2016-06

Using Confidence Data to Forecast the Canadian Business Cycle

Kevin Moran
Simplice Aimé Nono
Imad Rherrad

novembre / november 2016

Centre de recherche sur les risques les enjeux économiques et les politiques publiques

www.crrep.ca

Abstract

This paper assesses the contribution of confidence - or sentiment - data in predicting Canadian economic slowdowns. A probit framework is specified and applied to an indicator on the status of the Canadian business cycle produced by the OECD. Explanatory variables include all available Canadian data on sentiment (which arise from four different surveys) as well as various macroeconomic and financial data. The model is estimated via maximum likelihood and sentiment data are introduced either as individual variables, as simple averages (such as confidence indices) and as confidence factors extracted, via principal components' decompositions, from a larger dataset in which all available sentiment data have been collected. Our findings indicate that the full potential of sentiment data for forecasting future business cycles in Canada is attained when all data are used through the use of factor models.

Kevin Moran : Department of Economics, Universit.é Laval, kevin.moran@ecn.ulaval.ca

Simplice Aimé Nono : Department of Economics, Laval University, simplice-aime.nono.1@ulaval.ca

Imad Rherrad : Ministère des Finances du Québec. Email: imad.rherrad@finances.gouv.qc.ca

1 Introduction

Survey data on consumer and business confidence – or *sentiment*– play important roles in the decision processes of monetary and government policy makers worldwide.[1] Interest for this type of data arises because of their timeliness and the fact that they are seldom revised. In addition, it reflects the belief that these data provide signals about current and future economic developments that complements the information embodied in standard time-series from financial markets or national accounts.[2]

The growing interest for such data has manifested itself in the establishment of several different surveys. In Canada, four major separate surveys regularly examine various aspects of consumer and business sentiment. First, the *Business Confidence Survey*, established in 1977 by the Conference Board, is a quarterly survey that queries managers of Canada's business organizations. Second, the *Consumer Confidence Survey*, originated by the same Conference Board in 1979, scrutinizes consumers' attitudes and optimism about their current and future economic prospects. Next, the more recently (1997) established *Business Outlook Survey*, organized and managed by the Bank of Canada's regional offices, is a quarterly consultation with businesses across Canada that covers topics related to demand conditions, productive capacity, prices and inflation. Finally, the Bank of Canada's *Senior Loan Officers Survey*, quarterly and established in 1999, analyzes the business-lending practices of major Canadian financial institutions.

This expanding diversity in the available data on business and consumer sentiment holds the potential to improve forecasts of future Canadian economic developments and thus lead to better decision making. However, it calls into question the most efficient use

[1]Throughout the paper we refer to 'confidence' and 'sentiment' data interchangeably. See Murray (2013) for a general overview of the decision process at the Bank of Canada and how sentiment data contribute to that process.

[2]Sentiment data might be able to signal future economic developments because they reflect information about future fundamental shocks that is not contained in other time-series; alternatively it may be that confidence and its evolution has a causal impact on future economic developments. See Barsky and Sims (2012) for a discussion.

of al these various data. Should forecasters focus on individual, particularly promising survey questions to obtain a parsimonious forecasting equation? Or should they instead use all available information, even at the risk of overfitting their models?

This paper provides an analysis of this question. To this end, we test several specifications of probit forecasting equations for Canada's future business cycle turning points. Throughout our analysis, the variable to forecast is the status of Canada's business cycle, as measured by the OECD.[3] The explanatory variables drawn from sentiment data include individual time-series resulting from some specific survey questions, simple aggregates of these time-series (such as *Index of Consumer Confidence* produced by the Conference Board of Canada) and estimated factors extracted from all survey data available for Canada. We assess the forecasting ability of these models by comparing them to those using the 'classical predictors' popularized in the literature on predicting recessions[4] as well as information extracted from a large, 144-variables dataset of Canadian macroeconomic, financial and national accounts data.

Our results indicate that sentiment data has substantial forecasting power for future status of the Canadian business cycle. Specifically, we first show that within the class of single-predictor probits, models using sentiment data produce results comparable to the best performances obtained using the classical predictors popularized by Estrella and Mishkin (1998). Next, we report strong evidence in favor of using multiple-predictors frameworks with confidence 'factors' extracted from all available sentiment; this happens because such factors are orthogonal one to the other and including additional such factors can only improve the performance of a given model. Indeed, we find that the statistical significance of each factor used remains high in these multiple-predictors models and that

[3] This measures stems from a growth-cycle framework for understanding business cycles (Zarnowitz and Ozyldirim, 2006). We use this measure because the alternative, the recession dates established by the C.D. Howe Institute for Canada, include only one recession in the last 25 years (since the early 1990s). Accordingly, we refer to *slowdowns* in the Canadian economy instead of recessions when discussing our models and our results. See Section 4 for a complete description of all data used in the present study.

[4] The term spread and the return on stock markets are two such classical predictors (Estrella and Mishkin, 1998).

measures of model performance increase relative to cases where only individual variables are used. We show that these findings are robust to varying the forecasting horizon and the sample used and that they are strengthened in an out-of-sample experiment. Overall, our result indicate that the full potential of sentiment data for forecasting business cycles is likely to be attained when all such data are used and amalgamated through factor models.

The rest of the paper is organised as follows. Section 2 discusses the related literature using sentiment data for forecasting or analysing economic fluctuations. Next, Section 3 discusses our probit forecasting framework and Section 4 provides a detailed description of all data used. Section 5 describes our results and Section 6 concludes by suggesting likely avenues for future research.

2 Related Literature

The expanding availability of survey data on sentiment has generated a growing empirical literature, which has tended to fall under two general themes; we review them in turn.

Forecasting with Confidence Data

The first major direction along which this literature has progressed assesses the ability of sentiment data to forecast future economic developments. Christiansen et al. (2014) is a recent, representative contribution to this research agenda. In that paper, the authors examine the forecasting ability of the *Consumer Confidence Index* and of the *Purchasing Managers' Index* for US recessions.[5] They use the probit framework popularized by Estrella and Mishkin (1998), which aims to forecast a binary 'recession' variable indicating whether the economy is experiencing a downturn or an expansion.[6] Christiansen et al.

[5]The *Consumer Confidence Index* is constructed from answers to the University of Michigan Survey questions. The *Purchasing Managers' Index* is produced by the *Institute of Supply Management* by aggregating survey answers from managers and purchasers at important manufacturing companies in the US. Christiansen et al. (2014) use the NBER dates to measure U.S. recessions.

[6]The Estrella and Mishkin (1998) strategy of identifying predictors of future recessions has been ex-

(2014) report that sentiment variables have substantial power to predict the occurrence of future downturns, both in-sample and out-of-sample. Specifically, sentiment variables, when taken individually, predict future downturns better than the ‘classical’ recession predictors identified by Estrella and Mishkin (1998) (the term spread and stock market indices). In addition, when sentiment variables are combined with other explanatory variables (including estimated factors extracted from a large dataset of macroeconomic variables) the model attains a superior forecasting performance. Taylor and McNabb (2007) present a similar analysis, applied to data from the UK, the Netherlands, France and Italy. They also report that sentiment, particularly data drawn from business surveys, can contribute significantly to forecasting economic downturns.

Researchers have provided evidence about the substantial forecasting ability of sentiment variables in other contexts. For example, recent work by Ollivaud et al. (2016) shows that small forecasting models for various OECD countries that include sentiment data among a very limited list of explanatory variables have the capacity to predict future economic developments as well as larger models drawing their information from multiple explanatory variables. In a related way, Hansson et al. (2005) use a VAR framework to show that survey data from the Swedish Business Tendency Survey can help forecast the growth of Swedish GDP, particularly when all sentiment data are aggregated via a dynamic factor model and forecasting is undertaken using the estimated factors. Their results are confirmed and extended to the case of the Norwegian business cycle by Martinsen et al. (2014). Additionally, Bodo et al. (2000) show that forecasting Industrial Production in the Euro area using an error-correction system that includes a business confidence index produces good results and Batchelor and Dua (1998) report that the Blue Chip *Consensus forecast* could have better predicted the 1991 US recession if the information contained in the Conference Board’s Consumer Confidence index had been taken into account.

As described above, a rich variety of survey data exists in Canada, with four major

tended by various authors (Kauppi and Saikkonen, 2008; Nyberg, 2010; Chen et al., 2011; Fornaro, 2016; Kotchoni and Stevanovic, 2016).

surveys assessing the sentiment of businesses, consumers, and financial institutions. However, the ability of these data to help forecast Canadian economic developments has been the subject of only limited analysis. This analysis includes Pichette (2012), who studies how the Bank of Canada's *Business Outlook Survey* (BOS) data is correlated with future values of output, investment and consumption growth. The author reports that the BOS data helps predict future output and investment growth, but that results for consumption are weaker. More recent work on the same lines includes Pichette and Robitaille (2016), which again shows that the BOS data has important explanatory power in real-time forecasting exercises where the data vintages used reproduce the information known at the time, before the data revisions, which can be sizeable for GDP. The analysis in Pichette (2012) and Pichette and Robitaille (2016) is limited to the BOS survey however, and pertains to the growth rate of variables. As indicated above, we provide a generalization of that analysis by examining all the information contained in the four available surveys; in addition, our focus is on the prediction of a binary variable indicating whether the economy is experiencing a slowdown or an expansion. Other work includes Binette and Chang (2013), who analyse the performance of *Canada's Short-Term Indicator Model* (*CSI*) a forecasting framework for future Canadian GDP growth; the model includes some of the BOS Survey data among its explanatory variables. As indicated above, our paper provides a novel contribution to the literature, by showing how using all Canadian data on sentiment and organising it within a factor model provides a very promising avenue for predicting future Canadian economic developments.

Structural Impact of Confidence Shocks

A second major thrust of the research using survey data concerns its structural interpretation. This literature originated in the work of Matsusaka and Sbordone (1995) and is represented by recent contributions from Leduc and Sill (2013), Barsky and Sims (2012) and Lambertini et al. (2013). Using VAR frameworks that include sentiment variables, these authors first identify shocks to the sentiment variable (thus assigning a structural interpretation to its innovations) and then estimate the macroeconomic impact of those shocks.

In that context, Leduc and Sill (2013) report that declines in expected future unemployment rates (as measured by answers to the relevant question in the Livingstone and *SPF* surveys) have a positive and contemporaneous impact on economic activity, reducing current unemployment and increasing current inflation. In addition, they show that these declines trigger a tightening process of monetary policy, a result coherent with a worldview whereby monetary policy exerts a *stabilizing* influence on economic fluctuations, gradually tightening interest rates when the economy is affected by positive waves of optimism about future economic conditions. Lambertini et al. (2013) extend the findings in Leduc and Sill (2013) and document that positive confidence shocks also have positive impacts on real estate and housing prices. In a related contribution, Barsky and Sims (2012) show that shocks to the forward-looking questions in the Michigan Consumer Survey are associated with gradual and long-lasting rises in consumption and output in the US. They also argue that such effects are compatible with the view that these shocks represent signals about future fundamental innovations to technology.

The relative abundance of Canadian data on sentiment suggests that exercises where a structural identification is assigned to shocks in Canadian confidence data and the macroeconomic impact of such shocks is estimated would represent a fruitful avenue for future research. In addition, the extent to which US confidence data affect their Canadian counterparts, and how they then jointly affect overall Canadian economic fluctuations, are important open questions.

3 Model

We adopt the framework popularized by Estrella and Mishkin (1998) and used by much of the literature on forecasting recessions or economic slowdowns.[7] Denote by y_{t+h} the

[7]See for example, Taylor and McNabb (2007), Kauppi and Saikkonen (2008), Nyberg (2010) and Christiansen et al. (2014). These authors motivate their interest in forecasting a binary variable –whether the

binary variable indicating whether economic activity at time $t+h$ experiences a slowdown ($y_{t+h}=1$) or an expansion ($y_{t+h}=0$). Our aim is to forecast $P\left(y_{t+h}=1\right)$ on the basis of information available at time t.

To this end, consider the following Probit model:

$$y^{*}_{t+h}=\beta'\mathbf{X_t}-\epsilon_t,\ \ \epsilon_t\sim N(0,\sigma); \tag{1}$$

$$y_{t+h}=1(y^{*}_{t+h}\geq 0); \tag{2}$$

where the unobserved variable y^{*}_{t+h} is a function of the vector of explanatory variables $\mathbf{X_t}$, y_{t+h} is the indicator variable signalling the state of the business cycle at time $t+h$ and $P\left(y_{t+h}=1\right)=P\left(\epsilon_t\leq\beta'\mathbf{X_t}\right)$ is the model's probability of a slowdown at time $t+h$. The time-series of realized values for the indicator variable y_{t+h}, together with the vector of explanatory variables $\mathbf{X_t}$, can then be used to maximize the sample's likelihood

$$LogL=\sum_{t}\left[y_{t+h}log\Phi(\beta'\mathbf{X_t})+(1-y_{t+h})log(1-\Phi(\beta'\mathbf{X_t}))\right], \tag{3}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

Following Christiansen et al. (2014), we posit that three general types of variables are included in the explanatory bloc $\mathbf{X_t}$. First, we include specific, individual variables, such as the term spread or stock market returns, which have been identified elsewhere as valuable predictors for economic downturns (Estrella and Mishkin, 1998); below these variables are represented by the vector $\mathbf{f_t}$.

Second, sentiment variables are included and denoted by the vector $\mathbf{s_t}$. The variables included in $\mathbf{s_t}$ might be individual time-series (such as the response to one specific question in one survey), popular simple aggregates of these sentiment data (such as the *Index of Consumer Confidence* produced by the Conference Board from the answers to its consumer survey), or estimated factors extracted after merging all available sentiment data in a

economy is experiencing a slowdown or an expansion– in two ways. First, there exists legitimate interest from policy makers or market participants for this question. Second, a large empirical literature has documented the fact that time-series processes with regimes fit the evolution of economic activity well; in that context, interest for an underlying binary variable indicating the status of the business cycle is natural.

large dataset. Finally, general macroeconomic, financial and national accounts data are included and represented by the vector $\mathbf{Z_t}$. As is the case for sentiment variables, this block of explanatory variables is included through the use of factors extracted from a large dataset.

Specifying our probit model with these three blocks of explanatory variables leads equations (1)-(2) to be rewritten into the following:

$$y^*_{t+h} = \alpha' \mathbf{f_t} + \beta' \mathbf{s_t} + \gamma' \mathbf{Z_t} - \epsilon_t, \;\; \epsilon_t \sim N(0, \sigma), \tag{4}$$

$$y_{t+h} = 1(y^*_{t+h} \geq 0). \tag{5}$$

In addition, we follow Kauppi and Saikkonen (2008) and analyze an alternative specification that can include lagged value of the business cycle indicator $y_{t-s}, s \in (1, ...)$ is added to the explanatory block of the model. In that case, (4) becomes

$$y^*_{t+h} = \alpha' \mathbf{f_t} + \beta' \mathbf{s_t} + \gamma' \mathbf{Z_t} + \delta' y_{t-s} - \epsilon_t, \;\; \epsilon_t \sim N(0, \sigma). \tag{6}$$

4 Data

4.1 Canadian Business Cycles

A well-known chronology of US business cycles is constructed and maintained by the NBER's Business Cycle Dating Committee. This chronology identifies peaks and troughs in economic activity, defining a recession as the period from a peak to a trough. The resulting business cycles *dates* have served as the basis of an extensive empirical literature.

The C.D. Howe Institute performs a similar exercise for the Canadian economy and the Institute's *Business Cycle Council* has produced a list of all Canadian recessions since 1926 (Cross and Bergevin, 2012).[8] Importantly, these data imply that recessions have become increasingly rare events over time in Canada. According to the Institute's chronology, the

[8] The C.D. Howe Institute is an independent not-for-profit research institute whose objective is fostering economically sound public policies for Canada. See https://www.cdhowe.org/council/business-cycle-council for details about the Institute'S Business Cycle Council.

Canadian economy has experienced only two recessions since the early 1980s, with the last one occurring during the $2008-2009$ Great recession.[9] One practical implication from this feature of the C.D. Howe data is that empirical work using explanatory variables for which limited historical data are available will likely include only one recession (in 2008-2009), severely reducing the potential power of any econometric method.

Both the NBER and the C.D. Howe Institute interpret recessions as declines in the *level* of general economic activity. However, other conceptual frameworks view recessions as periods where economic activity, even if it is growing, is doing so at a rate below its long-term potential. Notably, the OECD uses a *growth cycle* framework to compute troughs and peaks, ie. turning points in these growth cycles, for all member countries.[10] Figure 1 below illustrates the implication of these two differing views of what constitutes an economic downturn. In the figure, the dark-shaded episodes are the C.D. Howe recessions dates for Canada while the light-shaded periods are the growth slowdowns for Canada as identified by the OECD growth-cycle methodology. Although both chronologies overlap to a considerable extent, notice that the OECD has identified several low-growth episodes since the 1991 recession, while the C.D. Howe methodology has only identified one recession, in $2008-2009$.[11]

[9] This is contrast to the US economy, which was affected by an additional recession in 2001, according to the NBER dates.

[10] See Zarnowitz and Ozyldirim (2006) for a discussion of growth cycles and a description of estimated growth cycles for the US. In addition, see Anderson and Vahid (2001) for an analysis where an alternative metric is used to defined recessions.

[11] Table 8 in Appendix B provides a detailed list of all peaks and troughs in the Canadian business cycle identified by the OECD methodology.

Figure 1: Canadian Recessions: OECD and C.D. Howe

Because both Bank of Canada sentiment surveys were only begun in the late 1990s, using data drawn from these surveys can only apply to the post-2000 period and as such, only one recession episode according to the C.D. Howe dates. We therefore choose to apply our methodology to the OECD *growth cycles* dates instead. This is in line with other work in the literature forecasting economic downturns. Notably, Taylor and McNabb (2007) apply their methodology to three alternative definition of recession episodes in each of the three countries analyzed (the U.K., France, Italy and the Netherlands); one of these definitions is coherent with the *growth cycle* view underpinning the OECD data. In view of the favourable results reported below about the forecasting ability of sentiment data, fruitful avenues for future research might include using expanded data applied to a narrower definition of recession, as the one embodied by the C.D. Howe chronology.

4.2 Explanatory Variables

We assess the forecasting power of three types of explanatory variables: (i) classical predictors, (ii) sentiment variables and (iii) general macroeconomic and financial variables.

We describe each block of explanatory variables in turn.

Classical Predictors

In their influential analysis, Estrella and Mishkin (1998) single out some specific variables likely to contain valuable forecasting power for future US recessions. Notably, they argue that the forward-looking characteristic of the term spread and of stock market returns may give them the ability to signal future economic developments. Estrella and Mishkin (1998) also test the signalling ability of monetary aggregates, housing permits and CPI inflation. Their results do confirm that the term spread and stock market returns, taken individually or when combined, are valuable indicators of future recessions, both in-sample and in out-of-sample experiments. Many contributions to the literature on forecasting recessions have since used these variables as benchmarks to assess their methods or choice of new variables. Christiansen et al. (2014), for example, analyze the ability of sentiment variables to forecast US recessions by benchmarking to such classical predictors.[12]

We follow this strategy and start our analysis by using the following Canadian equivalents to these classical predictors: the term spread (measured as the difference between the 10-year Canadian government bond yield and that of a 3-month Treasury Bill), and the return on the benchmark SP/TSX stock market index.[13] We also assess the forecasting ability of a monetary aggregate (the *M1+* definition), a short-term interest rate (the Bank of Canada's overnight rate) and the nominal CAN/US exchange rate.

Sentiment Variables

Next, we use the four Canadian surveys on sentiment described above. To this end, data from both Conference Board surveys (of households and business executives, respec-

[12]The importance of the term spread as a predictor for future business cycles is further documented in Duarte et al. (2005), Wright (2006), Rudebusch and Williams (2009) and Kotchoni and Stevanovic (2016).

[13]We approximate the return on the SP/TSX index by the log-difference in the index's level.

tively) are used, as well as data from both surveys from the Bank of Canada: the Business Outlook Survey and the Senior Loan Officers survey.

Since each survey contains several questions, a fairly large number of potentially useful sentiment variables are available for the analysis. One important goal of this paper is to identify the best manner in which the information contained in these data can be used to forecast economic slowdowns. To this end, our analysis first assesses the forecasting ability of all individual time-series available in the four surveys. Next, we study aggregates of survey answers, such as the *Index of Consumer Confidence*, a simple average of the ratios of positive to negative responses for the four questions in the Conference Board's Consumer Survey. The Conference Board also publishes the *Index of Business Confidence*, again a simple aggregate of the answers to its survey of business executives.[14]

Note that the simple-sum averages underlying these indices represent a simple, but very specific way to aggregate information in that survey; this begs the question of how to best identify the relevant information contained in all available sentiment data. One popular method of aggregating information from a large dataset of individual variables is to employ a factor model, in which all available variables are assumed to be affected by a given set of common components (or factors) and by idiosyncratic components. In the case of our sentiment data, that would imply the following:

$$s_{it} = \mathbf{\Lambda_i}'\mathbf{S_t} + e_{it}, \tag{7}$$

where s_{it}, $i = 1, ns$ represent the individual sentiment variables present in the dataset, $\mathbf{S_t}$ and $\mathbf{\Lambda_i}$ denote the $p \cdot 1$ ($p \leq ns$) vectors of common factors and 'loading' of these factors on each individual variables, respectively, while e_{it} represents the idiosyncratic component for each variable. The use of factor models to synthesize information contained in large dataset and help forecasting was popularized by contributions in Stock and Watson (2002a,b) and is now a standard part of the forecaster's toolkit.[15] Note that there are potentially as

[14]The Bank of Canada does not publish aggregates of answers to its Business Outlook Survey and its Senior Loan Officers Survey, but we construct our own such indices.

[15]Important contributions in this literature include Forni et al. (2005), Boivin and Ng (2006) and Bai

many factors p as variables ns in a given dataset: we estimate the factors by computing the principal components' decomposition of the covariance matrix of all sentiment data and examine the predictive ability of all these components.

Macroeconomic Variables

To assess the forecasting ability of general macroeconomic variables, we make use of a panel of 144 Canadian macroeconomic and financial series. This dataset is comprised of publicly available time series relevant for the Canadian economy, such as interest rates, commodity prices, exchange rates and NIPA Components (investment, private or government consumption, etc.). When necessary, data are transformed at a quarterly frequency by taking the quarterly average of monthly or daily values (for some financial variables). In addition, all data are standardized into zero mean, unit-variance indicators, as is standard in the factor model literature. Table 9 in Appendix C lists all time series contained in the database and how each variable from those raw data was transformed.

Including all 144 variables in the probit model in (4) and (6) is not feasible. Instead, we once again extract relevant information from these variables by employing a factor model similar to the one used above for the confidence variables, so that we have

$$z_{it} = \mathbf{\Gamma_i}'\mathbf{Z_t} + e_{it}, \tag{8}$$

where Z_{it} represent the macroeconomic variables present in the dataset, $\mathbf{Z_t}$ and $\mathbf{\Gamma_i}$ denote the $p{\cdot}1$ vectors of common factors and 'loading' of these factors on each individual variables, respectively, while e_{it} represents the idiosyncratic component. Again, we estimate the factors via the principal components' decomposition of the matrix for all z_{it} variables.

and Ng (2008). Stock and Watson (2006) review the literature on forecasting with factor models and describe the various methods to estimate underlying factors from a given dataset.

5 Results

We provide three sets of results to analyze the ability of our probit models to forecast future economic slowdowns. First, simple models with one single predictor are assessed. Next, multiple-predictor models are analyzed and finally a robustness analysis, which includes an out-of-sample experiment, is presented.

5.1 Single-predictor models

Table 1 presents the results from estimating the probit model (4) at the one-quarter-ahead horizon ($h = 1$) with only one explanatory variable at the time, using the sample 2002Q1 to 2014Q4.[16] That variable is either one of the classical predictors described above (Panel A of the table) or one of the confidence indices (Panel B). For each variable considered, the table reports results arrived at using the contemporaneous value of the variable or one of its first two lags. To compare their effectiveness as predictors, the table reports each variable's estimated coefficient, its p-value, the Estrella (1998) pseudo-R^2 measure (R^2_{es} thereafter) and the optimized log-likelihood.

The table shows that the stock market variable performs best among the classical predictors. Notably, its contemporaneous value has high significance and leads to the highest R^2_{es} of the table's top panel. The monetary aggregate and exchange rate variables also exhibit some significance but it is weaker. In addition, the best predictor of y_{t+1} is, in all cases, the time-t dated value of the variable and significance substantially declines when the $t-1$ and $t-2$ values are used. Finally, the term spread is found to have little explanatory power, in contrast with results in Estrella and Mishkin (1998) and Christiansen et al. (2014). Turning to confidence indices, the bottom panel of Table 1 reports that the Conference Board's *Consumer Confidence Index* holds much promise as a predictor of future economic downturns: its R^2_{es} is strongest when the contemporaneous value of the index is used but predictive power remains significant when lagged values are

[16] Our choice of sample is dictated by the earliest dates at which all data from the two Bank of Canada surveys become fully available.

employed instead. Other sentiment indices fare less strongly.

Aggregates like the Conference Board's *Consumer Confidence Index* represent one specific way to summarize the information contained in sentiment data and as such already embody hypotheses about how sentiment data should be transformed and used. One might instead be interested in assessing the predictive ability of the raw data underlying these indices. Such an assessment is conducted in Table 2, which studies the predictive ability of the raw individual times-series from the different surveys' answers. As was the case in Table 1, Table 2 indicates if contemporaneous or lagged values of the variable are used and reports the estimated coefficient and its significance, as well as the R^2_{es} and the optimized log-likelihood.[17] The table shows that several variables have important forecasting power, notably the variable $v1p$, which relates to the question about past sales growth in the Bank of Canada's BOS survey.[18] This variable appears in all three panels of Table 2, indicating that its contemporaneous as well as its two lagged values help predict future economic slowdowns in a statistically significant manner. Variables from the Conference Board's Consumer Survey (prefixes nq) also appear in the various panels of Table 2 and some exhibit high R^2_{es} values. However, the key takeaway from Table 2 is the order of magnitude of the best reported figures for the performance measure R^2_{es}; overall, these best figures are roughly comparable to the best comparable ones reported in Table 1. This suggests that within single-predictor models, the best 'classical' variables, the best confidence indices and the best individual sentiment variables have comparable performance.

We now examine the forecasting ability of the estimated factors. Recall that according to (7) and (8), the evolution of all confidence variables and all macroeconomic variables in our dataset can be decomposed into the influence of common factors and idiosyncratic shocks. For each group of variable, we estimate the factors via the principal component

[17] Variables with n and nq prefixes refer to the Conference Board's business and consumer surveys, respectively, while the v and w prefixes refer to the Bank of Canada's Business Outlook and Loan Officer surveys, respectively. Promising individual variables are identified by conducting a pre-experiment forecasting exercise across all variables and retaining those with high R^2_{es} for further analysis.

[18] Appendix 3 provides a detailed description of all questions in the BOS Survey.

decomposition of the covariance matrix of all variables in the group, which delivers a series of components ordered in decreasing importance for explaining that matrix.[19]

Researchers using factor analysis and principal component decompositions often focus on the first few components, arguing that they explain the majority of the dataset's variability. However, one principal component could explain a large fraction of a dataset's overall variability but still forecast poorly the variable of interest. In our context, this implies that the components most useful for forecasting economic slowdowns might explain only a small fraction of the overall sentiment data's covariance matrix. Our analysis therefore proceeds by keeping the full set of principal components and studying their forecasting ability one at the time.[20]

Table 3 reports our results and is divided in three panels: panel A first analyses the factors $\mathbf{S_t}$ recovered from the sentiment data (see equation 7), panel B reports those associated with the macroeconomic factors $\mathbf{Z_t}$ (equation 8) and panel C depicts results obtained when all macroeconomic and confidence variables are combined into one larger database from which a new set of factors, $\mathbf{W_t}$, are extracted. In each panel of the table, the forecasting ability of 10 factors is reported; these 10 factors have been chosen by keeping the best (as measured by the model's R^2_{es}) among those whose estimated coefficient was significant statistically in a pre-experiment analysis. These factors are labeled by their order in the principal component analysis and by whether the contemporaneous or lagged values are used: for example, $\mathbf{S_{37t}}$ in Panel A refers to the performance of the (contemporaneous value of the) 37th principal component extracted from the sentiment data.

The three panels of the table report significant ability in forecasting the future business cycle status. This is indicated by the high significance of the coefficients and the high R^2_{es} values: for example, Panel A of the table indicates that the confidence factors have R^2_{es}

[19] See Stock and Watson (2006) for a discussion about how factor models can be consistently estimated using principal component decompositions and other methods.

[20] Bai and Ng (2008) provide a systematic analysis of the tradeoff between explaining a significant proportion of a dataset's variability and the ability to forecast future values of the variable of interest.

values ranging from 0.175 to as high as 0.33, a performance slightly superior to the best one attained by an individual variable (the stock market returns) or a confidence index (the Conference Board's *Index of Business Confidence*) in Table 1. Next, Panel B of the table shows that some of the factors retrieved from the macroeconomic time series attain an even better performance, with R^2_{es} values rising as high as 0.40 (Z_{44t}). Panel C of the table, reflecting results obtained after merging the two dataset before extracting the factors, shows that this leads to a slight *deterioration* of the model's forecasting ability, a result consistent with the arguments advanced in Boivin and Ng (2006) whereby adding more variables to a dataset before extracting factors does not necessarily always lead to improvements in forecasting ability.

5.2 Multiple-predictor models

Results in Table 3 might appear to suggest that aggregating information using the factor models (7) or (8) provides little or no improvement in forecasting ability, when compared to results obtained by using individual variables, as in Table 1 and Table 2. Indeed, the best R^2_{es} were found to be roughly of the same order of magnitude in all tables analyzed so far: factor-based prediction models may sometimes exhibit slightly higher R^2_{es} in Table 3, but this small improvement might not be enough to arbitrage the challenges these models pose, notably with respect to communication of results. Taken together, therefore, results depicted in Table 1, Table 2 and Table 3 might be interpreted as suggesting that sentiment data, although incorporating valuable predictive power, do not appear to deliver a clear increase in such predictive power relative to classical predictors.

However, multiple-predictor probit models have the potential to modify this assessment. Because our estimated factors are orthogonal to each other, the predictive ability of each new added factor should always increase the overall performance of a given model; by contrast, no such assurance is present when specific individual variables are added to a model since any additional variable would not be orthogonal to those already in use.

To test this conjecture, Table 4 reports estimation results using several predictors at the same time. The table is divided in two panels: the standard model is assessed on the left-hand side while a model that includes the lagged value of the dependent variable, as in (6), is analyzed on the right-hand side. For each estimation, between four and six variables from each block considered so far are used: the blocks are the classical predictors (heading "*M1*" in the table), confidence indices ("*M2*"), factors drawn from our macroeconomic dataset ("*M3*"), from the sentiment data ("*M4*") and from the amalgamated database ("*M5*"). For each column, only estimated coefficients significant at the 10% level are reported: for example, the column M1 on the left-hand side panel of the table only reports the estimated coefficient for the stock market return variable since it is the only one statistically significant.[21]

As expected, the first two columns of the table's left-hand side (headings *M1* and *M2*) depict results very similar to those from Table 1: in each column, only one individual variable is statistically significant (the return to stock markets and the Conference Board's *Consumer Confidence Index*) and the R^2_{es} are very similar to those in Table 1.[22] These columns thus confirm that the addition of specific, additional variables to the probit does not significantly improve performance.

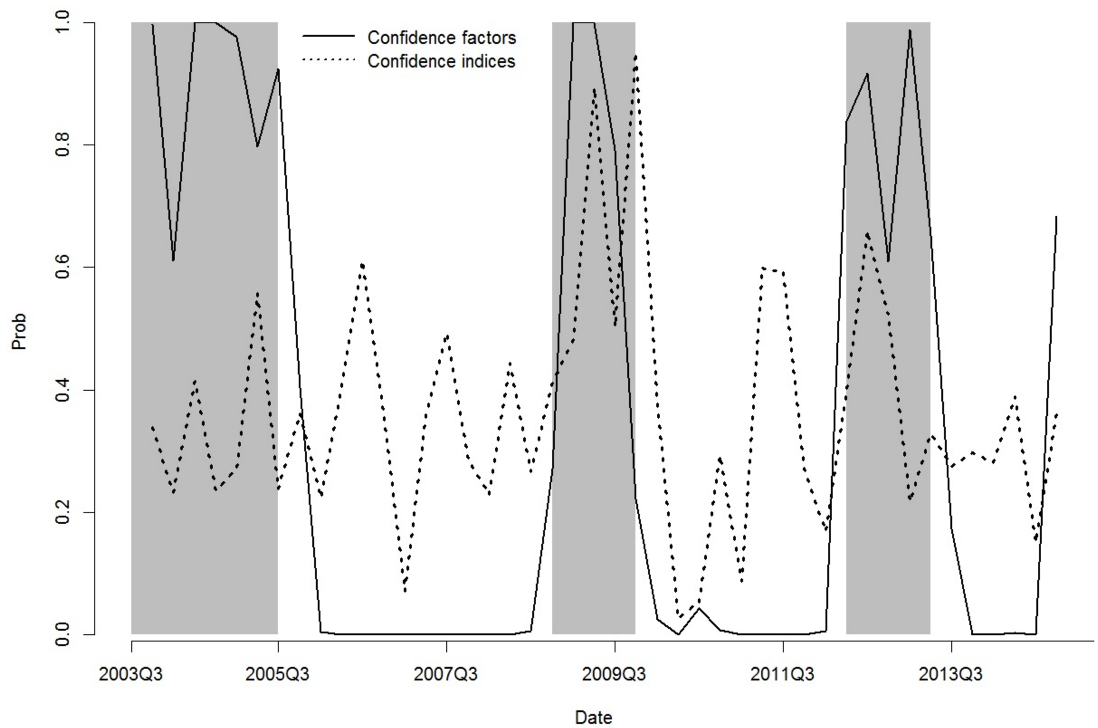
The next three columns (headings *M3*, *M4* and *M5*) paint a different picture. Here, the different factors included in the multiple-predictor probit appear to all add to the model's performance. Indeed, all factors are statistically significant and model performance is improved substantially: the R^2_{es} reported at the bottom of the three columns are significantly higher than those in Table 3 and attain a high mark of 0.956 (for the model using factors extracted from confidence variables, *M4*). The AIC and QPS scores concur with this result and also show big declines relative to the first two columns of the table.

[21] We identify the best model in each of columns *M3* to *M5* as follows: all factors significant at the 10% level in a single-predictor model are first identified (this usually selects around 10 potential predictors). These factors are then included in a multiple-predictor probit in all possible permutations to select the best model.

[22] Results are not exactly the same because the presence of other variables influences the estimation process even if not all coefficient estimates are depicted.

Recall that one important goal of the present paper is to identify the best way by which sentiment data can help forecast business cycle turning points. In that regard, the table shows that using all available data on sentiment, through the factor model (7), significantly increases performance relative to using summary variables like the Conference Board's *Consumer Confidence Index*. This is apparent when comparing columns *M2* and *M4*: the R^2_{es} measure is significantly higher in the latter column, and both AIC and QPS scores are much lower. In synthesis, the left-hand side panel of the table has two key findings: (i) using several orthogonal factors significantly improves the predicting ability of our probit framework and in particular (ii) including sentiment data through the factor model (7) provides substantial improvements relative to more straightforward summary measures of sentiment. This last finding is illustrated graphically in Figure 2 below, which shows the estimated probabilities for the best model arising from confidence indices (dashed lines) and from confidence factors (full lines): the predictions based on confidence factors have visibly better forward-indicator capabilities and have much more contrasted probability regions.

Figure 2: Estimated Probability of Slowdown: Confidence Indices and Confidence Factors



The right-hand side of Table 4 next analyzes a model similar to (6), where the lagged values of the dependent variable y_{t-1} is added to the estimation. As expected, the lagged status of the business cycle is very informative and its estimated coefficients are very-highly significant and positive –which implies positive auto-correlation in the business cycle– in all columns. Interestingly, the presence of y_{t-1} also significantly modifies which variable performs well and how much they do so. For example, the stock market return variable is now absent from the table (column *M1*), suggesting that this variable does not help predict future business cycle status once the current status is included; instead the money growth rate appears significantly. Similarly, the Conference Board's *Consumer Confidence Index* is now also absent (column *M2*) and is replaced by the Bank of Canada's Senior Loan Officers (SLO) survey data. This suggests that once y_{t-1} is included, a different type of information is emphasized by the estimation. In addition, the table shows that including y_{t-1} in the probit significantly increases the R^2_{es} performance of models *M1* and *M2*, relative to what was reported in the left-hand side panel of the Table.

The columns reporting results using factors (*M3* to *M5*) deliver a slightly different message. On the one hand, the estimated coefficient on y_{t-1} does remain highly significant and the number of other predictors statistically significant is lower that it was before. On the other hand, the improvements in overall model performance are now much more modest. Including y_{t-1} does increase slightly the R^2_{es} for models *M3* and *M5*, relative the the left-hand side of the table, but column *M4* reports a decrease overall performance, from $R^2_{es} = 0.956$ in the left-hand side of the table to 0.889 on the right-hand side. This suggests that using all available sentiment data in conjunction with the factor model (7) creates a powerful predictor for future business cycle turning points, which does not need lagged values of the business cycle status to perform well.

In short, Table 4 illustrates that the in-sample forecasting ability of our probit model is considerably improved when a framework with multiple predictors is used, and that these improvements are substantial when factors extracted from our bigger datasets are included as predictors. This is particularly so when such factors are drawn from sentiment data. In addition, the lagged state of the business cycle does have an influence on these results but while overwhelmingly positive for the *M1* and *M2* models, this influence is more modest and not consistently positive for models *M3*-*M5* where factors extracted from data are the relevant predictors.

5.3 Robustness

We now assess the robustness of our results. We first evaluate our framework's performance over longer forecasting horizons. Next, we reestimate our model over a different sample and finally, we conduct an out-of-sample experiment.

Longer Forecasting Horizons

Until now, our analysis has concentrated on the task of predicting the occurrence of economic slowdowns one-quarter ahead, ie the case $h = 1$ in our basic model (1). Table 5 next assesses the ability of our framework to forecast at longer horizons.

Accordingly, the table has four panels, one for each value of h. In each panel, several performance measures are reported for the best model in each block of predictor variables.[23] These measures, in addition to the usual AIC and QPS criteria, include the CML (which measures the aggregate of a model's mistakes) as well as the proportion of correctly predicted downturns ($shots^+$) and the proportion of missed expansions ($shots^-$).[24]

The results depicted in Table 5 largely accord with those discussed above. First, the table shows that using factors, extracted from either sentiment or macroeconomic variables, provides substantial improvement to the forecasting ability of the probit. To see this, compare the AIC or QPS metrics for the first two lines (classical predictors and confidences indices) with those in the last three. For each horizon $h = 1, 2, 3, 4$, these figures are, for the most part, noticeably smaller when factors are the main predictors. Further, the proportion of correctly predicted downturns ($shots^+$) is also (almost) uniformly better when such factors used. In addition, the table strongly suggests that the full potential of sentiment data for forecasting future business cycles is likely to be attained when all such data is used, and is amalgamated through a factor model like (7). This is evident when comparing the *Confidence indices* and *Confidences factors* lines in the table: in most cases and for most performance measures, the latter model performs best.

Note that one interesting caveat to this assessment occurs for the *CML* measure, especially at longer horizons: in such cases, predicting with confidence factors appears to

[23]The best model for each block of variables is identified using the procedure described above in the context of the results from Table 3.

[24]The CML is discussed in Buja et al. (2005). It aggregates the model's mistakes by summing the false positives (predicting an economic downturn that does not occur) and the false negatives (failing to predict a recession that actually occurs). It is computed as $CML = \frac{1}{T}\sum_{t=1}^{T}\left[(1-q)y_t(1-\mathbb{1}_{(\widehat{p}_t \geq 0.5)}) + q(1-y_t)(\mathbb{1}_{(\widehat{p}_t \geq 0.5)})\right]$, where $\mathbb{1}_{()}$ is 1 if its argument is true and 0 otherwise and q is the relative cost of the two different types of mistakes. We use $q = 1/3$, which penalizes false positives more heavily.

result in inferior performance. For example the *CML* at $h = 4$ for *Confidence indices* is 0.360 while it is higher (0.588) for *Confidence factors*. At the same time, the measure $shots^+$ (the proportion of correctly called recession episodes) clearly favors the factor-based model. This indicates that the confidence factors-based model results in a relatively high number of false positives (predicting an economic downturn that does not occur). In that sense, confidence factor-based prediction models might appear to be "nervous" indicators, not missing many recessions but announcing some that do not occur.[25]

Finally, as was the case above, the right-hand side of Table 5 shows that the addition of y_{t-1}, the lagged business-cycle status, modifies somewhat our general assessment without overturning its qualitative nature. Indeed, forecasting with factors, notably confidence factors, still increases the performance of the probit, but to a lesser extent to what was the case in the left-hand side of the table.

Different Estimation Sample

We now assess the robustness of our results to the choice of the estimation sample. To this end, Table 6 reports on a similar exercise than the one underlying Table 5, but where the different models are estimated using data stopping at 2010 Q1, before the last economic slowdown identified by the OECD methodology. Overall, the message identified by Table 5 is unchanged: confidence factors predict generally better than confidence indices, although as the forecasting horizon increases, this pattern may cease to be robust. As was the case above, instances where confidence factors have lower *CML* measures but higher $shots^+$ are present, indicating that factors may be "nervous" predictors.

An out-of-sample experiment

Finally, Table 7 presents our (recursively computed) out-of-sample experiment. To construct this table the entire sample (data from 2002Q3 to 2014Q4) is first used to

[25] Importantly, the out-of-sample experiment described below suggests that such inferior performance, as measured by the CML metric, disappears in out-of-sample experiments.

extract factors (the factors thus also cover the period 2002Q3 to 2014Q4). However, the best model for each block of variable is now identified recursively. Specifically, factor data from 2002Q3 to 2010Q1 is first used, all permutations are tried, and the best model from that period is identified and used to compute a forecast for 2010Q2. Next, the factor data from 2002Q3 to 2010Q2 is used to identify the best model for that period and to forecast a value for 2010Q3, and so on. This resembles the real process by which a central bank or a national statistical agency would use to forecast in real time.[26]

Results reported in Table 7 are very favorable to factor-based probit models, particularly those drawn from sentiment data. All performance measures now point to the *Confidence factors* line as possessing superior information for the forecasting of future business cycle turning points: the QPS and CML measures are now considerably lower than when simple confidence indices are used, whereas the $shots^+$ and $shots^-$ measures are now higher and lower, respectively. Even when the lagged business cycle status is used (right-hand side of the table) confidence factors now out-predict single specific variables. This most likely occurs because the factor-based strategy allows the flexibility to change the factors used in the exercise, at each point in time, every time the model is reestimated; by contrast, using specific variable only allows to compute new estimations of the coefficient assigned to a specific variable. Overall the evidence in Table 7 reinforces substantially the one discussed above, wherein all sentiment data ought to be used, and be amalgamated through factor models to achieve their full forecasting potential.

6 Conclusion

A rapidly expanding literature has documented that confidence –or sentiment– data can increase the performance of forecasting frameworks to signal the future occurrence of economic slowdowns. The present paper adds to this evolving body of knowledge by showing that Canadian data on sentiment can contribute substantially to the task of forecasting the Canadian business cycle, particularly when all available such sentiment

[26] Ideally, the factors themselves would be reestimated recursively, at each stage of the experiment.

data is used and amalgamated through factor models.

Specifically, we report that Canadian data on sentiment, as gathered from the answers to four different available surveys, significantly help forecast future slowdowns in the Canadian economy. Further we show that using all available such data, by amalgamating all time-series and extracting factors from the amalgamated dataset, provides substantially improved forecasts relative to those obtained using individual series or simple averages of sentiment data.

Possible avenues for fruitful future research include evaluating the relative performance of US and Canadian sentiment data for the Canadian business cycle, assessing the predictive ability of our framework over longer historical samples (even if doing so would result in having less variety in the available sentiment data) and implementing a structural identification exercise to study the causal impacts of confidence data.

Table 1: Single-predictor Probit: Classical predictors and Confidence indices

Variable		Estimate	Std Error	p-value	R^2_{es}	$ln\hat{L}$
		Panel A: Classical Predictors				
InterestRate						
	$lag0$	-0.70	1.04	0.50	0.020	-30.50
	$lag1$	-0.56	1.03	0.68	0.010	-29.67
	$lag2$	-1.13	1.11	0.31	0.050	-28.89
TermSpread						
	$lag0$	0.16	0.45	0.72	0.006	-31.00
	$lag1$	-0.04	0.46	0.93	0.001	-29.87
	$lag2$	0.04	0.48	0.93	0.001	-28.87
Stock Market						
	$lag0$	-7.92	3.46	0.02*	0.280	-28.76
	$lag1$	-2.30	2.47	0.35	0.040	-28.00
	$lag2$	-0.60	2.44	0.80	0.002	-28.69
Exchange Rate						
	$lag0$	9.39	5.49	0.09*	0.130	-30.53
	$lag1$	0.88	4.89	0.85	0.001	-29.57
	$lag2$	-4.24	5.29	0.42	0.028	-30.53
Money						
	$lag0$	14.34	9.00	0.055*	0.040	-29.31
	$lag1$	11.36	17.29	0.51	0.015	-29.22
	$lag2$	12.94	17.46	0.46	0.025	-28.70
		Panel B: Confidence Indices				
Business Conf. Index (Conference Board)						
	$lag0$	-3.48	2.45	0.16	0.090	-30.10
	$lag1$	-2.59	2.42	0.28	0.050	-28.90
	$lag2$	-2.21	2.41	0.36	0.036	-28.30
Consumer Conf. Index (Conference Board)						
	$lag0$	-6.47	2.63	0.01*	0.290	-29.13
	$lag1$	-4.15	2.22	0.06*	0.163	-26.51
	$lag2$	-2.55	2.08	0.22	0.054	-27.30
BOS (Bank of Canada)						
	$lag0$	-0.07	0.05	0.16	0.080	-29.21
	$lag1$	-0.04	0.05	0.37	0.034	-29.03
	$lag2$	-0.02	0.05	0.70	0.06	-28.65
Senior Loan Officers (Bank of Canada)						
	$lag0$	-0.004	0.01	0.73	0.008	-31.16
	$lag1$	-0.009	0.01	0.34	0.038	-30.71
	$lag2$	-0.002	0.01	0.82	0.002	-29.13

Notes: "*" indicates statistical significant at the 10% level. Reports estimate of probit model $y^*_{t+1} = \bar{\alpha} + \alpha f_{t-d} + \epsilon_t$, where f_t is either a classical predictor or a confidence index and $d = 0, 12$. $R^2_{es} \equiv 1 - (ln\hat{L}/lnL_0)^{-(2/T)lnL_0}$ is Estrella's (1998) pseudo-R^2, where $ln\hat{L}$ is the estimated likelihood and lnL_0 is the likelihood only with a constant term. Finally, T is the sample size and the sample runs from 2002Q1 to 2014Q4.

Table 2: Single-predictor Probit: Individual Sentiment Variables

Variable	Estimate	Std Error	p-value	R^2_{es}	$ln\hat{L}$
Variables entering contemporaneously					
$nq02c_q$	-0.04	0.024	0.088	0.114	-29.93
$nq02d_q$	0.05	0.024	0.035	0.191	-28.24
$nq02h_q$	-0.65	0.380	0.090	0.132	-29.65
$nq07d_q$	0.05	0.019	0.005	0.323	-25.46
$nq08a_q$	-0.14	0.085	0.094	0.115	-29.06
$nq10a_q$	0.04	0.022	0.095	0.113	-29.74
$nq10d_q$	0.13	0.060	0.036	0.189	-27.64
$nq10h_q$	-0.15	0.083	0.073	0.139	-30.34
$n1w_m$	0.19	0.088	0.035	0.180	-29.34
$n2w_m$	0.24	0.097	0.015	0.251	-29.91
$v1_p$	0.07	0.031	0.019	0.225	-28.43
$w1_n$	0.02	0.011	0.077	0.135	-29.47
Variables entering with one lag					
$nq02c_q$	-0.04	0.024	0.096	0.109	-28.68
$nq02d_q$	0.06	0.025	0.018	0.243	-27.36
$nq02h_q$	-0.86	0.444	0.052	0.191	-28.61
$nq04b_q$	0.04	0.021	0.041	0.181	-29.87
$nq05a_q$	-0.09	0.038	0.019	0.225	-29.35
$nq6d_q$	0.08	0.039	0.050	0.164	-29.94
$nq7d_q$	0.04	0.018	0.020	0.218	-25.71
$nq7f_q$	-0.13	0.060	0.030	0.213	-31.53
$nq9c_q$	-0.06	0.032	0.052	0.152	-29.63
$n1w_m$	0.19	0.091	0.032	0.186	-27.76
$n2w_m$	0.25	0.102	0.015	0.252	-26.88
$n4g_m$	-0.06	0.030	0.062	0.142	-28.86
$v1_p$	0.08	0.031	0.016	0.235	-27.10
$v6_p$	0.07	0.034	0.043	0.177	-28.13
Variables entering with two lags					
$nq03c_q$	0.04	0.021	0.074	0.132	-27.77
$nq03d_q$	-0.07	0.038	0.079	0.145	-28.34
$nq04c_q$	0.04	0.021	0.047	0.170	-26.60
$nq04d_q$	-0.05	0.028	0.097	0.121	-28.53
$nq06a_q$	0.07	0.039	0.056	0.157	-26.76
$nq7a_q$	-0.18	0.097	0.061	0.175	-29.14
$nq7d_q$	0.04	0.018	0.040	0.173	-26.06
$nq7f_q$	-0.16	0.066	0.017	0.254	-26.49
$nq10g_q$	-0.06	0.036	0.087	0.121	-27.78
$n1b_m$	-0.08	0.047	0.075	0.126	-29.04
$n1w_m$	0.15	0.089	0.090	0.119	-26.76
$n2w_m$	0.20	0.100	0.044	0.172	-25.91
$n3m_m$	-0.08	0.047	0.093	0.151	-27.87
$n4b_m$	0.18	0.096	0.065	0.151	-28.00
$n4g_m$	-0.06	0.031	0.062	0.143	-27.23
$v1_p$	0.07	0.032	0.025	0.204	-26.04

Notes: See Notes for Table 1.

Table 3: Single-predictor Probit: Factors

Panel A: Confidence Data

Factor	S_{37t}	S_{42t}	S_{43t}	S_{37t-1}	S_{42t-1}	S_{43t-1}	S_{44t-1}	S_{34t-2}	S_{43t-2}	S_{44t-2}
Intercept	-0.37	-0.37	-0.36	-0.42	-0.43	-0.42	-0.39	-0.49	-0.47	-0.46
	(0.062)	(0.058)	(0.064)	(0.038)	(0.036)	(0.039)	(0.049)	(0.020)	(0.026)	(0.029)
Coef.	0.44	0.43	-0.43	0.55	0.47	-0.53	0.43	0.56	-0.62	0.55
	(0.029)	(0.037)	(0.034)	(0.010)	(0.025)	(0.013)	(0.044)	(0.035)	(0.006)	(0.018)
R^2_{es}	0.197	0.185	0.193	0.280	0.218	0.271	0.175	0.194	0.331	0.255
$ln\hat{L}$	-29.81	-29.06	-29.59	-27.39	-27.70	-27.63	-28.22	-30.43	-25.00	-26.93

Panel B: Macroeconomic and Financial Data

Factor	Z_{44t}	Z_{46t}	Z_{35t-1}	Z_{37t-1}	Z_{46t-1}	Z_{6t-2}	Z_{18t-2}	Z_{34t-2}	Z_{37t-2}	Z_{46t-2}
Intercept	-0.41	-0.37	-0.41	-0.41	-0.42	-0.41	-0.42	-0.45	-0.47	-0.45
	(0.047)	(0.065)	(0.038)	(0.039)	(0.041)	(0.047)	(0.065)	(0.027)	(0.022)	(0.030)
Coef.	0.69	0.58	0.41	0.40	-0.61	0.69	0.58	-0.43	-0.43	0.61
	(0.003)	(0.018)	(0.048)	(0.048)	(0.015)	(0.003)	(0.091)	(0.048)	(0.038)	(0.015)
R^2_{es}	0.403	0.256	0.170	0.163	0.268	0.403	0.121	0.178	0.180	0.270
$ln\hat{L}$	-26.48	-30.00	-29.72	-28.78	-26.61	-24.81	-30.05	-28.78	-27.18	-25.32

Panel C: Amalgamated Dataset Factors

Factor	W_{27t}	W_{33t}	W_{37t}	W_{45t}	W_{46t}	W_{24t-1}	W_{30t-1}	W_{33t-1}	W_{46t-1}	W_{46t-2}
Intercept	-0.36	-0.36	-0.35	-0.37	-0.37	-0.39	-0.39	-0.42	-0.43	-0.47
	(0.060)	(0.061)	(0.070)	(0.061)	(0.064)	(0.046)	(0.048)	(0.034)	(0.038)	(0.028)
Coef.	-0.33	0.33	0.37	-0.48	0.58	0.40	-0.34	0.36	0.68	0.63
	(0.097)	(0.097)	(0.063)	(0.025)	(0.015)	(0.050)	(0.094)	(0.077)	(0.009)	(0.012)
R^2_{es}	0.115	0.116	0.145	0.224	0.268	0.164	0.117	0.130	0.309	0.293
$ln\hat{L}$	-29.67	-31.04	-28.89	-25.49	-31.08	-31.12	-27.46	-28.84	-26.46	-24.78

Notes: Each probit model estimates $y^*_{t+1} = \bar{\alpha} + \alpha f_{t-d} + \epsilon_t$, where f_t a factor (either of type S_{it}, Z_{it} or W_{it}) and $d = 0, 1, 2$, estimated using PCA. Parameter estimates' p-values are in parenthesis under estimated coefficients. Only the coefficients significant at the 10% level are reported. $R^2_{es} \equiv 1 - (ln\hat{L}/lnL_0)^{(2/T)lnL_0}$ is Estrella's (1998) pseudo-R^2, where $ln\hat{L}$ is the estimated likelihood and lnL_0 is the likelihood computed only with a constant term. Finally, T is the sample size and the sample runs from 2002Q1 to 2014Q4.

Table 4: Probit with Multiple predictors: In-Sample Results

Panel A: Standard Model

Variables	$M1$	$M2$	$M3$	$M4$	$M5$
$Intercept$	-0.20	-0.41	-0.85	-2.08	-1.61
	(0.350)	(0.047)	(0.008)	(0.028)	(0.020)
y_{t-1}					
$InterestRate_t$	-				
	(-)				
$Termspread_t$	-				
	(-)				
$StockMarket_t$	-8.90				
	(0.051)				
$Ex.Rate_t$	-				
	(-)				
$Money_t$	-				
	(-)				
BCI_t		-			
		(-)			
CCI_t		-6.22			
		(0.030)			
BOS_t		-			
		(-)			
SLO_t		-			
		(-)			
Z_{44t}			0.95		
			(0.005)		
Z_{37t-1}			-0.76		
			(0.029)		
Z_{46t-1}			1.05		
			(0.006)		
Z_{34t-2}			-0.65		
			(0.013)		
S_{37t}				1.48	
				(0.029)	
S_{43t}				-1.46	
				(0.027)	
S_{42t-1}				2.48	
				(0.016)	
S_{44t-1}				1.77	
				(0.031)	
S_{37t-2}				1.68	
				(0.017)	
W_{33t}					0.84
					(0.026)
W_{37t}					1.42
					(0.014)
W_{46t}					1.17
					(0.037)
W_{24t-1}					1.13
					(0.038)
W_{33t-1}					1.35
					(0.010)
W_{46t-2}					1.76
					(0.022)
R^2_{es}	0.341	0.292	0.844	0.956	0.936
AIC	56.69	57.00	36.03	27.37	32.04
QPS	0.450	0.454	0.311	0.282	0.324

Panel B: Dynamic Model

Variables	$M1$	$M2$	$M3$	$M4$	$M5$
$Intercept$	-1.86	-1.83	-1.76	-1.82	-1.97
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
y_{t-1}	3.11	2.91	2.33	2.96	3.37
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
$InterestRate_t$	-				
	(-)				
$Termspread_t$	-				
	(-)				
$StockMarket_t$	-				
	(-)				
$Ex.Rate_t$	-				
	(-)				
$Money_t$	65.21				
	(0.028)				
BCI_t		-			
		(-)			
CCI_t		-			
		(-)			
BOS_t		-			
		(-)			
SLO_t		-0.03			
		(0.054)			
Z_{44t}			0.70		
			(0.050)		
Z_{37t-1}			-		
			()		
Z_{46t-1}			-		
			(-)		
Z_{34t-2}			-0.64		
			(0.040)		
S_{37t}				-	
				(-)	
S_{43t}				-	
				(-)	
S_{42t-1}				-	
				(-)	
S_{34t-2}				1.10	
				(0.023)	
S_{37t-2}				-	
				(-)	
W_{33t}					-
					(-)
W_{37t}					-
					(-)
W_{46t}					0.72
					(0.044)
W_{30t-1}					-0.98
					(0.045)
W_{33t-1}					-
					(-)
W_{46t-2}					-
					(-)
R^2_{es}	0.903	0.872	0.901	0.889	0.940
AIC	30.54	31.70	29.53	28.64	26.91
QPS	0.060	0.053	0.077	0.075	0.099

Notes: The left panel presents results from the standard probit while the right panel analyzes its dynamic version, both estimated on 2002Q1 to 2014Q4. Variables considered are: $M1$ = Classical predictors, $M2$ = Confidence indices, $M3$ = Macroeconomic factors, $M4$ = Confidence factors, $M5$ = Amalgamated-dataset factors. Coefficients' p-values are in parenthesis under estimates. $R^2_{es} \equiv 1 - (ln\hat{L}/lnL_0)^{-(2/T)lnL_0}$ is Estrella's (1998) pseudo-R^2, where $ln\hat{L}$ is the estimated likelihood and lnL_0 is the likelihood only with a constant term. Other reported performance measures include Akaike's asymptotic information criterion AIC and the quadratic probability score QPS.

Table 5: Probit with Multiple Predictors: Longer Forecasting Horizons

	Standard Model						Dynamic Model					
Explanatory variables	AIC	QPS	CML	$shots^+$	$shots^-$	$Forecast$	AIC	QPS	CML	$shots^+$	$shots^-$	$Forecast$
Horizon h=1												
Classical predictors	56.49	0.445	0.310	0.47	0.14	0.34	30.54	0.059	0.069	0.76	0.10	0.03
Confidence indices	57.00	0.450	0.329	0.47	0.10	0.36	31.70	0.052	0.069	0.81	0.10	0.06
Macroeconomic factors	36.02	0.304	0.239	0.80	0.03	0.05	29.53	0.075	0.073	0.80	0.14	0.04
Confidence factors	27.37	0.275	0.287	0.93	0.07	0.05	28.64	0.073	0.078	0.87	0.07	0.00
Complete-data factors	32.03	0.318	0.599	0.93	0.14	0.13	26.91	0.099	0.086	0.88	0.03	0.25
Horizon h=2												
Classical predictors	62.98	0.458	0.303	0.47	0.14	0.45	50.84	0.117	0.138	0.88	0.10	0.15
Confidence indices	58.55	0.439	0.333	0.18	0.03	0.34	50.72	0.120	0.141	0.88	0.10	0.15
Macroeconomic factors	40.53	0.306	0.260	0.83	0.07	0.08	47.23	0.162	0.156	0.60	0.10	0.16
Confidence factors	20.94	0.378	0.581	0.80	0.07	0.01	42.31	0.186	0.173	0.73	0.00	0.00
Complete-data factors	35.97	0.294	0.542	0.87	0.07	0.16	46.05	0.185	0.168	0.81	0.07	0.47
Horizon h=3												
Classical predictors	62.72	0.476	0.311	0.47	0.14	0.49	58.33	0.264	0.234	0.59	0.03	0.23
Confidence indices	58.94	0.444	0.342	0.12	0.00	0.33	57.46	0.276	0.231	0.13	0.00	0.22
Macroeconomic factors	50.16	0.340	0.286	0.53	0.07	0.13	57.33	0.296	0.253	0.27	0.00	0.24
Confidence factors	32.21	0.461	0.435	0.67	0.07	0.40	55.65	0.297	0.254	0.13	0.00	0.12
Complete-data factors	49.27	0.368	0.407	0.60	0.07	0.37	53.32	0.342	0.280	0.63	0.03	0.59
Horizon h=4												
Classical predictors	61.30	0.476	0.310	0.47	0.14	0.48	60.22	0.467	0.360	0.00	0.00	0.32
Confidence indices	58.16	0.467	0.360	0.00	0.00	0.32	57.94	0.497	0.382	0.00	0.00	0.32
Macroeconomic factors	50.22	0.388	0.327	0.40	0.03	0.10	55.08	0.541	0.425	0.13	0.03	0.34
Confidence factors	35.05	0.353	0.588	0.40	0.14	0.52	55.02	0.537	0.416	0.00	0.00	0.20
Complete-data factors	55.08	0.391	0.368	0.33	0.03	0.30	52.93	0.594	0.484	0.25	0.14	0.71

Notes: The left panel presents results from the standard probit in (1) while the right panel analyzes its dynamic version (6) for horizons $h = 1, 2, 3, 4$, on the sample 2002Q3 - 2014Q1. In addition to the AIC criterion, performance measures include the Quadratic Probability Score $QPS = \frac{2}{T}\sum_{t=1}^{T}(y_t - \widehat{p}_t)^2$, ($\widehat{p}_t$ is the model's predicted recession probability), the cost-weighted misclassification loss $CML = \frac{1}{T}\sum_{t=1}^{T}\left[(1-q)y_t(1-\mathbb{1}_{(\widehat{p}_t \geq 0.5)}) + q(1-y_t)(\mathbb{1}_{(\widehat{p}_t \geq 0.5)})\right]$, where $\mathbb{1}_{()}$ equals 1 if its argument is true and 0 otherwise and q is the relative cost of mistakes ($q = 1/3$). Finally, $shots^+$ is the proportion of true positives (recession hit rate) whereas $shots^-$ is the proportion of false negatives (missed expansion rate).

Table 6: Probit with Multiple Predictors: Earlier Sample (2002Q3 - 2010Q1)

	Standard Model						Dynamic Model					
Explanatory variables	AIC	QPS	CML	*shots*$^+$	*shots*$^-$	*Forecast*	AIC	QPS	CML	*shots*$^+$	*shots*$^-$	*Forecast*
Horizon h=1												
Classical predictors	45.70	0.452	0.309	0.38	0.13	0.77	39.15	0.260	0.164	0.92	0.17	0.83
Confidence indices	45.29	0.436	0.320	0.31	0.13	0.62	24.14	0.028	0.051	0.92	0.09	0.14
Macroeconomic factors	27.13	0.279	0.208	0.85	0.04	0.83	20.95	0.053	0.061	0.83	0.04	0.13
Confidence factors	26.43	0.274	0.247	0.67	0.04	0.06	14.94	0.091	0.074	0.92	0.04	0.00
Complete-data factors	22.08	0.259	0.212	0.82	0.09	0.14	21.11	0.051	0.06	0.91	0.04	0.01
Horizon h=2												
Classical predictors	49.54	0.462	0.303	0.38	0.13	0.59	49.32	0.353	0.224	0.92	0.17	0.57
Confidence indices	46.55	0.429	0.329	0.15	0.00	0.49	35.86	0.074	0.011	0.92	0.09	0.11
Macroeconomic factors	34.58	0.329	0.226	0.83	0.09	0.75	29.21	0.152	0.156	0.75	0.00	0.24
Confidence factors	26.69	0.328	0.228	0.50	0.04	0.06	21.12	0.200	0.173	0.75	0.04	0.00
Complete-data factors	29.48	0.274	0.249	0.73	0.09	0.13	31.40	0.122	0.128	0.82	0.04	0.05
Horizon h=3												
Classical predictors	49.00	0.482	0.311	0.38	0.13	0.53	50.76	0.440	0.282	0.92	0.17	0.53
Confidence indices	46.05	0.443	0.345	0.08	0.00	0.40	42.02	0.167	0.175	0.92	0.09	0.15
Macroeconomic factors	39.22	0.347	0.233	0.83	0.09	0.70	34.91	0.256	0.261	0.67	0.00	0.33
Confidence factors	17.83	0.499	0.518	0.58	0.13	0.01	19.29	0.371	0.382	0.58	0.09	0.00
Complete-data factors	27.00	0.324	0.287	0.55	0.04	0.13	38.31	0.214	0.202	0.64	0.04	0.10
Horizon h=4												
Classical predictors	47.38	0.472	0.307	0.38	0.13	0.56	48.98	0.536	0.352	0.23	0.09	0.56
Confidence indices	44.43	0.471	0.369	0.00	0.00	0.32	44.63	0.303	0.263	0.00	0.00	0.19
Macroeconomic factors	41.33	0.351	0.235	0.83	0.09	0.66	36.05	0.417	0.471	0.42	0.04	0.46
Confidence factors	17.87	0.377	0.406	0.42	0.04	0.01	38.60	0.361	0.308	0.17	0.17	0.21
Complete-data factors	31.6	0.407	0.321	0.27	0.13	0.29	42.26	0.339	0.288	0.10	0.00	0.18

Notes: The left panel presents results from the standard probit in (1) while the right panel analyzes its dynamic version (6) for horizons $h = 1, 2, 3, 4$, on the sample 2002Q3 - 2014Q1. In addition to the AIC criterion, performance measures include the Quadratic Probability Score $QPS = \frac{2}{T}\sum_{t=1}^{T}(y_t - \widehat{p}_t)^2$, ($\widehat{p}_t$ is the model's predicted recession probability), the cost-weighted misclassification loss $CML =$ $\frac{1}{T}\sum_{t=1}^{T}\left[(1-q)y_t(1-\mathbb{1}_{(\widehat{p}_t \geq 0.5)}) + q(1-y_t)(\mathbb{1}_{(\widehat{p}_t \geq 0.5)})\right]$, where $\mathbb{1}_{()}$ equals 1 if its argument is true and 0 otherwise and q is the relative cost of mistakes ($q = 1/3$). Finally, $shots^+$ is the proportion of true positives (recession hit rate) whereas $shots^-$ is the proportion of false negatives (missed expansion rate).

Table 7: An Out-of-Sample Experiment (2010Q2 to 2014Q1)

	Standard Model				Dynamic Model			
Explanatory variables	QPS	CML	$shots^{+}$	$shots^{-}$	QPS	CML	$shots^{+}$	$shots^{-}$
Classical predictors	0.457	0.130	0.5	0.21	0.227	0.060	0.70	0.00
Confidence indices	0.499	0.204	0.00	0.21	0.257	0.067	0.70	0.00
Macroeconomic factors	0.421	0.111	0.75	0.29	0.135	0.019	1.00	0.07
Confidence factors	0.347	0.074	0.75	0.14	0.113	0.056	0.75	0.07
Complete-data factors	0.375	0.056	1.00	0.21	0.120	0.037	1.00	0.14

Notes: The left panel presents results from the standard probit in (1) while the right panel analyzes its dynamic version (6) for horizons $h = 1, 2, 3, 4$, on the sample 2002Q3 - 2014Q1. In addition to the AIC criterion, performance measures include the Quadratic Probability Score $QPS = \frac{2}{T}\sum_{t=1}^{T}(y_t - \widehat{p}_t)^2$, ($\widehat{p}_t$ is the model's predicted recession probability), the cost-weighted misclassification loss $CML =$ $\frac{1}{T}\sum_{t=1}^{T}\left[(1-q)y_t(1-\mathbb{1}_{(\widehat{p}_t \geq 0.5)}) + q(1-y_t)(\mathbb{1}_{(\widehat{p}_t \geq 0.5)})\right]$, where $\mathbb{1}_{()}$ equals 1 if its argument is true and 0 otherwise and q is the relative cost of mistakes ($q = 1/3$). Finally, $shots^{+}$ is the proportion of true positives (recession hit rate) whereas $shots^{-}$ is the proportion of false negatives (missed expansion rate).

References

Heather M Anderson and Farshid Vahid. Predicting the probability of a recession with nonlinear autoregressive leading-indicator models. *Macroeconomic Dynamics*, 5(4):482–505, 2001.

Jushan Bai and Serena Ng. Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2):304–317, 2008.

Robert B Barsky and Eric R Sims. Information, animal spirits, and the meaning of innovations in consumer confidence. *The American Economic Review*, 102:1343–1377, 2012.

Roy Batchelor and Pami Dua. Improving macro-economic forecasts: The role of consumer confidence. *International Journal of Forecasting*, 14(1):71–81, 1998.

A. Binette and J. Chang. CSI: A model for tracking short-term growth in Canadian real GDP. Bank of Canada Review, Summer 2013.

Giorgio Bodo, Roberto Golinelli, and Giuseppe Parigi. Forecasting industrial production in the Euro area. *Empirical economics*, 25(4):541–561, 2000.

Jean Boivin and Serena Ng. Are more data always better for factor analysis? *Journal of Econometrics*, 132(1):169–194, 2006.

Andreas Buja, Werner Stuetzle, and Yi Shen. Loss functions for binary class probability estimation and classification: Structure and applications. *Working draft, November*, 2005.

Zhihong Chen, Azhar Iqbal, and Huiwen Lai. Forecasting the probability of US recessions: a probit and dynamic factor modelling approach. *Canadian Journal of Economics/Revue canadienne d'économique*, 44(2):651–672, 2011.

Charlotte Christiansen, Jonas Nygaard Eriksen, and Stig Vinther Møller. Forecasting US recessions: The role of sentiment. *Journal of Banking & Finance*, 49:459–468, 2014.

Philip Cross and Philippe Bergevin. Turning points: Business cycles in Canada since 1926. C. D. Howe Institute Commentary No. 366, October 2012.

A. Duarte, I. A. Venetis, and I. Paya. Predicting real growth and the probability of recession in the Euro area using the yield spread. *International Journal of Forecasting*, 21:261–277, 2005.

Arturo Estrella. A new measure of fit for equations with dichotomous dependent variables. *Journal of Business & Economic Statistics*, 16(2):198–205, 1998.

Arturo Estrella and Frederic S Mishkin. Predicting US recessions: Financial variables as leading indicators. *Review of Economics and Statistics*, 80(1):45–61, 1998.

Paolo Fornaro. Forecasting US recessions with a large set of predictors. *Journal of Forecasting*, 2016.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic factor model: One sided estimation and forecasting. *Journal of the American Statistical Association*, 100:830–840, 2005.

Jesper Hansson, Per Jansson, and Marten Löf. Business survey data: do they help in forecasting GDP growth? *International Journal of Forecasting*, 21:377–389, 2005.

Heikki Kauppi and Pentti Saikkonen. Predicting US recessions with dynamic binary response models. *The Review of Economics and Statistics*, 90(4):777–791, 2008.

Rachidi Kotchoni and Dalibor Stevanovic. Forecasting U.S. recessions and economic activity. August 2016.

Luisa Lambertini, Caterina Mendicino, and Maria Teresa Punzi. Expectation-driven cycles in the housing market: Evidence from survey data. *Journal of Financial Stability*, 9(4): 518–529, 2013.

Sylvain Leduc and Keith Sill. Expectations and economic fluctuations: an analysis using survey data. *Review of Economics and Statistics*, 95(4):1352–1367, 2013.

Kjetil Martinsen, Francesco Ravazzolo, and Ralph Wulfsberg. Forecasting macroeconomic variables using disaggregate survey data. *International Journal of Forecasting*, 30:65–77, 2014.

John G Matsusaka and Argia M Sbordone. Consumer confidence and economic fluctuations. *Economic Inquiry*, 33(2):296–318, 1995.

John Murray. Monetary policy decision making at the Bank of Canada. *Bank of Canada Review*, pages 1–9, 2013. Autumn.

Henri Nyberg. Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting*, 29(1-2):215–230, 2010.

P. Ollivaud, P A Pionnier, E Rusticelli, C Schwellnus, and Seung Hee Koh. A contest between small-scale bridge and large-scale dynamic factor models. OECD Economics Department Working Papers No. 1313, July 2016.

Lise Pichette. Extracting information from the *Business Outlook Survey* using statistical approaches. Bank of Canada Discussion paper No. 2012-8, December 2012.

Lise Pichette and M. N. Robitaille. Assessing the *Business Outlook Survey* underlying indicator using real-time data. Bank of Canada Discussion Paper, 2016.

Glenn D Rudebusch and John C Williams. Forecasting recessions: the puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics*, 27: 492–503, 2009.

James H Stock and Mark W Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460): 1167–1179, 2002a.

James H Stock and Mark W Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162, 2002b.

James H Stock and Mark W Watson. Forecasting with many predictors. *Handbook of Economic Forecasting*, 1:515–554, 2006.

Karl Taylor and Robert McNabb. Business cycles and the role of confidence: Evidence for Europe. *Oxford Bulletin of Economics and Statistics*, 69(2):185–208, 2007.

Jonathan H Wright. The yield curve and predicting recessions. 2006.

V. Zarnowitz and A. Ozyldirim. Time series decomposition and measurement of business cycles, trends, and growth cycles. *Journal of Monetary Economics*, 53:1717–1739, 2006.

A Canadian Survey Data on Sentiment

A.1 Conference Board Consumer Confidence survey

The Conference Board of Canada has been operating a monthly survey of Canadian households since 1979, to measure levels of optimism regarding current and future economic conditions. Surveyed households are asked to give their views about their current and expected financial positions and employment outlook. In addition, they are also asked to assess whether now is a good time or a bad time to make a major purchase such as a house, car or other big-ticket items. Specifically, the four questions comprising the survey are as follows:

1. Considering everything, would you say that your family is better or worse off financially than six months ago?

2. Again, considering everything, do you think that your family will be better off, the same or worse off financially six months from now?

3. How do you feel the job situation and overall employment will be in this community six months from now?

4. Do you think that right now is a good or bad time for the average person to make a major outlay for items such as a home, car or other major item?

Each question is answered positively or negatively; for example, a surveyed household answering that his family is better off financially than six months ago (first question) will be labelled as having responded positively.

The Conference Board then aggregates answers by calculating the ratio of positive responses for each question and taking the simple average across the four questions to create the publicly available *Index of Consumer Confidence*. As such, this index represents one specific way to aggregate information contained in the survey. As indicated in the text, the present paper assesses whether more flexible aggregation methods arising from factor models can improve on the signaling ability of these data.

A.2 Conference Board Business Confidence Survey

The Conference Board of Canada has been operating a quarterly survey of Canadian business executives since 1977. The survey is meant to measure perceptions of the current economic environment and the investment intentions of business. The questions comprising the survey, as well as all categories for the answers, are detailed below. As was the case for the Consumer survey, the Conference Board constructs an aggregate of survey answers, the *Index of Business Confidences* by summing the net ratio of positive answers to the third, fifth and eight questions below.

List of questions in the Business Confidence Survey

1. Do you expect overall economic conditions in Canada six months from now to be:
    - Better,
    - Worse,
    - The same.
2. Do you expect prices, in general, in Canada to increase over the next six months at an annual rate of: (data since 1987Q3)
    - $< 1\%$
    - 1%
    - 2%
    - 3%
    - 4%
    - 5%
    - 6%
    - 7%
    - 8%
    - $> 8\%$
3. Over the next six months, do you expect your firm's financial position to:
    - Improve,
    - Worsen,
    - Remain the same.

4. Over the next six months, do you expect your firm's profitability to:
   - Improve,
   - Worsen,
   - Remain the same.
5. Would you say the present is a good or a bad time to undertake expenditures to expand your plant or add to your stock of machinery and equipment?
   - Good,
   - Bad,
   - Not sure.
6. What change in the level of your capital investment expenditures do you expect over the next 6 months?
   - Up 20%
   - Up 10% to 19%
   - Up 1% to 9%
   - No change,
   - Down 1% to 9%
   - Down 10% to 19%
   - Down 20% or more.
7. In which region(s) of the country do you expect the bulk of your planned investment expenditures for the next six months to take place? (data since 1987$Q3$)
   - Atlantic Provinces,
   - Quebec,
   - Ontario,
   - Prairie Provinces,
   - British Columbia,
   - United States, (data since 1994$Q4$)
   - International. (data since 1994$Q4$)
8. How do you assess your current level of operations relative to optimal capacity?

- Above capacity,
- At or close to capacity,
- At, close to, or above, capacity
- Slightly below capacity,
- Substantially below capacity.

9. Compared with six months ago, what is your current rate of return to invested capital?
   - Better than expected,
   - As expected,
   - Worse than expected.
10. What factors, if any, are currently adversely affecting the level of your planned expenditures in Canada?
    - Excess productive capacity,
    - Weak market demand,
    - Foreign competition,
    - Rising cost of capital goods,
    - Rising labour costs,
    - Overall corporate liquidity,
    - High interest rates,
    - Weak commodity prices,
    - Shortage of qualified staff,
    - Government policies,
    - Taxes,
    - More attractive opportunities outside Canada,
    - Appreciation of the Canadian dollar,
    - Depreciation of the Canadian dollar.

A.3 Bank of Canada Business Outlook Survey

The *Business Outlook Survey* is a quarterly survey of the senior management of 100 Canadian businesses that are selected with a view to produce a representative selection of Canada's gross domestic product. The survey's was initiated in the Fall of 1997 and its purpose is to "gather the perspectives of these businesses on topics of interest to the Bank of Canada (such as demand and capacity pressures) and their forward-looking views on economic activity" and is conducted by the staff of the regional offices of the Bank. It was created to extend and formalize the informal discussion that the Bank as always conducted with relevant Canadian economic actors.

Each question elicits a categorical response from a surveyed firm (see below for the list of all questions and answer categories) and the Bank of Canada makes the percentage of firms answering each answer category publicly available. In addition, the Bank emphasizes a "balance of opinion" synthesis for each question, which is constructed by substracting the percentage of negative responses from the percentage of positive ones (balance of opinion can thus vary between -100 and 100). The complete list of all questions and answer categories is as follows:

List of questions in the Business Outlook Survey

1. (PAST SALES GROWTH) Over the past 12 months, the rate of increase in your firms sales volume (compared with the previous 12 months) was
   - Greater,
   - Less,
   - The same.
2. (FUTURE SALES GROWTH) Over the next 12 months, the rate of increase in your firms sales volume (compared with the past 12 months) is expected to be
   - Greater,
   - Less,
   - The same.
3. (FUTURE SALES GROWTH) Compared with 12 months ago, have your recent indicators (order books, advanced bookings, sales inquiries, etc.)... (Data since 2003Q3)
   - Improved,
   - Deteriorated.

4. (INVESTMENT IN MACHINERY AND EQUIPMENT) Over the next 12 months, your firms investment spending on M & E (compared with the past 12 months) is expected to be
    - Higher,
    - Lower,
    - The same.
5. (FUTURE EMPLOYMENT LEVEL) Over the next 12 months, your firms level of employment is expected to be
    - Higher,
    - Lower,
    - The same.
6. (ABILITY TO MEET DEMAND) How would you rate the current ability of your firm to meet an unexpected increase in demand? (Data since 1999$Q3$)
    - Some difficulty,
    - Significant difficulty.
7. (LABOUR SHORTAGES) Does your firm face any shortages of labour that restrict your ability to meet demand?
    - Yes,
    - No.
8. (INTENSITY OF LABOUR SHORTAGES) Compared with 12 months ago, are labour shortages generally...(Data since 2001$Q1$)
    - More intense,
    - Less intense.
9. (INPUT PRICE INFLATION) Over the next 12 months, are prices of products/services purchased expected to increase at a greater, lesser, or the same rate as over the past year?
    - Greater,
    - Less,
    - The same.
10. (OUTPUT PRICE INFLATION) Over the next 12 months, are prices of products/services sold expected to increase at a greater, lesser, or the same rate as over the past year?

- Greater,
- Less,
- The same.

11. (INFLATION EXPECTATIONS) Over the next two years, what do you expect the annual rate of inflation to be, based on the consumer price index? (Data since 2001Q2)

    - Above 3%,
    - 2% to 3%,
    - 1% to 2%,
    - Below 1%.

12. (CREDIT CONDITIONS) Over the past 3 months, how have the terms and conditions for obtaining financing changed (compared with the previous 3 months)? (Data since 2001Q4)

    - Tightened,
    - Eased.

A.4 Bank of Canada Senior Loan Officer Survey

The bank of Canada has been conducting the Senior Loan Officer Survey since 1999. This quarterly survey assesses the business-lending practices of major Canadian financial institutions, gathering information both on price but also non-price terms of business lending.

Specifically, the survey asks the Senior Loan Officer of participating institutions the following question: How have your institution's general standards (i.e. your appetite for risk) and terms for approving credit changed in the past three months?

- Tightened,
- Eased,
- Remain unchanged.

Surveyed institutions condition their answer on the evolution of business lending conditions by taking into account each of the following conditions:

1. Pricing of credit (spreads over base rates, fees),
2. General standards,
3. Limit of capital allocation,
4. Terms of credit (collateral, covenants, etc.),

so that a given institution could, in principle, report tightening conditions on pricing of credit but easing them with respect to general standards, limits or terms of credit. Two balance of opinion times series are made publicly available by the Bank of Canada: the balance of opinion to the pricing of credit, as well as a non-price aggregate to the remaining three categories (a "tightening" is coded if the institution reports tightening either general standards, limits of capital allocation or terms of credit).[27]

[27] The questions are further detailed as to whether they pertain to loans provided to corporate, commercial and small business firms; responses for commercial and small business firms are further provided for five regions: British Columbia, the Prairies, Ontario, Quebec, and the Atlantic provinces.

B Canadian Business Cycles according to the OECD

The OECD maintains a database of the status of the business cycles for each member country. The classification follows a *growth-cycle* methodology and the following dates of peaks and troughs are obtained for Canada:

Table 8: Chronology of the Canadian Business Cycle Since 1961

Monthly Trough	Monthly Peak
1961 M3	1962 M2
1963 M6	1966 M4
1968 M1	1968 M12
1971 M2	1974 M1
1975 M5	1976 M6
1977 M7	1979 M11
1982 M11	1985 M11
1986 M11	1989 M5
1992 M5	1994 M12
1996 M8	2000 M6
2001 M10	2002 M7
2003 M10	2007 M8
2009 M7	2011 M11
2012 M11	2014 M10

Note: Source: OECD. http://www.oecd.org/std/leading-indicators/CLI-components-and-turning-points.pdf.

C Variables in the macroeconomic and financial database

The table below lists all variables contained in our macroeconomic, financial and national accounts database. The table reports the series number, a description, the short name in the database and the original frequency before its transformation in quarterly data. The data are publicly available and originate from Statistics Canada, The Bank of Canada and other statistical agencies. The database is managed and used by the Department of Finance of the Government of Quebec for analysis of business cycles.

Table 9: Variable Names

N.	Description	Short Name	Freq.
1	BAs (bankers' acceptances): First contract	C1RAB3	d
2	BAs: Second contract	C2RAB3	d
3	BAs: Third contract	C3RAB3	d
4	BAs Fourth contract	C4RAB3	d
5	Corporate Canadian 10-year bond (BBB)	CORP10BBB	d
6	CRB index: Spot Commodity prices	CRBSPOT	d
7	Monetary Conditions Index	ICM	d
8	Commodity price index: Aluminum	IMPALUM	d
9	Commodity price index: Silver	IMPARGENT	d
10	Commodity price index: Live cattle	IMPBETAIL	d
11	Commodity price index: Wheat	IMPBLE	d
12	Commodity price index: Lumber	IMPBOIS	d
13	Commodity price index: Copper	IMPCUIVRE	d
14	Commodity price index: Natural Gas	IMPGAZ	d
15	Commodity price index: Nickel	IMPNICKEL	d
16	Commodity price index: Barley	IMPORGE	d
17	Commodity price index: Crude Oil	IMPPETROLE	d
18	Commodity price index: Lead	IMPPLOMB	d
19	Commodity price index: Pork	IMPPORC	d
20	Commodity price index: Zinc	IMPZINC	d
21	Net assets of chartered banks in foreign currency	AVOIRETRANG	m
22	Total assets of chartered banks in Canadian dollars	AVOIRTOTAL	m
23	Resident assets in Canadian dollars to chartered banks	AVRESAVCAN	m
24	Resident assets in foreign currency to chartered banks	AVRESAVETR	m
25	Resident deposits in foreign currency to chartered banks	AVRESDEPETR	m
26	Resident loans in foreign currency to chartered banks	AVRESPREETR	m

Continued on next page

Table 9 – continued from previous page

N.	Description	Short Name	Freq.
27	Exchange Rates Australia/CAD	CANAU	m
28	Exchange rate US/CAD	CANEU	m
29	Exchange rate EURO/CAD	CANEURO	m
30	Exchange Rates Swiss/CAD	CANFS	m
31	Exchange Rates UK/CAD	CANLS	m
32	Exchange Rate Japan/CAD	CANYE	m
33	Credit in shares and others	CRDENTACTI	m
34	Other corporate loans	CRDENTAUT	m
35	Short-term credit to firms, seasonally adjusted	CRDENTCT	m
36	Short-term credit to firms by the chartered bank, s. adjusted	CRDENTCTBC	m
37	Credit bonds and debentures	CRDENTOBLI	m
38	Credit for consumption, not seasonally adjusted	CREDITCONS	m
39	Participation rate for 15 years and older, seasonally adjusted	EPAACTIV	m
40	Unemployment rate for 15 years and older, seasonally adjusted	EPACHOMAGE	m
41	Jobs for 15 years and older, seasonally adjusted	EPAEMPLOIS	m
42	Population for 15 years and older, not seasonally adjusted	EPAPOP	m
43	Yield is at Constant Maturity Treasury Securities 10 Years Of Uscen	FCM10	m
44	Yield is at Constant Maturity Treasury Securities Of 3 Month Uscen	FTBS3	m
45	Total new construction set, seasonally adjusted	MEC	m
46	Currency outside banks	MMHORSBANQ	m
47	M1 ++	MMM1PLPLUS	m
48	M1 +	MMM1PLUS	m
49	Unfilled orders	MNFCOMMCARN	m
50	Manufacturing Shipments	MNFLIVRAIS	m
51	New manufacturing orders	MNFNOUVCOMM	m
52	The inventory-to-shipment ratio	MNFRATIO	m
53	Total manufacturing inventories	MNFSTOCKS	m
54	Building permits, total	PERMBAT	m
55	Building permits, non-residential	PERMBATNONRES	m
56	Building permits, residential	PERMBATRES	m
57	Canadian spot rate at noon - US $ in $ CDN	PFX	m
58	Canadian spot rate at noon	PFXI	m
59	Bankers' Acceptance to 1 month Yield	RAB1	m
60	BAs to 12 months Yield	RAB12	m
61	Bankers' Acceptance to 3 months Yield	RAB3	m
62	Yield BAs to 6 months Yield	RAB6	m
63	Bank rate (Official discount rate: Last Wednesday of the month)	RBANK	m

Continued on next page

Table 9 – continued from previous page

N.	Description	Short Name	Freq.
64	Treasury Bills to 1 year Yield	RBT12	m
65	Treasury bills to 3 months Yield	RBT90	m
66	Canadian government bonds on 10 years and more Yield	RC10	m
67	Government of Canada bonds on 1-3 years Yield	RC13	m
68	Government of Canada bonds on 3-5 years Yield	RC35	m
69	Government of Canada bonds on 5-10 years Yield	RC510	m
70	Canadian government bonds of 10 years Yield	RCF10	m
71	Canadian government bonds of 2 years Yield	RCF2	m
72	Canadian government bonds to 3 years Yield	RCF3	m
73	Canadian government bonds (30 years) Average yield	RCF30	m
74	Canadian government bonds (5 years) Average yield	RCF5	m
75	Spot bank rate daily target (RBANK - 0.25)	RCIBLE	m
76	Guaranteed Investment Certificate rate to 5 years	RCPG5	m
77	Canada Savings Bonds rate	RCSB	m
78	Canadian dollars Euro-3 months Yield	RE3	m
79	Mortgage rate of Canadian banks to 1 year	RHYP1	m
80	Mortgage rate of Canadian banks to 3 years	RHYP3	m
81	Mortgage rate of Canadian banks to 5 years	RHYP5	m
82	Morgage daily spot rate	RJOUR	m
83	Upper limit rate of the operating band of the Bank of Canada	ROBHIGH	m
84	Lower limit rate of the operating band of the Bank of Canada	ROBLOW	m
85	90 day Commercial paper rate	RPC90	m
86	Loans chartered banks - prime business loans rate	RPRIME	m
87	Non-checkable saving deposits rate	RSDB	m
88	Toronto Stock Exchange Composite Share Price Index	TSX	m
89	Unit labor costs	CUM	q
90	Real GDP at market prices (growth contribution)	PIB_CC	q
91	Real GDP at market prices	PIB	q
92	Personal consumption expenditures (growth contribution)	PIBC_CC	q
93	Personal consumption expenditures	PIBC	q
94	Personal consumption expenditures on durables goods (growth contribution)	PIBCDUR_CC	q
95	Personal consumption expenditures on durable goods	PIBCDUR	q
96	Personal consumption expenditures on non-durable goods (growth contribution)	PIBCNONDUR_CC	q
97	Personal consumption expenditures on non-durable goods	PIBCNONDUR	q
98	Personal consumption expenditures on semi-durable goods (growth contribution)	PIBCSEMIDUR_CC	q
99	Personal consumption expenditures on semi-durable goods	PIBCSEMIDUR	q
100	Personal consumption expenditures on services (growth contribution)	PIBCSERV_CC	q

Continued on next page

Table 9 – continued from previous page

N.	Description	Short Name	Freq.
101	Personal consumption expenditures on services (growth contribution)	PIBCSERV	q
102	GDP deflator	PIBDEGONF	q
103	Final domestic demand (growth contribution)	PIBDIF_CC	q
104	Final domestic demand	PIBDIF	q
105	Government expenditures (growth contribution)	PIBG_CC	q
106	Government expenditure (GDP)	PIBG	q
107	Government current expenditures on G & S (growth contribution)	PIBGC_CC	q
108	Government current expenditures on G & S	PIBGC	q
109	Government investments except inventories (growth contribution)	PIBGI_CC	q
110	Government investments except inventories (growth contribution)	PIBGI	q
111	Government investments in machinery and equipment	PIBGIMM	q
112	Government investments in software	PIBGIMMLOG	q
113	Government investments in computers & office supplies	PIBGIMMORDI	q
114	Government investments in telecommunications	PIBGIMMTELECOM	q
115	Government investments in inventories (growth contribution)	PIBGSTOCKS_CC	q
116	Government investments ininventories	PIBGSTOCKS	q
117	Business investment except inventories (growth contribution)	PIBI_CC	q
118	Business investment except inventories	PIBI	q
119	Business investment in Non-residential construction (growth contribution)	PIBICNONRES_CC	q
120	Business investment in Non-residential construction	PIBICNONRES	q
121	Business investment in Non-residential construction (growth contribution)	PIBICNONRESCC	q
122	Business investment in Residential construction (growth contribution)	PIBICRES_CC	q
123	Business investment in Residential construction	PIBICRES	q
124	Business investment in machinery and equipment (growth contribution)	PIBIMM_CC	q
125	Business investment in machinery and equipment	PIBIMM	q
126	Business investment in software	PIBIMMLOG	q
127	Business investment in computers & office supplies	PIBIMMORDI	q
128	Business investment in telecommunications	PIBIMMTELECOM	q
129	Non-residential Business investment and equipment (growth contribution)	PIBINONRES_CC	q
130	Non-residential Business investment and equipment	PIBINONRES	q
131	Imports of goods and services (GDP) (growth contribution)	PIBM_CC	q
132	Imports of goods and services (GDP)	PIBM	q
133	Nominal GDP	PIBNOMINAL	q
134	Business investment in inventories (growth contribution)	PIBSTOCKS_CC	q
135	Business investment in inventories	PIBSTOCKS	q
136	Exports of goods and services (GDP) (growth contribution)	PIBX_CC	q
137	Exports of goods and services (GDP)	PIBX	q

Continued on next page

Table 9 – continued from previous page

N.	Description	Short Name	Freq.
138	Net exports (GDP) (growth contribution)	PIBXM_CC	q
139	Net exports (GDP)	PIBXM	q
140	Nationnal real gross product at market prices	PNB	q
141	Labour productivity	PRODUCTIVITE	q
142	Corporate profits after taxes	PROFITS_POSTTAX	q
143	Personal disposable income	RPD	q
144	Utilization rate of industrial capacity by (NAICS)	TUCI	q