2001. "A New Quarterly Forecasting Model. Part II: A Guide to Canadian NAOMI." Department of Finance Canada Working Paper No. 2001-25.

Rünstler, G. and F. Sédillot. 2003. "Short-Term Estimates of Euro Area Real GDP by Means of Monthly Data." ECB Working Paper No. 276.

Stock, J.H. and M.W. Watson. 1996. "Evidence on Structural Instability in Macro Time Series Relations." *Journal of Business and Economic Statistics* (January): 25–38.

Todd, R.R. 1984. "Improving Economic Forecasting with Bayesian Vector Autoregression." Federal Reserve Bank of Minneapolis *Quarterly Review* 8: 18–29.

Trehan, B. 1992. "Predicting Contemporaneous Output." Federal Reserve Bank of San Francisco *Economic Review* 2: 3–11.

Zadrozny, P.A. 1990. "Estimating a Multivariate ARMA Model with Mixed-Frequency Data: An Application to Forecasting U.S. GNP at Monthly Intervals." Federal Reserve Bank of Atlanta Working Paper No. 90-6.

Appendix A. Monthly Indicators Considered for Inclusion in the Quarterly GDP Model

Indicator	Transformation 1*	Transformation 2**	Publication lag (days after reference month)
Overnight rate	AVE	DLV	
T-bill 3 month	AVE	DLV	
T-bill 1 year	AVE	DLV	
Government of Canada Bond yield (10 year)	AVE	DLV	
TSX Index	AVE	DLN	
Bank of Canada Commodity Price Index	AVE	DLN	
Bank of Canada Non-Energy Commodity Price Index	AVE	DLN	
CAD/US exchange rate	AVE	DLN	
Consumer confidence (Conference Board of Canada)	AVE	LV	1 to 10
US Purchasing Manager's Index: Mfg. Total	AVE	LV	1 to 10
US Purchasing Manager's Index: Mfg. Production	AVE	LV	1 to 10
New Motor Vehicle Sales (G&M)	AVE	DLN	1 to 10
LFS: Total employment	AVE	DLN	1 to 10
LFS: Full-time employment	AVE	DLN	1 to 10
LFS: Per employee hours worked	AVE	DLN	1 to 10
LFS: Total hours worked	AVE	DLN	1 to 10
LFS: Public sector hours worked	AVE	DLN	1 to 10
LFS: Unemployment rate	AVE	LV	1 to 10
Housing Starts	AVE	DLN	1 to 10
MLS Existing Home Sales - 25 Majors Market	AVE	DLN	11 to 20
US Industrial Production: Total	AVE	DLN	11 to 20
US Industrial Production: excl. M.V. and Parts	AVE	DLN	11 to 20
Composite Index	AVE	DLN	21 to 30
M2	AVE	DLN	21 to 30
Building permits -values	AVE	DLN	31 to 40
Retail Trade (K$)	AVE	DLN	51 to 60
Retail Trade (K$), previous three months average at end of quarter	SP	DLN	21 to 30
Gross Domestic Product		DLN	51 to 60

*Transformation to yield quarterly figures
AVE = convert to quarterly average
END = take the end-of-quarter value as the quarterly value
SP = take the average of the three previous months at the end of the quarter

** Transformation to render stationarity
LV = level
DLV = first difference in level
DLN = first difference in the log

Appendix B. Summary Statistics of the Quarterly Bridge Equation

(Sample period: 1986Q3 to 2004Q2, all data non-annualized)

Table B-1. OLS regression statistics (dependent variable = GDP)

Variable	Coefficient	Standard error	*T*-statistics[3]
Consumer confidence[1]	0.002	0.000	4.057
Total hours worked	0.176	0.041	4.298
Total hours worked(*t*-1)	0.214	0.047	4.564
Housing starts	0.025	0.004	6.225
Retail sales[2]	0.146	0.021	6.801
US industrial production(*t*-1)	0.271	0.057	4.765
US industrial production(*t*-3)	-0.155	0.059	-2.615
GDP(t-3)	0.231	0.059	3.935

1. Consumer Confidence in levels, all other variables in quarter-over-quarter growth rates.
2. Quarterly level of Retail Sales is calculated as the average of the three previous months at the end of the quarter. For example, the level of Q3 is based on the average of June, July, and August.
3. All significant at 1% level.

Table B-2. Goodness-of-fit

Statistics	Value
Adjusted R^2	0.767
RMSE (in-sample) [1]	0.302

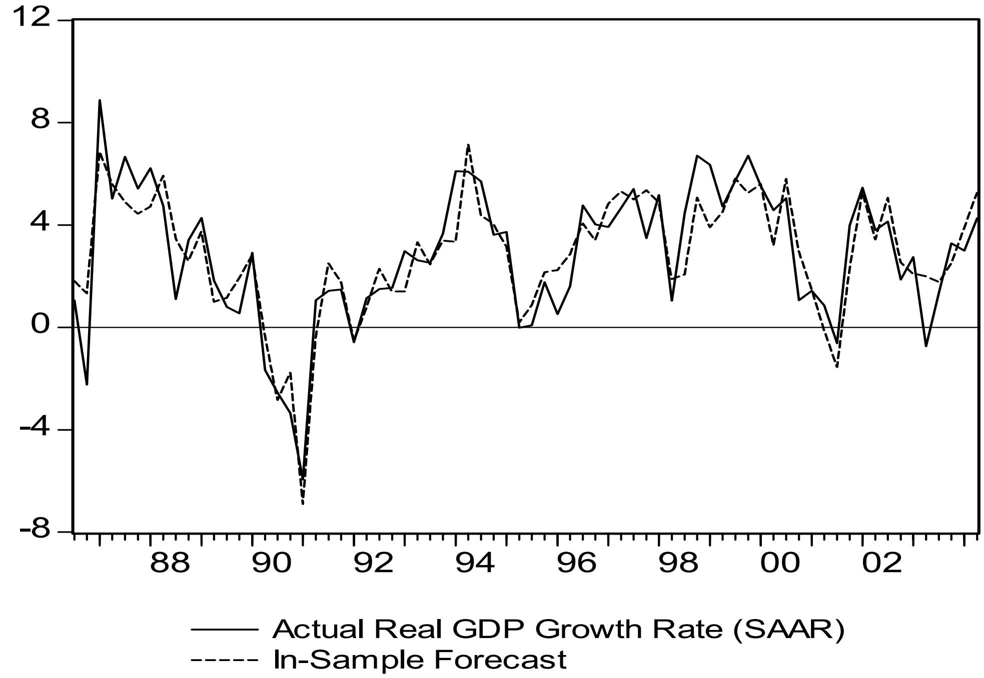
1. In comparison, the standard error of the dependent variable is 0.660.

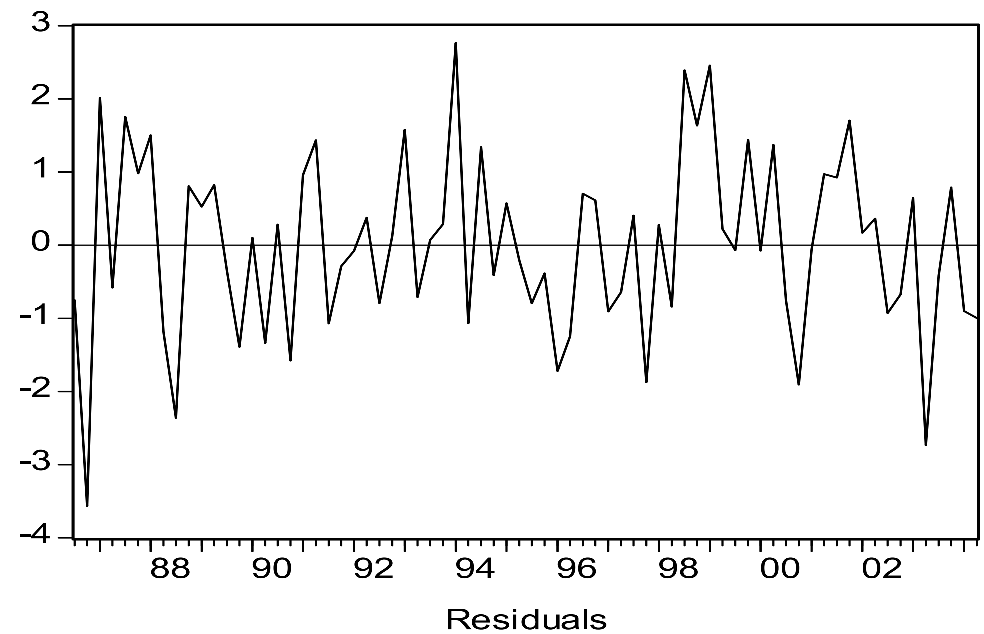
Table B-3. Residual diagnostics (lags in parentheses, unless otherwise noted)

Statistics	Value	*P*-value
Jarque-Bera normality test	0.309	0.857
Durbin Watson	2.022	
Breusch-Godfrey serial correlation LM(1)	0.462	0.497
Breusch-Godfrey serial correlation LM(4)	5.456	0.244
ARCH(1)	1.075	0.300
White heteroscedasticity test	44.480	0.451
Ramsey regression specification error test(2) [1]	1.186	0.553
Chow forecast test (from 1999Q3 to 2004Q2)	16.853	0.464

1. Number in parentheses indicates the number of fitted terms included in the test.

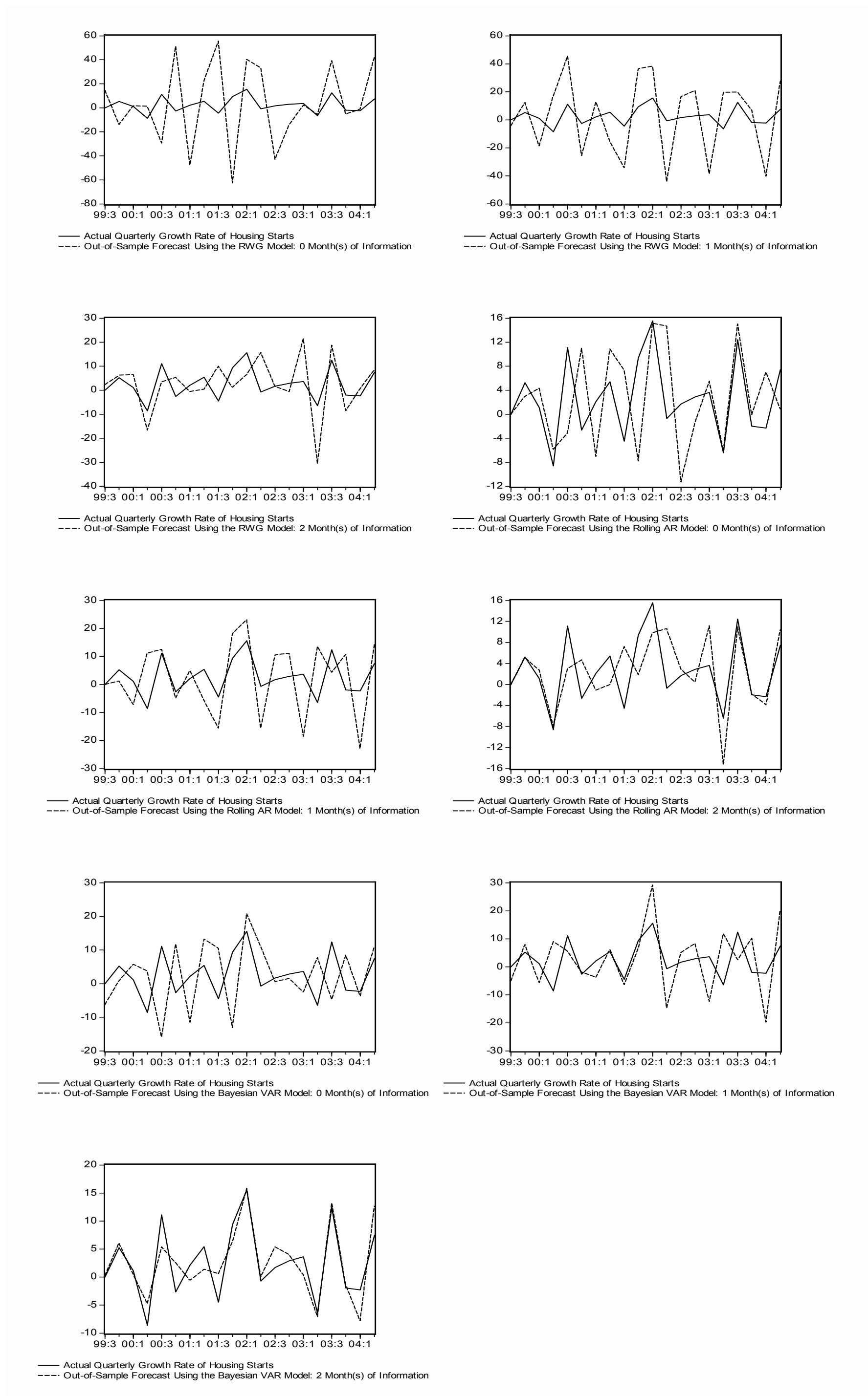
Figure B-1. In-sample forecasts of the GDP equation with full quarterly information





Appendix C. Forecasts of Housing Starts

Figure C-1. Housing starts: forecast vs. actual



Appendix D. Comparison with Benchmark Models

Table D-1. Test for equality of forecast accuracy

Model 1[1]	Model 2	HLN *p*-value[2]	H_0 = Model 1 encompasses Model 2[3]	H_0 = Model 2 encompasses Model 1[3]
RWG, 0 months	Duguay	0.108		
	Duguay Adjusted	0.099 *		
	NAOMI	0.105		
	Quarterly AR	0.070 *		
	TAR	0.077 *		
Rolling AR, 0 months	Duguay	0.694		Reject
	Duguay Adjusted	0.782		
	NAOMI	0.924	Reject	
	Quarterly AR	0.103		
	TAR	0.165		
BVAR, 0 months	Duguay	0.102	Reject	
	Duguay Adjusted	0.117	Reject	
	NAOMI	0.289	Reject	
	Quarterly AR	0.070 *	Reject	
	TAR	0.101		
RWG, 1 month	Duguay	0.103		Reject
	Duguay Adjusted	0.106		Reject
	NAOMI	0.146		
	Quarterly AR	0.073 *		
	TAR	0.084 *		
Rolling AR, 1 month	Duguay	0.542	Reject	Reject
	Duguay Adjusted	0.190	Reject	Reject
	NAOMI	0.512	Reject	Reject
	Quarterly AR	0.036 **	Reject	
	TAR	0.064 *	Reject	
BVAR, 1 month	Duguay	0.459	Reject	Reject
	Duguay Adjusted	0.258	Reject	Reject
	NAOMI	0.539		Reject
	Quarterly AR	0.072 *	Reject	
	TAR	0.110		
RWG, 2 months	Duguay	0.279		Reject
	Duguay Adjusted	0.384		Reject
	NAOMI	0.130		Reject
	Quarterly AR	0.065 *	Reject	
	TAR	0.176		
Rolling AR, 2 months	Duguay	0.212		Reject
	Duguay Adjusted	0.315		Reject
	NAOMI	0.126		Reject
	Quarterly AR	0.014 **	Reject	
	TAR	0.141	Reject	
BVAR, 2 months	Duguay	0.196		Reject
	Duguay Adjusted	0.310		Reject
	NAOMI	0.173		Reject

cont'd. . .

Table D-1 (Concluded)

BVAR, 2 months	Quarterly AR	0.038 **	Reject	
	TAR	0.213	Reject	
Full quarterly information	Duguay	0.165		Reject
	Duguay Adjusted	0.198		Reject
	NAOMI	0.081 *		Reject
	Quarterly AR	0.053 *	Reject	
	TAR	0.479	Reject	

1. Model 1 refers to the monthly indicator-based model.
2. HLN value is the Diebold-Mariano test result corrected for small-sample bias, proposed by Harvey, Leybourne, and Newbold (1997). The null hypothesis is: $mean(e^2_{1,t}) = mean(e^2_{2,t})$. One '*' indicates result is significant at the 10% level, two '**' at the 5% level.
3. Only indicated where the null hypothesis is rejected at the 5% level.

Appendix E. Real-Time Analyses

The terminology used in this paper is as follows:

A vintage for a variable, Y, is a time series of the variable available at any given time, S, in the past. Each vintage is customarily named for its last observation. For example, vintage 2004Q4 of GDP refers to the GDP time series that has its most recent available observation for 2004Q4 (which, in Canada, is released by the end of February 2005). Thus a new vintage contains not only the first estimate of a variable for the vintage date, but also potentially revised estimates for earlier dates.

In Table E-1, a total of T+1 vintages for Y are shown, with vintage T being the most recent, or *current vintage*. Estimates for $Y(0)$ through $Y(T)$ in this vintage are referred to as *current estimates*. In comparison, the estimates along the main diagonal (in bold type) are called *first estimates*. The first estimate of the growth rate of Y at time S is calculated as $(Y(S)/Y(S\text{-}1)\text{-}1)$ based on vintage S. The difference between the current and the first estimates of a particular dated observation, say $Y(0)$, indicates the cumulative magnitude of revision since $Y(0)$ is first released. This measure, however, may mask the true volatility of data revisions, if earlier revisions are at least partly reversed in a later vintage. Therefore, it is likely for the difference between largest and smallest realizations of $Y(0)$ across vintages to be bigger than that suggested by the cumulative revision.

Table E-1. Schema for data vintages

Vintage **0**	Vintage **1**	…	Vintage ***S***	…	Vintage ***T*-1**	Vintage ***T***
***Y*(0)**	*Y*(0)	…	*Y*(0)	…	*Y*(0)	*Y*(0)
	***Y*(1)**	…	*Y*(1)	…	*Y*(1)	*Y*(1)
		…	…	…	…	…
		…	*Y*(*S*-1)	…	…	…
			***Y*(*S*)**	…	*Y*(*S*)	*Y*(*S*)
				…	…	…
					***Y*(*T*-1)**	*Y*(*T*-1)
						***Y*(*T*)**

Table E-2. Descriptive statistics for revisions in real GDP and selected indicators (1998Q4 to 2004Q2)

Measured on the quarterly annualized growth rate of GDP and monthly growth rate of indicators

		Real GDP	Retail sales	US industrial production
Mean	First estimate	3.09	0.21	0.10
	Current estimate[1]	3.53	0.30	0.12
	Revision[2]	0.44	0.09	0.02
Median	First estimate	3.33	0.19	0.15
	Current estimate[1]	3.87	0.38	0.09
	Revision[2]	0.30	0.02	0.02
Max	First estimate	5.96	2.44	1.04
	Current estimate[1]	6.77	2.67	1.08
	Revision[2]	2.21	1.86	0.92
Min	First estimate	-0.80	-2.10	-1.07
	Current estimate[1]	-0.73	-1.54	-0.95
	Revision[2]	-1.53	-1.01	-1.09
Standard deviation	First estimate	1.80	0.96	0.49
	Current estimate[1]	2.24	0.92	0.49
	Revision[2]	1.05	0.53	0.37
Skewness	First estimate	-0.62	-0.05	-0.28
	Current estimate[1]	-0.33	0.26	-0.02
	Revision[2]	0.29	0.82	-0.24
Kurtosis	First estimate	2.51	2.77	2.66
	Current estimate[1]	2.12	2.85	2.01
	Revision[2]	2.35	4.04	3.80
Jarque-Bera probability	First estimate	0.43	0.92	0.54
	Current estimate[1]	0.56	0.67	0.24
	Revision[2]	0.70	0.00	0.29
Gap[3]	Mean	1.25	0.60	0.41
	Max	2.79	1.87	1.16

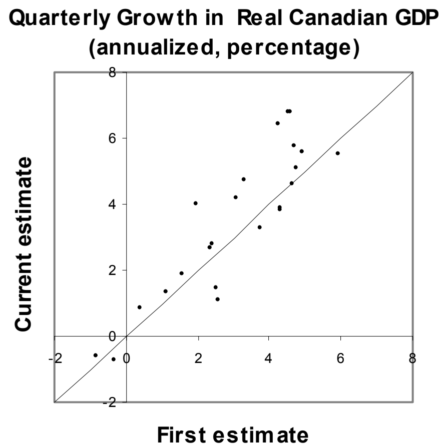
[1] Current estimate is calculated from the 2004Q3 vintage for real GDP, the 2004M9 vintage for retail sales, and the 2004M10 vintage for US industrial production.
[2] Revision = Current estimate – First estimate
[3] Gap is defined as the difference between the largest and the smallest value of a particular observation for a given time period, across all vintages of data.

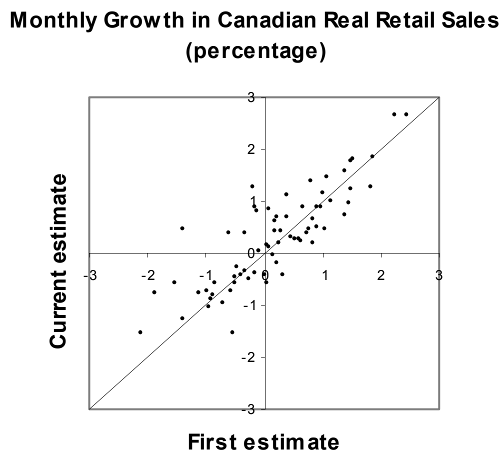
Table E-3. Test for equality of forecast accuracy: real-time forecasts by indicator model based on rolling AR

Model 1[1]	Model 2	HLN *p*-value[2]	H_0 = Model 1 encompasses Model 2[3]	H_0 = Model 2 encompasses Model 1[3]
I. Real-time forecasts assessed against current estimates of GDP, pseudo-real time against current estimates				
(real time)	(pseudo-real time)			
Rolling AR, 0 months	Rolling AR, 0 months	0.327		
Rolling AR, 1 months	Rolling AR, 1 months	0.001 **		
Rolling AR, 2 months	Rolling AR, 2 months	0.361		
Full quarterly information	Full quarterly information	0.274		
II. Real-time forecasts assessed against first estimates of GDP, pseudo-real time against current estimates				
(real time)	(pseudo-real time)			
Rolling AR, 0 months	Rolling AR, 0 months	0.027 **		Reject
Rolling AR, 1 months	Rolling AR, 1 months	0.386		
Rolling AR, 2 months	Rolling AR, 2 months	0.259		
Full quarterly information	Full quarterly information	0.600		
III. Real-time indicator model assessed against current estimates of GDP, Quarterly AR against current estimates				
(real time)	(real time)			
Rolling AR, 0 months	Quarterly AR	0.115		Reject
Rolling AR, 1 months	Quarterly AR	0.285		
Rolling AR, 2 months	Quarterly AR	0.015 **		Reject
Full quarterly information	Quarterly AR	0.037 **		Reject
IV. Real-time indicator model assessed against first estimates of GDP, Quarterly AR against first estimates				
(real time)	(real time)			
Rolling AR, 0 months	Quarterly AR	0.106		
Rolling AR, 1 months	Quarterly AR	0.001 **		
Rolling AR, 2 months	Quarterly AR	0.058 *		Reject
Full quarterly information	Quarterly AR	0.084 *		Reject

Note: For footnotes 1, 2, and 3, see the corresponding footnotes for Table D-1.

Figure E-1. Data revisions for real GDP, retail sales, and U.S. industrial production (1998Q4 to 2004Q2)





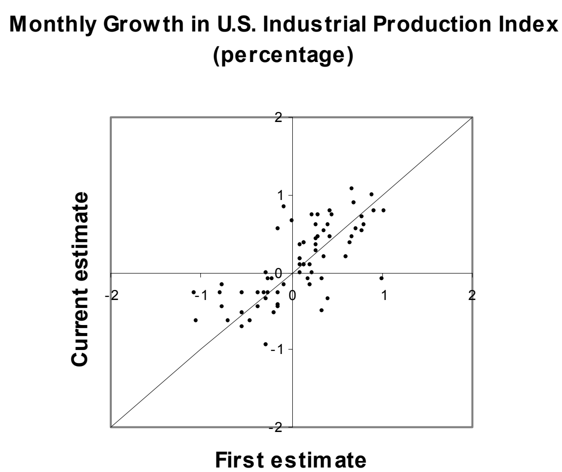
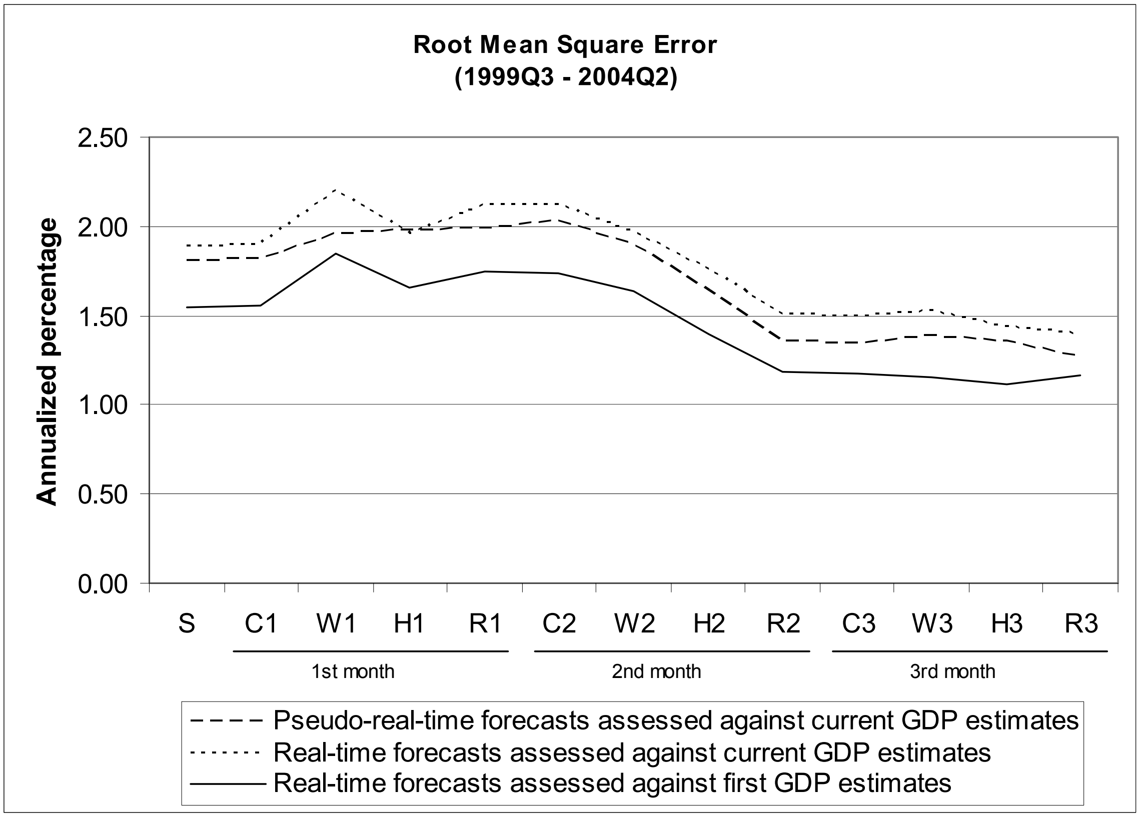
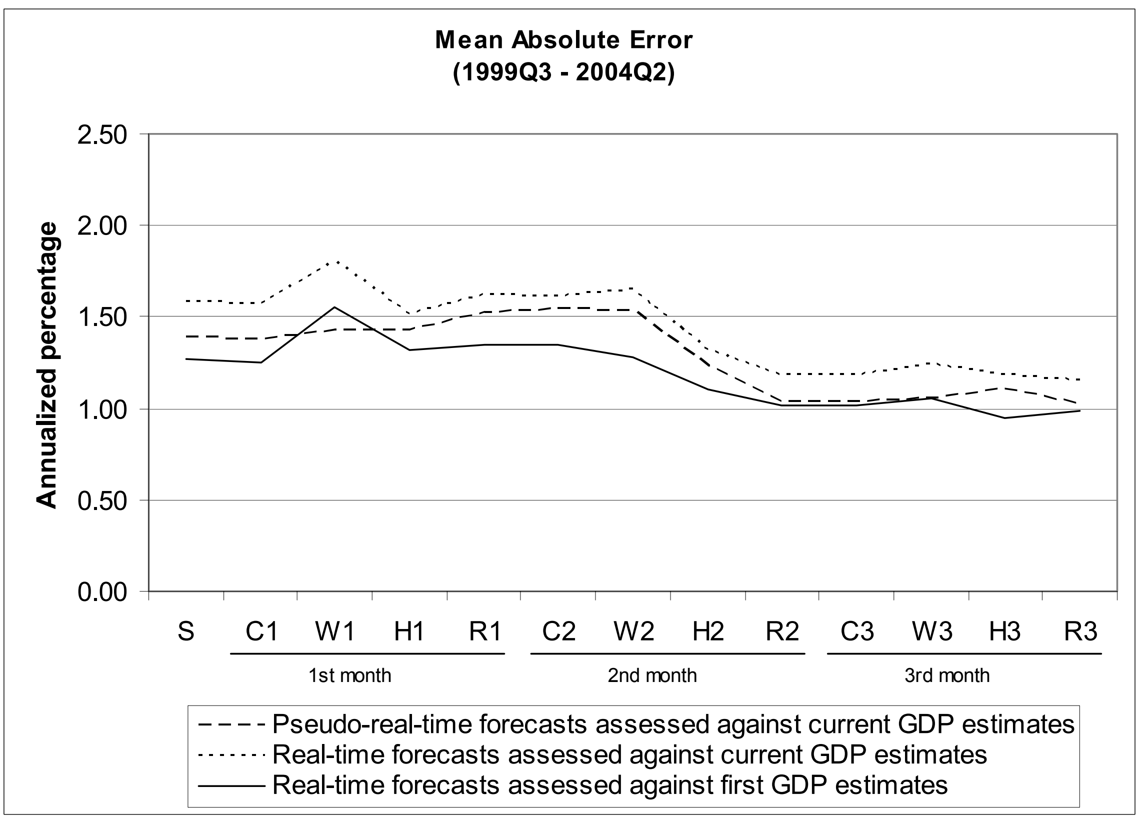


Figure E-2. Comparison of prediction errors based on the rolling AR approach for the monthly indicator model





Bank of Canada Working Papers
Documents de travail de la Banque du Canada

Working papers are generally published in the language of the author, with an abstract in both official languages, and are available on the Bank's website (see bottom of page). ***Les documents de travail sont généralement publiés dans la langue utilisée par les auteurs; ils sont cependant précédés d'un résumé bilingue. On peut les consulter dans le site Web de la Banque du Canada, dont l'adresse est indiquée au bas de la page.***

2006

2006-25	Linear and Threshold Forecasts of Output and Inflation with Stock and Housing Prices	G. Tkacz and C. Wilkins
2006-24	Are Average Growth Rate and Volatility Related?	P. Chatterjee and M. Shukayev
2006-23	Convergence in a Stochastic Dynamic Heckscher-Ohlin Model	P. Chatterjee and M. Shukayev
2006-22	Launching the NEUQ: The New European Union Quarterly Model, A Small Model of the Euro Area and the U.K. Economies	A. Piretti and C. St-Arnaud
2006-21	The International Monetary Fund's Balance-Sheet and Credit Risk	R. Felushko and E. Santor
2006-20	Examining the Trade-Off between Settlement Delay and Intraday Liquidity in Canada's LVTS: A Simulation Approach	N. Arjani
2006-19	Institutional Quality, Trade, and the Changing Distribution of World Income	B. Desroches and M. Francis
2006-18	Working Time over the 20th Century	A. Ueberfeldt
2006-17	Risk-Cost Frontier and Collateral Valuation in Securities Settlement Systems for Extreme Market Events	A. García and R. Gençay
2006-16	Benchmark Index of Risk Appetite	M. Misina
2006-15	LVTS, the Overnight Market, and Monetary Policy	N. Kamhi
2006-14	Forecasting Commodity Prices: GARCH, Jumps, and Mean Reversion	J.-T. Bernard, L. Khalaf, M. Kichian, and S. McMahon
2006-13	Guarding Against Large Policy Errors under Model Uncertainty	G. Cateau
2006-12	The Welfare Implications of Inflation versus Price-Level Targeting in a Two-Sector, Small Open Economy	E. Ortega and N. Rebei
2006-11	The Federal Reserve's Dual Mandate: A Time-Varying Monetary Policy Priority Index for the United States	R. Lalonde and N. Parent
2006-10	An Evaluation of Core Inflation Measures	J. Armour

Copies and a complete list of working papers are available from:
Pour obtenir des exemplaires et une liste complète des documents de travail, prière de s'adresser à :

Publications Distribution, Bank of Canada
234 Wellington Street, Ottawa, Ontario K1A 0G9
Telephone: 1 877 782-8248
(toll free in North America)
Email: publications@bankofcanada.ca
Website: http://www.bankofcanada.ca

Diffusion des publications, Banque du Canada
234, rue Wellington, Ottawa (Ontario) K1A 0G9
Téléphone : 1 877 782-8248 (sans frais en Amérique du Nord)
Adresse électronique : publications@banqueducanada.ca
Site Web : http://www.banqueducanada.ca