



CIRANO
Allier savoir et décision

2016RP-08

Prévision de l'activité économique au Québec

Maxime Leroux, Rachidi Kotchoni, Dalibor Stevanovic

Rapport de projet/Project report

2016RP-08

Prévision de l'activité économique au Québec

Maxime Leroux, Rachidi Kotchoni, Dalibor Stevanovic

Rapport de projet *Project report*

Montréal

Date du rapport : Février 2016

Date de publication : Juin 2016

© 2016 Maxime Leroux, Rachidi Kotchoni, Dalibor Stevanovic. Tous droits réservés. *All rights reserved.*

Reproduction partielle permise avec citation du document source, incluant la notice ©.

Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source



Centre interuniversitaire de recherche en analyse des organisations

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act.

Les partenaires du CIRANO

Partenaires corporatifs

Autorité des marchés financiers
Banque de développement du Canada
Banque du Canada
Banque Laurentienne du Canada
Banque Nationale du Canada
Bell Canada
BMO Groupe financier
Caisse de dépôt et placement du Québec
Fédération des caisses Desjardins du Québec
Financière Sun Life, Québec
Gaz Métro
Hydro-Québec
Industrie Canada
Intact
Investissements PSP
Ministère de l'Économie, de l'Innovation et des Exportations
Ministère des Finances du Québec
Power Corporation du Canada
Rio Tinto
Ville de Montréal

Partenaires universitaires

École de technologie supérieure (ÉTS)
École Polytechnique de Montréal
HEC Montréal
Institut national de la recherche scientifique (INRS)
McGill University
Université Concordia
Université de Montréal
Université de Sherbrooke
Université du Québec
Université du Québec à Montréal
Université Laval

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web.

ISSN 1499-8629 (Version en ligne)

Prévision de l'activité économique au Québec

Maxime Leroux^{}, Rachidi Kotchoni[†], Dalibor Stevanovic[‡]*

Sommaire

Nous évaluons la prévisibilité de l'activité économique du Québec dans un environnement riche en données. Notre approche consiste à voir la province du Québec comme une économie régionale soumise aux influences des économies canadienne et américaine. Trois grands ensembles d'information sont utilisés : les données québécoises, canadiennes et américaines, soit un total de 453 variables macroéconomiques. Nous comparons un grand ensemble de modèles dans le but d'identifier ceux qui sont les plus efficaces pour prédire les principaux agrégats de l'économie québécoise tels que le PIB, l'emploi, l'inflation, l'investissement, etc. Nos résultats suggèrent que le meilleur modèle en termes d'erreur quadratique moyenne dépend de la série à prédire et de l'horizon de prévision visé. Un modèle ayant de bonnes performances à court horizon peut devenir moins bon à long horizon. Un modèle bon pour prédire le PIB nominal ne l'est pas forcément pour prédire le PIB réel. Dans la catégorie des modèles standards, le modèle ARMA(1,1) s'est révélé un bon benchmark pour prédire le PIB nominal ou l'inflation. Les modèles riches en données se classent souvent comme les meilleurs individuellement. La moyenne des prévisions fournies par une sélection des 5, 10 ou 20 meilleurs modèles individuels délivre des performances encore plus robustes.

Mots clés : Prévision, agrégats macroéconomiques, grandes bases de données, modèles à facteurs, modèles de séries chronologiques

Summary

We evaluate the predictability of the economic activity of Quebec in a data-rich environment. In our framework, the province of Quebec is treated as a regional economy that is exposed to the influence of the Canadian and American economies. Three large information sets are used: data from Quebec, Canada and the US, for a total of 453 macroeconomic variables. We compare a large number of models for the purpose of identifying those that are most efficient at forecasting macroeconomic aggregates of Quebec such as the GDP, employment, inflation, investment, etc. Our results suggest that the best model in terms of mean squared error depends on the variable of interest and on the forecasting horizon. A model that performs well at short horizons does not necessarily perform well at long horizons. Likewise, the model that best predicts the nominal GDP does not necessarily win the race when it comes to predict the real GDP. The ARMA(1,1) is found to be one of the best standard models to predict the nominal GDP and inflation. Models exploiting rich data sets often rank best individually. The most robust performances are obtained by averaging the forecasts delivered by the 5, 10 or 20 best individual models.

Keywords: Forecasting, Macroeconomic aggregates, Large data sets, Factor models, Time series models

Codes JEL/JEL Codes: C22, C32, C53, C55, E17

^{*} Département des sciences économiques, École des sciences de la gestion, Université du Québec à Montréal. 320, rue Sainte-Catherine Est, Montréal, QC, H2X 1L7.

[†] African School of Economics, Abomey-Calavi, BENIN.

[‡] Département des sciences économiques, École des sciences de la gestion, Université du Québec à Montréal. 320, rue Sainte-Catherine Est, Montréal, QC, H2X 1L7 (dstevanovic.econ@gmail.com).

1 Introduction

Au Québec et dans le monde développé, d'immenses quantités de données économiques sont générées chaque jour dans l'espoir que les chercheurs pourront s'en servir pour améliorer notre compréhension du fonctionnement et de la dynamique de l'économie. Malgré ce fait, un grand nombre d'exercices de prévision macroéconomique continue de se baser sur des ensembles d'information assez restreints. Dans le modèle le plus simple, on se sert du passé d'une variable macroéconomique pour prédire son évolution future. Souvent, il arrive qu'on élargisse l'ensemble d'information sur lequel se base la prévision en ajoutant des variables dites exogènes dans membre de droite. Dans un environnement riche en donnée, on peut être tenté d'ajouter un grand nombre de variables exogènes dans le but d'avoir une performance prédictive maximale. Cependant, cet envi se heurte très vite à des obstacles bien connus que sont le sur-apprentissage (autrement dit, "overfitting") et la malédiction de la dimensionnalité.

Le sur-apprentissage découle du fait qu'en échantillon fini, tout modèle tend à s'ajuster de mieux en mieux aux données d'apprentissage au fur et à mesure qu'on augmente le nombre de régresseurs et ce, que ces régresseurs soient pertinents ou non. Or un modèle trop ajusté aux données d'apprentissage a souvent du mal à prédire les données hors échantillon. Quant à la malédiction de la dimensionnalité, elle traduit le fait qu'un modèle devient de moins en moins attrayant au fur et à mesure que le nombre de variables augmente. Ce manque d'attrait se traduit sur deux plans. D'abord, sur le plan numérique où l'optimisation du modèle devient difficile, voire impossible compte tenu de la cardinalité de l'espace des paramètres. Ensuite, sur le plan économétrique où l'accumulation des imprécisions d'estimation d'un trop grand nombre de paramètres fini par déboucher sur des prévisions sous-optimales. Ces deux problèmes - le sur-apprentissage et la malédiction de la dimensionnalité - invitent l'économètre à un effort de parcimonie.

L'approche classique pour éluder les problèmes sus-cités consiste à réduire la dimensionnalité des données avant de procéder à l'estimation du modèle de prévision proprement dit. Les modèles à indices de diffusion proposés par Stock et Watson (2002) s'inscrivent dans cette logique. Ces auteurs utilisent l'analyse en composante principale (ACP) pour extraire un petit nombre de facteurs qui résument l'information contenue dans un grand nombre de variables macroéconomiques. Les facteurs d'ACP (nommés indices de diffusion dans ce contexte) sont ensuite ajoutés comme variables exogènes à des modèles autorégressifs parcimonieux. Les modèles à indices de diffusion ont été testés empiriquement et les résultats suggèrent que ceux-ci ont de meilleures performances prédictives que les modèles autorégressifs univariés. Ces modèles ont été utilisés dans la prévision des agrégats macroéconomiques nationaux (surtout américains, mais aussi de l'économie canadienne et de la zone Euro) ainsi qu'en finance. Cependant, leur performance a rarement été testée dans le cadre des économies régionales telles que le Québec.¹

1. À notre connaissance, Kopoin *et al.* (2013) sont les seuls auteurs à avoir exploré la question.

Dans le présent article, nous passons en revue un ensemble de modèles de prévision macroéconomiques dont plusieurs sont à la frontière de la recherche empirique. Notre objectif est d'identifier les modèles les plus efficaces pour prédire les principaux agrégats de l'économie québécoise tels que le PIB, l'emploi, l'inflation, l'investissement, etc. Nous considérons trois ensembles d'information potentiellement pertinents pour notre exercice de prévision : les données québécoises (notées X_t^{QC}), canadiennes (X_t^{CA}) et américaines (X_t^{US}). En effet, le Québec peut être considéré comme une économie régionale soumise à l'influence de ses géants voisins. L'ensemble de nos analyses porte exclusivement sur des données de fréquence trimestrielle.

Deux grands ensembles de modèles sont considérés : les modèles standards (AR, ARMA, VAR, etc.) et les modèles riches en données. Cette dernière catégorie de modèles s'obtient le plus souvent en ajoutant des indices de diffusion à un modèle standard. Cependant, d'autres modèles riches en données (comme le *Complete Subset Regression*) s'obtiennent en combinant un grand nombre de prévisions ponctuelles basées sur des sous-ensembles d'informations distincts. Différentes spécifications de modèles et leurs variantes, différentes approches d'extraction des indices de diffusion et différentes possibilités de combinaisons de prévisions sont croisées pour obtenir 119 modèles de prévision macroéconomiques. Chaque modèle est estimé pour chacune des séries macroéconomiques à prédire. Pour chaque variable à prédire et pour chaque horizon de prévision, le meilleur modèle ou la meilleure combinaison de modèles est identifié par le biais d'une évaluation de performance hors échantillon.

Nos résultats suggèrent que le meilleur modèle dépend de la série à prédire et de l'horizon de prévision visé. Un modèle ayant de bonnes performances à court horizon peut devenir moins bon à long horizon. Un modèle bon pour prédire le PIB nominal ne l'est pas forcément pour prédire le PIB réel. Dans la catégorie des modèles standard, le modèle ARMA(1,1) s'est révélé un bon benchmark pour prédire le PIB nominal ou l'inflation. Les modèles riches en données se classent souvent comme les meilleurs modèles individuels. La moyenne des prévisions fournies par une sélection des 5, 10 ou 20 meilleurs modèles individuels délivrent des performances encore plus robustes.

La suite du document s'articule comme suit. La section 2 présente en détail le design de la course de modèles. La section 3 décrit les données utilisées tandis que la section 4 décrit les modèles employés. La section 5 présente les analyses factorielles préliminaires. La section 6 commente les résultats de la course de modèle. La section 7 présente la conclusion de l'article. L'annexe 1 présente la liste complète des variables employées dans nos analyses. Enfin, les résultats détaillés pour toutes les variables d'intérêt sont présentés sous forme de tableaux dans le matériel supplémentaire.

2 Méthode d'évaluation des modèles

La méthode utilisée pour évaluer la performance prédictive des différents modèles consiste à faire un exercice historique de prévision hors échantillon. Tous nos modèles sont donc entraînés sur la base de données historiques complètement révisées. Par exemple, si le prévisionniste se trouve à la fin du quatrième trimestre de l'année 1999, nous faisons l'hypothèse que son ensemble d'information se termine effectivement en 1999Q4. On peut adapter notre approche pour tenir compte des retards de publication en supposant que l'ensemble d'information du prévisionniste se termine à la dernière donnée historique disponible en temps réel. Par contre, notre approche ne peut s'appliquer à l'exercice de prévision en temps réel proprement dit, où les dernières observations disponibles au prévisionniste seraient encore sujets à révisions.²

Avant de commencer, nous divisons l'échantillon disponible en deux parties comme illustré sur la figure 1. La partie la plus ancienne est l'échantillon initial d'apprentissage (ou d'estimation) sur lequel les modèles seront estimés. La partie la plus récente est l'échantillon de validation sur lequel les modèles seront évalués en vue de la course de modèle. Sur la figure 1 par exemple, la période initiale d'estimation commence au premier trimestre de l'année 1981 et se termine au dernier trimestre de l'an 1999. La période de validation quant à elle part du premier trimestre de l'an 2000 au dernier trimestre de 2013.³

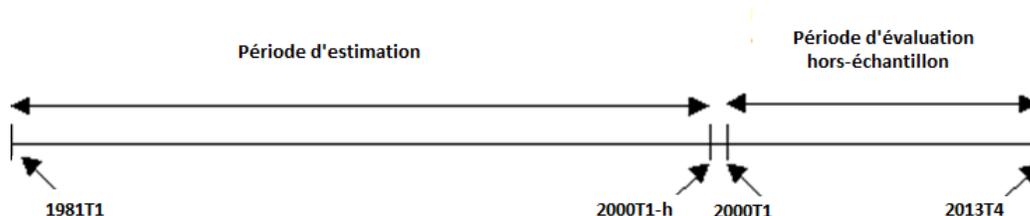


FIGURE 1 – Prévision hors-échantillon

L'exercice de course de modèles proprement dit se déroule comme suit. Premièrement, on estime tous les modèles sur la fenêtre initial d'apprentissage. Deuxièmement, on prédit la première observation de l'ensemble de validation selon chaque modèle. Troisièmement, on calcule et sauvegarde les erreurs des prévisions précédentes. Quatrièmement, on incrémente la fin de la fenêtre d'apprentissage d'une période et on répète les trois étapes précédentes, et

2. Nous ne disposons présentement pas de données en temps réel sur l'économie québécoise. La problématique de prévision en temps réel sera donc examinée dans nos recherches futures. Une autre possibilité serait d'incorporer l'information à plus haute fréquence, par exemple des indicateurs mensuels disponibles à l'intérieur du trimestre et avant que la donnée trimestrielle soit accessible. Cette approche de fréquences mixtes sera également examinée dans nos recherches futures.

3. Pour les modèles conçus pour prédire à un horizon $h > 1$, la fenêtre initial d'apprentissage s'arrête en réalité à $h-1$ trimestres avant 1999Q4, de sorte que la première observation à prédire se trouve en 2000Q1.

ainsi de suite jusqu'à ce que la dernière observation de l'ensemble de validation soit atteinte. À ce niveau, on a le choix entre une fenêtre croissante et une fenêtre glissante. Dans le premier cas, la fenêtre d'apprentissage s'élargit d'une période à chaque itération. Dans le second cas, la fenêtre d'apprentissage perd sa période la plus ancienne avant d'en gagner une nouvelle. Nous utiliserons la fenêtre glissante qui est plus robuste aux changements structurels plus fréquents. À la fin de ce processus, on dispose d'une série d'erreurs de prévision pouvant servir à évaluer les performances de chaque modèle.

Il est important de noter qu'à chaque itération, les modèles sont complètement recalibrés (nombre retards dans modèles autoregressifs, nombre de facteurs pour les modèles à diffusion d'indices, etc.). Une autre approche possible serait de spécifier un modèle pour chaque horizon donné, d'en estimer les paramètres une fois pour toute et de s'en servir tout au long de l'échantillon de validation. Cependant, nous privilégions la re-estimation à chaque itération afin de tenir compte des possibilités de ruptures structurelles.

Deux métriques sont utilisées pour mesurer la performance des modèles. La première métrique est l'erreur de prévision quadratique moyenne (*MSPE*) et est donnée par :

$$MSPE_{d,j} = \frac{\sum_{i=t^*}^T (\hat{y}_{d,j,i} - y_{d,i})^2}{T - t^*} \quad (1)$$

où $\hat{y}_{d,j}$ est la prévision de la variable numero d en utilisant le modèle j et y_d sa valeur réellement observée. La seconde métrique est le pseudo- R^2 qui se calcule comme suit :

$$pseudoR_{d,j}^2 = 1 - \frac{MSPE_{d,j}}{\overline{MSPE}_d} \quad (2)$$

où \overline{MSPE}_d est le *MSPE* du modèle ayant comme unique régresseur la variable constante. Ce dernier modèle équivaut à utiliser la moyenne de la variable sur l'échantillon d'apprentissage comme prévision.⁴

Le pseudo- R^2 n'est rien d'autre que l'analogie du R^2 calculé hors-échantillon. Les bons modèles devraient avoir des *MSPEs* faibles et des pseudo- R^2 élevés. Le pseudo- R^2 d'un modèle peut être négatif si les régresseurs inclus dans ce modèle n'ont pas de pouvoir prédictif réel sur la variable d'intérêt. À juste titre, cette métrique a été utilisée dans la littérature comme mesure de prévisibilité. Voir par exemple Galbraith (2003).

4. La moyenne historique se déplace alors en même temps que l'ensemble d'information utilisé pour les autres modèles.

3 Données

Les données utilisées pour ce travail proviennent de plusieurs sources : Statistic Canada (StatCan), Institut de Statistique du Québec (ISQ), Bloomberg et Federal Reserve Economic Data (FRED). Cette dernière base de données est alimentée et conservée par la Réserve fédérale de Saint-Louis. Nous disposons de 523 variables couvrant différents aspect de l'économie québécoise ainsi que de 620 variables canadiennes. Enfin, 157 variables américaines ont été récupéré de FRED en se basant largement sur l'ensemble de variables identifiés par Stock et Watson (2002) et mis à jour par McCracken et Ng (2014). Les données couvrent la période allant du premier trimestre de l'année 1981 jusqu'au quatrième trimestre de l'année 2013⁵.

Dans le cas des variables québécoises et canadiennes, un premier tri a été effectué pour retirer toutes les variables nominales lorsqu'une contrepartie réelle était disponible ainsi que pour conserver uniquement les variables désaisonnalisées lorsque la contrepartie avec saisonnalité était présente dans le lot d'origine⁶. Certaines variables ont également été supprimées étant donné leur comportement, par exemple les taux de taxes de vente et les taux de cotisation à la RRQ qui varient très peu à travers le temps et qui lorsqu'elles varient le font de manière très angulaire. Après cette sélection, il ne restait plus que 136 variables québécoises et 160 variables canadiennes.

La mise en oeuvre de l'ACP nécessite que les variables utilisées pour l'extraction des facteurs soient stationnaires. Cependant, il serait fastidieux d'effectuer des tests de stationnarité pour les 453 variables sélectionnées. Pour éluder le problème, nous avons adopté les transformations identifiées par McCracken et Ng pour des variables américaines et les avons appliqué aux variables semblables de notre étude. Pour les autres variables, nous avons inspecté visuellement les séries et utilisé notre jugement pour déterminer les transformations appropriées en vue d'obtenir des variables stationnaires. L'Annexe 1 présente la liste de toutes les variables ainsi que les transformations appliquées pour leur stationnarisation.

Enfin, nous avons ciblé 30 variables pour lesquelles nous souhaitons connaître la performance prévisionnelle de nos modèles. Ces 30 variables sont listées dans le Tableau 1.

4 Modèles

Dans cette section, nous énumérons tous les modèles utilisés dans l'étude. La variable à prédire est spécifiée comme suit :

5. Les données utilisées sont celles disponibles avant les révisions statistiques publiées le 10 novembre 2015 par Statistique Canada.

6. Les variables américaines étaient déjà triées.

Tableau 1 – Variables à prévoir

Nom	Transformation
Total indice des prix à la consommation du Québec	5
Demande intérieure finale, en \$ enchaînés	5
Produit intérieur brut aux prix du marché (millions de \$ enchaînés de 2007)	5
Dépenses personnelles en biens et services - \$ enchaînés	5
Investissements non résidentiels des entreprises, en \$ enchaînés	5
Formation brute de capital fixe Machines et matériel, en \$ enchaînés	5
Formation brute de capital fixe en produits de propriété intellectuelle, en \$ enchaînés	5
Formation brute de capital fixe des entreprises Construction, en \$ enchaînés	5
Construction domiciliaire totale, en \$ enchaînés	5
Exportations internationales de biens, en \$ enchaînés	5
Exportations internationales, en \$ enchaînés	5
Exportations interprovinciales, en \$ enchaînés	5
Importations internationales, en \$ enchaînés	5
Importations interprovinciales, en \$ enchaînés	5
Produit intérieur brut aux prix du marché (millions de \$)	5
Dépenses personnelles en biens et services, en \$ courants	5
Dépenses publiques en biens et services	5
Formation brute de capital fixe entreprises résidentiel	5
Investissements non résidentiels des entreprises, courants	5
Exportations totales en dollars courants	5
Exportations interprovinciales, en dollars courants	5
Importations internationales, en dollars courants	5
Importations interprovinciales, en dollars courants	5
Demande intérieure finale, nominal	5
Salaires et traitements	5
Excédent d'exploitation des sociétés au net, en dollars courants	5
Revenu personnel disponible	5
Emploi	5
Taux de chômage (%)	2
Dépenses publiques en biens et services, en \$ enchaînés	5

$$y_{t+h}^{(h)} = \frac{400}{h}(Y_{t+h} - Y_t), \quad (3)$$

où Y_t est la variable d'intérêt (possiblement en logarithme) et h est l'horizon de prévision. Par exemple, si Y_t est logarithme du PIB, la quantité de prévision, $y_{t+h}^{(h)}$ est le taux de croissance annualisé moyen sur h trimestres.

4.1 Modèles standards

Les 3 premiers modèles sont univariés et utilisent uniquement le passé d'une variable pour prédire son évolution future.

ARD Le premier est le modèle de référence pour la comparaison des performances : le modèle autorégressif direct d'ordre p (i.e., $ARD(p)$) qui est donné par :

$$y_{t+h}^{(h)} = \alpha^{(h)} + \rho(L)^{(h)}y_t + \epsilon_{t+h}, \quad t = 1, \dots, T, \quad (4)$$

où le régresseur est défini comme suit :

$$y_t = Y_t - Y_{t-1}. \quad (5)$$

Nous suivons Stock et Watson (2002) et McCracken et Ng (2014) dans ce choix de modélisation. La prévision à l’horizon h s’obtient facilement comme suit :

$$\hat{y}_{T+h|T}^{(h)} = \hat{\alpha}^{(h)} + \hat{\rho}(L)^{(h)} y_T. \quad (6)$$

ARI La version *itérative* du modèle autorégressif est également considérée :

$$\begin{aligned} y_{t+1} &= \alpha + \rho(L)y_t + \epsilon_{t+1}, \quad t = 1, \dots, T \\ \hat{y}_{T+h|T} &= \hat{\alpha} + \hat{\rho}(L)\hat{y}_{T+h-1|T} \end{aligned} \quad (7)$$

où la première équation de (7) est le modèle à estimer et la deuxième l’équation de la prévision. Notons que les deux modèles, ARD et ARI, donnent exactement la même prévision pour $h = 1$ et diffèrent uniquement lorsque $h > 1$. Dans ce dernier cas, la méthode de calcul de $\hat{y}_{T+h|T}^{(h)}$ n’est pas la même selon que l’on manipule un modèle de prévision direct ou un modèle de prévision itératif. Dans le modèle direct, $\hat{y}_{T+h|T}^{(h)}$ est l’output immédiat de la formule de prévision. En revanche, l’output immédiat du modèle itératif est $\hat{y}_{T+1|T}^{(1)}$, qui est identique à $\hat{y}_{T+1|T}$ selon nos notations. Il est donc nécessaire d’utiliser $\hat{y}_{T+1|T}$ pour prédire $\hat{y}_{T+2|T}$ et ainsi de suite, jusqu’à obtenir $\hat{y}_{T+h|T}$. Enfin, on construit $\hat{y}_{T+h|T}^{(h)}$ à partir de $\hat{y}_{T+h|T}$ selon la formule 3.

Il n’est pas clairement établi dans la littérature que l’une des deux spécifications domine l’autre. Ceci explique pourquoi nous considérons les deux modèles.⁷ Les ordres autorégressifs, p , sont déterminés par le critère d’information bayésien (BIC) pour les deux modèles, comme dans Stock et Watson (2002). Le BIC est privilégié dans notre étude car il a tendance à sélectionner des modèles plus parcimonieux que le AIC (Akaike Information Criterion).

ARMA Dans la catégorie des modèles univariés, le modèle autorégressif moyenne mobile (ARMA(1,1)) est également considéré :

$$\begin{aligned} y_{t+1} &= \alpha + \rho y_t + \Theta \epsilon_t + \epsilon_{t+1}, \quad t = 1, \dots, T \\ \hat{y}_{T+h|T} &= \hat{\alpha} + \hat{\rho} \hat{y}_{T+h-1|T} + \hat{\Theta} \hat{\epsilon}_{T+h-1|T} \end{aligned} \quad (8)$$

Ce modèle est employé dans l’optique d’une prévision itérative.

7. Marcellino *et al.* (2006) et Chevillon (2007) trouvent que le modèle direct est légèrement plus performant que le modèle itératif. En théorie, le modèle ARI serait plus efficace si la structure autorégressive ne change pas dans le temps, tandis que le modèle ARD serait plus robuste aux changements structurels.

Les deux modèles suivants incorporent davantage d'information que les précédents, mais ne font toujours pas partie de la catégorie des modèles riches en données. Il s'agit du modèle autorégressif à retards échelonnés (ADL(p)) et de sa contrepartie en prévision itérative, le modèle autorégressif vectoriel (VAR(p)).

ADL Dans le cas du ADL(p) on augmente le modèle (4) des retards d'autres variables pertinentes pour la prévision de la variable d'intérêt.

$$\begin{aligned} y_{t+h}^{(h)} &= \alpha^{(h)} + \rho(L)^{(h)} y_t + \beta(L)^{(h)} X_t + \epsilon_{t+h}, \quad t = 1, \dots, T \\ \hat{y}_{T+h|T}^{(h)} &= \hat{\alpha}^{(h)} + \hat{\rho}(L)^{(h)} y_T + \hat{\beta}(L)^{(h)} X_T \end{aligned} \quad (9)$$

où X_t contient le taux de chômage québécois, le taux d'intérêt des bons du trésor canadien à échéance de 3 mois, le taux de change entre le Canada et les États-Unis, une mesure du PIB canadien ou américain et le TSX ou le S&P500.⁸

VAR Le VAR(p) utilise les mêmes variables, mais dans ce cas-ci elles sont toutes prédites simultanément.

$$\begin{aligned} Z_{t+1} &= C + \Psi(L)Z_t + E_{t+1}, \quad t = 1, \dots, T \\ \hat{Z}_{T+h|T} &= \hat{C} + \hat{\Psi}(L)Z_{T+h-1|T} \end{aligned} \quad (10)$$

où $Z_t = [y_t' X_t']'$ et $\hat{y}_{T+h|T}$ est le premier élément de $\hat{Z}_{T+h-1|T}$. Dans les deux modèles, les ordres autorégressifs sont déterminés par BIC. La principale différence entre modèles ADL et VAR est que le premier est spécifié sous forme directe tandis que le deuxième est spécifié sous forme itérative.

4.2 Modèles riches en données

Les modèles présentés dans la suite sont tous considérés comme étant riches en données même si tous ne se basent pas sur la même méthodologie. La plupart s'obtiennent en ajoutant des indices de diffusion à des modèles standard. Certains sont obtenus par combinaison d'un grand nombre de prévisions basées sur des ensembles d'information différents.

8. Il s'agit donc de deux versions, l'une avec de l'information supplémentaire sur la conjoncture économique canadienne et l'autre sur celle des États-Unis.

ARDI Notre premier modèle riche en données est le ARDI, qui n'est autre que le ARD (4) augmenté d'un petit nombre de facteurs non directement observables, à l'instar des modèles à indices de diffusions de Stock et Watson (2002). Ceci suppose donc une première étape où l'on extrait les facteurs latents à partir d'un grand ensemble de données dont l'on souhaite réduire la dimensionalité. Pour l'instant, on suppose que les facteurs sont statiques. L'équation qui relie les facteurs aux variables peut être représentée comme suit :

$$X_t = \Lambda F_t + u_t \quad (11)$$

où X_t est un ensemble de N variables macroéconomiques, F_t est un ensemble de k facteurs, avec k beaucoup plus petit que N (d'où la réduction de la dimensionalité), Λ est la matrice d'exposition des variables aux facteurs et u_t est une matrice de mêmes dimensions que X_t contenant les composantes idiosyncratiques.

Dans la représentation ci-dessus, les facteurs F_t ne sont pas identifiables de manière unique sans restrictions supplémentaires. Pour y remédier, on fera l'hypothèse que les F_t sont des facteurs d'ACP. Ceci nous amène à postuler :

$$F = XU \quad (12)$$

où X contient les X_t pour toutes les périodes et donc est de taille (T, N) , U est une matrice de taille (N, k) qui contient les contributions des variables à la construction des facteurs et F est la matrice de taille (T, k) qui contient les facteurs.

Ensuite, on résout :

$$\begin{aligned} \max F'F &= U'X'XU \\ \text{s.c. } U'U &= I_k \end{aligned} \quad (13)$$

La résolution de ce problème débouche sur le fait que les colonnes de U représentent les k vecteurs propres de la matrice $X'X$ associé aux k plus grandes valeurs propres.⁹ Notons que dans une vision structurelle, les facteurs d'ACP que nous estimons ne représentent qu'une rotation des vrais facteurs de la représentation 14. Ceci rend l'interprétation des vrais facteurs impossibles. Cependant, on peut examiner les variables auxquelles les facteurs d'ACP sont le plus corrélés pour tenter de leur attribuer un sens économique.

9. Même si les formules ne le montre pas explicitement, il est important de standardiser les variables (par centrage et réduction) avant de procéder à l'extraction des facteurs. Ceci amène à voir les facteurs d'ACP comme ceux qui maximisent la variance expliquée.

Une fois les facteurs d'ACP extraits, on les ajoute au modèle autorégressif direct sous forme de variables exogènes. La prévision se calcule simplement par :

$$\begin{aligned} y_{t+h}^{(h)} &= \alpha^{(h)} + \rho(L)^{(h)}y_t + \beta(L)^{(h)}F_t + \epsilon_{t+h}, \quad t = 1, \dots, T \\ \hat{y}_{T+h|T}^{(h)} &= \hat{\alpha}^{(h)} + \hat{\rho}(L)^{(h)}y_T + \hat{\beta}(L)^{(h)}\hat{F}_T \end{aligned} \quad (14)$$

Le critère d'information BIC est utilisé afin de déterminer simultanément le nombre de retards autorégressif (donc, l'ordre du polynôme $\rho(L)^{(h)}$), le nombre de facteurs dans \hat{F}_t ainsi que le nombre de retards des facteurs (l'ordre des polynômes $\beta(L)^{(h)}$).

Plusieurs versions du modèle à indices de diffusion sont également considérées. Tout d'abord, il y a deux versions où des sous-ensembles de X_t sont choisis en fonction du pouvoir prédictif des variables prises individuellement. En effet, le modèle AR-DI de base considère *de facto* que toutes les variables utilisées dans l'extraction des facteurs sont pertinentes pour la prévision de la variable d'intérêt. Or, Boivin et Ng (2006) ont montré que ceci n'est pas toujours le cas. Nous utilisons donc deux méthodes développées par Bai et Ng (2008), l'un dit à seuil rigide et l'autre à seuil souple, pour sélectionner les variables qui contribueront à l'extraction des facteurs.

ARDI-hard La méthode à seuil rigide consiste à conserver uniquement les variables ayant un lien significatif avec la variable d'intérêt dans le modèle suivant :

$$y_{t+h}^{(h)} = \alpha^{(h)} + \sum_{j=0}^3 \rho_j^{(h)} y_{t-j} + \beta_i^{(h)} X_{i,t} + \epsilon_t \quad (15)$$

Seuls les variables pour lesquelles la statistique t de Student associée à β_i est plus grande qu'un certain seuil t_c sont sélectionnées¹⁰ :

$$X_t^* = \{X_i \in X_t | t_{X_i} > t_c\} \quad (16)$$

ARDI-soft La méthode à seuil ne se base pas sur une borne arbitraire comme dans le cas précédent, mais plutôt sur la régression LASSO¹¹ suivante :

10. Deux seuils sont utilisés, 1,28 et 1,65. À l'instar de Bai et Ng (2008), nous utilisons 3 retards dans 15

11. *Least Absolute Shrinkage and Selection Operator*.

$$\hat{\beta}^{lasso} = \arg \min_{\beta} \left[\sum_{i=1}^T (M_{ly} y_t^{(h)} - \sum_j \beta_j M_{ly} X_{j,t-h})^2 + \lambda \sum_j |\beta_j| \right] \quad (17)$$

où M_{ly} est la matrice de projection sur l'espace orthogonale à une constante et y_{t-i} avec $i = 0, 1, 2, 3$. Le LASSO est une méthode de sélection de variables forçant les coefficients les moins importants à converger vers 0. Ceci permet de ne conserver que les variables les plus pertinentes dans la formule de prévision.

$$X_t^* = \{X_i \in X_t \mid \beta_{X_i}^{lasso} \neq 0\} \quad (18)$$

En principe le nombre de variables à conserver pourrait être déterminé de façon optimale, mais une telle procédure serait extrêmement longue. Pour contourner ce problème, le nombre de variables est fixé à 30¹² comme le font Bai et Ng (2008)¹³. Après sélection des variables, les facteurs d'ACP sont estimés en utilisant le sous-ensemble X_t^* au lieu de X_t . Enfin, on estime le modèle à indices de diffusion de la même manière que le modèle AR-DI de base.

ARDI-tstat Dans le modèle AR-DI, le nombre de facteurs k est déterminé par un critère d'information. Mais cette façon de faire impose que les k premiers facteurs sont tous pertinents pour la prévision. Ceci peut être problématique si, par exemple, la variable à prévoir n'a aucun lien avec un sous-groupe \tilde{X}_t et que le premier facteur est très lié à ce sous-groupe. Pour contourner ce problème nous appliquons l'idée du ciblage rigide, mais cette fois aux facteurs estimés. De plus, nous limitons la structure du modèle en imposant qu'il ne contient pas de retard des facteurs retenus.¹⁴

$$\begin{aligned} y_{t+h}^{(h)} &= \alpha^{(h)} + \sum_{j=0}^3 \rho_j^{(h)} y_{t-j} + \sum_{i=1}^K \beta_i^{(h)} \hat{F}_{i,t} + \epsilon_{t+h} \\ K^* &= \{i \in 1, \dots, K \mid t_i > t_c\} \\ \hat{y}_{T+h|T}^{(h)} &= \hat{\alpha}^{(h)} + \hat{\rho}(L)^{(h)} y_T + \sum_{i \in K^*} \hat{\beta}_i(L)^{(h)} \hat{F}_{i,T} \end{aligned} \quad (19)$$

12. En réalité, la procédure ne mène pas nécessairement à 30 exactement puisqu'il peut arriver qu'au point où 31 variables sont conservées, l'itération suivante élimine 2 variables très corrélées entre-elles. La procédure trouve plutôt le nombre qui s'en rapproche le plus.

13. Les auteurs trouvent que le nombre optimal est généralement dans les alentours de 30 et que la performance prévisionnelle n'est pas affectée en fixant le nombre de variables à 30.

14. L'inférence sur les facteurs estimés est en principe sujette au problème de régresseurs générés ; mais nous supposons que la théorie asymptotique de Bai et Ng (2006) s'applique puisque nos ensembles d'information contiennent un très grand nombre de variables.

ARDI-Y1 Dans le modèle ARDI, le BIC est utilisé afin de déterminer simultanément le nombre de retards autorégressif, donc l'ordre du polynôme $\rho(L)^{(h)}$, le nombre de retards des facteurs (l'ordre des polynômes $\beta(L)^{(h)}$) ainsi que le nombre de facteurs dans \hat{F}_t . Afin de réduire la lourdeur de la procédure, nous pouvons restreindre le modèle en imposant seulement un retard autoregressif : $\rho(L)^h = \rho^h$.

ARDI-same Dans le modèle ARDI de base, le BIC détermine le nombre de retards pour chaque facteur individuellement. En effet, les facteurs proviennent de différents ensembles de données : québécois (QC), canadien (CA) et américain (US). Dans une spécification nous estimons les facteurs séparément pour chacun des ensembles et BIC détermine simultanément le nombre de retards pour chaque sous-groupe de facteurs. Ici, nous restreignons ce nombre à être le même pour tous les facteurs, indépendamment de quel ensemble d'information ils sont calculés.

Les différences apportées dans les modèles (ARDI-tstat), (ARDI-Y1) et (ARDI-same) semblent marginales, mais il serait intéressant de voir si elles sont empiriquement significatives, i.e. si la flexibilité du modèle ARDI domine les restrictions *ad hoc* dans les autres spécifications.

ARDI-DU Un dernier modèle à indice de diffusion est considéré, mais celui-ci diffère dans l'estimation des facteurs latents. Au lieu des facteurs *statiques* ce sont plutôt des facteurs *dynamiques* qui sont estimés comme dans Forni *et al.* (2005). Ils sont estimés comme composantes principales généralisées mais avec un filtre "one-sided" afin de les rendre utiles pour la prévision. Tout le reste est fait comme dans le modèle ARDI.

Jusqu'à présent, la représentation factorielle de X_t a servi seulement à réduire la dimension du problème : au lieu de considérer X_t comme prédicteurs, on les remplace par un petit nombre de facteurs. Par contre, la structure vectorielle de X_t peut aussi être mise à contribution. Ce type de modèle est populairement appelé FAVAR (pour Factor-Augmented VAR). Boivin et Ng (2005) identifient 2 façons de faire de la prévision en utilisant la représentation suivante :

$$X_t = \Lambda F_t + u_t \quad (20)$$

$$u_t = D(L)u_{t-1} + v_t \quad (21)$$

$$F_t = A(L)F_{t-1} + e_t \quad (22)$$

où $D(L) = \text{diag}(\delta_i(L))$, $i = 1, \dots, N$, est un polynôme matriciel fini, v_t et e_t sont des bruits blancs noncorrélés à tous les horizons. Les processus autorégressifs (21) et (22) sont supposés stationnaires.

En utilisant cette spécification, on peut construire la prévision de toutes les variables incluses dans X_t . On commence d'abord par faire la prévision de $F_{T+h|T}$ et de $u_{T+h|T}$. Ensuite on se sert de $\hat{\Lambda}$ pour construire la prévision de X_t . Cela nécessite de spécifier des dynamiques pour

les facteurs, (22), et pour la partie idiosyncratique (21). Si la structure à facteurs imposée est celle qui a réellement généré les données, le FAVAR est susceptible de produire de meilleures prévisions que les modèles précédents. Ici la prévision des deux composantes peut soit se faire de manière itérative ou bien de façon directe.

FAVAR-I Dans le cas itératif, la prévision complète est la suivante :

$$\begin{aligned}\hat{F}_{T+h|T} &= \hat{A}(L)\hat{F}_{T+h-1|T} \\ \hat{u}_{T+h|T,i} &= \hat{\delta}_i(L)\hat{u}_{T+h-1|T,i} \\ \hat{X}_{T+h|T,i} &= \hat{\Lambda}_i\hat{F}_{T+h|T} + \hat{u}_{T+h|T,i}\end{aligned}\tag{23}$$

où la prévision des facteurs latents se fait en utilisant un VAR(p) et celle de la composante idiosyncratique en utilisant un AR(p) itératif.

FAVAR-D Dans le cas direct, il s'agit d'utiliser les versions directes des modèles itératifs :

$$\begin{aligned}\hat{F}_{T+h|T} &= \hat{A}(L)^{(h)}\hat{F}_T \\ \hat{u}_{T+h|T,i} &= \hat{\delta}_i(L)^{(h)}\hat{u}_{T,i} \\ \hat{X}_{T+h|T,i} &= \hat{\Lambda}_i\hat{F}_{T+h|T} + \hat{u}_{T+h|T,i},\end{aligned}\tag{24}$$

donc d'utiliser un VAR(p) direct pour la prévision des facteurs et un AR(p) direct pour la prévision de la composante idiosyncratique.

FAVARMA Dans la même veine que le FAVAR itératif, on considère le FAVARMA de Dufour et Stevanović (2013) où une composante moyenne mobile vient s'ajouter à l'équation prédictive des facteurs ¹⁵.

$$\begin{aligned}\hat{F}_{T+h|T} &= \hat{A}(L)\hat{F}_{T+h-1|T} + \hat{\beta}(L)\hat{e}_{T+h-1|T} \\ \hat{u}_{T+h|T,i} &= \hat{\delta}_i(L)\hat{u}_{T+h-1|T,i} \\ \hat{X}_{T+h|T,i} &= \hat{\Lambda}_i\hat{F}_{T+h|T} + \hat{u}_{T+h|T,i}\end{aligned}\tag{25}$$

15. Le nombre de facteurs pour les FAVARs et le FAVARMA est choisi avec le critère de Bai et Ng (2002).

DFM Nous poursuivons avec le modèle à facteurs dynamiques de Forni *et al.* (2005) où une prévision non paramétrique de la composante commune est utilisée au lieu de prévoir des facteurs. Le nombre de facteurs dynamiques est choisi avec le test de Hallin et Liška (2007).

$$\begin{aligned}\hat{u}_{T+h|T,i} &= \hat{\delta}_i(L)\hat{u}_{T+h-1|T,i} \\ \hat{X}_{T+h|T} &= \hat{\chi}_{T+h|T} + \hat{u}_{T+h|T,i}\end{aligned}\tag{26}$$

Notons que pour tous les modèles où l'on fait la prévision de $X_{T+h|T}$, il est nécessaire de calculer la prévision finale de la même manière que dans les modèles itératifs. En effet, X_t ne contient pas $y_t^{(h)}$, mais plutôt y_t .

Les deux derniers modèles sont toujours dans la catégorie des modèles riches en données, mais diffèrent quant à la manière de résumer l'information contenue dans le panel de variables macroéconomiques.

CSR Le premier est le *Complete subset regression* (CSR) d'abord présenté et utilisé dans un contexte peu riche en données¹⁶ par Elliott *et al.* (2013) puis utilisé dans le contexte plus riche par Elliott *et al.* (2015). Ce modèle consiste à estimer un très grand nombre de modèles de prévision utilisant L régresseurs et un retard de la variable à prévoir, puis à combiner toutes les prévisions en prenant simplement leur moyenne. On a donc :

$$\begin{aligned}\hat{y}_{T+h|T,m}^{(h)} &= \hat{c} + \hat{\rho}y_t + \hat{\beta}X_{t,m} \\ \hat{y}_{T+h|T}^{(h)} &= \frac{\sum_{m=1}^M \hat{y}_{T+h|T,m}^{(h)}}{M}\end{aligned}\tag{27}$$

où les L variables comprises dans X_t sont choisies aléatoirement et M est le nombre maximal de combinaison de variables utilisées. 3 valeurs de L sont utilisées ; 1, 10 et 20¹⁷.

3PRF Le dernier modèle considéré est le filtre régressif triple de Kelly et Pruitt (2015) (Three-pass regression filter (3PRF)). Les étapes sont les suivantes :

1. Régression temporelle de X_i sur Z_t pour $i = 1, \dots, N$

$$X_{i,t} = \phi_{0,i} + Z_t' \phi_i + \varepsilon_{i,t}\tag{28}$$

16. Le panel utilisé dans l'application empirique ne contient que 12 variables

17. Le nombre de combinaisons possibles augmente très rapidement avec le nombre de variables dans X_t .

2. Régression transversale de X_t sur $\hat{\phi}_i$ pour $t = 1, \dots, T$

$$X_{i,t} = \varsigma_{0,t} + \hat{\phi}'_i f_t + \epsilon_{i,t} \quad (29)$$

3. Régression temporelle de $y_{t+h}^{(h)}$ sur \hat{f}_t

$$y_{t+h}^{(h)} = \beta_0 + \beta \hat{f}_t + \eta_{t+h} \quad (30)$$

4. Préviation

$$\hat{y}_{T+h|T}^{(h)} = \hat{\beta}_0 + \hat{\beta} \hat{f}_T \quad (31)$$

Ici Z_t contient ce que Kelly et Pruitt appellent des proxies. Les proxies sont des variables qui sont pertinentes pour la prévision de la variable d'intérêt. Quatre retards de y_t sont utilisés comme proxies pour la prévision de $y_{t+h}^{(h)}$.

4.3 Combinaison des panels

Rappelons que nous avons 3 panels de données : QC, CA et US. La plupart des modèles riches en données peuvent utiliser les panels macroéconomiques de deux façons différentes : (i) en les combinant pour n'en former qu'un seul ; (ii) en les considérant séparément. Dans le cas du AR-DI, ces alternatives reviendraient à estimer respectivement les modèles suivants :

$$\begin{aligned} X_t^{ALL} &= \Lambda F_t^{ALL} + u_t^{ALL} \\ \hat{y}_{T+h|T}^{(h)} &= \hat{\alpha}^{(h)} + \hat{\rho}(L)^{(h)} y_T + \hat{\beta}(L)^{(h)} \hat{F}_T^{ALL} \end{aligned} \quad (32)$$

et

$$\begin{aligned} X_t^{(1)} &= \Lambda F_t^{(1)} + \mu_t^{(1)} \\ &\dots \\ X_t^{(D)} &= \Lambda F_t^{(D)} + \mu_t^{(D)} \\ \hat{y}_{T+h|T}^{(h)} &= \hat{\alpha}^{(h)} + \hat{\rho}(L)^{(h)} y_T + \sum_{i=1}^D \hat{\beta}_i(L)^{(h)} \hat{F}_T^{(i)} \end{aligned} \quad (33)$$

où D est le nombre d'ensembles d'information distincts. La première option, possible pour tous les modèles riches en données, consiste de 4 ensembles différents :

$$X_t = \begin{cases} [X_t^{QC}, X_t^{CAN}, X_t^{US}] \\ [X_t^{QC}, X_t^{CAN}] \\ [X_t^{QC}, X_t^{US}] \\ [X_t^{QC}] \end{cases}$$

Dans la deuxième option qui n'est possible que pour les modèles à indices de diffusion et le filtre régressif triple, les facteurs sont estimés séparément de ces 3 ensembles de groupes de séries :

$$\begin{aligned} X_t^{(1)}, X_t^{(2)}, X_t^{(3)} &= X_t^{QC}, X_t^{CAN}, X_t^{US} \\ X_t^{(1)}, X_t^{(2)} &= X_t^{QC}, X_t^{CAN} \\ X_t^{(1)}, X_t^{(2)} &= X_t^{QC}, X_t^{US} \end{aligned}$$

4.4 Agrégation des prévisions

Avec tous les modèles et toutes les combinaisons de panels, nous pouvons produire 110 prévisions de chacune des variables et pour chaque horizon d'intérêt. Cela ouvre la porte à l'agrégation des prévisions individuelles. La première agrégation est la simple moyenne :

$$\hat{y}_{T+h|T}^{(h)*} = \frac{\sum_{i=1}^M \hat{y}_{i,T+h|T}^{(h)}}{M} \quad (34)$$

La moyenne est aussi utilisée après une sélection des 5, 10, 20, 40 et 60 meilleurs modèles. L'agrégation suivante est une moyenne pondérée où les poids sont déterminés en fonction du ratio du $MSPE^{-1}$ sur la somme des $MSPE^{-1}$:

$$\begin{aligned} \omega_i &= \frac{MSPE_i^{-1}}{\sum_{i=1}^M MSPE_i^{-1}} \\ \hat{y}_{T+h|T}^{(h)*} &= \sum_{i=1}^M \omega_i \hat{y}_{i,T+h|T}^{(h)} \end{aligned} \quad (35)$$

La dernière agrégation est une moyenne calculée sur un sous groupe de modèles choisi en fonction du niveau des prévisions une itération à la fois. Pour chaque itération les prévisions sont classées de la plus grande à la plus petite et une proportion λ est retirée de chacune des extrémités du vecteur de prévision. La moyenne est calculée sur les prévisions restantes.

Finalement, on peut aussi utiliser la prévision médiane de tous les modèles à toutes les itérations. En incluant toutes les spécifications et toutes les façons de prendre en comptes les panels macroéconomiques, nous obtenons 119 modèles à estimer pour chaque variables et chaque horizon de prévision.

5 Analyse préliminaire : structure factorielle des ensembles d'information

Avant d'estimer les modèles de prévision basés sur les grands ensembles d'information, il faut vérifier l'existence de la structure à facteurs dans ces derniers. Nous commençons par présenter les 'tests' visuels habituels : (i) "scree plot", les valeurs propres, en ordre décroissant, de la matrice de corrélations des données ; (ii) "trace" représentant le R^2 de k facteurs subséquents.

La Figure 2 montre ces quantités pour chacun des ensembles d'information, QC , CA et US , ainsi que pour tous les trois ensemble. Comme c'est habituellement le cas avec données macroéconomiques, les valeurs propres descendent à un rythme lent mais nous pouvons constater que 6-7 facteurs suffisent pour expliquer 50% de la variance des observables. Notons que la démarcation des valeurs propres les importantes est plus prononcée dans le cas de la base de données fusionnées.

Après cette vérification visuelle, nous procédons à l'estimation du nombre de facteurs (statiques et dynamiques) à l'aide des critères d'information et des tests formels. Le Tableau 2 présente les résultats pour tous les ensembles. L'explication et la performance des différents tests et critères d'information sont disponibles dans Mao Takongmo et Stevanović (2015). Les résultats suggèrent qu'une structure à facteurs existe dans nos ensembles de données.

Tableau 2 – Résultats des tests statistiques

	QC	CAN	US	ALL
Bai et Ng (2002)	5	6	9	7
Alessi et al. (2013)	12	9	13	7
Bai et Ng (2007)	3, 5, 9	4, 5, 9	2, 5, 7	3, 5, 9
Hallin et Liska (2007)	1	3	7	5
Amengual et Watson (2007)	9	5	6	5
Onatski (2010)	1	1	1	1
Ahn et Horenstein (2013)	1	1	1	1

Ce tableau représente l'estimation du nombre de facteurs statiques et dynamiques à l'aide de différents critères d'information et tests. Pour plus de détails, voir Mao Takongmo et Stevanović (2015).

La Figure 3 montre les 6 premiers facteurs estimés à partir de données québécoises seulement. Le Tableau 3 montre les cinq séries les plus corrélées avec chacun des facteurs estimés. Nous pouvons constater que le premier facteur est relié à l'activité économique réelle, le deuxième aux indices d'inflation et le troisième au marché immobilier. Le quatrième facteur est corrélé avec les taux de change, le cinquième à l'investissement et le sixième aux exportations. Ces résultats témoignent les aspects d'une petite économie ouverte. Les Figures 4-5 et Tableaux 4-5 présentent le même exercice pour le Canada et pour les États-Unis.

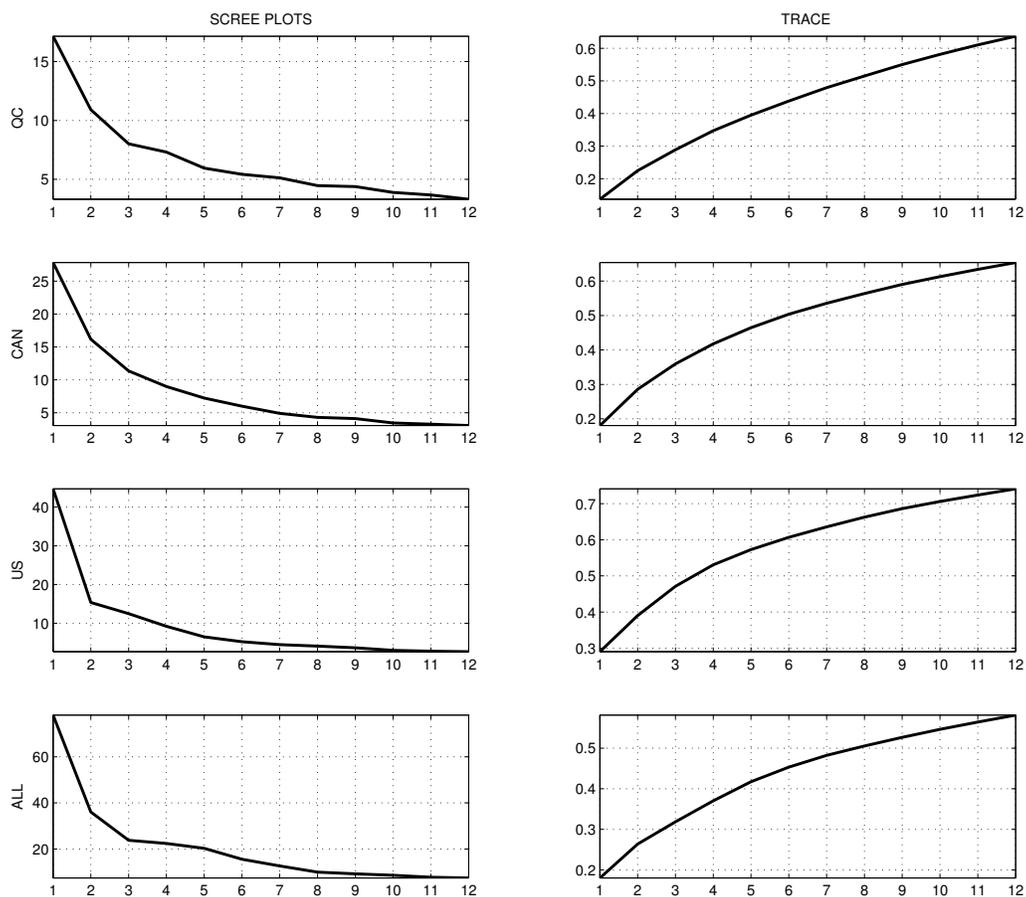


FIGURE 2 – Approche visuelle

“Scree plots” représente les valeurs propres, en ordre décroissant, de la matrice de corrélations des données. “Trace” représente le R^2 de k facteurs. QC indique que l’ensemble de données contient seulement les données québécoises, CA canadiennes, US américaines, tandis que ALL présente les résultats lorsque les trois ensembles sont combinés dans un seul panel.

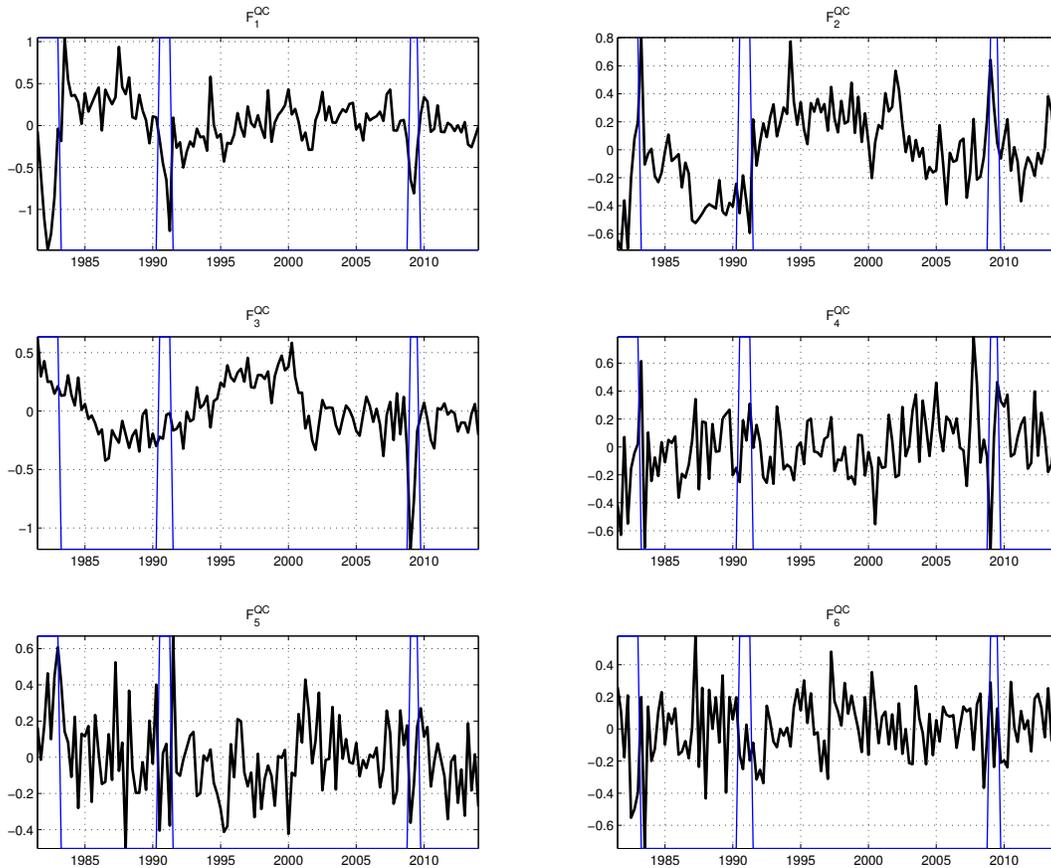


FIGURE 3 – Facteurs québécois

Cette figure montre les 6 facteurs québécois estimés sur tout l'échantillon. Les lignes bleues verticales délimitent les dates de récession canadiennes, déclarées par l'institut C.D. Howe.

Tableau 3 – Interprétation des facteurs du Québec

F1		F2	
0,74	Demande intérieure finale	0,41	Total indice des prix à la consommation
0,52	Produit intérieur brut aux prix du marché	0,39	Salaire implicite total
0,52	Dépenses personnelles en biens et services	0,39	Dépenses en loyers
0,51	Emploi	0,28	Indice implicite des prix à la consommation
0,51	Formation brute de capital fixe	0,28	IPC excluant aliments et énergie
F3		F4	
0,45	Mises en chantier - Locataires	0,44	United States dollar
0,41	Mises en chantier - Multiples	0,42	Japanese yen
0,40	Mises en chantier	0,42	Swiss franc
0,37	Stock de logements - multiples	0,41	Swedish krona
0,30	Mises en chantier - Propriétaires	0,29	Rémunération des salariés
F5		F6	
0,28	Construction domiciliaire totale	0,52	Exportations interprovinciales de services
0,28	Investissement total - construction résidentielle	0,50	Exportations interprovinciales
0,26	Investissements résidentiels autre que le neuf	0,31	Exportations interprovinciales de biens
0,23	Valeur des stocks - entreprises et gouvernements	0,25	Importations internationales de services
0,21	Investissements en rénovation	0,23	Dépenses publiques en biens et services

Ce tableau présente les cinq séries les plus corrélées, en chiffre et en valeur absolue, avec chacun des facteurs.

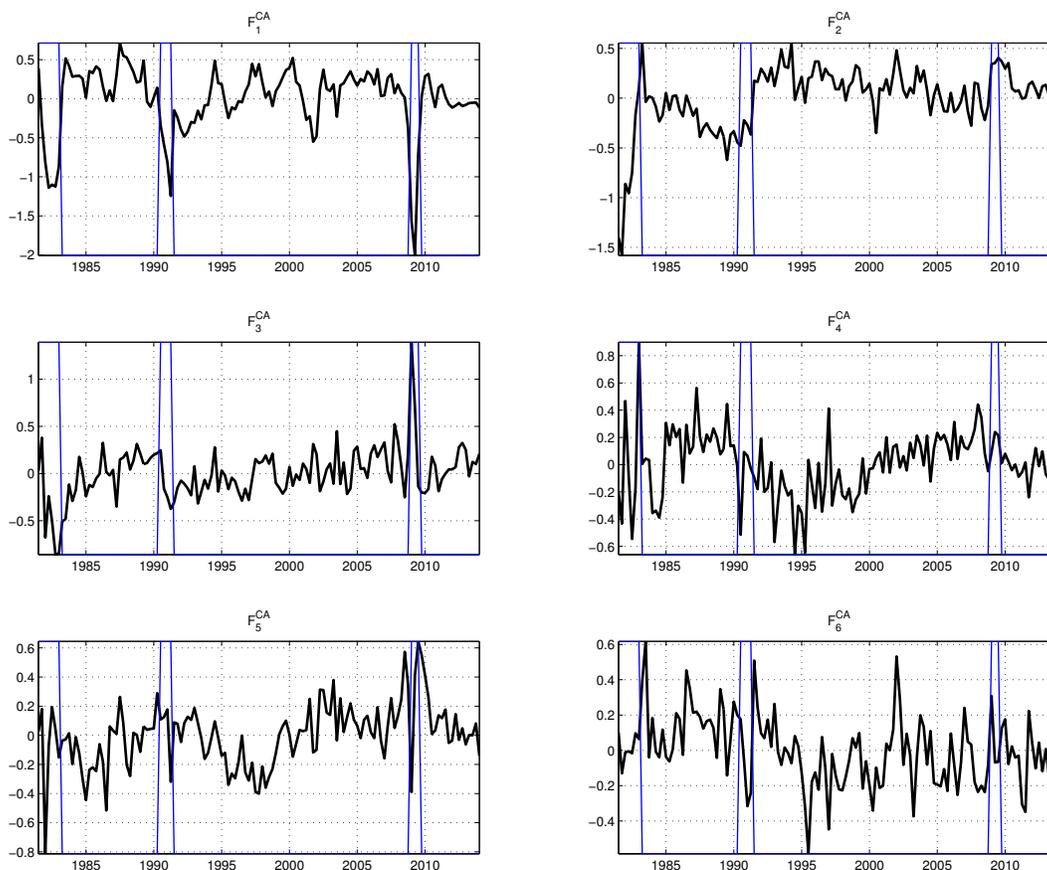


FIGURE 4 – Facteurs canadiens

Cette figure montre les 6 facteurs canadiens estimés sur tout l'échantillon. Les lignes bleues verticales délimitent les dates de récession canadienne, déclarées par l'institut C.D. Howe.

Tableau 4 – Interprétation des facteurs du Canada

F1		F2	
0,74	Taux d'emploi de la main-d'œuvre	0,60	Housing
0,74	Total employed, all industries	0,57	Services
0,73	Emploi dans le secteur privé non agricole	0,57	All-items CPI excluding food
0,68	Goods-producing sector	0,55	All-items CPI excluding energy
0,67	Demande intérieure finale	0,54	Salaire implicite total
F3		F4	
0,39	Revenu primaire des sociétés	0,36	Exportations - total
0,30	Bénéfices des sociétés non distribués	0,35	Exportations totales de marchandises
0,29	Goods	0,35	Exportations totales de marchandises
0,26	All-items CPI excluding food	0,31	Exportations nettes
0,26	Variation matérielle des stocks des entreprises	0,26	Exportations nettes de marchandises
F5		F6	
0,23	A CS : Long	0,28	Valeur des stocks des entreprises non agricoles
0,20	AA+ CS : Long	0,28	Valeur de la variation matérielle des stocks des entreprises
0,20	Total, all commodities	0,26	Variation matérielle des stocks des entreprises - non-agricoles
0,19	Gov of Canada bonds : over 10 years	0,24	Stock non agricoles
0,18	Food and energy	0,22	Prix dans l'industrie - pétrole et charbon

Ce tableau présente les cinq séries les plus corrélées, en chiffre et en valeur absolue, avec chacun des facteurs.

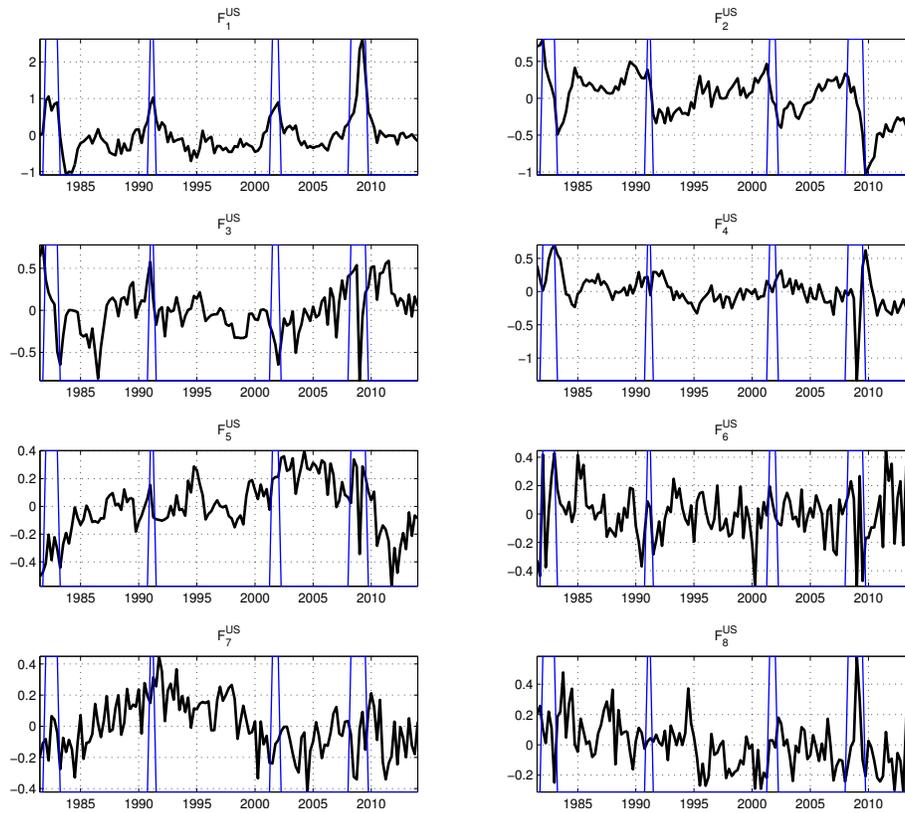


FIGURE 5 – Facteurs américains

Cette figure montre les 6 facteurs américains estimés sur tout l'échantillon. Les lignes bleues verticales délimitent les dates de récession américaines, déclarées par NBER.

Tableau 5 – Interprétation des facteurs des États-Unis

F1		F2	
0,87	All Employees : Total nonfarm	0,64	AAA-FEDFUNDS spread
0,86	All Employees : Goods-Producing Industries	0,57	BAA-FEDFUNDS spread
0,86	Nonfarm Business Sector : Employment	0,57	10-Year Treasury Maturity -FEDFUNDS spread
0,85	Nonfinancial Corporations Sector : Employment	0,49	6-Month Treasury Bill
0,81	Nonfarm Business Sector : Hours of All Persons	0,45	3-Month Treasury Bill - FEDFUNDS spread
F3		F4	
0,49	Producer Price Index : Finished Goods	0,29	CPI for All Urban Consumers : Medical Care
0,46	Producer Price Index : Finished Consumer Goods	0,28	CPI for All Urban Consumers : All Items Less Food
0,44	CPI for All Urban Consumers : Commodities	0,26	CPI for All Urban Consumers : All items less shelter
0,42	Producer Price Index : Industrial Commodities	0,25	CPI for All Urban Consumers : All Items
0,41	CPI for All Urban Consumers : Transportation	0,25	CPI for All Urban Consumers : Commodities
F5		F6	
0,31	Building Permits in the Midwest Census Region	0,46	Nonfarm Business Sector : Real Compensation Per Hour
0,25	Housing Starts in Midwest Census Region	0,46	Business Sector : Real Compensation Per Hour
0,21	Building Permits	0,39	Nonfarm Business Sector : Compensation Per Hour
0,19	6-Month Treasury Bill -FEDFUNDS spread	0,38	Business Sector : Compensation Per Hour
0,18	Housing Starts in West Census Region	0,23	10-Year Treasury Constant Maturity Rate
F7		F8	
0,34	Total Population : All Ages	0,23	Moodys Seasoned Baa Corporate Bond Yield
0,23	M2 Money Stock	0,18	Moodys Seasoned Aaa Corporate Bond Yield
0,23	Total Nonrevolving Credit Owned and Securitized	0,16	10-Year Treasury Constant Maturity Rate
0,17	Capacity Utilization : Manufacturing (NAICS)	0,15	S&P 500 Stock Price Index
0,16	Capacity Utilization : Manufacturing (SIC)	0,15	5-Year Treasury Constant Maturity Rate

Ce tableau présente les cinq séries les plus corrélées, en chiffre et en valeur absolue, avec chacun des facteurs.

6 Résultats

Dans cette section nous résumons les résultats les plus importants. Les résultats complets pour toutes les variables d'intérêt sont présentés en Annexe 2. Il serait bon de rappeler ici certains détails de l'exercice. Nous avons fait une course de modèles dans le cadre d'un exercice de prévision hors échantillon. Cet exercice a donné lieu à 119 prévisions du taux de croissance du PIB réel du Québec durant la période 2000Q1-2013Q4.¹⁸ Nous avons construit les prévisions pour les horizon $h=1, 2, 4$ et 8 trimestres. La métrique de comparaison est l'erreur quadratique moyenne (MSPE) durant la période d'évaluation : le modèle ayant la MSPE la plus petite en sort gagnant. Pour faciliter la comparaison, les résultats sont exprimées relativement à la MSPE du modèle AR direct qui est notre modèle de référence.

La notation des ensembles d'information se lit comme suite : 'Q' veut dire seulement X_t^{QC} a été utilisé ; 'QC' signifie X_t^{QC} et X_t^{CA} pris ensemble ; 'QU' signifie X_t^{QC} et X_t^{US} pris ensemble ; et 'QCU' dénote les trois bases de données combinées [X_t^{QC} , X_t^{CA} , X_t^{US}]. Par la suite, nous considérons la combinaison des facteurs estimés séparément des différentes bases de données : 'Q-C' veut dire que les facteurs ont été estimés séparément de X_t^{QC} et de X_t^{CA} ; 'Q-U' de X_t^{QC} et de X_t^{US} , et finalement 'Q-C-U' indique que les facteurs ont été estimés séparément de X_t^{QC} , X_t^{CA} et X_t^{US} .

6.1 Prévision du taux de croissance du PIB du Québec

PIB RÉEL Nous commençons par résumer les principaux résultats au niveau de la prévision du taux de croissance du PIB réel du Québec. L'analyse de l'exercice de prévision des autres variables suivra à la prochaine section. Le Tableau 6 montre la performance des 10 meilleurs modèles et combinaisons pour chaque horizon, en terme de MSPE relatif au modèle AR direct, ainsi que le pouvoir prévisionnel des modèles en terme du pseudo- R^2 . Pour prévoir le taux de croissance du PIB réel dans un trimestre, le meilleur modèle est la moyenne des 5 modèles ayant la plus petite erreur quadratique moyenne, suivi des moyennes de 10, 20 et 40 modèles. Le meilleur modèle individuel est le FAVARMA avec 1 seul facteur estimé à partir de la combinaison des données québécoises et canadiennes.

À deux trimestres en avance, le meilleur modèle semble être le DFM combinant les ensembles de données québécoises et canadiennes, suivi des moyennes (5 et 10), et du modèle ARDI utilisant les facteurs dynamiques calculés également de l'ensemble 'QC'. Lorsqu'on prévoit les taux de croissance moyens sur 1 et 2 ans, les moyennes des prévisions individuelles sont les meilleurs candidats ainsi que le modèle ARDI-tstat-QC.

En résumé, pour prévoir le taux de croissance moyen du PIB réel, deux ingrédients semblent importants : (i) moyenne sur un grand nombre de modèles de prévision, (ii) l'ensemble d'information combinant les données québécoises et canadiennes.

18. Comme l'analyse de sensibilité des résultats, nous avons refait l'exercice pour une autre période d'évaluation : 1995Q1-2013Q4. Les résultats sont très semblables.

La deuxième partie du Tableau 6 montre la mesure de prévisibilité, le pseudo- R^2 . Rappelons que si le chiffre est positif, le modèle se débrouille mieux en terme de RMSPE que la moyenne historique. Nous remarquons que la prévisibilité du taux de croissance moyen du PIB réel est relativement bonne, bien que le meilleur modèle ne réalise qu'un gain de 10% par rapport à la moyenne historique à l'horizon $h=2$ trimestres.

Tableau 6 – Préviation du taux de croissance du PIB réel

RMSPE								
#	Modèles	h=1	Modèles	h=2	Modèles	h=4	Modèles	h=8
1	Mean forecast 5	0.7568	DFM-QC	0.7607	Mean forecast 5	0.6247	ARDItstat-QC	0.4273
2	Mean forecast 10	0.7755	Mean forecast 5	0.7659	Mean forecast 10	0.6391	Mean forecast 5	0.4471
3	Mean forecast 20	0.7954	Mean forecast 10	0.7804	Mean forecast 20	0.6686	Mean forecast 10	0.4832
4	Mean forecast 40	0.8029	ARDIDU-QC	0.7805	Mean forecast 40	0.6822	Mean forecast 20	0.4960
5	FVRM-1fac-QC	0.8136	Mean forecast 20	0.7888	Mean forecast 60	0.6881	Mean forecast 40	0.5121
6	Mean forecast 60	0.8153	Mean forecast 40	0.7907	ARDIDU-QCU	0.6927	FAVARIT-Q	0.5313
7	ARDIDU-Q-C	0.8186	ARMA(1.1)	0.8066	FAVARDI-QC	0.6936	ARDIY1-QC	0.5369
8	ARDIY1-Q	0.8225	FAVARIT-Q	0.8165	ARDI-QC	0.7098	Mean forecast 60	0.5466
9	DFM-QC	0.8267	Mean forecast 60	0.8214	ARDIY1-QC	0.7136	Inversely prop forecast	0.5509
10	ARDIY1-QCU	0.8337	DFM-QCU	0.8234	ARDItstat-QC	0.7247	Median forecast	0.5598

Pseudo- R^2								
#	Modèles	h=1	Modèles	h=2	Modèles	h=4	Modèles	h=8
1	Mean forecast 5	0.2514	DFM-QC	0.1066	Mean forecast 5	0.2916	ARDItstat-QC	0.5563
2	Mean forecast 10	0.2329	Mean forecast 5	0.1005	Mean forecast 10	0.2752	Mean forecast 5	0.5358
3	Mean forecast 20	0.2132	Mean forecast 10	0.0835	Mean forecast 20	0.2418	Mean forecast 10	0.4982
4	Mean forecast 40	0.2058	ARDIDU-QC	0.0834	Mean forecast 40	0.2264	Mean forecast 20	0.4850
5	FVRM-1fac-QC	0.1952	Mean forecast 20	0.0736	Mean forecast 60	0.2198	Mean forecast 40	0.4682
6	Mean forecast 60	0.1935	Mean forecast 40	0.0714	ARDIDU-QCU	0.2145	FAVARIT-Q	0.4484
7	ARDIDU-Q-C	0.1903	ARMA(1.1)	0.0528	FAVARDI-QC	0.2135	ARDIY1-QC	0.4425
8	ARDIY1-Q	0.1864	FAVARIT-Q	0.0411	ARDI-QC	0.1951	Mean forecast 60	0.4325
9	DFM-QC	0.1823	Mean forecast 60	0.0354	ARDIY1-QC	0.1908	Inversely prop forecast	0.4280
10	ARDIY1-QCU	0.1753	DFM-QCU	0.0329	ARDItstat-QC	0.1782	Median forecast	0.4187

Ce tableau contient les 10 meilleurs modèles de préviation du taux de croissance du PIB réel. La première partie montre les erreurs quadratiques moyennes relatives au modèle de référence, AR direct. Si le chiffre est plus petit que 1, cela veut dire que le modèle en question produit un MSPE plus petit que le modèle AR direct. La deuxième partie du tableau montre les pseudo- R^2 . Si le chiffre est plus grand que 0, cela indique que le modèle fait mieux, en terme de l'erreur quadratique moyenne, que si les prévisions sont faites avec la simple moyenne historique.

PIB NOMINAL Il peut être intéressant de prévoir aussi le taux de croissance moyen du PIB en valeur nominale. Le Tableau 7 présente les résultats. Les meilleurs modèles sont en général les moyennes et ARMA(1,1), suivis du filtre à trois étapes avec combinaison 'QU' et le modèle à facteurs dynamiques. La grande différence par rapport au PIB réel est cependant la prévisibilité. Comme le témoigne la deuxième partie du tableau, les horizons de 1 et 2 trimestres sont difficiles à prévoir puisque l'amélioration par rapport à la préviation par moyenne historique est très modeste. La performance des meilleurs modèles augmente avec l'horizon mais n'atteint pas celle du PIB réel.

Tableau 7 – Prédiction du taux de croissance du PIB nominal

RMSPE								
#	Modèles	h=1	Modèles	h=2	Modèles	h=4	Modèles	h=8
1	Mean forecast 5	0.8787	ARMA(1.1)	0.8240	ARMA(1.1)	0.7211	Mean forecast 5	0.6466
2	Mean forecast 10	0.8790	Mean forecast 5	0.8422	DFM-QCU	0.7726	Mean forecast 10	0.6732
3	Mean forecast 20	0.8944	Mean forecast 10	0.8669	Mean forecast 5	0.7727	ARDItstat-QC	0.6902
4	3PRF-QU	0.8951	DFM-QC	0.8689	Mean forecast 10	0.7996	ARMA(1.1)	0.7174
5	ARDItstat-Q	0.9047	Mean forecast 20	0.8922	DFM-QU	0.8359	Mean forecast 20	0.7189
6	Mean forecast 40	0.9127	FAVARIT-Q	0.8923	Mean forecast 20	0.8408	ARDItstat-Q	0.7466
7	Mean forecast 60	0.9204	DFM-QU	0.9031	DFM-QC	0.8554	Median forecast	0.7686
8	DFM-QCU	0.9222	ARDItstat-Q	0.9124	FAVARIT-Q	0.8623	Mean forecast 40	0.7799
9	DFM-QU	0.9232	Mean forecast 40	0.9154	Mean forecast 40	0.8640	CRS-10-QC	0.7868
10	FAVARIT-Q	0.9258	DFM-QCU	0.9187	ARDItstat-Q	0.8769	Trimmed mean forecast	0.7930

Pseudo- R^2								
#	Modèles	h=1	Modèles	h=2	Modèles	h=4	Modèles	h=8
1	Mean forecast 5	0.0623	ARMA(1.1)	0.0186	ARMA(1.1)	0.1053	Mean forecast 5	0.1798
2	Mean forecast 10	0.0620	Mean forecast 5	-0.0030	DFM-QCU	0.0414	Mean forecast 10	0.1461
3	Mean forecast 20	0.0455	Mean forecast 10	-0.0324	Mean forecast 5	0.0412	ARDItstat-QC	0.1245
4	3PRF-QU	0.0448	DFM-QC	-0.0348	Mean forecast 10	0.0078	ARMA(1.1)	0.0901
5	ARDItstat-Q	0.0346	Mean forecast 20	-0.0626	DFM-QU	-0.0372	Mean forecast 20	0.0881
6	Mean forecast 40	0.0260	FAVARIT-Q	-0.0627	Mean forecast 20	-0.0433	ARDItstat-Q	0.0530
7	Mean forecast 60	0.0178	DFM-QU	-0.0756	DFM-QC	-0.0614	Median forecast	0.0251
8	DFM-QCU	0.0159	ARDItstat-Q	-0.0866	FAVARIT-Q	-0.0700	Mean forecast 40	0.0108
9	DFM-QU	0.0148	Mean forecast 40	-0.0903	Mean forecast 40	-0.0720	CRS-10-QC	0.0019
10	FAVARIT-Q	0.0120	DFM-QCU	-0.0942	ARDItstat-Q	-0.0881	Trimmed mean forecast	-0.0059

6.2 Prédiction du taux de croissance de l'emploi au Québec

Le Tableau 8 présente les résultats de prédiction du taux de croissance de l'emploi au Québec. Les meilleurs modèles s'avèrent être les moyennes des prévisions et des FAVARMA contenant les facteurs québécois seulement. Par contre, il s'agit d'une variable difficile à prévoir. La deuxième partie du tableau révèle qu'il est très difficile de battre la simple moyenne historique, sauf à l'horizon de 2 trimestres où le meilleur modèle améliore la prédiction par moyenne historique de 10%. Il serait intéressant d'étudier les causes de ce comportement et de trouver un meilleur modèle de prédiction.

6.3 Prédiction du taux d'inflation de l'IPC au Québec

Le Tableau 9 contient les résultats de prédiction du taux d'inflation tel que mesuré par l'indice des prix à la consommation au Québec. Comme c'est souvent le cas avec les taux d'inflation, ils sont difficiles à prévoir et le modèle ARMA(1,1) est toujours une très bonne option. Nous avons trouvé que les moyennes de prédiction performant le mieux aux horizons de 1, 2 et 4 trimestres, suivi des modèles FAVARMA et DFM utilisant les données québécoises et américaines. Les améliorations par rapport au modèle AR direct ne sont pas très grande à court terme mais deviennent plus intéressantes aux horizons de 1 et 2 ans. En particulier, ARMA(1,1) performe très bien à l'horizon de 4 trimestres. Les pseudo- R^2 montrent que la prévisibilité du taux d'inflation augmente grandement avec l'horizon de prédiction.

Tableau 8 – Prédiction du taux de croissance de l'emploi

RMSPE								
#	Modèles	h=1	Modèles	h=2	Modèles	h=4	Modèles	h=8
1	Mean forecast 5	0.8611	Mean forecast 5	0.8229	Mean forecast 5	0.8506	Mean forecast 20	0.7318
2	Mean forecast 10	0.8621	Mean forecast 20	0.8479	Mean forecast 10	0.8550	Mean forecast 10	0.7351
3	FVRM-AR3fac-Q	0.8745	Mean forecast 10	0.8496	FVRM-MA2fac-Q	0.8621	Mean forecast 5	0.7402
4	Mean forecast 20	0.8751	ARMA(1,1)	0.8656	DFM-QCU	0.8643	Mean forecast 60	0.7440
5	DFM-QC	0.8812	ARDItstat-Q	0.8839	Mean forecast 20	0.8733	Inversely prop fcst	0.7557
6	ARDIhard-1.28-Q	0.8819	Mean forecast 40	0.8846	Mean forecast 40	0.8752	Mean forecast 40	0.7616
7	DFM-QCU	0.8833	FAVARDI-QCU	0.8856	FVRM-MA3fac-Q	0.8771	Trimmed Mean fcst	0.7625
8	FVRM-MA3fac-QC	0.8882	Mean forecast 60	0.9084	Mean forecast 60	0.8838	Mean forecast	0.7664
9	FVRM-AR2fac-Q	0.8902	CRS-10-Q	0.9231	DFM-QU	0.8939	Median forecast	0.7740
10	Mean forecast 40	0.8918	CRS-10-QC	0.9274	Median forecast	0.8976	CRS-10-QC	0.7914

Pseudo- R^2								
#	Modèles	h=1	Modèles	h=2	Modèles	h=4	Modèles	h=8
1	Mean forecast 5	-0.0197	Mean forecast 5	0.1069	Mean forecast 5	-0.0830	Mean forecast 20	-0.0180
2	Mean forecast 10	-0.0209	Mean forecast 20	0.0798	Mean forecast 10	-0.0886	Mean forecast 10	-0.0226
3	FVRM-AR3fac-Q	-0.0356	Mean forecast 10	0.0780	FVRM-MA2fac-Q	-0.0977	Mean forecast 5	-0.0297
4	Mean forecast 20	-0.0362	ARMA(1,1)	0.0606	DFM-QCU	-0.1005	Mean forecast 60	-0.0349
5	DFM-QC	-0.0435	ARDItstat-Q	0.0407	Mean forecast 20	-0.1120	Inversely prop fcst	-0.0512
6	ARDIhard-1.28-Q	-0.0443	Mean forecast 40	0.0400	Mean forecast 40	-0.1143	Mean forecast 40	-0.0595
7	DFM-QCU	-0.0460	FAVARDI-QCU	0.0389	FVRM-MA3fac-Q	-0.1168	Trimmed Mean fcst	-0.0607
8	FVRM-MA3fac-QC	-0.0518	Mean forecast 60	0.0141	Mean forecast 60	-0.1253	Mean forecast	-0.0661
9	FVRM-AR2fac-Q	-0.0542	CRS-10-Q	-0.0018	DFM-QU	-0.1382	Median forecast	-0.0767
10	Mean forecast 40	-0.0561	CRS-10-QC	-0.0065	Median forecast	-0.1429	CRS-10-QC	-0.1008

Tableau 9 – Prédiction du taux d'inflation de l'IPC

RMSPE								
#	Modèles	h=1	Modèles	h=2	Modèles	h=4	Modèles	h=8
1	Mean forecast 10	0.9169	Mean forecast 5	0.9051	ARMA(1,1)	0.7854	Mean forecast 5	0.8484
2	Mean forecast 5	0.9218	Mean forecast 10	0.9078	Mean forecast 5	0.8487	DFM-QU	0.8659
3	Mean forecast 20	0.9405	ARMA(1,1)	0.9132	DFM-Q	0.9031	Mean forecast 10	0.9260
4	FVRM-MA3fac-Q	0.9462	Mean forecast 20	0.9211	Mean forecast 10	0.9424	Mean forecast 20	0.9411
5	CRS-20-QC	0.9566	Mean forecast 40	0.9489	AR(p) itératif	0.9444	DFM-Q	0.9481
6	ARDIsoft-Q-C	0.9642	ARDI-Q	0.9541	Mean forecast 20	0.9780	DFM-QCU	0.9706
7	3PRF-QC	0.9659	ARDIY1-Q	0.9572	3PRF-QC	0.9804	Median forecast	0.9980
8	Mean forecast 40	0.9682	ARDIDU-Q	0.9591	AR(p) direct	1.0000	AR(p) direct	1.0000
9	FAVARDI-Q	0.9725	ARDItstat-Q-C-U	0.9622	FVRM-1fac-QC	1.0002	AR(p) itératif	1.0046
10	ARDItstat-Q-C	0.9748	DFM-Q	0.9628	Mean forecast 40	1.0484	Inversely prop forecast	1.0097

Pseudo- R^2								
#	Modèles	h=1	Modèles	h=2	Modèles	h=4	Modèles	h=8
1	Mean forecast 10	0.0647	Mean forecast 5	0.0173	ARMA(1,1)	0.1764	Mean forecast 5	0.4383
2	Mean forecast 5	0.0598	Mean forecast 10	0.0143	Mean forecast 5	0.1101	DFM-QU	0.4267
3	Mean forecast 20	0.0407	ARMA(1,1)	0.0084	DFM-Q	0.0530	Mean forecast 10	0.3869
4	FVRM-MA3fac-Q	0.0349	Mean forecast 20	-0.0001	Mean forecast 10	0.0118	Mean forecast 20	0.3768
5	CRS-20-QC	0.0242	Mean forecast 40	-0.0303	AR(p) itératif	0.0097	DFM-Q	0.3722
6	ARDIsoft-Q-C	0.0165	ARDI-Q	-0.0360	Mean forecast 20	-0.0255	DFM-QCU	0.3573
7	3PRF-QC	0.0148	ARDIY1-Q	-0.0392	3PRF-QC	-0.0281	Median forecast	0.3392
8	Mean forecast 40	0.0124	ARDIDU-Q	-0.0413	AR(p) direct	-0.0486	AR(p) direct	0.3378
9	FAVARDI-Q	0.0080	ARDItstat-Q-C-U	-0.0447	FVRM-1fac-QC	-0.0488	AR(p) itératif	0.3348
10	ARDItstat-Q-C	0.0057	DFM-Q	-0.0453	Mean forecast 40	-0.0994	Inversely prop forecast	0.3314

6.4 Prédiction des autres variables

Dans cette section nous résumons les principales résultats pour les autres variables d'intérêt.

Demande intérieure finale Le Tableau 10 montre 3 meilleurs modèles de prévision, selon RMSPE, de la demande intérieure finale en versions réelle et nominale. Les meilleurs modèles sont les différentes moyennes de prévision, le VAR incluant quelque séries américaines, le modèle FAVARMA avec données québécoises seulement (dans la prévision 2 trimestres en avance) et le modèle à indices de diffusion dynamiques combinant les facteurs québécois et canadiens. La performance relative au modèle de référence est semblable entre séries réelle et nominale. Par contre, les valeurs du pseudo- R^2 sont très différentes. La série de demande intérieure finale en dollars courants semble bien plus difficile à prévoir puisque les modèles n'arrivent pas à battre la simple moyenne historique sauf à l'horizon $h=1$ trimestre.

Tableau 10 – Prédiction de la demande intérieure finale

h=1						h=2					
Réel			Nominal			Réel			Nominal		
Modèles	RMSE	p- R^2									
Mean forecast 5	0.88	0.17	Mean forecast 10	0.84	0.04	Mean forecast 10	0.80	0.13	Mean forecast 5	0.79	-0.04
Mean forecast 10	0.90	0.15	FVRM-MA2fac-Q	0.84	0.03	Mean forecast 20	0.82	0.11	Mean forecast 10	0.79	-0.05
ARDIDU-Q-C	0.91	0.14	Mean forecast 5	0.84	0.03	Mean forecast 40	0.86	0.07	Mean forecast 20	0.80	-0.06
h=4						h=8					
Réel			Nominal			Réel			Nominal		
Modèles	RMSE	p- R^2									
Mean forecast 20	0.85	0.06	Mean forecast 5	0.74	-0.27	Mean forecast 20	0.58	0.23	VAR-US	0.55	-0.75
Mean forecast 10	0.87	0.04	VAR-US	0.74	-0.27	Mean forecast 10	0.58	0.23	Mean forecast 10	0.55	-0.76
Mean forecast 5	0.88	0.03	Mean forecast 10	0.75	-0.29	Mean forecast 5	0.59	0.22	Mean forecast 5	0.56	-0.78

Ce tableau contient les 3 meilleurs modèles de prévision pour série nominale et série réelle. La colonne RMSE montre les erreurs quadratiques moyennes relatives au modèle de référence, AR direct. Si le chiffre est plus petit que 1, cela veut dire que le modèle en question produit un MSPE plus petit que le modèle AR direct. La colonne p- R^2 du tableau montre les pseudo- R^2 . Si le chiffre est plus grand que 0, cela indique que le modèle fait mieux, en terme de l'erreur quadratique moyenne, que si les prévisions sont faites avec la simple moyenne historique.

Dépenses personnelles en biens et services Le Tableau 11 compare les résultats de la prévision du taux de croissance de consommation en termes réel et nominal. Pour la série réelle, les meilleurs modèles aux horizons 1, 2 et 8 trimestres sont les moyennes des prévisions. À l'horizon de 1 an, le modèle Complete Subset Regression combinant les données québécoises et canadiennes (et américaines) performe le mieux. Dans le cas de la consommation en terme nominal, en plus des moyennes de prévisions, on retrouve les modèles VAR-US, ARMA et ARDI-tstat combinant les ensembles X_t^{QC} et X_t^{CA} . Au niveau du pseudo- R^2 , les chiffres sont semblables à l'exception du long terme où la série nominale semble plus facile à prévoir.

Investissements Nous avons considéré la prévision de plusieurs séries d'investissement. Le Tableau 12 montre les résultats pour le taux de croissance de l'investissement non résidentiels des entreprises en termes réel et nominal. Les moyennes des prévisions, le modèle VAR-US, le modèle à facteurs dynamiques combinant les trois ensembles de données ainsi que le

Tableau 11 – Prédiction de la consommation

h=1						h=2					
Réel			Nominal			Réel			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 20	0,88	0,00	Mean forecast 5	0,82	0,09	Mean forecast 5	0,80	0,02	Mean forecast 5	0,85	0,06
Mean forecast 10	0,88	0,00	Mean forecast 10	0,83	0,08	Mean forecast 10	0,81	0,01	VAR-US	0,85	0,05
Mean forecast 5	0,88	-0,01	Mean forecast 20	0,85	0,06	Mean forecast 20	0,84	-0,03	Mean forecast 10	0,85	0,05
h=4						h=8					
Réel			Nominal			Réel			Nominal		
Modèles	RMSE	p-R ²									
CSR-1-QC	0,83	-0,02	Mean forecast 10	0,79	0,03	Mean forecast 5	0,85	-0,01	Mean forecast 5	0,60	0,20
CSR-1-QCU	0,83	-0,03	ARMA(1,1)	0,81	0,01	Mean forecast 10	0,87	-0,04	Mean forecast 10	0,66	0,12
CSR-1-Q	0,83	-0,03	Mean forecast 20	0,81	0,01	CSR-10-QU	0,88	-0,05	ARDItstat-QC	0,68	0,08

modèle ARDI-tstat utilisant les trois ensembles de facteurs se présentent comme les meilleurs candidats pour prévoir la série réelle. Dans le cas de la variable nominale on y retrouve également le modèle ARMA(1,1). Tous ces modèles améliorent la prédiction par rapport au modèle de référence d'à peu près 20% à tous les horizons. En terme de prévisibilité, seulement la série nominale présente un pseudo- R^2 très faible à l'horizon de 2 ans.

Tableau 12 – Prédiction de l'investissement non résidentiels des entreprises

h=1						h=2					
Réel			Nominal			Réel			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 5	0,85	0,30	Mean forecast 5	0,75	0,40	VAR-US	0,83	0,18	Mean forecast 10	0,73	0,26
Mean forecast 10	0,87	0,29	Mean forecast 10	0,77	0,39	Mean forecast 5	0,84	0,17	Mean forecast 20	0,74	0,25
Mean forecast 20	0,88	0,28	Mean forecast 20	0,79	0,37	Mean forecast 10	0,86	0,16	Mean forecast 5	0,74	0,25
h=4						h=8					
Réel			Nominal			Réel			Nominal		
Modèles	RMSE	p-R ²									
DFM-QCU	0,82	0,10	Mean forecast 5	0,71	0,19	ARDItstat-Q-C-U	0,74	0,14	VAR-US	0,80	0,03
Mean forecast 5	0,85	0,07	DFM-QCU	0,72	0,17	Mean forecast 5	0,81	0,06	Mean forecast 5	0,83	-0,01
Mean forecast 10	0,85	0,07	Mean forecast 10	0,74	0,16	Mean forecast 10	0,85	0,01	ARMA(1.1)	0,84	-0,01

Le Tableau 13 résume les principaux résultats pour la construction domiciliaire. Nous remarquons qu'aux horizons de 1 et 2 trimestres, les meilleurs modèles arrivent à battre le modèle de référence mais de très peu : le meilleur étant la moyenne des 5 prévisions dans le cas de la série en terme réel et $h = 2$. Par contre, la performance des modèles, au niveau de l'erreur quadratique moyenne et du pseudo- R^2 , s'améliore grandement lorsqu'on prévoit à 1 et 2 ans en avance.

Tableau 13 – Prédiction de la construction domiciliaire

h=1						h=2					
Réel			Nominal			Réel			Nominal		
Modèles	RMSE	p-R ²									
ARDIY1-Q	0,94	0,18	Mean forecast 5	0,98	0,27	Mean forecast 5	0,90	0,08	Mean forecast 5	0,92	0,20
Mean forecast 5	0,95	0,17	Mean forecast 10	0,98	0,27	Mean forecast 10	0,91	0,08	ARMA(1.1)	0,92	0,19
DFM-Q	0,96	0,16	Mean forecast 20	0,99	0,27	ARMA(1.1)	0,91	0,08	Mean forecast 20	0,95	0,17
h=4						h=8					
Réel			Nominal			Réel			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 20	0,69	0,22	Mean forecast 5	0,64	0,33	Mean forecast 5	0,54	0,32	Mean forecast 5	0,55	0,30
Mean forecast 10	0,70	0,21	Mean forecast 10	0,64	0,33	Mean forecast 10	0,63	0,21	Mean forecast 10	0,63	0,19
Mean forecast 40	0,70	0,21	Mean forecast 20	0,71	0,26	Mean forecast 20	0,70	0,13	Mean forecast 20	0,69	0,11

Les Tableaux 14-16 présentent les résultats pour la formation brute de capital fixe dans les machines et matériel, en produits de propriétés intellectuelle et des entreprise en construction respectivement. Dans tous les cas il s'agit des séries en terme réel. Pour les deux premiers

types d'investissement les résultats indiquent une bonne performance des modèles de prévision : moyennes des prévisions, VAR-US, ARMA(1,1). Les modèles améliorent en général la précision entre 10 et 30%, sauf qu'à l'horizon de 8 trimestre dans le cas des machines et matériel. Le pseudo- R^2 est élevé pour la plupart des cas. Par contre, la formation du capital fixe dans le domaine des entreprises de construction est difficile à prévoir : les meilleurs modèles n'arrivent pratiquement pas à battre le modèle de référence et les pseudo- R^2 sont tous très proche de zéro.

Tableau 14 – Prévision de la Formation brute de capital fixe Machines et matériel, en \$ enchaînés

h=1			h=2			h=4			h=8		
Best models	RMSPE	p- R^2	Best models	RMSPE	p- R^2	Best models	RMSPE	p- R^2	Best models	RMSPE	p- R^2
Mean forecast 10	0.88	0.45	ARMA(1.1)	0.69	0.40	ARMA(1.1)	0.87	0.09	Mean forecast 5	0.99	-0.02
VAR-US	0.88	0.44	Mean forecast 5	0.79	0.31	Mean forecast 5	0.93	0.03	AR(p) direct	1.00	-0.03
Mean forecast 5	0.89	0.44	Mean forecast 10	0.82	0.28	Mean forecast 10	0.94	0.02	CRS-1-QCU	1.00	-0.03

Tableau 15 – Prévision de la Formation brute de capital fixe en produits de propriété intellectuelle, en \$ enchaînés

h=1			h=2			h=4			h=8		
Best models	RMSPE	p- R^2									
Mean forecast 10	0.90	0.33	Mean forecast 5	0.76	0.40	Mean forecast 10	0.82	0.28	Mean forecast 5	0.66	0.38
Mean forecast 20	0.91	0.32	Mean forecast 10	0.78	0.38	Mean forecast 20	0.82	0.27	ADL-US	0.69	0.35
Mean forecast 5	0.91	0.32	Mean forecast 20	0.80	0.37	Mean forecast 5	0.83	0.26	Mean forecast 10	0.71	0.33

Tableau 16 – Prévision de la Formation brute de capital fixe des entreprises Construction, en \$ enchaînés

h=1			h=2			h=4			h=8		
Best models	RMSPE	p- R^2	Best models	RMSPE	p- R^2	Best models	RMSPE	p- R^2	Best models	RMSPE	p- R^2
ARDItstat-Q-C	0.99	0.00	AR(p) direct	1.00	-0.03	DFM-QCU	0.96	-0.04	Mean forecast 10	0.99	0.03
Mean forecast 5	1.00	-0.01	CRS-1-Q	1.00	-0.03	Mean forecast 10	0.96	-0.04	AR(p) direct	1.00	0.02
Mean forecast 10	1.00	-0.01	Mean forecast 5	1.00	-0.03	Mean forecast 5	0.96	-0.05	Mean forecast 20	1.02	0.00

Exportations Nous avons évalué la prévisibilité de plusieurs séries d'exportations. Le Tableau 17 présente les résultats pour le taux de croissance des exportations totales en dollars courants. Les meilleurs modèles sont toujours les moyennes de 5, 10 et 20. La prévisibilité semble maximisée aux horizons de 2, 4 et 8 trimestres en avance. En particulier, l'amélioration par rapport au modèle AR direct est de presque 30% aux horizons de 1 et 2 ans.

Les Tableaux 18 - 20 résument les résultats pour trois catégories des exportations totales : exportations internationales seulement ; exportations internationales des biens et les exportations interprovinciales. Les exportations internationales sont en terme réel tandis que les interprovinciales sont considérées en termes réel et nominal. Les résultats pour les exportations internationales totale et des biens seulement sont très semblables : le pouvoir prévisionnel des modèles augmentent avec l'horizon, RMSE et pseudo- R^2 , et la performance relative par rapport à l'autorégressif atteint 47%. Les modèles retenus sont les moyennes de prévisions ainsi que des variantes du modèle à indices de diffusions. Il est intéressant de constater que

Tableau 17 – Prédiction des Exportations totales en dollars courants

h=1			h=2			h=4			h=8		
Best models	RMSPE	p-R ²									
Mean forecast 5	0.85	0.13	Mean forecast 5	0.76	0.23	Mean forecast 10	0.70	0.29	Mean forecast 5	0.71	0.27
Mean forecast 10	0.86	0.11	Mean forecast 10	0.79	0.20	Mean forecast 5	0.71	0.27	Mean forecast 10	0.71	0.27
Mean forecast 40	0.87	0.11	Mean forecast 20	0.81	0.18	Mean forecast 20	0.72	0.27	Mean forecast 20	0.72	0.27

ces derniers ont utilisé les trois ensembles d'information, QC, CA et US afin de mieux prévoir les exportations internationales. Alors, le choix des ensembles d'information est sensible à l'objectif de prédiction.

Tableau 18 – Prédiction des Exportations internationales, en \$ enchaînés

h=1			h=2			h=4			h=8		
Best models	RMSPE	p-R ²									
Mean forecast 10	0.82	0.18	Mean forecast 10	0.71	0.29	Mean forecast 5	0.60	0.37	ARDIssoft-QCU	0.53	0.49
Mean forecast 5	0.82	0.17	Mean forecast 5	0.71	0.28	Mean forecast 10	0.61	0.36	Mean forecast 20	0.56	0.46
Mean forecast 40	0.83	0.17	Mean forecast 20	0.72	0.27	Mean forecast 20	0.63	0.35	Mean forecast 5	0.58	0.44

Tableau 19 – Prédiction des Exportations internationales de biens, en \$ enchaînés

h=1			h=2			h=4			h=8		
Best models	RMSPE	p-R ²									
Mean forecast 20	0.78	0.21	Mean forecast 10	0.68	0.32	Mean forecast 5	0.61	0.38	Mean forecast 5	0.53	0.49
Mean forecast 5	0.78	0.21	Mean forecast 5	0.68	0.32	ARDIDU-Q-C-U	0.62	0.37	Mean forecast 10	0.55	0.46
Mean forecast 10	0.78	0.21	Mean forecast 20	0.69	0.31	Mean forecast 20	0.62	0.37	ARDIssoft-QCU	0.56	0.46

Importations Maintenant nous résumons les résultats de prédiction des importations internationales et interprovinciales. Les Tableaux 21-22 résume la performance des modèles pour les séries réelles et nominales. Dans le cas des importations internationales, les meilleurs modèles sont les moyennes de prévisions ainsi que le modèle ARDI-QU utilisant les ensembles d'information du Québec et des États-Unis. En terme réel, l'amélioration par rapport au modèle de référence est de 30% à court terme, mais baisse à 10% à l'horizon de 1 an. Les modèles semblent performer mieux dans la prédiction de la série nominale puisque les gains sont de 45% à 25% même à l'horizon de 2 ans.

Dans le cas des importations interprovinciales, nous retrouvons comme meilleurs modèles différentes spécifications des indices de diffusion construites surtout à partir des données québécoises. De plus, le modèle FAVARMA avec facteurs québécois performe le mieux à l'horizon de deux ans.

Tableau 20 – Préviation des Exportations interprovinciales

h=1						h=2					
Réal			Nominal			Réal			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 5	0,75	0,20	Mean forecast 5	0,91	0,00	ARMA(1,1)	0,93	0,07	ARMA(1,1)	0,77	0,06
Mean forecast 10	0,79	0,15	CRS-10-Q	0,91	0,00	Mean forecast 5	0,94	0,06	Mean forecast 5	0,81	0,01
Mean forecast 20	0,80	0,14	Mean forecast 10	0,92	-0,02	Mean forecast 20	0,94	0,06	Mean forecast 10	0,86	-0,05
h=4						h=8					
Réal			Nominal			Réal			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 5	0,94	0,11	ARMA(1,1)	0,74	0,03	Mean forecast 5	0,85	0,04	Mean forecast 20	0,81	-0,10
Mean forecast 10	0,97	0,08	Mean forecast 10	0,76	0,00	Mean forecast 10	0,87	0,02	Mean forecast 10	0,84	-0,13
ARMA(1,1)	0,97	0,08	Mean forecast 5	0,78	-0,03	FAVARIT-Q	0,88	0,01	CRS-1-QU	0,84	-0,14

Tableau 21 – Préviation des Importations internationales

h=1						h=2					
Réal			Nominal			Réal			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 5	0,70	0,33	Mean forecast 5	0,55	0,40	ARDIDU-Q-U	0,71	0,26	ARDIDU-QU	0,65	0,33
Mean forecast 10	0,71	0,32	Mean forecast 10	0,57	0,37	ARDIDU-QU	0,74	0,23	Mean forecast 5	0,68	0,30
CRS-10-QCU	0,71	0,32	Mean forecast 20	0,59	0,36	Mean forecast 5	0,75	0,22	Mean forecast 10	0,68	0,29
h=4						h=8					
Réal			Nominal			Réal			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 5	0,90	0,05	Mean forecast 5	0,57	0,42	Mean forecast 10	0,98	0,01	Mean forecast 10	0,75	0,26
3PRF-Q	0,90	0,05	Mean forecast 10	0,58	0,42	Mean forecast 5	0,98	0,01	Mean forecast 5	0,80	0,21
Mean forecast 10	0,90	0,05	Mean forecast 20	0,60	0,39	Mean forecast 20	0,98	0,00	Mean forecast 20	0,81	0,20

Tableau 22 – Préviation des Importations interprovinciales

h=1						h=2					
Réal			Nominal			Réal			Nominal		
Modèles	RMSE	p-R ²									
ARDIY1-Q	0,83	0,11	ARDIY1-Q	0,75	0,22	ARDIDU-Q	0,73	0,26	Mean forecast 5	0,60	0,38
Mean forecast 5	0,83	0,10	Mean forecast 10	0,76	0,22	Mean forecast 5	0,74	0,25	Mean forecast 10	0,60	0,37
Mean forecast 10	0,85	0,09	Mean forecast 5	0,76	0,21	Mean forecast 10	0,76	0,23	Mean forecast 20	0,63	0,34
h=4						h=8					
Réal			Nominal			Réal			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 5	0,73	0,29	Mean forecast 5	0,61	0,38	FVRM-MA2fac-Q	0,50	0,41	FVRM-MA2fac-Q	0,58	0,26
Mean forecast 10	0,74	0,29	ARDIhard-1.65-QC	0,62	0,36	Mean forecast 5	0,53	0,38	Mean forecast 5	0,61	0,23
ARDIDU-Q	0,75	0,28	Mean forecast 10	0,63	0,35	Mean forecast 20	0,55	0,36	Mean forecast 10	0,64	0,19

Dépenses publiques en biens et services Le Tableau 23 présente les résultats de prévision du taux de croissance des dépenses publiques, mesurées en termes réel et nominal. Cette variable inclut les dépenses faites au Québec par les gouvernements du Québec et du Canada ainsi que les gouvernements locaux. Les meilleurs modèles pour la série réelle sont les moyennes de prévisions, deux modèles d'indice de diffusion combinant les facteurs québécois, canadiens et américains, ainsi que le modèle FAVARDI utilisant les trois bases de données combinées. Cela indique que l'information externe est importante dans la prévision des dépenses publiques des gouvernements au Québec. La performance relative des modèles dans la prévision de la série nominale est bien plus petites surtout aux horizons de 1 et 2 trimestres.

Tableau 23 – Prévision des Dépenses publiques en biens et services

h=1			h=2								
Réal			Nominal			Réal			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 5	0.78	0.17	Mean forecast 5	0.99	-0.01	Mean forecast 10	0.61	0.36	Mean forecast 5	0.87	0.02
Mean forecast 10	0.78	0.17	Mean forecast 10	1.00	-0.02	Mean forecast 5	0.61	0.36	Mean forecast 10	0.88	0.01
ARDIDU-Q-C-U	0.82	0.13	AR(p) direct	1.00	-0.02	Mean forecast 20	0.67	0.30	Mean forecast 20	0.89	0.00
h=4			h=8								
Réal			Nominal			Réal			Nominal		
Modèles	RMSE	p-R ²									
Mean forecast 10	0.56	0.38	Mean forecast 5	0.59	0.05	Mean forecast 5	0.64	0.17	ARMA(1,1)	0.44	-0.06
Mean forecast 5	0.57	0.38	Mean forecast 10	0.61	0.02	Mean forecast 10	0.67	0.14	Mean forecast 5	0.50	-0.20
ARDIY1-Q-C-U	0.60	0.34	Mean forecast 20	0.65	-0.04	FAVARDI-QCU	0.68	0.12	Mean forecast 10	0.53	-0.27

Salaires et traitements Le Tableau 24 présente les résultats pour le taux de croissance des salaires en dollars courants. À court terme, les meilleurs modèles sont les moyennes de prévisions tandis que les modèles VAR s'avèrent de bons candidats aux horizons de 1 et 2 ans. Les améliorations par rapport au modèle de référence sont appréciables mais le pseudo-R² n'est pas élevé.

Tableau 24 – Prévision des Salaires et traitements en dollars courants

h=1			h=2			h=4			h=8		
Best models	RMSE	p-R ²	Best models	RMSE	p-R ²	Best models	RMSE	p-R ²	Best models	RMSE	p-R ²
Mean forecast 5	0.78	0.13	Mean forecast 20	0.82	0.02	VAR-US	0.66	0.10	VAR-US	0.60	-0.06
Mean forecast 10	0.80	0.11	Mean forecast 40	0.82	0.02	Mean forecast 5	0.66	0.10	Mean forecast 5	0.64	-0.13
Mean forecast 20	0.81	0.11	Mean forecast 10	0.82	0.02	Mean forecast 10	0.67	0.08	VAR-CAN	0.67	-0.19

Excédent d'exploitation des sociétés au net Le Tableau 25 présente des résultats de prévision de l'excédent d'exploitation des sociétés, mesurée en dollars courants. Le modèle FAVAR combinant l'information québécoise et canadienne est le meilleur à l'horizon de 1 trimestre, ARMA(1,1) à 2 trimestres et les moyennes de prévision à plus long terme. Nous notons en particulier la performance du modèle ARMA(1,1) à l'horizon $h = 2$ qui est nettement meilleur que les modèles suivants.

Revenu personnel disponible Le Tableau 26 résume les principaux résultats de prévision du taux de croissance du revenu personnel disponible en dollars courants. Il s'agit d'une

Tableau 25 – Prédiction des Excédents d'exploitation des sociétés au net, en dollars courants

h=1			h=2			h=4			h=8		
Best models	RMSE	p-R ²	Best models	RMSE	p-R ²	Best models	RMSE	p-R ²	Best models	RMSE	p-R ²
FAVARDI-QC	0.88	0.12	ARMA(1.1)	0.61	0.26	Mean forecast 5	0.64	0.16	Mean forecast 10	0.69	0.02
FAVARIT-QC	0.89	0.10	Mean forecast 5	0.76	0.08	ARDI-QC	0.67	0.12	Mean forecast 20	0.71	0.00
Mean forecast 5	0.90	0.10	Mean forecast 10	0.81	0.02	ARMA(1.1)	0.68	0.11	Mean forecast 5	0.72	-0.01

variable difficile à prévoir puisque les meilleurs modèles, parmi 119, arrivent difficilement à battre la simple moyenne historique. À court terme, certaines moyennes de prévision, un modèle à indice de diffusion et le modèle DFM améliorent légèrement la prévision de l'autorégressif directe et de la moyenne historique.

Tableau 26 – Prédiction du Revenu personnel disponible, en dollars courants

h=1			h=2			h=4			h=8		
Best models	RMSE	p-R ²	Best models	RMSE	p-R ²	Best models	RMSE	p-R ²	Best models	RMSE	p-R ²
Mean forecast 5	0.96	0.07	Mean forecast 5	0.89	0.07	AR(p) itératif	0.82	-0.11	AR(p) itératif	0.69	-0.31
ARDIsoft-Q	0.97	0.06	DFM-QC	0.90	0.06	Mean forecast 5	0.82	-0.12	Mean forecast 5	0.73	-0.40
Mean forecast 10	0.97	0.06	Mean forecast 10	0.91	0.06	Mean forecast 10	0.86	-0.17	VAR-CAN	0.74	-0.40

Taux de chômage Finalement, le Tableau 27 présente des résultats de prévision de la différence première du taux de chômage. Contrairement au cas du taux de croissance d'emploi, cet indicateur du marché d'emploi semble beaucoup plus facile à prévoir. Les améliorations par rapport à la simple moyenne historique sont appréciables même à long terme. Les meilleurs modèles améliorent la prévision par rapport au modèle de base entre 34% et 26% selon les horizons. Les meilleurs modèles sont en général les moyennes de prévisions individuelles, le modèle VAR-US ainsi que le modèle à facteurs dynamiques combinant les trois ensembles d'information.

Tableau 27 – Prédiction du Taux de chômage

h=1			h=2			h=4			h=8		
Best models	RMSE	p-R ²									
Mean forecast 5	0.74	0.24	VAR-US	0.67	0.29	Mean forecast 5	0.66	0.27	Mean forecast 5	0.72	0.16
Mean forecast 10	0.77	0.21	Mean forecast 5	0.68	0.29	DFM-QCU	0.73	0.19	Mean forecast 10	0.75	0.13
VAR-US	0.79	0.19	Mean forecast 20	0.71	0.26	Mean forecast 10	0.74	0.18	DFM-QU	0.75	0.12

7 Discussion

Cette section présente quelques éléments de réflexion sur les améliorations et extensions qu'on pourrait apporter à la présente étude.

Les résultats précédents suggèrent qu'en général, les séries nominales sont plus difficiles à prévoir que leurs contreparties réelles. Cependant, nous avons au départ écarté toutes les variables nominales lorsqu'une contrepartie réelle était disponible. Il est donc légitime de se demander si la présence de séries nominales dans les grands ensembles de données utilisés pour construire les facteurs pourrait permettre d'améliorer la prévisibilité des variables nominales. Intuitivement, il est plus difficile de prévoir une quantité nominale car son taux de croissance comporte une composante de l'inflation. Or, le taux d'inflation est habituellement difficile à prévoir. La meilleure façon de répondre à la question posée précédemment est de refaire l'exercice de prévision en incluant les séries nominales dans notre ensemble d'information.

L'examen de la corrélation entre les facteurs estimés et les variables observées a révélé que le premier facteur canadien et américain sont tous les deux reliés aux marchés de l'emploi des pays concernés. Si les marchés de l'emploi américain et canadien sont suffisamment indépendants l'un de l'autre, alors ces deux facteurs apportent différents types d'information. Dans le cas contraire, les deux marchés pourraient partager un facteur en commun. On aurait alors intérêt à estimer ce facteur commun de manière plus précise en combinant les bases de données US et CA.

De façon générale il est important de bien choisir les ensembles d'information servant à estimer les facteurs. Ils doivent être bien équilibrés : (i) représentatifs des économies considérées ; (ii) avoir une structure à facteurs assez forte pour permettre une estimation précise des facteurs sous-jacents et/ou la réduction de dimension ; (iii) pertinents pour les variables à prévoir. Il est donc tout-à-fait possible d'améliorer le pouvoir prédictif des modèles considérés dans cette étude en modifiant les ensembles d'information.

Étant donné le grand nombre de modèles considérés, il est important de préciser comment utiliser en pratique les méthodes développées dans cette étude. Nous avons déterminé les meilleurs modèles pour chaque variable et pour chaque horizon de prévision. Alors, notre recommandation est de suivre ces résultats : (i) faire la prévision pour chaque variable et chaque horizon à partir de tous les modèles étudiés ; (ii) pour la variable i et l'horizon h , choisir soit le meilleur modèle en terme de MSPE, soit une combinaison de 5 – 10 modèles les plus précis afin de construire la prévision finale.

Par exemple, considérons la prévision du PIB réel. Le Tableau 6 suggère que la meilleure façon de prévoir cette série un trimestre à l'avance est de considérer la moyenne des 5 modèles les plus précis, à savoir FVRM-1fac-QC, ARDIDU-Q-C, ARDIY1-Q, DFM-QC et ARDIY1-QCU. Si l'objectif est de connaître le taux de croissance moyen du PIB réel sur les 4 prochains trimestres, il faut regarder la colonne correspondante à $h = 4$.

Nous observons habituellement que la variance entre les modèles de prévision augmente durant les récessions, ce qui correspond aux périodes où la prévisibilité des variables macroéconomiques baisse de façon significative. Pour une variable et un horizon donné, la variance ainsi que l'écart entre le 5e et le 95e percentiles des prévisions produite par l'ensemble des modèles constituent des mesures *ad hoc* du pouvoir prédictif des modèles considérés.

Finalement, il est important de refaire régulièrement la course de modèles proposée dans la présente étude afin de tenir compte de possibles ruptures structurelles, mais aussi des dernières avancées en matière de prévision macroéconomique.

8 Conclusion

Le but du présent article est d'identifier les modèles économétriques les plus pertinents pour la prévision des principaux agrégats macroéconomiques du Québec dans un environnement riche en données. Notre approche consiste à considérer le Québec comme une économie régionale potentiellement soumise aux influences de ses deux géants voisins, à savoir le Canada et les États-Unis. Compte tenu de ceci, nous considérons trois grands groupes de variables pour mener nos exercices de prévisions : les variables québécoises (X_t^{QC}), canadiennes (X_t^{CA}) et américaines (X_t^{US}). Deux grands ensembles de modèles sont considérés dans cet exercice de prévision : les modèles standards (AR, ARMA, ADL et VAR) et les modèles riches en données (modèles à indice de diffusion, le *Complete Subset Regression*).

Nous considérons plusieurs options pour la construction des indices de diffusion à inclure dans les modèles riches en données. La première option consiste à extraire les facteurs séparément pour chacun des groupes de données (X_t^{QC} , X_t^{CA} et X_t^{US}). Les autres options consistent à extraire les facteurs après avoir fusionné plusieurs groupes de données. Ici, nous considérons tous les regroupements qui contiennent les données du Québec, c'est à dire : $[X_t^{QC}, X_t^{CA}]$ et $[X_t^{QC}, X_t^{US}]$ et $[X_t^{QC}, X_t^{CA}, X_t^{US}]$. Différentes spécifications de modèle et leurs variantes, différentes approches d'extraction des indices de diffusion et différentes possibilités de combinaisons de prévisions sont croisées pour obtenir 119 modèles de prévision macroéconomiques. Chaque modèle est estimé pour chacune des séries macroéconomiques à prédire. Le meilleur modèle est identifié pour chaque série dans une course de modèles sur la base de la qualité de prévision hors échantillon.

À l'examen des résultats, on constate que le meilleur modèle dépend de la série que l'on essaye de prédire et de l'horizon de prévision visé. Une famille de modèles qui donne de bon résultats à court horizon peut devenir moins attrayant à moyen et long horizon. Un modèle peut être bon pour prédire le PIB nominal et moins bon pour le PIB réel. Mais de façon générale, on peut affirmer que les moyennes des prévisions fournies par différents modèles délivrent les meilleurs performances.

Dans la catégorie des modèles standard, le modèle ARMA(1,1) s'est révélé un bon benchmark pour prédire certaines variables nominales (par exemple, le PIB nominal ou l'inflation). Les

modèle riches en données se classent souvent comme les meilleurs modèles individuels. Ceci démontre que les agrégats économiques du Québec peuvent être mieux prédits en associant les données canadiennes et américaines aux données québécoise.

Pour finir, notons que nos prévisions optimales ne sont pas conçues pour respecter les identités comptables, chaque agrégat étant prédit indépendamment des autres agrégats auxquels il est relié. En fait, prendre en compte les identités comptables reviendrait à ajouter de la structure à nos modèles de prévisions. Or, les modèles en forme réduite bénéficient de plus de liberté et de ce fait, génèrent en général de meilleures prévisions que les modèles structurelles.

Références

- Bai, J. et Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1), 191–221. Récupéré de <http://onlinelibrary.wiley.com/doi/10.1111/1468-0262.00273/abstract>
- Bai, J. et Ng, S. (2006). Confidence Intervals for Diffusion Index Forecasts and Inference for Factor Augmented Regressions. *Econometrica*, 74(4), 1133–1150.
- Bai, J. et Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2), 304–317. <http://dx.doi.org/10.1016/j.jeconom.2008.08.010>. Récupéré de <http://linkinghub.elsevier.com/retrieve/pii/S0304407608001085>
- Boivin, J. et Ng, S. (2005). Understanding and comparing factor-based forecasts. *International Journal of Central Banking*, 1 :3, p.117-152, 1(3), 117–152. Récupéré de <http://www.columbia.edu/~sn2294/pub/ijcb05.pdf>
- Boivin, J. et Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132(1), 169–194. <http://dx.doi.org/10.1016/j.jeconom.2005.01.027>. Récupéré de <http://linkinghub.elsevier.com/retrieve/pii/S030440760500045X>
- Chevillon, G. (2007). Direct Multi-Step Estimation and Forecasting. *Journal of Economic Surveys*, 21(4), 746–785.
- Dufour, J.-M. et Stevanović, D. (2013). Factor-Augmented VARMA Models With Macroeconomic Applications. *Journal of Business & Economic Statistics*, 31(4), 491–506. <http://dx.doi.org/10.1080/07350015.2013.818005>
- Elliott, G., Gargano, A. et Timmermann, A. (2013). Complete subset regressions. *Journal of Econometrics*. <http://dx.doi.org/10.1016/j.jeconom.2013.04.017>
- Elliott, G., Gargano, A. et Timmermann, A. (2015). Complete Subset Regressions with Large Dimensional Sets of Predictors. (2006).

- Forni, M., Hallin, M., Lippi, M. et Reichlin, L. (2005). The Generalized Dynamic Factor Model. *Journal of the American Statistical Association*, 100(471), 830–840. <http://dx.doi.org/10.1198/016214504000002050>
- Galbraith, J. (2003). Content horizons for univariate time series forecasts. *International Journal of Forecasting*, 19(1), 43–55.
- Hallin, M. et Liška, R. (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association*, 102(478), 603–617. <http://dx.doi.org/10.1198/016214506000001275>
- Kelly, B. et Pruitt, S. (2015). The three-pass regression filter : A new approach to forecasting using many predictors. *Journal of Econometrics*, 186(2), 294–316.
- Kopoin, A., Moran, K. et Paré, J.-P. (2013). Forecasting regional GDP with factor models : How useful are national and international data? *Economics Letters*, 121(2), 267–270. <http://dx.doi.org/10.1016/j.econlet.2013.08.007>. Récupéré de <http://linkinghub.elsevier.com/retrieve/pii/S0165176513003698>
- Mao Takongmo, C. O. et Stevanović, D. (2015). Selection of the Number of Factors in Presence of Structural Instability. *Actualité Économique, Special Issue on Identification, Simulation and Finite Sample Inference*.
- Marcellino, M., Stock, J. H. et Watson, M. W. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135, 499–526.
- McCracken, M. W. et Ng, S. (2014). FRED-MD : A monthly database for macroeconomic research. Récupéré de <http://www.columbia.edu/~sn2294/papers/freddata.pdf>
- Stock, J. H. et Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2), 147–162. <http://dx.doi.org/10.1198/073500102317351921>

Annexe 1 : Données

Les codes de transformations afin de rendre les variables stationnaires sont les suivants : (1) pas de transformation ; (2) première différence ; (4) logarithme ; (5) première différence des logarithmes. Les séries disponibles chez StatCan ont le code mnémonique qui commence par ‘v’. Les séries québécoises disponibles chez Institut de la Statistique du Québec ont un code alphabétique. Les séries québécoises et canadiennes dont le code mnémonique finit par ‘*’ sont disponibles chez StatCan et/ou ISQ mais pas leur code d’extraction. Les séries 269-280 sont disponibles à partir de *Datastream*. Finalement, les séries américaines ont été tirées de la base de données FRED.

Tableau 28 – Liste de variables

#	Code mnémonique	Nom	Transformation
Séries québécoises			
1	v41691783	Total indice des prix à la consommation du Québec	5
2	v41691784	IPC aliments, désaisonnalisé	5
3	v41691853	IPC automobile, désaisonnalisé	5
4	v41691854	IPC énergie, désaisonnalisé	5
5	v41691858	IPC essence, désaisonnalisé	5
6	v41691909	IPC excluant aliments et énergie, désaisonnalisé	5
7	v41691824	IPC gaz + mazout + électricité, désaisonnalisé	5
8	ISQ*	Total indice implicite des prix à la consommation	5
9	ISQ*	Prix du brent	5
10	v135428	Prix des matières premières - combustibles	5
11	v135431	Prix des matières premières - gaz naturel	5
12	v135417	Prix des matières premières - minéraux excluant énergie	5
13	ISQ*	Prix des matières premières - produits agricoles	5
14	v135341	Prix des matières premières - total	5
15	ISQ*	Demande finale, en \$ enchainés	5
16	ISQ*	Demande intérieure finale, en \$ enchainés	5
17	ISQ*	Produit intérieur brut aux prix du marché (millions de \$ enchainés de 2007)	5
18	ISQ*	Consommation finale, en \$ enchainés	5
19	ISQ*	Consommation des autres services, en \$ enchainés	5
20	ISQ*	Dépenses en aliments, en \$ enchainés	5
21	ISQ*	Dépenses en automobiles neuves, en \$ enchainés	5
22	ISQ*	Dépenses en automobiles usagées - constants	5
23	ISQ*	Dépenses en autres biens durables, en \$ enchainés	5
24	ISQ*	Dépenses en autres biens non durables, en \$ enchainés	5
25	ISQ*	Dépenses en autres biens semi-durables, en \$ enchainés	5
26	ISQ*	Dépenses en biens durables, en \$ enchainés	5
27	ISQ*	Dépenses en biens non durables, en \$ enchainés	5
28	ISQ*	Dépenses en biens semi-durables, en \$ enchainés	5
29	ISQ*	Dépenses en biens, en \$ enchainés	5
30	ISQ*	Dépenses en électricité, mazout et gaz, en \$ enchainés	5
31	ISQ*	Dépenses en énergie, en \$ enchainés	5
32	ISQ*	Dépenses en meubles et appareils, en \$ enchainés	5
33	ISQ*	Dépenses en loyers, en \$ enchainés	5
34	ISQ*	Dépenses en services, en \$ enchainés	5
35	ISQ*	Dépenses personnelles en biens et services - \$ enchainés	5
36	ISQ*	Investissement total - construction non résidentielle - en \$ enchainés	5
37	ISQ*	Investissement total - construction résidentielle - en \$ enchainés	5
38	ISQ*	Investissements en machines et matériel de bureau, en \$ enchainés	5
39	ISQ*	Investissements en rénovation, en \$ enchainés	5
40	ISQ*	Investissements non résidentiels des entreprises, en \$ enchainés	5
41	ISQ*	Investissements résidentiels autre que le neuf (commission + rénovation), en	5
42	ISQ*	Investissements résidentiels excluant les commissions (neuf + rénovation), en	5
43	ISQ*	Formation brute de capital fixe Machines et matériel, en \$ enchainés	5
44	ISQ*	Formation brute de capital fixe, en \$ enchainés	5
45	ISQ*	Formation brute de capital fixe en produits de propriété intellectuelle, en \$ enchainés	5
46	ISQ*	Formation brute de capital fixe des entreprises Construction, en \$ enchainés	5
47	ISQ*	Part de l'investissement dans le PIB, en \$ enchainés	2
48	ISQ*	Part de l'investissement total excluant la construction dans le PIB, reel	2
49	ISQ*	Excedent exploitation net des societes (% du produit intérieur net aux prix de base)	2
50	ISQ*	Rémunération des salariés	5
51	ISQ*	Rémunération des salariés (base nationale)	5
52	ISQ*	Revenu agricole	1
53	ISQ*	Revenu de la propriété payé par les ménages	5
54	ISQ*	Revenu de la propriété reçu par les ménages	5
55	ISQ*	revenu des ménages, \$ enchainés	5
56	ISQ*	Revenu personnel disponible réel (millions de \$ enchainés 1997)	5
57	ISQ*	Revenu mixte net des ménages	5
58	ISQ*	Revenu net de la propriété	5
59	ISQ*	Revenu net des entreprises non-agricoles, loyers compris	5

TABLE 28 – suite

#	Code mnémorique	Nom	Transformation
60	ISQ*	Revenus nets des entreprises individuelles non agricole excluant les loyers	5
61	ISQ*	Salaire implicite total	5
62	ISQ*	Salaire implicite, secteur privé	5
63	ISQ*	Salaire implicite, secteur public	5
64	ISQ*	Salaire privé	5
65	ISQ*	Salaires et traitements	5
66	ISQ*	Salaires et traitements, bruts	5
67	ISQ*	Construction domiciliaire neuve, en \$ enchaines	5
68	ISQ*	Construction domiciliaire totale, en \$ enchaines	5
69	ISQ*	Stock de logements (000 dunités)	5
70	ISQ*	Stock de logements (000 dunités) - détachés	5
71	ISQ*	Stock de logements (000 dunités) - multiples	5
72	ISQ*	Valeur du stock de logement total au Québec, désaisonnalisée en \$ enchaines	5
73	ISQ*	Mises en chantier (milliers)	4
74	ISQ*	Mises en chantier (milliers) - Locataires	4
75	ISQ*	Mises en chantier (milliers) - Multiples	4
76	ISQ*	Mises en chantier (milliers) - Propriétaires	4
77	ISQ*	Mises en chantier (milliers) - Simples	4
78	ISQ*	Exportations internationales - Électricité	5
79	ISQ*	Exportations internationales de biens : machine et matériel de bureau	5
80	ISQ*	Exportations internationales de biens, en \$ enchaines	5
81	ISQ*	Exportations internationales de machines excluant matériel de bureau	5
82	ISQ*	Exportations internationales de services, en \$ enchaines	5
83	ISQ*	Exportations internationales, en \$ enchaines	5
84	ISQ*	Exportations interprovinciales de biens, en \$ enchaines	5
85	ISQ*	Exportations interprovinciales de services, en \$ enchaines	5
86	ISQ*	Exportations interprovinciales, en \$ enchaines	5
87	ISQ*	Exportations nettes - en \$ enchaines	2
88	ISQ*	Exportations totales - en \$ enchaines	5
89	ISQ*	Importations internationales de biens, en \$ enchaines	5
90	ISQ*	Importations internationales de services, en \$ enchaines	5
91	ISQ*	Importations internationales, en \$ enchaines	5
92	ISQ*	Importations interprovinciales de biens, en \$ enchaines	5
93	ISQ*	Importations interprovinciales de services, en \$ enchaines	5
94	ISQ*	Importations interprovinciales, en \$ enchaines	5
95	ISQ*	Importations totales, en \$ enchaines	5
96	ISQ*	Variation matérielle des stocks, en \$ enchaines	1
97	ISQ*	Valeur des stocks - entreprises et gouvernements - en \$ enchaines	5
98	v2063756	Emploi	5
99	ISQ*	Emploi secteur privé	5
100	ISQ*	Emploi secteur public	5
101	v2063755	Population active (milliers)	5
102	v2063754	Population de 15 ans et plus, sur la base de l'enquête sur la population active	5
103	ISQ*	Population en âge de travailler	5
104	ISQ*	Population totale moyenne	5
105	v2063762	Rapport emploi-population	2
106	v2063761	Taux d'activité (%)	2
107	v2063760	Taux de chômage (%)	2
108	ISQ*	Moyenne mobile du taux hypothécaire 5 ans	2
109	ISQ*	Moyenne mobile, taux hypothécaire conventionnel à 5 ans	2
110	ISQ*	Taux hypothécaire conventionnel à 5 ans	2
111	ISQ*	Taux hypothécaire conventionnel à 5 ans, en termes réels	2
112	ISQ*	Dépenses publiques en biens et services, en \$ enchaines	5
113	ISQ*	Dépenses totales de l'administration publique, (Biens, services + investissements)	5
114	ISQ*	Impôts moins les subventions sur la production	5
115	ISQ*	Impôts sur le revenu du gouvernement du Québec	5
116	ISQ*	Impôts sur le revenu payés par les ménages	5
117	ISQ*	Investissement en construction non résidentielle des gouvernements	5
118	ISQ*	Investissements non résidentiels des administrations publiques, en \$ enchaines	5
119	ISQ*	Part de l'investissement du gouvernement dans le PIB, dollars enchaines	2
120	ISQ*	Transferts aux ménages de l'administration provinciale - Aide sociale	5
121	ISQ*	Transferts courants reçus par les ménages - Total	5
122	ISQ*	Transferts fédéraux aux ménages	5
123	ISQ*	Indice de volatilité des taux d'intérêt	2
124	ISQ*	Inflation anticipée	2
125	ISQ*	Ventes finales de biens	5
126	ISQ*	Taux d'épargne personnelle (%)	2
127	v37426	United States dollar, noon spot rate, average	5
128	v37456	United States dollar, 30-day forward closing rate	5
129	v37428	United States dollar, 180-day forward closing rate	5
130	v37429	United States dollar, 1-year forward closing rate	5
131	v37430	United Kingdom pound sterling, noon spot rate, average	5
132	v37431	United Kingdom pound sterling, 90-day forward noon rate	5
133	v37438	Swedish krona, noon spot rate, average	5
134	v37440	Swiss franc, noon spot rate, average	5
135	v37441	Japanese yen, noon spot rate, average	5
136	v41498903	Canadian dollar effective exchange rate index (CERI) (1992=100)	5
Séries canadiennes			
137	v41692942	All-items CPI excluding eight of the most volatile components (Bank of Canada definition)	5
138	v41691232	All-items CPI excluding food (2002=100)	5
139	v41691238	All-items CPI excluding energy (2002=100)	5
140	v41691237	Food and energy (2002=100)	5
141	v41691239	Energy (2002=100)	5

TABLE 28 – suite

#	Code mnémorique	Nom	Transformation
142	v41691219	Housing (1986 definition) (2002=100)	5
143	v41691222	Goods (2002=100)	5
144	v41691223	Durable goods (2002=100)	5
145	v41691225	Non-durable goods (2002=100)	5
146	v41691230	Services (2002=100)	5
147	v41691231	Services excluding shelter services (2002=100)	5
148	v54061220	Prix à l'importation de produits énergétiques	5
149	v79309137	Indice des prix industriels	5
150	v79309848	Indice des prix industriels - machinerie	5
151	v79309138	Prix dans l'industrie - pétrole et charbon	5
152	v52673496	Total, all commodities	5
153	v52673498	Energy	5
154	v52673499	Metals and Minerals	5
155	v52673502	Forestry	5
156	v62305752	Produit intérieur brut en \$ enchaines (millions de \$ enchainés de 1997)	5
157	v62305784	Demande finale - nominal	5
158	v62305723	Demande intérieure finale - en \$ enchaines	5
159	v62305724	Consommation - Total \$ enchainés	5
160	v62305725	Consommation - Biens \$ enchainés	5
161	v62305726	Consommation - Biens durables \$ enchainés	5
162	v62305727	Consommation - Biens semi-durables - en \$ enchaines	5
163	v62305728	Consommation - non durables - en \$ enchaines	5
164	v62305729	Consommation - services - en \$ enchaines	5
165	v62305731	Consommation finale - en \$ enchaines	5
166	StatCan*	Consommation personnelles en biens et services de consommation	5
167	StatCan*	Consommation en terme en \$ enchaines de biens durables excluant les voitures et les pièces	5
168	StatCan*	Consommation - Excluant aliments et loyers \$ enchaines	5
169	v62305736	Investissement non résidentiel des entreprises - total - en \$ enchaines	5
170	StatCan*	Investissement en outillage en dollars constant excluant bureau	5
171	StatCan*	Investissements non résidentiels des entreprises en construction - total - en \$ enchaines	5
172	v62305737	Investissements non résidentiels des entreprises en machines et outillages - total - en	5
173	v62295549	Rémunération des salariés	5
174	v62306158	Revenu des ménages	5
175	v63625369	Revenu disponible des ménages - en \$ enchaines	5
176	StatCan*	Revenu primaire des sociétés	2
177	StatCan*	Salaire implicite du secteur privé	5
178	v62468801	Salaire implicite public	5
179	WIMP*	Salaire implicite total	5
180	v62295550	Salaires et traitements	5
181	StatCan*	Salaires et traitements - privés	5
182	StatCan*	Salaires et traitements - publics	5
183	StatCan*	Bénéfices des sociétés non distribués	2
184	StatCan*	Bénéfices réinvestis nets	2
185	v62295552	Excédent d'exploitation brut des entreprises	5
186	v62295553	Excédent d'exploitation net des entreprises	5
187	v4667	Permis de bâtir : Total residential and non-residential	5
188	v4668	Permis de bâtir : Residential	5
189	v4670	Permis de bâtir : Industrial	5
190	v4671	Permis de bâtir : Commercial	5
191	v729949	Housing starts ; Total units	4
192	StatCan*	Mises en chantier (milliers)	4
193	StatCan*	Mises en chantier (milliers) - locataires	4
194	StatCan*	Mises en chantier (milliers) - Logements multiples	4
195	StatCan*	Mises en chantier (milliers) - Logements simples	4
196	StatCan*	Mises en chantier (milliers) - propriétaires	4
197	StatCan*	Dépenses - construction - total - nominal	5
198	StatCan*	Dépenses - construction résidentielle - neuf - en \$ enchaines	5
199	StatCan*	Dépenses - construction résidentielle - rénovation - en \$ enchaines	5
200	StatCan*	Dépenses - construction résidentielle - total - en \$ enchaines	5
201	StatCan*	Dépenses - construction résidentielle privé - autres que le neuf - en \$ enchaines	5
202	v61989250	Importations - total - en \$ enchaines	5
203	v61989276	Importations - machinerie, équipement et matériel de bureau - en \$ enchaines	5
204	v61989255	Importations de produits de lénergie - en \$ enchaines	5
205	v61989297	Importations de services - total - en \$ enchaines	5
206	v61989251	Importations totales de biens - en \$ enchaines	5
207	v61989281	Importation de produit de l'automobile - en \$ enchaines	5
208	v61989198	Exportations - total - en \$ enchaines	5
209	v61989224	Exportation de machinerie, équipement et matériel de bureau - en \$ enchaines	5
210	v61989203	Exportations dénergie - en \$ enchaines	5
211	v61989245	Exportations de services - en \$ enchaines	5
212	v61989199	Exportations totales de marchandises - en \$ enchaines	5
213	v61989229	Exportation de produit de l'automobile - en \$ enchaines	5
214	v62305748	Exportations nettes - en \$ enchaines	2
215	v62305749	Exportations nettes de marchandises - en \$ enchaines	2
216	v61989199	Exportations totales de marchandises - en \$ enchaines	5
217	v62425538	Impot sur la production - TOTAL	5
218	v62425530	Impôts directs payés par les ménages	5
219	v62355367	Impôts sur le revenu perçus par les administrations publiques - Total	5
220	v62425533	Total des impôts payés par les sociétés et entreprises publiques	5
221	v62425528	Recettes totales, ensemble des gouvernements	5
222	v62425563	Intérêts payés sur la dette de l'administration publique	5
223	StatCan*	Investissement total des gouvernements - en \$ enchaines	5
224	v62425551	Dépenses courantes brutes en biens et services de l'administration publique	5
225	StatCan*	Dépenses de biens et services et investissement des gouvernements	5
226	v62425550	Dépenses totales de l'administration publique	5

TABLE 28 – suite

#	Code mnémorique	Nom	Transformation
227	StatCan*	Dépenses totales des gouvernements - en \$ enchainées	5
228	v62425571	Prêts nets ou emprunts nets de l'administration publique	2
229	v62425564	Excédents ou déficits budgétaires de l'administration publique	2
230	v62425570	Stock non agricoles \$ enchainées	5
231	StatCan*	Valeur de la variation matérielle des stocks des entreprises - en \$ enchainées	1
232	StatCan*	Valeur des stocks des entreprises non agricoles	5
233	StatCan*	Variation matérielle des stocks des entreprises - non-agricoles \$ enchainées	1
234	v2062815	Unemployment rate (Rate); Both sexes; 15 years and over; Seasonally adjusted	2
235	v2057603	Total employed, all industries; Seasonally adjusted	5
236	v2057604	Goods-producing sector; Seasonally adjusted	5
237	v2057607	Utilities [22]; Seasonally adjusted	5
238	v2057608	Construction [23]; Seasonally adjusted	5
239	v2057609	Manufacturing [31-33]; Seasonally adjusted	5
240	v2057610	Services-producing sector; Seasonally adjusted	5
241	v2057611	Trade [41, 44-45]; Seasonally adjusted	5
242	v2057612	Transportation and warehousing [48-49]; Seasonally adjusted	5
243	v2057613	Finance, insurance, real estate and leasing [52-53]; Seasonally adjusted	5
244	v2057614	Professional, scientific and technical services [54]; Seasonally adjusted	5
245	v2057615	Business, building and other support services [55-56]; Seasonally adjusted	5
246	v2091051	Population active	5
247	v2091030	Population de plus de 15 ans	5
248	v2091048	Population de plus de 65 ans	5
249	StatCan*	Population en âge de travailler	5
250	StatCan*	Population totale (milliers)	5
251	StatCan*	Emploi dans le secteur privé non agricole	5
252	v2057607	Emploi dans le secteur public - niveau	5
253	v2091198	Taux d'activité (%)	2
254	v2091219	Taux d'emploi de la main-d'œuvre	2
255	v37412	Toronto Stock Exchange, value of shares traded (Dollars)	5
256	v37413	Toronto Stock Exchange, volume of shares traded (Shares)	5
257	v122620	Standard and Poors/Toronto Stock Exchange Composite Index, close (Index, 2000=1000)	5
258	v122628	Toronto Stock Exchange, stock dividend yields (composite), closing quotations	2
259	v122530	Bank rate (Percent)	2
260	v122505	Forward premium or discount (-), United States dollar in Canada : 3 month (Percent)	2
261	v122491	Prime corporate paper rate : 3 month (Percent)	2
262	v122558	Government of Canada marketable bonds, average yield : 1-3 year (Percent)	2
263	v122485	Government of Canada marketable bonds, average yield : 3-5 year (Percent)	2
264	v122486	Government of Canada marketable bonds, average yield : 5-10 year (Percent)	2
265	v122487	Government of Canada marketable bonds, average yield : over 10 years (Percent)	2
266	v122541	Treasury bill auction - average yields : 3 month (Percent)	2
267	v122552	Treasury bill auction - average yields : 6 month (Percent)	2
268	v122497	Average residential mortgage lending rate : 5 year (Percent)	2
269	DataStream	BBB CS : Long	1
270	DataStream	BBB CS : Mid	1
271	DataStream	BBB CS : Short	1
272	DataStream	BBB CS : Universe	1
273	DataStream	A CS : Long	1
274	DataStream	A CS : Mid	1
275	DataStream	A CS : Short	1
276	DataStream	A CS : Universe	1
277	DataStream	AA+ CS : Long	1
278	DataStream	AA+ CS : Mid	1
279	DataStream	AA+ CS : Short	1
280	DataStream	AA+ CS : Universe	1
281	v37153	Canadian dollar assets, total loans	5
282	v37119	Total personal loans	5
283	v37120	Business loans	5
284	v41552795	M1B (gross)	5
285	v37130	Residential mortgages	5
286	v41552798	M2+ (gross)	5
287	v37135	Chartered bank deposits, personal, term	5
288	v37140	Bankers acceptances	5
289	v36414	Total business and household credit; Seasonally adjusted	5
290	v36415	Household credit; Seasonally adjusted	5
291	v36416	Residential mortgage credit; Seasonally adjusted	5
292	v36417	Consumer credit; Seasonally adjusted	5
293	v36418	Business credit; Seasonally adjusted	5
294	v36420	Short-term business credit; Seasonally adjusted	5
295	v122396	Total, Canadas official international reserves	5
296	v122397	Convertible foreign currencies, United States dollars	5
Séries américaines			
297	CUMFN	Capacity Utilization : Manufacturing (NAICS)	1
298	GDPC1	Real Gross Domestic Product	5
299	PIECTR	Real personal income excluding current transfer receipts	5
300	CQRMTSPL	Real Manufacturing and Trade Industries Sales	5
301	RPI	Real Personal Income	5
302	INDPRO	Industrial Production Index	5
303	IPBUSEQ	Industrial Production : Business Equipment	5
304	IPCONGD	Industrial Production : Consumer Goods	5
305	IPDCONGD	Industrial Production : Durable Consumer Goods	5
306	IPDMAT	Industrial Production : Durable Materials	5
307	IPFINAL	Industrial Production : Final Products (Market Group)	5
308	IPFPNSS	Industrial Production : Final Products and Nonindustrial Supplies	5

TABLE 28 – suite

#	Code mnémonique	Nom	Transformation
309	IPFUELS	Industrial Production : Fuels	5
310	IPMANSICS	Industrial Production : Manufacturing (SIC)	5
311	IPMAT	Industrial Production : Materials	5
312	IPNCONGD	Industrial Production : Nondurable Consumer Goods	5
313	IPNMAT	Industrial Production : nondurable Materials	5
314	NAPMPI	ISM Manufacturing : Production Index ^f	1
315	CUMFNS	Capacity Utilization : Manufacturing (SIC)	1
316	CMRMTSPL	Real Manufacturing and Trade Industries Sales	5
317	TCU	Capacity Utilization : Total Industry	1
318	DDURRA3Q086SBEA	Real personal consumption expenditures : Durable goods (chain-typequantity index)	5
319	DNDGRA3Q086SBEA	Real personal consumption expenditures : Nondurable goods (chain-typequantity index)	5
320	DSERRA3Q086SBEA	Real personal consumption expenditures : Services (chain-type quantityindex)	5
321	GPDIC1	Real Gross Private Domestic Investment	5
322	A327RA3Q086SBEA	Real private fixed investment in equipment and software (chain-typequantity index)	5
323	A349RA3Q086SBEA	Real private fixed investment in structures (chain-type quantityindex)	5
324	A756RA3Q086SBEA	Real private fixed investment in new structures : Residentialstructures	5
325	A771RA3Q086SBEA	Real private fixed investment in new structures : Nonresidentialstructures	5
326	B007RA3Q086SBEA	Real private fixed investment (chain-type quantity index)	5
327	B008RA3Q086SBEA	Real private fixed investment : Nonresidential (chain-type quantity index)	5
328	B010RA3Q086SBEA	Real private fixed investment : Nonresidential : Equipment and software	5
329	B011RA3Q086SBEA	Real private fixed investment : Residential (chain-type quantity index)	5
330	B013RA3Q086SBEA	Real private fixed investment : Residential : Equipment (chain-typequantity index)	5
331	B139RA3Q086SBEA	Real auto output : Private fixed investment : New autos (chain-typequantity index)	5
332	B292RA3Q086SBEA	Real private fixed investment : Residential : Structures : Permanentsite : Multifamily	5
333	B680RA3Q086SBEA	Real private fixed investment : Nonresidential :Equipment : Industrial equipment	5
334	B681RA3Q086SBEA	Real private fixed investment : Nonresidential :Equipment : Transportation equipment	5
335	B862RA3Q086SBEA	Real private fixed investment : Nonresidential :Equipment : Other equipment	5
336	B863RA3Q086SBEA	Real private fixed investment : Residential : Other structures(chain-type quantity index)	5
337	B943RA3Q086SBEA	Real private fixed investment : Residential : Structures : Permanent site	5
338	B944RA3Q086SBEA	Real private fixed investment : Residential : Structures : Permanent site : Single family	5
339	B985RA3Q086SBEA	Real private fixed investment : Nonresidential :Intellectual property products : Software	5
340	E318RA3Q086SBEA	Real private fixed investment : Nonresidential : Structures : Miningexploration, shafts, and wells	5
341	RINV	Real Investment	5
342	W001RA3Q086SBEA	Real private fixed investment : Nonresidential : Structures : Commercialand health care	5
343	W003RA3Q086SBEA	Real private fixed investment : Nonresidential : Structures : Power andcommunication	5
344	W004RA3Q086SBEA	Real private fixed investment : Nonresidential : Other structures(chain-type quantity index)	5
345	W040RA3Q086SBEA	Real private fixed investment in new structures (chain-type quantity index)	5
346	PRS85006013	Nonfarm Business Sector : Employment	5
347	PRS88003013	Nonfinancial Corporations Sector : Employment	5
348	CES1021000001	All Employees : Mining and Logging : Mining	5
349	CLF16OV	Civilian Labor Force	5
350	DMANEMP	All Employees : Durable goods	5
351	ICNSA	Initial Claims	5
352	MANEMP	All Employees : Manufacturing	5
353	NAPMEI	ISM Manufacturing : Employment Index ^f	1
354	NDMANEMP	All Employees : Nondurable goods	5
355	PAYEMS	All Employees : Total nonfarm	5
356	SRVPRD	All Employees : Service-Providing Industries	5
357	UEMP15OV	Number of Civilians Unemployed for 15 Weeks & Over	5
358	UEMP15T26	Number of Civilians Unemployed for 15 to 26 Weeks	5
359	UEMP27OV	Number of Civilians Unemployed for 27 Weeks and Over	5
360	UEMP5TO14	Number of Civilians Unemployed for 5 to 14 Weeks	5
361	UEMPLT5	Number of Civilians Unemployed - Less Than 5 Weeks	5
362	UEMPMEAN	Average (Mean) Duration of Unemployment	5
363	UNRATE	Civilian Unemployment Rate	2
364	USFIRE	All Employees : Financial Activities	5
365	USGOOD	All Employees : Goods-Producing Industries	5
366	USGOVT	All Employees : Government	5
367	USTPU	All Employees : Trade, Transportation & Utilities	5
368	USTRAD	All Employees : Retail Trade	5
369	USWTRAD	All Employees : Wholesale Trade	5
370	CE16OV	Civilian Employment	5
371	USCONS	All Employees : Construction	5
372	HOABS	Business Sector : Hours of All Persons	5
373	HOANBS	Nonfarm Business Sector : Hours of All Persons	5
374	PRS84006023	Business Sector : Average Weekly Hours	2
375	PRS85006023	Nonfarm Business Sector : Average Weekly Hours	2
376	AWHMAN	Average Weekly Hours of Production and Nonsupervisory Employees :Manufacturing	2
377	AWOTMAN	Average Weekly Overtime Hours of Production and Nonsupervisory Employees : Manufacturing	2
378	HOUST	Housing Starts : Total : New Privately Owned Housing Units Started	4
379	HOUSTMW	Housing Starts in Midwest Census Region	4
380	HOUSTNE	Housing Starts in Northeast Census Region	4
381	HOUSTS	Housing Starts in South Census Region	4
382	HOUSTW	Housing Starts in West Census Region	4
383	PERMIT	New Private Housing Units Authorized by Building Permits	4
384	PERMITMW	New Private Housing Units Authorized by Building Permits in the Midwest Census Region	4
385	PERMITNE	New Private Housing Units Authorized by Building Permits in the Northeast Census Region	4
386	PERMITW	New Private Housing Units Authorized by Building Permits in the West Census Region	4
387	PERMITS	New Private Housing Units Authorized by Building Permits in the South Census Region	4
388	A371RX1Q020SBEA	Real private inventories	5
389	A373RX1Q020SBEA	Real private inventories : Nonfarm	5
390	B372RX1Q020SBEA	Real private inventories : Farm	5
391	NAPMII	ISM Manufacturing : Inventories Index ^f	1
392	NAPMNOI	ISM Manufacturing : New Orders Index ^f	1
393	NAPMSDI	ISM Manufacturing : Supplier Deliveries Index ^f	1

TABLE 28 – suite

#	Code mnémorique	Nom	Transformation
394	NAPM	ISM : PMI Composite Index	1
395	CMRMTSPL	Real Manufacturing and Trade Industries Sales	5
396	SP500	S&P 500 Stock Price Index	5
397	NETFI	Balance on Current Account, NIPAs	2
398	AAA	Moody's Seasoned Aaa Corporate Bond Yieldf	2
399	BAA	Moody's Seasoned Baa Corporate Bond Yieldf	2
400	FEDFUNDS	Effective Federal Funds Rate	2
401	GS1	1-Year Treasury Constant Maturity Rate	2
402	GS10	10-Year Treasury Constant Maturity Rate	2
403	GS5	5-Year Treasury Constant Maturity Rate	2
404	TB3MS	3-Month Treasury Bill : Secondary Market Rate	2
405	TB6MS	6-Month Treasury Bill : Secondary Market Rate	1
406	AAA-S	AAA-FEDFUNDS spread	1
407	BAA-S	BAA-FEDFUNDS spread	1
408	GS1-S	1-Year Treasury Constant Maturity Rate-FEDFUNDS spread	1
409	GS10-S	10-Year Treasury Constant Maturity Rate-FEDFUNDS spread	1
410	GS5-S	5-Year Treasury Constant Maturity Rate-FEDFUNDS spread	1
411	TB3MS-S	3-Month Treasury Bill : Secondary Market Rate-FEDFUNDS spread	1
412	TB6MS-S	6-Month Treasury Bill : Secondary Market Rate-FEDFUNDS spread	1
413	AMBSL	St. Louis Adjusted Monetary Base	5
414	BUSLOANS	Commercial and Industrial Loans, All Commercial Banks	5
415	M1SL	M1 Money Stock	5
416	M2REAL	Real M2 Money Stock	5
417	M2SL	M2 Money Stock	5
418	TOTRESNS	Total Reserves of Depository Institutions	5
419	NONREVSL	Total Nonrevolving Credit Owned and Securitized, Outstanding	5
420	REALLN	Real Estate Loans, All Commercial Banks	5
421	MZMSL	MZM Money Stock	5
422	DTCOLNVHFNM	Consumer Motor Vehicle Loans Owned by Finance Companies, Outstanding	5
423	DTCTHFNM	Total Consumer Loans and Leases Owned and Securitized by Finance Companies, Outstanding	5
424	INVEST	Securities in Bank Credit at All Commercial Banks	5
425	CPIAUCSL	Consumer Price Index for All Urban Consumers : All Items	5
426	CPIAPPSL	Consumer Price Index for All Urban Consumers : Apparel	5
427	CPITRNSL	Consumer Price Index for All Urban Consumers : Transportation	5
428	CPIMEDSL	Consumer Price Index for All Urban Consumers : Medical Care	5
429	CUSR0000SAC	Consumer Price Index for All Urban Consumers : Commodities	5
430	CUUR0000SAD	Consumer Price Index for All Urban Consumers : Durables	5
431	CUSR0000SAS	Consumer Price Index for All Urban Consumers : Services	5
432	CPIULFSL	Consumer Price Index for All Urban Consumers : All Items Less Food	5
433	CUUR0000SA0L2	Consumer Price Index for All Urban Consumers : All items less shelter	5
434	PPICMM	Producer Price Index : Commodities : Metals and metal products : Primary non ferrous metals	5
435	PPICRM	Producer Price Index : Crude Materials for Further Processing	5
436	PPIFCG	Producer Price Index : Finished Consumer Goods	5
437	PPIFGS	Producer Price Index : Finished Goods	5
438	PPIIDC	Producer Price Index : Industrial Commodities	5
439	NAPMPRI	ISM Manufacturing : Price	1
440	COMPNFB	Nonfarm Business Sector : Compensation Per Hour	5
441	HCOMPBS	Business Sector : Compensation Per Hour	5
442	COMPRNFB	Nonfarm Business Sector : Real Compensation Per Hour	5
443	RCPHBS	Business Sector : Real Compensation Per Hour	5
444	UMCSENT	University of Michigan : Consumer Sentimentf	1
445	GCEC96	Real Government Consumption Expenditures & Gross Investment	5
446	A553RC1Q027SBEA	National income : Compensation of employees : Wages and salaries : Government	5
447	A782RC1Q027SBEA	Gross government investment	5
448	A084RC1Q027SBEA	Government current transfer payments	5
449	W060RC1Q027SBEA	Government current transfer receipts	5
450	W066RC1Q027SBEA	Government total receipts	5
451	GDPDEF	Gross Domestic Product : Implicit Price Deflator	5
452	CNP16OV	Civilian Noninstitutional Population	5
453	POP	Total Population : All Ages including Armed Forces Overseas	5



1130, rue Sherbrooke Ouest, bureau 1400, Montréal (Québec) H3A 2M8

Tél. : 514-985-4000 • Téléc. : 514-985-4039

www.cirano.qc.ca • info@cirano.qc.ca

Centre interuniversitaire de recherche en analyse des organisations
Center for Interuniversity Research and Analysis on Organizations