

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

PRÉVISION DES PRIX DU LOGEMENT AVEC DES VAR :
L'IMPACT DE L'ADDITION DES EFFETS SPATIAUX

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN ÉCONOMIQUE

PAR
JOSEPH KOKOU

SEPTEMBRE 2008

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je tiens à exprimer toute ma gratitude à mes co-directeurs Douglas James Hodgson et Yvon Fauvel, professeurs au département des sciences économiques à l'Université du Québec à Montréal. Tous deux ont investi beaucoup de temps pour l'aboutissement de ce travail. Leurs conseils lucides et pertinents m'ont permis de donner le meilleur de moi-même. Je les remercie également pour leurs encouragements et leur sympathie.

J'adresse mes remerciements à toute l'équipe enseignante du département des sciences économiques pour la qualité d'enseignement dispensé, ainsi qu'au personnel administratif pour leur dévouement.

Je ne saurais suffisamment remercier mon épouse Valérie Bélanger pour son soutien tout au long de mes études et dans la réalisation de ce document. Je me sens privilégié d'avoir à mes côtés ma petite fille Joyce qui me donne le courage de compléter ce mémoire.

À tous mes amis et tous ceux que j'ai rencontrés durant mon parcours, merci pour votre présence et votre gentillesse.

TABLE DES MATIÈRES

LISTES DES FIGURES	v
LISTE DES TABLEAUX.....	vi
RÉSUMÉ	viii
INTRODUCTION	1
CHAPITRE I	
REVUE DE LITTÉRATURE.....	5
1.1 Déterminants du prix du logement.....	5
1.2 Effets de voisinage dans la détermination du prix du logement	7
1.3 Modélisation du prix du logement	9
CHAPITRE II	
DONNÉES	13
2.1 Variables d'intérêt	13
2.2 Source et échantillon des données.....	14
CHAPITRE III	
MÉTHODOLOGIES ÉCONOMÉTRIQUES.....	16
3.1 L'approche des vecteurs autorégressifs (VAR).....	16
3.2 L'approche des modèles vectoriels à correction d'erreur (VECM).....	20
3.3 L'approche des vecteurs autorégressifs spatiaux (VAR spatial).....	22
CHAPITRE IV	
TESTS DE RACINE UNITAIRE ET DE COINTÉGRATION	29
4.1 Analyse graphique.....	29
4.2 Tests de racine unitaire	32
4.3 Test de cointégration.....	40
CHAPITRE V	
RÉSULTATS	43
5.1 Analyse de la dynamique du prix du logement.....	43
5.2 Efficacité prédictive de trois modèles (VAR, VECM, VAR spatial)	54

CONCLUSION	63
BIBLIOGRAPHIE	65
ANNEXE A	
DONNÉES	65
ANNEXE B	
RMSE ET GAIN D'EFFICACITÉ	71

LISTE DES FIGURES

Figure		Page
3.1	Matrices de poids	25
4.1	L'évolution dans le temps des différentes variables en niveau.....	31
4.2	L'évolution dans le temps des différentes variables en première différence	35
5.1	Fonction de réponse du prix du logement au choc d'emploi	49
5.2	Fonction de réponse du prix du logement au choc de taux d'intérêt	51
5.3	RMSE, cas d'Hamilton	56
5.4	Gain d'efficacité, cas d'Hamilton	57
5.5	RMSE, cas de Montréal	58
5.6	Gain d'efficacité, cas de Montréal	58
5.7	RMSE, cas de Winnipeg	59
5.8	Gain d'efficacité, cas de Winnipeg	60
5.9	RMSE, Cas de Calgary	61
5.10	Gain d'efficacité, cas de Calgary	61

LISTE DES TABLEAUX

Tableau		Page
3.1	Distance en ligne directe entre Toronto et les différentes villes	24
4.1	Test ADF sur log_IPLN	33
4.2	Test ADF sur log_emploi.....	33
4.3	Test ADF sur le taux d'intérêt réel.....	34
4.4	Test ADF sur dlog_IPLN.....	36
4.5	Test ADF sur dlog_emploi.....	36
4.6	Test Phillips-Perron sur log_IPLN.....	37
4.7	Test Phillips-Perron sur log_emploi	37
4.8	Test Phillips-Perron sur le taux d'intérêt réel	37
4.9	Test ERS sur log_IPLN.....	38
4.10	Test ERS sur log_emploi	38
4.11	Test ERS sur le taux d'intérêt réel	38

4.12	Test ERS sur dlog_IPLN.....	39
4.13	Test ERS sur dlog_emploi	39
4.14	Test de cointégration de Engle-Granger	41
4.15	Estimation de la relation de cointégration entre log_IPLN et log_emploi	42
5.1	Estimation du VAR	44
5.2	Estimation du VECM.....	45
5.3	Estimation du VAR spatial.....	46
B.1	RMSE, cas de Montréal.....	71
B.2	Gain d'efficacité, cas de Montréal.....	71
B.3	RMSE, cas d'Hamilton.....	72
B.4	Gain d'efficacité, cas d'Hamilton.....	72
B.5	RMSE, cas de Winnipeg.....	73
B.6	Gain d'efficacité, cas de Winnipeg	73
B.7	RMSE, Cas de Calgary.....	74
B.8	Gain d'efficacité, cas de Calgary.....	74

RÉSUMÉ

Ce mémoire examine la qualité de prévision des prix immobiliers au Canada et dans cinq de ses régions métropolitaines, en utilisant trois modèles économétriques: VAR (Modèle d'autorégression vectorielle), VECM (Modèle vectoriel à correction d'erreur), VAR spatial (Modèle d'autorégression vectorielle spatiale). Le VAR spatial permet d'incorporer les effets spatiaux dans la modélisation.

Les résultats des estimations du VAR, du VECM et du VAR spatial et les fonctions de réponse permettent de faire les constats suivants:

- (i) un choc d'emploi affecte positivement le prix du logement;
- (ii) un choc de politique monétaire (taux d'intérêt) a un impact négatif sur le prix immobilier.

Pour mesurer la qualité prédictive des trois approches, nous avons fait appel à deux mesures de choix de modèle : la Racine de l'Erreur Quadratique Moyenne (*Root Mean Squared Error, RMSE*) et le gain d'efficacité (*Relative Root Mean Squared Error*). Les résultats indiquent que:

- (i) le VAR spatial donne une prévision de meilleure qualité comparativement au VECM et au VAR conventionnel;
- (ii) l'incorporation de l'interaction entre deux régions dans un modèle VAR permet de mieux prévoir le prix immobilier.

En somme, le taux d'intérêt et l'emploi demeurent des facteurs-clés dans la modélisation du prix du logement. De plus, l'approche VAR spatial permet d'améliorer la qualité de la prévision de cette variable d'intérêt.

Mots-clés : économétrie, prévision, prix immobilier, VAR spatial, VECM, VAR

INTRODUCTION

Les prix immobiliers ont connu depuis 1970 une montée fulgurante dans les pays de l'O.C.D.E. Le Canada n'est pas du reste. Selon Statistique Canada, le prix annuel moyen des logements neufs est passé de 150 720\$ CAN en 1995 à 227 219\$ CAN en 2004, soit une augmentation de 50,76%. Au cours de la même période, l'Indice des Prix à la Consommation (IPC) a crû de 19,20%, passant de 103,1 à 122,9.

Le prix moyen des logements existants progresse environ de 10% par année depuis 2002. Les prix immobiliers poursuivront leur progression en 2008 et dans les années à venir. Selon la revue *L'Observateur du Logement au Canada* (2007), le taux de croissance du prix des logements neufs atteint cette année 11,9%, son plus haut niveau depuis 17 ans. D'aucuns craignent, d'ailleurs, l'apparition d'une bulle dans les prix immobiliers.

De plus, l'augmentation des prix immobiliers n'est pas identique dans toutes les régions. Tsatsaronis et Zhu (2004) ont noté dans une étude sur la dynamique des prix immobiliers réalisée dans 17 pays qu'il existe des différences notables dans les prix immobiliers suivant les différents pays. Fauvel (2005), dans une étude sur les grandes agglomérations urbaines canadiennes préparée pour le compte de la Société Canadienne d'Hypothèque et de Logement (SCHL), est arrivé au même constat.

L'envolée des prix dans ce secteur a ainsi suscité un intérêt particulier chez les chercheurs. Des travaux ont donc cherché à répertorier les plus importants facteurs qui contrôlent les prix immobiliers, à expliquer les différences entre ces prix suivant les différentes régions, et à quantifier l'effet des chocs à travers les différents déterminants du prix du logement.

Parmi les facteurs influant sur les prix du logement, on trouve à travers la littérature: le taux d'intérêt, l'emploi ou le revenu disponible des ménages, la taille de la population, le coût du terrain, le coût de construction, etc.

Cependant, pour quantifier la part de chaque facteur dans la détermination du prix du logement, bien des travaux se sont heurtés à des résultats parfois surprenants, soit plus spécifiquement, la faible contribution de l'emploi dans la détermination du prix immobilier. Cette conclusion sur le rôle mineur de l'emploi a été corroborée par le travail de Tsatsaronis et Zhu (2004), puis confirmée par celui de Fauvel (2005). Par contre, d'autres études, comme celle de Baffoé-Bonie (1998) sur les données américaines, soutiennent que le marché immobilier est très sensible à la croissance de l'emploi.

Or, certains facteurs explicatifs des prix du logement, tel que la mobilité de l'emploi, ont une forte dimension géographique. Mais les chercheurs en font souvent abstraction. Il est cependant logique de prendre en compte le rôle du voisinage dans la détermination du prix immobilier. Cela pourrait vraisemblablement améliorer la robustesse des résultats. De plus, l'omission de l'effet spatial dans le modèle économétrique peut détériorer sa capacité de prévision.

L'objectif de ce travail est d'appliquer les modèles de l'économétrie spatiale au cas canadien, pour voir si la qualité de la modélisation du prix du logement peut être améliorée. Plus spécifiquement, il sera question *d'une part, d'analyser la dynamique du prix du logement en introduisant l'effet spatial, et d'autre part, de vérifier la capacité prédictive du VAR Spatial.*

Pour atteindre cet objectif, nous comparerons le VAR Spatial au VAR conventionnel et au VECM. L'étude se penche sur cinq régions métropolitaines du

Canada (Toronto, Hamilton, Montréal, Winnipeg et Calgary). Les données portent sur les prix du logement, le niveau de l'emploi, et le taux d'intérêt pour la période allant de 1969 à 2007. Les données sont trimestrielles.

Pour analyser la dynamique du prix, nous aurons recours aux estimations des VARs ainsi qu'aux fonctions de réponses suivant les différents chocs. Afin de mesurer la capacité de prévision, nous utiliserons la Racine de l'Erreur Quadratique Moyenne (*Root Mean Squared Error, RMSE*) et le gain d'efficacité (*Relative Root Mean Squared Error*) pour trois modèles économétriques: VAR (*Vector Autoregressive*), VECM (*Vector Error Correction Model*) et VAR Spatial (*Spatial Vector Autoregressive*).

Le présent document est subdivisé de la façon suivante :

Le chapitre 1 passe d'abord en revue les travaux sur les déterminants des prix immobiliers. Ensuite, sur la base d'intuition et de quelques travaux, elle souligne l'importance des effets du voisinage dans la détermination des prix du logement. Enfin, elle identifie les approches économétriques les plus utilisées dans la littérature pour la détermination de ces prix.

Le chapitre 2 détaille les caractéristiques des données qui seront utilisées. En effet, cette section précise la source et la nature des données, puis en quoi elles sont suffisantes pour saisir la problématique.

Au chapitre 3, nous décrivons les différents modèles économétriques que nous utiliserons. Dans cette dernière section, nous montrons les avantages, les inconvénients et les motivations qui ont guidé le choix de ces méthodes.

Au chapitre 4, nous examinons nos variables d'intérêt, en particulier nous effectuons les différents tests de racine unitaire et de cointégration.

Nous analysons les résultats au chapitre 5. L'attention sera portée particulièrement sur la réponse du prix du logement aux différents chocs et l'efficacité prédictive des différents modèles.

CHAPITRE I

REVUE DE LITTÉRATURE

1.1 Déterminants du prix du logement

Selon Statistique Canada, les dépenses des logements occupent une part substantielle du revenu disponible du ménage, et cette part ne cesse de croître au fil des ans. Cette progression des dépenses des logements est due, du moins en partie, à l'accroissement des prix du logement. Une importante littérature s'est donc intéressée à l'évolution de ces prix. En particulier, les études ont cherché à répertorier les différentes variables qui influenceraient la formation et l'évolution du prix immobilier. Parmi ces variables, le revenu réel ou l'emploi, le taux d'intérêt nominal sur les prêts hypothécaires et la démographie exercent une influence significative (Fortin et Leclerc, 2002).

Taux d'intérêt

Des *taux d'intérêt* bas feraient monter le prix de l'immobilier. En effet, des taux d'intérêt bas permettraient aux ménages d'avoir accès aux prêts et ceci pour de longues durées. Il en résulterait une augmentation de la demande des logements. Ce qui se traduirait par l'inflation des prix immobiliers. Englund et Ioannides (1997) documentent la relation négative entre le prix du logement et le taux d'intérêt. Quant à Fauvel (2005), en cherchant à apporter une lumière sur les variations du prix du logement pour la période 1972-2003 au niveau du Canada, il arrive à la conclusion qu'« une politique monétaire restrictive haussant les taux d'intérêt réels aura une incidence négative et substantielle sur les prix réels des logements ».

Emploi

Le lien entre le marché de *travail* et le développement des habitations est, sans aucun doute, évident. En France par exemple, historiquement, la création du logement social est consécutive à l'essor industriel. En effet, l'expansion de l'emploi se traduit par l'augmentation des revenus. Des revenus disponibles de plus en plus élevés, les ménages peuvent ainsi se procurer des maisons. Un accroissement de la demande et du prix du logement s'en suit. Il est donc plausible de noter l'évidence d'un lien entre l'emploi et le prix immobilier. C'est ce qui fait dire à Fortura et Kushner (1986) que l'emploi (le revenu réel) est un facteur explicatif non négligeable dans la détermination du prix du logement.

Démographie

La première étude qui a montré le fort impact de la *densité de la population* sur les prix du logement est de celle Mankiw et Weil (1989) sur les données américaines. En construisant un indice démographique, ils identifient une relation positive entre cet indice et les prix du logement. Mais, cette conclusion a été sévèrement critiquée par d'autres recherches (Hamilton, 1991). Récemment, Di Pasquale et Wheaton (1994), Fortin et Leclerc (2000) ont montré que la démographie a une part plausible dans les fluctuations des prix immobiliers.

Autres déterminants

Fauvel (2005) présente une littérature plus exhaustive sur les déterminants des prix des logements dans les régions métropolitaines. Ainsi, les mises en chantier, les coûts réels de construction, le volume des transactions entre acheteurs et les vendeurs de logements, la disponibilité des terrains, la sévérité des réglementations environnementales, les cours boursiers, etc sont identifiés comme d'autres variables explicatives des prix des logements.

1.2 Effets de voisinage dans la détermination du prix du logement

Parmi les déterminants précités du prix du logement, l'emploi aurait une dimension géographique. Par exemple, aux localisations proches les unes des autres, le développement de l'activité économique ou de l'emploi dans l'une affecterait l'autre.

Considérons au départ qu'il y a une répartition homogène des activités économiques dans différentes régions de l'espace. Un choc de l'emploi en un point donné de cet espace va se diffuser progressivement vers les autres régions. En effet, l'expansion de l'emploi à Calgary par exemple affecterait, certainement, le niveau du taux de chômage d'Edmonton. Dans une étude sur la convergence des économies européennes Beaumont (2000) souligne que certains facteurs de la croissance économique, entre autres, la mobilité du travail, peuvent expliquer la force des interactions entre les régions.

La proximité entre les régions influencera donc les comportements des consommateurs et des entreprises. Ainsi, les décisions d'offre d'emploi et de résidence des ménages, de même que les décisions de demande d'emploi et d'établissement des entreprises devraient dépendre des conditions du marché local d'une région relatives aux conditions dans les autres régions. À ce titre, l'incorporation de l'effet de proximité permettrait d'apporter un éclairage nouveau sur la contribution des différents facteurs dans la formation et l'évolution du prix des logements.

Dans la littérature, plusieurs auteurs ont pris en compte cet effet de proximité dans l'analyse de la dynamique du prix des logements.

C'est ainsi que Beenstock et Felsentein (2005), à partir d'un modèle VAR spatial, analysent les fonctions de réponses du prix de logement aux différents chocs dans une région et leur propagation dans d'autres régions d'Israël. Ils constatent que la structure spatiale du modèle permet de capter cette propagation et la matrice de poids détermine l'amplitude de la fonction de réponse dans les autres régions.

Case et al. (2004), pour expliquer le comportement des prix du logement, comparent quatre modèles dont certains prennent en compte les effets spatiaux. Cette étude rapporte que les modèles qui prennent en compte la dépendance spatiale ont une capacité de prédiction plus élevée.

Gelfand et al. (2004), motivés par l'impact que la proximité pourrait avoir sur l'évolution temporelle des prix du logement, proposent un modèle spatio-temporel pour expliquer la dynamique des prix du logement. Ils trouvent que la composante spatiale du modèle est très importante dans l'explication des prix du logement.

Holly et al. (2006) modélisent la dynamique d'ajustement du prix réel des logements sur des données américaines à des chocs macroéconomiques. Ils utilisent des modèles économiques qui prennent en compte la dépendance et l'hétérogénéité spatiales afin d'examiner le rôle des facteurs spatiaux de la dynamique du prix des logements.

Lesage et al. (2004), en utilisant des données qui exhibent une dépendance spatiale, comparent des modèles hédoniques qui ignorent cette dépendance spatiale à ceux qui l'incorporent. Ils arrivent au constat que les derniers améliorent la significativité des résultats des estimations. Cette démonstration a été faite aussi bien sur les données issues d'une expérience Monte Carlo que sur des données réelles des prix du logement.

Militino et al. (2004) explorent différents modèles faisant appel à des techniques de la statistique spatiale pour modéliser les prix du logement. Différentes structures spatiales, ainsi que différentes matrices de poids ont été utilisées.

Il existe ainsi une littérature exhaustive de l'impact des effets spatiaux sur l'explication et la dynamique des prix du logement. Des études ont également analysé ou pris en compte ces effets dans l'explication d'autres agrégats économiques. Ceci démontre l'intérêt de la littérature économique récente aux modèles de l'économétrie spatiale.

1.3 Modélisation du prix du logement

Simple régression linéaire

Dans la littérature, l'approche la plus utilisée pour évaluer les déterminants du prix du logement consiste à écrire ce dernier comme une fonction des facteurs d'offre et de demande de logements. Ainsi, les études sur les prix du logement font intervenir, en général, comme variables explicatives le taux d'intérêt réel, la population, le revenu (l'emploi), le stock de logements, le coût de construction.

VAR (Vector Autoregressive)

Des auteurs comme Smith et al (1988), Meen (2001) et Fauvel (2005) ont montré que les marchés immobiliers présentent, par contre, une certaine persistance dans le temps. Conséquence, il faut donc en tenir compte dans la spécification des modèles. Cependant, des modèles temporels à équation unique et à variables explicatives exogènes pour expliquer l'évolution des prix du logement comportent des lacunes. Ils ne tiennent pas compte des rétroactions des prix du logement sur les variables explicatives. Dans la littérature, une façon de tenir compte de ces particularités dans la détermination des prix du logement, consiste à utiliser les modèles d'autorégression vectorielle (VAR).

Ainsi, Tsatsaronis et Zhu (2004) ont eu recours à un modèle VAR pour expliquer la dynamique des prix immobiliers dans 17 pays industrialisés. Ce qui leur a permis de détecter l'importance d'un certain nombre de facteurs macroéconomiques influençant la dynamique de ces prix immobiliers.

Pour sa part, Iacoviello (2000), en se basant sur l'approche SVAR (méthode structurelle d'autorégression vectorielle), identifie les plus importants facteurs macroéconomiques qui contrôlent les fluctuations des prix du logement dans six pays européens (France, Allemagne, Italie, Espagne, Suède et Angleterre). Il souligne aussi l'importance des chocs monétaires dans les fluctuations des prix du logement et les mécanismes de propagation macroéconomique qui en découlent.

VECM (Vector Error Correction Model)

De nombreuses séries macroéconomiques et financières sont non stationnaires et peuvent partager des relations stables à long terme. Les modèles VAR sont visiblement incapables d'en rendre compte. Pour cela, une classe particulière de modèles, soient les modèles vectoriels à correction d'erreur (VECM) ont été largement utilisés. Meen (2001) a utilisé ce type de modèle pour expliquer les différences des prix réels de logements entre différents pays.

VAR Spatial (Spatial Vector Autoregressive)

Les précédents travaux basés soit sur une régression simple, soit sur l'approche VAR ou encore sur l'approche VECM, bien que révélateurs sur la dynamique des prix du logement, ignorent complètement l'effet de proximité et son importance sur la dynamique des prix du logement. De nouvelles approches qui exploitent plus ou moins les interactions spatiales entre les régions ont donc fait leur apparition.

Muellbauer et Murphy (1997) notent des différences dans l'évolution des prix du logement régionaux. Ce travail souligne que ces différences peuvent être attribuées aux variations dans les caractéristiques régionales. Ce travail, par contre, ne jette aucun regard sur l'effet de voisinage entre les régions.

Gelfand et al. (2004), motivés par l'idée d'expliquer efficacement la dynamique des prix des maisons, proposent quatre modèles dans lesquels ils introduisent l'effet de proximité. Ils arrivent à la conclusion que la composante spatiale a une part importante dans l'explication des prix immobiliers. Cette étude, par contre, est basée sur les microdonnées (distances entre les quartiers, accessibilité au centre d'emploi, qualité sociale du voisinage, etc.).

En utilisant toujours les microdonnées et des modèles contenant une composante spatiale, Case et al. (2004), rapportent le même résultat que précédemment.

D'autres travaux par contre, en se basant les données macroéconomiques, exploitent les interactions spatiales entre les régions. Carlino et Defina (1998), en analysant les effets de la politique monétaire sur l'output, concluent qu'il est important de tenir compte à la fois des canaux de transmission traditionnels, mais aussi de la transmission géographique.

Bhattacharjee et Jensen-Butler (2005), en utilisant la technique des équations économétriques simultanées, construisent un modèle dans lequel ils incorporent l'hétérogénéité et l'interaction spatiales observées à travers différentes régions. Ce modèle reproduit les variations passées de demande et de prix observées sur le marché de logements. De plus, cette approche a permis d'évaluer l'impact asymétrique des chocs sur le marché des logements suivant différentes régions.

Di Giacinto (2003) utilise l'approche SVAR (Structural Vector AutoRegressive) et incorpore l'information géographique en utilisant des techniques de l'économétrie spatiale.

Plus récemment, Beenstock et Felsentein (2005), en appliquant le modèle d'autorégression vectorielle spatiale (VAR Spatial), analysent la dynamique des prix du logement dans 9 régions de l'Israël. Leur travail conclue qu'il est capital de tenir compte de l'effet de proximité dans la détermination des prix immobiliers.

Au Canada, bien qu'il existe des travaux qui ont cherché à identifier les déterminants des prix des logements, en se basant soit sur une simple régression ou sur l'approche VAR conventionnel, il y a cependant peu ou presque pas de travaux qui ont utilisé des approches VECM ou VAR Spatial. La contribution de notre travail est d'appliquer les modèles trouvés utiles dans d'autres pays au cas canadien, pour voir si la qualité de la modélisation peut être améliorée.

CHAPITRE II

DONNÉES

2.1 Variables d'intérêt

Nous avons vu précédemment qu'il existe de nombreuses variables explicatives du prix du logement. Sur la base de la revue de littérature à ce sujet, l'évidence d'une relation tangible entre le prix du logement et l'emploi d'une part, et d'autre part avec le taux d'intérêt n'est plus à démontrer.

Sur ce, nos modèles économétriques contiendront trois variables d'intérêt: le prix nominal du logement, le taux d'intérêt réel et l'emploi. Pour notre application empirique, nous avons recours aux données dont nous présenterons la source et les caractéristiques dans la section suivante.

Il est important de justifier l'utilisation non-conventionnelle du prix nominal du logement. De façon générale, dans un modèle économique, toutes les variables devraient être exprimées en termes réels afin d'éliminer les effets de tendance et de pouvoir bien capter le mouvement des fluctuations. Toutefois, dans notre cas, nous avons plutôt choisi d'utiliser le prix nominal puisque un rapport de deux prix, soit IPLN sur IPC, ne témoigne pas réellement de la dynamique du prix immobilier. Par exemple, une hausse presque proportionnelle de l'IPLN et de l'IPC conduirait à un rapport de prix constant. Dans ce cas, l'apport explicatif des variables du modèle serait presque nul. A cet égard, le prix nominal du logement semble être le meilleur indicateur.

2.2 Source et échantillon des données

L'échantillon de données est constitué à partir des statistiques tirées de Statistique Canada. Cet organe recense et publie diverses données socio-économiques sur le Canada. L'échantillon pour la présente étude est construit à partir de la base de données CANSIM II.

L'étude est réalisée sur un échantillon de données canadiennes à l'échelle nationale, en plus des données de 5 régions métropolitaines du Canada. Il s'agit des régions de Toronto, Hamilton, Montréal, Winnipeg et Calgary. L'échantillon couvre la période de 1969-2007 et les données sont trimestrielles.

De cette base de données, quatre séries ont été extraites: l'indice des prix à la consommation (IPC), l'indice des prix des logements neufs (IPLN), le Bons du Trésor à 3 mois, l'emploi à moyenne mobile de 3 mois. Les tableaux de ces séries sont reportés à l'annexe A.

L'indice des prix à la consommation (IPC). L'IPC est tiré du tableau 326-0001 (v735319). Cet indice sert à calculer l'inflation qui est: $d(\log(\text{IPC}/\text{IPC}(-1)))^1$.

L'indice des prix des logements neufs (IPLN). L'IPLN est une série qui se veut une mesure de l'appréciation des logements neufs. Il est basé sur le prix des logements neufs compte tenu de leur qualité. Cette série permet de construire des mesures précises pour la variable d'intérêt, le prix du logement. Le prix nominal du logement est mesuré par le logarithme de l'indice des prix des logements neufs (\log_IPLN).

Le Bons du Trésor (rendement moyen à 3 mois). Le taux d'intérêt réel est mesuré par le taux de rendement sur les Bons du Trésor à 3 mois, auquel on soustrait

¹ $d(\log(\text{IPC}/\text{IPC}(-1)))$ est la différence première du logarithme de la série IPC. Ceci constitue une approximation du taux d'inflation (variation en pourcentage de la série IPC).

l'inflation précédemment calculée à partir de l'IPC.

L'emploi (moyennes mobiles de 3 mois, désaisonnalisées en milliers personnes). Il est tiré du tableau 282-0090 (v2067395). Nous prenons également le logarithme de l'emploi (`log_emploi`).

CHAPITRE III

MÉTHODOLOGIES ÉCONOMÉTRIQUES

Jusqu'à présent, nous avons porté l'attention sur les données et les variables. Nous passons à présent vers les modèles économétriques utilisés.

Le but fondamental de ce travail est d'appliquer les modèles de l'économétrie spatiale pour voir si la qualité de la modélisation du prix du logement peut être améliorée. Pour se faire, nous mesurerons la capacité prédictive du VAR spatial. Ce dernier sera ensuite comparé à deux autres approches des vecteurs autorégressifs : le VAR et Le VECM. Dans les lignes qui suivent, nous nous proposons de décrire ces trois approches de modélisation.

3.1 L'approche des vecteurs autorégressifs (VAR)

Un VAR: de quoi s'agit-il?

Un VAR est une forme de modèle dynamique linéaire. Il comporte plusieurs équations. Chaque équation dénote une relation linéaire où une variable est exprimée comme une combinaison de ses propres valeurs passées et des valeurs passées des autres variables. L'ensemble de ces variables du modèle sont endogènes. Chaque équation est complétée par un terme d'erreur qui est soit endogène, soit exogène.

Avant de présenter les avantages liés à l'utilisation de la méthodologie VAR pour le marché immobilier, nous avons jugé important de spécifier un modèle VAR.

Formulation économétrique d'un modèle VAR

Dans un modèle VAR conventionnel, chaque variable est fonction de ses propres valeurs passées mais aussi des valeurs passées des autres variables. Dénotons par y_{jkt} , la variable j dans la région k au temps t , où $j = 1, 2, \dots, n$; $k = 1, 2, \dots, N$ et $t = 1, 2, \dots, T$. Un modèle VAR d'ordre q prend la forme suivante:

$$y_{jkt} = \alpha_{jk} + \sum_{i=1}^q \beta_{1i} y_{1kt-i} + \sum_{i=1}^q \beta_{2i} y_{2kt-i} + \dots + \sum_{i=1}^q \beta_{ni} y_{nkt-i} + \varepsilon_{jkt}.$$

De façon plus simplifiée:

$$y_{jkt} = \alpha_{jk} + \sum_{j=1}^n \sum_{i=1}^q \beta_{ji} y_{jkt-i} + \varepsilon_{jkt}.$$

Le modèle VAR suppose que les effets dynamiques sont les mêmes dans les k régions et que les effets interrégionaux sont absents. En effet, il existe deux façons d'inclure les effets interrégionaux. La première façon est d'introduire une ou plusieurs variables appelées variables décalées² dans le modèle. La seconde façon est de spécifier un processus spatial pour les erreurs. Par exemple, on peut supposer que les erreurs suivent un processus autorégressif.

Or, dans un VAR conventionnel, il n'y a pas de variables décalées. De plus, les termes d'erreur sont des bruits blancs non corrélés (dans un modèle de prévision) ou des chocs corrélés (dans un modèle structurel).

Pourquoi un modèle VAR pour le marché immobilier?

La modélisation VAR a été introduite par Sims (1980) et elle est utilisée dans la prévision, l'inférence structurelle et dans l'analyse des politiques (Stock et Watson, 2001).

² Une description détaillée de cette variable est présentée plus loin dans le document.

Une première qualité des modèles VAR est qu'ils permettent d'appréhender le comportement dynamique de variables dépendantes linéairement du passé. De plus en tenant compte des relations entre plusieurs variables, les modèles VAR permettent ainsi d'expliquer, et non plus uniquement de décrire l'évolution d'une série.

La seconde qualité des VAR est qu'ils évitent d'avoir à décider quelles sont les variables exogènes et endogènes du modèle. Ils comportent seulement des variables endogènes. Le troisième avantage est que les VAR sont des modèles empiriques. Ainsi, seules les données décident des relations possibles entre les variables. L'intervention du chercheur pour poser des restrictions ou des hypothèses est quasi absente (Sims, 1980). Les seules et uniques restrictions dans un VAR sont le choix de variables à inclure dans le modèle et le nombre de retards. Notons que ces deux derniers avantages sont essentiellement vrais pour les modèles de prévision.

Des modèles structurels, quant à eux, reposent sur les restrictions théoriques a priori des chercheurs. Mais l'efficacité des estimateurs et la précision des prévisions sont améliorées si on impose des restrictions raisonnables. Outre sa capacité prédictive, le modèle structurel permet aussi d'identifier et de quantifier la contribution des différents chocs dans les fluctuations des variables incluses dans le modèle. Ainsi, Christiano, Eichenbaum et Evans (2005) appliquent l'approche VAR pour dériver les faits stylisés suivant un choc monétaire contractionniste. Gali (1999) utilise un VAR dans un célèbre article sur la relation entre le progrès technologique et la dynamique du marché de travail, pour produire des preuves empiriques sur l'effet d'un choc technologique positif sur les variables réelles.

Il existe une large littérature où les VAR sont utilisés pour modéliser les prix immobiliers. Outre les articles précédemment cités dans la section de la revue de

littérature, rappelons en plus les travaux de Baffoe-Bonie (1998) et de Lastrapes (2002). Baffoe-Bonie (1998), à partir d'un modèle VAR, analyse les effets de l'emploi, du taux d'intérêt hypothécaire, de la masse monétaire et de l'indice des prix à la consommation sur les prix des logements et le stock des maisons. Il constate que le marché immobilier est très sensible au taux d'intérêt hypothécaire et à la croissance de l'emploi, tant au niveau national que régional. Lastrapes (2002) dans un modèle VAR comprenant les prix réels des logements, les ventes de maisons (neuves et existantes), la masse monétaire, le taux de rendement sur les bons du Trésor à 3 mois, le taux hypothécaire à long terme, la production industrielle et un indice des prix des matières premières, identifie les canaux de transmission de la politique monétaire au prix des logements.

Inconvénients des modèles VAR

Nous avons présenté précédemment les avantages qui sont liés aux modèles VAR. Cependant, l'utilisation des méthodes VAR n'est pas dépourvue d'inconvénients. Outre les difficultés liées au choix des variables, la principale limite des VARs se réfère à l'identification des chocs.

En principe, les modèles VAR comportent plusieurs variables et chaque variable est fonction de ses propres valeurs passées, des valeurs passées et présentes des autres variables mais aussi d'un terme d'erreur. Un choc sur un terme d'erreur affectera immédiatement la variable à laquelle il est associé puis, par la suite, les autres variables. Cette relation choc-terme d'erreur n'est possible que si le présent terme d'erreur n'est pas corrélé de façon contemporaine avec les autres termes d'erreur du système.

Cependant, les termes d'erreur sont, en pratique, généralement corrélés. Ils ont donc une composante commune qui ne peut pas être associée à une variable spécifique. Afin d'analyser la réponse d'une variable aux différents chocs et la part

de chaque choc dans les mouvements de la variable, il faut que les termes d'erreur soient non corrélés entre eux. Si ce n'est pas le cas, il faut répartir les parties communes des termes d'erreur entre les différentes variables. Sinon, l'analyse de la propagation d'un choc n'est pas possible.

Heureusement, la littérature fournit plusieurs méthodes d'identification des chocs. La plus utilisée est celle de la décomposition de Cholesky. Il s'agit d'une méthode mathématique qui consiste à imposer un nombre de restrictions quant à l'impact contemporain du choc sur certaines variables du système. Cette méthode ne requiert que le choix de l'ordre des séries. L'ordonnancement devrait répondre à des a priori économiques forts ou à des faits stylisés robustes. On trouve ici la principale faiblesse de cette méthode. On peut en effet se demander pourquoi, par exemple, un choc sur la demande ne doit pas être corrélé à un choc sur l'offre. Malheureusement, la réponse donnée par cette méthode est parfois arbitraire et sans fondements économiques.

Dans le cadre de ce travail, notre objectif est de définir principalement un modèle de prévision des prix du logement et non un modèle structurel de marché. L'analyse des VAR portera ainsi essentiellement sur l'efficacité et la qualité des estimations. De ce fait, le problème d'identification ne se pose pas. Une autre manière de contourner le problème d'identification, est d'utiliser le VECM. Nous décrivons cette méthode dans la section suivante.

3.2 L'approche des modèles vectoriels à correction d'erreur (VECM)

Un VECM: de quoi s'agit-il

Les modèles VAR font l'hypothèse que les séries sont stationnaires. En général, les séries macroéconomiques et financières sont non stationnaires. Les différencier suffisamment permet de les stationnariser. Cette opération comporte,

cependant, des limites en particulier si les variables partagent une ou plusieurs relations stables à long terme. Dans ce cas, une classe particulière de modèles est utilisée : les modèles vectoriels à correction d'erreurs (VECM).

Formulation économétrique d'un modèle VECM

Contrairement au modèle VAR, un VECM intègre, en plus les relations de court terme entre les variables, une relation de long terme. La différence avec le VAR se trouve dans l'introduction d'un terme supplémentaire pour capter la relation de long terme.

Les séries doivent être non stationnaires d'ordre un. Leur combinaison linéaire, en revanche, doit être stationnaire³.

À titre d'exemple, nous reproduisons un VECM à deux variables x_t et y_t comme suit. Nous supposons les deux variables intégrées d'ordre un et cointégrées.

$$\begin{aligned}\Delta x_t &= \lambda_1 z_{t-1} + \sum_i \alpha_i \Delta x_{t-i} + \sum_j \beta_j \Delta y_{t-j} + \varepsilon_{x_t} \\ \Delta y_t &= \lambda_2 z_{t-1} + \sum_i \alpha'_i \Delta x_{t-i} + \sum_j \beta'_j \Delta y_{t-j} + \varepsilon_{y_t} .\end{aligned}$$

Les termes ε_t sont deux bruits blancs. z_{t-1} est le résidu de la relation de cointégration et représente la variable de long terme. λ est la force de rappel vers la cible de long terme.

Comment détecte-t-on une relation de cointégration? Deux approches sont utilisées : le test à deux étapes de Engle et Granger (1987) et les tests de Johansen (1988). La première approche convient à un modèle où il existe une seule relation de cointégration. La seconde est plus générale.

³ De façon générale, deux séries I(d) sont dites cointégrées CI(d,b), si leur combinaison linéaire est I(d-b), avec $0 < b \leq d$.

Pourquoi un modèle VECM pour le marché immobilier?

Depuis 1970 au moins, les prix réels des logements ont fluctué autour d'une hausse tendancielle. Cette évolution est généralement due aux facteurs influant l'offre de logements, à la croissance démographique et à l'augmentation du revenu par habitant (OCDE). On peut alors supposer a priori l'existence d'une relation de long terme entre les prix de logements et le revenu par habitant (emploi). Dans les données canadiennes, les séries des prix immobiliers et de l'emploi (que nous avons représenté plus loin) indiquent aussi une tendance à la hausse. Il se pourrait qu'elles partagent une relation de long terme. Il convient donc de faire une étude de cointégration entre ces variables. Un modèle VECM peut donc tenir compte de cette relation.

Dans la littérature qui concerne la modélisation des variables liées au marché immobilier, des modèles à correction d'erreur (ECM) ont été largement utilisés. Sa variante vectorielle, le VECM, reste cependant peu employée. Meen (2000) développe un modèle VECM dans lequel il introduit les variables endogènes suivantes : les mises en chantier, le prix des logements, le taux d'intérêt de court terme et les coûts de construction. Il dénombre deux relations de cointégration. Son modèle permet d'expliquer les différences internationales observées dans la tendance des prix réels des logements, ainsi que les propriétés cycliques des variables du marché des logements.

À présent, passons au dernier modèle qui serait considéré dans notre étude, soit le VAR Spatial.

3.3 L'approche des vecteurs autorégressifs spatiaux (VAR Spatial)

Autocorrélation spatiale

On parle d'autocorrélation spatiale lorsqu'une variable possède une relation fonctionnelle entre deux points de l'espace. Pour l'emploi, par exemple, et pour deux

régions, l'augmentation du nombre d'emplois dans l'une des régions influencerait possiblement l'emploi dans l'autre région. Il en est de même pour la variable prix du logement.

L'autocorrélation spatiale diffère de l'autocorrélation temporelle. La seconde est unidirectionnelle: seul le passé influence le futur dans un vase clos. En revanche, l'autocorrélation spatiale laisse entrevoir la possibilité que le lien temporel puisse être affecté par le lien spatial. Intuitivement, les circonstances économiques en un lieu donné peuvent être déterminées par des conditions en d'autres lieux.

La littérature sur l'économétrie spatiale fournit une documentation sur la prise en compte de cette interaction. Puisque la réalisation d'une variable donnée d'une région est affectée par l'intensité de l'interaction entre cette région et une autre, il convient d'en tenir compte par des mesures de proximité et de transmission. Dans la littérature, pour capter cette idée, on utilise les concepts de *distance*, *population*, etc. Ainsi, l'autocorrélation spatiale est formalisée à partir de matrices de poids.

Matrice de poids

La matrice de poids est une matrice dont chaque élément représente le niveau de l'interaction entre deux régions. Ses éléments sont exogènes, donc définis a priori par le chercheur. La matrice de poids ne contient pas d'éléments à estimer.

Plusieurs formes fonctionnelles des matrices de poids sont proposées dans la littérature. Beenstock (2005) utilise une forme fonctionnelle qui combine à la fois les mesures de distance et de taille (population) des régions. Plus spécifiquement, il définit un élément de la matrice de poids comme suit:

$$w_{ij} = \frac{1}{d_{ij}} \frac{p_i}{p_i + p_j}$$

où d_{ij} est la distance entre les régions i et j , et p dénote la taille de la population. Cette équation suppose que l'interaction entre deux régions est inversement proportionnelle à la distance entre les deux régions, et dépend directement de la population relative de la région i par rapport à j .

Bavaud (1998) teste une forme de matrice de poids basée uniquement sur les distances entre les régions. Cette formulation suppose que l'interaction entre deux régions est inversement proportionnelle à la distance qui les sépare. Pour rendre comparables les modèles spatiaux, la matrice de poids est standardisée de sorte que les poids soient alors compris entre 0 et 1. Vu la simplicité et la cohérence des résultats que Bavaud (1998) obtient, nous utiliserons cette forme fonctionnelle dans le cadre de cette étude.

Nous définissons ici la distance comme la distance à vol d'oiseau entre la ville de référence (Toronto) et les autres villes considérées dans la présente étude. Pour obtenir les poids, les distances sont alors standardisées⁴. Les distances ainsi que les poids standardisés sont présentés dans le tableau suivant :

Tableau 3.1 Distance en ligne directe entre Toronto et les différentes villes

Villes	Hamilton	Montréal	Winnipeg	Calgary
Distance (en km)	74	494	1510	2706
Poids	0.8	0.6	0.4	0.2

Source: Ressources naturelles Canada

⁴ En fonction de la distance d_i , on suppose la standardisation du poids w_i de la façon suivante : $d_i \leq 100$, $w_i = 0.2$; $100 < d_i \leq 1000$, $w_i = 0.4$; $1000 < d_i \leq 2000$, $w_i = 0.6$; $d_i > 2000$, $w_i = 0.8$.

En se basant sur ces poids⁵, nous définissons comme suit les matrices de poids W suivantes :

Figure 3.1 Matrices de poids

<i>Hamilton</i>	<i>Montréal</i>
log_emploi_ham	log_emploi_mtl
log_emploi_tor	log_emploi_tor
$txreel$	$txreel$
log_ipln_ham	log_ipln_mtl
<i>Winnipeg</i>	<i>Calgary</i>
log_emploi_win	log_emploi_cal
log_emploi_tor	log_emploi_tor
$txreel$	$txreel$
log_ipln_win	log_ipln_cal

,

où log_emploi_x représente le logarithme de l'emploi de la ville x , $txreel$ est le taux d'intérêt réel et log_ipln_x est le logarithme de l'indice des prix des logements neufs (IPLN) de la ville x .

Formulation économétrique d'un modèle VAR Spatial

L'utilité de la matrice de poids est qu'elle permet de relier une observation de la variable dans une région avec les observations de cette variable pour d'autres régions. Explicitement, la matrice de poids permet de générer un nouvelle variable

⁵ Dans notre étude, les poids définis sont attribués à la variable log_emploi . Pour des raisons techniques le poids des autres variables est constant et égal à 0.1. Notons également que les variables qui apparaissent à la verticale de la matrice sont les mêmes qu'à l'horizontale.

appelée *variable spatiale décalée*. La variable spatiale décalée⁶ est définie: $y^* = Wy$. Par exemple, la variable décalée de l'emploi serait la variable emploi ajustée par un coefficient qui représente la distance. Cette nouvelle variable permet d'introduire l'autocorrélation spatiale dans les modèles économétriques.

Contrairement à la spécification VAR conventionnel, le modèle VAR Spatial inclut à la fois les dépendances aussi bien temporelle que spatiale entre les différentes variables du système :

$$y_{jkt} = \alpha_{jk} + \sum_{j=1}^n \sum_{i=1}^q \beta_{ji} y_{jkt-i} + \sum_{j=1}^n \sum_{i=1}^q \lambda_{ji} y_{jkt-i}^* + \varepsilon_{jkt} \quad ,$$

où α_{jk} est une constante et représente l'effet fixe pour la variable j dans la région. Les β et λ sont respectivement les coefficients temporels et spatiaux. λ est le paramètre spatial indiquant l'ampleur de l'interaction entre les différentes régions. Enfin, ε_{jk} dénote les perturbations. L'autocorrélation est introduite dans le modèle par le biais de la variable spatiale décalée y^* .

Il existe d'autres façons d'incorporer l'autocorrélation spatiale dans un modèle économétrique. Par exemple, au lieu de chercher à construire une variable décalée, on peut spécifier un processus spatial pour les erreurs. D'autres auteurs, comme Anselin (1980), ont produit des modèles qui contiennent à la fois une variable décalée et une autocorrélation des erreurs.

⁶ Dans notre projet, la variable décalée y^* est construite avec la matrice de poids W et y . La variable y est \log_emploi puisque nous avons supposé que l'interaction s'effectue au travers de l'emploi. En raison de la mobilité plus prononcée au niveau de l'emploi que nous avons mentionné plus tôt dans la revue de la littérature, nous supposons que les chocs entre régions sont corrélés par une seule variable, l'emploi. Ceci signifie qu'il existe une autocorrélation spatiale uniquement par la variable emploi, mais pas d'autocorrélation spatiale au niveau des autres variables.

Pourquoi un modèle VAR Spatial pour le marché immobilier?

Jusque dans les années 80, les modèles de l'économétrie spatiale étaient réservés et appliqués à des problèmes d'économie régionale. Citons, par exemple, Anselin et Can (1986) qui ont traité des problèmes de développement urbain à partir des techniques de l'économétrie spatiale. Plus récemment, ces modèles ont été étendus à d'autres sujets tels que l'économie publique (Case et al, 1993), l'économie internationale (Aten, 1996), la politique monétaire (Di Giacinto, 2003), la croissance économique (Wilfrid, 2004) et l'emploi (Hernandez-Murillo et al., 2005). Le champ de l'économétrie spatiale devient très vaste. Le Gallo (2000) précisait justement que les méthodes de l'économétrie spatiale sont potentiellement applicables à toutes les études empiriques nécessitant l'utilisation de données spatiales.

L'économie immobilière n'est pas du reste. Dans la section sur revue de la littérature, nous avons mentionné quelques références à ce sujet: Muellbauer et Murphy (1997), Case et al. (2004), Gelfand et al. (2004), Beenstock et Felsentein (2005), etc.

La popularisation des méthodes de l'économétrie spatiale est due aux qualités qu'elles comportent. Outre les qualités pré-citées d'un modèle VAR conventionnel, le modèle VAR Spatial porte en lui deux avantages supplémentaires. Le premier, d'ordre théorique, est l'exploitation de l'interaction spatiale qui pourrait se trouver dans les données régionales pour la détermination du prix du logement. Le second, d'ordre empirique, est la disponibilité d'outils statistiques d'analyse que rend possible la méthode VAR Spatial.

Un exemple introductif

Considérons à titre d'exemple deux villes métropolitaines: Montréal (M) et

Hamilton (H) et une seule variable d'intérêt y . Nous supposons également que l'ordre temporel est de 1. En comparaison avec la forme générale mentionnée plus tôt, $N=2$ et $n=q=1$. Nous supposons de même que la matrice de poids W est de 1. Ce qui implique que la variable spatiale décalée $y^*=y$. L'équation de la variable d'intérêt pour la ville de Montréal est :

$$y_{Mt} = \alpha_M + \beta_M y_{Mt-1} + \lambda_H y_{Ht-1} + \varepsilon_{Mt} .$$

De la même manière, pour la ville de Hamilton la forme de la variable se détermine comme suit:

$$y_{Ht} = \alpha_H + \beta_H y_{Ht-1} + \lambda_M y_{Mt-1} + \varepsilon_{Ht} .$$

Nous supposons que ε_{Mt} et ε_{Ht} sont de moyenne nulle, de variance constante et non corrélées. C'est le cas dans un modèle de prévision mais pas dans un modèle structurel. Si les λ étaient zéro, alors les deux équations constituent le VAR conventionnel. Par contre, si λ sont différents de zéro, le VAR conventionnel devient le VAR Spatial.

CHAPITRE IV

TESTS DE RACINE UNITAIRE ET DE COINTÉGRATION

4.1 Analyse graphique

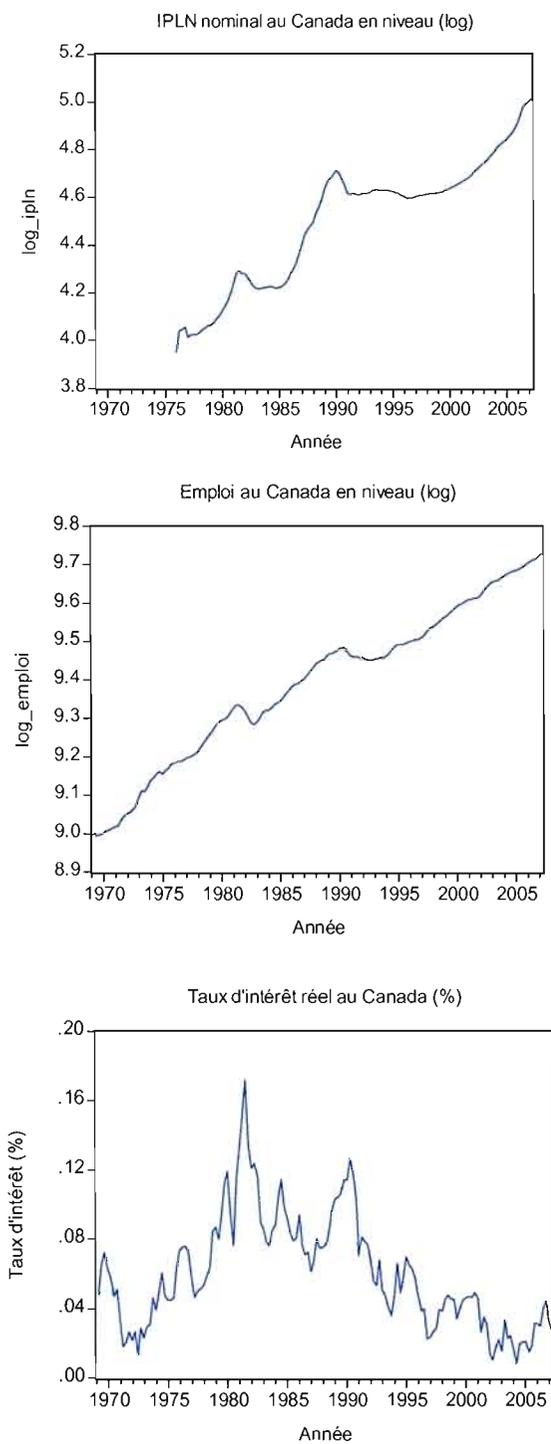
La modélisation des séries temporelles requiert, pour bien des cas, que celles-ci soient stationnaires. En effet, la non-stationnarité a des conséquences sur le plan économétrique. Notamment, les procédures d'inférence statistique ne sont pas valables en cas de non-stationnarité. En plus, les lois asymptotiques standards ne sont plus applicables. Il convient ainsi de vérifier la présence de la non-stationnarité, et dans le cas échéant, la corriger.

Il est possible de détecter la non-stationnarité par des techniques d'analyse graphiques de la série par rapport au temps. Pour cela, nous allons représenter graphiquement les séries. Si on y observe une tendance, alors ces séries sont non stationnaires. Cependant, une telle analyse est préliminaire et informelle, et doit précéder un test formel.

Nous représentons les séries en niveau suivantes au Canada: le logarithme de l'Indice des Prix des Logements Neufs nominal (\log_IPLN), le logarithme de l'Emploi (\log_emploi) et le taux d'intérêt réel.

De la figure 4.1, les séries représentées exhibent une tendance à la hausse ou à la baisse. Elles pourraient être soit non stationnaires de nature déterministe ou soit non stationnaires de nature stochastique. Ces constats doivent être renforcés

par des tests de racine unitaire. Nous utiliserons ainsi trois différents tests de racine unitaire (test de *Augmented Dickey-Fuller*, test de *Phillips-Perron* et test de *Elliot-Rothenberg-Stock*).

Figure 4.1 L'évolution dans le temps des différentes variables en niveau

4.2 Tests de racine unitaire

4.2.1 Test de *Augmented Dickey-Fuller (ADF)*

Le test ADF est un test paramétrique qui repose sur l'estimation d'un processus autorégressif. Le modèle général ADF s'écrit comme suit :

$$\Delta x_t = \lambda + \delta t + \alpha x_{t-1} + \sum_{j=1}^p \gamma_j \Delta x_{t-j} + \varepsilon_t,$$

où λ est la constante, t la mesure la tendance et p le nombre de retards optimal qui permet de blanchir le terme d'erreur.

En fait, Dickey et Fuller (1981) considèrent trois modèles de base: modèle sans constante ni tendance déterministe, modèle avec constante sans tendance déterministe, et modèle avec constante et tendance déterministe. À partir de ces équations, on teste l'hypothèse nulle de racine unitaire contre l'hypothèse alternative d'absence de racine unitaire. L'application du test ADF nécessite la sélection du nombre p de retards.

Sur la base de ce test, nous nous proposons de vérifier la stationnarité des trois variables au niveau national et au niveau des 5 régions métropolitaines. Le nombre de retards est choisi sur la base du critère d'information d'Akaike (AIC).

Les statistiques ADF calculées pour la série \log_IPLN sont reportées au tableau 4.1. Les valeurs critiques de Mackinnon à 1% est de -2.58, -3.49 et -4.04 respectivement pour les modèles sans constante et tendance, avec constante sans tendance et avec constante et tendance. Toutes les statistiques ADF sont supérieures à la valeur critique. Nous ne rejetons donc pas l'hypothèse nulle de racine unitaire, au niveau de significativité de 1%. La série \log_IPLN est non stationnaire.

Pour la série \log_emploi (tableau 4.2), les statistiques ADF calculées sont supérieures aux valeurs critiques. Nous ne rejetons pas l'hypothèse nulle de racine unitaire, au niveau de significativité de 1%. La série \log_emploi est également non stationnaire.

Quant au taux d'intérêt réel, nous rejetons l'hypothèse nulle de racine unitaire. La série de taux d'intérêt réel est donc stationnaire (tableau 4.3).

Tableau 4.1 Test ADF sur \log_IPLN

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Sans constante et tendance	-0.97	-2.57	-2.1	-1.41	-2.24	-1.49
Avec constante sans tendance	-1.85	-3.23	-2.76	-1.72	-3.43	-2.32
Avec constante et tendance	-1.80	-3.33	-2.77	-1.42	-2.84	-1.45

Tableau 4.2 Test ADF sur \log_emploi

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Sans constante et tendance	-0.98	-1.37	-3.36	-1.72	-2.43	-1.74
Avec constante et sans tendance	-2.37	-3.48	-3.16	-2.69	-3.59	-3.10
Avec constante et tendance	-2.42	-3.45	-3.78	-2.77	-3.90	-3.75

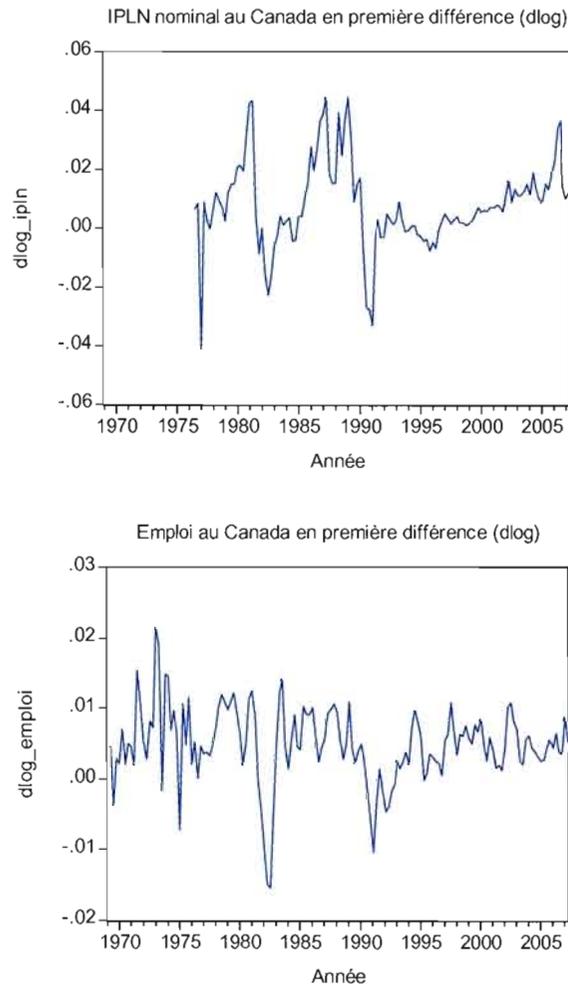
Tableau 4.3 Test ADF sur le taux d'intérêt réel

	Canada
Sans constante ni tendance	-11.42
Avec constante	-11.39
Avec constante et tendance	-11.36

En conclusion, les séries \log_IPLN et \log_emploi sont non stationnaires. La série de taux d'intérêt réel est par contre stationnaire.

Nous allons ainsi vérifier la stationnarité des séries \log_IPLN et \log_emploi en première différence, soient $d\log_IPLN$ et $d\log_emploi$. Comme précédemment, une première intuition peut être fournie par l'étude du graphique. De la figure 4.2, il ressort une absence de tendance dans les deux séries. Ce qui laisse présager que $d\log_IPLN$ et $d\log_emploi$ pourraient être stationnaires.

Figure 4.2 L'évolution dans le temps des différentes variables en première différence



Nous nous proposons de vérifier cette intuition avec le test ADF. Les valeurs critiques de Mackinnon à 1% sont de -2.58, -3.49 et -4.04 respectivement pour les modèles sans constante et tendance, avec constante sans tendance et avec constante et tendance. À la lumière des résultats dans les tableaux ci-dessous, nous rejetons l'hypothèse nulle de racine unitaire au niveau de 1% des séries $dlog_IPLN$ et $dlog_emploi$. En conclusion, les deux séries en première différence sont stationnaires au niveau de significativité de 1%.

Tableau 4.4 Test ADF sur dlog_IPLN

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Sans constante et tendance	-3.54	-4.35	-12.70	-4.39	-3.32	-3.01
Avec constante et sans tendance	-3.53	-4.33	-12.65	-4.41	-4.24	-3.90
Avec constante et tendance	-5.87	-4.28	-12.59	-4.47	-4.82	-4.26

Tableau 4.5 Test ADF sur dlog_emploi

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Sans constante et tendance	-3.13	-3.89	-6.12	-4.55	-3.27	-3.04
Avec constante et sans tendance	-4.82	-3.88	-6.09	-4.51	-3.85	-5.25
Avec constante et tendance	-4.76	-4.11	-6.06	-4.44	-4.54	-5.27

4.2.2 Test de *Phillips-Perron*

Le test de Dickey-Fuller souffre d'un certain nombre de problèmes. Phillips et Perron (1988) proposent une correction non paramétrique afin de régler le problème de l'autocorrélation et/ou de l'hétéroscédasticité des erreurs.

Les valeurs calculées de la statistique de Phillips-Perron pour les séries log_IPLN, log_Emploi et taux d'intérêt réel sont toutes inférieures aux valeurs critiques de Mackinnon à 1% (-3.47 pour le modèle avec constante et -4.01 pour le modèle avec constante et tendance). Nous rejetons ainsi l'hypothèse nulle de racine unitaire, au niveau de significativité de 1%, quel que soit le nombre de retards p .

Sur la base du test de Phillips-Perron, les trois séries en niveau sont stationnaires. Les tableaux ci-contre donnent les valeurs des statistiques pour un test avec 4 retards, à titre d'exemple.

Tableau 4.6 Test Phillips-Perron sur log_IPLN

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Avec constante	-7.08	-6.19	-4.73	-4.22	-8.78	-5.02
Avec constante et tendance	-7.06	-6.34	-4.74	-4.68	-9.04	-5.03

Tableau 4.7 Test Phillips-Perron sur log_emploi

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Avec constante	-6.16	-10.27	-9.47	-10.97	-10.80	-9.56
Avec constante et tendance	-6.21	-10.27	-9.47	-11.04	-10.82	-9.58

Tableau 4.8 Test Phillips-Perron sur le taux d'intérêt réel

	Canada
Avec constante	-11.44
Avec constante et tendance	-11.45

4.2.3 Test *Elliot-Rothenberg-Stock (ERS)*

C'est le dernier test de racine unitaire dans le cadre de notre étude. Le nombre de retards est choisi également avec le critère d'information d'Akaike. Le test est, comme précédemment, appliqué d'abord sur les trois séries en niveau, soient le log_IPLN, le log_emploi et le taux d'intérêt réel. Les valeurs critiques de Mackinnon à 1% sont de -2.58 dans le modèle avec constante et de -3.48 dans le modèle avec constante et tendance. À 5%, elles sont de -1.95 et de -2.89. Pour les variables log_IPLN et log_emploi, les tests ne rejettent pas l'hypothèse nulle de racine unitaire, au niveau de significativité de 1%. En conséquence, log_IPLN et log_emploi sont non stationnaires à 1% et le taux d'intérêt réel est stationnaire à 5%.

Tableau 4.9 Test ERS sur log_IPLN

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Avec constante	-0.71	-2.41	-1.71	-1.39	-1.04	-1.71
Avec constante et tendance	-1.59	-2.96	-1.74	-1.75	-2.38	-2.61

Tableau 4.10 Test ERS sur log_emploi

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Avec constante	-2.38	-0.03	-0.34	-0.61	-0.79	-1.25
Avec constante et tendance	-2.44	-0.87	-0.94	-1.07	-1.10	-3.14

Tableau 4.11 Test ERS sur le taux d'intérêt réel

	Canada
Avec constante	-2.652
Avec constante et tendance	-3.305

Nous reprenons le test ERS sur $dlog_IPLN$ et $dlog_emploi$. Seules les statistiques calculées pour le Canada sont supérieures aux valeurs critiques de Mackinnon de à 1%. Globalement, nous rejetons l'hypothèse nulle de racine unitaire. Les séries $dlog_IPLN$ et $dlog_emploi$ sont donc stationnaires.

Tableau 4.12 Test ERS sur $dlog_IPLN$

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Avec constante	-0.53	-11.70	-8.70	-5.57	-9.22	-10.48
Avec constante et tendance	-1.42	-12.90	-11.17	-7.66	-13.13	-10.27

Tableau 4.13 Test ERS sur $dlog_emploi$

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Avec constante	-3.71	-8.35	-14.83	-4.79	-6.04	-9.08
Avec constante et tendance	-5.98	-12.79	-15.84	-8.26	-9.78	-11.59

À la lumière des résultats obtenus, le test de Phillips-Perron rejette largement l'hypothèse nulle de racine unitaire pour les variables log_IPLN et log_emploi . Par contre, les tests ADF et ERS concluent en général en faveur de la non-stationnarité pour les deux variables en niveau (log_IPLN et log_emploi) et en faveur de la stationnarité pour leurs premières différences ($dlog_IPLN$ et $dlog_emploi$). Quant au taux d'intérêt réel, les trois tests indiquent qu'il est stationnaire en niveau.

En conclusion, l'image générale qui émerge de ces tests est que les variables \log_IPLN et \log_emploi sont intégrées d'ordre un, $I(1)$ alors que le taux d'intérêt est intégré d'ordre zéro, $I(0)$.

4.3 Test de cointégration

La cointégration permet de spécifier les relations stables à long terme tout en analysant la dynamique à court terme entre les variables. Il y a de la cointégration lorsque la combinaison linéaire des séries intégrées d'ordre un, $I(1)$, est $I(0)$. Puisque les variables \log_IPLN et \log_emploi sont $I(1)$, il existerait possiblement un lien de cointégration entre elles.

Dans la littérature, diverses approches sont utilisées afin de déterminer le nombre de relations de cointégration, entre autre, l'approche Engle-Granger (1987) et celle de Johansen (1991). La première se fonde sur les tests de racine unitaire de Dickey-Fuller et la seconde est basée sur deux statistiques : le test de la trace et le test de la valeur propre (eigenvalue). L'approche de Engle-Granger permet d'obtenir qu'une seule relation de cointégration alors que celle de Johansen permet de distinguer plusieurs vecteurs de cointégration. Dans le cadre de ce travail, il existerait au plus une relation de cointégration. Nous allons ainsi adopter l'approche de Engle-Granger.

Puisque nous nous intéressons uniquement à la dynamique du prix du logement, nous reportons uniquement les résultats concernant la variable \log_IPLN comme variable dépendante.

Tableau 4.14 Test de cointégration de Engle-Granger

	Canada	Toronto	Hamilton	Montréal	Winnipeg	Calgary
Statistique calculée	-3.72	-5.02	-2.80	-3.50	-3.79	-5.00

La valeur critique est de -3.38 au niveau de significativité de 5%. Nous rejetons ainsi l'hypothèse nulle de non stationnarité pour ce qui est du Canada, Toronto, Montréal, Winnipeg et Calgary puisque les statistiques de Engle-Granger calculées sont inférieures à la valeur critique. En revanche, à Hamilton, nous ne rejetons pas l'hypothèse de non stationnarité. On en déduit que globalement les résidus de la relation entre \log_IPLN et \log_emploi sont stationnaires. En conséquence, les séries \log_IPLN et \log_emploi sont donc cointégrées.

La conclusion serait qu'il existe globalement une relation de cointégration, donc une relation de long terme entre les variables. Le tableau suivant donne les résultats de l'estimation de cette relation. Les coefficients de l'estimation représentent la force de rappel vers la cible de long terme.

Tableau 4.15 Estimation de la relation de cointégration entre \log_IPLN et \log_emploi

	Variable	Coefficient	Écart-type	t-statistique	probabilité
Canada	log_emploi_can	1,466	0.259	5.657	0.000
	constante	0.002	0.001	1.203	0.231
Toronto	log_emploi_tor	0.434	0.132	3.282	0.001
	constante	0.006	0.002	2.840	0.005
Hamilton	log_emploi_ham	0.116	0.046	2.482	0.014
	constante	0.008	0.001	7.104	0.000
Montréal	log_emploi_mtl	0.139	0.057	2.437	0.016
	constante	0.011	0.001	11.338	0.000
Winnipeg	log_emploi_win	0.049	0.064	0.765	0.446
	constante	0.009	0.001	8.804	0.000
Calgary	log_emploi_cal	0.121	0.108	1.121	0.264
	constante	0.014	0.0024	5.495	0.000

CHAPITRE V

RÉSULTATS

5.1 Analyse de la dynamique du prix du logement

5.1.1 Estimation du VAR, du VECM et du VAR Spatial

Nous ne reportons pas ici les résultats des estimations des VAR de toutes les villes pour des raisons d'espace. Toutefois, nous allons présenter uniquement les résultats de Winnipeg à titre indicatif et pour une discussion sommaire.

Les séries du $dlog_IPLN$, du $dlog_emploi$ et du taux d'intérêt réel étant stationnaires, il est donc possible de les modéliser par des processus VAR. Il faut a priori déterminer l'ordre p du processus. Pour cela, nous avons estimé plusieurs modèles VAR avec l'ordre p allant 1 à 20. Nous avons calculé le critère d'information Akaike (AIC) pour chacun des modèles et nous retenons le modèle qui minimise ce critère. Dans le cas où il existe un vecteur de cointégration, le VAR précédemment estimé est mal spécifié. Il faut alors considérer un VECM. Puisque l'idée principale est de tenir compte de l'interaction spatiale, nous estimons également un VAR spatial. Nous reproduisons dans les tableaux 5.1, 5.2 et 5.3 suivants les résultats pour la variable dépendante $dlog_IPLN$.

Tableau 5.1 Estimation du VAR

Modèle d'autorégression vectorielle				
Variable dépendante = dlog_ipln_win				
R-carré = 0.6571				
R-carré ajusté = 0.5667				
Variable		Coefficient	t-statistique	probabilité
txreel	lag1	0.029004	0.409412	0.683200
txreel	lag2	-0.001970	-0.019231	0.984699
txreel	lag3	-0.069144	-0.669960	0.504580
txreel	lag4	0.014883	0.144594	0.885351
txreel	lag5	0.075487	0.742646	0.459609
txreel	lag6	-0.080354	-0.804037	0.423471
txreel	lag7	-0.066666	-0.690476	0.491652
txreel	lag8	0.128561	1.806841	0.074092
dlog_ipln_win	lag1	0.632307	5.671659	0.000000
dlog_ipln_win	lag2	0.134623	1.030549	0.305483
dlog_ipln_win	lag3	0.086409	0.674730	0.501558
dlog_ipln_win	lag4	-0.080338	-0.619380	0.537214
dlog_ipln_win	lag5	-0.142100	-1.076593	0.284508
dlog_ipln_win	lag6	0.245810	1.903796	0.060097
dlog_ipln_win	lag7	-0.086360	-0.665697	0.507289
dlog_ipln_win	lag8	-0.058509	-0.526906	0.599541
dlog_emploi_win	lag1	-0.018443	-0.267573	0.789634
dlog_emploi_win	lag2	0.081127	1.186478	0.238522
dlog_emploi_win	lag3	-0.013314	-0.178425	0.858786
dlog_emploi_win	lag4	0.004807	0.064662	0.948585
dlog_emploi_win	lag5	0.078674	1.097492	0.275322
dlog_emploi_win	lag6	-0.009348	-0.138739	0.889963
dlog_emploi_win	lag7	0.032306	0.500199	0.618143
dlog_emploi_win	lag8	0.131739	2.233431	0.027972
constant		0.000093	0.044769	0.964389

Tableau 5.2 Estimation du VECM

Modèle vectoriel à correction d'erreur				
Variable dépendante = dlog_ipln_win				
R-carré = 0.3828				
R-carré ajusté = 0.2133				
Variable		Coefficient	t-statistique	probabilité
txreel	lag1	-0.062016	-0.712269	0.478121
txreel	lag2	0.021380	0.252639	0.801117
txreel	lag3	-0.153601	-1.797684	0.075544
txreel	lag4	-0.064762	-0.754813	0.452310
txreel	lag5	-0.010823	-0.127336	0.898955
txreel	lag6	-0.101293	-1.212818	0.228339
txreel	lag7	-0.056261	-0.687508	0.493512
txreel	lag8	-0.042918	-0.520356	0.604079
dlog_ipln_win	lag1	0.386493	3.681615	0.000393
dlog_ipln_win	lag2	0.069249	0.606945	0.545399
dlog_ipln_win	lag3	0.008990	0.078652	0.937482
dlog_ipln_win	lag4	0.197968	1.742577	0.084786
dlog_ipln_win	lag5	-0.006053	-0.053812	0.957203
dlog_ipln_win	lag6	0.024487	0.217194	0.828543
dlog_ipln_win	lag7	0.078016	0.712268	0.478121
dlog_ipln_win	lag8	-0.041113	-0.392494	0.695611
dlog_emploi_win	lag1	0.039444	0.477235	0.634339
dlog_emploi_win	lag2	-0.150091	-1.771211	0.079874
dlog_emploi_win	lag3	0.127573	1.435523	0.154567
dlog_emploi_win	lag4	0.101934	1.161859	0.248331
dlog_emploi_win	lag5	-0.113565	-1.305595	0.194982
dlog_emploi_win	lag6	0.079986	0.905349	0.367671
dlog_emploi_win	lag7	0.024121	0.297234	0.766966
dlog_emploi_win	lag8	0.027146	0.344244	0.731457
ec term		0.001426	1.728404	0.087307
constant		-0.278121	-1.715543	0.089649

Tableau 5.3 Estimation du VAR spatial

Modèle d'autorégression vectorielle spatiale			
Variable dépendante = dlog_ipln_win			
R-carré = 0.9987			
R-carré ajusté = 0.9982			
Variable	Coefficient	t-statistique	probabilité
dlog_emploi_win lag1	-0.013731	-0.297452	0.766649
dlog_emploi_win lag2	0.015222	0.433980	0.665102
dlog_emploi_win lag3	0.029948	1.203561	0.231189
dlog_emploi_win lag4	-0.000935	-0.048104	0.961715
dlog_emploi_win lag5	-0.009771	-0.615946	0.539126
dlog_emploi_win lag6	-0.000711	-0.053213	0.957653
dlog_emploi_win lag7	0.001466	0.127112	0.899070
dlog_emploi_win lag8	-0.002106	-0.207889	0.835677
dlog_emploi_tor lag1	0.003588	0.081580	0.935120
dlog_emploi_tor lag2	0.010523	0.300377	0.764423
dlog_emploi_tor lag3	0.008076	0.327004	0.744249
dlog_emploi_tor lag4	-0.005278	-0.277925	0.781561
dlog_emploi_tor lag5	-0.002781	-0.180343	0.857195
dlog_emploi_tor lag6	0.003477	0.269282	0.788187
dlog_emploi_tor lag7	0.006473	0.582435	0.561395
dlog_emploi_tor lag8	0.003036	0.311670	0.755846
txreel lag1	-0.012979	-0.297329	0.766743
txreel lag2	-0.002469	-0.069342	0.944836
txreel lag3	-0.013064	-0.516752	0.606305
txreel lag4	-0.000368	-0.018751	0.985072
txreel lag5	0.001893	0.118702	0.905715
txreel lag6	-0.000550	-0.040890	0.967453
txreel lag7	0.000853	0.073492	0.941540
txreel lag8	0.000740	0.072566	0.942276
dlog_ipln_win lag1	1.291960	14.568231	0.000000
dlog_ipln_win lag2	-0.278528	-2.002861	0.047504
dlog_ipln_win lag3	0.018882	0.141377	0.887815
dlog_ipln_win lag4	0.112124	0.922161	0.358342
dlog_ipln_win lag5	-0.142173	-1.276399	0.204342
dlog_ipln_win lag6	-0.029297	-0.286783	0.774786
dlog_ipln_win lag7	0.039382	0.420563	0.674846
dlog_ipln_win lag8	-0.039682	-0.614593	0.540017
constant	-0.191661	-1.505917	0.134784

Selon la théorie économique, une hausse du taux d'intérêt réel devrait faire ralentir l'activité économique et donc exercer une pression à la baisse sur le prix général des biens. Cette théorie pourrait ne pas être similaire dans le cas des logements puisque ceux-ci sont moins sensibles à une variation du taux d'intérêt à très court terme, et l'effet devrait se faire sentir plus à long terme. De façon générale, les résultats des estimations du VAR, du VECM et du VAR spatial donnent des résultats très mitigés, à partir desquels il est difficile de pouvoir dresser un portrait général de l'impact du taux d'intérêt sur les prix des logements.

Concernant l'emploi, un choc positif de l'emploi devrait, théoriquement, faire augmenter les prix des logements, à cause de la stimulation de l'activité économique. Les résultats obtenus ne sont pas non plus clairs sur cette relation positive, avec des estimations pour les trois modèles qui donnent des coefficients tantôt positifs, tantôt négatifs. Ce fait est normal, à cause de la complexité de la dynamique.

Les p-value sont supérieures à 5%, montrant que les résultats obtenus ne sont pas significatifs. Cependant, les modèles demeurent toujours bons puisque des tests de significativité jointe qui suppose nuls les coefficients liés soit au taux d'intérêt, soit à l'emploi, donne des statistiques de Fisher calculées supérieures aux valeurs théoriques pour les trois modèles.

Un autre élément important qui ressort des estimations est que le R^2 ajusté du VAR spatial est élevé et supérieur à ceux du VECM et du VAR traditionnel. Le VAR spatial explique donc mieux la variabilité du prix du logement.

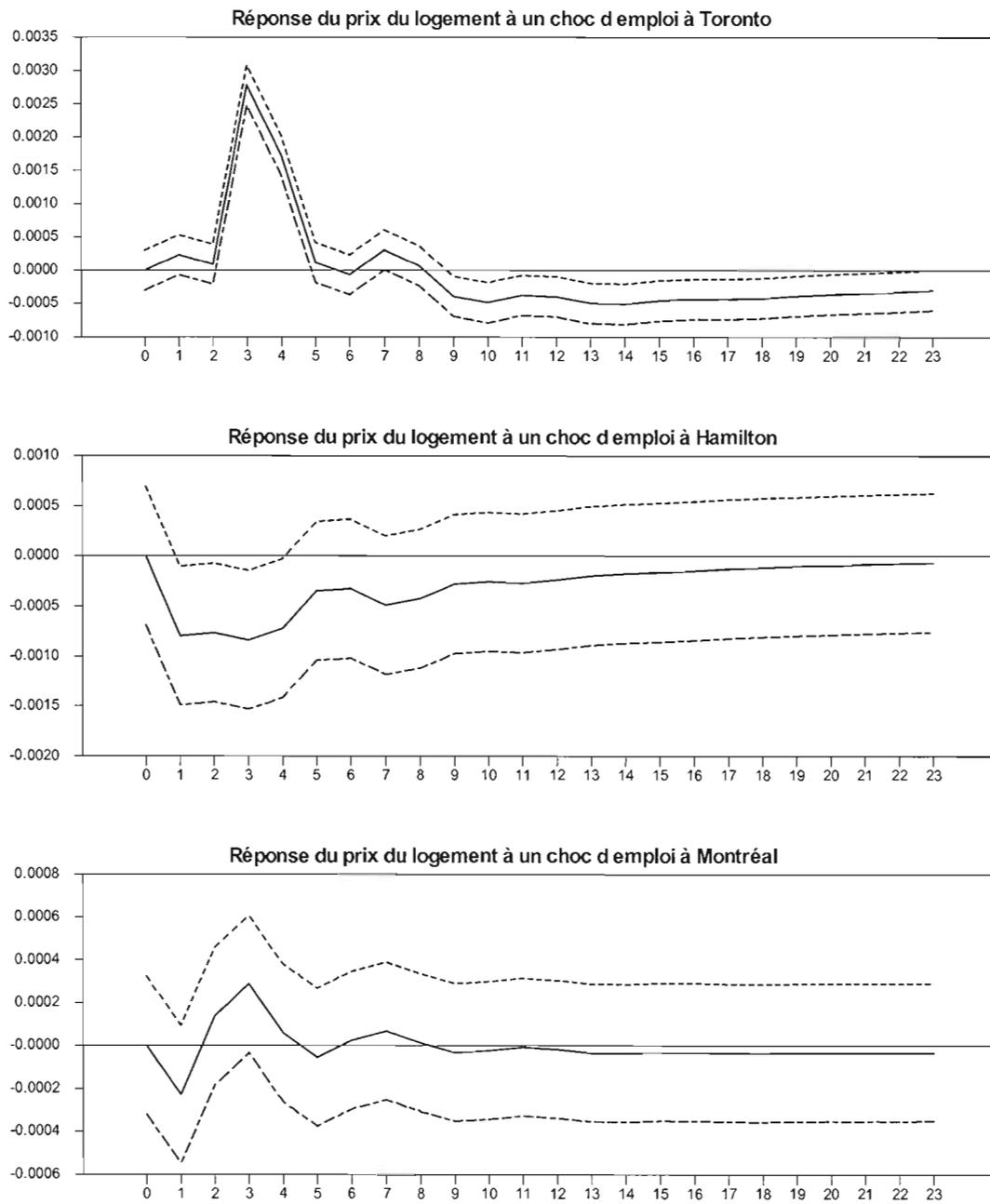
5.1.2 Fonctions de réponses

La réaction dynamique d'une variable endogène aux chocs est mieux explicitée par les représentations graphiques comme le suggère Sims (1980). Nous utiliserons ainsi la représentation graphique des fonctions de réponse, avec leurs intervalles de confiance au niveau de 95%. Du fait que nous ne disposons pas de techniques adéquates pour représenter les fonctions de réponses à partir de VAR Spatial, nous allons nous limiter ici à l'interprétation des fonctions de réponses du VAR traditionnel. De plus, l'identification des chocs n'est pas la principale préoccupation, puisqu'il s'agit ici d'un modèle prédictif plutôt que structurel.

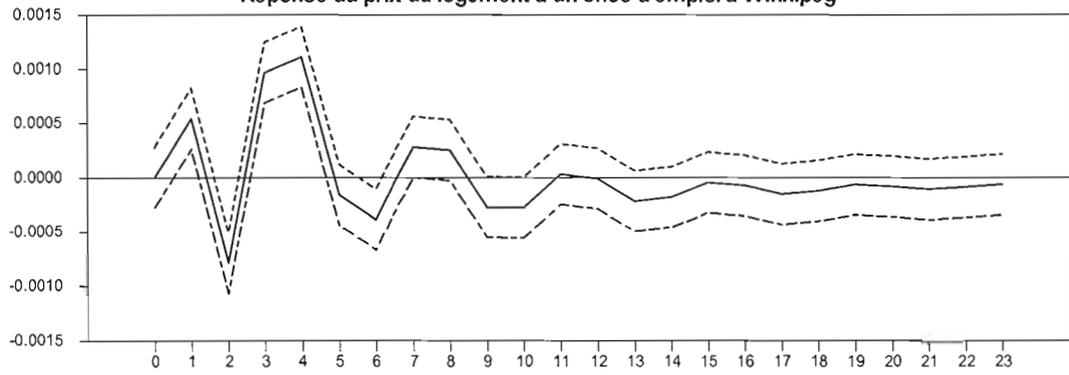
Réponses du prix à l'emploi

En vertu de la théorie économique et des résultats des estimations obtenues précédemment, un choc positif de l'emploi devrait se traduire par l'augmentation des prix du logement. Qu'en est-il réellement des fonctions de réponse? Comme il est possible de le voir dans les graphiques de la figure 5.1, aux 2 premiers trimestres, les effets sont négligeables. La réponse maximale est atteinte entre le 3^{ème} et le 6^{ème} trimestre. Les conséquences de ce choc diminuent progressivement et s'estompent quasiment entre le 8^{ème} et le 10^{ème} trimestre. Par la suite, dans certains cas, la réponse au choc devient négative.

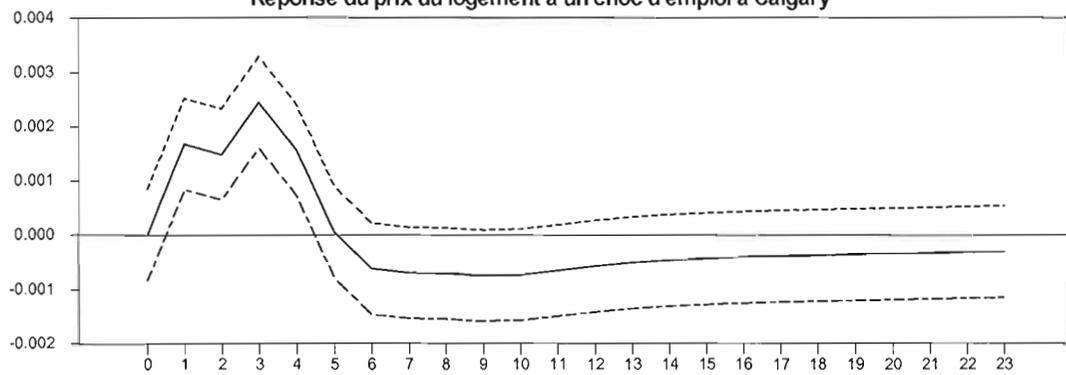
Figure 5.1 Fonction de réponse du prix du logement au choc d'emploi



Réponse du prix du logement à un choc d emploi à Winnipeg



Réponse du prix du logement à un choc d emploi à Calgary

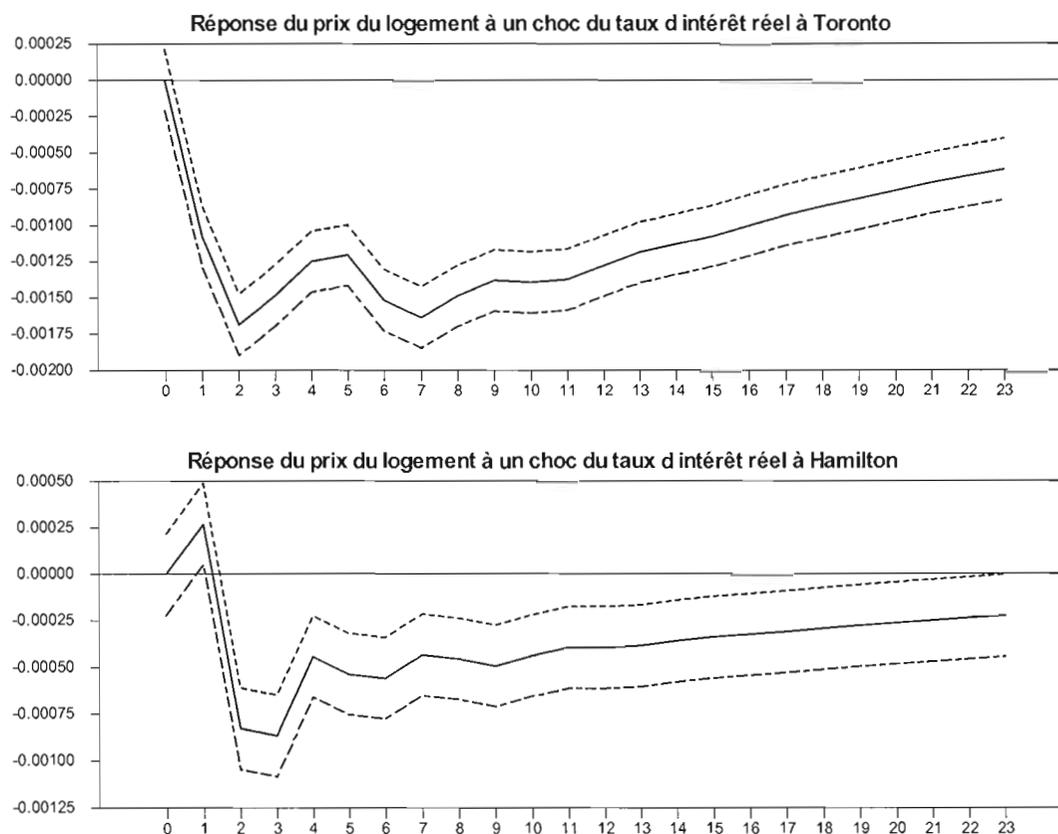


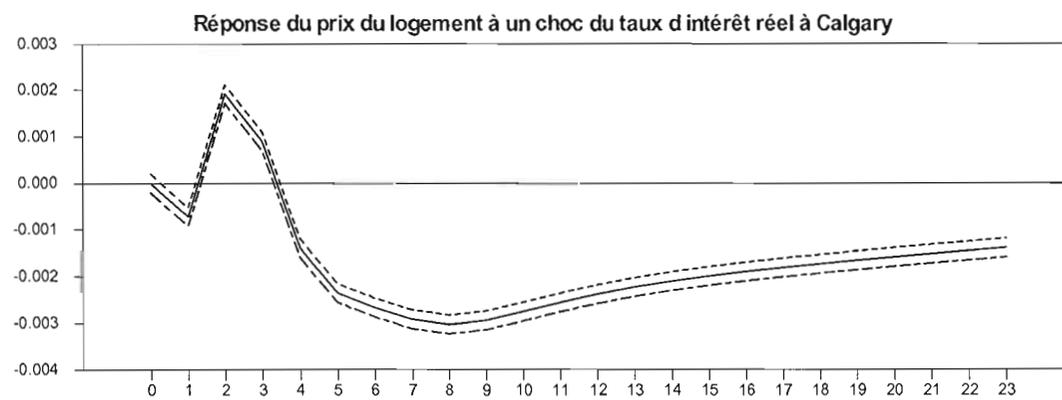
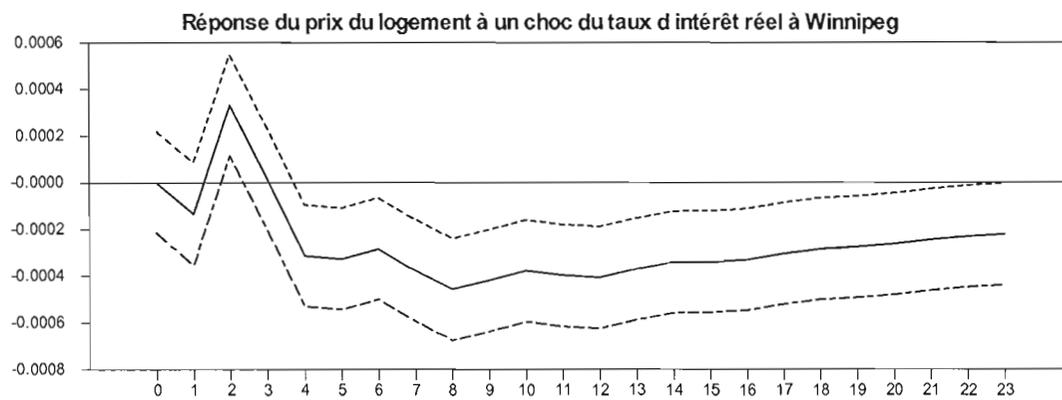
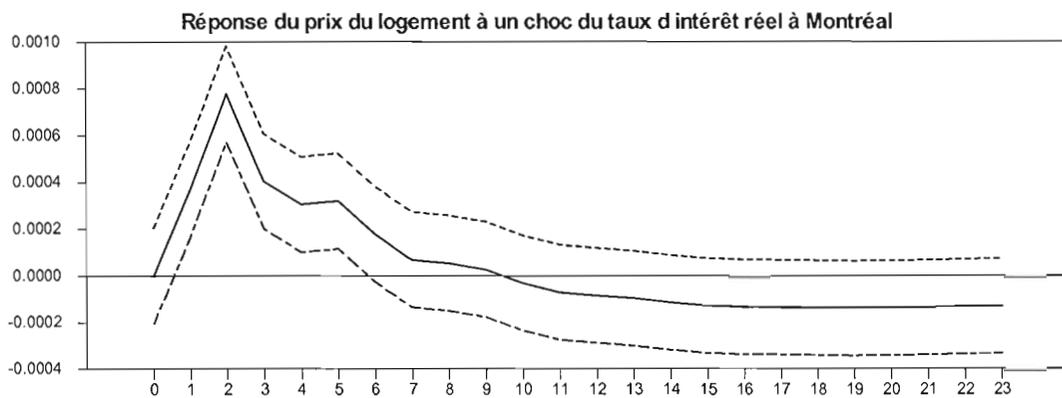
Réponses du prix au taux d'intérêt réel

Suivant les résultats des estimations des VAR et de la théorie économique, un choc positif de taux d'intérêt aurait comme effet de réduire les prix immobiliers.

Comment réagissent les prix de logements au choc de taux d'intérêt en se basant sur les fonctions de réponse? Dans presque tous les cas, un choc positif du taux d'intérêt réel a un impact négatif. Comme on peut l'observer dans la figure 5.2, à l'impact du choc, les prix de logements ne diminuent pas aux deux 1^{ers} trimestres dans tous les cas. La réponse maximale se produit entre le 3^{ème} et le 7^{ème} trimestre. Les effets diminuent vers le 6^{ème} trimestre. Dans certains cas l'effet reste persistant pour s'estomper vers le 22^{ème} trimestre.

Figure 5.2 Fonction de réponse du prix du logement au choc de taux d'intérêt





5.1.3 Interprétation globale de la dynamique du prix du logement

Les estimations du VAR, du VECM et du VAR Spatial ainsi que l'analyse des fonctions de réponse permettent de donner une idée globale de la dynamique du prix du logement.

Les résultats des estimations des VAR sont cohérents avec la théorie économique ainsi qu'avec les résultats des études antérieures. Le prix du logement dépend négativement du taux d'intérêt et positivement de l'emploi.

En effet, suivant la théorie économique, un choc positif du taux d'intérêt a pour effet de réduire la quantité de logements demandés, faisant ainsi déplacer la courbe de demande vers la gauche. Ce qui exerce une pression à la baisse des prix du logement. Les prix du logement devraient être ainsi reliés négativement au taux d'intérêt. Ce qu'on retrouve effectivement dans la présente étude à travers les estimations des VAR. De plus, on remarque, suivant l'analyse de la fonction de réponse, qu'il y a un délai de réponse des prix du logement au choc du taux d'intérêt. Ce qui est conforme à la littérature empirique sur la réaction des prix du logement aux effets du taux d'intérêt.

Du côté de l'emploi, nos résultats sont également compatibles avec la théorie. En effet, un choc de l'emploi produit un impact positif sur les prix du logement dans tous cas à l'exception du cas d'Hamilton. Les effets sont moins persistants que dans le cas du choc de taux d'intérêt et deviennent même négatifs dans tous les cas après 10 trimestres.

5.2 Efficacité prédictive de trois modèles (VAR, VECM, VAR Spatial)

5.2.1 Présentation des résultats au niveau des villes

Pour évaluer l'efficacité prédictive du prix du logement de chacun des 3 modèles utilisés, nous avons eu recours à deux mesures : la Racine de l'Erreur Quadratique Moyenne (*Root Mean Squared Error, RMSE*) et le gain d'efficacité (*Relative Root Mean Squared Error*)

La période de prévision est du premier trimestre 1996 au troisième trimestre 2005. Nous générons une prévision du prix du logement pour chacune des villes sur un horizon de 1 à 8 trimestres. Cette prévision est effectuée à l'intérieur de l'échantillon (*in-sample*).

La racine de l'erreur quadratique moyenne (RMSE)

La racine de l'erreur quadratique moyenne se fonde sur le calcul de l'erreur de prévision. Rappelons ici son expression :

$$RMSE = \sqrt{\frac{1}{T} \sum_i (Valeur_{observée_i} - Valeur_{prédite})^2},$$

où T est la taille de l'échantillon.

De cette formule, on peut facilement remarquer que plus la prévision est bonne, plus est petite cette mesure. On peut donc comparer la capacité prédictive des trois modèles sur cette base.

La mesure du gain d'efficacité prédictive

La comparaison basée sur le RMSE peut être complétée par un calcul de gain d'efficacité des modèles VAR, VECM et VAR spatial par rapport à un modèle de marche aléatoire.

Cette mesure se calcule à partir des RMSE obtenues pour chacun des modèles, selon les équations suivantes :

$$\text{Gain_d'efficacité} = 1 - \frac{RMSE_i}{RMSE_{ma}}$$

où $RMSE_i$ représente le RMSE des modèles VAR, VECM et VAR spatial et $RMSE_{ma}$, le RMSE du modèle de marche aléatoire.

Cette mesure révèle le pourcentage de réduction ou d'augmentation du RMSE relative à l'utilisation du VAR, VECM ou du VAR spatial par rapport à un modèle de marche aléatoire. Il nous permet de voir rapidement si les modèles font mieux que la marche aléatoire en ce qui concerne la prévision du prix du logement. En résumé, si le gain d'efficacité est positif, le modèle (VAR, VECM ou VAR Spatial) prédit mieux que le modèle de marche aléatoire. Plus ce gain est élevé, meilleur est le modèle. Par contre, si ce gain est négatif, le modèle de marche aléatoire est plus efficace.

Dans les pages qui suivent, nous reportons les résultats obtenus pour chaque ville : Hamilton, Montréal, Winnipeg et Calgary.

Cas de Hamilton

À tous les horizons, le VAR spatial a le RMSE le plus petit. Ce qui indique que sa qualité de prévision est meilleure par rapport au VAR et au VECM. Remarquons ici la piètre performance du VECM. Cela s'explique par le fait que nous n'avons trouvé aucune relation de cointégration avec le test de Engle-Granger, mais nous en avons supposé une afin d'estimer le VECM. Donc, il est normal que le VECM ne puisse pas bien expliquer le prix du logement s'il n'existe aucune relation de cointégration.

Quant au gain d'efficacité, le VAR spatial donne la meilleure prévision par rapport à la marche aléatoire. À court terme, le gain d'efficacité des trois modèles est très faible par rapport au modèle de référence. Avec l'horizon, ce gain augmente.

Figure 5.3 RMSE, cas d'Hamilton

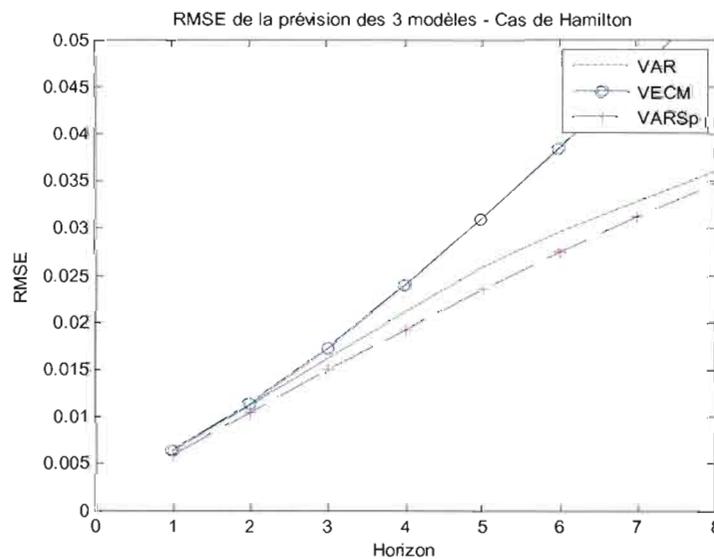
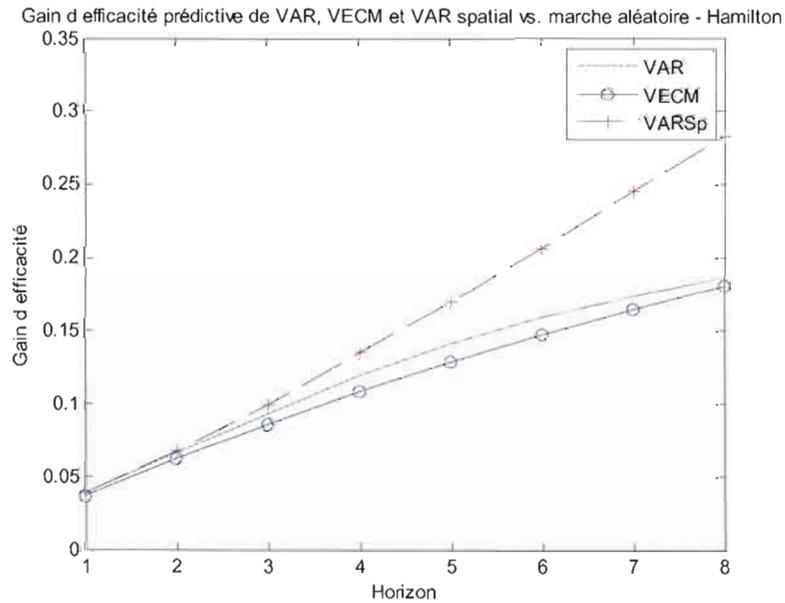


Figure 5.4 Gain d'efficacité, cas d'Hamilton



Cas de Montréal

De façon globale, le VAR spatial minimise le RMSE pour chaque horizon de prévision. À court terme, soit avant le 3^e horizon, les valeurs de RMSE sont similaires dans les trois modèles. Par la suite, la capacité prédictive du VECM se détériore de façon appréciable par rapport aux autres modèles. À l'horizon 8, cette faiblesse du VECM est d'autant plus accentuée.

Le VAR spatial est visiblement le plus efficace des trois modèles, lorsqu'on les compare au modèle de marche aléatoire. En effet, le gain d'efficacité du VAR spatial est supérieur à celui du VAR et du VECM, cet écart augmentant avec l'horizon de prévision. On remarque ici le VECM fait mieux que le VAR, contrairement au cas d'Hamilton.

Figure 5.5 RMSE, cas de Montréal

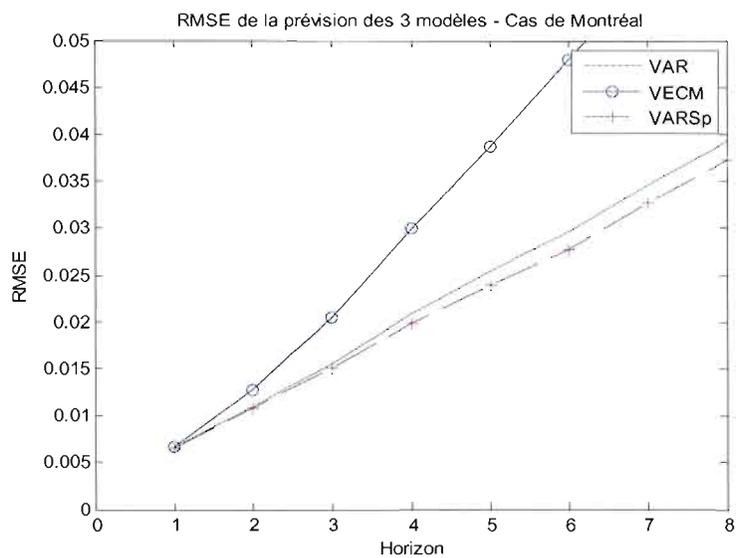
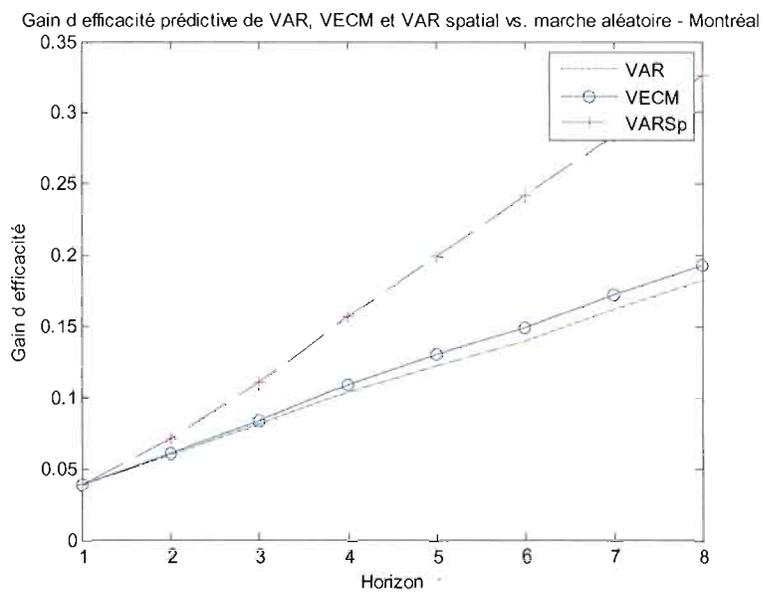


Figure 5.6 Gain d'efficacité, cas de Montréal



Cas de Winnipeg

Le RMSE du VAR Spatial est le plus petit à tous les horizons. Cependant, l'écart entre les RMSE des trois modèles n'est pas important. Dans ce cas, il est difficile de juger de la meilleure capacité prédictive d'un modèle en particulier.

La mesure de gain d'efficacité nous permet de voir que le VAR spatial et le VECM sont plus performants que le VAR conventionnel, mais le VAR spatial demeure le meilleur modèle. On constate que, plus l'horizon est grand, plus est performant le VECM par rapport au VAR. Ceci pourrait s'expliquer par le terme à correction d'erreur introduit dans le VECM, qui capte la force de rappel de long terme.

Figure 5.7 RMSE, cas de Winnipeg

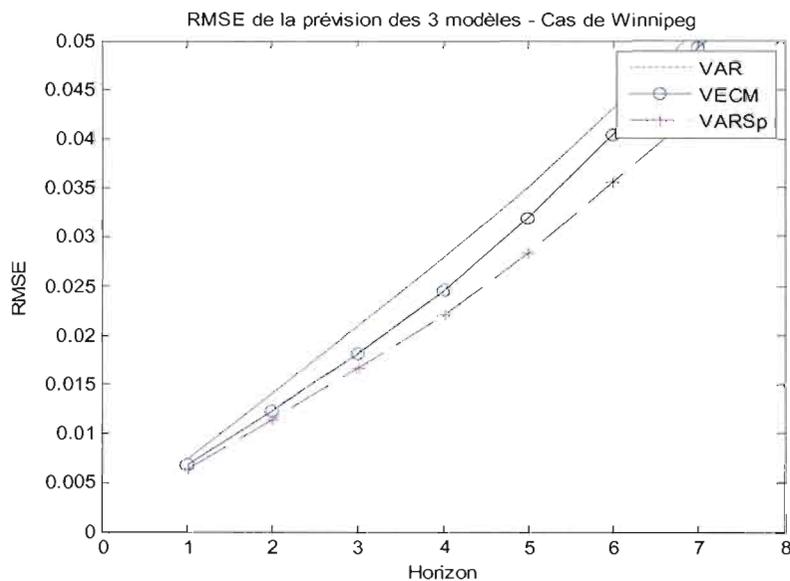
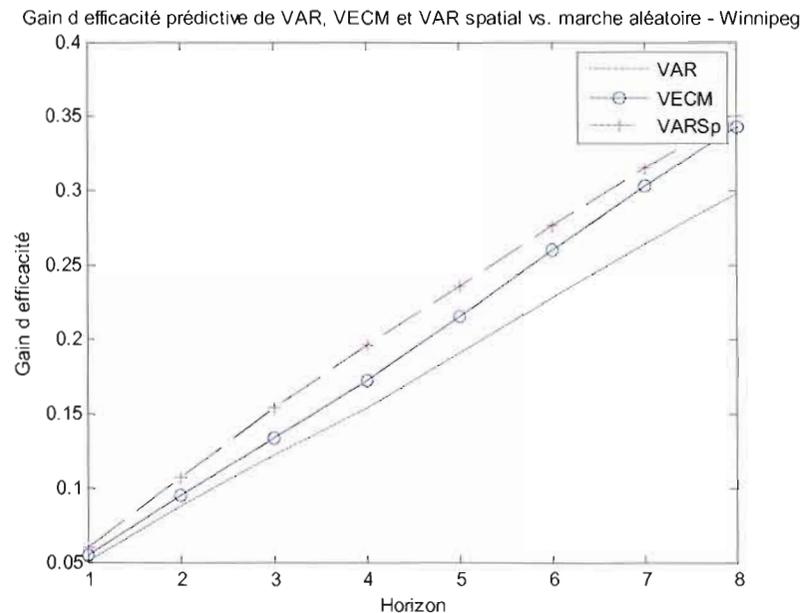


Figure 5.8 Gain d'efficacité, cas de Winnipeg



Cas de Calgary

Le VAR Spatial minimise plus le RMSE, quelque soit l'horizon. Dans ce cas-ci, il est aussi clair que c'est ce modèle qui est le plus performant. Le VECM fait mieux que le VAR aux horizons supérieurs à deux.

Encore une fois, le gain d'efficacité prédictive du VAR spatial est supérieur aux deux autres modèles, le VAR conventionnel étant le moins efficace. De plus, par rapport à la marche aléatoire, les trois modèles sont plus efficaces et cette efficacité croît avec l'horizon.

Figure 5.9 RMSE, Cas de Calgary

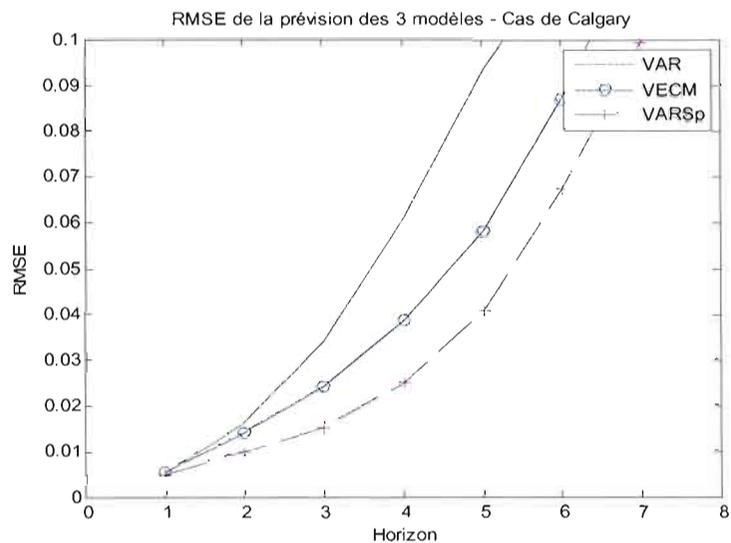
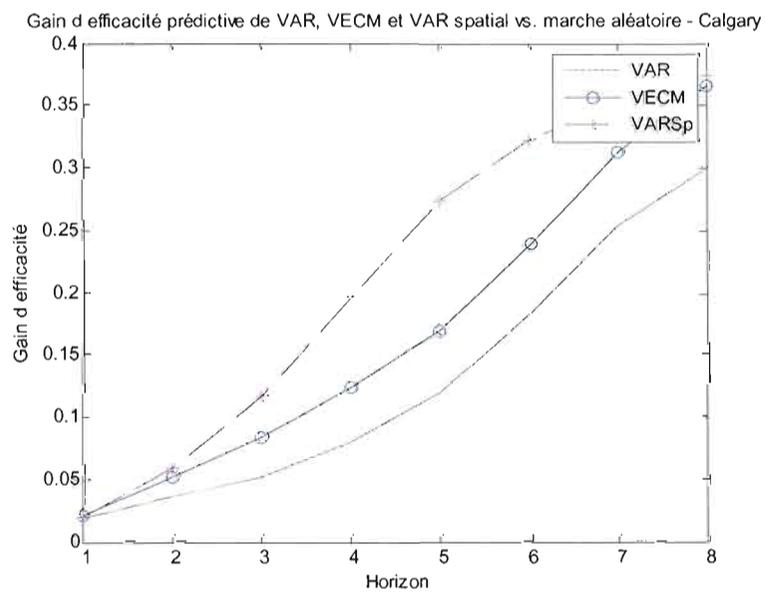


Figure 5.10 Gain d'efficacité, cas de Calgary



5.2.2 Quel modèle a la meilleure efficacité prédictive ?

À la suite de l'analyse des résultats que nous venons d'effectuer, le VAR spatial est le plus efficient des trois modèles. Par la suite, vient le VECM qui fait mieux que le VAR conventionnel dans trois cas sur quatre.

Nous remarquons que le gain d'efficacité par rapport au modèle de marche augmente avec l'horizon dans chaque cas et pour chaque modèle. Ce qui ressort de cette analyse est que la prise à compte de l'interaction entre les villes par le biais du VAR spatial améliore significativement l'efficacité prédictive. À l'horizon 8 par exemple, l'efficacité prédictive d'un VAR spatial par rapport à la marche aléatoire se situe entre 25% et 40%, comparativement à un intervalle de 15% et 35% pour les VECM et VAR conventionnel.

CONCLUSION

Dans cette étude, nous avons examiné l'importance de l'addition des effets spatiaux dans la prévision des prix immobiliers au Canada et dans cinq de ses régions métropolitaines, en utilisant trois modèles économétriques (VAR, VECM, VAR spatial).

Au Canada, bien qu'il existe des travaux qui ont modélisé le prix du logement, par l'approche VAR, il y a cependant peu d'études qui ont utilisé des approches VECM ou VAR spatial. Ce travail permet donc de faire une synthèse portant sur la capacité prédictive du prix immobilier par le VAR spatial.

L'objectif de ce travail était d'appliquer les modèles de l'économétrie spatiale au cas canadien, pour voir si la qualité de la modélisation du prix du logement peut être améliorée. Plus spécifiquement, il était question d'une part, d'analyser la dynamique du prix du logement en introduisant l'effet spatial et d'autre part, de vérifier la capacité prédictive du VAR spatial.

Pour atteindre cet objectif, nous avons eu recours à l'estimation du VAR, du VECM et du VAR spatial et aux fonctions de réponse. L'efficacité prédictive a été vérifiée grâce à deux mesures de choix de modèle : la Racine de l'Erreur Quadratique Moyenne (*Root Mean Squared Error, RMSE*) et le gain d'efficacité (*Relative Root Mean Squared Error*).

Suivant cette démarche, nous avons pu obtenir deux principaux résultats. Premièrement, l'estimation et l'analyse des fonctions de réponse du VAR permettent de conclure qu'un choc d'emploi affecte positivement le prix du logement. Le choc de politique monétaire (taux d'intérêt), quant à lui, a un impact négatif sur le prix

immobilier. Ceci est donc cohérent avec la théorie économique. Deuxièmement, l'analyse de RMSE et du gain d'efficacité prédictive nous a permis de voir que c'est le VAR spatial qui donne une prévision de meilleure qualité. L'incorporation de l'interaction entre deux régions dans un modèle VAR permet donc de mieux prévoir le prix immobilier.

Ce travail constitue une alternative aux travaux qui font abstraction des facteurs géographiques dans la prévision du prix immobilier. Une poursuite de la recherche dans ce sens pourrait améliorer davantage les qualités prédictives du VAR spatial, surtout en ce qui concerne la construction de la matrice de poids.

BIBLIOGRAPHIE

- Anselin, L. (1980), « Estimation Methods for Spatial Autoregressive Structures », *Cornell University, Regional Science Dissertation and Monograph*, Series no.8, Ithaca, NY.
- Anselin L. et Can A. (1986), « Model comparison and model validation issues in empirical work on urban density models », *Geographical Analysis*, 18, 179-197.
- Aten, B. (1996), « Evidence of Spatial Autocorrelation in International Prices », *Review of Income and Wealth*, 42, 147-163.
- Baffoé-Bonie, J. (1998), « The Dynamic Impact of Macroeconomic Aggregates on Housing Prices and Stock of Houses: A National and Regional Analysis », *Review of Real Estate Finance and Economics*, 178-198.
- Baumont, C. et al. (2000), « Convergence des régions européennes, une approche par l'économétrie spatiale », *Pôle d'Économie et de Gestion*, 1-24.
- Bavaud, F. (1998), « Models for Spatial Weights: A Systematic Look », *Geographical Analysis*, 30, 153-171.
- Bhattacharjee, A. et Jensen-Butler, C.* (2005), « A Model of Regional Housing Markets in England and Wales », *School of Economics and finance, University of St. Andrews*, 1-24.
- Beenstock, M. et Felsentein, D. (2005), « Spatial Vector Autoregressions », *International Regional Science Review*, 1-26.
- Carlino, G.A. et Defina, R. (1998), « The differential regional effects of monetary policy », *Review of Economics and Statistics* 80:572-87.
- Case, B. et al (2004), « Modeling Spatial and Temporal House Price Patterns: A Comparison of Four Models », *Review of Real Estate Finance and Economics*, vol.29, no.2, 167-191.
- Case, A.C. et al, (1993), « Budget spillovers and fiscal policy interdependence: Evidence from the states », *Journal of Public Economics*, 52, 285-307.

- Christiano, L.J., Eichenbaun, M. et Evans, C. (2005), « Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy », *Journal of Political Economy*, 113(3), 1-45.
- Demers, F. (2005), « Modelling and Forecasting Housing Investment: The Case of Canada », *Bank of Canada*, working paper 2005-41.
- Dickey, D.A et Fuller, W.A (1981), « Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root », *Econometrica*, 49, 1057-1072.
- Di Giacinto, V. (2003), « Differential Regional effects of Monetary Policy: a Geographical SVAR Approach », *International Regional Science Review*, 26, 3:313-341.
- Di Pasquale, D. et Wheaton, W.C. (1994), « Housing market dynamics and the future housing prices », *Journal of Urban Economics*, 35:1-27.
- Engle, R.F et Granger, C.W.J. (1987) « Cointegration and Error Correction: Representation, Estimation and Testing », *Econometrica*, 55, 251-276.
- Englund, P. et Ioannides, Y. M. (1997), « House Price Dynamics: an International Empirical Perspective », *Journal of Housing Economics*, 6, 119-136.
- Fauvel, Yvon, (2005), « La variation des prix des logements au Canada », *Société d'Hypothèques et de Logement*, Référence Dossier no 6625-43, 97 pages.
- Fortin, M. et Leclerc, A. (2002), « Déterminants du prix réel des logements au Canada », *L'Actualité Économique, revue de l'analyse économique*, vol.78, no.3.
- Fortura, P. et Kushner, J. (1986), « Canadian Inter-City House Prices Differentials » *AREUEA Journal*, 14(4): 525-536.
- Gali, J. (1999), « Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? », *American Economic Review*, 89, 249-271.
- Gelfand, A. E. et al. (2004), « The Dynamics of Locations in Home Price », *The Journal of Real Estate Finance and Economics*, vol. 29, no.2, 149-166.
- Hamilton, B. W. (1991), « The baby-boom. The baby bust and the housing market: the second look », *Journal of Regional science and Urban Economics*, 21: 547-552.

Hernandez-Murillo, R. et Owyang, M., T. (2006), « The information content of regional employment data for forecasting aggregate conditions », *Economics Letters*, vol.90, no.3, 335-339.

Holly, S. et al. (2006), « A Spatio-Temporal Model of House Prices in the US », *Faculty of Economics and Centre for International Macroeconomics and Finance, University of Cambridge*.

Iacoviello, M. (2000), « House Prices and the Macroeconomy in Europe: Results from a Structural VAR Analysis », *European Central Bank*, working paper no.18.

Johansen, S. (1988), « Statistical Analysis of Cointegration Vector » *Journal of Economic Dynamics and Control*, 12, 232-254.

Lastrapes, W.D. (2002), « The Real Price of Housing and Money Supply Shocks: Time Series Evidence and Theoretical Simulations », *Journal of Housing Economics*, 11: 40-74.

LeGallo, J. (2000), « Économétrie Spatiale: l'autocorrélation spatiale », <http://bibliotheques.uni-lille1.fr/grisemine>, 1-45.

Lesage, J. P. et Pace, K. R. (2004), « Models for spatial dependence Missing data », *The Journal of Real Estate Finance and Economics*, vol. 29, no.2, 233-254

Mankiw, N. G. et Weil, D.N. (1989), « The baby-boom. The baby bust and the housing market: the second look », *Journal of Regional science and Urban Economics* », 19: 235-258.

Meen, G. (2001), « Modelling spatial markets. Theory, Analysis and Policy », *Kluwer Academic Publishers / Dordrecht / London*.

Militino, A. F. et al. (2004), « Alternative Models for Describing Spatial Dependence among Dwelling Selling Prices », *The Journal of Real Estate Finance and Economics*, vol. 29, no.2, 193-209.

Muellbauer, J. et Murphy, A. (1997), « Booms and busts in the UK housing market », *Economic Journal* 107, 1720-1746.

OCDE (2004), « Marchés du logement, patrimoine et cycle économique », *Perspectives économiques de l'OCDE*, no 75, pages 157-178.

Phillips, P.C.B et Perron P (1988), « Testing for a Unit Root in a Time Series Regression », *Biometrika*, 75, 335-346.

Ressources naturelles Canada, *Distance en ligne directe*, http://geonames.nrcan.gc.ca/search/search_f.php, page consultée le 10 janvier 2008.

SCHL, (2004), *L'observateur du Logement au Canada*, édition 2004, 59 pages.

SCHL, (2005), *L'observateur du Logement au Canada*, édition 2005, 78 pages.

SCHL, (2007), « Aperçu du marché canadien », *Actualités Habitation Canada*, juin, 16 pages.

Sims, C. (1980), « Macroeconomics and Reality », *Econometrica*, 1-48.

Smith, L. B. et al (1988), « Recent Developments in Economic Models of Housing Market », *Journal of Economic Literature*, 26(1): 29-64.

Stock J.H. et M.W.Watson (2001), « Vector Autoregressions », *Journal of Economic Perspectives*, 15(4) :101-115.

Tsatsaronis, K. et Zhu, H. (2004), « Dynamique des prix immobiliers: études sur 17 pays », 14 pages.

Wilfrid, K. (2004), « Effets de voisinage dans le modèle de Solow avec des externalités spatiales », *Université de Bourgogne, Pôle d'Économie et de Gestion*, 1-19.

ANNEXE A

DONNÉES

Tableau: Indice des prix à la consommation

Tableau 326-0001 (v735319): Indice des prix à la consommation. Le contenu du panier de 2001; Canada; Ensemble (1992=100) [P100000].

Tableaux: Indices des prix des logements neufs

-V734237 (Tableau 327-0005): Indices des prix des logements neufs; Canada; Total (maison et terrain); 1992=100 (indice) [P10248].

-V734245 (Tableau 327-0005): Indices des prix des logements neufs; Montréal, Québec [24462]; Total (maison et terrain); 1992=100 (indice) [P10256].

-V734249 (Tableau 327-0005): Indices des prix des logements neufs; Toronto et Oshawa, Ontario [35535, 35532]; Total (maison et terrain); 1992=100 (indice) [P10259].

-V734250 (Tableau 327-0005): Indices des prix des logements neufs; Hamilton, Ontario [35537]; Total (maison et terrain); 1992=100 (indice) [P10260].

-V734256 (Tableau 327-0005): Indices des prix des logements neufs; Winnipeg, Manitoba [46602]; Total (maison et terrain); 1992=100 (indice) [P10267].

-V734259 (Tableau 327-0005): Indices des prix des logements neufs; Calgary, Alberta [48825]; Total (maison et terrain); 1992=100 (indice) [P10270].

Tableau: Bons du Trésor

Tableau 176-0043 (V122541): Statistiques du marché financier au dernier mercredi

sauf indication contraire; Canada; Adjudication de Bons du Trésor - rendement moyen : à 3 mois (pourcentage) [B14007].

Tableau: Emploi

Tableau 282-0090 (v2067395): EPA estimations selon la région métropolitaine de recensement. Moyennes mobiles de 3 mois; Canada; Emploi; Désaisonnalisées (personnes - milliers).

ANNEXE B

RMSE ET GAIN D'EFFICACITÉ

Tableau B.1 RMSE, cas de Montréal

Horizon	VAR	VECM	VAR Sp
1	0.0066	0.0066	0.0065
2	0.0109	0.0126	0.0107
3	0.0154	0.0203	0.0150
4	0.0208	0.0299	0.0198
5	0.0254	0.0387	0.0239
6	0.0296	0.0479	0.0277
7	0.0346	0.0571	0.0326
8	0.0392	0.0663	0.0372

Tableau B.2 Gain d'efficacité, cas de Montréal

Horizon	VAR	VECM	VAR Sp
1	0.0380	0.0387	0.0387
2	0.0599	0.0610	0.0708
3	0.0810	0.0835	0.1101
4	0.1039	0.1095	0.1570
5	0.1226	0.1307	0.1989
6	0.1396	0.1494	0.2419
7	0.1620	0.1720	0.2836
8	0.1827	0.1926	0.3256

Tableau B.3 RMSE, cas d'Hamilton

Horizon	VAR	VECM	VAR Sp
1	0.0063	0.0064	0.0059
2	0.0112	0.0113	0.0105
3	0.0162	0.0172	0.0149
4	0.0212	0.0239	0.0193
5	0.0257	0.0309	0.0235
6	0.0297	0.0384	0.0274
7	0.0328	0.0464	0.0312
8	0.0361	0.0545	0.0348

Tableau B.4 Gain d'efficacité, cas d'Hamilton

Horizon	VAR	VECM	VAR Sp
1	0.0385	0.0361	0.0386
2	0.0659	0.0618	0.0667
3	0.0930	0.0857	0.0989
4	0.1188	0.1082	0.1342
5	0.1410	0.1289	0.1696
6	0.1594	0.1471	0.2064
7	0.1731	0.1644	0.2449
8	0.1869	0.1803	0.2826

Tableau B.5 RMSE, cas de Winnipeg

Horizon	VAR	VECM	VAR Sp
1	0.0074	0.0068	0.0063
2	0.0139	0.0122	0.0114
3	0.0209	0.0181	0.0166
4	0.0279	0.0245	0.0219
5	0.0350	0.0320	0.0284
6	0.0431	0.0404	0.0356
7	0.0512	0.0492	0.0428
8	0.0591	0.0578	0.0503

Tableau B.6 Gain d'efficacité, cas de Winnipeg

Horizon	VAR	VECM	VAR Sp
1	0.0507	0.0549	0.0592
2	0.0875	0.0942	0.1069
3	0.1219	0.1331	0.1536
4	0.1543	0.1720	0.1958
5	0.1905	0.2151	0.2355
6	0.2289	0.2599	0.2769
7	0.2640	0.3030	0.3158
8	0.2983	0.3429	0.3507

Tableau B.7 RMSE, Cas de Calgary

Horizon	VAR	VECM	VAR Sp
1	0.0053	0.0057	0.0050
2	0.0163	0.0142	0.0101
3	0.0337	0.0242	0.0153
4	0.0610	0.0384	0.0250
5	0.0931	0.0579	0.0406
6	0.1177	0.0869	0.0671
7	0.1372	0.1217	0.0992
8	0.1542	0.1508	0.1239

Tableau B.8 Gain d'efficacité, cas de Calgary

Horizon	VAR	VECM	VAR Sp
1	0.0189	0.0216	0.0202
2	0.0367	0.0515	0.0592
3	0.0524	0.0831	0.1159
4	0.0801	0.1229	0.1953
5	0.1194	0.1704	0.2739
6	0.1835	0.2377	0.3219
7	0.2547	0.3125	0.3523
8	0.3004	0.3658	0.3739