Forecasting Recessions in Canada: An Autoregressive Probit Model Approach

Antoine Poulin-Moore* Kerem Tuzcuoglu†

February 12, 2024

*Bank of Canada. E-mail: APoulin-Moore@bank-banque-canada.ca
†Bank of Canada. E-mail: ktuzcuoglu@bank-banque-canada.ca

Acknowledgements

We would like to express our gratitude to Daniel de Munnik, Christopher Hajzler, and Rodrigo Sekkel for their precious comments and suggestions. We would also like to thank seminars participants at the Bank of Canada for their useful comments. Finally, we are indebted to Connor Neff and LJ Valencia for their valuable research assistance. The mistakes are, of course, our own. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Bank of Canada.

Abstract

We forecast recessions in Canada using an autoregressive (AR) probit model. The presence of the lagged latent variable in this model results in an intractable likelihood with a high dimensional integral. We employ composite likelihood methods, which facilitates the estimation of this complex model, and provide their asymptotic results. We perform a variable selection procedure on a large variety of Canadian and foreign macro-financial variables by using the area under the receiver operating characteristic curve (AUROC) as the performance criterion. Our findings suggest that the AR model improves meaningfully the forecasting performance of Canadian recessions – relative to its static counterpart and a variety of probit models proposed in the Canadian literature. This finding is robust to changes in the performance criteria or the sample considered. Our results also highlight the short-term predictive power of the US economic activity and confirm that financial indicators are reliable predictors of Canadian recessions, aligning with the existing literature.

Topics: Business fluctuations and cycles, Econometric and statistical methods

JEL codes: E32, C53, C51

1 Introduction

Forecasting recessions always has been of great interest in macroeconomics, given how such episodes can have a significant, pervasive and persistent impact on various sectors of the economy. Foreseeing the different phases of the business cycles is also critical for policymakers, as it may influence their ability to conduct appropriate monetary and fiscal policies. A large body of literature, following Estrella and Hardouvelis (1991) and Estrella and Mishkin (1998), use static probit models to predict recessions. Several studies find the yield curve, defined as the spread between yields from government bonds with longer and shorter maturities, to be a useful recession predictor. An inversion of the yield curve, that is a negative bond yield spread, is nearly a perfect signal of recessions in the United-States. An inverted yield curve is also found to be a reliable leading indicator of Canadian recessions (Atta-Mensah and Tkacz, 1998).[1] Other financial and macroeconomic leading indicators, such as stock prices (Estrella and Mishkin, 1998), credit market activity (Levanon et al., 2015), credit spreads, and certain employment indicators (Ng, 2014), were found to be good predictors of recessions.

Chauvet and Potter (2005) extend the static probit model of Estrella and Mishkin (1998) by incorporating breakpoints and autocorrelated unobservables. However, Kauppi and Saikkonen (2008) suggest that the Bayesian estimation in Chauvet and Potter (2005) can be computationally intensive. Therefore, they propose ad hoc dynamic probit models that can simply be estimated by maximum likelihood estimation. The authors show that the inclusion of the lagged binary and lagged latent variables can improve both the in-sample and out-of-sample recessions forecasting performance of a static probit model in the United-States. Hao and Ng (2011) replicated this analysis for the Canadian economy, but found mixed forecasting performances across the variety of static and dynamic probit models they considered. The authors also highlight that the bond yield spread, housing starts, real money supply, and the composite index of leading indicators can be helpful to predict recessions in Canada. Finally, Fossati et al. (2018) showed that factor-augmented

[1] This stylized fact can be seen in Figure 2 in the Appendix B.1, which shows the evolution of the bond yield spread in Canada and in the United States over the last sixty years.

static probit models outperform their counterparts which are based solely on observed data. In addition, they find the real activity factor to be a good predictor of Canadian recessions, especially at shorter-term forecast horizons.

This paper contributes to the literature in two important respects. Foremost, it is the first one to propose an autoregressive (AR) probit model to predict Canadian recessions while providing its formal underlying econometric and asymptotic theories. Compared to a static probit model, the latent variable in this paper is modeled as an autoregressive process, which captures the persistence in the underlying state of the business cycle. The presence of the lagged latent variable smooths the recession probabilities and prevents false-positive predictions, which makes it a great candidate for a recession prediction model. However, the AR probit model, which is similar to one of the specifications of Chauvet and Potter (2005), has a complex likelihood function, containing high-dimensional integrals. This paper avoids computationally intensive Bayesian or other simulation-based techniques and use composite likelihood (CL) estimation methods (Lindsay, 1988; Varin et al., 2011). In particular, we use *marginal* composite likelihood (MCL) that significantly reduces the complexity of the *full* likelihood. MCL estimation takes only a few seconds, which makes our method much more attractive for practitioners. Additionally, in contrast to Kauppi and Saikkonen (2008) and Hao and Ng (2011), we provide the asymptotic theory for our estimators.[2] While asymptotic results on MCL estimation are available for more general models, such as general state space models (Varin and Vidoni, 2008; Ng et al., 2011) or AR panel probit model with correlated random effects (Tuzcuoglu, 2023), to the best of our knowledge, we are the first to provide them for the single-equation, time series version of the AR probit. In addition, we also compare our empirical results to those obtained using *pairwise* composite likelihood (PCL) estimators.

Second, the variable selection in this paper is performed from a broader and more up-to-date variety of Canadian and foreign macro-financial variables relative to the existing

[2]The statistical properties of some of the proposed methods in Kauppi and Saikkonen (2008) are not known. They rely on the results of de Jong and Woutersen (2011) that are valid only for the probit model with lagged observed variable. However, for the model with a lagged latent variable, which is more akin to our model, Kauppi and Saikkonen (2008) state that there is no formal proof of the validity for their (asymptotic) results.

Canadian literature. It includes various real economic activity indicators for both Canada and the United-States, several commodity prices indices, measures of international trade, and financial variables. We also consider for each of these twenty-six variables up to twelve lags, which amounts to a total of three hundreds and twelve potential leading indicators. By virtue of our fast estimation technique, we are able to run tens of thousands of probit models to select the best predictors. This paper also incorporates the latest recession dates published by the C.D.Howe Institute Business Cycle Council, which now includes the COVID-19 recession and changes to previous recession dates. The model selection and its performance assessment are conducted in this paper using the area under the receiver operating characteristic curve (AUROC) as a performance criteria. This metric can provide an objective assessment of a model's classification accuracy improving on broadly-used goodness of fit measures and scoring rules (Berge and Jordà, 2011).[3]

Our empirical analysis focuses on the Canadian recessions which occurred between June 1973 and December 2022.[4] We perform a variable selection procedure based on both in-sample and pseudo out-of-sample performances. Our procedure finds the following variables as the best predictors: the Canadian bond yield spread, the Chicago Fed National Activity Index, and the TSX Composite Index. While the bond yield spread tends to predict recessions well in advance, the CFNAI and TSX are usually better at shorter-term forecasts. The reliability of US economic indicators to predict Canadian activity aligns with the existing literature (e.g. Bragoli and Modugno (2017)), highlighting the fact that Canada has a small and open economy who's international trade rely heavily on its southern neighbour. Finally, our estimation also yields a high autocorrelation coefficient capturing the persistence of the recession indicator.

Our empirical results also highlight the superiority of the AR model over its static counterpart, both in-sample and out-of-sample. Compared with the static probit model, the AR probit has a significantly better fit to the data, forecasts the turning points of

[3]Our results are robust to a variety of criteria such as the pseudo R^2, the quadratic probability score and the root-mean-square error.

[4]The C.D. Howe Institute Business Cycle Council is an arbiter of business cycle dates in Canada – similar to NBER's Business Cycle Dating Committee in the US. Since the publication of Hao and Ng (2011), the C.D. Howe Institute changed a few recession dates. Thus, by itself, this calls for the need of reassessing the findings of the literature on Canadian recessions.

business cycles more accurately, and yields much smoother probability forecasts that result in a sizeable reduction in recession false signals. We also compare our model to those existing in the Canadian recession forecasting literature, specifically the four static and dynamic models considered in Hao and Ng (2011). The forecasting results show clear evidence for significant upper hand by the AR probit model. This indicates that using a model with an autoregressive latent component and a large set of potential predictors yield enhancement in forecasting Canadian recessions. Our results are also robust to a variety of changes, such as using a different sample period that excludes the COVID-19 episode and using different performance criteria.

Finally, it is also worth mentioning another commonly used probit model for recession prediction: the *dynamic* probit model that incorporates the lagged observed dependent variable as a predictor. While this model has a satisfactory in-sample fit performance, it is doing rather a poor job when it comes to out-of-sample forecasting. For instance, Hao and Ng (2011) in-sample results show that this model almost always predicts the recessions with a one period lag. Similarly, Kauppi and Saikkonen (2008) results show a significant deterioration in out-of-sample forecasting power of this model after one period. One obvious reason is that the lagged binary variable, which appears to be the main determinant of the current state of the economy in empirical studies, is not known in the out-of-sample forecasting exercise since recessions are announced with a significant lag.[5] Therefore, we do not consider this model in our analysis.

The rest of this paper is organized as follow. Section 2 introduces our model, its estimation, its asymptotic properties and the AUROC that we use to assess its performance. Section 3 presents the data. Section 4 shows our in-sample and pseudo out-of-sample empirical results. Section 5 shows their robustness. Finally, section 6 concludes.

[5]The average lag in NBER announcements for the start and end of US recessions is 7 and 15 months, respectively.

2 Methodology

In this section, we introduce the autoregressive probit model and its estimation by composite likelihood methods. In particular, we provide the marginal composite likelihood estimator of the AR probit model and discuss its asymptotic results and forecasting procedure.

2.1 Autoregressive Probit Model

For $t = 1, \ldots, T$, we consider the following AR probit model

$$\begin{aligned} y_t &= \mathbb{1}[y_t^* \geq 0], \\ y_t^* &= \mu + \rho y_{t-1}^* + \boldsymbol{\beta}' \mathbf{x}_{t-m} + \varepsilon_t, \end{aligned} \tag{1}$$

where y_t is the binary outcome, y_t^* is the underlying continuous latent process, $\mathbf{x}_{t-m}$ is a $K-$dimensional vector of lagged observable covariates, ε_t is the unobservable error term, ρ is the autocorrelation coefficient of the latent process, $\boldsymbol{\beta}$ is the coefficient vector, and μ is the constant term. Note that m is the employed lag order of $\mathbf{x}$, which means that the data of $\mathbf{x}$ is assumed to be available for $t = -m+1, \ldots, T$. The model also allows for different lags for each covariate, but for simplicity we use a common lag notation. In practice, one chooses $m \geq H$, where H is the forecast horizon. Let $\boldsymbol{\theta} = (\rho, \mu, \boldsymbol{\beta}')'$ is the $(K+2)$-dimensional vector of parameters. For stationarity of the latent process y_t^*, we assume $|\rho| < 1$. The error term ε_t is assumed to be independent and identically distributed with $\mathcal{N}(0, (1-\rho^2)\sigma_\varepsilon^2)$ where $\sigma_\varepsilon^2 = 1$ is assumed for identification purposes – a typical assumption in probit and logit models (Greene, 2003). The multiplication of the error distribution by $\sqrt{1-\rho^2}$ is just a reparametrization that facilitates the mathematical terms in the distribution of y_t^*.

Note that the AR probit model differs from the static (ST) probit only by the the term ρy_{t-1}^*, which generates persistence both in y_t^* and y_t. However, its presence significantly complicates the likelihood of the AR probit model since y_{t-1}^* is unobserved and needs to be integrated out. This results in the following likelihood function containing a T-dimensional

integral

$$L(\mathbf{y}|\mathbf{x};\boldsymbol{\theta}) = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_T}^{b_T}\boldsymbol{\phi}_T(y_1^*,\ldots,y_T^*|\mathbf{x};\boldsymbol{\theta})\mathrm{d}y_1^*\cdots\mathrm{d}y_T^*,$$

where the limits of integration are time-varying such that $(a_t, b_t) = (-\infty, 0)$ if $y_t = 0$, and $(a_t, b_t) = (0, \infty)$ if $y_t = 1$, for all $t = 1, \ldots, T$; $\boldsymbol{\phi}_T(\cdot)$ is the T-dimensional joint Gaussian density; and, $\mathbf{x} = (\mathbf{x}'_{-m+1}, \ldots, \mathbf{x}'_{T-m})'$ and $\mathbf{y} = (y_1, \ldots, y_T)'$.[6] Calculating the likelihood function requires evaluation of the T-dimensional joint Gaussian which is computationally demanding even for moderate T.

2.2 Marginal composite likelihood estimator

Composite likelihoods reduces the number of integrals by ignoring dependencies between certain subsets of $(y_1^*, \ldots, y_T^*)$. In this paper, we focus on the Marginal Composite Likelihood (MCL) that utilizes univariate distributions of y_t^*. For this, we use backward-substitution in equation (1):

$$\begin{aligned}
y_t^* &= \mu + \rho y_{t-1}^* + \boldsymbol{\beta}'\mathbf{x}_{t-m} + \varepsilon_t, \\
&= (1+\rho)\mu + \rho^2 y_{t-2}^* + \boldsymbol{\beta}'\mathbf{x}_{t-m} + \rho\boldsymbol{\beta}'\mathbf{x}_{t-m-1} + \varepsilon_t + \rho\varepsilon_{t-1}, \\
&\vdots \\
&= \frac{1-\rho^t}{1-\rho}\mu + \rho^t y_0^* + \sum_{k=0}^{t-1}\rho^k\boldsymbol{\beta}'\mathbf{x}_{t-m-k} + \sum_{k=0}^{t-1}\rho^k\varepsilon_{t-k}. \qquad (2)
\end{aligned}$$

Next, we need to make an assumption on the initial latent value y_0^*. There are various possibilities here: (i) assuming a particular non-random value for it such as $y_0^* = 0$ as in Chauvet and Potter (2005) or $y_0^* = (\mu + \boldsymbol{\beta}'\bar{\mathbf{x}})/(1-\rho)$ as in Kauppi and Saikkonen (2008) and Hao and Ng (2011), (ii) treating it as another parameter to be estimated as in Müller and Czado (2005), (iii) drawing it from a stationary distribution as (indirectly assumed) in Varin and Vidoni (2006). These are ad hoc assumptions on y_0^*, but different choices do not

[6] Note that we ignored the initial value y_0^* in this formula for simplicity. Depending on the modeling choice of y_0^*, the dimension of the integral could increase to $T+1$.

significantly affect the parameter estimates since the importance of y_0^* vanish in a large T setting.[7] However, the choice matters for the analytical derivations. Note that the first and second conditional moments of y_t^* will depend on y_0^*, and thus in general on time period t. To achieve time-homogeneity in the conditional mean, variance and covariances, we follow the option (iii) and assume that the initial latent variable is drawn from a stationary distribution such that

$$y_0^* = \frac{1}{1-\rho}\mu + \frac{1}{\sqrt{1-\rho^2}}\varepsilon_0.$$

The underlying assumption here is that data generating process has started a long time ago and reached its stationary distribution before the initial date of our observed data.

Incorporating y_0^* into (2), we obtain

$$y_t^* = \frac{1}{1-\rho}\mu + \sum_{k=0}^{t-1}\rho^k \boldsymbol{\beta}'\mathbf{x}_{t-m-k} + \frac{\rho^t}{\sqrt{1-\rho^2}}\varepsilon_0 + \sum_{k=0}^{t-1}\rho^k\varepsilon_{t-k}. \tag{3}$$

Hence, the first and second conditional moments of y_t^* can be derived as follows:

$$\begin{aligned}
\mathrm{E}[y_t^*|\mathbf{x}_{-m+1},\ldots,\mathbf{x}_{t-m};\boldsymbol{\theta}] &= \frac{1}{1-\rho}\mu + \sum_{k=0}^{t-1}\rho^k\boldsymbol{\beta}'\mathbf{x}_{t-m-k} \\
\mathrm{Var}[y_t^*|\mathbf{x}_{-m+1},\ldots,\mathbf{x}_{t-m};\boldsymbol{\theta}] &= \frac{\rho^{2t}\mathrm{Var}(\varepsilon_0)}{1-\rho^2} + \sum_{k=0}^{t-1}\rho^{2k}\mathrm{Var}(\varepsilon_{t-k}) \\
&= \frac{\rho^{2t}(1-\rho^2)\sigma_\varepsilon^2}{1-\rho^2} + \frac{1-\rho^{2t}}{1-\rho^2}(1-\rho^2)\sigma_\varepsilon^2 \\
&= 1 \\
\mathrm{Cov}[y_t^*, y_{t-j}^*|\mathbf{x}_{-m+1},\ldots,\mathbf{x}_{t-m};\boldsymbol{\theta}] &= \frac{\rho^{2t-j}\mathrm{Var}(\varepsilon_0)}{1-\rho^2} + \rho^j\sum_{k=0}^{t-j-1}\rho^{2k}\mathrm{Var}(\varepsilon_{t-j-k}) \\
&= \frac{\rho^{2t-j}\mathrm{Var}(\varepsilon_0)}{1-\rho^2} + \rho^j\frac{1-\rho^{2t-2j}\mathrm{Var}(\varepsilon_t)}{1-\rho^2} \\
&= \rho^j.
\end{aligned}$$

[7] Although not reported here, we verify this claim in Monte Carlo simulations and our estimations.

The univariate conditional distribution of y_t^* and the associated conditional recession probabilities can be written as

$$y_t^*|\mathbf{x} \sim \mathcal{N}\left(\frac{\mu}{1-\rho} + \sum_{k=0}^{t-1} \rho^k \boldsymbol{\beta}' \mathbf{x}_{t-m-k},\ 1\right),$$
$$\mathbb{P}(y_t = 1|\boldsymbol{x}) = \mathbb{P}(y_t^* \geq 0|\boldsymbol{x}) = \boldsymbol{\Phi}\left(\frac{\mu}{1-\rho} + \sum_{k=0}^{t-1} \rho^k \boldsymbol{\beta}' \mathbf{x}_{t-m-k}\right), \tag{4}$$

where $\mathbb{P}$ is the probability operator. Finally, we define the marginal composite log-likelihood as

$$\begin{aligned}
\mathcal{L}_{MCL}(\boldsymbol{\theta}|\mathbf{y}, \mathbf{x}) &= \frac{1}{T}\sum_{t=1}^{T} \ln f(y_t|\mathbf{x}; \boldsymbol{\theta}) \\
&= \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}(y_t = 1) \ln \mathbb{P}(y_t = 1|\mathbf{x}; \boldsymbol{\theta}) + \mathbb{1}(y_t = 0) \ln \mathbb{P}(y_t = 0|\mathbf{x}; \boldsymbol{\theta}) \\
&= \frac{1}{T}\sum_{t=1}^{T} y_t \ln \boldsymbol{\Phi}\left(\frac{\mu}{1-\rho} + \sum_{k=0}^{t-1} \rho^k \boldsymbol{\beta}' \mathbf{x}_{t-m-k}\right) \\
&\qquad + (1 - y_t) \ln \boldsymbol{\Phi}\left(-\frac{\mu}{1-\rho} - \sum_{k=0}^{t-1} \rho^k \boldsymbol{\beta}' \mathbf{x}_{t-m-k}\right),
\end{aligned} \tag{5}$$

where $\boldsymbol{\Phi}(\cdot)$ is the cumulative distribution function of a standard normal distribution.

One can also use other CL functions, such as a *Pairwise* Composite Likelihood (PCL) that utilizes bivariate distributions of $(y_1^*, \ldots, y_T^*)$. Note that in this case we have the following bivariate conditional distribution:

$$\begin{bmatrix} y_t^* \\ y_{t-j}^* \end{bmatrix} \Big| \mathbf{x} \sim \mathcal{N}\left(\begin{bmatrix} \frac{\mu}{1-\rho} + \sum_{k=0}^{t-1} \rho^k \boldsymbol{\beta}' \mathbf{x}_{t-k} \\ \frac{\mu}{1-\rho} + \sum_{k=0}^{t-j-1} \rho^k \boldsymbol{\beta}' \mathbf{x}_{t-j-k} \end{bmatrix}, \begin{bmatrix} 1 & \rho^j \\ \rho^j & 1 \end{bmatrix}\right). \tag{6}$$

The PCL relies on conditional bivariate probabilities of the form $\mathbb{P}\left(y_t = s_1,\ y_{t-j} = s_2|\mathbf{x}; \boldsymbol{\theta}\right)$ where $s_1, s_2 \in \{0, 1\}$. In the empirical analysis, we will compare the results of MCL and PCL estimations, but our main focus in this paper will be on the former.

2.3 Asymptotic properties

Let's define the MCL estimator as $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathcal{L}_{MCL}(\boldsymbol{\theta}|\mathbf{y},\mathbf{x})$, where $\boldsymbol{\Theta}$ is the compact parameter space containing the true parameter $\boldsymbol{\theta}_*$ in its interior. We further assume that the covariates, $\boldsymbol{x}$, have finite fourth moments, are strictly exogenous, satisfy the non-singularity condition for $E[\boldsymbol{x}_t\boldsymbol{x}_t']$.[8] Following standard asymptotic literature (Amemiya, 1985; Newey and McFadden, 1994) and similar results in the CL literature (Lindsay, 1988; Varin et al., 2011; Ng et al., 2011; Tuzcuoglu, 2023), the asymptotic distribution of the MCL estimator, as $T \to \infty$, can be given by

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) \to_d \mathcal{N}\left(0, \boldsymbol{\mathcal{H}}^{-1}(\boldsymbol{\theta}_*)\boldsymbol{\Omega}(\boldsymbol{\theta}_*)\boldsymbol{\mathcal{H}}^{-1}(\boldsymbol{\theta}_*)\right),$$

where $\boldsymbol{\mathcal{H}}(\boldsymbol{\theta})$ is the Hessian matrix and $\boldsymbol{\Omega}(\boldsymbol{\theta})$ is the long-run variance of the score function. A detailed proof of the asymptotic result, consistent estimators for $\boldsymbol{\mathcal{H}}(\boldsymbol{\theta})$ and $\boldsymbol{\Omega}(\boldsymbol{\theta})$ matrices, and further details can be found in the Technical Appendix A.

2.4 Forecasting

The h-period ahead latent variable conditional on information at time T, where $m \geq h$, can be written as

$$\begin{aligned} y^*_{T+h} &= \mu + \rho y^*_{T+h-1} + \boldsymbol{\beta}'\mathbf{x}_{T+h-m} + \varepsilon_{T+h}, \\ &= \frac{1-\rho^h}{1-\rho}\mu + \rho^h y^*_T + \sum_{k=0}^{h-1}\rho^k\boldsymbol{\beta}'\mathbf{x}_{T+h-m-k} + \sum_{k=0}^{h-1}\rho^k\varepsilon_{T+h-k}. \end{aligned}$$

Note that the distribution of the composite error term is not a standard Gaussian distribution, instead, it is $\sum_{k=0}^{h-1}\rho^k\varepsilon_{T+h-k} \sim \mathcal{N}\left(0, 1-\rho^{2h}\right)$. Hence, when computing the h-period ahead

[8] These are commonly used assumptions in nonlinear dynamic panel data models (Honoré and Kyriazidou, 2000; Wooldridge, 2005; Bartolucci and Nigro, 2010, 2012).

forecasts, we need to scale the latent variable by $\sqrt{1-\rho^{2h}}$:

$$\mathbb{P}_T(y_{T+h}=1)=\mathbb{P}_T(y^*_{T+h}\geq 0)=\mathbf{\Phi}\left(\frac{\frac{1-\rho^h}{1-\rho}\mu+\rho^h y^*_T+\sum_{k=0}^{h-1}\rho^k\boldsymbol{\beta}'\mathbf{x}_{T+h-m-k}}{\sqrt{1-\rho^{2h}}}\right),$$

where $\mathbb{P}_T$ is the probability function conditional on information at time T.

Even though our MCL resembles the likelihood of the second model proposed by Kauppi and Saikkonen (2008) (also utilized by Hao and Ng (2011)), we deviate from their model in the forecasting procedure by taking into account the distribution of the moving-average error term, i.e., by including the scaling factor of $\sqrt{1-\rho^{2h}}$. The larger is ρ and the shorter is the forecast horizon h, the more important becomes the scaling. Moreover, the shorter the forecasting horizon the more impact the scaling has. However, for small ρ and distant forecasting horizons, the scaling factor gets close to one and becomes unimportant.

2.5 Evaluation Criterion

Various measures have been proposed in the literature to assess the goodness of fit of models with dichotomous dependent variables. First, many alternatives to the standard linear regression R^2, such as *pseudo* R^2 measures, have been suggested (see, for example, McFadden (1974), Cragg and Uhler (1970), Efron (1978), Estrella and Hardouvelis (1991) and Estrella and Mishkin (1998)). Likelihood-based goodness of fit criteria, such as those of Akaike (1973) and Schwarz (1978) (AIC and BIC), have also been used in the literature. While these measures provide an evaluation of dichotomous models' quality of fit, they do not assess the models' abilities to accurately classify binary outcomes. Another branch of the economics literature, such as Diebold and Rudebusch (1989) or more recently Kauppi and Saikkonen (2008), has rather leveraged scoring rules which do provide an assessment of a model classification accuracy. Some examples include Brier et al. (1950) quadratic probability score (QPS), the log probability score, and the root mean square error (RMSE). However, the performance assessment based on these scores relies on the intrinsic structure of their loss function as highlighted in Berge and Jordà (2011).

To overcome these limitations, we select the area under the receiver operating characteristic

curve (AUROC) to assess the performance of our probit models. This measurement has the advantage to provide a performance assessment that is independent of any loss function since it is constructed by solely using true and false positive rates. Berge and Jordà (2011) argues that this criterion has a major advantage over alternatives when it comes to prediction or forecast accuracy for binary outcomes. While some studies use the ROC analysis for US recessions (e.g. Liu and Moench (2016)), to the best of our knowledge, our paper is the first one to use it to assess the classification accuracy of recessions in Canada.

The ROC curve is a simple graphical representation which summarizes the classification ability of a model with a binary dependent variable. Let's define the estimated recession binary variable as $\hat{y}_t(\tau) \equiv \mathbb{1}\left[\mathbb{P}(y_t = 1|\boldsymbol{x}) > \tau\right]$, where τ is a threshold value and $\mathbb{P}(y_t = 1|\boldsymbol{x})$ is the model-based conditional recession probability (showed in Equation (4) for the AR probit model). Given a set of observed and estimated recession binary variables, we can express a model's true positive rate $TPR(\tau)$ and false positive rate $FPR(\tau)$ at any threshold $\tau \in [0, 1]$ as

$$TPR(\tau) = \frac{\sum_{t=1}^{T} \mathbb{1}[\hat{y}_t = 1,\ y_t = 1]}{\sum_{t=1}^{T} \mathbb{1}[y_t = 1]} \qquad \text{and} \qquad FPR(\tau) = \frac{\sum_{t=1}^{T} \mathbb{1}[\hat{y}_t = 1,\ y_t = 0]}{\sum_{t=1}^{T} \mathbb{1}[y_t = 0]}.$$

Note that, for any given level of threshold τ, the model provides a different set of $\hat{y}_t$'s, and thus, different TPRs and FPRs. The ROC curve plots TPR against FPR for all $\tau \in [0, 1]$, representing the trade-off between them. Finally, the AUROC is obtained by calculating the area under the ROC curve, which is independent from τ. The higher is the AUROC value, the better is the classification performance of a model. Furthermore, we can statistically test the difference in two AUROC values by using bootstrap methods (Carpenter and Bithell, 2000), which is important for comparing model performances.

Note that studies in the recession forecasting literature commonly use $\tau = 25\%$ or $\tau = 50\%$ as the threshold values to define a recession. Consequently, we provide our OOS classification results also under these two fixed threshold values. As further robustness checks, we also present performance comparisons based on other criteria, such as pseudo R^2, QPS, and RMSE. The AR probit model has a superior forecasting performance under any classification criterion.

3 Data

This section discusses the data underlying the binary recession variables and the leading indicators. Our dataset covers the time period between June 1973 and December 2022. More details such as data sources, variable abbreviations and their detailed descriptions can be found in Table 9 of Appendix B.2.

3.1 Canadian Recessions

In this analysis, the binary variable y_t is set equal to one if the economy is experiencing a recession in period t and to zero otherwise. To define this binary variable, we use the definition from the C.D.Howe Institute Business Cycle Council.[9] The council defines recessions as periods of pronounced, pervasive, and persistent decline in aggregate economic activity. For instance, the Great Recession is a good example of such a period of contraction. During that episode, the fall in real GDP and employment were sharp, broad-based across industries and lasted for an extended period. Over the last fifty years, other economic downturns have occurred in Canada, without necessarily meeting all the magnitude, length, and scope criteria required to be considered a recession. For instance, the burst of the dotcom bubble in 2001 and the 2014-15 oil price shock resulted in economic downturns, but they were not classified as a recession by the C.D. Howe Institute. The list of the recessions considered in this paper are provided in Table 1.

It is worth noting that the commonly recognized definition of a recession has greatly evolved since the latest publications in the Canadian recession forecasting literature. For instance, Hao and Ng (2011) define a recession in terms of the cumulative absence of positive growth over two consecutive quarters, in line with Cross (1996) and Cross (2001). More recent papers, such as Fossati et al. (2018), use a definition of a recession consistent with the first report from the C.D. Howe Institute Business Cycle Coucil (Cross and Bergevin, 2012). However, the council has revised notably their recession dates since their first report, reflecting various factors such as Statistics Canada's expansion of the expenditure-based

[9] In this paper, we use their latest 2021 update. See Cross and Bergevin (2012) for a more comprehensive review of their original methodology.

Table 1: Recessions in Canada between 1974 and 2022

Peak	Trough	Description
1974-October	1975-March	1973 oil crisis
1981-June	1982-October	Monetary policy tightening
1990-March	1992-May	Gulf War
2008-October	2009-May	Great Financial Crisis
2020-February	2020-April	COVID-19

Notes: The recession dates presented in this table reflect the definition of the C.D. Howe Business Cycle Council, consistent with their ninth and most recent report (C.D.Howe, 2021).

GDP time series back to 1961 from 1981 (C.D.Howe, 2017). Methodological changes have also been implemented to better reflect the importance of the breadth to determine recessions (C.D.Howe, 2019). By itself, these changes in the definition of a recession call for the need of reassessing the findings of the literature on Canadian recessions, which we undertake in this paper.

3.2 Explanatory variables

To select potential explanatory variables, we start by considering macro-financial indicators that were found to be informative to predict recessions in Canada based on Hao and Ng (2011), which include the bond yield spread (SP), housing starts (HS), real money supply (M1) and a composite leading indicator (CLI).[10] We build on the existing literature by considering additional domestic indicators.

First, we add the Canadian version of the Sahm rule (SAHM) based on Sahm (2019). This measurement represents the difference between the 3-month moving average of the unemployment rate relative to its prior 12-month low. While this indicator tends to lag slightly the peaks in the business cycles, it is a highly reliable signal of recessions in the US, which also happen to be the case for Canada (see Figure 3 in Appendix B.1). Second, we consider other domestic indicators included in the CLI and considered in the literature (e.g. Liu and Moench (2016)). This includes the consumer confidence index (CCI), the

[10]Given that the original CLI series used by used Hao and Ng (2011) is discontinued, we replace it using the OECD Composite Leading Indicator.

S&P/TSX composite index (TSX) and building permits (BP).

In order to better reflect some of the specificities of the Canadian economy, we also consider a variety of global and commodity prices indicators. As highlighted in Binette et al. (2017), exports play a particularly important role for a small open economy like Canada. In fact, exports represent about one third of Canadian GDP and are highly volatile. Therefore, they have a substantial impact on quarterly GDP dynamics.[11] Accordingly, we incorporate in our pool of potential leading indicators various foreign economic activity measures, with a particular focus on US economic indicators. This prevalence of the US variables stems from the fact that about three-quarters of Canadian merchandise exports are shipped towards their southern neighbour. In particular, we embed in our pool of variables: global exports (WEX), real Canadian merchandise exports (EX) and imports (IM), US employment (USE), US industrial production (USIP), the US Purchasing Manager Index (USPMI), and the Chicago Fed National Activity Index 3-month moving average (CFNAI).

The commodity-related sector also plays an important role in the Canadian economy, which motivates the inclusion of various energy and non-energy commodity prices indicators. In 2019, the nominal share of production related to this sector was about 7.5 percent.[12] Canada is also a major commodity exporter, with commodities representing more than 55% of Canadian goods exports. We measure commodity prices using the Bank of Canada Commodity Price Index (BCPI), which is a chain Fischer index of the spot price of 26 commodities produced in Canada (Kolet and Macdonald, 2010). We also include seven key BCPI subcomponents, such as oil and metal prices (BCPI-O & BCPI-M), and the non-energy commodity prices index (BCNE).

4 Empirical Analysis

In this section, we discuss the best predictors for Canadian recessions and present our in-sample and pseudo out-of-sample results.

[11] For instance, the average absolute annualized contribution to GDP growth from exports stands close to three percentage points between 1961 and 2022.

[12] See the annual GDP by industry account from Statistics Canada for agriculture, forestry, fishing and hunting, mining, quarrying, and oil and gas extraction.

4.1 In-sample results

In this section, we present the variable selection procedure and provide an in-sample performance comparison of the autoregressive (AR) and static (ST) probit models, formally defined in section 2. In order to identify the optimal set of covariates (and their lags) for these models, we follow a procedure similar to Hao and Ng (2011):

1. First, we conduct a pre-screening of the potential explanatory variables. To do so, we estimate single-regressor static equations for each variables identified in Table 9 of Appendix B.2, allowing their lags to vary between one and twelve-months ahead. This results in twelve different models per explanatory variable. We estimate these models using MCL with data going from June 1973 to December 2022, and compute their in-sample AUROC values (Table 10 of Appendix C). [13] Then, we create a shortlist by selecting the ten regressors having the highest average AUROC values across these forecast horizons (Table 2).[14]

2. Second, we combine the best regressor, at its optimal lag, with any two remaining variables from Table 2, by allowing the lags of these two additional regressors to vary between one and twelve months ahead. We assess the performance of the resulting three-variables AR or ST probit model using their in-sample AUROC values.

3. Third, from the best thirty models in-sample identified in step 2, we select the model having the highest AUROC value obtained from the pseudo out-of-sample (OOS) exercise (see section 4.2 for more details on the OOS exercise). This step ensures a balance between the in-sample and out-of-sample performances.

Table 2 shows the best ten explanatory variables based on their average in-sample AUROC values. Consistent with the literature, the bond yield spread stands out as the best single predictor at longer forecast horizons (above six months-ahead). The solid performance of this indicator at a longer range partly reflect the forward-looking information

[13] Note that we assess all models in-sample performance over a fix period going from November 1974 to December 2022.

[14] This step allows us to conduct the combined optimization procedure on a smaller list of variables, which is less computationally intensive.

Table 2: Shortlist of potential explanatory variables

		Forecast horizon (h)												
Order	Mnemonic	1	2	3	4	5	6	7	8	9	10	11	12	*Avg.*
1	CFNAI	0.874	0.864	0.864	0.853	0.838	0.809	0.785	0.760	0.755	0.744	0.741	0.727	0.801
2	SP	0.660	0.697	0.728	0.755	0.782	0.803	0.821	0.822	0.826	0.832	0.837	0.842	0.784
3	SAHM	0.854	0.823	0.797	0.771	0.744	0.713	0.682	0.653	0.627	0.615	0.614	0.614	0.709
4	CLI	0.602	0.649	0.706	0.740	0.754	0.756	0.747	0.728	0.708	0.691	0.680	0.672	0.703
5	USE	0.754	0.753	0.738	0.701	0.681	0.654	0.649	0.627	0.631	0.608	0.615	0.586	0.666
6	USIP	0.711	0.692	0.709	0.679	0.666	0.658	0.631	0.612	0.638	0.610	0.621	0.614	0.653
7	BCNE	0.650	0.655	0.662	0.654	0.677	0.683	0.672	0.667	0.635	0.611	0.610	0.597	0.648
8	TSX	0.651	0.661	0.693	0.671	0.671	0.684	0.668	0.654	0.618	0.614	0.593	0.584	0.647
9	BCPI-M	0.622	0.640	0.676	0.710	0.738	0.703	0.666	0.629	0.595	0.571	0.596	0.576	0.644
10	M1	0.615	0.641	0.647	0.627	0.643	0.648	0.625	0.610	0.611	0.595	0.614	0.600	0.623

Notes: this table shows the in-sample AUROC values calculated from single-variable static probit models, with lags varying between one and twelve-months ahead. The last column provides the average AUROC values across these forecast horizons. The models are estimated using MCL with data going from June 1973 to December 2022. This table contains the ten best explanatory variables based on their average in-sample AUROC. See Table 10 in Appendix C for the complete list.

it conveys about market participants perceptions of risk and expectations of future interest rates. The CFNAI, however, appears to be the best single predictor of economic downturns at shorter horizons, outperforming the other indicators when considering forecasts conducted between one and six-month ahead. US employment and industrial production are also strong predictors at this forecast horizon, although to a lower extent. The strong leading properties of these US economic indicators on Canadian downturns is quite intuitive given that Canada has a small and open economy which international trade relies heavily on its southern neighbour.[15] Other domestic variables, such as the Canadian Sahm indicator and the composite leading indicator are also particularly informative for Canadian recessions at shorter forecast horizons. Finally, commodity prices offer a moderate performance which generally peaks before the seven-month ahead forecast horizon. Commodity-prices excluding energy (BCNE) and metal prices (BCPI-M) stands out as the best commodity prices subindices.

Table 3 shows the regression results for the optimal AR and ST model specifications that were obtained using the variable selection procedure previously described. The bond yield spread and the CFNAI appear among the best predictors, as they are selected in both

[15]In 2022, about 75% of Canadian nominal exports were shipped towards the United-States.

the AR and the ST equations. In addition, the TSX composite index and real money supply are selected as the third regressors in the AR and in the ST models, respectively. All these regressors' coefficients have negative values, suggesting that an improvement in these indicators imply a lower recession probability. It is worth noting that the AR model has a high and statistically significant autoregressive coefficient ρ, highlighting the strong auto-correlation component of the latent process y_t^*, and thus that of y_t. Finally, we can see that the AR model outperforms the ST model, with a significant five percentage points advantage in its AUROC value. Table 11 in the Appendix C also shows that this advantage is robust to a variety of performance measurements, such as the pseudo R^2, the QPS and the RMSE.

Table 3: In-sample regression results

	Specifications	
	AR_{MCL}	ST
SP_{t-7}	-0.116^{***}	
	(0.022)	
$CFNAI_{t-1}$	-0.052	-0.640^{***}
	(0.040)	(0.140)
TSX_{t-1}	-0.091^{***}	
	(0.024)	
SP_{t-12}		-0.503^{***}
		(0.066)
$M1_{t-2}$		-0.426^{***}
		(0.117)
y_{t-1}^*	0.922^{***}	
	(0.016)	
Constant	-0.071^{*}	-1.090^{***}
	(0.039)	(0.108)
AUROC	0.987	0.9379
DIFF	0.049^{***}	

Notes: This regression table contains the parameters of the optimal AR and ST probit models, estimated by MCL using data from June 1974 to December 2022. The last line shows the difference between both models AUROC values, along with its test of significance (Carpenter and Bithell, 2000). *p<0.1; **p<0.05; ***p<0.01.

4.2 Out-of-sample results

We now compare the performance of the AR and ST probit models by assessing their forecasting abilities in a more realistic situation. Given that real-time data vintages are not available for every regressors contained in our models, we limit our analysis to a pseudo out-of-sample forecasting exercise. To proceed, we first estimate the models using a sample going from June 1973 to March 1989 and obtain a first one-month ahead forecast for April 1989. Next, we add this last observation to our estimation sample and obtain a second one-month ahead forecast for May 1989. We repeat this expanding-window procedure recursively until we obtain a forecast for December 2022, which corresponds to the last month of our sample. The choice of this forecasting period allows us to test the models on three recessions, namely the 1990's recession, the GFC, and the COVID-19 pandemic. For the sake of simplicity, we assume that the values of the explanatory variables x_t are know at the same time as the binary variable y_t. Note that relaxing this assumption, to better account for the recession dates publication delays, would only have a marginal impact on the relative performance of the ST and AR models.[16]

Table 4 compares the OOS performance of the two models considered at a one-month ahead forecast horizon. Such as for the in-sample results, we can see that the AR model outperforms the ST model based on their AUROC values. This advantage holds also for other performance measurements. As a sanity check, we also verify in Table 4 that our AR model outperforms reproduced versions of the models proposed in Hao and Ng (2011).[17]

Let's have a closer look at how the models perform during the three recessions contained in the simulation period. Figure 1 plots the one-month ahead forecasted recession probabilities for both the AR and ST probit models.[18] First, we can observe that both the AR and ST models provide a meaningful response during the three recessions analyzed, although

[16] Certain papers, such as Hao and Ng (2011), assume that the values of the explanatory variables x_t are known six months in advance relative to the recession binary variable y_t.

[17] The four probit models presented by Hao and Ng (2011) are labelled as: static (ST_{HN}), dynamic (DYN_{HN}), autoregressive (AR_{HN}) and dynamic autoregressive ($DYNAR_{HN}$). Note that methodological differences, such as the forecasts scaling, differentiate AR_{MCL} from AR_{HN}. It is also worth noting that the authors use a different sample and recession definition to select their optimal models, which makes this exercise an imperfect comparison.

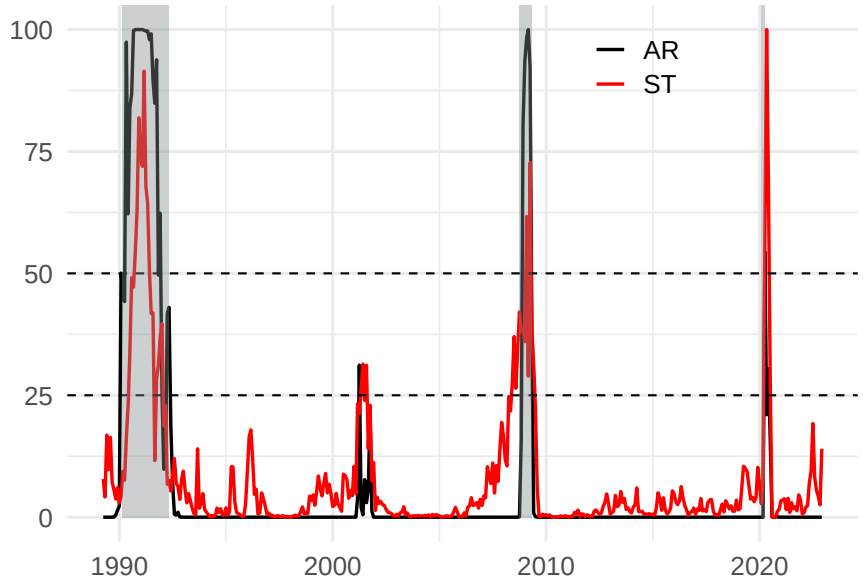
[18] We only present the ST model in Figure 1, as it represents the best benchmark for the OOS exercise.

Table 4: Models OOS performance comparison ($h = 1$)

Model	AUROC	DIFF	R2	QPS	RMSE
AR_{MCL}	98.2%		78.3%	3.4%	13.1%
ST	96.6%	1.5%	45.9%	8.5%	20.7%
ST_{HN}	93.6%	4.5%	46.3%	8.5%	20.6%
DYN_{HN}	87.1%	11.1%***	44.7%	8.7%	20.9%
AR_{HN}	91.3%	6.8%***	43.0%	9.0%	21.2%
$DYNAR_{HN}$	87.3%	10.9%***	43.9%	8.9%	21.0%

Notes: The second column shows the difference between the AR_{MCL} and the specified model AUROC values, along with its test of significance. *p<0.1; **p<0.05; ***p<0.01. $R2$, QPS and $RMSE$ stand for the pseudo R^2, quadratic probability score and root-mean squared error, respectively. The first two rows reflect the performance of the AR and ST probit models presented in this paper. The last four rows reflect the performance of the static (ST_{HN}), dynamic (DYN_{HN}), autoregressive (AR_{HN}) and dynamic autoregressive ($DYNAR_{HN}$) probit models replicated from Hao and Ng (2011).

Figure 1: Forecasted probabilities from OOS exercise ($h = 1$)



Notes: this figure compares the one-month ahead forecasts ($h = 1$) between the AR_{MCL} and ST probit models over the OOS exercise simulation period (April-1989 to December 2022). The two dashed lines represent the 25% and 50% thresholds, respectively.

not always identifying precisely the turning points. Both models also yield a significant response in 2001, which corresponds to the burst of the dotcom bubble. Despite this episode not being considered an official recession, the 2001 pick-up in the AR model forecasted probabilities mostly reflect the deterioration in US economic activity and in

Table 5: Turning points identification (OOS)

	25% threshold			*25% threshold*	
Peaks	AR_{MCL}	ST	Troughs	AR_{MCL}	ST
March 1990	-2	+3	May 1992	-4	-4
October 2008	+1	-5	May 2009	0	+1
February 2020	+1	+1	April 2020	0	+3

Notes: $(+a)$ or $(-a)$ imply that a model identifies a turning point *"a"* months too late or too early. A peak is identified too early (late) if the model's forecasted probability exceed a given threshold before (after) the actual peak. A trough is identified too early (late) if the model's forecasted probability fell below a given threshold before (after) the actual trough. The lags and leads in this table are assessed at the 25% threshold.

the TSX composite index. It is worth noting that the ST model's forecasts are much more volatile, and therefore, provide more false signals at lower thresholds relative to the AR model. For instance, the ST models' FPR at the 12.5% threshold is about 8%, almost four times larger than the AR model's rate (Table 12 in Appendix C). In contrast, we can see that the AR model's forecasts are much smoother than the ST model outside of the recessions, reflecting the model's higher persistence due to its autoregressive feature.

In Table 5, we assess more precisely the models' accuracy to identify the turning points in the business cycles over the simulation period. First, we can see that the AR model is generally more precise than the ST to signal peaks in the business cycles. While the AR model is a bit early in calling the 1990s recession, it is only one month late in identifying the peaks preceding the GFC and the COVID-19 pandemic. In contrast, the ST model is generally off by a larger amount, especially in the case of the GFC, where the model identifies the peak five months too early. The AR advantage over the ST equation is also obvious when looking at the identification of the troughs: the AR model perfectly identifies the turning points at the 25% threshold for the last two recessions of the OOS exercise. Note that the choice of the threshold can impact the models ability to identify the turning points, although the relative advantage of the AR model at other thresholds remains (see Table 13 in Appendix C).

Finally, we compare in Table 6 the in-sample and OOS performance between the autoregressive probit model estimated using the MCL and PCL estimators (AR_{MCL} and

AR_{PCL}, respectively). While the estimated parameters using PCL should theoretically be more efficient than those obtained by using MCL, our results in Table 6 suggest that no empirical advantage is obtained from using the former estimator. In addition, it is worth noting that the PCL estimation takes significantly longer that the MCL estimation. However, there are two main caveats preventing us to reach robust conclusions from this exercise. First, the optimal AR model presented in this paper was selected using the MCL estimators, resulting into an obvious advantage in favour of AR_{MCL}. Second, this comparison is conducted using one model only. A more thorough comparison, including several model specifications, would be required to generalize these findings.

Table 6: Performance comparison between AR_{MCL} and AR_{PCL} (AUROC)

Model	**In-sample**	**OOS**
AR_{MCL}	0.9871	0.9817
AR_{PCL}	0.9865	0.9807
Difference	0.0006	0.0010

Notes: This table shows the in-sample and OOS AUROC values obtained from the MCL and PCL estimations.

5 Robustness

In this section, we assess the sensitivity of our empirical results to the exclusion of the COVID-19 recession from both the in-sample and pseudo out-of-sample analysis. Given the historical data volatility observed during this period, it is reasonable to assess how the relative performance of the models respond to this change. Table 7 presents the regression results of the AR_{MCL} and of the ST model estimated with a sample going from June 1973 to December 2019, and using the same selection procedure as described in subsection 4.1. For the AR_{MCL} model, the variable selection is fairly close to the complete in-sample analysis, with the only difference being that the TSX variable is now lagged by two months instead of one. For the ST model, $M1_{t-2}$ is now replaced by $SAHM_{t-9}$ and the bond yield spread has now a shorter lag (seven months instead of twelve). We can see that both models in-sample AUROC values are better relative to those obtained when including the COVID-19 pandemic (see Table 14 in Appendix C for the complete in-sample results).

Although the difference between the AR_{MCL} and the ST models' AUROC values is smaller, the former model continues to have a significant 1.7 percentage points advantage over its static counterpart. Table 8 presents the results of the pseudo out-of-sample exercise conducted between April 1989 and December 2019. Similar to the in-sample results, all AUROC values are still better relative to those obtained using the complete sample. Again, the advantage of the AR model over the ST model reduces when looking at the AUROC values, but remains substantial when looking at the other performance criteria. Finally, the AR model significantly outperforms the replicated models from Hao and Ng (2011) through all performance criteria considered in the OOS exercise, supporting the robustness of our results.

Table 7: In-sample regression results (ex. COVID)

	Specifications	
	AR_{MCL}	ST
SP_{t-7}	-0.138^{***}	-0.564^{***}
	(0.022)	(0.080)
$CFNAI_{t-1}$	-0.056	-1.507^{***}
	(0.044)	(0.195)
TSX_{t-2}	-0.102^{***}	
	(0.019)	
$SAHM_{t-9}$		0.795^{***}
		(0.124)
y^*_{t-1}	0.909^{***}	
	(0.021)	
Constant	-0.099^{**}	-1.978^{***}
	(0.049)	(0.191)
AUROC	0.993	0.975
DIFF	0.017^{***}	

Notes: This regression table contains the parameters of the optimal AR and ST probit models, estimated by MCL using data from June 1974 to December 2019. The last line shows the difference between both models AUROC values, along with its test of significance (Carpenter and Bithell, 2000). *p<0.1; **p<0.05; ***p<0.01.

Table 8: Models OOS performance comparison ($h = 1$, ex. COVID)

Model	AUROC	DIFF	R2	QPS	RMSE
AR_{MCL}	99.5%		75.4%	4.0%	14.1%
ST	99.1%	0.3 %	51.9%	7.8%	19.8%
ST_{HN}	94.6%	4.8%*	48.2%	8.4%	20.5%
DYN_{HN}	87.5%	11.9%***	50.3%	8.1%	20.1%
AR_{HN}	92.0%	7.4%**	48.9%	8.3%	20.4%
$DYNAR_{HN}$	87.7%	11.8%***	49.4%	8.2%	20.3%

Notes: The second column shows the difference between the AR_{MCL} and the specified model AUROC values, along with its test of significance.*p<0.1; **p<0.05; ***p<0.01. $R2$, QPS and $RMSE$ stand for the pseudo R^2, quadratic probability score and root-mean squared error, respectively. The first two rows reflect the performance of the AR and ST probit models presented in this paper. The last four rows reflect the performance of the static (ST_{HN}), dynamic (DYN_{HN}), autoregressive (AR_{HN}) and dynamic autoregressive ($DYNAR_{HN}$) probit models replicated from Hao and Ng (2011).

6 Conclusion

Recessions have significant, pervasive and persistent economic consequences, and therefore, their prediction has attracted a great interest from both academicians and practitioners. In this paper, we use an autoregressive probit model to forecast recessions in Canada. Compared to its static counterpart, the AR model contains a lagged latent variable, which helps capturing the autocorrelation in the recession binary variable. However, the AR probit model results in an intractable likelihood function containing a high dimensional integral. Therefore, we propose using composite likelihood methods which yield consistent, asymptotically normally distributed, and less computationally-intensive estimators. This fast estimation method allows us to perform a variable selection procedure on a large variety of Canadian and foreign macro-financial variables. We use the AUROC as the classification performance criterion, although our results are robust to a variety of performance measurements.

Our results suggest that the best leading indicators of Canadian recessions are the Chicago Fed National Activity Index (CFNAI), the Canadian government bond yield spread, and the the S&P/TSX composite index. In particular, the CFNAI has a short-term predictive power on Canadian recessions reflecting the interconnectedness between both economies. In contrast, the bond yield spread is a reliable long-term predictor for recessions,

aligning with the existing literature. Our model comparison shows that the AR probit model provides superior in-sample and pseudo out-of-sample performances relative to its static version. The AR model fits the data better, identifies the business cycle turning points more accurately, and yields much smoother recession probability forecasts that results in much fewer false signals. Finally, our AR probit model has a superior forecasting performance compared to a variety of static and dynamic probit models proposed in the Canadian recession literature.

References

Akaike, Hirotogu, "Information theory and an extension of the maximum likelihood principle," in B N Petrov and F Caski, eds., *Proc. Second International Symposium on Information Theory*, Budapest: Akademiai Kiado 1973, pp. 267–281.

Amemiya, Takeshi, *Advanced Econometrics*, Harvard University Press, 1985.

Atta-Mensah, Joseph and Greg Tkacz, "Predicting Canadian recessions using financial variables: A probit approach," Technical Report, Bank of Canada 1998.

Bartolucci, Francesco and Valentina Nigro, "A Dynamic model for binary panel data with unobserved heterogeneity admitting a $\sqrt{n}$-consistent conditional estimator," *Econometrica*, 2010, *78* (2), 719–733.

__ **and** __ , "Pseudo conditional maximum likelihood estimation of the dynamic logit model for binary panel data," *Journal of Econometrics*, 2012, *170* (1), 102–116.

Berge, Travis J and Òscar Jordà, "Evaluating the classification of economic activity into recessions and expansions," *American Economic Journal: Macroeconomics*, 2011, *3* (2), 246–77.

Binette, André, Tony Chernis, and Daniel De Munnik, "Global Real Activity for Canadian Exports: GRACE," Technical Report, Bank of Canada Staff Discussion Paper 2017.

Bragoli, Daniela and Michele Modugno, "A now-casting model for Canada: Do US variables matter?," *International Journal of Forecasting*, 2017, *33* (4), 786–800.

Brier, Glenn W et al., "Verification of forecasts expressed in terms of probability," *Monthly weather review*, 1950, *78* (1), 1–3.

Carpenter, James and John Bithell, "Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians," *Statistics in medicine*, 2000, *19* (9), 1141–1164.

C.D.Howe, *Business Cycle Council Communique - December 2017, C.D. Howe Institute*, 2017.

__ , *cdhowe.org*, 2019.

__ , *cdhowe.org*, 2021.

Chauvet, Marcelle and Simon Potter, "Forecasting recessions using the yield curve," *Journal of Forecasting*, 2005, *24* (2), 77–103.

Cragg, John G and Russell S Uhler, "The demand for automobiles," *The Canadian Journal of Economics/Revue Canadienne d'Economique*, 1970, *3* (3), 386–406.

Cross, Philip, "Alternative measures of business cycles in Canada: 1947-1992," *Canadian Economic Observer*, 1996, *9* (2), 3–1.

__ **and Philippe Bergevin**, "Turning points: Business cycles in Canada since 1926," *CD Howe Institute*, 2012, *366*.

Cross, Phillip, "Tracking the Business Cycle: Monthly Analysis of the Economy at Statistics Canada, 1926-2001," *Canadian Economic Observer*, 2001, *14* (12), 3–1.

de Jong, Robert M and Tiemen Woutersen, "Dynamic time series binary choice," *Econometric*

Theory, 2011, *27* (04), 673–702.

Diebold, Francis X and Glenn D Rudebusch, "Scoring the leading indicators," *Journal of business*, 1989, pp. 369–391.

Efron, Bradley, "Regression and ANOVA with zero-one data: Measures of residual variation," *Journal of the American Statistical Association*, 1978, *73* (361), 113–121.

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

__ **and Gikas A Hardouvelis**, "The term structure as a predictor of real economic activity," *Journal of Finance*, 1991, *46* (2), 555–576.

Fossati, Sebastian, Rodrigo Sekkel, and Max Sties, "Forecasting Recessions in Canada," 2018.

Gallant, A Ronald, "Nonlinear statistical models," 1987.

Greene, William H, *Econometric analysis*, Pearson Education India, 2003.

Hao, Lili and Eric CY Ng, "Predicting Canadian recessions using dynamic probit modelling approaches," *Canadian Journal of Economics/Revue canadienne d'économique*, 2011, *44* (4), 1297–1330.

Honoré, Bo E and Ekaterini Kyriazidou, "Panel data discrete choice models with lagged dependent variables," *Econometrica*, 2000, *68* (4), 839–874.

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

Kolet, Ilan and Ryan Macdonald, "The fisher bcpi: The bank of canadaâs new commodity price index," Technical Report, Bank of Canada 2010.

Levanon, Gad, Jean-Claude Manini, Ataman Ozyildirim, Brian Schaitkin, and Jennelyn Tanchua, "Using financial indicators to predict turning points in the business cycle: The case of the leading economic index for the United States," *International Journal of Forecasting*, 2015, *31* (2), 426–445.

Lindsay, Bruce G, "Composite likelihood methods," *Contemporary Mathematics*, 1988, *80* (1), 221–239.

Liu, Weiling and Emanuel Moench, "What predicts US recessions?," *International Journal of Forecasting*, 2016, *32* (4), 1138–1150.

McFadden, Daniel, "L.(1973)â Conditional Logit Analysis of Qualitative Choice Behavior,â," *Frontiers in econometrics*, 1974, *2*.

Müller, Gernot and Claudia Czado, "An autoregressive ordered probit model with application to high-frequency financial data," *Journal of Computational and Graphical Statistics*, 2005, *14* (2), 320–338.

Newey, Whitney K and Daniel McFadden, "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, 1994, *4*, 2111–2245.

__ **and Kenneth D West**, "Automatic lag selection in covariance matrix estimation," *The Review of Economic Studies*, 1994, *61* (4), 631–653.

Ng, Chi Tim, Harry Joe, Dimitris Karlis, and Juxin Liu, "Composite likelihood for time series models with a latent autoregressive process," *Statistica Sinica*, 2011, pp. 279–305.

Ng, Serena, "Boosting recessions," *Canadian Journal of Economics/Revue canadienne d'économique*, 2014, *47* (1), 1–34.

Sahm, Claudia, "Direct stimulus payments to individuals," *Recession ready: Fiscal policies to stabilize the American economy*, 2019, pp. 67–92.

Schwarz, Gideon, "Estimating the dimension of a model," *The annals of statistics*, 1978, pp. 461–464.

Tuzcuoglu, Kerem, "Composite Likelihood Estimation of an Autoregressive Panel Ordered Probit Model with Random Effects," *Journal of Business & Economic Statistics*, 2023, *41* (2), 593–607.

Varin, Cristiano and Paolo Vidoni, "Pairwise likelihood inference for ordinal categorical time series," *Computational Statistics & Data Analysis*, 2006, *51* (4), 2365–2373.

__ and __ , "Pairwise likelihood inference for general state space models," *Econometric Reviews*, 2008, *28* (1–3), 170–185.

__ , Nancy Reid, and David Firth, "An overview of composite likelihood methods," *Statistica Sinica*, 2011, *21*, 5–42.

Wooldridge, Jeffrey M, "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity," *Journal of Applied Econometrics*, 2005, *20* (1), 39–54.

A Technical appendix

In this technical section, we provide the mathematical details for the asymptotic distribution of the MCL estimator. In particular, we compute the first and second derivatives of the MCL, derive the score and its long-run variance, calculate the Hessian matrix, and provide the asymptotic variance of the estimator. First, we introduce some notation: let $m_t \equiv m_t(\boldsymbol{\theta}) = m_t(\mathbf{x}, \boldsymbol{\theta}) = \mu/(1-\rho) + \sum_{k=0}^{t-1} \rho^k \boldsymbol{\beta}' \mathbf{x}_{t-m-k}$ such that $y_t^* | \mathbf{x} \sim \mathcal{N}(m_t, 1)$. We may suppress the dependency on $\boldsymbol{\theta}$ or $\mathbf{x}$ for notational simplicity throughout mathematical derivations.

Derivatives of $m_t(\boldsymbol{\theta})$

The derivatives of $m_t(\boldsymbol{\theta})$ are needed to calculate the asymptotic variance. Note that $m_t(\boldsymbol{\theta})$ can be also written as $m_t(\boldsymbol{\theta}) = \mu + \rho m_{t-1}(\boldsymbol{\theta}) + \boldsymbol{\beta}' \mathbf{x}_{t-m}$ with $m_0(\boldsymbol{\theta}) = \mu/(1-\rho)$. Then, the first derivative of $m_t(\boldsymbol{\theta})$ can be recursively computed as

$$\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \rho \frac{\partial m_{t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \begin{bmatrix} m_{t-1}(\boldsymbol{\theta}) \\ 1 \\ \mathbf{x}_{t-m} \end{bmatrix} \qquad \text{with} \qquad \frac{\partial m_0(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \mu/(1-\rho)^2 \\ 1/(1-\rho) \\ \mathbf{0} \end{bmatrix}.$$

The second derivative of $m_t(\boldsymbol{\theta})$ can be computed as

$$\frac{\partial^2 m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \rho \frac{\partial m_{t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} + \mathbf{e}_1 \frac{\partial m_{t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{\partial m_{t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbf{e}_1' \quad \text{with} \quad \frac{\partial m_0(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \begin{bmatrix} \frac{2\mu}{(1-\rho)^3} & \frac{1}{(1-\rho)^2} & \mathbf{0} \\ \frac{1}{(1-\rho)^2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $\mathbf{e}_1 = (1, 0, \ldots, 0)'$ is of dimension $(K+2) \times 1$.

Asymptotic distribution

Let us rewrite the marginal composite log-likelihood as

$$\mathcal{L}_{MCL}(\theta|\mathbf{y},\mathbf{x}) = \frac{1}{T}\sum_{t=1}^{T} \ln f(y_t|\mathbf{x};\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T} y_t \ln \boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right) + (1-y_t)\ln \boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right).$$

The score of the individual log-likelihood is $s_t(\boldsymbol{\theta}) = s(\boldsymbol{\theta}|y_t,\mathbf{x}) = \partial \ln f(y_t|\mathbf{x};\theta)/\partial\boldsymbol{\theta}$, where

$$\begin{aligned} s(\boldsymbol{\theta}|y_t,\mathbf{x}) &= y_t \frac{\partial \ln \boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} + (1-y_t)\frac{\partial \ln \boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} \\ &= y_t \frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right)} - (1-y_t)\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right)}. \end{aligned} \tag{7}$$

Note that $\phi(m_t(\boldsymbol{\theta})) = \phi(-m_t(\boldsymbol{\theta}))$. The hessian of the individual log-likelihood is $h_t(\boldsymbol{\theta}) = h_t(\boldsymbol{\theta}|y_t,\mathbf{x}) = \partial^2 \ln f(y_t|\mathbf{x};\theta)/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'$, where

$$\begin{aligned} h_t(\boldsymbol{\theta}|y_t,\mathbf{x}) &= y_t \frac{\partial^2 \ln \boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}\partial\boldsymbol{\theta}'} + (1-y_t)\frac{\partial^2 \ln \boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}\partial\boldsymbol{\theta}'} \\ &= y_t \frac{\partial}{\partial \boldsymbol{\theta}'}\left[\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right)}\right] - (1-y_t)\frac{\partial}{\partial \boldsymbol{\theta}'}\left[\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right)}\right] \\ &= y_t\left[\frac{\partial^2 m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial\boldsymbol{\theta}'} \frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right)} + \frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\left(\frac{-m_t(\boldsymbol{\theta})\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right)} + \frac{-\phi(m_t(\boldsymbol{\theta}))^2}{\boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right)^2}\right)\right] \\ &\quad - (1-y_t)\left[\frac{\partial^2 m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial\boldsymbol{\theta}'} \frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right)} + \frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\left(\frac{-m_t(\boldsymbol{\theta})\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right)} + \frac{\phi(m_t(\boldsymbol{\theta}))^2}{\boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right)^2}\right)\right] \\ &= y_t \frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right)}\left[\frac{\partial^2 m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial\boldsymbol{\theta}'} - \frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\left(m_t(\boldsymbol{\theta}) + \frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(m_t(\boldsymbol{\theta})\right)}\right)\right] \\ &\quad - (1-y_t)\frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right)}\left[\frac{\partial^2 m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial\boldsymbol{\theta}'} - \frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\left(m_t(\boldsymbol{\theta}) - \frac{\phi(m_t(\boldsymbol{\theta}))}{\boldsymbol{\Phi}\left(-m_t(\boldsymbol{\theta})\right)}\right)\right] \end{aligned} \tag{8}$$

One can easily show the uniform boundedness of the score and Hessian. First, by following the steps in the online appendix of Tuzcuoglu (2023), we can show that $|\partial m_t(\mathbf{x},\boldsymbol{\theta})/\partial\boldsymbol{\theta}|$ and $|\partial^2 m_t(\mathbf{x},\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'|$ are both uniformly bounded as long as $|\mathbf{x}| = |(\mathbf{x}_{-m+1},\ldots,\mathbf{x}_{T-m})'|$ is bounded, where $|\cdot|$ is the Euclidean norm. Second, note that the inverse Mills ratio $|\phi(\nu)/\Phi(\nu)|$ is bounded by a linear function of ν for any $\nu \in \boldsymbol{R}$. This implies that $|\phi(m_t(\mathbf{x},\boldsymbol{\theta}))/\Phi(m_t(\mathbf{x},\boldsymbol{\theta}))|$ is uniformly bounded as long as $|\mathbf{x}|$ is bounded. Finally, these two results imply that both the variance of the score and the Hessian matrix is uniformly

bounded by a linear function of the fourth moment of $\mathbf{x}$.

Under the assumptions given in Section 2 and following the asymptotic results of Newey and McFadden (1994) and Tuzcuoglu (2023), one can show that the score and Hessian satisfy the following asymptotics:

$$
\begin{aligned}
s(\boldsymbol{\theta}) &= \frac{1}{T}\sum_{t=1}^{T} s(\boldsymbol{\theta}|y_t,\mathbf{x}) = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial \ln f(y_t|\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \rightarrow_d \mathcal{N}\left(0,\boldsymbol{\Omega}(\boldsymbol{\theta})\right),\\
h(\boldsymbol{\theta}) &= \frac{1}{T}\sum_{t=1}^{T} h(\boldsymbol{\theta}|y_t,\mathbf{x}) \rightarrow_p \boldsymbol{\mathcal{H}}(\boldsymbol{\theta}),
\end{aligned}
$$

where $\boldsymbol{\mathcal{H}}(\boldsymbol{\theta})$ is the Hessian matrix and $\boldsymbol{\Omega}(\boldsymbol{\theta})$ is the long-run variance of the score that can be computed as $\boldsymbol{\Omega}(\boldsymbol{\theta}) = \boldsymbol{\Omega}_0(\boldsymbol{\theta}) + \sum_{l=1}^{\infty}\boldsymbol{\Omega}_l(\boldsymbol{\theta}) + \boldsymbol{\Omega}_l'(\boldsymbol{\theta})$ with $\boldsymbol{\Omega}_0(\boldsymbol{\theta}) = \text{Var}(s_t(\boldsymbol{\theta}))$ and $\boldsymbol{\Omega}_l(\boldsymbol{\theta}) = \text{Cov}(s_t(\boldsymbol{\theta}), s_{t-l}(\boldsymbol{\theta}))$. Then, for the true parameter $\boldsymbol{\theta}_*$, the asymptotic distribution of the MCL estimator can be given by

$$
\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) \rightarrow_d \mathcal{N}\left(0, \boldsymbol{\mathcal{H}}^{-1}(\boldsymbol{\theta}_*)\boldsymbol{\Omega}(\boldsymbol{\theta}_*)\boldsymbol{\mathcal{H}}^{-1}(\boldsymbol{\theta}_*)\right).
$$

An estimator for the Hessian matrix is

$$
\widehat{\boldsymbol{\mathcal{H}}}(\hat{\boldsymbol{\theta}}) = \frac{1}{T}\sum_{t=1}^{T}\left[y_t\frac{\partial^2 \ln \boldsymbol{\Phi}(m_t(\hat{\boldsymbol{\theta}}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} + (1-y_t)\frac{\partial^2 \ln \boldsymbol{\Phi}(-m_t(\hat{\boldsymbol{\theta}}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right],
$$

where the second derivatives are given in (8) in detail. Moreover, the derivatives of $m_t(\boldsymbol{\theta})$ are given at the end of this section. To obtain an estimator of the long-run matrix $\boldsymbol{\Omega}(\boldsymbol{\theta})$, we rely on Newey and West (1994) for the bandwidth selection and Gallant (1987) for the Parzen kernel weights. Then, an estimator of $\boldsymbol{\Omega}(\boldsymbol{\theta})$ can be given by

$$
\begin{aligned}
\hat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}}) &= \hat{\boldsymbol{\Omega}}_0(\hat{\boldsymbol{\theta}}) + \sum_{l=1}^{M_T}\hat{\boldsymbol{\Omega}}_l(\hat{\boldsymbol{\theta}}) + \hat{\boldsymbol{\Omega}}_l'(\hat{\boldsymbol{\theta}}),\\
&= \frac{1}{T}\sum_{t=1}^{T} s_t(\hat{\boldsymbol{\theta}})s_t'(\hat{\boldsymbol{\theta}}) + \frac{1}{T}\sum_{l=1}^{M_T}\kappa(l/M_T)\sum_{t=l+1}^{T} s_t(\hat{\boldsymbol{\theta}})s_{t-l}'(\hat{\boldsymbol{\theta}}) + s_{t-l}(\hat{\boldsymbol{\theta}})s_t'(\hat{\boldsymbol{\theta}}),
\end{aligned}
$$

where the bandwidth is $M_T = \lfloor 12(T/100)^{4/25} \rfloor$ and the Parzen kernel is

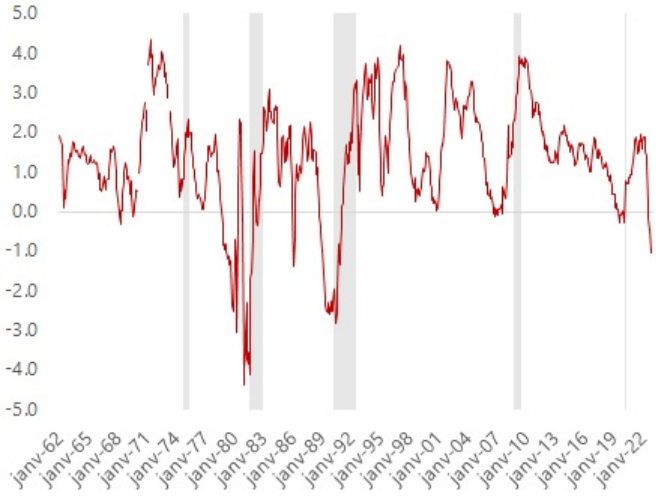
$$\kappa(l/M_T) = \begin{cases} 1 - 6|l/M_T|^2 + 6|l/M_T|^3 & \text{if } 0 \leq l/M_T \leq 0.5, \\ 2(1 - |l/M_T|)^3 & \text{if } 0.5 < l/M_T \leq 1. \end{cases}$$

Note that $\lfloor \cdot \rfloor$ denotes the floor function. This choice of bandwidth provides better asymptotic results – in terms of correct sizes of nominal t-tests – in the Monte Carlo simulations compared with the bandwidth choice of $\lfloor 4(T/100)^{2/9} \rfloor$ as in Kauppi and Saikkonen (2008) that results in smaller estimated standard errors, thus, higher over-rejection rates. The results are available upon request.

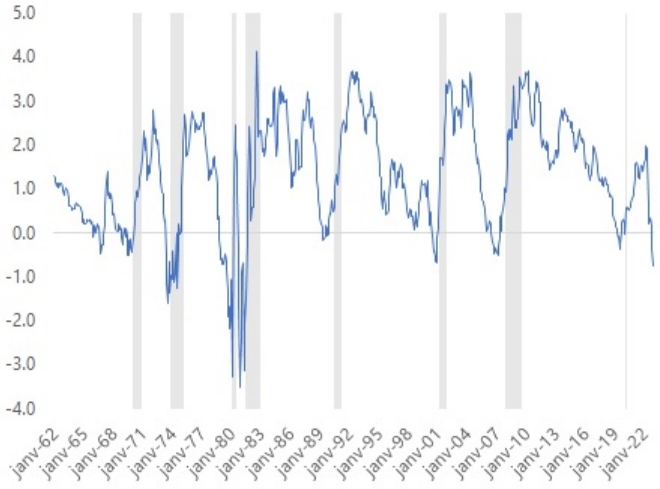
B Data appendix

B.1 Figures

Figure 2: Evolution of the yield curve in Canada and in the United-States over history



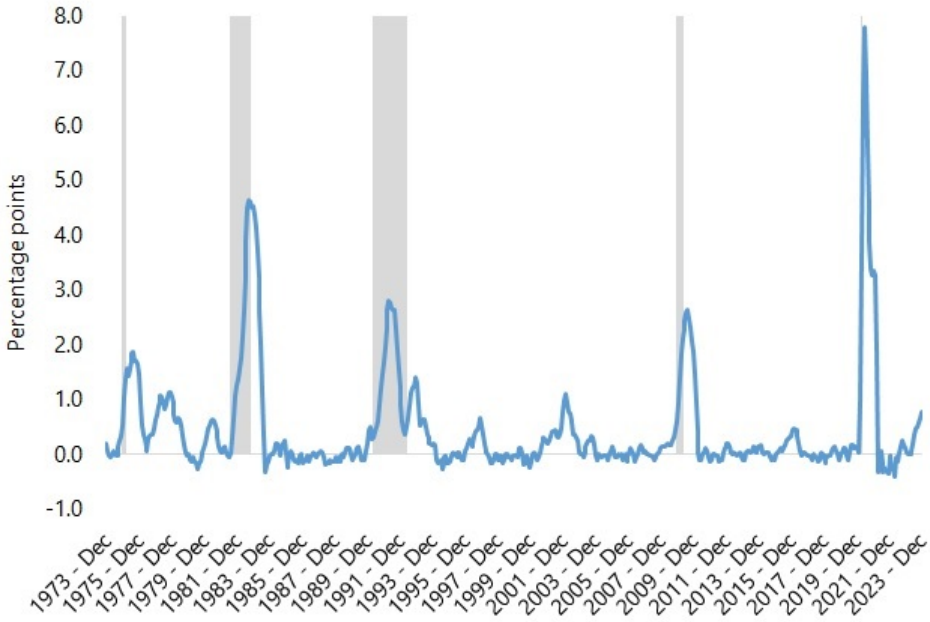
(a) Canada



(b) United-States

Notes: The bond yield spreads are shown in percentage points on the y-axis. The United-States spread reflects the difference between the 10-year and 3-month treasuries yields, while the Canadian spread is the difference between the 10-year and over government of Canada marketable bonds and the 3-month treasury bills yields. Recession dates are represented by the shaded grey area in Figures 2a and 2b. Those are defined by the C.D. Howe Institute in Canada and by the National Bureau of Economic Research (NBER) in the United-States.

Figure 3: Canadian version of the Sahm rule



Notes: the Canadian Sahm rule is consistent with the measurement proposed in Sahm (2019) (see Figure 2). This indicator represents the percentage points difference between the 3-month moving average of the unemployment rate and its prior year minimum. We use the seasonally adjusted unemployment rate from Statistics Canada Labour Force Survey to replicate this indicator for Canada.

B.2 Potential explanatory variables

Table 9: List of potential explanatory variables

Variable	Mnemonic	Description	Source	Period
Canadian Sahm rule	SR	Unemployment rate (3-month moving average) relative to prior 12-month low	Sahm (2019), Statistics Canada Labour Force Survey	Mar 1956 - Dec 2022
Housing starts	HS	Month-to-month log difference of housing starts	Canada Mortgage and Housing Corporation (CMHC)	Feb 1956 - Dec 2022
Consumer confidence	CCI	Month-to-month growth rate of consumer confidence	Conference Board of Canada and Bank of Canada calculations	Apr 1961 - Dec 2022
Railway Carloadings	RAIL	Month-to-month log difference of rail carloads, seasonally adjusted	Statistics Canada and Bank of Canada Calculations	Feb 1970 - Dec 2022
Building Permits	BP	Month-to-month log difference of building permits, seasonally adjusted	Statistics Canada and Bank of Canada Calculations	Feb 1948 - Dec 2022
Composite Leading Indicator	CLI	Month-to-month growth rate of the Composite Leading Indicator, seasonally adjusted, amplitude adjusted	OECD	Feb 1956 - Dec 2022
M1	M1	Month-to-month growth rate of money supply (M1++), CPI deflated, seasonally adjusted	Bank of Canada	February 1968 - Dec 2022
Exchange rate	CAD	Month-to-month growth rate of exchange rate (US$/CAD$)	Federal Reserve Board and Haver	Feb 1947 - Dec 2022
TSX Composite Index	TSX	Month-to-month growth rate of the TSX Composite Index, CPI deflated	S&P	Feb 1921 - Dec 2022
Bond yield spread	SP	Difference between the over 10 years government of Canada marketable bonds yield and 3-months treasury bills yield	Bank of Canada	Jan 1962 - Dec 2022

Variable	Mnemonic	Description	Source	Period
Drilling rigs	RIGS	Month-to-month log difference of drilling rigs count, seasonally adjusted	Baker Hughes and Bank of Canada Calculation	Feb 1968 - Dec 2022
Goods exports	EX	Month-to-month log difference of Canadian merchandise exports, seasonally adjusted	Statistics Canada and Bank of Canada Calculations	Feb 1968 - Dec 2022
Goods imports	IM	Month-to-month log difference of Canadian merchandise imports, seasonally adjusted	Statistics Canada and Bank of Canada Calculations	Feb 1968 - Dec 2022
BCPI - Total	BCPI	Month-to-month growth rate of the BCPI, CPI deflated	Bank of Canada	Feb 1972 - Dec 2022
BCPI - Crude oil	BCPI-O	Month-to-month growth rate of the crude oil BCPI, CPI deflated)	Bank of Canada	Feb 1972 - Dec 2022
BCPI - Natural gas	BCPI-G	Month-to-month growth rate of the natural gas BCPI, CPI deflated	Bank of Canada	Feb 1972 - Dec 2022
BCPI - Agriculture	BCPI-A	Month-to-month growth rate of the agriculture BCPI, CPI deflated	Bank of Canada	Feb 1972 - Dec 2022
BCPI - Metals	BCPI-M	Month-to-month growth rate of the metals BCPI, CPI deflated	Bank of Canada	Feb 1972 - Dec 2022
BCPI - Forestry	BCPI-F	Month-to-month growth rate of the forestry BCPI, CPI deflated	Bank of Canada	Feb 1972 - Dec 2022
BCPI - Energy	BCPI-E	Month-to-month growth rate of the energy BCPI, CPI deflated	Bank of Canada	Feb 1972 - Dec 2022
BCNE	BCNE	Month-to-month growth rate of the BCNE, CPI deflated	Bank of Canada	Feb 1972 - Dec 2022
CFNAI	CFNAI	Chicago Fed National Activity index, 3-months moving average	Federal Reserve Bank of Chicago	May 1967 - Dec 2022

Variable	Mnemonic	Description	Source	Period
US industrial production	USIP	Month-to-month growth rate of the industrial production index, seasonally adjusted	Board of Governors of the Federal Reserve System	Feb 1921 - Dec 2022
US employment	USE	Month-to-month growth rate of US employment, seasonally adjusted	Bureau of Labor Statistics	Feb 1948 - Dec 2022
World exports	WEX	Month-to-month growth rate of world export, CPI deflated, bil. US$	IMF	Feb 1960 - Dec 2022
US PMI	USPMI	Month-to-month growth rate of US Purchasing Manager Index, seasonally adjusted	ISM	Feb 1948 - Dec 2022

C Empirical analysis appendix

Table 10: In-sample performance of explanatory variables (AUROC)

		Forecast horizon (h)												
Order	Mnemonic	1	2	3	4	5	6	7	8	9	10	11	12	*Avg.*
1	CFNAI	0.874	0.864	0.864	0.853	0.838	0.809	0.785	0.760	0.755	0.744	0.741	0.727	0.801
2	SP	0.660	0.697	0.728	0.755	0.782	0.803	0.821	0.822	0.826	0.832	0.837	0.842	0.784
3	SR	0.854	0.823	0.797	0.771	0.744	0.713	0.682	0.653	0.627	0.615	0.614	0.614	0.709
4	CLI	0.602	0.649	0.706	0.740	0.754	0.756	0.747	0.728	0.708	0.691	0.680	0.672	0.703
5	USE	0.754	0.753	0.738	0.701	0.681	0.654	0.649	0.627	0.631	0.608	0.615	0.586	0.666
6	USIP	0.711	0.692	0.709	0.679	0.666	0.658	0.631	0.612	0.638	0.610	0.621	0.614	0.653
7	BCNE	0.650	0.655	0.662	0.654	0.677	0.683	0.672	0.667	0.635	0.611	0.610	0.597	0.648
8	TSX	0.651	0.661	0.693	0.671	0.671	0.684	0.668	0.654	0.618	0.614	0.593	0.584	0.647
9	BCPI-M	0.622	0.640	0.676	0.710	0.738	0.703	0.666	0.629	0.595	0.571	0.596	0.576	0.644
10	M1	0.615	0.641	0.647	0.627	0.643	0.648	0.625	0.610	0.611	0.595	0.614	0.600	0.623
11	BCPI	0.651	0.670	0.672	0.645	0.649	0.629	0.603	0.594	0.575	0.558	0.532	0.504	0.607
12	CCI	0.582	0.587	0.615	0.620	0.642	0.653	0.635	0.597	0.576	0.572	0.597	0.593	0.606
13	BCPI-A	0.609	0.620	0.597	0.581	0.577	0.614	0.648	0.662	0.629	0.608	0.560	0.531	0.603
14	BCPI-F	0.609	0.613	0.611	0.584	0.598	0.593	0.579	0.559	0.535	0.533	0.555	0.575	0.579
15	BCPI-O	0.620	0.645	0.644	0.612	0.610	0.574	0.535	0.521	0.513	0.505	0.493	0.544	0.568
16	RIGS	0.605	0.613	0.607	0.585	0.583	0.557	0.543	0.536	0.556	0.538	0.539	0.553	0.568
17	RAIL	0.586	0.613	0.574	0.572	0.589	0.584	0.566	0.536	0.556	0.543	0.535	0.552	0.567
18	BCPI-E	0.618	0.640	0.638	0.607	0.593	0.561	0.529	0.522	0.514	0.501	0.519	0.551	0.566
19	HS	0.616	0.602	0.584	0.582	0.597	0.565	0.535	0.541	0.517	0.496	0.521	0.503	0.555
20	IM	0.601	0.579	0.584	0.583	0.581	0.543	0.529	0.510	0.539	0.519	0.547	0.525	0.553
21	USPMI	0.555	0.560	0.567	0.567	0.577	0.554	0.570	0.476	0.533	0.537	0.554	0.533	0.549
22	EX	0.541	0.555	0.542	0.549	0.549	0.546	0.543	0.544	0.542	0.520	0.515	0.513	0.538
23	IM	0.558	0.571	0.594	0.563	0.543	0.520	0.494	0.494	0.509	0.539	0.502	0.505	0.533
24	BP	0.583	0.563	0.552	0.549	0.551	0.534	0.522	0.501	0.528	0.496	0.498	0.496	0.531
25	CAD	0.540	0.559	0.559	0.545	0.544	0.507	0.512	0.503	0.543	0.541	0.486	0.517	0.530
26	BCPI-G	0.578	0.591	0.568	0.553	0.525	0.511	0.509	0.519	0.508	0.484	0.467	0.471	0.524

Notes: this table shows the in-sample AUROC values calculated from single-variable static probit models, with lags varying between one and twelve-months ahead. The last column provides the average AUROC values across these forecast horizons. The models are estimated using MCL with data going from June 1973 to December 2022. The variables are sorted in descending order of performance. See Table 9 for the variable abbreviations.

Table 11: Models in-sample performance comparison ($h = 1$)

Model	AUROC	TPR	FPR	R2	QPS	RMSE
AR_{MCL}	98.7%	89.3%	3.1%	75.4%	4.3%	14.7%
ST	93.8%	66.1%	2.9%	42.9%	10.0%	22.4%
ST_{HN}	91.9%	71.4%	4.6%	44.1%	9.8%	22.1%
DYN_{HN}	98.4%	92.9%	1.0%	82.4%	3.1%	12.4%
AR_{HN}	98.9%	94.6%	2.9%	70.8%	5.1%	16.0%
$DYNAR_{HN}$	98.5%	92.9%	1.0%	82.4%	3.1%	12.4%

Notes: $R2$, QPS and $RMSE$ stand for the pseudo R^2, quadratic probability score and root-mean squared error, respectively. The true positive rates (TPR) and the false positive rates (FPR) are calculated at the 25% threshold. The first two rows reflect the performance of the AR and ST probit models presented in this paper. The last four rows reflect the performance of the static (ST_{HN}), dynamic (DYN_{HN}), autoregressive (AR_{HN}) and dynamic autoregressive ($DYNAR_{HN}$) probit models replicated from Hao and Ng (2011).

Table 12: Models OOS TPR and FPR comparison ($h = 1$)

Model	$TPR_{12.5}$	$FPR_{12.5}$	TPR_{25}	TPR_{25}	TPR_{50}	TPR_{50}
AR_{MCL}	91.4%	2.2%	88.6%	1.1%	71.4%	0.3%
ST	85.7%	8.1%	74.3%	3.2%	34.3%	0.5%
ST_{HN}	85.7%	8.1%	65.7%	1.4%	37.1%	0.8%
DYN_{HN}	77.1%	5.7%	68.6%	3.2%	37.1%	0.5%
AR_{HN}	91.4%	19.2%	71.4%	6.8%	48.6%	0.8%
$DYNAR_{HN}$	77.1%	5.4%	62.9%	4.1%	37.1%	0.5%

Notes: The true positive rates (TPR) and the false positive rates (FPR) are calculated at the 12.5%, 25% and 50% thresholds. The first two rows reflect the performance of the AR and ST probit models presented in this paper. The last four rows reflect the performance of the static (ST_{HN}), dynamic (DYN_{HN}), autoregressive (AR_{HN}) and dynamic autoregressive ($DYNAR_{HN}$) probit models replicated from Hao and Ng (2011).

Table 13: Turning points identification (OOS)

	50% threshold			*50% threshold*	
Peaks	AR_{MCL}	ST	Troughs	AR_{MCL}	ST
March 1990	-2	6	May 1992	-5	-11
October 2008	+1	3	May 2009	-1	-3
February 2020	+1	+1	April 2020	0	2

Notes: $(+a)$ or $(-a)$ imply that a model identifies a turning point *"a"* months too late or too early. A peak is identified too early (late) if the model's forecasted probability exceed a given threshold before (after) the actual peak. A trough is identified too early (late) if the model's forecasted probability fell below a given threshold before (after) the actual trough. The lags and leads in this table are assessed at the 50% threshold.

Table 14: Models in-sample performance comparison ($h = 1$, ex. COVID)

Model	AUROC	DIFF	R2	QPS	RMSE
AR_{MCL}	99.3%		77.8%	4.0%	14.1%
ST	97.5%	1.7%***	52.5%	8.5%	20.6%
ST_{HN}	92.6%	6.7%***	45.3%	9.8%	22.2%
DYN_{HN}	98.7%	0.6%	85.6%	2.6%	11.4%
AR_{HN}	99.2%	0.1%	73.7%	4.7%	15.4%
$DYNAR_{HN}$	98.8%	0.5%	85.7%	2.6%	11.3%

Notes: The second column shows the difference between the AR_{MCL} and the specified model AUROC values, along with its test of significance. *p<0.1; **p<0.05; ***p<0.01. $R2$, QPS and $RMSE$ stand for the pseudo R^2, quadratic probability score and root-mean squared error, respectively. The first two rows reflect the performance of the AR and ST probit models presented in this paper. The last four rows reflect the performance of the static (ST_{HN}), dynamic (DYN_{HN}), autoregressive (AR_{HN}) and dynamic autoregressive ($DYNAR_{HN}$) probit models replicated from Hao and Ng (2011).