



**Développement d'un modèle
économétrique à variable latente
afin d'estimer les échanges économiques :
une expérimentation appliquée au Québec**



**SOCIOÉCONOMIE
DES TRANSPORTS**



**ÉTUDES ET RECHERCHES
EN TRANSPORT**

**Développement d'un modèle
économétrique à variable latente
afin d'estimer les échanges économiques :
une expérimentation appliquée au Québec**

**SOCIOÉCONOMIE
DES TRANSPORTS**

**Jean Dubé
Marc Dupéré**

ÉTUDES ET RECHERCHES
EN TRANSPORT

**DÉVELOPPEMENT D'UN MODÈLE ÉCONOMÉTRIQUE
À VARIABLE LATENTE AFIN D'ESTIMER
LES ÉCHANGES ÉCONOMIQUES :
UNE EXPÉRIMENTATION APPLIQUÉE AU QUÉBEC**

Jean Dubé et Marc Dupéré
JDMD Groupe conseils inc.

Réalisé pour le compte du ministère des Transports

Mars 2006

La présente étude a été réalisée à la demande du ministère des Transports du Québec et a été financée par la Direction du partenariat, de la modélisation et de la géomatique.

Les opinions exprimées dans le présent rapport n'engagent que la responsabilité de leurs auteurs et ne reflètent pas nécessairement les positions du ministère des Transports du Québec.

Collaborateur

France-Serge Julien, Service de la modélisation des systèmes de transport

Dépôt légal
Bibliothèque nationale du Québec, 2006
ISBN 2-550-46754-X (pdf)

Titre et sous-titre du rapport Développement d'un modèle économétrique à variable latente afin d'estimer les échanges économiques : une expérimentation appliquée au Québec		N° du rapport Transports Québec RTQ-06-02	
		Date de publication du rapport (année – mois) 2006-03	
Titre du projet de recherche Développement d'un modèle économétrique à variable latente afin d'estimer les échanges économiques : une expérimentation appliquée au Québec		N° du contrat (RRDD-AA-CCXX) N/A	N° de projet ou dossier N/A
Responsable de recherche		Date du début de la recherche 2005- 01	Date de fin de la recherche 2006-02
Auteurs du rapport Jean Dubé et Marc Dupéré			
Chargé de projet, direction France-Serge Julien, Direction de la planification		Coût total de l'étude 14 150 \$	
Étude ou recherche réalisée par (nom et adresse de l'organisme) JDMD Groupe conseils inc. 344, rue de Lammenais Sainte-Foy (Québec) G2G 2M5		Étude ou recherche financée par (nom et adresse de l'organisme) Ministère des Transports du Québec Service de la modélisation des systèmes de transport 35, rue de Port-Royal Est, 4 ^e étage Montréal (Québec) H3L 3T1	
Problématique Le ministère des Transports du Québec a entrepris différentes démarches pour améliorer la connaissance de la demande de transport des marchandises entre le Québec et ses partenaires économiques et au sein de son marché domestique. Une de ces démarches consiste à estimer la demande par l'entremise de modèles économétriques qui pourraient reproduire des flux interrégionaux de transport.			
Objectifs L'objectif de cette recherche est de confectionner, selon un modèle économétrique, une matrice de flux économiques, tous modes, à partir de différents indicateurs : production manufacturière, démographie, emplois et revenus.			
Méthodologie Un modèle économétrique « à variable latente à indicateurs et causes multiples » est développé afin d'estimer l'évolution des échanges économiques du Québec avec le reste du Canada et les États-Unis. Une modélisation de type panel est utilisée car elle permet de tenir compte de l'information transversale et longitudinale. Afin de déterminer le niveau de précision des résultats, ceux-ci sont comparés à des données réelles.			
Résultats et recommandations L'application du modèle MIMIC a donné de bons résultats au chapitre des exportations du Québec, puisque les écarts économiques entre les échanges estimés et les échanges réels sont très faibles. Ce n'est toutefois pas le cas pour les importations, car il existe une différence significative entre les taux de la demande estimée et de la demande observée. L'évaluation de la demande de transport par ce modèle comporte donc des limites.			
Mots-clés Flux d'échanges économiques; flux de transport de marchandises; modèle de transport; demande de transport, modèle économétrique.	Nombre de pages 49 pages	Nombre de références bibliographiques 19	Langue du document <input checked="" type="checkbox"/> Français <input type="checkbox"/> Anglais Autre (spécifier) :

TABLE DES MATIÈRES

1. INTRODUCTION.....	9
2. LE MODÈLE THÉORIQUE.....	11
3. LES DONNÉES MANUFACTURIÈRES.....	17
3.1 Définitions des concepts.....	17
3.2 Les données canadiennes.....	18
3.3 Les données américaines.....	20
3.4 Survol du secteur manufacturier nord-américain.....	22
3.5 Survol des autres variables.....	26
4. RÉSULTATS D'ESTIMATIONS.....	29
4.1 Le modèle des importations.....	31
4.2 Le modèle des exportations.....	33
5. ANALYSE COMPARATIVE DES RÉSULTATS.....	37
5.1 Le cas des importations.....	37
5.2 Le cas des exportations.....	40
6. CONCLUSION.....	43
BIBLIOGRAPHIE.....	45
ANNEXE 1 TRANSFORMATION DE HAUSER ET GOLDBERGER.....	49

LISTE DES TABLEAUX

Tableau 1	Importance relative dans l'activité manufacturière nord-américaine, 1992-2001	23
Tableau 2	Indice de concentration des industries manufacturières nord-américaines, 1992-2001	23
Tableau 3	Importance relative dans l'activité manufacturière canadienne, 1992-2001	24
Tableau 4	Croissance moyenne des principaux indicateurs manufacturiers, 1992-2001 (en %)	25
Tableau 5	Importance relative des Etats-Unis par rapport au Canada, 1992-2001 – Ratio des indicateurs (Etats-Unis sur Canada)....	27
Tableau 6	Importance relative des régions nord-américaines selon les caractéristiques, 1992-2001.....	27
Tableau 7	Revenu personnel et revenu personnel disponible per capita (15 ans et plus), 1992-2001 – (en \$ CA).....	28
Tableau 8	Résultats d'estimations du modèle MIMIC appliqué à la demande de transport (importation).....	35
Tableau 9	Résultats d'estimations du modèle MIMIC appliqué à la demande de transport (exportation).....	36

LISTE DES SCHÉMAS

Schéma 1	Illustration de la chaîne de relations du modèle MIMIC	12
Schéma 2	Comportement et persistance du processus de production agrégé.....	31
Schéma 3	Facteurs influençant le comportement de la variable latente (importation).....	38
Schéma 4	Facteurs influençant le comportement de la variable latente (exportation).....	40

LISTE DES GRAPHIQUES

Graphique 1	Évolution de la valeur ajoutée nord-américaine, 1992-2001	22
Graphique 2	Évolution de la demande d'importation du Québec pour tous modes, 1992-2001 – (en M\$ CA).....	39
Graphique 3	Évolution de la demande d'exportation du Québec pour tous modes, 1992-2001 – (en M\$ CA).....	41

1. INTRODUCTION

La connaissance des facteurs multiples et complexes qui conditionnent la demande de transport est un atout important lorsque vient le temps de conseiller les décideurs publics. Elle permet, notamment, de mieux comprendre les liens entre l'économie, l'environnement et la disponibilité et la qualité des réseaux de transport. Or, comme il existe peu de données dans le domaine du transport des marchandises, la demande qui en résulte est généralement difficile à cerner. C'est pourquoi le ministère des Transports est à analyser différentes méthodes qui lui permettraient d'obtenir un portrait de cette demande. Certains travaux sur le sujet ont déjà été effectués sans qu'il en ressorte des résultats probants, et ce, pour diverses raisons.

Les derniers travaux portant sur ce sujet consistaient à répartir les flux économiques en partant des informations disponibles sur les flux de transport et les échanges économiques. Cette analyse a permis de faire ressortir l'existence d'un phénomène de transit ou « plaques tournantes » dans les données portant sur les flux de transport¹, lequel phénomène complique énormément l'exercice.

Dans le présent rapport, nous proposons une méthode macroéconomique afin de déterminer la demande de transport au Québec. Plus précisément, nous développons un modèle économétrique « à variable latente à indicateurs et causes multiples »² (MIMIC) afin d'estimer l'évolution des flux d'échanges économiques tous modes du Québec avec les régions du Canada et des États-Unis à partir de données portant sur l'industrie manufacturière, l'évolution démographique et l'évolution des revenus.

Nous optons pour une modélisation de type panel afin de tenir compte à la fois de l'information transversale et longitudinale. Cette méthodologie est en soi une innovation puisque, à notre connaissance, elle n'a pas encore été exploitée en relation avec des modèles MIMIC dans le domaine de l'économie.

L'objectif du projet est donc de comparer l'évolution de la demande de transport qui découle de l'application du modèle MIMIC à des données économiques liées à des entités géographiques pour lesquelles l'information est disponible auprès des organismes statistiques officiels. Dans le cas qui nous concerne, l'Institut de la statistique du Québec (ISQ) compile des données sur le commerce entre le Québec, le Canada et les États américains. Nous pourrions ainsi confronter les résultats obtenus selon le modèle aux statistiques officielles afin de déterminer les différences et les rapprochements entre les deux.

¹ À ce sujet, voir Dubé et Dupéré (2004).

² Traduction de l'expression anglaise « *Multiple Indicators and Multiple Causes* » (MIMIC).

Si les valeurs des marchandises transportées (ou flux économiques tous modes) dérivées du modèle concordent avec les flux économiques réels fournis par les organismes de statistiques, le modèle pourrait alors être subséquentement développé pour en dériver une matrice de déplacements (par mode) entre les régions du Québec et ses partenaires économiques de l'Amérique du Nord. Encore là, les résultats de cette démarche pourraient être comparés aux données réelles disponibles.

Le présent rapport est divisé en cinq grandes sections. D'abord, nous décrivons le cadre théorique de la modélisation proposée. Pour ce faire, nous commençons par présenter les caractéristiques d'une modélisation de type panel avant d'en montrer l'application dans les modèles MIMIC. Par la suite, nous présentons les données qui serviront à l'analyse et aux estimations. Il sera notamment question des définitions utilisées ainsi que des changements méthodologiques apportés dans les enquêtes effectuées auprès des manufacturiers. Nous dresserons également un portrait rapide du secteur manufacturier nord-américain ainsi que de l'évolution de la population et du revenu personnel.

Après avoir indiqué les particularités des données, nous présentons les résultats obtenus lors des estimations économétriques, autant pour les importations que pour les exportations. Nous utilisons, par la suite, ces résultats afin de reconstituer l'évolution des flux d'échanges et de les confronter aux données officielles disponibles. Finalement, nous concluons le rapport en énonçant les principaux constats tirés de l'analyse et les recommandations visant à orienter les travaux futurs.

2. LE MODÈLE THÉORIQUE

Le modèle théorique retenu s'inspire des travaux de Zellner (1970), Hauser et Goldberger (1971), Goldberger (1972), Jöreskog (1972) et Jöreskog et Goldberger (1975). Il est basé sur des hypothèses comportementales reliées à une variable inobservable. Nous avons décidé d'utiliser cette méthodologie en l'adaptant à un univers jumelant les données transversales et longitudinales.

Les données combinant les informations longitudinales (temporelles) et transversales (caractéristiques individuelles) sont appelées **données de panel**. L'avantage à utiliser de telles données tient à la taille de l'échantillon disponible qui est nettement plus grande, ce qui permet, notamment, d'obtenir des résultats d'estimation plus fiables. En effet, en utilisant la notation N pour désigner le nombre total d'individus observés et T pour désigner le nombre total de périodes recensées, les données de panel nous procurent un échantillon total disponible aux estimations de NT observations.

De plus, les données de panel permettent d'avoir recours aux procédures à composantes d'erreurs qui reposent sur la décomposition du terme d'erreur en deux effets spécifiques : un effet représentant le terme aléatoire classique d'un modèle de régression linéaire et un effet propre à chaque individu n . En termes techniques, cela signifie que ces types de modèles permettent de contrôler l'hétérogénéité individuelle, c'est-à-dire les différences entre les individus. Cette spécificité est particulièrement importante puisqu'elle entraîne des biais dans les résultats d'estimation lorsqu'elle est ignorée.

Cette méthodologie, largement répandue et documentée en économique³, a l'avantage de permettre l'utilisation d'une source de variation plus riche ainsi qu'une estimation plus efficace des paramètres. L'interaction entre les données longitudinales et transversales permet de mieux déceler et estimer des effets qui ne sont normalement pas détectables dans des données à une seule dimension.

Le corps du modèle repose sur la théorie des modèles à variable latente (plus précisément les modèles MIMIC) qui permet de quantifier une variable non observable et d'en décrire l'évolution. Ce type de modélisation nous semble parfaitement approprié dans le cas présent puisque nous tentons d'estimer les flux d'échanges économiques du Québec en nous basant sur des informations provenant du secteur manufacturier⁴.

³ Notamment dans Greene (2003), Woolridge (2000, 2002) et Baltagi (1999).

⁴ En aucun cas nous ne faisons référence à des données portant spécifiquement sur les transports.

En science économique, les modèles MIMIC ont, jusqu'à présent, surtout été exploités pour étudier l'évolution de l'économie souterraine (Giles 1999a, 1999b; Giles et Tedds 2002; Giles, Tedds et Werknehg 2002; Tedds 1998, 2004). L'idée principale sous-jacente à ce genre de modèle est d'utiliser des variables observables afin de statuer sur la forme et la grandeur de la variable latente. Deux types de variables sont considérés dans cette modélisation.

On a d'abord les variables de « causes » ou *causales* qui provoquent la variable latente, c'est-à-dire qu'elles expliquent le comportement de la variable non observable. Viennent ensuite les variables « *indicatrices*⁵ » qui sont la résultante de la variable latente. De manière générale, la relation unissant les différents types de variables peut être illustrée par un simple schéma (schéma 1) :

Schéma 1 Illustration de la chaîne de relations du modèle MIMIC



Il est à noter que les résultats du modèle dépendent, en grande partie, du choix des variables de *causes* et des variables *indicatrices*. Étant donné que la variable latente ne peut être mesurée, il n'est pas possible de démontrer formellement que la séquence hypothétique de cause à effet tienne. C'est pour cette raison que cette étape est particulièrement cruciale dans ce type de modélisation.

Formellement, la chaîne d'événements s'exprime, en première étape, pour un individu n et une période t par une équation de comportement de la variable latente $y_{n,t}^*$. Cette équation, que nous désignerons par le terme « modèle structurel » (ou équation structurelle), est décrite par la relation (1).

$$y_{n,t}^* = x_{1,n,t}\alpha_1 + \dots + x_{k,n,t}\alpha_k + \dots + x_{K,n,t}\alpha_K + \varepsilon_{n,t} \quad (1)$$

où $y_{n,t}^*$ est la variable que l'on cherche à expliquer, $x_{k,n,t}$ est la k ième variable de *cause* et les α_k sont les paramètres à estimer. Le terme d'erreur de la relation, $\varepsilon_{n,t}$, est supposé de moyenne nulle et de variance σ_ε^2 .

⁵ Le terme « indicatrice », dans le cas présent, ne fait pas référence aux variables muettes (binaires) souvent utilisées dans les analyses de régressions. Il s'agit d'une interprétation différente.

La relation entre les variables *indicatrices* et la variable non observable, quant à elle, nous procure un système d'équations que nous désignerons par l'expression « modèle de mesure » (ou équations de mesure). Nous exprimons cette relation par l'équation (2) :

$$\begin{aligned}
 y_{1,n,t} &= y_{n,t}^* \beta_1 + u_{1,n,t} \\
 \dots & \\
 y_{m,n,t} &= y_{n,t}^* \beta_m + u_{m,n,t} \\
 \dots & \\
 y_{M,n,t} &= y_{n,t}^* \beta_M + u_{M,n,t}
 \end{aligned} \tag{2}$$

où $y_{m,n,t}$ est la m ième variable indiquant le comportement de la variable latente et les β_m sont les paramètres à estimer.

Nous spécifions un terme d'erreur propre à chaque variable *indicatrice*, $u_{m,n,t}$, et supposons, pour le moment, que ces termes sont de moyenne nulle et sont indépendants entre eux. Le terme d'erreur se décompose en deux effets distincts afin d'incorporer la dimension panel au modèle, soit un effet propre à chaque individu et un effet aléatoire global, ce qu'illustre l'équation (3) :

$$u_{m,n,t} = \mu_{m,n} + v_{m,n,t} \tag{3}$$

où $\mu_{m,n}$ est une composante propre à chaque variable *indicatrice* et à chaque individu, mais fixe dans le temps⁶, et $v_{m,n,t}$ est l'équivalent du terme d'erreur traditionnel des modèles de régressions linéaires multiples.

Dans la littérature, le terme μ peut être traité de différentes façons, selon que l'on considère cet effet, propre à un individu, comme étant aléatoire ou fixe. Si nous avons de bonnes raisons de croire que ce paramètre peut être estimé, alors nous devrions opter pour une modélisation avec effets fixes. Cependant, lorsque l'on considère que le paramètre μ est le résultat d'une variable aléatoire, alors nous devrions opter pour une modélisation avec effets aléatoires.

⁶ D'où la notation n référant seulement à l'individu et l'absence d'indice marquant l'espace temporel (t).

Dans le cadre de la présente analyse, l'effet sera considéré comme étant fixe dans le temps et représentera une hétérogénéité non observable entre les régions⁷. Cette hétérogénéité peut être causée par des différences structurelles, industrielles ou historiques.

En substituant les équations (1) et (3) dans l'équation (2), il est possible d'écrire le modèle sous sa forme réduite, comme le montre l'équation (4). De cette façon, le système d'équations exprime les variables *indicatrices* en fonction des variables de *causes* :

$$\begin{aligned}
 y_{1,n,t} &= x_{1,n,t}\Pi_{11} + \dots + x_{k,n,t}\Pi_{1k} + \dots + x_{K,n,t}\Pi_{1K} + \mu_{1,n} + \xi_{1,n,t} \\
 \dots & \\
 y_{m,n,t} &= x_{1,n,t}\Pi_{m1} + \dots + x_{k,n,t}\Pi_{mk} + \dots + x_{K,n,t}\Pi_{mK} + \mu_{m,n} + \xi_{m,n,t} \quad (4) \\
 \dots & \\
 y_{M,n,t} &= x_{1,n,t}\Pi_{M1} + \dots + x_{k,n,t}\Pi_{Mk} + \dots + x_{K,n,t}\Pi_{MK} + \mu_{M,n} + \xi_{M,n,t}
 \end{aligned}$$

où Π_{mk} sont les coefficients à estimer que l'on peut exprimer sous forme matricielle :

$$\Pi = \begin{bmatrix} \Pi_{11} & \dots & \Pi_{m1} & \dots & \Pi_{M1} \\ \vdots & \ddots & \vdots & & \vdots \\ \Pi_{1k} & \dots & \Pi_{mk} & \dots & \Pi_{Mk} \\ \vdots & & \vdots & \ddots & \vdots \\ \Pi_{1K} & \dots & \Pi_{mK} & \dots & \Pi_{MK} \end{bmatrix} \quad (5)$$

Les coefficients de la forme réduite sont donc formés par une interaction entre les coefficients du modèle structurel et ceux du modèle de mesure.

Formellement, la relation entre les coefficients α_k et β_m afin de former les coefficients Π_{mk} est résumée par l'équation (6) :

$$\Pi_{mk} = \beta_m \alpha_k \quad (6)$$

Le terme d'erreur de l'équation (4), quant à lui, est défini par l'équation (7) :

$$\xi_{m,n,t} = \varepsilon_{n,t} \beta_m + \nu_{m,n,t} \quad (7)$$

⁷ La modélisation à effets fixes est classiquement retenue lorsqu'on utilise des données portant sur des entités géographiques.

La matrice de termes d'erreurs (7) est de moyenne nulle et sa matrice de variance-covariance est donnée par l'équation (8) :

$$E(\xi'\xi) = \begin{bmatrix} \beta_1^2 \sigma_\varepsilon^2 + \sigma_{v_1}^2 & \cdots & \beta_1 \beta_m \sigma_\varepsilon^2 & \cdots & \beta_1 \beta_M \sigma_\varepsilon^2 \\ \vdots & \ddots & \vdots & & \vdots \\ \beta_1 \beta_m \sigma_\varepsilon^2 & \cdots & \beta_m^2 \sigma_\varepsilon^2 + \sigma_{v_m}^2 & \cdots & \beta_m \beta_M \sigma_\varepsilon^2 \\ \vdots & & \vdots & \ddots & \vdots \\ \beta_1 \beta_M \sigma_\varepsilon^2 & \cdots & \beta_m \beta_M \sigma_\varepsilon^2 & \cdots & \beta_M^2 \sigma_\varepsilon^2 + \sigma_{v_M}^2 \end{bmatrix} \quad (8)$$

La forme réduite finale du modèle peut s'écrire, de façon matricielle, comme étant :

$$y = x\Pi + Z_\mu \mu + \xi \quad (9)$$

où y est une matrice de variables *indicatrices* de dimension $(NT \times M)$, x est une matrice de variables de *causes* de dimension $(NT \times K)$, Π est la matrice de coefficients à estimer de dimension $(K \times M)$, ξ est une matrice de termes d'erreurs de dimension $(NT \times M)$, μ est une matrice d'effets fixes par individu et par variable de *causes* de dimension $(N \times M)$ et Z_μ est une matrice indicatrice⁸ de dimension $(NT \times N)$.

L'étude de cette forme réduite finale révèle que le modèle est suridentifié puisque la matrice de coefficient de la forme réduite, de dimension $(K \times M)$, est en fait une composition du produit du vecteur α , de dimension $(K \times 1)$, et du vecteur β , de dimension $(1 \times M)$. Autrement dit, les $K \times M$ paramètres à estimer sont exprimés en termes de seulement $K + M$ paramètres.

De plus, les $\frac{1}{2}M(M+1)$ paramètres de la matrice variance-covariance peuvent être exprimés en termes d'un scalaire, σ_ε^2 , du vecteur β et des M éléments non nuls de la matrice de variance-covariance du modèle de mesure

(2). En somme, les $\frac{1}{2}M(M+1)$ paramètres distincts sont exprimés en termes de seulement $1 + 2M$ paramètres.

⁸ Formée de uns et de zéros.

Au moment des estimations, nous adopterons une approche nous permettant de composer avec ce problème de suridentification. Cette approche sera expliquée plus en détail au cours des discussions portant sur les estimations empiriques.

3. LES DONNÉES MANUFACTURIÈRES

Cette section comprend l'analyse des données que nous avons considérées lors de l'estimation du modèle MIMIC. Celles-ci se divisent en deux grandes catégories : les variables reliées à la consommation, c'est-à-dire la population et le revenu personnel, et les variables reliées à la production, c'est-à-dire les données manufacturières. Les trois premières sous-sections porteront plus spécialement sur les concepts des variables utilisées ainsi que sur la description des enquêtes canadiennes et américaines et les modifications méthodologiques apportées aux enquêtes au fil des années. Les deux dernières sous-sections dresseront un portrait du territoire canado-américain tout en portant une attention particulière à la situation du Québec.

3.1 Définitions des concepts

Avant même de présenter les données utilisées lors de l'analyse statistique, il nous semble important de s'arrêter sur la définition des concepts de chacune des variables retenues. Cette sous-section vise donc la description des variables que nous avons considérées lors de la construction des indices de déplacements de marchandises.

La définition des statistiques est semblable pour le Canada et les États-Unis. Plusieurs concepts pertinents sont utilisés dans les deux sources de données. À titre indicatif, nous définissons, pour le reste de l'analyse, le **coût du matériel, des fournitures et des marchandises** (ci-après noté **coût des intrants**⁹) comme faisant référence à toute charge directe payée, ou payable, pour des éléments consommés ou mis en production durant l'année incluant notamment les frais de transport, les coûts en carburant et en électricité. Autrement dit, il s'agit de la charge des matériaux et du coût des équipements servant à la transformation. Les **expéditions manufacturières** couvrent la valeur de vente nette reçue, ou à recevoir, de tous les produits expédiés, que ce soit les produits primaires ou secondaires, aussi bien que toute autre entrée telle que des paiements pour des contrats de travail effectués pour d'autres ou la vente de rebus. La **valeur ajoutée**, quant à elle, est une mesure de la production nette, c'est-à-dire qu'elle soustrait les achats d'intrants (incluant le coût des combustibles et de l'électricité) de la production brute.

Les statistiques manufacturières sont divisées en deux secteurs lors des analyses. Le **secteur non productif** correspond aux employés de l'établissement manufacturier qui ne sont pas directement engagés dans la production tels ceux affectés à la supervision de l'usine à un niveau supérieur, aux superviseurs de lignes, aux employés du service des ventes, de la publicité, de l'administration, et ainsi de suite. Le **secteur productif**, quant à lui, correspond aux employés de l'établissement manufacturier qui sont

⁹ Cette classe de coût des intrants fait référence au capital physique uniquement. Il n'est pas question du coût du capital humain.

directement engagés dans la production à un niveau inférieur, ou égal, aux superviseurs de lignes, par exemple les employés engagés dans la fabrication, le développement et l'assemblage du produit.

Pour chacun de ces secteurs, certains indicateurs, tels que l'emploi et la masse salariale, sont disponibles. L'**emploi manufacturier** correspond aux travailleurs, à temps plein et à temps partiel, sur la liste de paie d'une entreprise manufacturière au cours de certains mois déterminés dans l'année¹⁰. La **masse salariale** se définit par les revenus bruts des employés sur la liste de paie d'un établissement manufacturier. Les revenus bruts incluent toutes formes de compensations telles que, notamment, le salaire, les commissions, les bonus et les paies de vacances.

Pour ce qui est des individus, l'**emploi total** se définit comme le nombre de personnes qui effectuent un travail quelconque contre une rémunération ou en vue d'un bénéfice et la **population active** comme la population civile des personnes de 15 ans et plus (16 ans et plus pour les États-Unis), qui ont un emploi ou sont au chômage, à l'exclusion des pensionnaires d'établissements.

Finalement, le **revenu personnel** correspond à la somme de tous les revenus touchés par les individus, y compris la rémunération du travail et le rendement des investissements ainsi que les transferts des administrations et d'autres secteurs (ce qui inclut les prestations de la sécurité de la vieillesse et d'assurance-emploi). Le **revenu personnel disponible**, quant à lui, est la portion de revenu qui reste après le paiement des impôts directs personnels, y compris l'impôt sur le revenu, les cotisations aux régimes d'assurance sociale et autres frais. En bref, il s'agit d'un bon indicateur des fonds disponibles pour les dépenses personnelles en biens et services et l'épargne personnelle aux fins d'investissements.

3.2 Les données canadiennes

Les données canadiennes proviennent de *Statistique Canada*, l'organisme officiel de statistiques du gouvernement du Canada.

Les **données manufacturières** proviennent de l'enquête annuelle des manufactures (EAM). Il s'agit d'un sondage mené tous les ans depuis 1917 auprès des entreprises de l'industrie manufacturière du Canada. Cette enquête, qui procède par échantillonnage, est administrée dans le cadre du Programme des enquêtes unifiées auprès des entreprises (EUE) et vise tous les établissements incorporés et non incorporés où l'activité principale est la fabrication (environ 100 000 établissements). Au fil des années, certaines modifications ont été intégrées, ce qui affecte, par le fait même, la comparabilité chronologique des données.

¹⁰ Mars, mai, août et novembre pour les travailleurs à la production et mars pour les travailleurs non reliés à la production.

La première modification conceptuelle et méthodologique significative, visant notre échantillon, a été intégrée en 2000. Avant cette année, l'EAM fournissait des estimations des principales statistiques financières de toutes les entreprises manufacturières constituées en société qui avaient des employés et dont les ventes de biens manufacturés étaient égales ou supérieures à 30 000 \$. L'univers d'enquête s'est depuis étendu à l'ensemble des unités manufacturières, ce qui signifie un ajout d'environ 60 000 unités. Ces unités supplémentaires représentent toutefois moins de 5 % des recettes totales provenant de la vente de biens manufacturés.

De plus, à compter de l'année 2000, les données des sièges sociaux ne sont plus incluses, ce qui a des incidences notamment sur le nombre d'employés de l'administration, le coût des intrants, la valeur des livraisons, les autres revenus et la valeur ajoutée totale¹¹.

Les **données sur l'emploi** proviennent de l'enquête sur la population active (EPA). Ce sondage fut introduit à la suite de la Deuxième Guerre mondiale afin de répondre à un besoin d'information sur le marché du travail. Cette enquête, qui est fréquemment utilisée par plusieurs organismes, sert notamment à produire certains indicateurs macroéconomiques importants tels que le taux de chômage, le taux d'emploi et le taux d'activité.

La version actuelle du questionnaire a été introduite en 1997 afin de combler certaines lacunes statistiques et d'accroître la qualité des données. Les modifications apportées comprenaient l'ajout de nombreuses questions améliorant ainsi la précision des statistiques. L'EPA procède par échantillonnage et est représentative de la population canadienne civile non institutionnalisée de 15 ans et plus. Notons que l'enquête exclut les résidents du Yukon, des Territoires du Nord-Ouest et du Nunavut, les personnes qui vivent dans les réserves indiennes, les membres à temps plein des Forces armées canadiennes et les pensionnaires d'établissements.

Les **statistiques démographiques** proviennent de deux sources. Tout d'abord, le recensement, réalisé tous les cinq ans, nous permet d'obtenir les chiffres réels de population pour ces années. Entre les deux recensements, la population est estimée en utilisant comme base le recensement le plus récent auquel il faut ajouter le nombre de naissances, le nombre d'immigrants et le solde des résidents non permanents, puis retrancher le nombre des décès et d'émigrants. Ensuite, pour chaque province, il faut en outre ajouter le solde migratoire interprovincial afin d'aboutir à une estimation postcensitaire de la population.

¹¹ Toutefois, Statistique Canada ne chiffre pas l'ampleur de cette modification sur les variables mentionnées.

Finalement, les **données sur les revenus** proviennent des comptes économiques provinciaux. Ces statistiques, qui servent à produire des estimations annuelles de certains agrégats des comptes des revenus et dépenses des provinces et des territoires, proviennent de l'agrégation des résultats de plusieurs enquêtes.

3.3 Les données américaines

Les données américaines, quant à elles, proviennent de trois organismes différents : le *Bureau of Labor Statistics (BLS)*, le *US Census Bureau* et le *Bureau of Economic Analysis (BEA)*.

Les **données manufacturières américaines** proviennent du *US Census Bureau* et sont colligées grâce à deux sources. Celle qui permet d'obtenir les données comportant le plus de détails sur le secteur manufacturier est le recensement, qui a lieu tous les cinq ans. Cet exercice est la source principale d'information à propos de la structure et du fonctionnement de l'économie américaine. Lors du recensement, un formulaire est envoyé à plus de 5 millions de compagnies américaines selon 600 versions de ce formulaire, chacune étant particulière à une industrie.

Pour chacune des quatre années séparant les recensements, un sondage annuel sur les établissements manufacturiers est mené. Ce sondage, appelé *Annual Survey of Manufactures* (ci-après ASM) procède à l'aide d'un échantillon qui est sélectionné tous les cinq ans et vise tous les établissements d'au moins un employé¹². Un nouvel échantillon est introduit la deuxième année suivant chacun des recensements.

Un changement important de méthodologie est survenu en 1997, ce qui affecte la comparabilité des données temporelles. Afin de se conformer à une classification standard à l'échelle nord-américaine, le *Census Bureau* a réparti les différentes industries selon le code SCIAN (Système de classification des industries nord-américaines) plutôt que selon le code SIC (Standard Industrial Classification). En conséquence, plus de la moitié des industries dans le secteur manufacturier a changé de classification à partir de cette date.

Bien que la majorité des changements ont été effectués à l'intérieur de la classe manufacturière, certaines nouvelles entreprises sont entrées dans le secteur manufacturier alors que d'autres l'ont quitté. Par exemple, l'exploitation forestière (*logging*) et une partie de l'édition (*publishing*) ont quitté le secteur manufacturier alors que les boulangeries (*bakery*), les confiseries (*candy store*) et les tailleurs (*tailor*) l'ont intégré. L'effet net du changement de classification

¹² Un établissement manufacturier est défini comme un établissement qui transforme mécaniquement, physiquement ou chimiquement des matériaux ou substances en un nouveau produit.

sur les expéditions manufacturières totales est une baisse estimée à 3 %¹³ du niveau.

Cette modification entraîne également des conséquences sur certains autres indicateurs manufacturiers. À ce sujet, nous pouvons observer certains bris majeurs dans les séries chronologiques. À titre d'exemple, les établissements manufacturiers auxiliaires¹⁴ ne sont plus inclus dans les données manufacturières. Cette restructuration a eu pour effet de faire diminuer le nombre d'entreprises dans le secteur manufacturier.

La diminution du nombre d'entreprises recensées a, incidemment, affecté à la baisse les emplois non productifs et, par ricochet, la masse salariale de cette catégorie. Bien que la tendance des séries chronologiques semble relativement stable pour les années subséquentes à 1997, il demeure que nous observons un bris important entre 1996 et 1997 dans les séries de données portant le secteur « non productif ».

Les **données sur l'emploi** sont compilées par le *BLS* et proviennent du *Current Employment Statistics (CES)*, un sondage couvrant les listes de paie de plus de 390 000 entreprises portant sur l'emploi, les heures travaillées et les salaires et traitements des travailleurs. Les statistiques issues de cette source de données sont représentatives de la population civile non institutionnalisée. À noter que, contrairement au Canada, les statistiques américaines incluent les individus de 16 ans et plus¹⁵ et non de 15 ans et plus.

Les **statistiques démographiques** proviennent du *US Census Bureau*. Tout comme pour le Canada, les données reposent sur deux sources. La première est le recensement, qui a lieu tous les cinq ans. Entre ces années, la population est estimée sur la base du résultat du recensement le plus récent auquel il faut ajouter ou soustraire selon le cas, les naissances, les décès, la migration domestique et internationale nette et le mouvement net des forces armées.

Finalement, **les données de revenus** sont récoltées par le *BEA* et proviennent essentiellement des registres administratifs et du recensement, ou de sondages, du même type. À titre d'exemple, les données servant à l'estimation du revenu et du revenu disponible proviennent d'organismes aussi variés que le département du trésor américain, le département de la santé et des services sociaux ainsi que du programme fédéral d'impôt sur le revenu.

¹³ Cet effet est récurrent pour les années subséquentes à 1997.

¹⁴ Les établissements auxiliaires peuvent être définis comme des entités servant principalement un autre établissement de la même entreprise. Dans la classification SIC, ces établissements étaient classés dans la même industrie que l'établissement qu'ils desservait alors qu'avec le SCIAN, ils sont plutôt classés en fonction du service qu'ils offrent.

¹⁵ L'âge minimal pour travailler, aux États-Unis, est fixé à 16 ans.

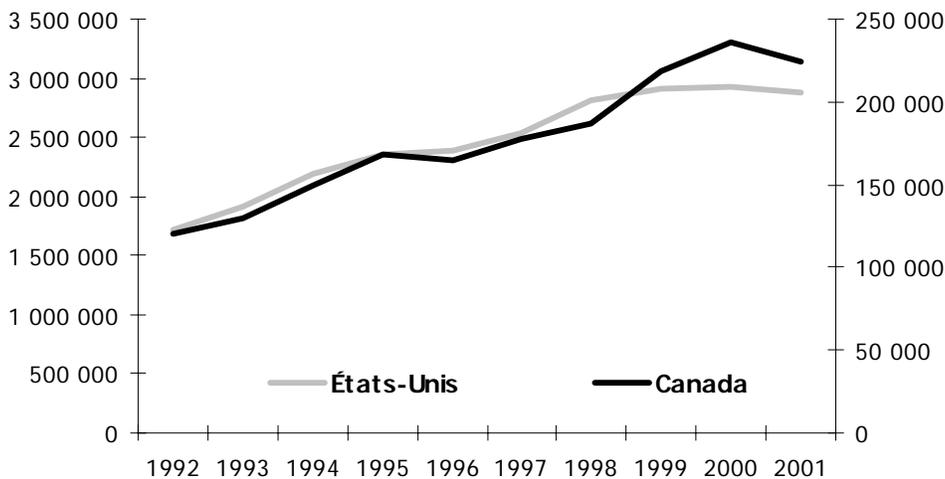
3.4 Survol du secteur manufacturier nord-américain

Le portrait du secteur manufacturier nord-américain a considérablement changé au cours de la décennie 1990-2000. Ce phénomène est principalement dû à l'ouverture des frontières économiques et à la mondialisation des marchés, les producteurs ayant tendance à s'orienter vers des bassins de production possédant le plus d'avantages comparatifs. Aux États-Unis, ce mouvement est encore plus prononcé qu'au Canada puisque la croissance de la valeur ajoutée manufacturière est moins importante (graphique 1).

Cependant, la valeur ajoutée de ce secteur n'est qu'un indicateur des performances de l'industrie. Afin de dresser un portrait plus exhaustif, il importe de tenir compte de l'évolution du nombre d'emplois, des salaires versés, du coût des intrants et des expéditions manufacturières. Pour ce faire, nous concentrerons l'analyse des données sur la désagrégation régionale retenue lors de la modélisation.

Comme première analyse, nous pouvons constater que le Québec se classe dans les derniers échelons en ce qui concerne son importance dans l'ensemble du secteur manufacturier nord-américain, sur la base du découpage géographique retenu ici. Seules les régions de l'Atlantique, de l'Ouest canadien du New Hampshire/Maine et du Vermont sont d'importance relative moindre que le Québec dans l'ensemble nord-américain (tableau 1).

Graphique 1 Évolution de la valeur ajoutée nord-américaine, 1992-2001
En M\$ CA (échelle de droite – Canada; échelle de gauche – Etats-Unis)



Sources : Statistique Canada et US Census Bureau.

Tableau 1 Importance relative dans l'activité manufacturière nord-américaine, 1992-2001

Régions	Emplois à la production	Salaires à la production	Coût des intrants	Expéditions manufacturières	Valeur ajoutée
Ouest américain	22,5 %	21,9 %	23,2 %	23,7 %	24,3 %
Midwest	21,1 %	24,6 %	22,1 %	21,8 %	21,5 %
Sud	15,5 %	14,6 %	18,9 %	17,5 %	16,0 %
Sud-Est	15,3 %	13,8 %	13,5 %	14,1 %	14,8 %
New York/New Jersey	6,2 %	6,3 %	5,4 %	6,3 %	7,2 %
Pennsylvanie	4,4 %	4,6 %	3,9 %	4,1 %	4,5 %
Nouvelle-Angleterre	3,5 %	3,8 %	2,7 %	3,3 %	4,0 %
Ontario	5,1 %	4,9 %	5,4 %	4,6 %	3,5 %
Québec	2,8 %	2,2 %	2,1 %	2,0 %	1,7 %
Ouest canadien	2,0 %	1,8 %	1,7 %	1,5 %	1,2 %
New Hampshire/Maine	1,0 %	1,0 %	0,7 %	0,8 %	0,9 %
Atlantique	0,5 %	0,4 %	0,4 %	0,3 %	0,3 %
Vermont	0,2 %	0,2 %	0,2 %	0,2 %	0,2 %
	100 %	100 %	100 %	100 %	100 %

Sources : Statistique Canada et US Census Bureau

Les deux plus gros joueurs du secteur manufacturier représentent, *grosso modo*, près de 45 % du total des grands indicateurs. L'activité manufacturière nord-américaine est surtout concentrée dans l'Ouest et le Midwest américain et laisse peu de place aux autres régions. Une façon d'étudier cette facette est de mesurer l'indice de concentration des quatre (C4), des huit (C8) et des douze (C12) régions les plus importantes¹⁶.

Les quatre régions les plus importantes en termes d'activité manufacturière englobent environ 75 % des emplois des salaires totaux versés à la production. Cette proportion monte à environ 77 % quand on considère le coût des intrants, les expéditions manufacturières et la valeur ajoutée (tableau 2).

Tableau 2 Indice de concentration des industries manufacturières nord-américaines, 1992-2001

	Emplois à la production	Salaires à la production	Coût des intrants	Expéditions manufacturières	Valeur ajoutée
C4	74,3 %	74,8 %	77,6 %	77,0 %	76,6 %
C8	93,5 %	94,4 %	95,0 %	95,3 %	95,7 %
C12	99,8 %	99,8 %	99,8 %	99,8 %	99,8 %

Sources : Statistique Canada et US Census Bureau
Calculs : JDMD Groupe conseils inc.

¹⁶ Cet indice de concentration s'inspire des mesures C4 et C8 connues classiquement en organisation industrielle. Voir notamment Carlton et Perloff (2000), p. 247-250, à cet effet.

Qui plus est, les huit régions les plus importantes en termes d'activité manufacturière représentent environ 95 % des activités nord-américaines totales. Les cinq régions les moins importantes dans le secteur manufacturier, dont fait partie le Québec, représentent donc environ 5 % des activités totales, soit en moyenne 1 % par région, ce qui indique un fort taux de concentration des industries.

À l'intérieur du Canada, le Québec représente environ le quart de la valeur ajoutée totale (tableau 3). L'Ontario représente, pour sa part, plus de la moitié de la valeur ajoutée totale canadienne.

L'étude transversale permet de déterminer l'importance relative du Québec dans un contexte nord-américain et canadien. Cependant, elle ne nous apprend rien sur la tendance des principaux indicateurs du secteur manufacturier pour les périodes recensées. L'évolution temporelle est particulièrement intéressante puisqu'elle nous permet de capter une autre dimension de la problématique.

Tableau 3 Importance relative dans l'activité manufacturière canadienne, 1992-2001

Régions	Emplois à la production	Salaires à la production	Coût des intrants	Expéditions manufacturières	Valeur ajoutée
Ontario	49,0 %	52,5 %	56,9 %	54,8 %	52,3 %
Québec	26,6 %	23,9 %	21,6 %	23,3 %	25,7 %
Ouest canadien	19,3 %	19,4 %	17,5 %	17,9 %	18,2 %
Atlantique	5,0 %	4,3 %	4,0 %	3,9 %	3,7 %
	100 %	100 %	100 %	100 %	100 %

Sources : Statistique Canada

Le Québec a été la région qui a connu le plus fort taux de croissance en ce qui concerne les expéditions manufacturières entre 1992 et 2001 (tableau 4¹⁷). Cette forte croissance peut, en partie, s'expliquer par l'apparition de l'ALENA¹⁸ en janvier 1994. Cette hypothèse est encore plus probable si l'on considère que les deux seules régions ayant connu un taux de croissance des expéditions manufacturières supérieur à 7 % sont le Québec (7,39 %) et l'Ouest canadien (7,19 %). Parmi les six régions ayant connu le plus fort taux de croissance des expéditions manufacturières, on trouve trois régions canadiennes.

¹⁷ Évidemment, les résultats obtenus ne sont pas indépendants de la période sélectionnée.

¹⁸ ALENA : Accord de libre-échange nord-américain.

Tableau 4 Croissance moyenne des principaux indicateurs manufacturiers, 1992-2001 (en %)

Régions	Emplois à la production	Salaires à la production	Coût des intrants	Expéditions manufacturières	Valeur ajoutée
Vermont	-0,19	5,64	5,05	6,50	8,52
Québec	3,09	4,33	7,59	7,39	7,43
Ontario	1,90	4,39	4,80	5,72	7,33
Ouest canadien	3,35	4,40	7,16	7,19	7,27
New Hampshire/Maine	-0,75	4,98	6,68	6,91	7,03
Ouest américain	-0,62	3,94	6,02	6,37	6,78
Atlantique	2,95	3,66	6,37	6,40	6,76
Sud-Est	-2,20	2,63	5,58	5,93	6,22
Sud	-0,77	4,42	6,88	6,62	6,20
Midwest	-1,43	2,85	5,94	5,86	5,72
Pennsylvanie	-2,60	1,77	5,60	5,41	5,17
Nouvelle-Angleterre	-3,40	1,41	4,82	4,74	4,72
New York/New Jersey	-5,17	-1,96	3,39	2,51	1,92

Sources : Statistique Canada et US Census Bureau
Calculs : JDMD Groupe conseils inc.¹⁹

Le Québec a également enregistré, en importance, le second taux de croissance de la valeur ajoutée manufacturière nord-américaine pour la période 1992-2001. Seul l'État du Vermont a affiché un meilleur taux de croissance moyen que le Québec. À l'exception de l'Atlantique, les autres régions canadiennes ont offert un taux de croissance de la valeur ajoutée manufacturière supérieur à celui des régions américaines (sauf le Vermont).

Afin de soutenir cette croissance du secteur manufacturier au Québec, le processus de production a nécessité une croissance des intrants de production. Le Québec est la région ayant compilé le taux le plus élevé en ce qui a trait à la croissance du coût des intrants. L'Ouest canadien est la deuxième région où le taux de croissance a été le plus important.

L'emploi à la production a également connu une forte progression au Québec, montrant, en importance, le second taux de croissance des régions considérées, juste derrière l'Ouest canadien. La masse salariale des travailleurs à la production a aussi connu une progression, mais relativement moins importante que les autres indicateurs. En termes d'importance, le Québec se classe sixième derrière le Vermont, le New Hampshire/Maine, le sud des États-Unis, l'Ouest canadien et l'Ontario.

¹⁹ L'estimation du taux de croissance moyen est basée sur une capitalisation pour une période infinie. À ce sujet, on peut consulter notamment Charest, Lusztig et Schwab (2000), p. 55.

Cette brève description des données manufacturières permet de constater que, malgré leur importance relative faible, les régions du Québec et de l'Ouest canadien ont connu une croissance importante du secteur manufacturier entre 1992 et 2001. Or, la croissance simultanée de l'ensemble des indicateurs manufacturiers cache certaines subtilités.

La forte croissance du coût des intrants a fait en sorte que la croissance de l'indice du coût unitaire de production²⁰ a été plus élevée au Québec que dans l'ensemble des régions canadiennes. L'importance de la croissance de l'emploi à la production indique que le Québec a connu une croissance de la productivité du travail (dans le secteur manufacturier) plus faible que l'ensemble des autres régions nord-américaines. La croissance plus modérée des salaires à la production, comparativement à la croissance des emplois à la production, nous apprend que le Québec a connu une croissance du salaire moyen deux fois plus faible que celle de l'Ontario. Il va sans dire que ces constats ne sont pas sans importance pour le comportement futur de l'industrie manufacturière québécoise.

3.5 Survol des autres variables

Sur les aspects de la démographie, de l'emploi et de la population active, les États-Unis représentent environ neuf fois la taille du Canada (tableau 5). Le bassin de consommation ou de demande potentielle est donc nettement plus grand chez nos voisins du Sud. De plus, la population américaine a connu une croissance démographique légèrement supérieure à celle du Canada pour la période analysée, ce qui a fait croître lentement l'écart de population entre les deux pays.

²⁰ Défini comme le rapport du coût des intrants sur la valeur ajoutée.

**Tableau 5 Importance relative des Etats-Unis par rapport au Canada
1992-2001
Ratio des indicateurs (Etats-Unis sur Canada)**

	Population totale	Population 15 ans et +	Population 25-54 ans	Emplois totaux	Population active
1992	9,1	8,9	8,6	9,3	9,0
1993	9,1	8,9	8,6	9,5	9,0
1994	9,1	8,9	8,6	9,5	9,0
1995	9,1	8,9	8,6	9,5	9,1
1996	9,1	8,8	8,6	9,6	9,1
1997	9,1	8,8	8,6	9,6	9,1
1998	9,2	8,8	8,6	9,5	9,1
1999	9,2	8,8	8,6	9,4	9,0
2000	9,2	9,0	8,8	9,3	9,0
2001	9,2	9,0	8,8	9,2	8,9

Sources : Statistique Canada, US Census Bureau et Bureau of Labor Statistics
Calculs : JDMD Groupe conseils inc.

De plus, il est possible de constater que la concentration de la population et de l'emploi se répartit de manière semblable à l'activité manufacturière sur le territoire canado-américain (tableau 6). Ainsi, les régions les plus intensives en activité manufacturière sont également celles où l'on trouve les plus grands bassins de population, d'emploi et de population active.

**Tableau 6 Importance relative des régions nord-américaines
selon les caractéristiques, 1992-2001**

Régions	Population	Population 25-54 ans	Population 15 ans et +	Emploi	Population active	Expéditions manufacturières	Valeur ajoutée
Ouest américain	25,3 %	25,3 %	25,0 %	25,6 %	25,6 %	23,7 %	24,3 %
Midwest	14,8 %	14,7 %	14,8 %	15,2 %	15,1 %	21,8 %	21,5 %
Sud	16,1 %	15,7 %	15,8 %	15,6 %	15,6 %	17,5 %	16,0 %
Sud-Est	16,4 %	16,4 %	16,5 %	16,5 %	16,3 %	14,1 %	14,8 %
New York/New Jersey	8,9 %	9,0 %	9,0 %	8,6 %	8,7 %	6,3 %	7,2 %
Pennsylvanie	4,1 %	3,9 %	4,2 %	4,0 %	4,0 %	4,1 %	4,5 %
Nouvelle-Angleterre	3,5 %	3,6 %	3,6 %	3,7 %	3,7 %	3,3 %	4,0 %
Ontario	3,7 %	3,9 %	3,8 %	3,7 %	3,8 %	4,6 %	3,5 %
Québec	2,4 %	2,6 %	2,5 %	2,2 %	2,4 %	2,0 %	1,7 %
Ouest canadien	2,9 %	3,0 %	2,9 %	3,0 %	3,0 %	1,5 %	1,2 %
New Hampshire/Maine	0,8 %	0,8 %	0,8 %	0,9 %	0,9 %	0,8 %	0,9 %
Atlantique	0,8 %	0,8 %	0,8 %	0,7 %	0,7 %	0,3 %	0,3 %
Vermont	0,2 %	0,2 %	0,2 %	0,2 %	0,2 %	0,2 %	0,2 %

Sources : Statistique Canada, US Census Bureau et Bureau of Labor Statistics
Calculs : JDMD Groupe conseils inc.

Finalement, le revenu personnel et le revenu personnel disponible ont aussi une distribution semblable aux principaux indicateurs manufacturiers. Toutefois, lorsque l'on regarde la répartition de ces revenus per capita (15 ans et plus²¹), le portrait change légèrement (tableau 7).

En effet, on peut remarquer que les régions de la Nouvelle-Angleterre et de New York/New Jersey sont celles qui possèdent le niveau de revenu per capita le plus élevé, alors que le Québec arrive avant-dernier, tout juste devant l'Atlantique. Le fait que les régions les plus riches se retrouvent près du Québec peut possiblement avoir un effet positif sur les exportations du Québec vers les États-Unis.

Tableau 7 Revenu personnel et revenu personnel disponible per capita (15 ans et plus), 1992-2001 (en \$ CA)

Régions	Revenu personnel	Revenu personnel disponible
Nouvelle-Angleterre	57 196	47 782
New York/New Jersey	55 952	47 344
Ouest américain	47 587	41 239
Midwest	46 503	40 284
Pennsylvanie	45 860	39 926
Sud-Est	45 413	39 565
New Hampshire/Maine	45 044	39 533
Vermont	42 347	37 203
Sud	41 720	37 152
Ontario	33 638	25 781
Ouest canadien	31 658	24 438
Québec	28 443	21 543
Atlantique	25 831	20 570

Sources : Statistique Canada, Bureau of Labor Statistics
Calculs : JDMD Groupe conseils inc.

²¹ Dans le cas des États-Unis, on considère les 16 ans et plus.

4. RÉSULTATS D'ESTIMATIONS

Dans la présente section, nous développons deux types de modèles visant à décrire le comportement des importations et des exportations du Québec et qui nous permettront de construire un indice propre aux flux de chacune de ces composantes. Ces indices serviront ensuite à obtenir la demande de transport québécoise (en termes monétaires) à l'aide des paramètres du modèle structurel.

Avant de parler des modèles proprement dits, il est nécessaire de faire quelques précisions méthodologiques. En premier lieu, lors de l'estimation d'un modèle MIMIC tel que retenu dans notre analyse, il faut porter une attention particulière au problème de suridentification présent dans la forme réduite du modèle. Une façon de contourner ce problème est simplement de laisser tomber l'hypothèse selon laquelle les termes d'erreurs du modèle de mesure (*indicateurs*) sont mutuellement indépendants, comme le supposent Hauser et Goldberger (1971)²². De cette façon, nous éliminons le problème de suridentification lié à la forme de la matrice variance-covariance donnée en (8). Il ne reste donc plus que la matrice des coefficients de la forme réduite, donnée par (5), qui est encore suridentifiée. Afin de solutionner ce problème, nous imposons les contraintes données en (10) sur le système à estimer.

$$\frac{\Pi_{m1}}{\Pi_{11}} = \frac{\Pi_{m2}}{\Pi_{12}} = \dots = \frac{\Pi_{mk}}{\Pi_{1k}} = C_m, \forall m = 2, \dots, M \quad (10)$$

Dans les contraintes portant sur les coefficients, C_m représente un scalaire déterminé par itérations. Ces contraintes permettent d'exprimer les $K \times M$ paramètres à estimer en termes des $K + M$ paramètres du modèle structurel et du modèle de mesure à condition de fixer la valeur d'un paramètre puisque ceux-ci ne sont pas identifiables séparément. Dans le cas présent, nous avons décidé de normaliser à l'unité le paramètre associé à la variable dépendante de la valeur ajoutée manufacturière (schémas 3 et 4).

Pour l'ensemble des modèles estimés, nous avons eu recours à une forme logarithmique linéaire. Nous avons privilégié cette forme pour trois raisons majeures. Premièrement, l'utilisation de la forme logarithmique linéaire permet de capturer une relation possiblement non linéaire²³ (fonction de type exponentielle). Deuxièmement, le recours à la forme logarithmique nous procure des séries de données plus lisses sur l'horizon longitudinal de l'échantillon. Finalement, cette transformation nous donne une approximation directe du taux de croissance annuel de l'indice à calculer lorsque nous effectuons une différence première.

²² Pour les détails de la transformation, consultez l'annexe 1.

²³ Cependant, la relation à estimer demeure linéaire dans les paramètres.

Nous avons également eu recours à la transformation *within*, bien connue en analyse panel. Cette technique vise simplement à retirer d'une observation donnée, sa moyenne historique (ou longitudinale). La transformation *within* nous permet d'éliminer les paramètres liés aux effets fixes lors des estimations. De plus, elle nous procure des données qui ont une corrélation beaucoup moins forte entre elles (colinéarité), ce qui est souhaitable lorsqu'on veut obtenir des estimateurs les plus précis possible.

Nous avons aussi modélisé l'hétérogénéité temporelle en introduisant une variable dichotomique par année à l'exception de la première année²⁴. Pour cette observation, nous avons considéré la constante, ce qui est équivalent²⁵. L'avantage de cette procédure est qu'elle nous permet de nous fier aux statistiques R^2 puisque celles-ci ne donnent lieu à aucune interprétation lorsque le modèle ne possède pas de constante. Il faut aussi noter que les coefficients d'explication de régressions sont ordinairement plus faibles lorsqu'on applique les transformations *within* que lorsqu'on considère la relation avec les valeurs métriques originales.

Puisque le modèle MIMIC permet d'obtenir un indice de comportement de la variable latente, il est nécessaire de fixer le niveau de celle-ci à l'aide d'un point de référence. Celui-ci peut être endogène ou exogène dépendamment des informations disponibles sur la variable jugée inobservable. Dans le cas présent, nous avons recouru à un point de référence exogène donné par les informations de 1997.

Les résultats d'estimations pour les modèles des importations et des exportations sont présentés dans les sections qui suivent. Pour chacun des modèles, nous commençons par présenter les variables retenues comme causes et indicateurs tout en présentant les raisons nous portant à considérer ces variables plutôt que d'autres. Par la suite, nous présentons et discutons les résultats d'estimations obtenus.

²⁴ Cette procédure équivaut à introduire une variable de tendance propre à chaque année.

²⁵ Seule l'interprétation des coefficients est différente.

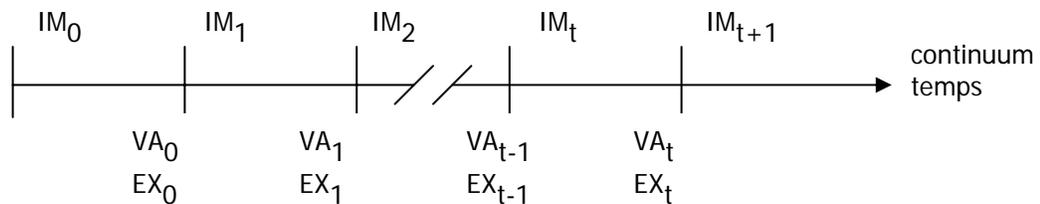
4.1 Le modèle des importations

En ce qui concerne les variables *causales* pour le modèle des importations, nous considérons que les firmes d'une entité géographique donnée choisissent, à chaque période, le nombre d'employés de chaque type d'emploi qui leur permet de minimiser leurs coûts d'exploitation (en considérant que le capital physique est fixe). La présence de certaines rigidités²⁶ rend les ajustements sur le choix des emplois moins élastiques. Le choix final des firmes représente l'ajustement optimal étant donné les conditions du marché. De façon simultanée, les firmes importent le niveau d'intrants leur permettant de maximiser leurs profits.

Du côté des consommateurs, nous considérons que le revenu disponible de l'entité géographique influence le niveau des importations demandées à chaque année. Une entité plus riche possède un pouvoir d'achat plus grand pouvant se refléter par une diversité de la demande de biens et services et, par conséquent, sur la demande d'importation de la région. Ainsi, de ce point de vue, la richesse cause une pression sur la demande intérieure ainsi que sur la demande extérieure.

En ce qui concerne les variables *indicatrices*, nous avons retenu les expéditions manufacturières et la valeur ajoutée manufacturière de la période précédente ainsi que le coût des intrants de la période présente. Bien que l'idée du coût des intrants soit assez claire pour indiquer la demande de transport en importation (puisque une bonne partie des intrants sont importés), elle semble cependant un peu moins évidente en ce qui concerne les expéditions manufacturières et la valeur ajoutée de la période précédente. Notre argumentaire repose sur la persistance du processus de production (schéma 2) :

Schéma 2 Comportement et persistance du processus de production agrégé



²⁶ En raison de la présence de conventions de travail notamment.

Nous considérons que la firme acquiert ses intrants en début de période sur le marché domestique ou étranger (importation). Puis, au cours de la période, cette même firme produit une certaine quantité de biens qui résultera en expéditions manufacturières et en valeur ajoutée. Ensuite, au début de la période suivante, la firme importera à nouveau une certaine quantité de fournitures dont l'intensité sera modulée, du moins en partie, sur la valeur de la production de la période précédente et ainsi de suite pour les périodes subséquentes.

En suivant ce raisonnement, nous pouvons faire l'estimation du modèle sous sa forme réduite en plus de l'estimation du modèle en imposant la contrainte de suridentification (tableau 8). L'analyse nous montre que les paramètres estimés sans imposer les contraintes de suridentification et ceux estimés lorsqu'on les impose sont relativement semblables²⁷. Ils sont légèrement différents puisque le modèle non contraint ne satisfait pas les restrictions nécessaires. Cependant, leur similitude nous indique que le modèle MIMIC est approprié pour la modélisation du phénomène.

Un test statistique basé sur un ratio de vraisemblance permet de vérifier si le modèle de causalité est correctement spécifié. Une statistique LR nous permet de comparer les éléments de la diagonale des matrices variance-covariance (modèles contraint et non contraint). L'idée derrière ce test repose sur la comparaison des composantes des deux matrices. Si celles-ci sont relativement semblables, nous ne pouvons rejeter l'hypothèse que celles-ci sont identiques. Dans ce cas, le pouvoir prédictif du modèle contraint se détériore peu lorsque les restrictions de suridentification sont imposées²⁸.

Dans le cas du modèle des importations, le calcul de la statistique LR ne nous permet pas de rejeter l'hypothèse selon laquelle le modèle de causalité est correctement spécifié, et ce, à un seuil supérieur au seuil critique de 5 %.

Les coefficients d'explication des régressions sont très élevés (entre 89 % et 96 %), ce qui nous indique qu'une grande proportion des variations est captée par la forme fonctionnelle adoptée. Les statistiques χ^2 élevées nous indiquent également que le modèle est nettement significatif dans son ensemble. Qui plus est, l'ensemble des paramètres estimés est significatif à des seuils en deçà des zones critiques normalement utilisées. Cette remarque est encore plus vraie lorsqu'on considère le modèle contraint.

²⁷ À l'exception du coefficient associé à la variable de la masse salariale des employés non productifs qui est non significatif dans le modèle non contraint.

²⁸ Ainsi que le suggèrent Hauser et Goldberger (1971), p. 108.

Des tests de routine portant sur le comportement du terme d'erreur global ont également été menés afin de s'assurer que ceux-ci sont homoscedastiques et non corrélés sériellement. En ce qui concerne la détection d'hétéroscedasticité, nous avons effectué les tests de White (W), de Breusch-Pagan (BP) ainsi que le test de Goldfeld-Quandt (GQ). Les deux premiers sont distribués selon une loi de χ^2 alors que le test GQ suit une loi de Fisher.

Le test BP ne nous permet pas de rejeter l'hypothèse d'homoscedasticité des termes d'erreurs à un seuil de 5 % à l'exception des termes portant sur l'équation avec la variable des expéditions manufacturières de l'année précédente. Cependant, les tests W et GQ nous indiquent que nous ne pouvons rejeter l'hypothèse d'homoscedasticité pour l'ensemble des modèles, ce qui montre que les variables dichotomiques chargées de modéliser le comportement hétéroscedastique sont appropriées. Le test GQ est plus puissant que les tests W et BP, ce qui laisse présager que les termes d'erreurs ont des comportements homoscedastiques.

En ce qui concerne la présence d'autocorrélation, nous avons opté pour le test alternatif de Durbin-Watson (DWa) afin de vérifier si les termes aléatoires sont corrélés entre eux. Ce test est plus général que le test classique de Durbin-Watson puisqu'il est valide lorsque les variables explicatives ne sont pas strictement exogènes. Le test DWa nous permet de ne pas rejeter, à un seuil de 5 %, l'hypothèse nulle selon laquelle les termes d'erreurs ne sont pas corrélés entre eux dans le temps.

4.2 Le modèle des exportations

Le modèle des exportations est essentiellement basé sur le même raisonnement que celui des importations en ce qui concerne les variables *causales*. Ainsi, nous considérons que les emplois manufacturiers productifs et non productifs agissent à titre de variables indépendantes. De plus, nous supposons que la population active de la région peut être une variable affectant le niveau des exportations de la région puisqu'elle représente un bassin de travailleurs potentiels. Ces travailleurs potentiels peuvent être impliqués dans la production d'un secteur autre que le manufacturier. Il est simple de s'imaginer qu'une bonne partie de la production est par la suite expédiée vers d'autres entités géographiques étant donné une demande intérieure limitée et une ouverture des marchés de plus en plus grande.

En ce qui concerne les variables *indicatrices*, nous avons retenu les variables de valeur ajoutée manufacturière et d'expéditions manufacturières. Il nous semble tout à fait naturel de penser que les expéditions manufacturières et la valeur ajoutée du secteur manufacturier puissent nous donner une indication des exportations totales. Cependant, contrairement au modèle des importations, nous avons considéré les variables pour une même période plutôt que d'introduire des retards (tableau 9).

Tout comme dans le modèle des importations, nous notons que les paramètres du modèle contraint et ceux du modèle non contraint sont semblables. Cette remarque nous indique que le modèle MIMIC est approprié dans ce cas. De plus, la diagonale de la matrice de variance-covariance contrainte est semblable à celle non contrainte, ce qui nous indique que le modèle de causalité peut être approprié. À un seuil de 5 %, la statistique LR nous indique que nous ne pouvons pas rejeter l'hypothèse selon laquelle le modèle de causalité est correctement spécifié.

Les coefficients d'explication des régressions ainsi que les statistiques χ^2 sont très élevés, ce qui indique que le modèle explique la majeure partie des variations et qu'il est significatif dans son ensemble. La totalité des paramètres estimés sont significatifs, autant sous la forme contrainte que sous la forme non contrainte.

Tout comme pour le modèle des importations, des tests de routine ont été menés afin de détecter la présence d'hétéroscédasticité et d'autocorrélation des termes aléatoires. Les tests W, BP et GQ ne nous permettent pas de rejeter l'hypothèse d'homoscédasticité. En ce qui concerne la présence d'autocorrélation temporelle, le test DWa ne nous permet pas de rejeter l'hypothèse nulle voulant que les termes d'erreurs soient non corrélés. Ainsi, les tests statistiques nous révèlent que les termes d'erreurs des relations estimées respectent les hypothèses d'homoscédasticité et d'absence de corrélation sérielle, ce qui rend valides les écarts-types des coefficients estimés.

Tableau 8 Résultats d'estimations du modèle MIMIC appliqué à la demande de transport (importation)

Variable dépendante : Coûts des intrants	Modèles		Variable dépendante : Expéditions manufacturières	Modèles		Variable dépendante : Valeur ajoutée	Modèles	
	Non contraint	Contraint		Non contraint	Contraint		Non contraint	Contraint
Emplois manufacturiers reliés à la production	0,585 <i>4,72</i>	0,479 <i>7,06</i>	Emplois manufacturiers reliés à la production	0,512 <i>5,22</i>	0,528 <i>7,06</i>	Emplois manufacturiers reliés à la production	0,561 <i>4,43</i>	0,649 <i>7,06</i>
Emplois manufacturiers non reliés à la production	-0,022 <i>-0,37</i>	0,072 <i>2,28</i>	Emplois manufacturiers non reliés à la production	0,105 <i>2,30</i>	0,079 <i>2,28</i>	Emplois manufacturiers non reliés à la production	0,172 <i>2,93</i>	0,097 <i>2,28</i>
Revenu personnel disponible	0,325 <i>2,30</i>	0,447 <i>5,77</i>	Revenus personnels disponibles	0,583 <i>5,21</i>	0,493 <i>5,77</i>	Revenus personnels disponibles	0,690 <i>4,78</i>	0,605 <i>5,77</i>
année 1994	0,103 <i>4,65</i>	0,096 <i>4,73</i>	année 1994	0,036 <i>2,04</i>	0,042 <i>2,55</i>	année 1994	0,020 <i>0,87</i>	0,024 <i>1,15</i>
année 1995	0,187 <i>6,72</i>	0,177 <i>7,92</i>	année 1995	0,129 <i>5,85</i>	0,139 <i>7,18</i>	année 1995	0,114 <i>4,00</i>	0,121 <i>4,90</i>
année 1996	0,193 <i>5,96</i>	0,180 <i>7,47</i>	année 1996	0,202 <i>7,87</i>	0,215 <i>9,85</i>	année 1996	0,174 <i>5,26</i>	0,183 <i>6,64</i>
année 1997	0,185 <i>4,50</i>	0,185 <i>6,63</i>	année 1997	0,219 <i>6,71</i>	0,231 <i>8,67</i>	année 1997	0,187 <i>4,44</i>	0,184 <i>5,53</i>
année 1998	0,208 <i>3,72</i>	0,197 <i>5,66</i>	année 1998	0,171 <i>3,86</i>	0,193 <i>5,48</i>	année 1998	0,149 <i>2,60</i>	0,154 <i>3,52</i>
année 1999	0,224 <i>3,68</i>	0,206 <i>5,53</i>	année 1999	0,230 <i>4,77</i>	0,256 <i>6,73</i>	année 1999	0,216 <i>3,48</i>	0,227 <i>4,80</i>
année 2000	0,268 <i>3,69</i>	0,245 <i>5,70</i>	année 2000	0,218 <i>3,78</i>	0,249 <i>5,55</i>	année 2000	0,213 <i>2,86</i>	0,225 <i>4,06</i>
année 2001	0,247 <i>3,13</i>	0,212 <i>4,59</i>	année 2001	0,260 <i>4,16</i>	0,298 <i>6,14</i>	année 2001	0,232 <i>2,88</i>	0,254 <i>4,23</i>
Constante	-0,163 <i>-4,19</i>	-0,149 <i>-6,15</i>	Constante	-0,207 <i>-6,72</i>	-0,225 <i>-9,16</i>	Constante	-0,192 <i>-4,82</i>	-0,200 <i>-6,58</i>
Observations	117	117		117	117		117	117
R2	89,75 %	89,32 %		95,50 %	95,44 %		93,36 %	93,20 %
Statistique chi2	1024,29	1039,06		2483,64	2507,67		1644,89	1666,59
Test de la validité du modèle de causalité								
Statistique de test	8,15							
Valeur critique (5 %)	9,49							
* Les variables sont en logarithmes. Les statistiques t sont inscrites en italique sous le coefficient estimé.								

Tableau 9 Résultats d'estimations du modèle MIMIC appliqué à la demande de transport (exportation)

Variable dépendante :	Modèles		Variable dépendante :	Modèles	
	Non contraint	Contraint		Non contraint	Contraint
Expéditions manufacturières			Valeur ajoutée		
Population active	0,512 <i>2,07</i>	0,619 <i>2,94</i>	Population active	0,936 <i>3,05</i>	0,852 <i>2,94</i>
Emplois manufacturiers reliés à la production	0,267 <i>4,04</i>	0,183 <i>3,24</i>	Emplois manufacturiers reliés à la production	0,184 <i>2,24</i>	0,251 <i>3,24</i>
Emplois manufacturiers non reliés à la production	0,115 <i>2,67</i>	0,153 <i>4,17</i>	Emplois manufacturiers non reliés à la production	0,241 <i>4,49</i>	0,211 <i>4,17</i>
année 1993	0,085 <i>5,86</i>	0,084 <i>5,80</i>	année 1993	0,076 <i>4,22</i>	0,076 <i>4,23</i>
année 1994	0,212 <i>14,26</i>	0,212 <i>14,29</i>	année 1994	0,208 <i>11,29</i>	0,208 <i>11,30</i>
année 1995	0,299 <i>19,29</i>	0,300 <i>19,64</i>	année 1995	0,286 <i>14,80</i>	0,285 <i>14,85</i>
année 1996	0,312 <i>18,53</i>	0,313 <i>19,21</i>	année 1996	0,281 <i>13,42</i>	0,281 <i>13,58</i>
année 1997	0,356 <i>15,76</i>	0,362 <i>17,48</i>	année 1997	0,366 <i>13,05</i>	0,361 <i>13,30</i>
année 1998	0,424 <i>17,31</i>	0,431 <i>19,38</i>	année 1998	0,444 <i>14,60</i>	0,439 <i>14,96</i>
année 1999	0,460 <i>17,45</i>	0,465 <i>19,58</i>	année 1999	0,491 <i>14,98</i>	0,487 <i>15,47</i>
année 2000	0,515 <i>17,90</i>	0,522 <i>20,30</i>	année 2000	0,529 <i>14,77</i>	0,524 <i>15,27</i>
année 2001	0,503 <i>16,54</i>	0,505 <i>18,70</i>	année 2001	0,502 <i>13,28</i>	0,500 <i>13,83</i>
Constante	-0,330 <i>-20,11</i>	-0,333 <i>-22,18</i>	Constante	-0,333 <i>-16,29</i>	-0,331 <i>-16,76</i>
Observations	130	130		130	130
R2	96,01 %	95,95 %		94,48 %	94,46 %
Statistique chi2	3124,09	3107,09		2227,14	2230,95
<u>Test de la validité du modèle de causalité</u>					
Statistique de test	6,12				
Valeur critique (5 %)	7,81				
* Les variables sont en logarithmes.					
Les statistiques t sont inscrites en italique sous le coefficient estimé.					

5. ANALYSE COMPARATIVE DES RÉSULTATS

En partant des résultats d'estimations (tableaux 5 et 6), nous pouvons reconstituer le modèle structurel en standardisant un des coefficients du modèle de mesure. Cette procédure nous permet de retracer les valeurs des paramètres du modèle structurel ainsi que les paramètres restants du modèle de mesure. Elle est un prélude à l'analyse comportementale du modèle MIMIC et nous permettra de comparer les résultats dégagés avec une source de données officielles, soit les données sur le commerce international du Québec avec les États-Unis et le Canada provenant de l'ISQ.

Un des principaux inconvénients de la méthodologie développée est qu'elle ne nous apprend rien quant à la répartition géographique des importations et des exportations d'une région donnée. Ainsi, les résultats obtenus visent les agrégats d'échanges. Nous avons désagrégé les résultats selon les pays à partir d'une répartition exogène²⁹. Nous avons arrêté la désagrégation des flux d'échanges économiques au niveau national (Canada et États-Unis).

La présente section se divise en deux sous-sections dans lesquelles nous étudions, de façon séparée, les estimations des flux d'échanges économiques du Québec avec le Canada et les États-Unis. La première section est consacrée à l'étude comportementale du modèle MIMIC appliqué au cas des importations québécoises alors que la seconde partie sert à expliciter les résultats d'estimations propres à la demande d'exportations provenant des marchés extérieurs.

5.1 Le cas des importations

La reconstruction de l'indice de croissance des importations repose sur les variables *causales* et sur la tendance³⁰ associée au modèle de la valeur ajoutée manufacturière. La reconstruction de l'indice se fait à partir du schéma comportemental (schéma 3). Chacun des nombres liant les variables *causales* à la variable latente indique l'impact direct qu'a la variable sur la demande de transport³¹. Les nombres liant les variables causales indiquent, quant à eux, l'ampleur des effets indirects.

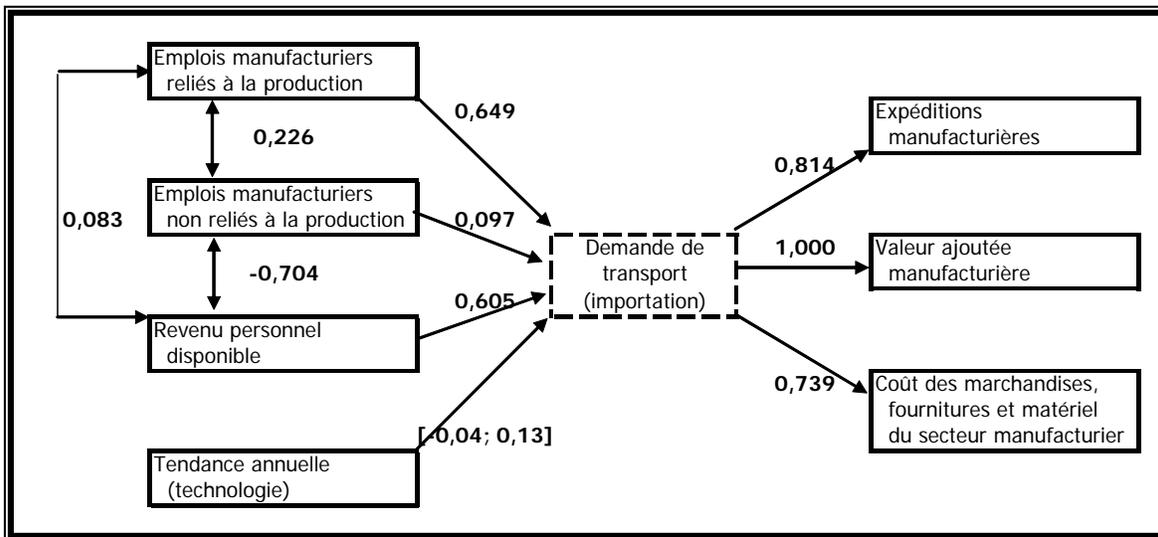
²⁹ Évidemment, ce scénario pourrait être exogénéisé dans le cas d'applications futures.

³⁰ La tendance est représentée par l'ensemble des variables binaires indiquant une année particulière. Il s'agit des variables contrôlant l'hétérogénéité temporelle.

³¹ Dans le cas de la forme retenue (logarithmique linéaire), il s'agit d'une approximation directe de l'élasticité puisque le rapport d'un logarithme est une approximation du taux de croissance.

À titre d'exemple, si les emplois manufacturiers reliés à la production augmentent d'un pourcent, alors, la croissance de la demande de transport (importation) augmente de 0,649 % (effet direct). La croissance des emplois manufacturiers reliés à la production influence également la croissance des emplois non reliés à la production et le revenu personnel disponible qui, à leur tour, ont un impact (indirect) sur la demande de transport. La croissance totale de la demande de transport pour une croissance d'un pourcent des emplois reliés à la production est donc de 0,721 %³².

Schéma 3 Facteurs influençant le comportement de la variable latente (importation)

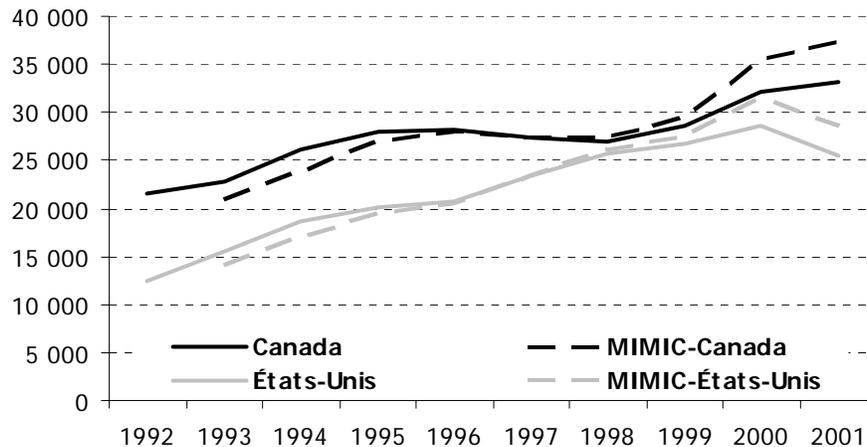


La variable de tendance annuelle repose sur l'estimation des coefficients d'hétérogénéité temporelle. Les variables d'hétérogénéité temporelle captent une partie des changements qui ne peuvent être interprétés par l'une ou l'autre des variables explicatives du modèle. Elles permettent, en outre, de capturer une partie des changements méthodologiques survenus dans l'enquête ASM de 1997. Puisque nous avons laissé les coefficients libres de s'ajuster selon l'équation retenue, nous devons faire un choix quant au scénario à retenir pour l'analyse. Dans le cas des importations, nous optons pour le scénario de croissance médian lors de la reconstitution de la demande de transport pour le Québec (graphique 2).

³² $0,649 + (0,226 \times 0,097) + (0,083 \times 0,605) = 0,721$.

Les estimations obtenues, tant pour le Canada que pour les États-Unis, oscillent autour de la valeur observée. Les flux sont sous-estimés pour la période antérieure à 1997 alors qu'ils sont surestimés pour la période postérieure à 1997. La croissance des échanges estimée pour les États-Unis est relativement bonne puisque les deux séries évoluent de façon parallèle. Qui plus est, nous ne pouvons pas rejeter l'hypothèse selon laquelle le taux de croissance des deux séries est identique, et ce, jusqu'à un seuil de plus de 8 %. Cependant, nous ne pouvons pas tirer la même conclusion en ce qui a trait à la demande d'importation du Québec en provenance du reste du Canada. Statistiquement, nous notons une différence significative entre le taux de croissance de la demande estimée et celui de la demande observée.

Graphique 2 Évolution de la demande d'importation du Québec pour tous modes, 1992-2001 (en M\$ CA)



Source : Institut de la statistique du Québec
Estimations : JDMD Groupe conseils inc.

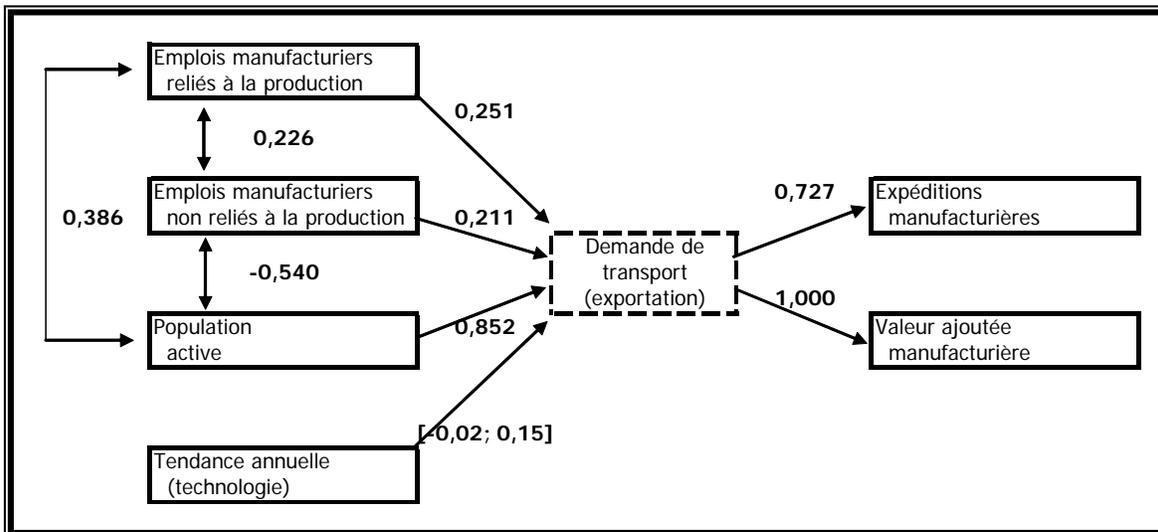
Ainsi, sur l'horizon temporel disponible, le modèle développé a tendance à déterminer correctement l'évolution, en termes monétaires, de la demande d'importation du Québec provenant des États-Unis, mais ce n'est pas le cas pour le Canada.

Le modèle MIMIC nous permet d'estimer l'évolution de la demande d'importation du Québec avec des écarts d'environ plus ou moins 10 % sur l'ensemble de la période recensée, à l'exception des deux dernières années. Bien que l'écart de prévision soit un peu plus élevé que le seuil standard dans ce type d'analyse (soit plus ou moins 5 %), il ne faut pas perdre de vue que ces estimations sont faites sans même recourir aux données comparatives, le but de cette analyse étant de vérifier le comportement du modèle lorsqu'il peut être confronté avec une source de données externe servant de balise. Ainsi, le modèle développé explique environ 90 % des variations des demandes d'importation en ayant recours aux données manufacturières et de revenus disponibles seulement. Bien que ce niveau de précision soit relativement bon, il n'en reste pas moins que la tendance de l'évolution des flux est statistiquement différente en ce qui concerne les importations provenant du reste du Canada.

5.2 Le cas des exportations

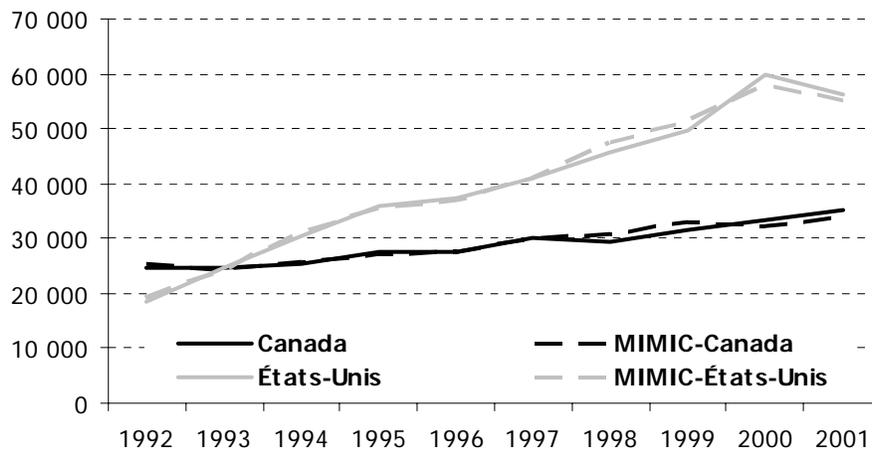
Dans le cas des exportations du Québec, les flux monétaires estimés reposent sur une moyenne des coefficients de tendance. La reconstruction de l'indice peut être résumée par le schéma comportemental (schéma 4). Tout comme dans le cas des importations, les variables *causales* et les coefficients du modèle structurel (apparaissant sur les flèches reliant les variables *causales* à la variable latente) permettent de tracer l'évolution des flux d'échanges estimés et de les comparer avec les flux d'échanges observés.

Schéma 4 Facteurs influençant le comportement de la variable latente (exportation)



Le modèle MIMIC appliqué aux exportations du Québec nous permet d'obtenir une demande estimée qui est sensiblement la même que la demande observée, autant pour le Canada que pour les États-Unis (graphique 3). Les prévisions du modèle oscillent autour des statistiques enregistrées. Contrairement au modèle des importations, nous arrivons à la conclusion que les taux de croissance moyens des flux réels et des flux estimés sont semblables. En termes techniques, les taux de croissance estimés sont statistiquement identiques aux taux de croissance enregistrés, et ce, à un seuil de signification de plus de 40 %.

Graphique 3 Évolution de la demande d'exportation du Québec pour tous modes, 1992-2001 (en M\$ CA)



Source : Institut de la statistique du Québec
Estimations : JDMD Groupe conseils inc.

Globalement, les écarts de prévision entre les échanges calculés et les échanges réels (en termes monétaires) sont très faibles. Ces écarts laissent entrevoir la capacité du modèle à prévoir les échanges économiques du Québec à l'aide des indicateurs sélectionnés, et ce, dans un intervalle inférieur à plus ou moins 5 %. Le modèle semble donc très bien performer pour l'ensemble des années considérées.

6. CONCLUSION

Le modèle MIMIC appliqué au cas de la demande de transport en termes monétaires nous a permis de soulever certains constats intéressants. Nous avons réussi à construire des indices d'évolution permettant de prévoir l'évolution d'une variable qui est considérée inobservable en se basant sur des indices manufacturiers, démographiques et de revenus. En aucun cas dans les estimations nous ne faisons référence aux séries chronologiques elles-mêmes.

Le modèle des exportations nous montre que lorsque celui-ci est spécifié correctement, le recours aux modèles MIMIC nous permet d'obtenir une estimation des flux d'échanges économiques évoluant dans une marge d'erreur inférieure à 5 %. De plus, la robustesse de la forme fonctionnelle a été testée et les résultats préliminaires nous indiquent que la forme retenue est peu sensible aux modifications qui lui ont été imposées. Cette situation nous laisse donc penser que la forme retenue est robuste et que nous devrions accorder une bonne confiance aux résultats obtenus.

La situation n'est toutefois pas aussi bonne dans le cas des importations. En effet, bien que le modèle estime correctement l'amplitude des flux d'échanges provenant des États-Unis, il est important de noter que les flux économiques provenant du reste du Canada montrent certaines lacunes puisque les estimations ne respectent pas la tendance observée sur l'horizon étudié. De plus, il faut également mentionner que l'application de quelques tests ne nous a pas permis de confirmer la robustesse du modèle comme dans le cas des exportations. Cette situation permet de croire que le modèle retenu est peut-être bon localement, mais qu'il n'est pas approprié globalement. Si tel est le cas, il faudrait peut-être repenser la relation de causalité postulée dans le modèle.

Un des indices nous orientant vers cette hypothèse réside dans l'amplitude des scénarios de comportement qui sont, pour certaines périodes, significativement différents du scénario moyen. Ainsi, bien que le test statistique nous indique que le modèle de causalité est correctement spécifié, nous croyons qu'un ajout d'information en ce qui concerne la spécification des variables causales ou indicatrices ou une forme fonctionnelle alternative nous permettrait de capter certains effets présentement non contrôlés. Cette hypothèse devrait être étudiée de façon plus approfondie avant de valider l'approche retenue dans le cas des importations.

Finalement, d'autres avenues pourraient être explorées au cours de la modélisation en considérant un découpage alternatif des régions géographiques ou en ajoutant des données chronologiques. Dans l'optique où cette expérience s'avérerait profitable, il serait intéressant par la suite de se pencher sur la problématique de répartition des flux totaux selon la provenance et la destination.

BIBLIOGRAPHIE

- BALTAGI, B.H. (1999), *Econometrics*, New York, Springer, 2^e édition, 396 pages.
- CARLTON, D.W. et J.M. PERLOFF (2000), *Modern Industrial Organization*, 3^e édition, édition Addison-Wesley, New York, 780 pages.
- CHAREST, G., P. LUSZTIG et B. SCHWAB (2000), *Gestion financière*, 2^e édition, Renouveau pédagogique, Montréal, 1107 pages.
- DUBÉ, J. et M. DUPÉRÉ (2004), « Expérimentation de la méthode d'entropie-croisée pour l'estimation d'échanges économiques à partir de flux de transport », *Études et recherches en transport*, ministère des Transports du Québec, 38 pages.
- GILES, D.E.A. (1999a), « Measuring the Hidden Economy: Implications for Econometric Modelling », *Economic Journal*, no. 109, p. F370-F380.
- GILES, D.E.A. (1999b), « Modelling the Hidden Economy and the Tax-Gap in New-Zealand », *Empirical Economics*, no. 24, p. 621-640.
- GILES, D.E.A. et L.M. TEDDS (2002), « Taxes and the Canadian Underground Economy », Toronto, *Canadian Tax Paper*, no. 106, 270 pages.
- GILES, D.E.A., L.M. TEDDS et G. WERKNEH (2002), « The Canadian Underground and Measured Economies: Granger Causality Results », *Applied Economics*, no. 34, p. 2347-2352.
- GOLDBERGER A.S. (1972), « Maximum Likelihood Estimation of Regressions Containing Unobservable Independent Variables », *International Economic Review*, vol. 13, p. 1-15.
- GOLDBERGER, A.S. (1972), *Structural Equation Models in the Social Sciences*, John Hopkins University, Baltimore, p. 85-112.
- GREENE, W.H. (2003), *Econometric Analysis*, New Jersey, Prentice Hall, 5^e édition, New Jersey, 1026 pages.
- GRIFFITH, W.E., R.C. HILL et G.C. JUDGE (1993), *Learning and Practicing Econometrics*, New York, John Willey & Sons, 866 pages.
- HAUSER, R.M. et A.S. GOLDBERGER (1971), « The Treatment of Unobservable Variables in Path Analysis », chapitre 4 dans *Sociological Methodology*, Jossey-Bass, San Fransisco, p. 81-117.

- JÖRESKOG K. G. et A.S. GOLDBERGER (1975), « Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable », *Journal of the American Statistical Association*, vol. 70, no. 351, p. 631-639.
- TEDDS, L.M. (2004), *Revisiting the Latent Variable/MIMIC Approach to Measuring the Size of the Canadian Underground Economy*, Hamilton, Département d'économie de l'Université McMaster, 27 pages.
- TEDDS, L.M. (1998), *Measuring the Size of the Hidden Economy in Canada: A Latent Variable/MIMIC Model Approach*, Victoria, Département d'économie de l'Université de Victoria, 16 pages.
- WOOLDRIDGE, J.M. (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, 752 pages.
- WOOLDRIDGE, J.M. (2000), *Introductory Econometrics: A Modern Approach*, Cincinnati, South-Western College Publishing, 824 pages.
- ZELLENER, A. (1970), « Estimation of Regression Relationship Containing Unobservable Independent Variables », *International Economic Review*, vol. 11, no. 3, p. 441-454.

Annexe 1

ANNEXE 1 TRANSFORMATION DE HAUSER ET GOLDBERGER

Cette transformation, suggérée par Hauser et Goldberger (1971), permet aux termes $V_{m,n,t}$ de l'équation (3) d'être corrélés librement entre eux³³. Il s'ensuit logiquement que la forme de la matrice variance-covariance a une structure différente de celle spécifiée à l'équation (8). Dans le cas où les termes d'erreurs du modèle de mesure sont librement corrélés, on peut réécrire le modèle de façon à définir la variable latente comme une fonction **exacte** de ses causes, absorbant le terme d'erreur d'origine, $\varepsilon_{n,t}$, dans les termes $V_{m,n,t}$, que nous renommerons $W_{m,n,t}$. Ainsi, nous pouvons réécrire le système de la façon suivante :

$$y_{n,t}^* = x_{1,n,t}\alpha_1 + \dots + x_{k,n,t}\alpha_k + \dots + x_{K,n,t}\alpha_K \quad (A1)$$

$$y_{m,n,t} = y_{n,t}^*\beta_m + \mu_{m,n} + w_{m,n,t} \quad (A2)$$

où les variables ont la même signification que précédemment et la nouvelle matrice variance-covariance du modèle réduit est donnée par :

$$E(w'w) = \Sigma = \begin{bmatrix} w_{11} & \cdots & w_{1m} & \cdots & w_{1M} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{1m} & \cdots & w_{mm} & \cdots & w_{mM} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{1M} & \cdots & w_{mM} & \cdots & w_{MM} \end{bmatrix} \quad (A3)$$

Les termes w_{mm} ne sont pas contraints à une forme particulière, ce qui fait en sorte que les restrictions de suridentification visent maintenant les coefficients des vecteurs α et β seulement, soit les coefficients de la matrice Π .

³³ C'est-à-dire qu'on ne pose aucune restriction a priori sur la forme de la matrice.

