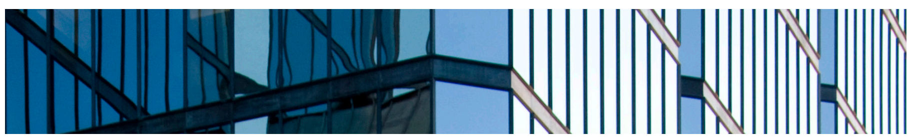
BANK OF CANADA
BANQUE DU CANADA

Bank of Canada Review

Summer 2013

Special Issue:
Tools for Current Analysis



Articles

The *Bank of Canada Review* is published four times a year. Articles undergo a thorough review process. The views expressed in the articles are those of the authors and do not necessarily reflect the views of the Bank.

For further information, contact:

Public Information
Communications Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9

Telephone: **613 782-8111**; **1 800 303-1282** (toll free in North America)
Email: **info@bankofcanada.ca**
Website: **bankofcanada.ca**

ISSN 1483-8303

Introduction: Tools for Current Analysis at the Bank of Canada

Don Coletti, International Economic Analysis, and
Sharon Kozicki, Canadian Economic Analysis

Monetary policy decisions draw on analysis from a variety of perspectives. The type and frequency of the analysis range from evaluations of financial conditions and new economic data, which may be done on a daily basis, to the Bank of Canada's medium-term economic projection, which staff update on a quarterly basis, to the Bank's estimates of potential output growth, which are typically revisited annually. The information underlying such analysis includes data published by Statistics Canada and other agencies, surveys conducted by the Bank and other institutions, assessments of regional and industry conditions based on media reports, as well as the intelligence received from meetings with business and government representatives.

The focus of this special issue of the *Bank of Canada Review* is on *current analysis*, which is the process of collecting and analyzing a broad spectrum of information to form a view of current economic activity. This type of analysis includes the *monitoring* of key indicators of macroeconomic conditions, including real gross domestic product (GDP) and its components. The term "monitoring" refers to staff short-term forecasts of economic activity that focus primarily on the two quarters of activity following the most recently published data.[1] The outputs of statistical models using information on economic indicators are combined with judgment, which may include, for example, estimates of the economic impact of special factors, such as large weather events or work stoppages, that may result in temporary disruptions to production or shift activity between quarters.[2]

There are a number of data-related challenges to producing frequent, timely forecasts. Analysts must deal with data that can be volatile and noisy, that are published at different frequencies (daily, weekly, monthly or

1 Since data are published with a lag, the term "backcasting" is often used to refer to predictions of activity for the previous quarter, while "nowcasting" refers to predictions of activity for the current quarter.

2 A key input into the monetary policy decision-making process is a quarterly projection of the global and Canadian economic conditions over a three-year horizon. The staff economic projection combines the monitoring for the short-term outlook (i.e., current analysis) with the longer-run predictions of structural macro models of the Canadian and global economies, and incorporates the staff's best judgment on various issues. The Canadian projection model traces the link from the policy rate to inflation, and is used to predict the consequences of various economic developments or "shocks" for the Canadian economy. (See T. Macklem, "Information and Analysis for Monetary Policy: Coming to a Decision," *Bank of Canada Review* (Summer 2002): 11–18.)

quarterly) or at different times during the year—with varying lags relative to the period over which the measures were taken—and that may be subsequently revised. These challenges motivate the use of multiple approaches, including various statistical techniques that draw on different information.

This special issue begins with a presentation of a new state-of-the-art indicator model for quarterly Canadian real GDP that provides a data-intensive, judgment-free approach to short-term forecasting. In "CSI: A Model for Tracking Short-Term Growth in Canadian Real GDP," André Binette and Jae Chang describe a model designed to address many of the challenges to producing forecasts. Canada's Short-Term Indicator model can provide a timely update to short-term forecasts each time new data on one of the indicators become available or when historical data are revised. The model's forecast accuracy is encouraging, and the Bank considers it an informative input to the staff monitoring of growth in real GDP.

As noted in the first article, current analysis does not rely only on the predictions of a single model. Analysts combine the results of several models with their own judgment when making forecasts. In the second article, "The Accuracy of Short-Term Forecast Combinations," Eleonora Granziera, Corinne Luu and Pierre St-Amant present the key findings of a recent research project that assessed the potential of combining the forecasts from different models (that is, taking a weighted average of models) to improve the accuracy and robustness of forecasts. Consistent with results in the academic literature, the authors find that averaging forecasts from several different models generally improves forecast accuracy. In addition, the authors find that unequal weighting of forecasts based on the past forecast performance of models tends to improve accuracy relative to an equal-weighted combination.

The choice of approach for current analysis is the focus of the third article, as well. In "Monitoring Short-Term Economic Developments in Foreign Economies," Russell Barnett and Pierre Guérin review the approaches used at the Bank of Canada for monitoring several key foreign economies—the United States, the euro area, Japan and China—and highlight how the specific challenges posed by policy needs and data availability influenced the respective modelling choices. As with the Canadian economy, the monitoring of foreign economies combines forecasts from different models with judgment to incorporate information not directly reflected in the most recent indicators.

In the final article in this special issue—"Big Data Analysis: The Next Frontier"—Nii Ayi Armah describes possible new, complementary sources of information for current analysis. With advances in technology, vast amounts of digital data are now available from business transactions, social media and networked computers. The combination of all of these data is called "big data." The availability of big data and the tools developed to analyze it could have a large impact on current analysis. Since this information is generally available on a more timely basis than traditional data, it may provide new insights into economic activity. However, several challenges related to methodological constraints, accessibility and privacy concerns limit the potential of big data, and the development of analytical tools for use with big data is still in its infancy.

CSI: A Model for Tracking Short-Term Growth in Canadian Real GDP

André Binette and Jae Chang, Canadian Economic Analysis

- The formulation of monetary policy requires central banks to assess the current state of the economy in a timely fashion. A variety of tools can be used to conduct this current analysis.
- Forecasting short-term growth in real GDP is a challenging task, given the wide range of potentially useful economic indicators and delays in the availability of data. Factor models offer a way to summarize the predictive content of many indicators without abandoning useful explanatory information in any of the series.
- Canada's Short-Term Indicator (CSI) is a new state-of-the-art indicator model for Canada that exploits the information content of 32 indicators to produce daily updates of real GDP growth forecasts for the two quarters following the latest release of official data.
- Although the forecast accuracy of this new model is encouraging, current analysis should not rely mechanically on predictions from a single model. Indeed, the Bank of Canada uses a wide range of models and information sources, as well as expert judgment, in producing its short-term forecasts.

The formulation of monetary policy relies, in part, on analysis of a variety of information about current economic conditions. Through current analysis,[1] economists try to understand and gauge the implications of the most recent economic conditions, including the impact of unpredictable events, such as natural disasters and work stoppages. Consequently, timely and accurate data are important for current analysis, since a clear understanding of current events is critical to better predict future developments. This in turn allows for the appropriate monetary policy response, given the forward-looking nature of the monetary policy approach.

The well-known maxim, "We need to know where we have been to know where we are going," highlights the value of short-term forecasting and early assessment, which are key facets of current analysis. To guide its monetary policy actions, the Bank of Canada devotes considerable time and

1 Current analysis is the process of collecting and analyzing a large amount of current information. This process is essential for monitoring and predicting short-term economic activity (see Coletti and Kozicki in this issue). Throughout this article, the terms "monitoring" and "short-term forecasting" are used interchangeably.

resources to monitoring and predicting short-term economic activity, as measured by real gross domestic product (GDP), and inflation. The Bank is continually developing new tools to improve its ability to predict economic developments over the short term, which is typically two quarters after the latest release of official GDP data.

The Bank of Canada is continually developing new tools to improve its ability to predict economic developments over the short term

Forecasting short-term growth in real GDP presents a number of challenges. Economists have a large number of data series at their disposal, ranging from National Accounts data to credit aggregates. From this profusion of data, they must extract the right information. As well, many indicators are published with lags, some of which are as long as two months. Economists need to find the best way to address the problems caused by these delays in the publication and revision of data. Another challenge is to develop tools that can use series with different frequencies, since data are published at daily, weekly, monthly or quarterly frequencies. The high-frequency data could provide useful information; for example, if economists consider quarterly data only, other information available over the course of the quarter, such as daily data on stock market indexes, could be lost. Another challenge involves truncated series resulting largely from redefinitions of variables.

The statistical agencies that produce these data face a trade-off between timeliness and the accuracy of the initial release. A lack of both timely and accurate economic data can lead to inaccurate conclusions about the state of the economy. Taken together, these challenges make current analysis a complex process. This article focuses on the forecasting aspect of current analysis and describes a recently developed state-of-the-art indicator model for tracking short-term growth in Canadian real GDP. This model can accommodate most challenges in current analysis at the same time and can complement other information, since the Bank uses a wide range of models and information sources as well as expert judgment in producing its short-term forecasts.[2]

The Bank uses a wide range of models and information sources as well as expert judgment in producing its short-term forecasts

Factor Models as Tools for Monitoring Economic Developments

Several statistical tools assist Bank staff in monitoring short-term economic developments. These tools are often econometric models, which are simplified mathematical approximations of a complicated and evolving reality. The variables included in these models are based on economic theory, and statistical techniques are used to identify relationships among them.

Research has demonstrated the potential of factor models to address the main challenge of current analysis—extracting useful information from abundant data on multiple indicators. Factor models describe the relationship among observed correlated variables in terms of a few unobserved variables, called factors. The premise of these models is that the factors explain the variation and common movement in a large number of observed variables. For example, movements in real GDP are correlated with changes in other measured variables such as employment and consumer confidence. Factor models formalize the idea that the true business cycle is not directly observed and is best measured by estimating the common movements of various economic time series (Burns and Mitchell 1946; Lucas 1977) (**Box 1**). By uncovering the underlying common movements, information from a variety of indicators can be used to forecast growth in real GDP.

Research has demonstrated the potential of factor models to extract useful information from abundant data on multiple indicators

2 Granziera, Luu and St-Amant (this issue) find that combining forecasts from different models generally improves forecast accuracy when compared with various benchmarks.

Box 1

Factor Models: Specification and Estimation

In general, factor models can be specified as:

$$x_{it} = \lambda_i f_t + e_{it},$$

where

$$f_t = \sum_{j=0}^{q} \varphi_j f_{t-j} + u_t.$$

x_{it} is one of N observed variables in the model and t represents the time period. Each variable $\{x_{it}\}_{i=1}^{N}$ is assumed to depend on a latent (unobserved) factor, f_t, and an idiosyncratic component, e_{it}. The term $\lambda_i f_t$ is the common component underlying x_{it}, with λ_i being the corresponding factor loading for variable i. The factor loading can be defined as the marginal effect of the unobserved factor f_t on x_{it}. All idiosyncratic components are assumed to be uncorrelated with each other and also uncorrelated with the unobserved common component. The factor f_t is assumed to follow a covariance-stationary autoregressive process. If f_t were known, parameter estimates of λ_i and φ_j could be obtained from regression analysis. Unfortunately, f_t, λ_i and φ_j are all unknown. The only known elements in the above system of equations are the observed data in $\{x_{it}\}_{i=1}^{N}$.

Since both λ_i and f_t are unobserved, the factor model is not identified, in the sense that some restrictions have to be imposed in order to estimate the model. For a relatively small number of variables, and assuming that residuals are normally distributed, maximum likelihood and the Kalman filter can be used to obtain estimates of the factor loadings and the common factor (Stock and Watson 1991; Kalman 1960). An alternative methodology for estimating the common latent factor is principal-component analysis (PCA), created in 1901 by British mathematician Karl Pearson (Pearson 1901). Empirical evidence suggests that, for purely predictive purposes on a given data set, factor models estimated by the Kalman filter generally have a similar performance to those estimated by PCA (Boivin and Ng 2005).

Employing these models to uncover useful patterns is called factor analysis, which was first developed in 1904 by the British psychologist Charles Spearman in the field of intelligence research (Spearman 1904). Spearman theorized that seemingly disparate cognitive test scores could be explained by a single general intelligence factor. Geweke (1977) and Sargent and Sims (1977) were among the earliest researchers to model economic time series with factor models.

Factor models offer a way to summarize the predictive content of many indicators without abandoning the relevant information in any of the series. If all indicators at a given point in time move together, the model will easily discover the general upward or downward trend in the series. When many indicators move in different directions and there is no obvious upward or downward trend, the conflicting signals are resolved by a weighted average, with series that are more informative (based on historical correlations) receiving more weight than less-informative series. Essentially, more-volatile series are often given less weight.

Canada's Short-Term Indicator Model

The Bank of Canada's factor model—Canada's Short-Term Indicator, or CSI—closely follows the approach of Camacho and Perez-Quiros (2010), which accommodates missing observations resulting from delays in the release of data, as well as data samples that represent short time spans, monthly and quarterly indicators (mixed frequencies), and different transformations of the data (monthly, quarterly and year-over-year growth rates). In addition, Camacho and Perez-Quiros (2010) include a way to deal with multiple GDP releases, which allows the information in the monthly real GDP figures for Canada to be exploited. The model is self-contained, since it makes internal predictions for each indicator, enabling assessment of the impact of each new data release on the model's forecast of real GDP growth.

Forecasting with CSI follows three important steps: (i) collect information for a wide variety of economic indicators;[3] (ii) conduct a complete evaluation of the available information (at this stage, CSI analyzes the indicators and determines weights to assign to each of them); and (iii) calculate the common component and the forecast of real GDP growth.

Main features of CSI

CSI is a monthly, dynamic, single-factor model built on the principle that any series can be divided into two components: a component that is common to all variables in the model and an idiosyncratic component. All indicators in CSI are projected based on a common component and on their own individual dynamics, as described by autoregressive (AR) processes in which the current values of the indicators are explained by using only their past values. The empirical analysis uses data available from 1982 through to 2012.

CSI is built on the principle that any series can be divided into two components: a component that is common to all variables in the model and an idiosyncratic component

Although CSI is a monthly model, its indicators include quarterly variables. The model simply considers these variables to be monthly series with missing observations. The quarterly indicators are linked to the monthly factor using a mathematical relationship that expresses quarterly growth rates as monthly growth rates in both the current quarter (the quarter being measured) and the previous quarter (Statistics Canada 2011).[4] This relationship implies that about 66 per cent of the quarterly growth rate of a series is known after the release of the first month of a given quarter, and about 90 per cent is known after two months.

Unlike statistical agencies in the United States and some other countries, which publish preliminary and advance estimates of quarterly real GDP, Statistics Canada provides monthly GDP figures for Canada. Monthly and quarterly real GDP series are not conceptually identical: monthly figures are published at basic prices, while quarterly real GDP is expressed at market prices, which include net taxes on products. Notwithstanding the conceptual difference, the growth rates of the two measures of real GDP often exhibit a similar dynamic at a quarterly frequency. Consequently, it is assumed that, after the first and second month of the quarter, information on monthly real GDP reflects early estimates of quarterly real GDP at market prices. Thus, this key monthly indicator is treated in the same way that preliminary and advance estimates of quarterly GDP are dealt with by researchers using data for other countries. As noted in Camacho and Perez-Quiros (2010), these early estimates of GDP are incomplete and the difference between them and the final quarterly release is unpredictable.

CSI indicators

While factor models, in theory, could process the information content of a very large number of indicators, Boivin and Ng (2006) show that larger data sets do not necessarily generate more-accurate forecasts in empirical applications. The choice of indicators used in CSI has therefore been guided

3 Armah (this issue) discusses the rapid growth in the number of potential indicators resulting from the development of information technology.

4 The first month of the current quarter has the largest impact, with a weight of 1; the previous month and the following month have the next-largest impact (weight = 2/3); and the last month of the current quarter and the second month of the previous quarter have the least impact (weight = 1/3).

by the following criteria: (i) the variables should be directly related to the Canadian economy; and (ii) forecasts over the past decade should be more accurate than simple benchmarks found in the literature.[5]

Over time, Bank staff have evaluated the ability of various indicators to predict the growth rate of real GDP. The current specification of CSI includes 32 indicators (**Appendix 1**),[6] most of which are well-known statistics for Canada, such as total hours worked (from the Labour Force Survey), retail trade and housing starts. Other indicators include soft information (such as consumer confidence), financial data and international variables. U.S. data series and the global purchasing managers' index (PMI) for manufacturing are used to proxy foreign demand for Canadian exports.[7] As previously mentioned, after the first and second month, monthly GDP information provides early estimates of quarterly GDP. A quarterly momentum indicator related to monthly GDP is also incorporated in the model to capture the early dynamic of a new quarter.[8] The addition of timely soft information and financial indicators gives the model early information about the quarter of interest and potentially improves forecast accuracy.[9]

The current specification of CSI includes 32 indicators, most of which are well-known statistics for Canada

CSI performance

CSI is based on the premise that common movements (i.e., the common factor) that affect all indicators are linked to the business cycle, as measured by growth in real GDP. The common factor should therefore have a profile similar to GDP growth. In fact, the model performs relatively well, since it explains about 75 per cent of the variation in the quarterly growth rate of real GDP over the 1982–2012 period (**Chart 1**).

CSI explains about 75 per cent of the variation in the quarterly growth rate of real GDP over the 1982–2012 period

The estimation results (factor loadings) also suggest that all of the indicators retained in the model exhibit a positive correlation with the common factor.[10] Nevertheless, the strength of the correlation varies among indicators (**Chart 2**). As expected, the momentum indicator, as well as early estimates of GDP and quarterly GDP, present the strongest relationships with the common factor and therefore have the greatest impact on the model's forecast. For monthly variables, the link with the common factor varies by the type of indicator (hard, soft or financial). Most of the variables with above-average correlation are standard statistics (hard indicators), with the exception of the global PMI for manufacturing. While 15 monthly indicators have a below-average relationship with the common factor, 8 of them are timely, with very short publication lags. These indicators have been included in an attempt to improve the forecast performance early in the forecast cycle.

5 Forecasts generated by an AR model and the unconditional mean of the series are the benchmarks against which we have compared CSI forecasts.

6 We initially considered about 50 indicators and retained in the model only the indicators that provide information not available in other series.

7 Morel (2012) presents a measure of foreign activity that tracks historical export data relatively well. The measure includes U.S. consumption, U.S. residential investment, U.S. business investment and foreign GDP outside of the United States. As proxies for U.S. activity, CSI uses U.S. retail sales, U.S. car sales, U.S. housing starts and U.S. industrial production.

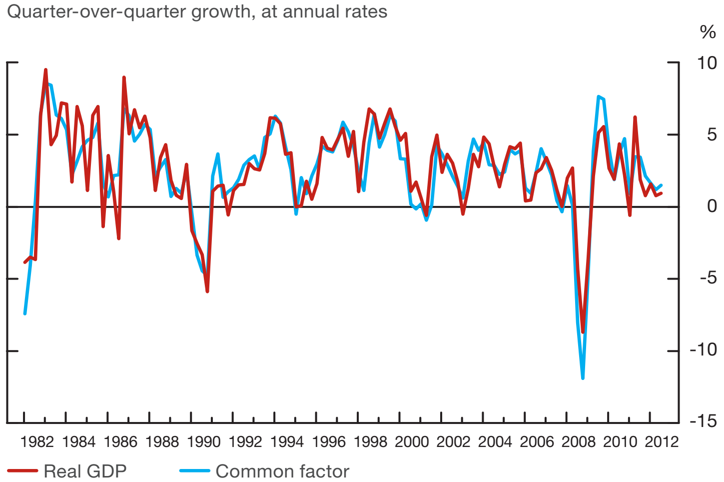
8 The momentum indicator is a quarterly series that exploits monthly GDP information from the last two months of the previous quarter to assess the vigour of economic activity at the start of a new quarter. The unpublished months of the momentum indicator and the early estimates are currently forecast with a moving average of the growth rate from the previous three months.

9 Most indicators are included in a difference-of-log format (growth rate), while some are incorporated in log-level format. The data transformations ensure that all indicators used in the model are stationary series and provide the best forecasting performance beyond the sample period.

10 A factor loading measures the change in an observed variable following a one-unit change in the common factor.

Chart 1: Quarterly growth in real GDP and the common factor

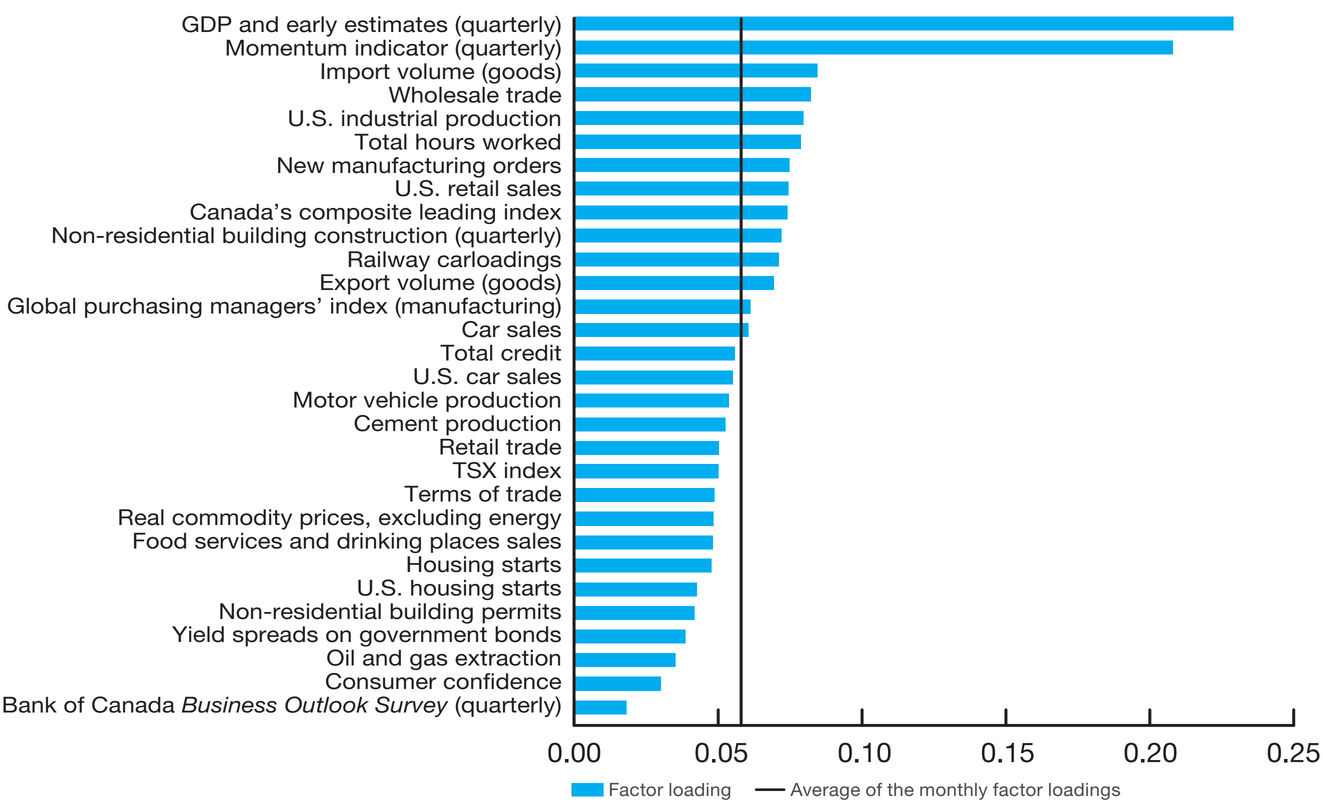
Quarter-over-quarter growth, at annual rates



Sources: Statistics Canada and Bank of Canada calculations

Last observation: 2012Q4

Chart 2: CSI factor loadings

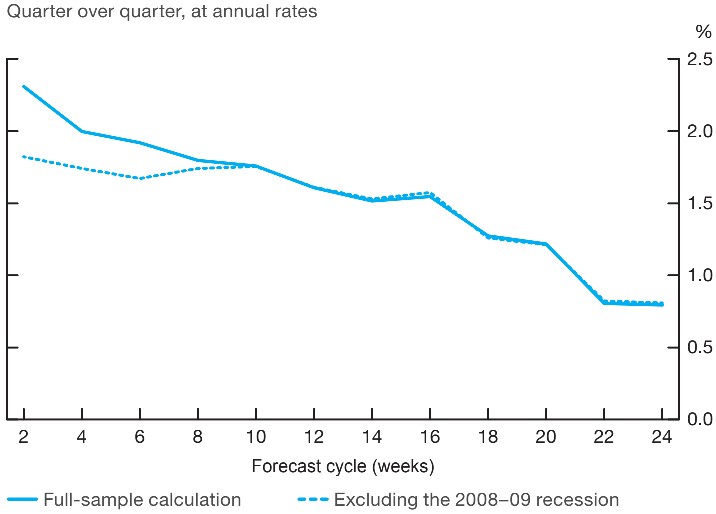


Note: See Box 1 for a discussion of factor loadings.
Source: Bank of Canada calculations

To assess the performance of CSI beyond the sample period, a quasi-real-time exercise is performed in which the model uses only the information available at the time that it makes its predictions. This approach mimics the actual conditions faced by analysts at the Bank. The exercise is conducted in quasi-real time, since the original unrevised data are not available for

Chart 3: Root-mean-square prediction errors, 2000Q1–2012Q4

Quarter over quarter, at annual rates



Source: Bank of Canada calculations

all the indicators.[11] The model's performance is assessed 12 times over the forecast cycle. For a given quarter, the cycle covers six months, representing a prediction every two weeks. For example, the initial forecast for the fourth quarter of 2012 was made in early September 2012, while the last prediction was made in the second half of February 2013 (i.e., just before the release of real GDP growth for the fourth quarter).[12]

Overall, the CSI model performs as anticipated. The initial forecasts are not very accurate, with root-mean-square prediction errors (RMSPEs) above 2 per cent (**Chart 3**), in part, because of the model's inability to predict the severe economic downturn in 2008–09. The accuracy of CSI increases, however, as more information becomes available, and significant improvements occur in weeks 18 and 22 with the release of the monthly GDP data for the first two months of the quarter (early estimates). This should come as no surprise, since GDP at basic prices and GDP at market prices are highly correlated at the quarterly frequency, despite the small conceptual difference.

Overall, the CSI model performs as anticipated, with accuracy increasing as more information becomes available

Another insightful measure of performance is the model's forecast memory, which is the ratio of the RMSPE to the standard deviation of quarterly GDP growth. When this ratio is above one, model forecasts are less accurate than a simple forecast that assumes GDP growth will equal the average of the series (i.e., the unconditional mean). Thus, the forecast memory indicates the horizon at which the indicators provide useful signals. As **Chart 4** shows, CSI provides valuable information above the unconditional mean as early as one month before the start of the quarter under consideration. For example, forecasts made by CSI in September (weeks 2 and 4) for the fourth quarter of a given year are, on average, more accurate than a forecast based on the unconditional mean of real GDP growth.

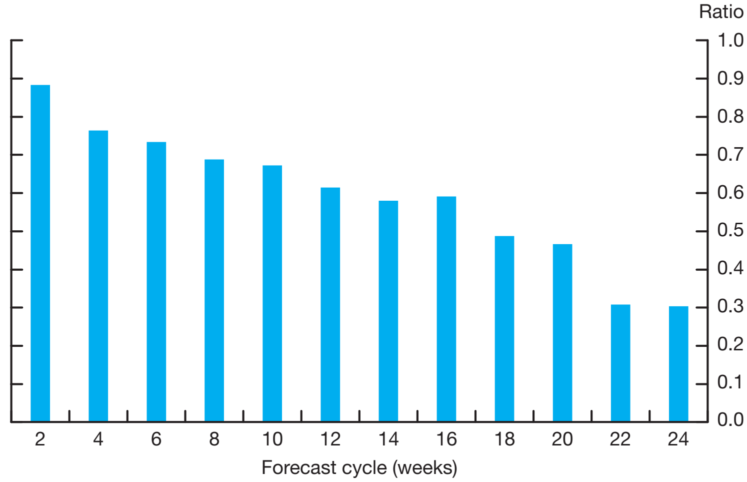
While forecast precision is important, the direction of a prediction is also crucial. The "hit ratio" indicates how often a model correctly predicts an increase or a decrease in the growth rate of any series. Although the CSI

11 Although real-time data would provide a better sense of the model's performance, this type of analysis is left for future work.

12 The first four weeks of the forecast cycle (i.e., before the start of the quarter being considered) are often called the forecasting period. Predictions made during the quarter (weeks 5 to 16) are referred to as "nowcasting," while those made after the end of the quarter (but before the quarterly data are released) are called "backcasting."

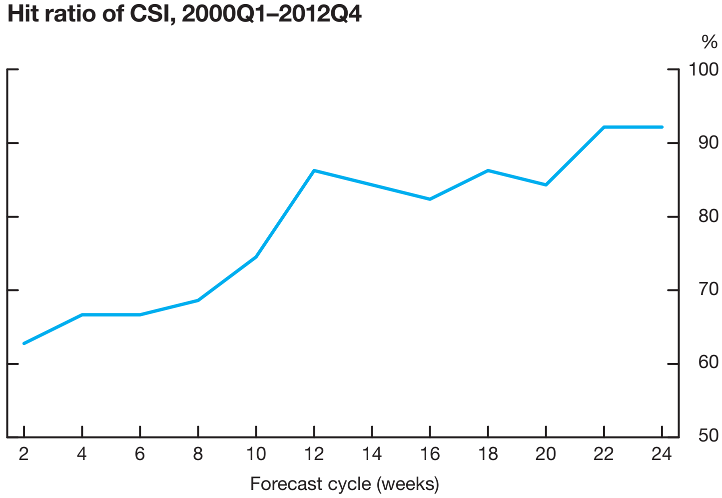
Chart 4: Forecast memory of CSI, 2000Q1–2012Q4

Ratio of the root-mean-square prediction error to the standard deviation of quarterly GDP growth



Source: Bank of Canada calculations

Chart 5: Hit ratio of CSI, 2000Q1–2012Q4



Source: Bank of Canada calculations

forecasts are not very accurate early in the forecasting cycle (**Chart 3**), their direction is correct more than 60 per cent of the time (**Chart 5**). Furthermore, as more data become available during the forecast cycle, CSI correctly predicts an increase or a decrease in real GDP growth with a hit ratio of about 90 per cent.

Conclusion

The main objective of CSI is to offer a data-intensive, judgment-free approach to short-term forecasting. CSI provides a way to extract information more systematically from some indicators that had been previously used with judgment to forecast GDP. A factor model can process a large number of indicators and, as implemented in CSI, is able to produce a new prediction of real GDP growth almost immediately following the latest release of data for an indicator. While these results are encouraging, current analysis should not rely mechanically on predictions from a single model. The Bank of Canada uses a wide range of models and information sources, as well as expert judgment, in producing its short-term forecasts. As such, CSI is considered to be a good complement to other forecasting tools, providing valuable information about the direction of economic growth during the current quarter and the next. Further assessment of its real-time performance is needed, however, to better ascertain the model's full potential.

Appendix 1

CSI Indicators

Indicator	Source[a]	Frequency
1. Early estimate 1 (first month of GDP)†	STC	Quarterly
2. Early estimate 2 (second month of GDP)†	STC	Quarterly
3. Quarterly GDP†	STC	Quarterly
4. Momentum indicator†	STC	Quarterly
5. U.S. industrial production†	FED	Monthly
6. Total hours worked (Labour Force Survey)†	STC	Monthly
7. Canada's composite leading index†	STC and MLI	Monthly
8. U.S. retail sales†	USCB	Monthly
9. Global purchasing managers' index (manufacturing)††	J.P. Morgan	Monthly
10. Real commodity prices, excluding energy prices†††	BoC	Monthly
11. Terms of trade†	STC	Monthly
12. TSX index†††	STC	Monthly
13. Wholesale trade†	STC	Monthly
14. Consumer confidence††	CBoC	Monthly
15. Car sales†	STC	Monthly
16. Import volume (goods)†	STC	Monthly
17. Export volume (goods)†	STC	Monthly
18. Retail trade†	STC	Monthly
19. New manufacturing orders†	STC	Monthly
20. U.S. car sales†	WA	Monthly
21. Food services and drinking places sales†	STC	Monthly
22. Oil and gas extraction†	STC	Monthly
23. Bank of Canada *Business Outlook Survey*†† (average balance of opinion on past sales growth, future sales growth, investment in machinery and equipment, and output price pressures; some or significant difficulty in meeting demand; and labour shortages)	BoC	Quarterly
24. Railway carloadings†	STC	Monthly
25. Housing starts†	CMHC	Monthly
26. U.S. housing starts†	USCB	Monthly
27. Motor vehicle production†	WA	Monthly
28. Cement production†	STC	Monthly
29. Non-residential building construction†	STC	Quarterly
30. Total credit (household and business)†††	BoC	Monthly
31. Non-residential building permits†	STC	Monthly
32. Yield spreads on government bonds††† (Government of Canada bond yields: yield of a 5–year bond minus the yield of a 3-month treasury bill)	STC	Monthly

† Hard indicator †† Soft indicator ††† Financial indicator

a. The indicators used in Canada's Short-Term Indicator (CSI) model are taken from the following sources: Statistics Canada (STC), Bank of Canada (BoC), Canada Mortgage and Housing Corporation (CMHC), Conference Board of Canada (CBoC), Macdonald-Laurier Institute (MLI), WardsAuto (WA), United States Census Bureau (USCB), Board of Governors of the Federal Reserve System (FED) and J.P. Morgan.

Literature Cited

Armah, N. A. 2013. "Big Data Analysis: The Next Frontier." *Bank of Canada Review* (Summer): 32–39.

Boivin, J. and S. Ng. 2005. "Understanding and Comparing Factor-Based Forecasts." *International Journal of Central Banking* 1 (3): 117–51.

———. 2006. "Are More Data Always Better for Factor Analysis?" *Journal of Econometrics* 132 (1): 169–94.

Burns, A. F. and W. C. Mitchell. 1946. *Measuring Business Cycles*. Cambridge, MA: National Bureau of Economic Research.

Camacho, M. and G. Perez-Quiros. 2010. "Introducing the Euro-Sting: Short-Term Indicator of Euro Area Growth." *Journal of Applied Econometrics* 25 (4): 663–94.

Coletti, D. and S. Kozicki. 2013. "Introduction: Tools for Current Analysis at the Bank of Canada." *Bank of Canada Review* (Summer): 1–2.

Geweke, J. 1977. "The Dynamic Factor Analysis of Economic Time Series Models." In *Latent Variables in Socioeconomic Models*, 365–83. Edited by D. Aigner and A. Goldberger. Amsterdam: North Holland.

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

Kalman, R. E. 1960. "A New Approach to Linear Filtering and Prediction Problems." *Journal of Basic Engineering* 82 (1): 35–45.

Lucas, R. E. 1977. "Understanding Business Cycles." *Carnegie-Rochester Conference Series on Public Policy* 5 (1): 7–29.

Morel, L. 2012. "A Foreign Activity Measure for Predicting Canadian Exports." Bank of Canada Discussion Paper No. 2012-1.

Pearson, K. 1901. "On Lines and Planes of Closest Fit to Systems of Points in Space." *Philosophical Magazine* 2 (6): 559–72.

Sargent, T. J. and C. A. Sims. 1977. "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory." Federal Reserve Bank of Minneapolis Working Paper No. 55.

Spearman, C. 1904. "The Proof and Measurement of Association Between Two Things." *American Journal of Psychology* 15 (1): 72–101.

Statistics Canada. 2011. "The Relationship Between Monthly, Quarterly, and Annual Growth Rates." *Canadian Economic Observer* 24 (6), Cat. No. 11-010-X (June).

Stock, J. H. and M. W. Watson. 1991. "A Probability Model of the Coincident Economic Indicators." In *Leading Economic Indicators: New Approaches and Forecasting Records*, 63–90. Edited by K. Lahiri and G. H. Moore. Cambridge, U.K.: Cambridge University Press.

The Accuracy of Short-Term Forecast Combinations

Eleonora Granziera, Corinne Luu and Pierre St-Amant, Canadian Economic Analysis

- This article examines whether, and under what circumstances, combining forecasts of real GDP from different models can improve forecast accuracy. It also considers which model-combination methods provide the best performance.
- In line with the previous literature, we find that combining forecasts from different models generally improves forecast accuracy when compared with various benchmarks.
- Unlike several previous studies, we find that assigning equal weights to each model is not always the best weighting scheme. Unequal weighting based on the past forecast performance of models tends to improve accuracy when forecasts across models are substantially different.

In conducting monetary policy, central banks need to regularly assess the current and the future state of the economy. To do this, they combine expert judgment with the results of several models, since no single model can provide the most-accurate results in all circumstances and at all forecasting horizons. For example, while some models do well at forecasting the current period, others do well at forecasting one or two quarters ahead. In addition, with the flow of new data, structural changes in the economy and the introduction of new modelling techniques, the relative usefulness of individual models tends to change over time. Economists at the Bank of Canada therefore regularly update the set of models they use in their current analysis and short-term forecasting.

Uncertainty about the appropriateness of individual models has led researchers to propose using combinations of forecasts from different models, i.e., a diversification strategy, since this strategy may produce forecasts that are less vulnerable to structural breaks and may mitigate the risk that decisions are based on the results of poorly performing models. Indeed, researchers have often found that forecasts generated by combinations of models are more accurate and more robust than those of individual models (Stock and Watson 2004).

This article presents the key findings of a recent project that assessed the potential for various combinations of models to improve the accuracy and robustness of forecasts. The project focused on models for Canadian real gross domestic product (GDP) that the Bank of Canada has used to backcast (predict the previous quarter before data for that quarter are released by Statistics Canada), nowcast (predict the current period) and forecast over short horizons (typically one or two quarters ahead).[1] We first briefly describe the models and explain how these models were estimated and their forecasts produced. We then explain how forecasts were combined and present the results from those combinations.

Models: Descriptions and Forecasts

To assess the benefits of combining forecasts, this article focuses on a group of simple models, as well as more-complex forecasting tools that the Bank has used to predict quarterly growth in Canadian real GDP, measured at market prices, from Statistics Canada's National Income and Expenditure Accounts. Some of the models in our sample are built to forecast quarterly growth in real GDP over the very short term, while others are designed to produce more accurate forecasts at longer horizons, up to four quarters following the latest release of real GDP (**Table 1**). The individual predictions from these tools are combined to forecast quarter-over-quarter annualized real GDP growth over the short term (i.e., the two quarters following the release by Statistics Canada of the latest quarterly data on real GDP), as well as over slightly longer horizons (i.e., the third or fourth quarter after the latest available data).

To assess and combine the forecasts from the various models, we need to generate predictions from these models in a manner similar to how they would have been generated in practice when forecasting real GDP. The

Table 1: Models of real GDP used in the forecast combinations

Name	Type of model	Forecast horizon[a]	Variables used
State-Space Nowcasting (SSN) Model	Factor model	1–2 quarters	Weekly financial data and monthly data (including total hours worked, monthly real GDP and housing starts)
Bayesian Vector Autoregression (BVAR) Model	Bayesian vector autoregression model	1–4 quarters	Key Canadian and U.S. macroeconomic variables (including U.S. real GDP growth, core inflation and interest rates)
Regional Aggregate Model (RAM)	Univariate models for each region in Canada, aggregated to the national level	1–2 quarters	Provincial-level indicators (e.g., provincial economic accounts, manufacturing sales, employment and retail sales)
Supply-Side Bridge Equation (SSBE) Model	Linear univariate model	1–2 quarters	Wholesale trade, housing starts, interest rates, U.S. retail sales and U.S. personal consumption
Investment-Saving (IS) Curve Models (two models)	Linear univariate models	1–4 quarters	Global output, interest rates, exchange rates and commodity prices. One model includes consumer confidence.
Yield Curve Model (YCM)	Linear univariate model	1–4 quarters	Lagged yield curve (difference between the overnight rate and the rate for 10-year Government of Canada bonds)
Canadian Composite Leading Indicator (CLI)[b]	Linear univariate model	1–2 quarters	Based on the CLI of real activity, composed of indicators of real GDP (e.g., housing index, money supply (M1), TSE 300 stock price index and U.S. Conference Board leading indicator)
Hours Model (HM)	Linear univariate model	1–4 quarters	Based on growth in total hours worked
Autoregressive Distributed Lag (ADL) Model	Linear univariate model	1–4 quarters	Financial variables, composite leading indicator, business credit, employment and U.S. real GDP growth
Narrow Money Model (MM)	Linear univariate model	1–2 quarters	Money supply (M1+)

a. Beyond the latest release of data on real GDP growth

b. Since the project was conducted, Statistics Canada has discontinued publication of the CLI; this model has therefore been modified to incorporate a different, albeit similar, measure of activity.

1 While the Bank still uses some of these models, others have been dropped and new ones have been added.

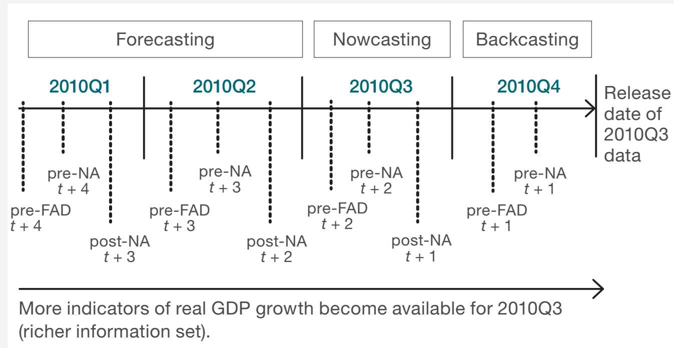
Box 1

Timeline for Real GDP Forecasts

To illustrate the timeline in our analysis, **Figure 1-A** shows the dates when forecasts were made for real GDP growth in the third quarter of 2010. The first forecast (pre-FAD *t* + 4) took place in January 2010. Owing to publication lags for the National Accounts (NA), at this time, quarterly data for real GDP were available only up until 2009Q3; consequently, this prediction for 2010Q3 is considered a four-step-ahead (*t* + 4) forecast. This forecast coincided with a major briefing provided by Bank staff to senior management before the January fixed announcement date (FAD) for a monetary policy decision. The next forecast was produced immediately before the release of real GDP data for 2009Q4 (pre-NA *t* + 4). This second forecast would therefore be at the end of February, which would still be considered a four-step-ahead forecast. Even though no new information regarding quarterly real GDP was released between the January and February forecasts, new weekly and monthly data would have become available, and therefore the forecasts for 2010Q3 may change in some models.

After the release of real GDP data for 2009Q4, the estimation period for the models was extended to include the new data and a new forecast for 2010Q3 (post-NA *t* + 3) was made, which would be a three-step-ahead forecast (*t* + 3). This process was continued until immediately before the actual release of the 2010Q3 real GDP data in November 2010. Eleven forecasts, including those preceding the first FAD, were made for each quarter. The forecast horizons ranged from four steps ahead to one step ahead.

Figure 1-A: Timeline of forecasting the real GDP growth rate in the third quarter of 2010[a]



pre-FAD: forecasts produced about one week before the first fixed announcement date in each quarter

pre-NA: forecasts produced immediately before the release of real GDP (also referred to as the National Accounts)

post-NA: forecasts produced immediately after a real GDP release

a. This is the timeline of the production of forecasts for 2010Q3 in this project. However, the Bank would have made forecasts for 2010Q3 much earlier based on other modelling methods, since this quarter would have been part of the longer-term projection using the Terms-of-Trade Economic Model (ToTEM) (see Coletti and Kozicki (this issue) for information on the projection process).

2011Q2 vintage of National Accounts data[2] is used for all estimations (albeit with a sample that lengthens over time), while the initial observation used to estimate each model varies depending on the model. At each point in time, predictions are produced for up to four quarters beyond the latest release of quarterly real GDP (**Box 1**). The models start to produce forecasts as early as 10.5 months before the actual release of real GDP data for the quarter of interest. A total of 11 forecasts are made: a week before the first fixed announcement date (FAD) in each quarter, as well as immediately before and immediately after the release of National Accounts (NA) data.[3] The forecasts are based on the information available up to that period (initial estimates use data up to 1999) for quarterly real GDP and for the monthly or weekly variables included in the models. When data for a new quarter are available, the models are re-estimated and the forecast cycle is repeated with forecasts produced up to 2011. The models are then evaluated by comparing these forecasts with the actual quarterly growth in real GDP (using the 2011Q2 vintage of data). Since data tend to be revised over time, this exercise should be considered a proxy for forecast accuracy in real time, and the results of this analysis should be interpreted with caution.

2 A vintage is the latest estimate for a given series at a particular time.

3 Some models produce only very short-term forecasts, i.e., the first or second quarter following the latest release of real GDP data (Table 1); they would therefore produce fewer than 11 forecasts for a given quarter.

Forecast Combinations

An initial objective of the forecast combination exercise is to assess whether and under what circumstances competing forecasts may be combined to produce a pooled forecast that performs better than the individual benchmark models. A second objective is to determine whether the relative success of combination methods changes with the forecast horizon.

To evaluate the accuracy of the forecasts (of each model as well as the combined forecasts), we use the root-mean-square prediction error (RMSPE), which measures the discrepancy between the model forecasts and the actual realizations. A lower RMSPE indicates a better performance, or, equivalently, more accurate model forecasts.

To combine forecasts, it is necessary to assign a weight to each model, and the success of the combination may depend on how these weights are assigned. There are several combination schemes with different degrees of sophistication, from simple averaging to complex methods where weights change over time.[4] This article considers the methodologies used most often in the literature, which can be distinguished according to the importance they assign to the past forecast performance of the models. **Box 2** provides a technical description of these combination schemes.

The **simple-average (SA) scheme**, which weighs forecasts equally, has the major advantage of not requiring the use of statistical methods to estimate the weights, since they are determined simply by the number of models.

Box 2

Forecast Combination Schemes

The combined h-step-ahead forecast, y^C_{t+h}, is constructed as a weighted average of the N single-model forecasts:

$$y^C_{t+h} = w^1 y^1_{t+h} + w^2 y^2_{t+h} + \dots + w^N y^N_{t+h},$$

where $y^i_{t+h}\ i = 1, \dots N$ is the forecast from model i.

The weights are computed using forecasts up to time t and differ according to the combination scheme:

(i) Simple average: $w^i = 1/N$

(ii) Inverse RMSPE: $w^i = \frac{RMSPE_{i,t+h}{}^{-1}}{\sum_{j=1}^{N} RMSPE_{j,t+h}{}^{-1}}$, where $RMSPE_{i,t+h} = \sqrt{\frac{1}{t}\sum_{\tau=1}^{t}\left(y_{\tau+h} - y^i_{\tau+h}\right)^2}$

(iii) Inverse rank: $w^i = \frac{RANK_{i,t+h}{}^{-1}}{\sum_{j=1}^{N} RANK_{j,t+h}{}^{-1}}$, where $RANK_{i,t+h} = 1$ if model i has the lowest RMSPE up to time t; $RANK_{i,t+h} = 2$ if model i has the second-lowest RMSPE; etc.

(iv) Least squares: weights are the ordinary least-squares coefficients estimated from

$$y_{t+h} = w^1 y^1_{t+h} + w^2 y^2_{t+h} + \dots + w^N y^N_{t+h} + \varepsilon_{t+h}.$$

4 See Timmermann (2006) for a comprehensive review.

Several studies document the success of the SA scheme relative to more sophisticated weighting schemes, at least when forecasting beyond the current quarter (e.g., Stock and Watson 1999, 2004). The imprecision of statistical methods when estimating weights with short samples of data is cited as the reason for this phenomenon. Theoretically, the SA scheme can be shown to be the optimal combination scheme if the models included in the combination have the same predictive accuracy (as measured by the RMSPE) and the correlations between forecasts from any two models are identical (Smith and Wallis 2009).[5] Intuitively, under these conditions, taking into account the past performance of models and the correlation of forecasts across models does not compensate for the imprecision introduced by estimating the weights.

Other combination schemes weigh the models based on their past performance. In the **inverse-RMSPE (I-RMSPE) scheme**, larger weights are assigned to models that have better forecast accuracy over the forecast sample. Similarly, the **inverse-rank (I-Rank) scheme** assigns to each model a weight inversely proportional to its rank, which is based on the performance of the model over the forecast sample. These schemes do not require the estimation of the correlation between single-model forecasts and do well in practice: for example, a recent Bank of Norway study finds that the I-RMSPE methodology is superior to the simple average (Bjørnland et al. 2012).

Finally, the **least-squares weighting** scheme takes into account not only the past performance of the models but also the correlation of forecasts from different models. The weights assigned in this combination tend to be larger for more accurate forecasts that are less correlated with other forecasts. Although, theoretically, the estimated weights obtained through this method are optimal (Timmermann 2006), they may be biased, especially when the sample size is small.[6] We consider three variants of this scheme: (i) weights are unconstrained (LS); (ii) combined weights sum to one but can take negative values (LSnp); and (iii) combined weights sum to one and must be positive (LSp).

The RMSPE of each weighted combination of forecasts is assessed against that of several benchmarks. The first benchmark is a simple autoregressive (AR) model of the quarterly real GDP growth rate that forecasts future real GDP growth based only on its past values. Although this model is the most commonly used benchmark model in the literature, it is not likely to be very successful at short forecasting horizons, since it does not use higher-frequency information from monthly indicators of economic activity. The second benchmark is an AR model that forecasts quarterly real GDP at market prices based on the monthly series for real GDP at basic prices.[7] The third benchmark is a forecasting strategy in which, at each period, a researcher would select the model that has been most accurate up to that period (best ex ante) and use it to forecast the next quarter.

5 The correlations between forecasts are identical if, for example, the forecasts from model A and model B have a correlation of 0.7 and the forecasts from model A and model C as well as forecasts from model B and model C also have a correlation of 0.7.

6 When the observations available to the researcher are few, the estimates might not be precise, i.e., they might be biased.

7 Basic prices exclude taxes and subsidies on products. The two measures of real GDP are highly correlated on a quarterly basis (at 0.99 since 2007). The benchmark model forecasts the growth rate of real GDP at basic prices on a monthly basis using its own lags. It then aggregates the monthly forecasts to quarterly forecasts. The quarterly growth rate of real GDP at basic prices is taken as a forecast of the quarterly growth rate of real GDP at market prices.

Results

Table 2 shows the RMSPE for forecast horizon $h = 1, \ldots, 4$, for each benchmark model and for the most accurate combination of models. The weights assigned to the models are initially computed on the sample from the 1999Q4–2005Q1 period. The benchmarks and combinations are then evaluated, based on their RMSPEs, over the remaining quarters, from 2005Q2 to 2011Q2.

For each forecasting model and combination, the RMSPE increases with the forecast horizon. Conversely, forecasts become more accurate as the release date approaches, since more information is available.

Combined forecasts are substantially more accurate than the AR benchmark model for real GDP growth at any forecast horizon considered, and the relative performance of the best combination improves as the release date approaches.

The comparison of the best combination and the AR model based on monthly real GDP at basic prices is limited to backcasting and nowcasting, where the relative performance of the quarterly AR model was least successful. Overall, the best combination proves more accurate than the AR monthly model, with relative gains rising as the forecast horizon increases. When backcasting immediately before the release of the quarterly National Accounts, however, the simple benchmark based on monthly GDP data has a slightly lower RMSPE than the combination. This is because two out of the three monthly figures of real GDP at basic prices are available and aggregating them to the quarterly frequency provides a very accurate indicator of real GDP at market prices (see Binette and Chang in this issue).

A comparison of the RMSPE of the best ex ante model with the RMSPE from the best combination indicates that relying on a single forecasting model will typically decrease accuracy: the best model is difficult to identify in advance, and choosing only the model that has performed best up to the time of the forecast produces systematically worse results than combining models.

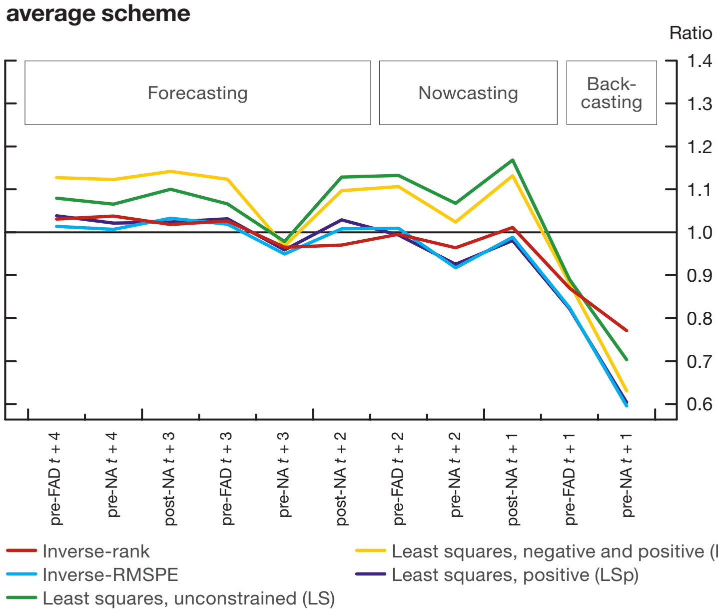
Choosing only the model that has performed best up to the time of the forecast produces systematically worse results than combining models

Table 2: Root-mean-square prediction errors (RMSPEs) from benchmark models and combinations, 2005Q2–2011Q2

Horizon (steps ahead)	Timing of forecast	Benchmark			Best combination[a]	
		Autoregressive	Autoregressive monthly	Best ex ante	RMSPE	Scheme
$t + 4$	pre-FAD	3.92		3.14	2.89	SA
$t + 4$	pre-NA	3.92		3.15	2.91	SA
$t + 3$	post-NA	3.92		2.96	2.75	SA
$t + 3$	pre-FAD	3.60		2.89	2.62	SA
$t + 3$	pre-NA	3.60		2.52	2.46	I-RMSPE
$t + 2$	post-NA	3.60		2.38	2.30	I-RANK
$t + 2$	pre-FAD	2.97	4.73	2.31	2.15	I-RANK, LSp
$t + 2$	pre-NA	2.97	3.08	2.12	1.78	I-RMSPE
$t + 1$	post-NA	2.97	1.87	2.19	1.71	LSp
$t + 1$	pre-FAD	2.90	1.42	1.29	1.27	I-RSMPE, LSp
$t + 1$	pre-NA	2.90	0.68	0.73	0.78	I-RMSPE

pre-FAD: forecasts produced about one week before the first fixed announcement date in a quarter
pre-NA: forecasts produced immediately before the release of real GDP (also referred to as the National Accounts)
post-NA: forecasts produced immediately after a real GDP release
a. These two columns show the results from the combination with lowest RMSPE for each horizon.

Chart 1: Relative RMSPEs of performance-based schemes over the simple-average scheme



pre-FAD: forecasts produced about one week before the first fixed announcement date in a quarter
pre-NA: forecasts produced immediately before the release of real GDP (also referred to as the National Accounts)
post-NA: forecasts produced immediately after a real GDP release
Source: Bank of Canada calculations

The performance of different combination schemes is now compared. Because of the documented success of the simple average, the performance-based combination methods are evaluated against the easier-to-implement SA scheme. **Chart 1** shows, for each forecast horizon, the RMSPE of each performance-based combination method relative to the RSMPE of the simple average. A number above (below) one suggests that a performance-based combination method is less (more) accurate than the SA scheme.

To begin, we focus on longer forecast horizons. In line with the previous literature, we find that the SA scheme generally performs well at longer horizons; however, the improvements in accuracy over the other combination schemes are not uniform. The increased accuracy in forecasts using a simple average compared with two of the least-squares variants is substantial (for example, the ratio of the LSnp RMSPE to the SA RMSPE is 1.14), while it is more modest with respect to other methods. This supports the finding that pooling techniques that take into account the correlation of the forecasts (i.e., LS) might be affected by small-sample estimation bias. As **Chart 1** shows, this bias is reduced by imposing the constraint that weights are positive and sum to one (LSp), or by using combination techniques that do not estimate the correlation across forecasts (I-RMSPE or I-Rank). These weighting schemes reduce the uncertainty around the estimates of the weights and can therefore improve the performance of the combination.

Performance-based weights deliver more accurate combinations than equal weights when nowcasting or backcasting

Performance-based weights deliver more accurate combinations than equal weights when nowcasting or backcasting.[8] The improvement is particularly substantial when backcasting, with the relative RMSPE plunging to 0.6 (**Chart 1**). Consistent with the previous discussion, there are significant gains from unequal weighting, because the forecast accuracy of individual models varies greatly at

8 At these shorter horizons, the improvements in accuracy obtained by allowing for performance-based weights more than compensate for the uncertainty introduced by computing the weights in the I-RMSPE or I-Rank schemes.

this horizon (with their RMSPEs ranging from 0.73 to about 2.79) and, among the models, the forecasts tend to be very different (correlation across individual models can be as low as 0.2 and as high as 0.87 for the pre-NA $t + 1$ horizon). In contrast, the optimal weights are close to equal weights for longer horizons, since the model forecasts tend to converge to the mean of real GDP growth and forecasts across models have similar correlation.[9]

Overall, the performance-based schemes that do not take into account the correlations between forecast errors, in particular, the I-RMSPE, are the most robust combination schemes across forecast horizons, since they achieve increased forecast accuracy at shorter horizons and their performance is comparable with the simple average at longer horizons.

The performance-based schemes that do not take into account the correlations between forecast errors are the most robust across forecast horizons

Conclusion

Combining forecasts from several models is more accurate than relying on a single model at all horizons. At longer horizons (three to four quarters ahead), the simple-average scheme outperforms, or does as well as, more sophisticated weighting schemes that take into account the past performance of the models. These results are in line with previous studies.

However, in contrast to much of the existing literature, combined forecasts using performance-based weights significantly increase accuracy at shorter horizons. This result occurs because the models we consider produce very different forecasts at these horizons. Some models are more accurate than others and therefore receive more weight.

Our results support the Bank's approach of using a wide range of models in a flexible manner rather than relying solely on a single model. Although the set of models the Bank uses changes over time, the finding that there are gains from combining forecasts is likely to be an enduring result.

There are certain caveats associated with our work, however. First, because some of the required real-time data were not available, our assessment of model combinations does not take into account the implications of data revisions for forecast accuracy. That is, when simulating the models, we use the 2011Q2 vintage of data rather than the data that were available at the time forecasts were made. Second, the sample we use for estimating models and assessing forecasts is small. The accumulation of progressively longer time series will help to address this limitation in future work.

9 For example, for the pre-FAD $t + 4$ horizon, the difference between the lowest and highest RMSPE is 0.20 and the correlation across model forecasts ranges from 0.87 to 0.99.

Literature Cited

Binette, A. and J. Chang. 2013. "CSI: A Model for Tracking Short-Term Growth in Canadian Real GDP." *Bank of Canada Review* (Summer): 3–12.

Bjørnland, H. C., K. Gerdrup, A. S. Jore, C. Smith and L. A. Thorsrud. 2012. "Does Forecast Combination Improve Norges Bank Inflation Forecasts?" *Oxford Bulletin of Economics and Statistics* 74 (2): 163–79.

Coletti, D. and S. Kozicki. 2013. "Introduction: Tools for Current Analysis at the Bank of Canada." *Bank of Canada Review* (Summer): 1–2.

Smith, J. and K. F. Wallis. 2009. "A Simple Explanation of the Forecast Combination Puzzle." *Oxford Bulletin of Economics and Statistics* 71 (3): 331–55.

Stock, J. H. and M. W. Watson. 1999. "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series." In *Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive W. J. Granger*, 1–44. Edited by R. F. Engle and H. White. Oxford: Oxford University Press.

———. 2004. "Combination Forecasts of Output Growth in a Seven-Country Data Set." *Journal of Forecasting* 23: 405–30.

Timmermann, A. 2006. "Forecast Combinations." *Handbook of Economic Forecasting*, 135–96. Edited by G. Elliott, C. W. J. Granger and A. Timmermann. Amsterdam: North Holland.

Monitoring Short-Term Economic Developments in Foreign Economies

Russell Barnett and Pierre Guérin, International Economic Analysis

- Assessing the economic prospects of key foreign economies—the United States, the euro area, Japan and China—is necessary because of their important direct and indirect links to the Canadian economy.
- The forecasting models constructed for each of these economies take into account the level of detail required for each region and key features of the data, such as timeliness and volatility.
- Forecasts from different models are typically combined to mitigate model uncertainty, and judgment is applied to the model forecasts to incorporate information that is not directly reflected in the most recent indicators.

Current and future developments in foreign economies can have important consequences for the conduct of domestic monetary policy in Canada because of the extensive linkages between the Canadian economy and the rest of the world through trade, commodity price, confidence and financial channels.[1] The International Economic Analysis Department at the Bank of Canada therefore carefully assesses the economic prospects of key foreign economies, specifically, the United States, the euro area, Japan and China.[2]

The Bank faces a number of challenges when building short-term forecasting models for these economies (i.e., for the current quarter and the next),[3] including the timeliness of data releases, reliance on data that may be volatile and subject to historical revisions, and the short sample periods for some variables. The level of detail required in the forecasts is another important consideration. Since the characteristics of the available data and the Bank's forecasting requirements differ across economies, a tailored

1 The importance of foreign shocks is illustrated by noting that roughly one-third of the variables used in Canada's Short-Term Indicator (CSI) model are related to foreign economic indicators, commodity prices or Canada's terms of trade (see Binette and Chang in this issue).

2 The Bank also monitors a number of other economies/regions and is currently expanding its coverage to include other major emerging markets.

3 Coletti and Kozicki (this issue) discuss the role of short-term forecasts in the Bank's economic projections. Macklem (2002) describes how economic projections, in conjunction with other Bank analysis and information, influence monetary policy decisions in Canada.

approach generates more-accurate forecasts. To this end, the Bank's choice of short-term forecasting models is aimed at meeting the challenges and needs for each country or region.

This article discusses the Bank's approach to assessing the short-term economic prospects of each of the four key foreign economies that it monitors closely. The influence of specific challenges on the choice of modelling approach for each economy is also highlighted. The concluding remarks suggest avenues for future work.

Monitoring the Global Economy

United States

The United States is Canada's largest trading partner, accounting for roughly 75 per cent of Canadian exports. It also plays an important role in the determination of commodity prices and financial conditions in the global economy, which can have a significant impact on the Canadian economy. Analyzing and forecasting U.S. economic conditions in detail is therefore essential, and this requirement is a key consideration in producing short-term forecasts of the U.S. economy. Unlike the other economies discussed in this article, where the primary focus is on aggregate real GDP growth and inflation, the Bank analyzes the U.S. economy on a disaggregated basis by producing individual forecasts for the major components of GDP (i.e., consumption, residential investment, business investment, inventory investment, government spending, exports and imports). This level of granularity is important, since forecasting Canadian exports is significantly improved by focusing on the components of U.S. GDP (Morel 2012). For example, the Bank's foreign activity measure, which captures the composition of foreign demand for Canadian exports, attaches a much larger weight to U.S. business and residential investment than that implied by their respective nominal shares of U.S. GDP (**Table 1**).[4]

The Bank of Canada analyzes the U.S. economy on a disaggregated basis by producing individual forecasts for the major components of GDP

Although this greater level of detail is necessary, it creates an additional challenge for Bank staff, since forecasts for individual components of GDP must also be consistent with a coherent view of overall economic conditions in the United States. To meet this challenge, the Bank uses a combination of error-correction (EC) and indicator models, as well as staff judgment,[5] to produce short-term forecasts for most components of U.S. GDP (see **Box 1** for a description of the models). The EC model, which incorporates a long-run behavioural relationship between variables, allows economic theory to help guide the short-term forecast, in particular, during the early part of any given quarter when few, if any, published monthly economic indicators are available.

Table 1: The composition of the foreign activity measure and U.S. GDP

	Weight in the foreign activity measure	Nominal share of GDP (2012)
U.S. consumption	0.207	0.686
U.S. residential investment	0.175	0.027
U.S. business investment	0.486	0.121

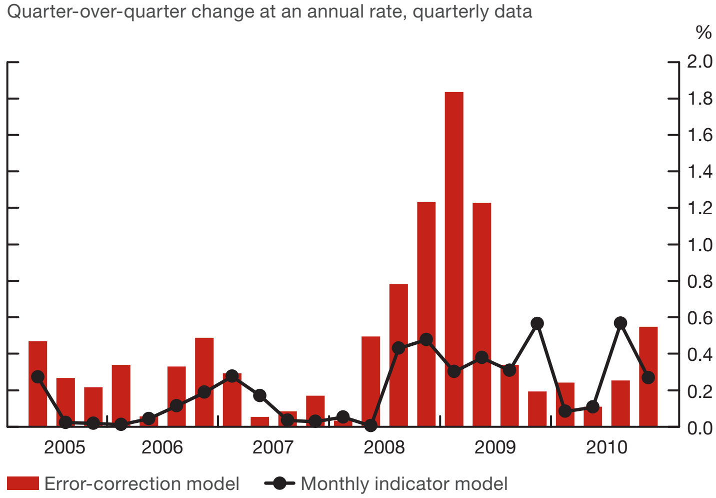
Sources: U.S. Bureau of Economic Analysis and Bank of Canada calculations

4 The foreign activity measure also assigns a weight of 13.2 per cent to foreign GDP outside the United States.

5 The relevance of judgment for macroeconomic forecasting is illustrated in Wright (2013), who finds that incorporating macroeconomic predictions from a survey of experts (i.e., judgment-based forecasts) in standard macroeconomic forecasting models yields substantial forecasting gains.

Chart 1: Absolute forecast errors in models of U.S. personal consumption expenditures at the end of each quarter, 2005Q2–2010Q4

Quarter-over-quarter change at an annual rate, quarterly data



Sources: U.S. Bureau of Economic Analysis and Bank of Canada calculations   Last observation: 2010Q4

At the same time, the indicator models allow staff to incorporate the most recent high-frequency data that are released throughout the quarter. The short-term forecast is typically generated using a weighted average of forecasts from different models (plus judgment), with the weights updated as new information becomes available during the quarter. Recently published data are also subject to revision, which often leads staff to reassess current economic conditions to reflect this additional information (see **Box 2** on page 29 for a discussion of data revisions to selected indicators).

To illustrate, consider how the short-term forecast for U.S. personal consumption expenditures (PCE) is produced. Bank staff have found that, early in the quarter, the indicator models add little additional information relative to the EC model. Therefore, at the beginning of a quarter, the short-term PCE forecast is heavily weighted toward an EC model that captures the behavioural response of consumption to movements in income, wealth and interest rates (Gosselin and Lalonde 2005). As the quarter proceeds, however, a number of important high-frequency indicators for consumption—such as motor vehicle sales, retail trade, consumer confidence and electricity output—become available. Each of these data series feeds into at least one of the indicator models used to generate forecasts of monthly and quarterly real PCE. The forecasts from these indicator models are then assigned a weight relative to the forecast from the EC model. This weight is continually updated throughout the quarter as more monthly data become available. Eventually, to minimize forecast errors, most of the weight will be given to the monthly data and indicator models, leaving the EC model with only a small weight. **Chart 1** shows how, over the 2005Q2–2010Q4 period, the forecast errors in one of the Bank's monthly indicator models of PCE were substantially lower than those of the EC model at the end of most quarters, once all the monthly indicators were available.[6]

Each of the components of U.S. real GDP is analyzed using a similar framework, although the weights assigned to the EC and the indicator models differ across components. Once the short-term forecasts of each

6 The 2005Q2–2010Q4 period was chosen to approximate the evaluation samples in the studies cited in Table 1-A of Box 1. The root-mean-square prediction error (RMSPE) of the indicator model was a little less than half that of the EC model over the sample period.

Box 1

Short-Term Forecasting Models for Monitoring Foreign Economies

The Bank currently uses three main types of model to monitor foreign economies:

(i) Error-correction (EC) models: In an EC model, economic variables are tied together by a long-term behavioural relationship, which is based on economic theory. This long-term relationship is complemented by other indicator variables to capture shorter-term dynamics. EC models can be particularly beneficial when few monthly indicators for the current quarter are available, and over medium-term horizons when economic theory can serve as an accurate anchor for forecasting.

For example, the EC model of U.S. personal consumption expenditures includes an estimated long-term relationship between consumption (C) and human wealth (HW), real housing wealth ($HOUSING$), real financial wealth (FIN) and the real interest rate (R), where the variables are included in this relationship based on economic theory[1]:

$$lC_t^* = -2.47 + 0.74 * lHW_t + 0.09 * lHOUSING_t + 0.17 * lFIN_t - 0.40 * R_t .$$

This long-term relationship is complemented with lags in consumption growth (C_{t-i}), expected growth in the desired level of consumption (C^*), movements in oil prices (OIL) and growth in real disposable income ($YPDI$) to capture some of the remaining short-run variation[2]:

$$dlC_t = \underbrace{0.34 * dlC_{t-1} + 0.11 * dlC_{t-2} + 0.55 * dlC_t^* + 0.19 * dlYPDI_t + 0.01 * dlOIL_t}_{\boxed{\text{Short-term dynamics}}}$$

$$\underbrace{- 0.09 * (lC_{t-1} - lC^*_{t-1})}_{\boxed{\text{Error correction}}}$$

The long-term relationship affects the dynamic short-term forecast through the error-correction term in the above equation. The error-correction term causes current consumption to adjust to the value suggested by the long-term relationship based on economic fundamentals.

(ii) Factor models: These models are based on the idea that the information in a large number of data series and indicators can be summarized in a few factors that describe the underlying trend in the data (see **Box 1** in Binette and Chang on page 5 of this issue).

(iii) Indicator models: These models are more parsimonious in terms of information content, since they use just a few indicators to forecast the variable of interest. For example, models based on industrial production or survey indicators such as the purchasing managers' index (PMI) are typically used to predict GDP growth. Often, the results of several indicator models are combined to form a single forecast.

(continued...)

1 For more details, see Gosselin and Lalonde (2005).

2 The EC model is estimated in log-level (l) and difference-of-log (dl) format.

Box 1 (*continued*)

Table 1-A shows the performance of factor and indicator models for predicting GDP growth in the euro area, Japan and China. Forecasts are evaluated using the root-mean-square prediction error (RMSPE) criterion relative to the RMSPE obtained from a univariate autoregressive (AR) model. As the forecasting horizon shortens, the forecasting performance of both the factor model and the indicator model frequently improves relative to the AR model. For Japan and China, the factor model exhibits a substantially lower RMSPE than the indicator model, while, for the euro area, the forecasting performance of the indicator model is similar to that of the factor model.

Table 1-A: Forecasting GDP growth—root-mean-square prediction error relative to the autoregressive model

Forecast horizon (months)	0	1	2	3	4	5
Factor model						
Euro area	0.58	0.74	0.80	0.79	0.94	0.98
Japan	0.37	0.57	0.81	1.28	1.47	1.12
China	0.47	0.55	0.82	0.42	0.52	0.86
Indicator model						
Euro area	0.64	0.68	0.87	0.88	0.94	0.96
Japan	0.75	0.92	1.30	1.18	1.20	1.24
China	0.81	0.82	0.87	0.90	0.90	0.95

Note: This table shows the RMSPE of a factor model and an indicator model relative to an AR model (results are from Lombardi and Maier (2010) for the euro area, Godbout and Lombardi (2012) for Japan, and Maier (2011) for China). The evaluation samples are from 2005Q2 to 2010Q1 for the euro area, 2006Q2 to 2010Q2 for Japan and 2008Q2 to 2010Q4 for China. Forecasts with horizon $h = \{0, 1, 2\}$ refer to forecasts for the current quarter, and forecasts with horizon $h = \{3, 4, 5\}$ refer to forecasts for the next quarter. The indicator is the headline composite PMI for the euro area, PMI manufacturing for Japan and the Hong Kong Monetary Authority indicator for China. Additional details on the models are provided in the respective studies.

component are produced, they are compiled to produce an aggregate GDP forecast as well as a forecast of the Canadian foreign activity measure. Although a similar approach is used for each component, it is important to note that judgment is sometimes added to complement the model's forecast. Many unexpected events occur that the models cannot capture in real time, for example, Hurricane Sandy in October 2012, the 2012 drought, and the Los Angeles and Long Beach port strikes late last year. When such events take place, analysts need to combine their judgment with the results of monitoring models to produce a more reliable short-term forecast.

Euro area

The monitoring tools developed for the euro area focus on the challenges related to the substantial lag in the publication of important "hard" indicators. Variables such as industrial production and GDP are released about 45 days following the end of the reference period (**Table 2**).[7] In comparison, the first estimates of GDP in China and the United States are released only 15 days and 30 days, respectively, after the end of the reference period. The survey data (or "soft" indicators) that are readily available at the end of the month (e.g., the purchasing managers' index (PMI)) deliver timely updates on the current business-cycle conditions and therefore receive greater weight in the monitoring of the euro area.[8]

The monitoring tools developed for the euro area focus on the challenges related to the substantial lag in the publication of important "hard" indicators

7 The reference period is one quarter for GDP and one month for industrial production.

8 Angelini et al. (2011) also underline the importance of survey data for predicting euro-area GDP growth, owing to the timeliness of the data.

Table 2: Publication lags for key indicators in the United States, the euro area, Japan and China

	GDP growth				Industrial production			
	United States	Euro area	Japan	China	United States	Euro area	Japan	China
Publication lag (days)	30	45	45	15	15	45	30	10–15

Sources: U.S. Bureau of Economic Analysis; Eurostat; Cabinet Office, Government of Japan; and National Bureau of Statistics of China

The Bank's monitoring models for the euro area build on the work of Lombardi and Maier (2010), which compares the forecasting performance of dynamic factor models[9] that exploit the informational content of a wide range of variables with that of indicator models that are based solely on the composite PMI (**Box 1**). Lombardi and Maier find that the PMI indicator model provides better forecasting results for the growth in euro-area GDP during the Great Recession of 2008–09 than the factor model, although the latter provides better forecasts during the preceding period, from 2000 to 2007. While this comparison highlights the general value of using information from many sources, it also suggests that models based exclusively on survey data may adjust more quickly to rapidly changing economic conditions. Factor models, which summarize information across a large number of indicators (including hard indicators released with a substantial delay), may be more sluggish to react to rapidly evolving economic conditions. This observation suggests that the weights allocated to different models often require adjustment based on judgment in the face of quickly changing economic conditions.

Japan

Forecasting growth in Japanese GDP is more challenging than for the other economies under consideration, owing to the volatility of Japanese macroeconomic aggregates, large data revisions (**Box 2**) and the substantial shocks to the Japanese economy observed over the past few decades (see, for example, Stock and Watson (2005)). In such an environment, factor models are often considered to be useful forecasting tools, since they summarize information from a large set of indicators, thereby potentially mitigating problems related to data volatility and revisions to individual series. Godbout and Lombardi (2012) compare the forecasting performance of two factor models for Japan with that of an indicator model based on the PMI and a simple model where GDP growth depends only on its own previous values (i.e., an autoregressive (AR) model). They find that factor models provide a greater number of accurate forecasts than both the AR benchmark model and the PMI model (**Box 1**). **Chart 2** shows that the most important explanatory factor extracted from principal-component analysis is related primarily to industrial production, real exports, survey data indicators (PMI manufacturing and its new orders subindexes), industrial activity and the Chinese PMI indicator for manufacturing output.[10] These empirical findings highlight the importance of foreign trade to the Japanese real economy and the relevance of survey data for monitoring its evolution.

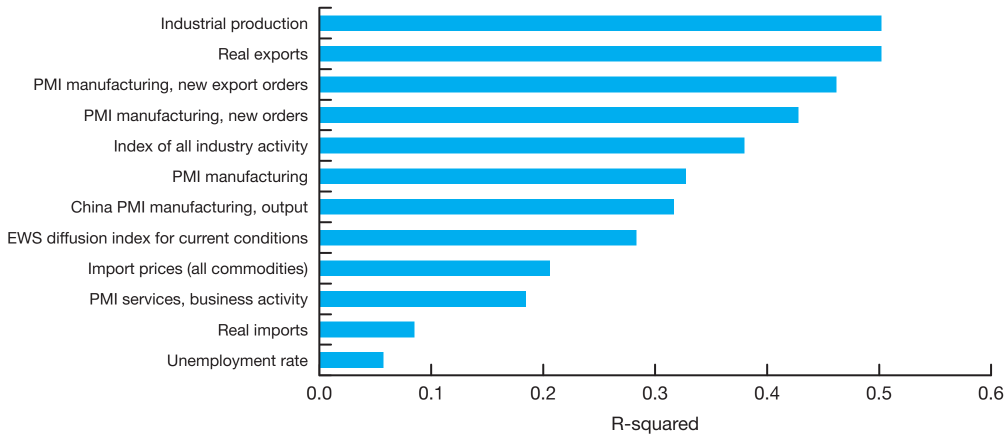
Forecasting growth in Japanese GDP is more challenging, owing to the volatility of Japanese macroeconomic aggregates, large data revisions and the substantial shocks to the Japanese economy observed over the past few decades

9 Dynamic factor models are described in Binette and Chang (this issue).

10 The R-squared in Chart 2 indicates the relative importance of selected individual time series in explaining the variations in the factor.

Chart 2: Relative importance of variables in forecasting growth in Japanese real GDP

R-squared between factor and selected individual time series



Note: The estimation sample is from October 2001 to December 2010.
PMI = purchasing managers' index; EWS = Economy Watchers Survey
Source: Bank of Canada calculations

China

The swift emergence of China as a major player in the global economy and, in particular, its significant influence on commodity prices (IMF 2011)[11] have triggered the need to develop monitoring tools that address a number of challenges specific to the Chinese economy. These challenges include such things as the shortness of available time-series data and inconsistencies in statistics (e.g., the quarterly GDP growth series does not necessarily add up to the annual GDP growth series). In addition, the rapid changes to the Chinese economy are likely to give rise to structural breaks in the data, which complicate attempts to design robust forecasting models. Maier (2011) evaluates the forecasting performance of a factor model and a set of indicator models comprising 33 indicators. He finds that both the factor model and a weighted average of the forecasts from the indicator models strongly outperform a standard AR benchmark model in forecasting Chinese GDP growth (**Box 1**). Indeed, the factor model closely tracks Chinese GDP growth in the current quarter (**Chart 3**). Indicators such as electricity and industrial production, as well as the PMI manufacturing component and Chinese equity prices, prove to be the most relevant for the indicator models. Forecasting performance is further improved when the forecasts from the factor model are combined with the forecasts from the indicator models. Overall, Maier's results suggest that there are significant gains in forecasting accuracy when forecasting methods are combined, likely because of the significant structural changes to the Chinese economy over the past few decades. Other studies of forecasting performance also find that model combinations or model averaging can improve forecasting performance (see Granziera, Luu and St-Amant in this issue).[12]

The swift emergence of China as a major player in the global economy has triggered the need to develop monitoring tools that address a number of challenges specific to the Chinese economy

11 According to estimates by the International Monetary Fund (IMF) of the purchasing-power-parity valuation of Chinese GDP, China accounted for 14.3 per cent of global GDP in 2011, compared with 7.5 per cent in 2001 and 4.1 per cent in 1991. Likewise, based on U.S. Energy Information Administration estimates, China's share of global primary energy consumption was 20.1 per cent in 2011, compared with 10.6 per cent in 2001 and 7.9 per cent in 1991.

12 Using data for six industrialized countries, Kuzin, Marcellino and Schumacher (2013) find that pooling nowcasts provides more-stable forecasts than selecting a single forecasting model sequentially, based on statistical information criteria.

Box 2

Volatility and Revisions

Data volatility and revisions present challenges for assessing the state of an economy. The volatility of macroeconomic indicators may lead to a greater reliance on a specific class of models, since it is more difficult to extract the underlying trend in economic conditions when data are highly volatile. Thus, factor models are often seen as useful devices to mitigate the volatility of indicators. Data revisions complicate monitoring because the historic data used in the analysis are not known with certainty.

Table 2-A provides statistics on the mean growth and volatility (standard deviation) of quarterly GDP and of monthly industrial production in the United States, the euro area and Japan from late 2001 to the end of 2010. Revisions to growth estimates are also included. Data revisions compared with the estimates available 4 months, 8 months and 16 months after the reference period are reported.[1] Following Giannone et al. (2012), we use the estimates available 24 months after the end of the reference period as a "true value" to calculate revision statistics.

A few key observations are worth noting. First, the volatility of GDP and industrial production varies across regions, with the Japanese data exhibiting the largest standard deviation for both series. Second, on average, over the sample examined, GDP growth tends to be revised down in the United States and Japan, and revised up in the euro area. Third, the standard deviation of the revisions to Japanese GDP estimates available four months after the end of the quarter (2.56 per cent) is about four times larger than the standard deviation of the revisions to euro-area GDP (0.57 per cent), and about twice as large as that for U.S. GDP (1.09 per cent). Similarly, industrial production, a monthly series, is also subject to substantial revisions.[2] Survey data, such as the purchasing managers' index, have an advantage over hard indicators, such as industrial production, since they are available on a more timely basis and are typically not subject to revisions.[3]

Table 2-A: Revision statistics for key indicators in the United States, the euro area and Japan

	GDP growth (per cent, quarter over quarter, seasonally adjusted annual rate)			Industrial production (per cent, year over year)		
	United States	Euro area	Japan	United States	Euro area	Japan
Mean	2.05	1.01	0.72	-0.01	0.16	0.53
Mean of revisions						
$Y_{t\|t+24} - Y_{t\|t+4}$	-0.30	0.09	-0.57	-0.44	0.13	0.13
$Y_{t\|t+24} - Y_{t\|t+8}$	-0.23	0.04	-0.32	-0.45	0.10	0.09
$Y_{t\|t+24} - Y_{t\|t+16}$	-0.23	0.09	-0.38	-0.34	0.02	0.02
Standard deviation (volatility)	2.74	2.61	5.18	5.15	6.32	10.96
Standard deviation of revisions						
$Y_{t\|t+24} - Y_{t\|t+4}$	1.09	0.57	2.56	0.78	0.50	0.83
$Y_{t\|t+24} - Y_{t\|t+8}$	1.11	0.44	2.36	0.72	0.37	0.67
$Y_{t\|t+24} - Y_{t\|t+16}$	0.73	0.21	1.70	0.77	0.26	0.21

Note: We calculate revision statistics as in Giannone et al. (2012). We report the mean and the standard deviation of the revisions defined as $(Y_{t|t+24} - Y_{t|t+i})$, where t indicates the reference period, while $t+i$ for i = {4, 8, 16} is the time in which the value of the series is observed. We consider the observation available two years later $(Y_{t|t+24})$ as the "true" value. We report the mean and standard deviation of this series to provide a benchmark for assessing the mean and standard deviation of the revisions. For quarterly GDP growth, all statistics refer to the 2001Q3–2010Q4 period. For industrial production, all statistics refer to the September 2001–December 2010 period.

Sources: Organisation for Economic Co-operation and Development real-time database, except euro-area industrial production, from the European Central Bank real-time database

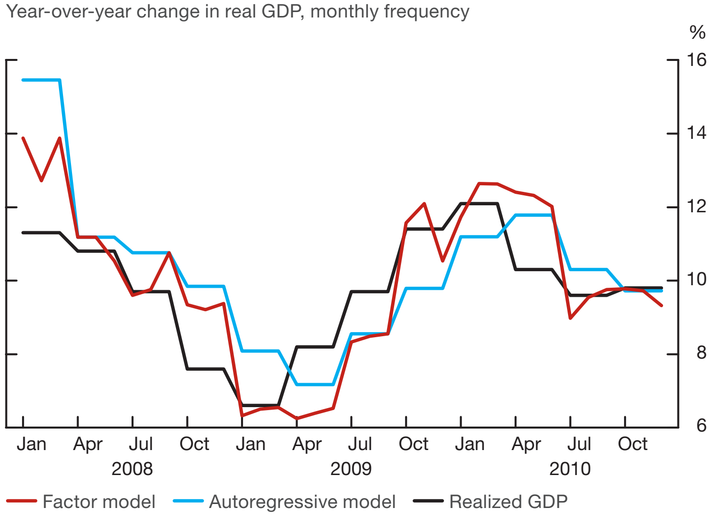
1 Revision statistics are not relative to the initial estimates, since these data are not fully available for the euro area and Japan over the full estimation sample. However, calculating revision statistics using the vintages available 4 months, 8 months and 16 months after the end of the reference period illustrates the trends of data revisions.

2 However, unlike revisions to GDP growth, the standard deviation of the revisions to the estimates of Japanese industrial production is similar to, or lower than, those for the United States or the euro area.

3 In most cases, all relevant information is available at the time of the publication of survey data. In contrast, revisions to hard indicators typically reflect the fact that additional (more accurate) information has become available.

Chart 3: Chinese GDP growth compared with factor-model and autoregressive-model forecasts

Year-over-year change in real GDP, monthly frequency



Sources: National Bureau of Statistics of China and Bank of Canada calculations

Last observation: December 2010

Conclusion

To better understand the evolution of foreign economies in the short term, Bank of Canada staff analyze an extensive set of indicators using a wide range of models that are selected based on the circumstances of the specific country and the level of detail required. Staff also use judgment in constructing forecasts to incorporate information that may not be directly reflected in the most timely high-frequency indicators. The Bank of Canada strives to improve the forecast accuracy of its short-term forecasting models. Avenues for future work include, but are not limited to, incorporating time variations in the parameters of the monitoring models to better account for the substantial volatility in some macroeconomic data, and incorporating density forecasts (i.e., a measure of uncertainty around mean forecasts).

Literature Cited

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin and G. Rünstler. 2011. "Short-Term Forecasts of Euro Area GDP Growth." *Econometrics Journal* 14 (1): C25–C44.

Binette, A. and J. Chang. 2013. "CSI: A Model for Tracking Short-Term Growth in Canadian Real GDP." *Bank of Canada Review* (Summer): 3–12.

Coletti, D. and S. Kozicki. 2013. "Introduction: Tools for Current Analysis at the Bank of Canada." *Bank of Canada Review* (Summer): 1–2.

Giannone, D., J. Henry, M. Lalik and M. Modugno. 2012. "An Area-Wide Real-Time Database for the Euro Area." *Review of Economics and Statistics* 94 (4): 1000–13.

Godbout, C. and M. J. Lombardi. 2012. "Short-Term Forecasting of the Japanese Economy Using Factor Models." Bank of Canada Working Paper No. 2012-7.

Gosselin, M.-A. and R. Lalonde. 2005. "MUSE: The Bank of Canada's New Projection Model of the U.S. Economy." Bank of Canada Technical Report No. 96.

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

International Monetary Fund (IMF). 2011. "Tensions from the Two-Speed Recovery: Unemployment, Commodities, and Capital Flows." *World Economic Outlook* (April).

Kuzin, V., M. Marcellino and C. Schumacher. 2013. "Pooling Versus Model Selection for Nowcasting GDP with Many Predictors: Empirical Evidence for Six Industrialized Countries." *Journal of Applied Econometrics* 28 (3): 392–411.

Lombardi, M. J. and P. Maier. 2010. "'Lean' Versus 'Rich' Data Sets: Forecasting During the Great Moderation and the Great Recession." Bank of Canada Working Paper No. 2010-37.

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

Maier, P. 2011. "Mixed Frequency Forecasts for Chinese GDP." Bank of Canada Working Paper No. 2011-11.

Morel, L. 2012. "A Foreign Activity Measure for Predicting Canadian Exports." Bank of Canada Discussion Paper No. 2012-1.

Stock, J. H. and M. W. Watson. 2005. "Understanding Changes in International Business Cycle Dynamics." *Journal of the European Economic Association* 3 (5): 968–1006.

Wright, J. H. 2013. "Evaluating Real-Time VAR Forecasts with an Informative Democratic Prior." *Journal of Applied Econometrics* 28 (5): 762–76.

Big Data Analysis: The Next Frontier

Nii Ayi Armah, Canadian Economic Analysis

- Current analysis is heavily dependent on data. The more timely, accurate and relevant the data, the better our assessment of the current state of economic activity.
- Technological advancements have provided an opportunity to exploit digital data from business transactions, social media and networked computers. The combination of all of these data is called "big data."
- Analysis of the vast quantities of digital information contained in big data can offer fresh insight for the monitoring of economic activity and inflation. Moreover, the timeliness of big data could improve real-time decision making for monetary policy.
- The potential of big data is, however, limited by challenges related to methodological constraints, a lack of easy access to the data and privacy concerns.

By providing an assessment of the present state of the economy, current analysis[1] contributes to the Bank of Canada's long-term macroeconomic projections, which in turn help to inform monetary policy decisions. Immediate and complete information about every economic and financial transaction within a country would improve current analysis by facilitating accurate and timely measurement of important macroeconomic indicators. Unfortunately, this ideal data set does not exist. The macroeconomic data produced by official statistical agencies are published with a lag and are subject to revision. Gross domestic product (GDP), for example, is a quarterly series that is published with a two-month lag and revised over the next four years. The consumer price index (CPI) is a monthly series that, although not subject to revision, is published three weeks after the end of the reporting month.

These issues with official data have led some researchers to explore the possibility of complementing official data with the use of "non-official" data that may be more timely.[2] An example of the early use of such data for current analysis is Lamont (1997), who finds that counting the frequency of appearances of the word "shortage" in print newspapers can be a good predictor of inflation in the United States. The Bank of Canada also uses non-official data to monitor the economy. For example, the Bank's regional

1 See Coletti and Kozicki (this issue) for a discussion of the role of current analysis in monetary policy.

2 There is a trade-off between the timely publication of official data and the accuracy of those data.

offices collect and analyze data obtained from quarterly consultations with businesses across Canada to gather their perspectives on such topics as demand and capacity pressures, as well as their views on economic activity in the future. These data, summarized in the *Business Outlook Survey*, provide a source of timely information that augments views gleaned from official data. With advances in technology, the proliferation of digital data and the declining cost of digital storage, another type of non-official data has recently emerged and is growing quickly—"big data."

Big data is a large-scale collection of information, some of which, such as business transactions, has always existed in corporate ledgers in the form of daily sales or inventory levels. Rich micro-level administrative data maintained by government agencies have also existed for some time. The high cost of retrieving and organizing all this information en masse has slowed the exploitation of these complementary sources of data for current analysis. However, digitization of previously paper-based data sources has made data much more accessible and easier to organize and analyze. Data emanating from services offered by public institutions or government agencies are a rich source of information on the behaviour of citizens. In addition, the rapid development of computer networking and the Internet has led to new sources of information in social media and web searches, as well as in electronic payments data such as credit card and debit card transactions. Since these data are ubiquitous and can be gathered quickly, they could provide more timely and detailed information about economic and financial transactions. Big data could therefore be another non-official resource and represents the next frontier in advancing current analysis. By providing an opportunity to exploit vast quantities of digital information, big data offers fresh insight that is relevant to the monitoring of economic activity and inflation. Its timeliness could augment official data to improve real-time decision making in current analysis, and it could also be an input in the construction of official statistics (Daas and van der Loo 2013).

This article describes the most important attributes of big data and discusses its possible applications and advantages for current analysis. Challenges that limit the full potential of big data, as well as initiatives to address these challenges, are then explored. Finally, the article concludes with the prospects for the use of big data in future current analysis.

What Is Big Data?

Big data refers to large and diverse digital data sets generated from economic transactions and social media interactions of billions of people around the world.

Big data refers to large and diverse digital data sets generated from economic transactions and social media interactions

The four Vs of big data

Big data has four main defining attributes: volume, variety, velocity and value. The **volume** of big data is typically much larger than that of traditional data sets. Manyika et al. (2011) describe the size of these data sets as being beyond the ability of typical database software tools to capture, store, manage and analyze. **Box 1** provides a sense of the magnitude of big data.

The types of information that constitute big data come from a **variety** of sources. Only about 10 per cent of big data is structured data (Gens 2011), the type that fits neatly into the rows and columns of relational databases. To be processed by traditional data-management tools and warehouses, and meaningfully interpreted by analysts, data must be in structured form. Examples of structured data are the transactional data that companies

Box 1

The Magnitude of Big Data

- A 2011 study by International Data Corporation (IDC) indicates that 1.8 zettabytes (1.8 trillion gigabytes) of data would be created that year (Gantz and Reinsel 2011). This amount of data would fill 57.5 billion 32-gigabyte iPads (EMC[2] 2011).
- Brands and organizations featured on Facebook receive 34,722 "likes" every minute of the day (Wikibon 2012).
- IDC estimates that the number of transactions between firms and those between firms and consumers will reach 450 billion per day by 2020 (Wikibon 2012).
- Walmart processes more than 1 million customer transactions every hour. These transactions are stored in databases that are estimated to contain more than 2.5 petabytes (2.5 million gigabytes) of data. This information would fill 167 times the number of books in the U.S. Library of Congress (Talend 2012).
- The Canadian Payments Association processed 6.3 billion individual retail payments in 2011 alone (Canadian Payments Association 2012).

collect on their customers, and the time-series data that statistical agencies collect on various macroeconomic and financial indicators. Unstructured data, which make up the remaining 90 per cent of big data, include emails, tweets, Facebook posts, road traffic information and audiovisual data. Traditional data warehouses strain under the load of unstructured data and typically cannot process them.

Velocity refers to the fact that data generated from some big data sources such as social media, mobile devices, Internet transactions and networked devices are updated very quickly. This creates an avalanche of data flows that overwhelms most traditional data-analysis hardware and software. Extracting value in real time from rapidly generated data requires specialized skills and data-analysis systems.

The ability to leverage insights and create significant **value** is the most important attribute of big data. Combining big data with sophisticated analytics could provide novel insights into household behaviour, firm expectations, financial market stability and economic activity that would support effective decision making. For example, these advanced methodologies are capable of analyzing patterns in a social network (which may be interconnected in highly complex ways) to determine how these interconnections could influence consumer expectations about inflation or other economic variables (see Einav and Levin 2013 for more details).

Combining big data with sophisticated analytics could provide novel insights into household behaviour, firm expectations, financial market stability and economic activity

Big Data and Current Analysis: A Glimpse into the Future

Since accurate and timely information about the current state of economic activity is important for monetary policy decisions, big data provides the opportunity to improve current analysis by exploiting digital data from economic transactions as well as by measuring consumer sentiment from social media and Internet searches. For example, existing monthly indicators could be combined with big data to predict GDP growth before official National Accounts data are released for a given quarter.[3]

An advantage to using big data is the ability to construct metrics that evolve quickly over time. The Billion Prices Project (BPP)[4] at the Massachusetts Institute of Technology, led by economists Alberto Cavallo and Roberto Rigobon,

3 Binette and Chang (this issue) describe a forecasting tool that uses a data set that, although large, is not of the magnitude of big data.

4 For more information on the Billion Prices Project, see http://bpp.mit.edu/.

calculates a daily inflation index from a continuously evolving basket of goods. Data for the BPP are collected with software that scours the websites of online retailers for their prices on a wide array of products.[5] The index is then calculated as an average of individual price changes. This virtual real-time inflation index could offer policy-makers and statistical agencies a glimpse of what is happening to inflation in real time. For example, BPP data show that, after Lehmann Brothers collapsed in September 2008, businesses started cutting prices almost immediately, suggesting that aggregate demand had weakened (Surowiecki 2011). In contrast, the official inflation numbers released by the statistical agencies did not show this deflationary pressure until that November, when October CPI data were released.

Canadians are increasingly moving away from traditional methods of payment, such as cash and cheques, toward a variety of electronic payment methods (Canadian Payments Association 2012). Analysis of these timely electronic data could help to predict economic activity and assess possible revisions to official retail and consumption data. Other research provides some evidence that payment-system data could be useful for studying the economic effects of occasional extreme events. For example, Galbraith and Tkacz (2013) use daily data on Canadian debit transaction volumes, as well as data on cheque transaction volumes and values, to investigate the impact on personal consumer expenditures of the 11 September 2001 terrorist attacks, the Severe Acute Respiratory Syndrome (SARS) epidemic in the spring of 2003, and the August 2003 electrical blackout in Ontario and in parts of Northeastern and Midwestern United States. Contrary to initial perceptions of these events, the authors find only small and temporary effects.

Big data could also be used to study developments in the labour and housing markets. Assessments of these markets have been carried out using data on the number of Internet searches. Choi and Varian (2009) find that unemployment and welfare-related searches can improve predictions of initial claims for unemployment benefits. Askitas and Zimmermann (2009), D'Amuri (2009), and Suhoy (2009) also find that Internet searches can be relevant for predicting labour market conditions in Germany, Italy and Israel, respectively. Choi and Varian (2011) as well as Wu and Brynjolfsson (2009) find that housing-related searches can improve on traditional models for predicting housing sales in the United States. Furthermore, Webb (2009) suggests that the high degree of correlation between the number of searches for "foreclosure" and the actual number of foreclosures can be the basis for an early-warning system to predict problems in the U.S. housing market.

McLaren and Shanbhogue (2011) examine the importance of online searches for predicting activity in the labour and housing markets in the United Kingdom. The authors specify two separate models in which either the growth in U.K. unemployment or growth in house prices is a function of previous growth rates. Their results indicate that the inclusion of Internet searches in these models improves the models' forecasting performance. McLaren and Shanbhogue (2011) point out that these data are particularly helpful for analyzing the impact of unexpected developments, such as temporary plant closures, epidemics and labour strikes. While survey data must be collected based on predetermined questions, Internet search data are more flexible and can be used to assess these special circumstances.

Finally, big data could be an input in the construction of official statistics. For example, some European countries are using point-of-sale scanner data in the compilation of their CPIs. Statistics Norway exploits scanner data to

◂ *Big data could be an input in the construction of official statistics*

5 The prices of services are not included in this data set.

compute a subindex for food and non-alcoholic beverages (Rodriguez and Haraldsen 2006). In June 2002, Statistics Netherlands introduced supermarket scanner data into its CPI (Schut 2002), and the Swiss Federal Statistical Office replaced the prices formerly collected in retail outlets with prices taken from scanner data to calculate its price indexes (Müller et al. 2006).

Challenges and Initiatives

Despite the innovation that has materialized from big data thus far, several factors limit its full potential, chief among them methodological constraints, the lack of easy access to data sets and privacy concerns.

Methodological constraints

Although strides have been made in developing methodologies for extracting value from big data, the implementation of these methodologies for current analysis is still evolving. Specifically, it remains unclear how best to select, organize and aggregate unstructured data so that they provide meaningful signals about economic conditions, and what analytical tools need to be developed to integrate those signals with information from conventional data sources. In addition, subsets of populations covered by big data are at times not necessarily representative of a relevant target population used for official statistics. Assessing how representative big data samples are could prove to be problematic for standard methodologies.

Although strides have been made in developing methodologies for extracting value from big data, the implementation of these methodologies for current analysis is still evolving

Lack of access to the data

Much of the data that constitute big data currently exist in silos. To unleash the full potential of big data, there is the need to first integrate the fragmented data sets so that they can be accessed easily and quickly by interested parties. The advent of cloud computing has enabled the creation of data centres that house massive amounts of data in one location. Since pooling of data sets is of paramount importance, a number of initiatives are under way to enhance access to big data. For example, the Government of Canada and the Ontario government have collaborated with IBM and a consortium of seven universities to establish a new Ontario-based, $210 million research project and data centre to help university and economic researchers use high-performance cloud-computing infrastructure to better exploit big data. Another initiative is an agreement between the U.S. Library of Congress and Twitter in 2010 to release 170 billion archived tweets to researchers and other interested parties exploring topics ranging from tracking vaccination rates to predicting stock market activity (Osterberg 2013).

Big data or Big Brother?

Information at the level of individual households and businesses can provide important insights into current economic conditions. By uncovering hidden connections between seemingly unrelated pieces of data, big data analysis can reveal personal information that some might deem too sensitive to share. The reasons for collecting and the need to protect these data are becoming more prominent issues in the debate over privacy and the appropriate use of personal data. Nevertheless, as much as institutions and individuals need to be careful about how invasive they are in their efforts to collect big data, it is difficult to deny that big data analysis has the potential to offer valuable information on economic growth and to improve current analysis. A balanced regulatory framework is therefore necessary to

A balanced regulatory framework is necessary to effectively address concerns about privacy, while still benefiting from technological advances effectively address concerns about privacy and the use of personal information, while still benefiting from technological advances and a thriving data-driven economy.

Conclusion

Reliable information about the current state of the economy is an important component for conducting monetary policy and, since data are the main resource driving current analysis, their accuracy and timeliness are key. A better real-time gauge of current economic conditions can improve assessments of economic momentum and forecasts of future growth. Digitization and the advent of the Internet have exponentially increased the amount of data available and also created new, viable sources. Practical applications of these data include the construction of timely price metrics from online retail prices, the use of electronic payment methods to help predict economic activity, and the use of Internet searches to assess labour and housing markets. Harnessing the full potential of this profusion of data is challenging. While some progress in unlocking the value of these data has been made with traditional data-analysis methods, the use of big data for current analysis is still in its infancy.

Literature Cited

Askitas, N. and K. F. Zimmermann. 2009. "Google Econometrics and Unemployment Forecasting." *Applied Economics Quarterly* 55 (2): 107–20.

Binette, A. and J. Chang. 2013. "CSI: A Model for Tracking Short-Term Growth in Canadian Real GDP." *Bank of Canada Review* (Summer): 3–12.

Canadian Payments Association. 2012. "Examining Canadian Payment Methods and Trends." (October).

Choi, H. and H. Varian. 2009. "Predicting Initial Claims for Unemployment Benefits." Google Inc. Available at http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/fr//archive/papers/initialclaimsUS.pdf.

———. 2011. "Predicting the Present with Google Trends." Google Inc. Available at http://people.ischool.berkeley.edu/~hal/Papers/2011/ptp.pdf.

Coletti, D. and S. Kozicki. 2013. "Introduction: Tools for Current Analysis at the Bank of Canada." *Bank of Canada Review* (Summer): 1–2.

Daas, P. and M. van der Loo. 2013. "Big Data (and Official Statistics)." Statistics Netherlands. Working Paper presented at the Meeting on the Management of Statistical Information Systems, Paris and Bangkok, 23–25 April.

D'Amuri, F. 2009. "Predicting Unemployment in Short Samples with Internet Job Search Query Data." Munich Personal RePEc Archive Paper No. 18403.

Einav, L. and J. D. Levin. 2013. “The Data Revolution and Economic Analysis.” National Bureau of Economic Research Working Paper No. 19035.

EMC[2]. 2011. “World’s Data More Than Doubling Every Two Years—Driving Big Data Opportunity, New IT Roles.” Press Release, 28 June.

Galbraith, J. W. and G. Tkacz. 2013. “Analyzing Economic Effects of September 11 and Other Extreme Events Using Debit and Payments System Data.” *Canadian Public Policy* 39 (1): 119–34.

Gantz, J. and D. Reinsel. 2011. “Extracting Value from Chaos.” International Data Corporation (IDC) Digital Universe. Available at http://www.emc.com/leadership/programs/digital-universe.htm.

Gens, F. 2011. “IDC Predictions 2012: Competing for 2020.” IDC Analyze the Future (December). Available at http://cdn.idc.com/research/Predictions12/Main/downloads/IDCTOP10Predictions2012.pdf.

Lamont, O. 1997. “Do ‘Shortages’ Cause Inflation?” In *Reducing Inflation: Motivation and Strategy*, 281–306. Edited by C. D. Romer and D. H. Romer. Chicago: University of Chicago Press.

Manyika, J., M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh and A. Hung Byers. 2011. “Big Data: The Next Frontier for Innovation, Competition, and Productivity.” McKinsey Global Institute (May).

McLaren, N. and R. Shanbhogue. 2011. “Using Internet Search Data as Economic Indicators.” Bank of England *Quarterly Bulletin* (Q2): 134–40.

Müller, R., H. M. Herren, S. Röthlisberger and C. Becker Vermeulen. 2006. “Recent Developments in the Swiss CPI: Scanner Data, Telecommunications and Health Price Collection.” Swiss Federal Statistical Office. Paper presented at the 9th meeting of the Ottawa Group, London, 14–16 May.

Osterberg, G. 2013. “Update on the Twitter Archive at the Library of Congress.” Library of Congress, 4 January.

Rodriguez, J. and F. Haraldsen. 2006. “The Use of Scanner Data in the Norwegian CPI: The ‘New’ Index for Food and Non-Alcoholic Beverages.” *Economic Survey* 4: 21–28.

Schut, C. (ed.). 2002. “Gebruik van Scannerdata van Supermarkten in de Consumentenprijsindex.” Statistics Netherlands, 4 July.

Suhoy, T. 2009. “Query Indices and a 2008 Downturn: Israeli Data.” Bank of Israel Discussion Paper No. 2009–06.

Surowiecki, J. 2011. “A Billion Prices Now.” *The New Yorker*, 30 May.

Talend. 2012. “How Big Is Big Data Adoption? Survey Results.” Available at http://info.talend.com/rs/talend/images/WP_EN_BD_Talend_SurveyResults_BigDataAdoption.pdf.

Webb, G. K. 2009. "Internet Search Statistics as a Source of Business Intelligence: Searches on Foreclosure as an Estimate of Actual Home Foreclosures." *Issues in Information Systems* 10 (2): 82–87.

Wikibon. 2012. "A Comprehensive List of Big Data Statistics." Wikibon Blog, 1 August. Available at http://wikibon.org/blog/big-data-statistics/.

Wu, L. and E. Brynjolfsson. 2009. "The Future of Prediction: How Google Searches Foreshadow Housing Prices and Sales." Sloan School of Management, Massachusetts Institute of Technology.