

# CHOIX OCCUPATIONNELS ET OFFRE DE TRAVAIL LORSQUE LES SALAIRES SONT FONCTION DES HEURES TRAVAILLÉES

Thomas Lemieux

Volume 96, numéro 4, décembre 2020

URI : <https://id.erudit.org/iderudit/1087023ar>

DOI : <https://doi.org/10.7202/1087023ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

HEC Montréal

ISSN

0001-771X (imprimé)

1710-3991 (numérique)

[Découvrir la revue](#)

Citer cet article

Lemieux, T. (2020). CHOIX OCCUPATIONNELS ET OFFRE DE TRAVAIL LORSQUE LES SALAIRES SONT FONCTION DES HEURES TRAVAILLÉES. *L'Actualité économique*, 96(4), 691–720. <https://doi.org/10.7202/1087023ar>

Résumé de l'article

Ce papier présente un modèle économétrique déterminant la décision de travailler dans une ou deux occupations différentes lorsque le salaire est fonction des heures travaillées. Le modèle est construit à partir d'un principe de base de la théorie de la demande avec contraintes de rationnement qui ne requiert pas les restrictions d'exclusions standards. L'approche économétrique est appliquée à la décision de travail dans le secteur formel et informel. Les estimations suggèrent une élasticité d'offre de travail de l'ordre de l'unité. De même elles montrent que le salaire dans le secteur informel diminue lorsque les heures travaillées dans ce secteur augmentent, tandis que le phénomène inverse se produit dans le secteur formel.

## CHOIX OCCUPATIONNELS ET OFFRE DE TRAVAIL LORSQUE LES SALAIRES SONT FONCTION DES HEURES TRAVAILLÉES\*

Thomas LEMIEUX

*Département de sciences économiques*

*Université de Montréal et*

*CIRANO*

RÉSUMÉ – Ce papier présente un modèle économétrique déterminant la décision de travailler dans une ou deux occupations différentes lorsque le salaire est fonction des heures travaillées. Le modèle est construit à partir d'un principe de base de la théorie de la demande avec contraintes de rationnement qui ne requiert pas les restrictions d'exclusions standards. L'approche économétrique est appliquée à la décision de travail dans le secteur formel et informel. Les estimations suggèrent une élasticité d'offre de travail de l'ordre de l'unité. De même elles montrent que le salaire dans le secteur informel diminue lorsque les heures travaillées dans ce secteur augmentent, tandis que le phénomène inverse se produit dans le secteur formel.

ABSTRACT – This paper develops an econometric model of the decision to work in one or two different occupations when wages depend on hours of work. The model is identified using a basic principle of demand theory with rationing constraints which does not require standard exclusion restrictions. The econometric approach is applied to the decision of working in the formal and in the informal sector. The estimates suggest standard labor supply elasticities of the order of one. The estimates also suggest that the informal sector wage declines as hours worked in that sector increase, while the opposite is true in the formal sector.

---

\* Une version préliminaire de ce travail a été présentée au sixième Congrès mondial de la société d'économétrie à Barcelone en Espagne au mois d'août 1990. L'auteur remercie spécialement Bernard Fortin et Pierre Fréchette pour les données qu'ils lui ont fournies, ainsi que Caroline Chapain pour son aide précieuse dans la préparation de la version française de ce texte. Il remercie aussi David Card, Nicole Fortin, Jeffrey Zabel et les participants à des séminaires à l'Université de Princeton et au MIT pour leurs remarques fort utiles. L'auteur se tient pour responsable de toutes erreurs pouvant subsister dans ce travail.

## INTRODUCTION

Dans la théorie standard des choix occupationnels, les emplois sont différenciés sur la base à la fois du taux de salaire et d'autres avantages non pécuniers. Les travailleurs se spécialisent dans un ou deux emplois sur la base d'avantages comparatifs<sup>1</sup>. Lorsque les travailleurs sont libres de choisir leur nombre d'heures travaillées et que leurs salaires ne dépendent pas de ce nombre, le choix occupationnel et l'offre de travail peuvent être analysés de manière séquentielle. Une fois l'occupation choisie, un travailleur offre des heures de travail jusqu'à ce que l'utilité marginale du loisir soit égale au taux de salaire qui prévaut dans cette occupation.

Le choix occupationnel et l'offre de travail ne sont cependant déterminés de manière séquentielle que dans le cas précis où le taux de salaire ne dépend pas des heures travaillées et où le travailleur est libre de choisir ses heures de travail. Le choix de l'occupation et l'offre de travail ne sont plus déterminés séquentiellement lorsque plusieurs occupations offrent des taux de salaire différents et requièrent également un nombre d'heures travaillées différent. Plus généralement, lorsque dans diverses occupations les salaires dépendent différemment des heures travaillées, de très nombreuses combinaisons de choix occupationnels et d'offre de travail deviennent possibles. Les travailleurs peuvent, en particulier, trouver optimal d'offrir de travailler dans plus d'une occupation à la fois, au lieu de se spécialiser dans une seule activité. Cette possibilité a depuis longtemps été reconnue dans la littérature concernant la production domestique et le travail sur le marché<sup>2</sup>. De même, elle a été explorée dans la littérature croissante sur les ménages agricoles<sup>3</sup>.

L'objectif de ce travail est de présenter une étude empirique, basée sur un modèle de maximisation de l'utilité, pour expliquer la décision de travailler dans une ou deux occupations différentes ainsi que le choix des heures travaillées et la détermination des salaires dans chacune de ces occupations. Cette étude empirique est appliquée au choix de travailler dans les secteurs formel et informel et utilise des données provenant de la région de Québec<sup>4</sup>. La structure du modèle consiste en trois équations : une fonction de gain dans le secteur informel, une fonction de gain dans le secteur formel et une fonction d'utilité. Les paramètres de ces trois équations sont estimés conjointement par la méthode du maximum de vraisemblance. La décision de travailler dans un ou deux secteurs a pour résultat

---

1. La littérature moderne débute avec le célèbre modèle d'autosélection et de distribution des revenus de Roy, 1951. Des applications récentes incluent le choix d'un secteur syndiqué ou non syndiqué (Lee, 1978), la décision d'aller au collège (Willis et Rosen, 1979), le choix du secteur industriel (Heckman et Sedlacek, 1990), le choix occupationnel (Robertson et Symons, 1990) et le choix de travailler dans le secteur informel (Magnac, 1991).

2. Voir Becker (1965) et Gronau (1977).

3. Sur les marchés de travail agricole des pays en voie de développement, on observe couramment que les membres d'un même ménage travaillent à la fois sur la ferme familiale et sur le marché. Voir Rosenzweig (1980) et Benjamin (1992).

4. Voir Fortin et Fréchette (1986).

l'existence de quatre régimes de travail possibles : ne pas avoir d'emploi, travailler uniquement dans le secteur formel ou dans le secteur informel et travailler dans les deux secteurs à la fois. Le modèle statistique est ainsi un modèle de régression à plusieurs régimes, observés mais endogènes.

D'une part, l'estimation jointe de l'élasticité de l'offre de travail et de la relation entre les salaires et les heures travaillées dans les deux secteurs est compliquée, ce qui peut expliquer que cela n'a jamais été fait auparavant. D'autre part, considérer de manière conjointe l'offre de travail et la détermination des salaires dans les deux secteurs fournit des restrictions additionnelles sur les paramètres des équations dans les différents régimes, ce qui peut permettre d'identifier les paramètres intéressants. Dans le cas usuel, lorsque le travailleur se spécialise uniquement dans un secteur, l'identification de l'élasticité d'offre de travail et de la relation entre les salaires et les heures travaillées est fortement reliée aux hypothèses de distributions et aux restrictions d'exclusions<sup>5</sup>. Dans le modèle considéré ici, l'identification a pu être établie, en utilisant un principe de base de l'analyse de la demande en présence de rationnement (Samuelson, 1947; Tobin et Houthakker, 1950). L'idée est simplement que la magnitude de l'élasticité de l'offre de travail dépend de la décision du travailleur d'offrir, ou non, des heures de travail à un autre secteur. Cela génère une série de restrictions sur les équations des heures travaillées et du salaire horaire moyen dans les différents régimes. Ces restrictions peuvent nous permettre d'identifier les paramètres intéressants du modèle, même en l'absence de restrictions d'exclusion standards.

Le travail est organisé de la manière suivante : le problème du choix du travailleur ainsi que la spécification des composantes aléatoires du modèle sont présentés à la section 1. L'identification et l'estimation du modèle statistique par la méthode du maximum de vraisemblance sont exposées à la section 2. Deux estimations  $y$  sont discutées : celle d'un modèle à choix discret déterminant le choix du secteur, et celle d'un modèle déterminant le choix du secteur, de même que les heures de travail et les salaires. Ces deux modèles sont estimés à la section 3. Les paramètres structurels estimés sont ensuite utilisés pour obtenir, entre autres, l'élasticité standard d'offre de travail et celle de l'impact des taxes sur la décision de travailler dans le secteur informel. La dernière section conclut par une présentation des autres applications potentielles du modèle.

#### 1. UN MODÈLE DÉTERMINANT LE CHOIX OCCUPATIONNEL, LES HEURES DE TRAVAIL ET LES SALAIRES

Deux idées essentielles du modèle déterminant le choix occupationnel, les heures de travail et les salaires sont proposées ici. Premièrement, il existe une

---

5. Pour une discussion détaillée de l'identification du modèle standard d'autosélection et de détermination des salaires (modèle de Roy), voir Heckman et Honoré (1990). Pour les modèles d'offre de travail statique, voir Mroz (1987). Des problèmes additionnels sur les systèmes structurels hédonistes d'estimation (Rosen, 1974) ont aussi été signalés par Brown et Rosen (1982), Epple (1987) et Kahn et Lang (1988) entre autres.

relation entre les salaires et les heures travaillées dans au moins un des deux secteurs. Deuxièmement, les heures de travail dans les deux secteurs sont parfaitement substituables<sup>6</sup>. Cela signifie qu'un travailleur est indifférent entre travailler dans le secteur formel ou dans le secteur informel, lorsque le taux de salaire est le même dans l'un et l'autre secteur. Si les heures de travail dans les deux secteurs sont parfaitement substituables, les salaires doivent dépendre dans au moins un secteur des heures travaillées afin que le travailleur offre à la fois des heures de travail aux secteurs formel et informel. Avec des salaires fixes dans les deux secteurs, le travailleur offrirait seulement des heures de travail au secteur offrant le meilleur salaire.

Le fait que, dans au moins un secteur, il existe un lien entre les salaires et les heures de travail a été établi par Lemieux, Fortin et Fréchette (1994) qui montrent que les gains dans le secteur informel sont une fonction concave des heures travaillées dans ce secteur. Le salaire  $y$  est alors négativement relié au heures travaillées, peut être parce que la nature informelle de l'activité de ce secteur limite la capacité de son marché. Les heures travaillées dans les deux secteurs sont supposées être parfaitement substituables afin de simplifier le modèle. Bien que cette hypothèse soit arbitraire, elle a aussi été utilisée dans le contexte de la production domestique par Gronau (1977) ainsi que dans le contexte du marché du travail agricole (rural) dans les pays en voie de développement par Rosenzweig (1980) et Benjamin (1992)<sup>7</sup>. En théorie, cette hypothèse pourrait être relâchée tout en permettant aux salaires de dépendre du nombre d'heures de travail. Cependant, en pratique, un tel modèle serait difficile à estimer et aurait peu d'implications intéressantes.

### 1.1 Le problème du choix du travailleur

Considérons la fonction exponentielle suivante représentant les gains résultant du travail dans les secteurs formel et informel :

$$Y_0 = W_0 \{ [1 - \exp(-\theta_0 h_0)] / \theta_0 \} \quad (1)$$

$$Y_1 = W_1 \{ [1 - \exp(-\theta_1 h_1)] / \theta_1 \} \quad (2)$$

où  $Y_k$  et  $h_k$  représentent respectivement les gains (revenu de travail) et les heures travaillées dans le secteur formel ( $k = 0$ ) et dans le secteur informel ( $k = 1$ ). La variable  $W_0(W_1)$  est le taux de salaire (gain marginal) correspondant à zéro heure

6. De la théorie du consommateur, on tire que les courbes d'indifférence entre les heures de travail dans les deux secteurs sont des fonctions linéaires lorsque les heures sont de parfaits substituts. La pente des courbes d'indifférence n'est pas nécessairement égale à un, cependant. En général, la fonction d'utilité prend la forme  $u(c, h_0 + ah_1)$ , bien que l'hypothèse  $a = 1$  soit maintenue pour simplifier l'analyse.

7. L'hypothèse selon laquelle les heures de travail domestique et celles sur le marché sont parfaitement substituables a été critiquée par Pollack et Wachter (1975). Lacroix et Fortin (1992) teste et rejette la parfaite substituabilité des heures de travail dans les secteurs formel et informel. Ils ne tiennent toutefois, pas compte du fait que les salaires peuvent dépendre des heures de travail.

travaillée dans le secteur formel (informel). Ce salaire reflète les capacités productives du travailleur dans chaque secteur. Il diffère typiquement du taux de salaire observé lorsque les salaires dépendent des heures de travail. Les paramètres  $\theta_0$  et  $\theta_1$  déterminent la courbure des deux fonctions de gains. La fonction de gains est une fonction concave des heures travaillées lorsque  $\theta_k > 0$ , linéaire lorsque  $\theta_k = 0$  et convexe lorsque  $\theta_k < 0$ . Les préférences sont représentées par une fonction d'utilité fortement séparable :

$$U(C, \ell) = C + W_R \{ [1 - \exp(\psi(T - \ell))] / \psi \} . \tag{3}$$

$C$  est un bien de consommation,  $\ell$  représente les heures de loisir et  $T$  est le nombre d'heures totales disponibles. La variable  $W_R$  représente le salaire de réserve. Comme  $T - \ell = h_0 + h_1$ , il en découle que  $h_0$  et  $h_1$  sont des substituts parfaits. De plus, le taux marginal de substitution dépend seulement du loisir ce qui signifie qu'il n'y a pas d'effet revenu<sup>8</sup>. Les hypothèses sur les formes fonctionnelles des équations (1), (2), (3) sont nécessaires pour produire un modèle économétrique soluble. Le nombre de biens consommés  $C$  est obtenu à l'aide de la contrainte budgétaire :

$$C = I + Y_0(h_0) + Y_1(h_1), \tag{4}$$

où les fonctions  $Y_0(h_0)$  et  $Y_1(h_1)$  sont données par les équations (1) et (2), et où  $I$  est le revenu hors-travail. Une fois l'équation (4) substituée dans la fonction d'utilité, le problème du consommateur devient le suivant :

$$\begin{aligned} \text{Max} \quad & U(I + Y_0(h_0) + Y_1(h_1), \ell) \text{ sujet à } T = \ell + h_1 + h_0, h_0 \geq 0 \text{ et } h_1 \geq 0 . \\ & \{ \ell, h_0, h_1 \} \end{aligned}$$

Considérons le Lagrangien :

$$\begin{aligned} \mathcal{L} = & I + Y_0(h_0) + Y_1(h_1) + W_R \{ [1 - \exp(\psi(T - \ell))] / \psi \} + \lambda(T - \ell - h_0 - h_1) \\ & - \varphi_0 h_0 - \varphi_1 h_1 . \end{aligned}$$

Les conditions de premier ordre (conditions nécessaires) pour l'existence d'un optimum sont :

$$\partial \mathcal{L} / \partial h_0 = W_0 \exp(-\theta_0 h_0) - \lambda - \varphi_0 = 0 \tag{5a}$$

$$\partial \mathcal{L} / \partial h_1 = W_1 \exp(-\theta_1 h_1) - \lambda - \varphi_1 = 0 \tag{5b}$$

$$\partial \mathcal{L} / \partial \ell = W_R \exp(\psi(T - \ell)) - \lambda = 0 \tag{5c}$$

$$\partial \mathcal{L} / \partial \lambda = T - \ell - h_0 - h_1 = 0 \tag{5d}$$

Les conditions de saturations complémentaires,  $\varphi_0 h_0 = 0$  et  $\varphi_1 h_1 = 0$  donnent quatre solutions (régimes) possibles au problème du consommateur :

---

8. Notons que l'essentiel du modèle reste inchangé lorsque l'effet revenu est différent de zéro, sauf que l'on se trouve en présence du problème habituel qui est de trouver un salaire virtuel approprié, avec des ensembles de budget non linéaires.

Régime 1 (ne pas avoir d'emploi) :  $h_0 = 0, h_1 = 0$ .

Régime 2 (travailler dans le secteur formel) :  $h_0 > 0, h_1 = 0$ .

Dans ce régime, les conditions de saturations complémentaires font en sorte que  $\varphi_0 = 0$ . En prenant les logarithmes des équations (5a) et (5c) et en résolvant avec  $\ln \lambda$  on obtient :

$$h_0 = (\theta_0 + \psi)^{-1} (\ln W_0 - \ln W_R). \quad (6)$$

Régime 3 (travailler dans le secteur informel) :  $h_0 = 0, h_1 > 0$ .

En utilisant la même procédure que le régime 2 cela donne :

$$h_1 = (\theta_1 + \psi)^{-1} (\ln W_1 - \ln W_R). \quad (7)$$

Régime 4 (travailler dans les deux secteurs) :  $h_0 > 0, h_1 > 0$ .

Dans ce régime, les conditions de saturations complémentaires font en sorte que  $\varphi_0 = \varphi_1 = 0$ . En prenant les logarithmes des équations (5a), (5b) et (5c) et en solutionnant pour  $\ln \lambda$  on obtient :

$$h_0 = (\theta_0 \psi + \theta_1 \psi + \theta_0 \theta_1)^{-1} [(\psi + \theta_1)(\ln W_0 - \ln W_R) - \psi(\ln W_1 - \ln W_R)], \quad (8a)$$

$$h_1 = (\theta_0 \psi + \theta_1 \psi + \theta_0 \theta_1)^{-1} [(\psi + \theta_0)(\ln W_1 - \ln W_R) - \psi(\ln W_0 - \ln W_R)]. \quad (8b)$$

Le modèle pourrait facilement être utilisé dans le cas où le temps est aussi alloué à des activités de production domestique. Dans ce cas, les équations déterminant les heures de travail prendraient la même forme que les équations (6), (7) et (8). Seule l'interprétation de  $\psi$  et  $W_R$  changerait<sup>9</sup>. Les conditions de second ordre (conditions suffisantes) pour un optimum supposent les restrictions suivantes sur les paramètres :  $\theta_0 + \theta_1 > 0$ ,  $\theta_0 + \psi > 0$ ,  $\theta_1 + \psi > 0$  et  $\theta_0 \theta_1 + \theta_0 \psi + \theta_1 \psi > 0$ . Ces restrictions sont toujours satisfaites lorsque à la fois les fonctions de gains et la fonction d'utilité sont concaves, soient, respectivement,  $\theta_0 > 0$ ,  $\theta_1 > 0$  et  $\psi > 0$ . De plus, ces restrictions ne sont pas en contradiction avec le fait d'avoir une fonction de gain convexe si celle-ci ne l'est pas trop. Pour que le modèle statistique décrit ci-après soit cohérent, les conditions de second ordre sont aussi suffisantes, mais pas toujours nécessaires.

### 1.2 Statique comparative

Un des avantages supplémentaires d'utiliser une spécification exponentielle pour les fonctions de gains et pour la fonction d'utilité est que cette forme fonctionnelle conduit à des résultats de statique comparative simple. En utilisant les équations (6), (7) et (8), il est facile de montrer les résultats suivants :

9. L'avantage d'une spécification exponentielle est que cela conduit à des spécifications semi-logarithmiques des heures de travail pour n'importe quel secteur (ou activité) où le temps est alloué. Additionner plus d'activités change simplement les coefficients des équations semi-logarithmiques mais ajoute, par contre, des complications en présence de solutions de coin dans ces autres activités.

Régime 2 (secteur formel seulement) :  $\partial h_0/\partial \ln W_0 = -\partial h_0/\partial \ln W_R = 1/(\psi + \theta_0)$

Régime 3 (secteur informel seulement) :  $\partial h_1/\partial \ln W_1 = -\partial h_1/\partial \ln W_R = 1/(\psi + \theta_1)$

Régime 4 (les deux secteurs à la fois) :

$$\partial h_0/\partial \ln W_0 = c(\psi + \theta_1), \partial h_0/\partial \ln W_1 = -c\psi, \partial h_0/\partial \ln W_R = -c\theta_1$$

$$\partial h_1/\partial \ln W_0 = -c\psi, \partial h_1/\partial \ln W_1 = c(\psi + \theta_0), \partial h_1/\partial \ln W_R = -c\theta_0,$$

où  $c = (\theta_0\theta_1 + \theta_0\psi + \theta_1\psi)^{-1}$ . La semi-élasticité des heures de travail par rapport à un taux de salaire correspondant à zéro heure travaillée est toujours plus grande pour les travailleurs du régime 4 (travailler dans les deux secteurs à la fois) que pour les travailleurs offrant des heures de travail à un secteur uniquement, parce que, à la marge, il existe une substitution possible entre  $h_0$  et  $h_1$ . Par exemple,  $\partial h_0/\partial \ln W_0$  est égal à  $c(\psi + \theta_1)$  pour les travailleurs du régime 4, ce qui est plus grand que  $1/(\psi + \theta_0)$ , la semi-élasticité dans le régime 2 (secteur formel uniquement) puisque  $c(\psi + \theta_0) = 1/[(\psi + \theta_0) - \psi^2/(\psi + \theta_1)] > 1/(\psi + \theta_0)$  (résultat qui provient des conditions de second ordre). Ce résultat est l'application de ce que Samuelson appelait le principe de Le Chatelier. Ce principe nous sera très utile lorsque nous discuterons de l'identification du modèle à la section 2.

### 1.3 Spécification aléatoire

Le facteur aléatoire est introduit dans le modèle en faisant l'hypothèse que  $\ln W_0$ ,  $\ln W_1$  et  $\ln W_R$  sont des fonctions linéaires des vecteurs des caractéristiques observées  $x_0$ ,  $x_1$  et  $x_R$  et des termes aléatoires  $\varepsilon_0$ ,  $\varepsilon_1$  et  $\varepsilon_R$  :

$$\ln W_0 = x_0\beta_0 + \varepsilon_0, \ln W_1 = x_0\beta_1 + \varepsilon_1 \text{ et } \ln W_R = x_1\beta_R + \varepsilon_R.$$

Comme il est difficile d'exclure, *a priori*, des variables explicatives telles que l'éducation et l'expérience sur le marché du travail dans les trois vecteurs  $x_0$ ,  $x_1$  et  $x_R$ , ces trois vecteurs sont supposés identiques et égaux à un vecteur  $x$ . Le modèle statistique est alors facile à exprimer en fonction des paramètres suivants :

$$\sigma_0^2 = \text{Var}[(\varepsilon_0 - \varepsilon_R)/(\psi + \theta_0)], \sigma_1^2 = \text{Var}[(\varepsilon_1 - \varepsilon_R)/(\psi + \theta_1)],$$

$$\gamma_0 = (\beta_0 - \beta_R)/[(\psi + \theta_0)\sigma_0], \gamma_1 = (\beta_1 - \beta_R)/[(\psi + \theta_1)\sigma_1],$$

$$\psi_0 = [\psi/(\psi + \theta_0)]/(\sigma_1/\sigma_0), \psi_1 = [\psi/(\psi + \theta_1)]/(\sigma_0/\sigma_1)$$

et en fonction des termes aléatoires

$$u_0 = (\varepsilon_0 - \varepsilon_R)/[(\psi + \theta_0)\sigma_0], u_1 = (\varepsilon_1 - \varepsilon_R)/[(\psi + \theta_1)\sigma_1],$$

où  $u_0$  et  $u_1$  sont des variables normalement distribuées, de variance 1 et dont le coefficient de corrélation est  $\rho$ . Les équations (6), (7) et (8), qui sont les équations de détermination des heures de travail peuvent aussi être réécrites comme suit :



$$h_{0i} = \sigma_0(x_i' \gamma_0 + u_{0i}), \quad (6')$$

$$h_{1i} = \sigma_1(x_i' \gamma_1 + u_{1i}), \quad (7')$$

$$h_{0i} = (1 - \psi_0 \psi_1)^{-1} \sigma_1[x_i'(\gamma_0 - \psi_0 \gamma_1) + u_{0i} - \psi_0 u_{1i}], \quad (8a')$$

$$h_{1i} = (1 - \psi_0 \psi_1)^{-1} \sigma_0[x_i'(\gamma_1 - \psi_1 \gamma_0) + u_{1i} - \psi_1 u_{0i}]. \quad (8b')$$

Pour que le modèle statistique soit cohérent, les conditions suivantes doivent être respectées :

Régime 2 (secteur formel uniquement) :  $u_{0i} > -x_i' \gamma_0$ .

Régime 3 (secteur informel uniquement) :  $u_{1i} > -x_i' \gamma_1$ .

Régime 4 (les deux secteurs à la fois) :  $u_{0i} - \psi_0 u_{1i} > -x_i'(\gamma_0 - \psi_0 \gamma_1)$  et  
 $u_{1i} - \psi_1 u_{0i} > -x_i'(\gamma_1 - \psi_1 \gamma_0)$ .

Il est facile de montrer que ces inégalités conduisent aux conditions de participation :

Régime 1 (ne pas avoir d'emploi) :  $u_{0i} < -x_i' \gamma_0$  et  $u_{1i} < -x_i' \gamma_1$ . (9a)

Régime 2 (secteur formel uniquement) :  $u_{0i} > -x_i' \gamma_0$  et  $u_{1i} - \psi_1 u_{0i} < -x_i'(\gamma_1 - \psi_1 \gamma_0)$ . (9b)

Régime 3 (secteur informel uniquement) :  $u_{1i} > -x_i' \gamma_1$  et  $u_{0i} - \psi_0 u_{1i} < -x_i'(\gamma_0 - \psi_0 \gamma_1)$ . (9c)

Régime 4 (les deux secteurs à la fois) :  $u_{0i} - \psi_0 u_{1i} > -x_i'(\gamma_0 - \psi_0 \gamma_1)$  et  
 $u_{1i} - \psi_1 u_{0i} > -x_i'(\gamma_1 - \psi_1 \gamma_0)$ . (9d)

Ces conditions de participation sont illustrées à la figure 1. Les conditions de second ordre garantissent que  $\psi_0 > 0$ ,  $\psi_1 > 0$  et que  $1/\psi_0 > \psi_1$ ; cette dernière condition fait en sorte que la probabilité d'être dans le régime 4 est positive.

Un inconvénient de ce modèle est qu'il n'existe pas de forme linéaire fermée pour représenter la fonction de gain ou du salaire horaire moyen. Les salaires horaires moyens dans les deux secteurs  $R_{0i}$  et  $R_{1i}$  sont donnés par :

$$R_0 = Y_0/h_0 = W_0\{[1 - \exp(-\theta_0 h_0)]/\theta_0 h_0\}$$

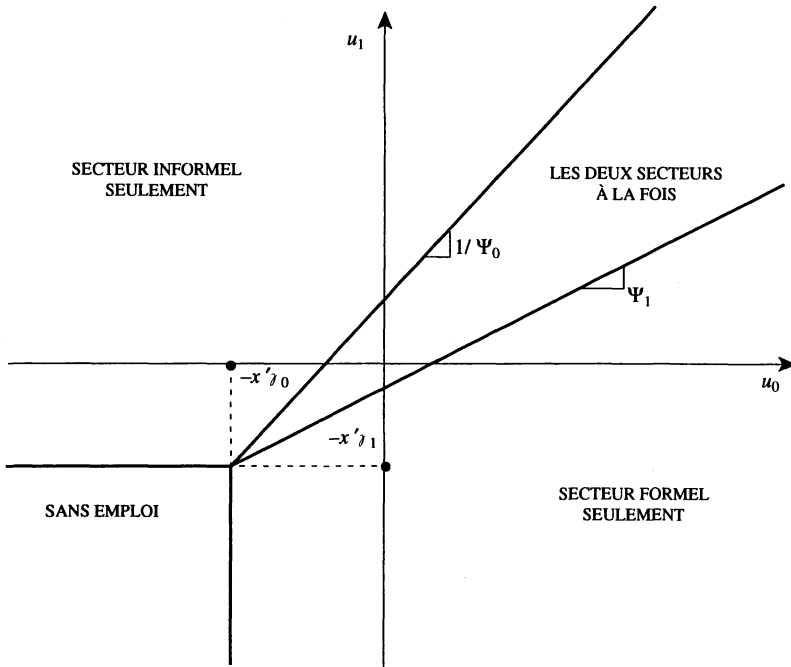
$$R_1 = Y_1/h_1 = W_1\{[1 - \exp(-\theta_1 h_1)]/\theta_1 h_1\}$$

La forme fonctionnelle représentant le salaire horaire moyen est ainsi  $R(x) = W(1 - \exp(-x))/x$ , avec  $x = \theta_0 h_0$  et  $x = \theta_1 h_1$ . Une approximation de premier ordre de  $\ln[R(x)]$  autour de  $x = 0$  donne<sup>10</sup> :

$$\ln(R) = \ln(W) - (1/2)x$$

10. Cette approximation est à peu près exacte pour une échelle de valeur des paramètres économiquement raisonnable. Par exemple, si le salaire horaire moyen d'un travail à temps complet était 10% plus élevé que le salaire horaire moyen d'un emploi à temps partiel, l'approximation prédirait une différence de 9,39%, ce qui est presque exact.

FIGURE 1  
PARTICIPATION À CHACUN DES RÉGIMES



Cette approximation fournit les spécifications suivantes en ce qui a trait au logarithme du salaire horaire moyen :

$$\ln R_{0i} = \ln W_{0i} - \theta_0 h_{0i} / 2, \quad (10a)$$

$$\ln R_{1i} = \ln W_{1i} - \theta_1 h_{1i} / 2. \quad (10b)$$

Le modèle statistique au complet est résumé comme suit :

Régime 1 (ne pas avoir d'emploi) :  $u_{0i} < -x_i' \gamma_0$  et  $u_{1i} < -x_i' \gamma_1$ .

Régime 2 (secteur formel uniquement) :  $u_{0i} > -x_i' \gamma_0$  et  $u_{1i} - \Psi_1 u_{0i} < -x_i' (\gamma_1 - \Psi_1 \gamma_0)$ ,

$$h_{0i} = \sigma_0 (x_i' \gamma_0 + u_{0i}) = \hat{y}_{1i} + e_{1i}, \quad (11a)$$

$$\ln R_{0i} = x_i' \beta_0 + \varepsilon_{0i} - \theta_0 \sigma_0 (x_i' \gamma_0 + u_{0i}) / 2 = \hat{y}_{2i} + e_{2i}. \quad (11b)$$

Régime 3 (secteur informel uniquement) :  $u_{1i} > -x_i' \gamma_1$  et  $u_{0i} - \Psi_0 u_{1i} < -x_i' (\gamma_0 - \Psi_0 \gamma_1)$ ,

$$h_{1i} = \sigma_1 (x_i' \gamma_1 + u_{1i}) = \hat{y}_{3i} + e_{3i}, \quad (11c)$$

$$\ln R_{1i} = x_i' \beta_1 + \varepsilon_{1i} - \theta_1 \sigma_1 (x_i' \gamma_1 + u_{1i}) / 2 = \hat{y}_{4i} + e_{4i}. \quad (11d)$$

Régime 4 (les deux secteurs à la fois) :  $u_{0i} - \Psi_0 u_{1i} > -x_i'(\gamma_0 - \Psi_0 \gamma_1)$  et  
 $u_{1i} - \Psi_1 u_{0i} > -x_i'(\gamma_1 - \Psi_1 \gamma_0)$ ,

$$h_{0i} = (1 - \Psi_0 \Psi_1) \sigma_1 [x_i'(\gamma_0 - \Psi_0 \gamma_1)^{-1} + u_{0i} - \Psi_0 u_{1i}] = \hat{y}_{5i} + e_{5i}, \quad (11e)$$

$$h_{1i} = (1 - \Psi_0 \Psi_1) \sigma_0 [x_i'(\gamma_1 - \Psi_1 \gamma_0)^{-1} + u_{1i} - \Psi_1 u_{0i}] = \hat{y}_{6i} + e_{6i}, \quad (11f)$$

$$\ln R_{0i} = x_i' \beta_0 + \varepsilon_{0i} - \theta_0 (1 - \Psi_0 \Psi_1)^{-1} \sigma_1 [x_i'(\gamma_0 - \Psi_0 \gamma_1) + u_{0i} - \Psi_0 u_{1i}] \\ = \hat{y}_{7i} + e_{7i}, \quad (11g)$$

$$\ln R_{1i} = x_i' \beta_1 + \varepsilon_{1i} - \theta_1 (1 - \Psi_0 \Psi_1)^{-1} \sigma_0 [x_i'(\gamma_1 - \Psi_1 \gamma_0) + u_{1i} - \Psi_1 u_{0i}] \\ = \hat{y}_{8i} + e_{8i}, \quad (11h)$$

où  $\hat{y}_{1i} = \sigma_0(x_i' \gamma_0)$ ,  $e_{1i} = \sigma_0 u_{0i}$ , etc... Enfin, un modèle statistique similaire peut être dérivé en remplaçant la spécification exponentielle par une spécification quadratique<sup>11</sup>. Cependant, la spécification exponentielle est préférée, car elle donne une spécification semi-logarithmique standard de Mincer du salaire horaire moyen.

## 2. IDENTIFICATION ET ESTIMATION

Cette section présente l'estimation, par la méthode du maximum de vraisemblance, du modèle du choix occupationnel, des heures de travail et du salaire horaire moyen présenté dans la section précédente. La décision qui détermine le choix de l'emploi est une question intéressante en soi; elle est applicable à toute une variété d'autres problèmes économiques<sup>12</sup>. L'estimation du modèle de choix discret qui caractérise le choix occupationnel est présentée séparément de celle du modèle général. Notons au passage qu'un modèle intermédiaire qui incorporerait à la fois les décisions concernant le choix occupationnel et les heures de travail pourrait aussi être estimé même si le salaire, ou plus généralement, la valeur de la production n'était pas observable. Un tel modèle serait donc utile pour analyser toute une série de problèmes relatifs à la théorie d'allocation du temps.

L'avantage du modèle général est qu'il permet d'identifier tous les paramètres intéressants du modèle même sans les restrictions d'exclusions standards. Le modèle pourrait aussi être estimé en deux étapes, en utilisant une généralisation de la procédure d'Heckman (1976)<sup>13</sup>.

11. Les gains quadratiques et les fonctions d'utilité prennent la forme  $Y_k = W_k h_k + \theta_k h_k^2/2$  pour laquelle le niveau des  $W$  (contrairement à leurs logarithmes) est une fonction linéaire des variables de  $x$  et des  $\varepsilon$  ( $W_k = x' \beta_k + \varepsilon_k$ ). Les salaires horaires moyens,  $R_k$ , sont exactement égaux à  $W_k + \theta_k h_k/2$ . La seule différence entre les deux modèles est que le modèle exponentiel conduit à une équation semi-logarithmique pour le salaire horaire moyen tandis que le modèle quadratique donne une spécification linéaire.

12. Le modèle est approprié pour estimer la décision discrète d'acheter deux biens de consommation, qu'ils soient substitués ou compléments (Voir DiNardo et Lemieux, 1992).

13. Voir Lemieux (1989) pour une application de cette technique d'estimation.

### 2.1 *Modèle du choix occupationnel*

La fonction de vraisemblance logarithmique du modèle du choix occupationnel est formée en utilisant les conditions de participation (9a) à (9d) :

$$\begin{aligned}
 L = & \sum_1 \ln \Phi_2(-x'_i \gamma_0, -x'_i \gamma_1; \rho) + \sum_2 \ln \Phi_2\left(x'_i \gamma_0, \frac{x'_i(-\gamma_1 + \psi_1 \gamma_0)}{s_1}; -\frac{\rho - \psi_0}{s_1}\right) \\
 & + \sum_3 \ln \Phi_2\left(x'_i \gamma_1, \frac{x'_i(-\gamma_0 + \psi_0 \gamma_1)}{s_0}; -\frac{\rho - \psi_1}{s_0}\right) \\
 & + \sum_4 \ln \Phi_2\left(-\frac{x'_i(-\gamma_0 + \psi_0 \gamma_1)}{s_0}, \frac{x'_i(-\gamma_1 + \psi_1 \gamma_0)}{s_1}; -\frac{(1 + \psi_0 \psi_1)\rho - \psi_0 - \psi_1}{s_0 s_1}\right)
 \end{aligned} \quad (12)$$

où  $\Phi_2(x, y; r)$  est une fonction de répartition normale standard à deux variables, où  $\Sigma$  indique la sommation de toutes les personnes dans le régime  $k$ , et où  $(s_0)^2 = 1 + (\psi_0)^2 - 2\rho\psi_0$  et  $(s_1)^2 = 1 + (\psi_1)^2 - 2\rho\psi_1$ . Les paramètres  $\gamma_0$ ,  $\gamma_1$ ,  $\psi_0$ ,  $\psi_1$  et  $\rho$  sont identifiés dans le modèle du choix occupationnel. Cela signifie qu'aucun des paramètres  $\beta_0$ ,  $\beta_1$ ,  $\beta_R$  ne sont identifiés à moins que des restrictions d'exclusions ne soient imposées sur  $x_0$ ,  $x_1$  et  $x_R$ . La fonction de vraisemblance est relativement facile à maximiser puisqu'il s'agit seulement d'intégrer numériquement des fonctions de densité normale à deux variables.

Dans le cas particulier où  $\psi_0 = \psi_1 = 0$ , le modèle du choix occupationnel est un probit bivarié standard. Dans ce cas, les décisions concernant la participation et les heures de travail dans les secteurs formel et informel sont prises séparément. Or, ceci est en contradiction avec l'hypothèse selon laquelle les heures de travail dans les deux secteurs sont parfaitement substituables. Savoir si la condition  $\psi_0 = \psi_1 = 0$  est réalisée, ou non, représente un test de spécification du modèle proposé ici.

### 2.2 *Modèle caractérisant le choix occupationnel, les heures de travail et le salaire horaire moyen*

La fonction de vraisemblance logarithmique du modèle caractérisant le choix occupationnel, les heures de travail et le salaire horaire moyen est formée sur la base de l'ensemble des équations (11) :

$$\begin{aligned}
 L = & \sum_1 \ln \Phi_2(-x'_i \gamma_0, -x'_i \gamma_1; \rho) + \sum_2 \ln f_{12}(e_{1i}, e_{2i}) + \sum_3 \ln f_{34}(e_{3i}, e_{4i}) \\
 & + \sum_2 \ln \Phi_1\left(\left[x'_i(-\gamma_1 + \psi_1 \gamma_0) - E(u_{1i} - \psi_1 u_{0i} | e_{1i}, e_{2i})\right] \left[Var(u_{1i} - \psi_1 u_{0i} | e_{1i}, e_{2i})\right]^{1/2}\right) \\
 & + \sum_3 \ln \Phi_1\left(\left[x'_i(-\gamma_0 + \psi_0 \gamma_1) - E(u_{0i} - \psi_0 u_{1i} | e_{3i}, e_{4i})\right] \left[Var(u_{0i} - \psi_0 u_{1i} | e_{3i}, e_{4i})\right]^{1/2}\right) \\
 & + \sum_4 \ln f_{5678}(e_{5i}, e_{6i}, e_{7i}, e_{8i})
 \end{aligned} \quad (13)$$

où  $f_{12}(\cdot)$ ,  $f_{34}(\cdot)$  et  $f_{5678}(\cdot)$  sont les fonctions de densité de  $(e_{1i}, e_{2i})$ ,  $(e_{3i}, e_{4i})$  et  $(e_{5i}, e_{6i}, e_{7i}, e_{8i})$ , respectivement, tandis que  $\Phi_1(\cdot)$  est une fonction de répartition normale standard à une variable. L'estimation du modèle général exploite toutes les informations disponibles dans les données. Les paramètres structurels  $\beta_0$ ,  $\beta_1$ ,  $\beta_R$ ,  $\psi$ ,  $\theta_0$ , et  $\theta_1$ , de même que la matrice de variance-covariance de  $(\varepsilon_0, \varepsilon_1, \varepsilon_R)$  sont identifiés. Il est ainsi possible de calculer les résultats de statique comparative dérivés à la section 1, à partir des estimations obtenues en maximisant l'équation (13).

Les paramètres structurels du modèle sont identifiés sans restrictions d'exclusions, en dépit du fait que les taux de salaire dans les secteurs formel et informel sont tous deux endogènes. Considérons le taux de salaire (le salaire horaire moyen) dans le secteur formel,  $\ln R_0$  : ce taux de salaire est endogène dans le secteur formel même s'il ne dépend pas des heures travaillées ( $\theta_0 = 0$ ), puisque  $\varepsilon_0$  et  $\varepsilon_R$  sont spécifiquement corrélés. De plus, les salaires et les heures de travail sont déterminés conjointement dans le cas où les salaires dépendent des heures de travail ( $\theta_0 \neq 0$ ). Le modèle est néanmoins identifié lorsque les équations des heures de travail et du salaire horaire moyen sont estimées conjointement dans tous les régimes. La raison en est que les équations des heures de travail dans le régime 2 (secteur formel uniquement) et dans le régime 3 (secteur informel uniquement) sont des équations d'offre de travail contrainte, tandis que les équations des heures de travail dans le régime 4 (les deux secteurs à la fois) sont des équations d'offre de travail non contrainte. Le fait que les élasticités d'offre de travail contrainte soient plus petites que les élasticités d'offre de travail non contrainte, lorsque les heures de travail dans les deux secteurs, fournit la source de variation nécessaire pour identifier le modèle. L'identification du modèle est discutée en de plus amples détails à l'Annexe A.

### 3. LES RÉSULTATS

#### 3.1 *Les données*

Les données utilisées dans ce papier proviennent d'une enquête réalisée dans la région de Québec au printemps 1986. Cette enquête posait des questions rétrospectives à propos des activités du marché du travail durant l'année 1985, en sus de questions générales sur l'attitude des gens vis à vis de la taxation et de l'évasion fiscale. Les activités du marché du travail ont alors été classées selon deux catégories, activités du secteur formel et activités du secteur informel, ceci en regardant si le salaire provenant du travail était reporté ou non dans le rapport d'impôt du travailleur. Des détails plus spécifiques sur l'enquête peuvent être trouvés dans Fortin et Fréchette (1987) ou Lemieux, Fortin et Fréchette (1994).

Le modèle non linéaire estimé dans ce papier est sensible à la présence d'observations extrêmes. Cinq pour-cent des observations originales ont donc été éliminées en utilisant une procédure simple d'effacement de ces observations<sup>14</sup>. L'échantillon final comprend 1 772 hommes et femmes âgés de 18 à 55 ans. Sur ces 1 772 personnes, 393 sont classés dans le régime 1 (pas d'emploi), 1 247 dans le régime 2 (travailler dans le secteur formel uniquement), 48 dans le régime 3 (travailler dans le secteur informel uniquement) et 84 dans le régime 4 (travailler dans les deux secteurs à la fois). Dans cet échantillon, 75,1% des gens travaillent dans le secteur formel et 7,5% dans le secteur informel. Les caractéristiques moyennes des personnes dans les différents régimes sont reportées au tableau 1. Elles sont composées des heures de travail, du salaire horaire moyen, du taux marginal d'imposition et d'autres variables caractéristiques incluses dans le vecteur  $x_i$  (expérience potentielle et expérience au carré, une variable dichotomique représentant le sexe et le nombre d'années d'éducation).

Les variables mesurant les activités de travail dans l'économie informelle risquent fort d'être contaminées par un biais substantiel de sous-évaluation. Lemieux, Fortin et Fréchette (1994) examinent la validité de ces mesures et concluent que le problème de sous-évaluation est substantiel, mais n'affecte pas, de manière significative, les mesures de covariance entre les heures travaillées, les revenus et le taux de salaire dans le secteur informel. D'une part, les résultats reportés ci-dessous sont tous basés sur les estimations des formes réduites. Ils devraient donc être exempts de l'erreur classique de mesure concernant les heures travaillées et le salaire horaire moyen. D'autre part, l'erreur de mesure affectera spécifiquement la matrice de variance-covariance des résidus du salaire horaire moyen et des heures de travail<sup>15</sup>.

### 3.2 Traitement des taxes

Un des facteurs clés expliquant la décision de travailler dans le secteur informel est le taux d'imposition associé, soit aux taxes sur la masse salariale et à l'impôt sur le revenu, soit à la taxation implicite des programmes de soutien du revenu. L'effet des taxes et des programmes de soutien du revenu est résumé implicitement par un taux marginal d'imposition unique. Un problème bien connu relatif à ce taux d'imposition est que c'est une variable endogène de l'équation d'offre de travail du secteur formel. Ceci est dû à la nature non linéaire des taxes implicites et explicites (voir Hausman, 1980). Examiner entièrement le problème de taxes endogènes dans un modèle comprenant une fonction de gain non linéaire,

---

14. La procédure consiste à éliminer les observations pour lesquelles la valeur  $(d_i - d)' \Sigma^{-1} (d_i - d)$  excède un seuil approprié, sont éliminées.  $d_i$  est un vecteur qui contient des variables endogènes et exogènes de l'observation  $i$ , tandis que  $d$  et  $\Sigma$  sont, respectivement, la moyenne empirique et la matrice de variance-covariance de ces mêmes variables.

15. Magnac (1991) montre comment incorporer une erreur de mesure classique dans la fonction de vraisemblance pour un modèle partageant des caractéristiques identiques à celui utilisé dans ce papier.

TABLEAU 1

MOYENNES DES VARIABLES UTILISÉES DANS L'ANALYSE EMPIRIQUE  
(ÉCARTS-TYPES ENTRE PARENTHESES)

	Échantillon complet	Régime 1 ( $h_0=0, h_1=0$ )	Régime 2 ( $h_0>0, h_1=0$ )	Régime 3 ( $h_0=0, h_1>0$ )	Régime 4 ( $h_0>0, h_1>0$ )
Salaire horaire moyen dans le secteur formel	–	–	13,935 (7,521)	–	8,823 (5,266)
Salaire horaire moyen dans le secteur informel	–	–	–	4,735 (2,685)	7,465 (3,852)
Taux marginal d'impôt implicite	0,453 (0,226)	0,488 (0,327)	0,435 (0,159)	0,688 (0,373)	0,402 (0,287)
Âge	36,62 (11,24)	41,30 (12,59)	36,10 (10,25)	26,63 (10,35)	28,37 (8,53)
Sexe (Femme = 1)	0,527 (0,499)	0,824 (0,381)	0,444 (0,497)	0,612 (0,492)	0,321 (0,470)
Éducation	11,344 (3,550)	9,216 (3,376)	11,90 (3,40)	11,694 (3,029)	12,845 (2,783)
Heures dans le secteur formel	1 089,92 (814,26)	–	1 481,93 (578,44)	–	1 010,20 (699,24)
Heures dans le secteur informel	18,36 (81,41)	–	–	311,78 (195,59)	205,52 (162,32)
Observations : (% du total)	1 772 100,0	393 22,1	1 247 70,4	49 2,7	84 4,7

dépasse le cadre de ce papier. C'est pourquoi un traitement simple du taux d'imposition est adopté ici, suivant ainsi une approche utilisée par Hall (1973) qui consiste en une approximation linéaire de la contrainte budgétaire autour d'une valeur observée du revenu imposable  $I + Y_0$  :

$$T(I + Y_0) \approx T_0 + \tau(I + Y_0)$$

où  $\tau(y)$  est le taux marginal implicite d'imposition lorsque le revenu est égal à  $y$ . Le taux marginal implicite d'imposition, utilisé dans ce papier, est calculé en utilisant les tables d'impôt et les réglementations des programmes de soutien du revenu qui prévalaient dans la province de Québec en 1985<sup>16</sup>.

Les équations des heures de travail dérivées à la section 2 sont modifiées pour prendre en compte les taxes, en remplaçant le salaire du secteur formel correspondant à zéro heures travaillées  $W_0$  par le taux de salaire net  $W_0(1 - \tau)$ <sup>17</sup>. Les équations de forme réduite pour les heures de travail (11a, 11c, 11e, 11f) sont ensuite modifiées, en remplaçant  $x'\gamma_0$  par  $x'\gamma_0 - \gamma_{0\tau} \ln(1 - \tau)$ . Le paramètre  $\gamma_{0\tau}$  devrait être égal à un, mais sa valeur estimée pourrait être inférieure à cause de l'erreur de mesure du taux marginal implicite d'imposition ou de l'illusion fiscale. Un paramètre  $\gamma_{1\tau}$  est introduit pour capturer la possibilité que le taux d'imposition puisse affecter le retour net espéré à travailler dans le secteur informel. Par exemple, le retour net espéré à travailler dans le secteur informel serait affecté par le taux d'imposition si la probabilité d'être pris par les autorités était positive, sachant que les travailleurs qui sont pris doivent payer leurs taxes plus une pénalité.

### 3.3 Estimation par maximum de vraisemblance du modèle caractérisant le choix occupationnel

Le modèle du choix occupationnel est estimé en maximisant, numériquement, la fonction de vraisemblance (12). Ce modèle qui caractérise le choix de l'emploi est appelé modèle «probit structurel» ; les estimations de ce modèle sont reportées aux colonnes 1 et 2 du tableau 2. Le taux marginal d'imposition est inclus dans la spécification de la colonne 1, mais exclu de la spécification de la colonne 2. La valeur estimée de  $\psi_0$  que l'on trouve dans la colonne 1 est positive (2,37) et statistiquement différente de zéro. Celle de  $\psi_1$  (0,0176) est aussi positive mais pas statistiquement différente de zéro. De plus, le modèle est assez sensible aux variables incluses dans les variables explicatives  $x_i$ . Par exemple, lorsque le taux marginal d'imposition est exclu de  $x_i$ , à la colonne 2, le paramètre  $\psi_1$  devient négatif ce qui viole les conditions de second ordre pour un optimum.

16. L'auteur tient à remercier Louis Beauséjour pour lui avoir fourni le programme informatique utilisé pour calculer le taux marginal implicite d'imposition.

17. Ce résultat a été obtenu en remplaçant le salaire brut du secteur du secteur formel,  $I + Y_0$ , par le revenu régulier net,  $I + Y_0 - T(I + Y_0)$ , dans le Lagrangien.



TABLEAU 2

ESTIMÉS DES MODELES DE PARTICIPATION PAR MAXIMUM DE VRAISEMBLANCE

	Probit structurel		Probit bivarié	
	(1)	(2)	(3)	(4)
Paramètres dans $\gamma_0$				
Constante	1,5834 (0,2607)	0,5008 (0,2312)	1,2156 (0,2415)	0,0770 (0,2030)
Expérience	-0,0010 (0,0110)	0,0282 (0,0105)	0,0196 (0,0096)	0,0501 (0,0088)
Expérience au carré divisée par 100	-0,0608 (0,0209)	-0,0998 (0,0203)	-0,0909 (0,0190)	-0,1322 (0,0183)
Sexe (femme = 1)	-1,1550 (0,0918)	-1,0790 (0,0835)	-1,0638 (0,0861)	-0,9805 (0,0791)
Éducation	0,0678 (0,0143)	0,0879 (0,0139)	0,0683 (0,0141)	0,0893 (0,0137)
Paramètres dans $\gamma_1$				
Constante	-0,5668 (0,2475)	-0,4527 (0,5637)	-0,6446 (0,2846)	-0,2785 (0,2634)
Expérience	-0,0653 (0,0138)	-0,0923 (0,0194)	-0,0659 (0,0143)	-0,0752 (0,0136)
Expérience au carré divisée par 100	0,0743 (0,0377)	0,1309 (0,0566)	0,0761 (0,0391)	0,0900 (0,0381)
Sexe (femme = 1)	-0,2894 (0,1052)	0,1702 (0,6412)	-0,2144 (0,1013)	-0,1724 (0,0980)
Éducation	-0,0136 (0,0191)	-0,0598 (0,0515)	-0,0134 (0,0179)	-0,0225 (0,0175)
Effet du taux marginal d'imposition implicite [ $-\ln(1 - \tau)$ ]				
Secteur formel ( $\gamma_{0\tau}$ )	-0,5311 (0,0541)	–	-0,5752 (0,0437)	–
Secteur informel ( $\gamma_{0\tau}$ )	0,1998 (0,0467)	–	0,2141 (0,0505)	–
$\Psi_0$	2,3703 (0,7192)	54,3431 ( $2,15 \cdot 10^3$ )	–	–
$\Psi_1$	0,0176 (0,0285)	-0,4142 (0,6461)	–	–
Corr( $u_0, u_1$ )	0,3727 (0,2379)	-0,0850 (0,4853)	-0,2193 (0,0709)	-0,3160 (0,0653)
Log-vraisemblance	-1 121,35	-1 200,03	-1 126,98	-1 207,55
Nombre de paramètres	15	13	13	11

Un modèle probit bivarié est un cas particulier du modèle probit structurel, dans lequel  $\psi_0 = \psi_1 = 0$ . Les estimations du modèle probit bivarié sont données dans les colonnes 3 et 4 du tableau 2. D'une part, la restriction  $\psi_0 = \psi_1 = 0$  est rejetée, en utilisant le test du ratio de vraisemblance avec ou sans le taux marginal d'imposition, soient, respectivement,  $(11,26 \sim \chi^2(2))$  et  $(15,04 \sim \chi^2(2))$ . D'autre part, les paramètres estimés du modèle probit bivarié sont moins sensibles à l'inclusion du taux d'imposition que les paramètres estimés du modèle probit structurel<sup>18</sup>.

Les paramètres estimés  $\gamma_0$  et  $\gamma_1$  du modèle probit structurel sont difficiles à interpréter. Dans le tableau 3, les valeurs estimées des paramètres, reportées dans la colonne 1 du tableau 2, sont donc reformulées en terme d'effets marginaux de variables exogènes sur les probabilités de participer à chaque régime. Le résultat important reporté au tableau 3 est qu'une hausse d'1 % du taux d'imposition augmente la probabilité de ne pas travailler (régime 1) de 0,13 % et réduit la probabilité d'appartenir au régime 2 (secteur formel uniquement) de 0,15 %. Par contre cette hausse a un effet moins important sur la probabilité de travailler dans le secteur informel (0,016 % pour le régime 3 et 0,007 % pour le régime 4). Bien que ces derniers effets estimés soient faibles, ils sont importants en terme d'élasticité puisque peu de gens font partie des régimes 3 et 4. Les élasticités de participation par rapport au taux d'imposition sont égales respectivement à 0,57, -0,21, 0,57 et 0,16 pour les régimes 1 à 4.

TABLEAU 3

IMPACT MARGINAL SUR LES PROBABILITÉS DANS LE MODÈLE PROBIT STRUCTUREL

	(1) Régime 1	(2) Régime 2	(3) Régime 3	(4) Régime 4
Expérience	0,0006	0,0066	-0,0025	-0,0046
(Exp. <sup>2</sup> )/100	0,0142	-0,0224	0,0038	0,0045
Sexe	0,2777	-0,2482	0,0053	-0,0349
Éducation	-0,0161	0,0178	-0,0015	-0,0001
Taux d'imposition	0,1261	-0,1489	0,0155	0,0074

NOTE : Chiffres basés sur les résultats rapportés au tableau 2, colonne 1. Ces impacts sont évalués à la valeur moyenne des variables explicatives.

18. Une partie du problème est que le modèle probit structurel est cohérent même avec des valeurs limites des paramètres (par exemple la corrélation entre  $u_0$  et  $u_1$  égale à un). Tel n'est pas le cas du modèle probit bivarié (une corrélation de un implique une probabilité nulle d'appartenir à au moins un régime). Ce problème est discuté plus en détails par Lemieux (1989).

### 3.4 Estimation par maximum de vraisemblance du modèle caractérisant le choix de l'emploi, des heures de travail et du salaire horaire moyen

Le tableau 4 présente les estimations de plusieurs spécifications du modèle général caractérisant le choix de l'emploi, les heures de travail et le salaire horaire moyen. La colonne 1 présente les estimations du modèle qui inclut le taux marginal d'imposition et dans lequel le taux de salaire dans le secteur formel ne dépend pas des heures travaillées ( $\theta_0 = 0$ ). Contrairement au modèle probit structural, les estimations de ce modèle donnent les résultats attendus, lorsque le taux d'imposition n'est pas inclus parmi les variables explicatives (colonne 2). De plus, la valeur estimée de  $\gamma_{0r}$  est positive et significative (voir la colonne 1), alors que celle de  $\gamma_{1r}$  est petite et statistiquement non différente de zéro. Cela suggère, comme attendu, que les travailleurs choisissent leur secteur et leurs heures de travail en fonction du gain net dans le secteur formel et du gain brut dans le secteur informel.

TABLEAU 4

ESTIMÉS DU MODELE AVEC HEURES DE TRAVAIL ET SALAIRE HORAIRE MOYEN  
PAR LA MÉTHODE DU MAXIMUM DE VRAISEMBLANCE

	(1)	(2)	(3)
Paramètres dans $\gamma_0$			
Constante	1,1690 (0,1464)	0,6625 (0,1448)	1,0750 (0,1448)
Expérience	0,0682 (0,0065)	0,0736 (0,0067)	0,0741 (0,0063)
Expérience carrée divisée par 100	-0,1642 (0,0142)	-0,1708 (0,0135)	-0,1741 (0,0140)
Sexe (femme = 1)	-0,9635 (0,0526)	-0,9190 (0,0550)	-0,9219 (0,0531)
Éducation	0,0306 (0,0085)	0,0348 (0,0089)	0,0292 (0,0084)
Paramètres dans $\gamma_1$			
Constante	0,2648 (0,2626)	0,7487 (0,2049)	-0,9176 (0,3405)
Expérience	0,0047 (0,0125)	0,0363 (0,0205)	-0,0484 (0,0142)
Expérience carrée divisée par 100	-0,0609 (0,0286)	-0,1153 (0,0368)	0,0433 (0,0342)
Sexe (femme = 1)	-0,8559 (0,0959)	-0,9468 (0,0796)	-0,2276 (0,1311)
Éducation	0,0083 (0,0160)	0,0158 (0,0151)	-0,0105 (0,0211)

TABLEAU 4 (suite)

	(1)	(2)	(3)
Paramètres dans $\gamma_R$			
Constante	0,3820 (0,1748)	1,1764 (0,1885)	1,5895 (0,1517)
Expérience	0,0706 (0,0118)	0,1312 (0,0131)	0,1605 (0,0075)
Expérience carrée divisée par 100	-0,0705 (0,0233)	-0,1854 (0,0269)	-0,2660 (0,0168)
Sexe (femme = 1)	0,2208 (0,0794)	-0,2389 (0,1043)	-0,7340 (0,0606)
Éducation	0,1176 (0,0127)	0,1714 (0,0128)	0,1870 (0,0098)
Effet du taux marginal d'imposition $[-\ln(1 - \tau)]$			
Secteur formel ( $\gamma_{0\tau}$ )	-0,4356 (0,0410)	–	-0,4972 (0,0399)
Secteur informel ( $\gamma_{1\tau}$ )	0,0562 (0,0456)	–	0,2724 (0,0495)
$\theta_0$	0	0	-0,1459 (0,1160)
$\theta_1$	1,4365 (0,2493)	0,3563 (0,2250)	0,9142 (0,3536)
$\psi$	0,6457 (0,0488)	0,3580 (0,0605)	0,0682 (0,0561)
$\sigma_0$	0,8480 (0,0173)	0,8742 (0,0225)	0,8753 (0,0195)
$\sigma_1$	0,3194 (0,0146)	0,4228 (0,0846)	0,3153 (0,0123)
$\sigma_R$	0,5998 (0,0334)	0,4647 (0,0188)	0,4747 (0,0144)
$\rho_{01}$	0,6622 (0,0610)	0,9085 (0,0846)	-0,4125 (0,0952)
$\rho_{0R}$	-0,6734 (0,0454)	-0,2866 (0,0881)	0,4088 (0,0565)
$\rho_{1R}$	-0,5683 (0,0314)	-0,4008 (0,1205)	-0,3951 (0,0435)
Log-vraisemblance	-3 220,7	-3 311,4	-3 149,3
Nombre de paramètres	25	23	26

NOTE : Les heures de travail  $h_0$  et  $h_1$  sont mesurées en milliers (par année).

La contrainte  $\theta_0 = 0$  est relâchée dans la spécification rapportée dans la colonne 3. Les valeurs estimées de  $\theta_0$  et de  $\theta_1$  sont tout à fait raisonnables : celle de  $\theta_0$  est négative (-0,146) ce qui implique que la fonction de gain dans le secteur formel est une fonction convexe des heures de travail. Le résultat est cohérent avec le fait que le salaire d'un travail à temps partiel est plus bas que le salaire de n'importe quel autre travail similaire à temps plein. De plus, la valeur estimée du paramètre de courbure,  $\theta_1$ , est positive (0,914), ce qui implique que la fonction de gain dans le secteur informel est une fonction concave des heures travaillées, comme le supposait les hypothèses. D'autre part, la valeur estimée du paramètre  $\psi$  est plus faible, en valeur absolue, que la valeur estimée de  $\theta_1$  ce qui est incompatible avec les conditions de second ordre ( $\psi + \theta_1 > 0$ ). La discussion des résultats va donc se limiter au cas où le taux de salaire dans le secteur formel ne dépend pas des heures travaillées (colonne 1), comme dans le modèle d'offre de travail standard.

La valeur estimée du paramètre de courbure,  $\theta_1$ , reportée à la colonne 1 est positive (1,437) et significative. La magnitude de ce paramètre estimé est facilement interprétable en utilisant l'équation (10b) qui pose  $\ln(R_1) = \ln(W_1) - \theta_1 h_1/2$ . Cela signifie que le salaire horaire moyen dans le secteur informel décroît de 7,2 % pour chaque 100 heures supplémentaires de travail (par année) dans le secteur informel<sup>19</sup>. Les autres paramètres structurels du modèle sont interprétés à l'aide d'une série de transformation présentées au tableau 5. La partie A du tableau 5 donne, par exemple, les effets marginaux des variables explicatives sur les probabilités de participer à chaque régime. Comparés aux effets marginaux calculés pour le modèle du choix occupationnel, les effets marginaux des taxes dans le modèle général sont plus importants dans le régime 4 (les deux secteurs à la fois), plus faibles dans le régime 1 (pas d'emploi) et comparables dans les régimes 2 (secteur formel uniquement) et 3 (secteur informel uniquement). Les élasticités implicites de participation par rapport au taux marginal d'imposition sont maintenant égales à 0,38, -0,19, 0,53 et 0,73, respectivement, pour les régimes 1 à 4. De surcroît, il est maintenant possible de calculer l'effet marginal des taux de salaires correspondant à zéro heures travaillées  $\ln W_0$  et  $\ln W_1$ , et du salaire de réserve,  $\ln W_R$ , sur la probabilité de participation. Les signes de ces effets marginaux sont d'ailleurs tels qu'attendus. L'effet des variables exogènes sur  $\ln W_0$ ,  $\ln W_1$  et  $\ln W_R$  sont reportés dans la partie B du tableau 5. Ces effets sont simplement les paramètres  $\beta_0$ ,  $\beta_1$  et  $\beta_R$  du modèle. Il est intéressant de noter que le taux de rendements sur les mesures standards de capital humain, telles que l'expérience et l'éducation sont plus importants dans le secteur formel que dans le secteur informel. De plus, les femmes se trouvent à avoir des salaires et des salaires de réserve plus faibles que ceux des hommes. Leur offre de travail moins grande est donc attribuable à un mélange de ces deux effets.

19. Les heures sont exprimées en milliers (par année) dans le modèle estimé.

TABLEAU 5

## SOMMAIRE DES EFFETS DES VARIABLES EXPLICATIVES

A. Impact marginal sur les probabilités : modèle caractérisant le choix occupationnel, les heures de travail et les salaires horaires				
	(1) Régime 1	(2) Régime 2	(3) Régime 3	(4) Régime 4
Expérience	-0,0135	0,0196	-0,0019	-0,0042
(Exp <sup>2</sup> )/100	0,0332	-0,0412	0,0032	0,0048
Sexe	0,2030	-0,1819	0,0054	-0,0265
Éducation	-0,0061	0,0081	-0,0007	-0,0012
Taux d'imposition	0,0847	-0,1356	0,0144	0,0365
Log(W <sub>0</sub> )	-0,3590	0,5402	-0,0540	-0,1272
Log(W <sub>1</sub> )	-0,0239	-0,1812	0,0406	0,1645
Log(W <sub>R</sub> )	0,3829	-0,3590	0,0134	-0,0372
B. Effet sur les salaires et sur le salaire réserve lorsque les heures sont égales à zéro				
	(5) Log(W <sub>0</sub> )	(6) Log(W <sub>1</sub> )	(7) Log(W <sub>R</sub> )	
Expérience	0,0796	0,0455	0,0423	
(Exp <sup>2</sup> )/100	-0,1321	-0,0828	-0,0423	
Sexe	-0,3951	-0,4368	0,1324	
Éducation	0,0873	0,0760	0,0705	
C. Élasticité des heures dans chacun des régimes				
	(8) Régime 2 : Secteur formel (h <sub>0</sub> )	(9) Régime 3 : Secteur infor. (h <sub>1</sub> )	(10) Régime 4 : Secteur formel (h <sub>0</sub> )	(11) Régime 4 : Secteur infor. (h <sub>1</sub> )
Expérience	0,0390	0,0048	0,0808	-0,1158
(Exp <sup>2</sup> )/100	-0,0939	-0,0624	-0,1718	0,1672
Sexe	-0,5514	-0,8768	-0,7802	-0,1410
Éducation	0,0175	0,0085	0,0334	-0,0381
Taux d'imposition	-0,2493	0,0576	-0,5558	0,9346
Log(W <sub>0</sub> )	1,0451	0	2,2222	-3,3873
Log(W <sub>1</sub> )	0	1,5404	-0,6891	3,3873
Log(W <sub>R</sub> )	1,0451	-1,5404	-1,5331	0

NOTE : Calculs basés sur les résultats rapportés à la colonne 1 du tableau 4. W<sub>0</sub> (W<sub>1</sub>) est le salaire dans le secteur formel (informel) lorsque les heures de travail sont égales à zéro. W<sub>R</sub> est le salaire de réserve.

Finalement, la partie C du tableau 5 est constituée des semi-élasticités des heures de travail par rapport aux variables contenues dans le vecteur  $x_i$ , au taux marginal d'imposition  $\ln W_0$ ,  $\ln W_1$  et  $\ln W_R$ . Les trois dernières rangées du tableau 5 représentent donc les élasticités d'offre de travail qui découlent des estimations du modèle. Celles-ci correspondent aux résultats de statique comparative de la section 3 exprimés en terme d'élasticité. Elles sont assez grandes et ont le signe attendu. L'élasticité des heures travaillées dans le secteur informel par rapport au taux de salaire correspondant à zéro heure travaillée ( $\ln W_1$ ) est, en particulier, positive pour les régimes 3 (1,54) et 4 (3,39)<sup>20</sup>. Les élasticités des heures travaillées dans le secteur formel par rapport au salaire du secteur formel ( $\ln W_0$ ) sont égales à 1,05, dans le régime 2, et à 2,22, dans le régime 4. Les élasticités du régime 4 sont plus grandes que celles des régimes 2 et 3, comme prévu par la théorie de la demande avec contraintes de rationnement. Ces élasticités sont assez grandes, mais des réserves doivent être émises à leur sujet pour plusieurs raisons. Premièrement, seule l'élasticité des heures travaillées dans le secteur formel du régime 2 (1,05) est strictement comparable aux élasticités obtenues dans les études sur l'offre de travail standard. De surcroît, les hommes et les femmes sont inclus dans l'échantillon et les élasticités d'offre de travail sont typiquement plus grandes pour les femmes que pour les hommes<sup>21</sup>.

Finalement, le même modèle est utilisé pour expliquer les décisions de participation et les décisions des heures de travail ; il se pourrait que les décisions de participation soient plus sensibles aux changements dans le taux de salaire que les heures de travail. Les estimations obtenues ici sont donc seulement suggestives puisqu'elles sont basées sur peu d'observations disponibles quant aux personnes travaillant dans le régime 4. Une stratégie possible pour de futurs travaux serait d'appliquer les méthodes développées dans ce papier à des problèmes plus standards tels que la décision d'offre de travail à l'intérieur de ménages sur lesquels on dispose de plus d'informations.

#### CONCLUSION

Nous avons présenté un modèle caractérisant le choix de l'emploi, des heures de travail et des salaires, lorsqu'il est possible pour les travailleurs de travailler dans plus d'une occupation en même temps. Le modèle est basé sur l'idée que les gains provenant du travail sont des fonctions non linéaires des heures travaillées et que les préférences dépendent seulement de la consommation et du loisir. Le modèle est appliqué à la décision de travailler dans le secteur formel et informel. Les résultats obtenus peuvent être résumés comme suit :

20. Des élasticités importantes sont obtenues également par Lemieux, Fortin et Fréchette (1994), en utilisant des procédures d'estimation très différentes, basées sur l'exclusion de certaines variables de l'équation de  $\ln W_1$ , telle que le statut syndical dans le secteur formel. Ils trouvent aussi que la fonction de gain dans le secteur informel est concave, en utilisant une procédure de variables instrumentales.

21. Par exemple, voir la recension de littérature de Killingsworth (1983).

- 1) Les implications du modèle du choix occupationnel sont imbriquées dans un modèle probit structurel. Ce modèle de choix discret est préféré à un modèle probit bivarié standard.
- 2) Les élasticités estimées des heures de travail sont plus grandes, comparées aux résultats habituellement obtenus dans les études sur l'offre de travail. Il existe, en particulier, un haut degré de substituabilité entre les heures travaillées dans les secteurs formel et informel pour les travailleurs offrant des heures de travail aux deux secteurs à la fois.
- 3) La valeur estimée du paramètre représentant la courbure,  $\theta_1$ , est positive ce qui signifie que la fonction de gain dans le secteur informel est une fonction concave des heures de travail.
- 4) Une hausse du taux marginal d'imposition augmente toujours la probabilité de travailler dans le secteur informel, comme elle augmente le nombre d'heures travaillées dans ce secteur si on y participe déjà.

Sur le plan méthodologique, l'approche proposée ici fournit une nouvelle façon d'estimer les élasticités d'offre de travail lorsque le taux de salaire est corrélié avec les préférences et qu'il n'existe pas de variables instrumentales disponibles. Cela suggère également une stratégie empirique pour estimer la relation entre le salaire horaire moyen et le nombre d'heures travaillées. Finalement, le modèle et les techniques d'estimation présentées dans ce papier sont applicables à toute une variété de problèmes économiques connexes incluant le double emploi, les travailleurs autonomes et l'allocation du temps pour les agriculteurs dans les pays en voie de développement.



## ANNEXE A

## IDENTIFICATION DU MODELE AVEC HEURES DE TRAVAIL ET SALAIRE HORAIRE MOYEN

Sous l'hypothèse de normalité, la procédure du maximum de vraisemblance conduit à des estimations efficaces des paramètres structurels du modèle et correctement ajustés pour l'autosélection. Dans le modèle général de la section 2.2, les paramètres structurels sont identifiés uniquement grâce aux restrictions sur les équations de forme réduite des heures de travail et du salaire horaire moyen dans les différents régimes<sup>22</sup>. Le but de cette annexe est d'expliquer en terme économiques d'où viennent ces restrictions en montrant qu'elles découlent du fait que les élasticités d'offre de travail dans un secteur sont plus grandes lorsque le travailleur offre des heures de travail à l'autre secteur que lorsqu'il ne le fait pas. Ce principe de base, bien connu de la littérature concernant l'analyse de la demande avec rationnement, est une application du célèbre principe de Le Chatelier de Samuelson<sup>23</sup>.

Pour simplifier l'exposé, cette annexe ignore les considérations d'efficacité et d'autosélection, pour montrer comment le modèle est identifié dans un contexte de régression linéaire standard. Les problèmes d'efficacité et d'autosélection pourraient être abordés, en construisant des estimateurs en deux étapes des formes réduites et ensuite, en utilisant la méthode de distance minimale pour imposer les restrictions entre régimes<sup>24</sup>.

**Identification de  $\theta_0$  et  $\theta_1$** 

Substituons  $\ln(W_1) = x'\beta_1 + \varepsilon_1$  dans l'équation du salaire horaire moyen (10b) dans le secteur informel :

$$\ln R_1 = x'\beta_1 - (\theta_1/2)h_1 + \varepsilon_1 \quad (\text{A1})$$

Les formes réduites des heures de travail,  $h_1$ , dans le secteur informel sont différentes combinaisons linéaires du vecteur  $x$ , ceci selon que l'individu travaille (11f), ou non (11c), dans le secteur formel. La raison en est que chaque élément de  $x$ , comme par exemple l'éducation, affecte la productivité des travailleurs du secteur informel ( $\ln(W_1)$ ) dans le régime 3 (secteur informel uniquement) et dans le régime 4 (les deux secteurs à la fois), de la même manière, mais produit des ajustements dans les heures travaillées différents. En d'autres mots, la forme réduite des heures de travail est une fonction d'éléments de  $x$  qui interagissent

22. On pourrait utiliser les estimations obtenues sans restrictions d'exclusion comme repère, pour évaluer, à l'aide de tests de spécification, l'à-propos de ces restrictions.

23. Voir Samuelson (1947 : 168) et Tobin et Houthakker (1950).

24. Une approche de ce type est utilisée par Lemieux (1989). L'estimation en deux étapes de la forme réduite est basée sur la méthode de généralisation d'Heckman (1976) qui prend correctement en compte le fait que le modèle de choix discret est le modèle probit structurel, présenté précédemment dans ce papier.

avec une variable indicatrice du régime 4. En utilisant la notation des équations (11c) et (11f), on peut écrire :

$$h_1 = D_3 \hat{y}_3 + D_4 \hat{y}_6 + \varepsilon \quad (A2)$$

où  $D_3$  et  $D_4$  sont des variables indicatrices du régime 3 (secteur informel seulement) et du régime 4 (les deux secteurs à la fois), respectivement, tandis que  $\hat{y}_3$  et  $\hat{y}_6$  sont des fonctions linéaires du vecteur  $x$  définies dans l'équation (11).  $D_4 x$  représente aussi un ensemble de variables, corrélées avec  $h_1$ , qui ne sont pas incluses dans la partie droite de l'équation (A1) et qui ne peuvent être corrélées avec le terme d'erreur de cette équation que pour des raisons d'autosélection. Le paramètre structurel  $\theta_1$  peut donc être estimé à l'aide de la méthode des variables instrumentales. Le paramètre  $\theta_0$  est identifié de manière similaire. Les estimations par MCO (moindres carrés ordinaires) et par 2MC (doubles moindres carrés) des équations du salaire horaire moyen dans les secteurs formel et informel sont reportées au tableau A1. Le vecteur  $x$  est composé d'une constante, d'une variable dichotomique représentant le sexe, des années d'éducation, de l'expérience potentielle et de l'expérience potentielle au carré; les variables instrumentales sont simplement  $D_4 x$ .

Les deux estimateurs (MCO et 2MC) conduisent à un paramètre de courbure dans le secteur formel,  $\theta_0$ , qui est négatif, suggérant ainsi que la fonction de gain est convexe. De même, les deux estimations du paramètre de courbure dans le secteur informel,  $\theta_1$ , sont positives, suggérant que la fonction de gain est concave. De plus, les estimations par 2MC sont plus grandes, en valeur absolue, que les estimations par MCO et statistiquement différentes de zéro.

TABLEAU A1

ESTIMÉS DES FONCTIONS DE GAINS MOYENS<sup>a</sup>

Variable dépendante :	Gains moyens dans le secteur formel ( $R_0$ )		Gains moyens dans le secteur informel ( $R_1$ )	
	MCO (1)	2MC <sup>a</sup> (2)	MCO (3)	2MC <sup>a</sup> (4)
Heures de travail divisées par 1000	0,3868 (0,0844)	2,2516 (0,5460)	-3,7724 (1,0980)	-7,4876 (3,4682)
Paramètre de courbure( $\theta_0$ ou $\theta_1$ )	-0,1934	-1,1258	1,8862	3,7438
Années d'éducation	0,0855 (0,0038)	0,0896 (0,0046)	-0,0105 (0,0201)	-0,0171 (0,0217)
Années d'expérience potentielle	0,0654 (0,0032)	0,0420 (0,0077)	0,0515 (0,0142)	0,0485 (0,0150)
Expérience potent, au carré (/100)	-0,0975 (0,0072)	-0,0542 (0,0151)	-0,1217 (0,0441)	-0,1106 (0,0470)
Sexe (femme =1)	-0,2058 (0,0241)	-0,0226 (0,0598)	-0,5091 (0,1014)	-0,5295 (0,1074)
Moyenne de la variable dépendante	2,4559	2,4559	1,6789	1,6789
R-carré ajusté	0,5018	0,3179	0,2596	0,1928
Écart-type de la régression	0,4032	0,4718	0,5668	0,5918
No d'observations	1 331	1 331	133	133

NOTE : a. Les variables instrumentales utilisées sont les variables explicatives (sauf les heures) multipliées par une variable binaire indiquant si l'individu se trouve dans le régime 4 (travaille dans les deux secteurs).

### Interprétation de $\psi$

L'identification du paramètre de courbure de la fonction d'utilité,  $\psi$ , est mieux comprise en terme du principe de Le Chatelier. Pour voir cela, notons que les fonctions d'offre de travail dans le régime 2 (secteur formel uniquement) et dans le régime 3 (secteur informel uniquement) sont des fonctions d'offre de travail

contrainte, tandis que les fonctions d'offre de travail dans le régime 4 (les deux secteurs à la fois) sont des fonctions d'offre de travail non contrainte. Appelons ces fonctions d'offre de travail  $h^c$  et  $h^u$  :

$$h^c = [h_0^2, h_1^3]' \text{ et } h^u = [h_0^4, h_1^4]'$$

où  $h_s^r$  représente les heures travaillées dans le secteur  $s = (0,1)$ , par une personne du régime  $r = (2,3,4)$ . En utilisant les équations (6), (7), (8a) et (8b), il est facile de montrer que

$$h^u = Ah^c$$

$$\text{où } A = \begin{bmatrix} 1 + c\psi^2 & -c\psi(\psi + \theta_1) \\ -c\psi(\psi + \theta_0) & 1 + c\psi^2 \end{bmatrix}$$

et où  $c = (\theta_0\theta_1 + \theta_0\psi + \theta_1\psi)^{-1}$ . Selon les conditions de second ordre du problème d'optimisation, les éléments de la diagonale devraient être plus grands que un. L'intuition pour ce résultat est que lorsqu'un facteur sous-jacent, tel que l'éducation, change, il affecte toujours les heures de travail non contraintes d'un secteur de manière plus importante que les heures contraintes de ce même secteur. Les éléments diagonaux de la matrice  $A$  devraient donc être plus grands que un.

De manière similaire, il peut être montré que les éléments hors de la diagonale dans la matrice  $A$  sont négatifs, puisque les heures de travail dans les deux secteurs sont de parfaits substituts. La matrice  $A$  est facilement estimée en deux étapes : premièrement on estime les formes réduites des équations d'offre de travail contrainte  $h^c = [h_0^2, h_1^3]$  et ensuite on calcule l'offre de travail non contrainte  $h^u = [h_0^4, h_1^4]$  à partir des valeurs prédites des équations de la première étape. Notons que cette approche diffère légèrement d'un double moindre carré puisque  $h^c$  et  $h^u$  sont observées dans deux échantillons différents de travailleurs. Une fois de plus, il existe un problème implicite d'autosélection dans cette procédure. Ce problème n'est pas discuté ici, mais il est traité correctement dans la procédure du maximum de vraisemblance de la section 3. Les résultats sont reportés au tableau B2. Tandis que les estimations des équations des heures de travail du secteur formel (première rangée de  $A$ , 1,20 et -1,73) sont en accord avec la théorie, les estimations des équations des heures de travail du secteur informel (seconde rangée de  $A$ , 0,10 et 0,19) ne le sont pas. Une explication potentielle de ce résultat est que l'équation trouvée à la première étape pour  $h_1^3$  est basée sur 49 observations uniquement, et a donc un faible pouvoir de prédiction<sup>25</sup>. Ce résultat peut être aussi expliqué par le fait que les conditions de second ordre ne sont pas satisfaites lorsque la version la moins contrainte du modèle est estimée (tableau 4, colonne 3).

25. Le R-carré est seulement de 0,062, tandis que l'équation de  $h_0^2$  ( $h_0$  dans le régime 2) est basée sur 1 246 observations et a un R-carré de 0,224.

TABLEAU A2

LIEN ENTRE LES HEURES CONTRAITES ET NON CONTRAITES

---

---

$$h^u = Ah^c$$

$$\text{où } h^u = [h_0^4, h_1^4]', h^c = [h_0^2, h_1^3]'$$

$$\text{et où } A = \begin{bmatrix} 1 + c\psi^2 & -c\psi(\psi + \theta_1) \\ -c\psi(\psi + \theta_0) & 1 + c\psi^2 \end{bmatrix}$$

$$\hat{A} = \begin{bmatrix} 1.1996 & -1.7343 \\ (0.2368) & (0.9398) \\ 0.0994 & 0.1906 \\ (0.0665) & (0.2640) \end{bmatrix}$$

---

## BIBLIOGRAPHIE

- BECKER, GARY (1965), «A Theory of the Allocation of Time», *Economic Journal*, 75 : 493-517.
- BENJAMIN, DWAYNE (1992), «Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models», *Econometrica*, 60 : 287-322.
- BROWN, JAMES, et HARVEY ROSEN (1982), «On Estimation of Structural Hedonic Price Models», *Econometrica*, 50 : 765-768.
- DiNARDO, JOHN, et THOMAS LEMIEUX. (1992), «Alcohol, Marijuana, and American Youth : the Unintended Consequences of Government Intervention», NBER Working Paper no. 4212.
- EPPLE, D. (1987), «Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products», *Journal of Political Economy*, 95 : 59-80.
- FORTIN, BERNARD, et PIERRE FRÉCHETTE (1986), «Enquête sur les incidences et les perceptions de la taxation dans la région de Québec», Banque de données, Université Laval.
- FORTIN, BERNARD, et PIERRE FRÉCHETTE (1987), «The Size and Determinants of the Underground Economy in Quebec», Polycopié, Université Laval.
- GRONAU, REUBEN (1977), «Leisure, Home Production, and Work – The Theory of the Allocation of Time Revisited», *Journal of Political Economy*, 85 : 1099-1123.
- HALL, ROBERT E. (1973), «Wages, Income and Hours of Work in the U.S. Labor Force», in CAIN et WATTS eds. *Income Maintenance and Labor Supply*. New York : Academic Press.
- HAUSMAN, JERRY (1980), «Labor Supply» in H. AARON et J. PECHMAN, eds. *How Taxes Affect Economic Behavior*. Washington D.C. : The Brookings Institution, 27-82.
- HECKMAN, JAMES (1976), «The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models», *Annals of Economic and Social Measurement* : 475-492.
- HECKMAN, JAMES, et BO HONORÉ (1990), «The Empirical Content of the Roy Model», *Econometrica*, 58 : 1121-1150
- HECKMAN, JAMES, et G. SEDLACEK (1990), «Self-Selection and the Distribution of Hourly Wages», *Journal of Labor Economics*, 8 : S329-S363.
- KAHN, SHULAMIT, et KEVIN LANG (1988), «Efficient Estimation of Structural Hedonic Systems». *International Economic Review* 29 : 157-166
- KILLINGSWORTH, MARK (1983), *Labor Supply*. Cambridge : Cambridge University Press.
- LACROIX, GUY, et BERNARD FORTIN (1992), «Utility-Based Estimation of Labour Supply Functions in the Regular and Irregular Sector», *Economic Journal* 102.

- LEE, LUNG-FEI (1978), «Unionism and Wage Rates : A Simultaneous Equations Models with Qualitative and Limited Dependent Variables». *International Economic Review* 19 : 415-433.
- LEMIEUX, THOMAS (1989), *An Econometric Analysis of Labor Supply in the Irregular Economy*, Thèse de Ph.D. non publiée, Princeton University.
- LEMIEUX, THOMAS, BERNARD FORTIN, et PIERRE FRÉCHETTE (1994), «The Effect of Taxes on Labor Supply in the Underground Economy», *American Economic Review*, 84 : 231-254.
- MAGNAC, THIERRY (1991), «Segmented or Competitive Labor Markets?», *Econometrica*, 59 : 165-188.
- MROZ, THOMAS (1987), «The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions», *Econometrica*, 55 : 765-799
- POLLACK, R.A., et M.L. WACHTER (1975), «The Relevance of the Household Production Function and its Implications for the Allocation of Time», *Journal of Political Economy*, 83 : 255-277.
- ROBERTSON, DONALD, et J. SYMONS (1990), «The Occupational Choice of British Children», *Economic Journal*, 100 : 828-841
- ROSEN, SHERWIN (1974), «Hedonic Prices and Implicit Markets : Product Differentiation in Pure Competition», *Journal of Political Economy*, 82 : 34-55
- ROSENZWEIG, MARK (1980), «Neoclassical Theory and the Optimizing Peasant : An Econometric Analysis of Market Family Labor Supply in a Developing Country», *Quarterly Journal of Economics*, 94 : 31-55.
- ROY, A.D. (1951), «Some Thoughts on the Distribution of Earnings», *Oxford Economics Papers*, 3 : 135-146.
- SAMUELSON, PAUL A. (1947), *Foundations of Economic Analysis*, Cambridge: Harvard University Press.
- TOBIN, JAMES, et H.S. HOUTHAKKER. (1950), «The Effects of Rationing on Demand Elasticities», *Review of Economic Studies*, 18 : 140-153.
- WILLIS, R., et S. ROSEN (1979), «Education and Self-Selection», *Journal of Political Economy*, 87 : 507-536.