

Chapitre 3

Fiscalité et offre individuelle de travail au noir

«Les individus n'aiment pas payer d'impôts, ils engagent une multitude d'actions pour réduire leur charge fiscale et, en de nombreuses occasions, ils y parviennent. (...) Pour sa part, le gouvernement prend des mesures pour inciter les individus à se conformer aux lois fiscales.»

-James Alm et Jorge Martinez-Vasquez (2000, p.1)-

3.1 Introduction¹

L'influence des modalités d'imposition sur l'offre de travail est largement reconnue. De nombreux développements théoriques et empiriques² ont, en effet, démontré le caractère désincitatif au travail du système fiscal. De ce fait, la fiscalité est susceptible d'avoir de multiples implications en termes d'offre de travail au noir.

En premier lieu, la complexité du système fiscal peut inciter l'individu à mettre en place des procédures légales d'exemptions ou d'allègement d'impôt (Schneider et Neck, 1993). Il rend alors plus profitables les efforts déployés par l'individu pour éviter l'imposition. Il favorise le travail déclaré au détriment du travail non déclaré puisque la réduction du fardeau fiscal rend la fraude (avec la probabilité d'être contrôlé et puni) moins attractive. Par conséquent, il renforce le phénomène d'évasion fiscale et réduit

¹Ce chapitre repose sur des travaux menés conjointement avec Bernard Fortin et Guy Lacroix.

²Se reporter à Blundell (1992) pour une revue de la littérature à ce sujet.

le travail au noir. De nombreuses études se sont penchées sur ce problème.³ De leur point de vue, l'élargissement de la base fiscale et la suppression des exemptions devrait augmenter la taille de l'économie souterraine. À l'inverse, une diminution des impôts directs devrait réduire son ampleur. Toutefois, l'analyse ne porte pas véritablement sur l'offre de travail au noir. Le choix de l'activité et de son volume horaire ne sont pas modélisés. Ils ne sont que la conséquence indirecte du problème de «*tax avoidance*» qui caractérise la volonté pour l'individu de ne pas se soumettre aux réglementations fiscales en matière de déclaration de revenus d'activité.

En second lieu, le système fiscal peut induire une dissimulation des revenus d'activité ou une dissimulation de l'activité elle-même. En effet, la théorie standard de la fraude fiscale avec revenus endogènes montre que, sous certaines conditions, le système fiscal peut favoriser la fraude et entraîner une participation accrue à l'économie souterraine (*e.g.* Cowell, 1985). L'individu peut être incité à travailler sur un marché souterrain de façon à réduire son fardeau fiscal, ce faisant il peut également conserver ses prestations sociales. Les travaux de Lacroix et Fortin (1992) ont confirmé l'impact positif du taux d'imposition effectif des revenus de travail sur l'offre de travail au noir. Ces auteurs ont également souligné l'impact négatif de la probabilité de détection et du taux d'amende sur cette décision. Un moyen de réduire la fraude consisterait donc en une baisse des taux d'imposition. Un argument avancé en défaveur de cette politique est qu'elle entraîne une diminution du rôle du système fiscal dans la redistribution des revenus. Une autre solution préconisée est d'augmenter la fréquence des contrôles et le montant de l'amende. Bien évidemment l'équilibre approprié entre la probabilité de détection et le taux de pénalité soulève des problèmes additionnels à la fois d'équité et d'efficacité (Sandmo, 1981).

La fiscalité apparaît, dès lors, comme un des facteurs déterminants de l'offre de travail au noir, qu'il s'agisse du volume de l'activité ou du choix de l'activité elle-même. En ce sens, elle mérite une attention toute particulière. Elle soulève toutefois deux problèmes sensiblement différents. Il s'agit d'une part, de l'évasion fiscale et d'autre part, de la fraude fiscale. Ces deux aspects génèrent des approches théoriques distinctes et pour lesquelles la participation à l'économie souterraine est abordée différemment. Dans le premier cas, il s'agit essentiellement de déterminer le niveau d'imposition optimal, en termes d'efficacité et d'équité, afin d'éviter tout comportement déviant. Dans le second cas, le problème soulevé est davantage celui de l'impact de la fiscalité sur l'allocation du temps de travail entre un marché officiel et un marché non officiel. C'est pourquoi, nous nous inscrivons pleinement dans la seconde perspective.

³Dans un article paru en 1998 au *Journal of Economic Literature*, Andréoni *et al.* proposent une recension des contributions sur l'évasion fiscale.

L'objectif de ce chapitre est d'élargir les travaux réalisés jusqu'alors. L'offre de travail sur les marchés officiel et souterrain est analysée simultanément à partir d'un modèle structurel fondé sur la maximisation de l'espérance de l'utilité. Le modèle tient compte des solutions de coin sur les deux marchés et de l'endogénéité des taux de rémunération bruts. Par ailleurs, l'estimation du modèle est réalisée par la méthode du maximum de vraisemblance à information complète, ce qui permet d'obtenir des estimateurs asymptotiquement convergents de la matrice de variance-covariance des coefficients estimés du modèle. Cela permet, en outre, d'effectuer des tests d'hypothèse non biaisés sur les paramètres du modèle.

Enfin, notre approche économétrique permet d'endogénéiser les variables subjectives de probabilité et d'amende. Ces deux variables peuvent être sujettes à un biais potentiellement important de dissonance cognitive. En effet, les travailleurs au noir sont susceptibles de sous-estimer les probabilités de sanctions de la fraude afin de réduire leur sentiment d'insécurité. Le risque de détection et l'amende associée sont dès lors endogènes à la participation au marché noir et seront traitées comme telles dans la procédure économétrique développée dans ce chapitre.

La structure du chapitre est la suivante. A la *section 3.2*, nous proposons une brève revue de la littérature consacrée au lien entre fiscalité et participation au marché souterrain. Nous exposons à la *section 3.3* un modèle microéconométrique d'offre de travail sur les marchés officiel et/ou souterrain. Nous présentons successivement le modèle théorique et la spécification économétrique retenue. La procédure d'estimation est développée à la *section 3.4*. Nous commentons les principaux résultats à la *section 3.5* et nous traitons spécifiquement de l'impact de la fiscalité dans les décisions individuelles et de la substituabilité des heures de travail sur chacun des deux marchés.

3.2 Littérature relative au lien entre fiscalité et travail au noir

L'intérêt sans cesse grandissant à l'égard de l'économie souterraine a donné lieu à une vaste littérature sur le sujet. Depuis les articles fondateurs de Allingham et Sandmo (1972) puis de Yitzhaki (1974) notamment, de nombreux développements sont apparus dans l'appréhension de l'influence de la fiscalité sur la participation à l'économie souterraine.

L'objet de cette section est de dresser un état de lieux de la littérature. Il n'est pas possible d'évoquer la question du travail au noir sans s'appuyer sur une théorie de la fraude fiscale et de ses déterminants. Il faut, en effet, que les paramètres fiscaux

soient clairement présents à l'esprit pour pouvoir prétendre enrichir la compréhension des interactions entre la fiscalité et l'offre de travail au noir.

Ainsi, de façon quelque peu schématique, nous proposons une dichotomie des contributions théoriques qui s'intéressent à la fraude fiscale selon qu'il s'agisse d'une sous-déclaration de revenus ou d'une dissimulation d'activité. Dans les paragraphes qui suivent, nous explicitons les modèles de fraude fiscale pour lesquels les revenus sont successivement exogènes puis endogènes. Nous poursuivons l'analyse en proposant un modèle de fraude fiscale avec revenus endogènes qui tient compte, à la fois, de la non linéarité du système fiscal et des programmes sociaux.

3.2.1 Modèles de fraude fiscale avec revenus exogènes

L'analyse théorique de la relation entre fiscalité et décision à risque a été formulée en termes d'effet du système fiscal sur le choix de portefeuille des individus (Mossin, 1968b; Stiglitz, 1969). La décision de participer au marché noir est traditionnellement assimilée à un choix risqué. Elle est appréhendée par l'addition d'un bien risqué au portefeuille de l'individu. L'économie du crime (Becker, 1968) et les modèles standard de choix de portefeuille en environnement incertain (*e.g.* Mossin, 1968a; Arrow, 1970) constituent les fondements de cette approche. Dans cette perspective, les travaux de Allingham et Sandmo (1972) ont véritablement posé le problème de la fraude fiscale. Ils ont notamment ouvert le large débat de l'impact des taux d'imposition sur la décision de fraude.

Leurs travaux reposent sur la modélisation du choix auquel est confronté un individu qui doit déclarer ses revenus aux autorités fiscales. En premier lieu, l'agent décide s'il fraude ou non. En second lieu, si la décision de frauder est retenue, il détermine quelle proportion de son revenu ne sera pas déclarée. Un modèle d'utilité espérée associée à la déclaration de revenus aux autorités fiscales est alors développé afin de déterminer si un taux d'imposition plus faible conduit l'individu à déclarer la totalité de ses revenus.

Considérons un contribuable avec des revenus exogènes W dont le montant réel n'est connu que de lui seul. Il fait face à un taux d'imposition τ proportionnel aux revenus imposables. Il peut néanmoins dissimuler tout ou partie de ses revenus aux autorités fiscales. Il déclare un niveau de revenu X , où $X \leq W$. En choisissant de frauder, il réduit son fardeau fiscal. Mais, il a une probabilité p (indépendante de X) d'être contrôlé. Dans ce cas, il est supposé que les autorités fiscales découvrent le véritable montant des revenus de l'individu. Le contribuable fraudeur devra alors

s'acquitter d'une amende au taux de pénalité θ (avec $\theta > \tau$) sur les revenus dissimulés ($W - X$).

Nous sommes ici typiquement en présence d'un choix de portefeuille en environnement incertain. L'incertitude provient du risque de détection de la fraude par les autorités fiscales. Le contribuable est supposé se conformer aux axiomes de von Neumann-Morgenstern de comportement sous incertitude. L'unique argument de la fonction d'utilité espérée est le revenu net, de sorte que celle-ci s'écrit :

$$EU = (1 - p)U(W - \tau X) + pU(W - \tau X - \theta[W - X])$$

L'utilité marginale est supposée positive et strictement décroissante en tout point, ce qui implique que l'individu est averse au risque.

Sur la base de ce modèle théorique de fraude fiscale, une augmentation de la probabilité de détection p ou du taux de pénalité θ pousse l'individu à déclarer davantage de revenus. En revanche, l'impact d'une hausse du taux marginal d'imposition τ sur la part du revenu déclaré est ambigu. En effet, ce modèle met en évidence deux effets opposés : un effet de substitution et un effet de revenu. Le premier incite l'agent à frauder davantage alors que le second agit en sens opposé.⁴ L'effet de substitution est positif dans la mesure où une augmentation du taux d'imposition rend la fraude plus profitable à la marge. L'effet revenu est négatif puisque, quel que soit l'état de la nature, une hausse du taux d'imposition diminue le niveau de richesse de l'individu. Par conséquent, sous l'hypothèse d'aversion au risque absolue décroissante avec le revenu, l'individu tend à réduire le montant de la fraude.

Le modèle ne permet donc pas de déterminer a priori l'effet du taux marginal d'imposition sur la déclaration de revenus de l'individu. Certains aménagements ont alors été proposés. Ainsi, Yitzhaki (1974) poursuit l'analyse en appliquant le redressement fiscal sur l'impôt fraudé et non plus sur les revenus fraudés. La fonction d'utilité espérée s'écrit désormais :

$$EU = (1 - p)U(W - \tau X) + pU(W - \tau X - F\tau[W - X])$$

où F désigne l'amende ($F > 1$).⁵

Le modèle prédit que l'effet revenu domine et par conséquent, qu'une hausse du taux d'imposition a pour effet de dissuader la fraude fiscale. Cette conclusion est quelque peu surprenante. Elle implique qu'il suffirait aux gouvernements d'imposer

⁴Ce résultat est obtenu sous l'hypothèse d'aversion absolue pour le risque décroissant.

⁵Notons que si $\theta = F\tau$, les conditions sont identiques à celle de Allingham et Sandmo (1972).

des taux d'imposition suffisamment élevés pour réduire la fraude. Or, en pratique, ces taux ne peuvent être très élevés, que ce soit pour limiter le risque de faillite, ou pour répondre à des considérations d'équité ou de morale.⁶ Par ailleurs, la relation négative entre le taux d'imposition et la fraude, obtenue par Yitzhaki (1974), est contraire à ce qui est observé empiriquement (Clotfelter, 1983; Poterba, 1987).

Ce résultat contre-intuitif a poussé les économistes à démontrer théoriquement que la relation entre le taux d'imposition et la fraude fiscale était positive. Plusieurs auteurs, dont Gordon (1989) ou Myles et Naylor (1996), ont imposé davantage de restrictions sur les préférences afin d'obtenir des conditions sous lesquelles une hausse de l'imposition incite à la fraude fiscale. Cela a certes permis de lever une ambiguïté, mais cela n'explique pas la présence aussi importante des solutions de coin.⁷

Nous saisissons ici l'intérêt de prolonger l'analyse par l'introduction d'un système fiscal plus élaboré. Il importe, en effet, de rendre compte de la non linéarité de la contrainte budgétaire à laquelle les individus sont confrontés en raison des nombreuses tranches d'imposition. Cela implique que le salaire horaire net d'impôt n'est pas constant lorsque le nombre d'heures de travail varie. Pour évaluer les élasticités de l'offre aux modifications du barème fiscal, il faut déterminer dans quelle mesure l'offre de travail s'ajuste aux variations de salaires résultant des changements de l'imposition. Pour cela, un taux d'impôt marginal doit être déterminé pour chaque individu. Cette différenciation doit permettre de juger de l'impact d'une variation du taux d'imposition, propre à l'individu, sur son offre de travail au noir. Un environnement fiscal suffisamment sévère pourrait inciter certains individus à modifier leur comportement, dans le but de diminuer leur taux marginal d'imposition. Dans ce cas, le travail au noir pourrait s'avérer une alternative intéressante. Nous entreprenons une telle tâche dans la *section 3.3* de ce chapitre et précisons ainsi la nature de la relation entre fiscalité, offre de travail sur le marché déclaré et offre de travail au noir.

Par ailleurs, les modèles de fraude fiscale, présentés ici, se heurtent à des hypothèses très restrictives qui limitent la portée de l'analyse. En particulier, les revenus sont supposés exogènes. Or, l'on doit s'attendre à ce que les individus adaptent leur niveau de revenus -et par conséquent de leur activité- afin de réduire leur fardeau fiscal. Ils décident donc simultanément de l'intensité de leur offre de travail et de leurs revenus.

L'hypothèse d'exogénéité des revenus, très peu réaliste, a incité de nombreux auteurs à proposer des modèles de fraude où les revenus sont endogénéisés (*e.g.* Pencavel,

⁶Se reporter à Gary S. Becker (1968). Voir également Andréoni (1991), Shavell (1987) et Mookherjee (1989, 1992) pour une discussion des contraintes qui, à l'optimum, bornent l'ampleur des pénalités.

⁷Par solution de coin, nous entendons le cas où un individu ne travaille pas sur le marché officiel et/ou le marché noir.

1979; Sandmo, 1981; Cowell, 1985; Fluet, 1987). Cela s'est traduit par l'introduction de la décision d'offre de travail dans les modèles. L'objet de la section suivante est précisément de développer les modèles de fraude fiscale avec revenus endogènes.

3.2.2 Modèles de fraude avec revenus endogènes

Pencavel (1979), le premier, souligne l'importance du processus de création de richesses dans l'analyse du comportement de fraude. Les modèles de fraude fiscale avec revenus exogènes sont, certes, très utiles pour l'étude de certains types de fraude. Mais, ils ne rendent compte ni du choix de l'activité, ni de l'intensité de celle-ci. Ils sont donc inappropriés au problème du travail au noir (Cowell, 1985). Ils permettent d'examiner la décision d'un individu de dissimuler une partie de son revenu, mais ce revenu est indissocié de sa source. Or, la fraude que nous souhaitons étudier s'apparente davantage à l'exercice d'une activité non déclarée. Celle-ci peut se substituer à une activité déclarée ou être cumulée.

La possibilité de formes diverses de fraude fiscale doit alors être introduite dans les modèles précédents. Le problème soulevé est quelque peu différent dans la mesure où il s'agit d'analyser l'impact des paramètres fiscaux sur l'incitation à travailler sur les marchés officiel et/ou souterrain. Les opportunités de fraude pouvant différer selon les emplois, elles peuvent influencer les choix d'activités (Pestieau *et al.*, 1991).⁸ Les modèles sont enrichis par la décision d'offre de travail de l'individu, de sorte que les revenus sont endogénéisés. Néanmoins, l'introduction de l'offre de travail dans les modèles de fraude fiscale complique singulièrement l'analyse.

Dans cette approche, les heures offertes dépendent des taux de rémunération sur les marchés officiel et souterrain, du revenu hors-travail, des paramètres de la fiscalité et des programmes sociaux, de la probabilité d'être contrôlé par les autorités gouvernementales et du taux de pénalité en cas de détection de la fraude. Dans ce contexte, les effets des paramètres fiscaux deviennent tous ambigus. Une augmentation du taux d'imposition réduit le taux de salaire effectif de l'individu, ce qui devrait diminuer son offre de travail officiel. Toutefois, si la courbe d'offre est à rebroussement, un niveau d'imposition plus élevé peut accroître l'offre de travail officiel et le montant des revenus non déclarés. Le degré d'aversion au risque peut également interagir avec l'offre de travail de façon inattendue. Les incitations à la fraude peuvent encourager l'individu à travailler davantage afin de s'assurer contre les pertes éventuelles dues aux contrôles et pénalités à payer (Weiss, 1976). Cela se traduirait alors par une augmentation de la

⁸ Il est intéressant de noter que certains effets incitatifs de la fraude peuvent conduire à des situations plus optimales au sens de Pareto (Weiss, 1976; Stiglitz, 1982). Cela apporte une justification à la présence de tels effets comme partie intégrante d'un système fiscal optimal.

fraude. Mais, une fois encore, la complexité du système fiscal est généralement sous-estimée, de sorte que l'on ne tient pas compte des nombreuses tranches d'imposition auxquelles le contribuable fait face. De même, le système de redistribution est généralement absent alors que l'on peut supposer un impact important sur le choix de l'activité.

L'analyse économétrique de l'offre de travail au noir pose, en outre, plusieurs questions très délicates à résoudre. Au-delà de la difficulté de récolte de données fiables, certaines variables telles que la probabilité de contrôle et le taux de pénalité sont difficilement mesurables. Pour cette raison, la presque totalité des études économétriques (*e.g.* Clotfelter, 1983 ; Lemieux, Fortin et Fréchette, 1994 ; Graversen et Smith, 2001) ignore ces variables ou les remplace par une forme réduite en fonction de variables observables. En outre, il faut s'attendre à ce que l'individu réagisse non pas aux valeurs objectives de la probabilité de contrôle et du taux de pénalité, mais plutôt à la perception subjective qu'il en a. En effet, il est probablement influencé par son appréciation des moyens mis en œuvre par les services fiscaux pour détecter la fraude et par son sentiment quant au caractère identifiable de son propre comportement (Andréoni, Erard et Feinstein 1998). Par ailleurs, en raison notamment de l'opprobre social associé à un comportement de fraude, il importe de tenir compte de la substituabilité imparfaite entre les heures de travail officielles et non officielles dans les préférences du travailleur. Enfin, le modèle économétrique doit tenir compte de plusieurs régimes (quatre) associés aux décisions de travailler ou non sur les marchés officiel et au noir. La présence de ces quatre régimes rend la fonction de vraisemblance à maximiser particulièrement complexe.

Un certain nombre d'études microéconométriques ont tenté d'estimer l'impact des politiques fiscales sur les heures de travail non déclarées. Lemieux, Fortin et Fréchette (1994) ont estimé un modèle d'allocation du temps entre les marchés du travail officiel et souterrain. Leur modèle est inspiré de celui de Gronau (1977). Ils concluent que l'impôt sur le revenu détourne les heures travaillées du marché officiel vers le marché noir. Cette distorsion est particulièrement importante pour certaines catégories de la population telles que les bénéficiaires de minima sociaux (Fortin, Lacroix et Montmarquette, 2000). Lacroix et Fortin (1992) ont estimé un système d'équations simultanées d'offre de travail sur les marchés officiel et souterrain, basé sur un modèle de fraude fiscale (*e.g.* Sandmo, 1981 ; Cowell, 1985). Leur modèle tient compte des solutions de coin sur les deux marchés et de l'endogénéité des taux de rémunération bruts. Les données sont obtenues à partir d'une enquête réalisée en 1986 sur 2 134 personnes de la ville de Québec (Canada). Ces données incluent des questions sur la probabilité subjective d'être contrôlé et le taux de pénalité en cas de fraude. Leurs résultats

confirment l'impact positif du taux de taxation effectif des revenus de travail sur l'offre de travail au noir ainsi que l'impact négatif de la probabilité de détection et du taux d'amende sur cette décision.

Le présent chapitre a pour but d'élargir cette dernière étude dans plusieurs directions. En premier lieu, l'exploitation d'une base de données semblable à celle utilisée par ces auteurs, mais plus récente et étendue à trois régions du Québec (Montréal, Québec, Bas-du-Fleuve) et comprenant près du double d'individus, devrait permettre d'améliorer la fiabilité des résultats. En second lieu, les données incluent une variable subjective indiquant la proportion des personnes travaillant sur le marché noir dans l'entourage de l'individu. Cette variable sera utilisée pour évaluer l'importance des effets de voisinage dans les décisions de travailler de l'individu. En troisième lieu, Lacroix et Fortin utilisent une méthode en deux étapes pour estimer leur modèle. Ils estiment d'abord les taux de salaire sur chacun des deux marchés (avec correction pour biais de sélection) et ils estiment ensuite les fonctions d'offre de travail en tenant compte des quatre régimes après avoir remplacé les taux de salaire par les taux de salaire prédits dans la première étape. Or, cette méthode soulève certaines difficultés en raison de la non linéarité de ces fonctions. Dans ce cas, les coefficients estimés ne sont, en général, pas convergents (Wooldridge, 2001). Dans ce chapitre, nous estimons le modèle à partir d'une méthode de maximum de vraisemblance à information complète, ce qui permet d'obtenir des estimateurs plus précis et convergents des paramètres de préférence et de leur écart-type. Cela permet également d'effectuer des tests d'hypothèses non biaisés sur les paramètres du modèle.

Enfin, notre approche économétrique permet d'endogénéiser les variables subjectives de probabilité et d'amende. De nombreux travaux empiriques confirment, en effet, l'endogénéité de ces deux variables (*e.g.* Beron, Tauchen et Witte, 1992; Durbin, Graetz et Wilde, 1990). Par conséquent, des biais importants sont susceptibles de se produire lorsqu'elles sont supposées exogènes. Selon la théorie de la dissonance cognitive empruntée à Akerlof et Dickens (1982), les individus ont des préférences quant à l'état de la nature et quant à leur croyance en l'état de la nature. Cette approche nous enseigne, en outre, qu'ils peuvent exercer un certain contrôle sur leurs croyances. Ils peuvent manipuler leurs croyances en sélectionnant les sources d'information qui confirment leurs désirs. Tout comme les travailleurs d'emplois dangereux sous-estiment généralement le danger inhérent à leur activité, les participants à l'économie souterraine peuvent se convaincre du caractère peu risqué de l'activité souterraine. Ils peuvent réduire le sentiment d'insécurité en rapportant de faibles probabilités de sanctions. Ainsi, le risque de détection et l'amende associée sont endogènes à la participation au marché noir et doivent être traitées comme telles. Or, toutes les études

ayant recours à de telles variables, dans leur analyse de l'offre de travail au noir, ont supposé qu'elles étaient exogènes (*e.g.* Isachsen et Strom, 1980 ; Isachsen, Klovland et Strom, 1982 ; Lacroix et Fortin, 1992). La procédure économétrique que nous développons dans ce chapitre tient compte, pour la première fois, de l'endogénéité de ces variables.

Le modèle prend la forme d'un système de six équations simultanées avec possibilité de solutions de coin sur l'un et/ou l'autre des marchés. Les préférences sont représentées par une fonction d'utilité quadratique. Les variables endogènes comprennent les heures de travail sur chacun des deux marchés, le revenu «virtuel» hors-travail, le salaire net sur le marché officiel, le salaire espéré sur le marché noir et une mesure du risque lié au marché noir. La technique économétrique traite de la non linéarité de la contrainte budgétaire associée à la présence de la fiscalité et des programmes sociaux et tient compte de l'endogénéité des taux marginaux effectifs de taxation.

3.3 Un modèle d'offre de travail sur les marchés officiel et souterrain

Notre approche se fonde sur les modèles de fraude fiscale avec revenus endogènes (Cowell, 1990). Sur le plan économétrique, il généralise les travaux de Lacroix et Fortin (1992).

3.3.1 Le modèle théorique

Nous considérons que les heures de travail sur le marché noir apportent une perte d'utilité à l'individu différente de celle découlant des heures de travail sur le marché officiel. Or, de nombreuses études imposent la substitution parfaite entre les heures de travail sur chacun des deux marchés. Dans ce cas, la désutilité marginale des heures de travail sur les deux marchés est supposée la même. La fonction d'utilité s'écrit $U = U(h_1 + h_2, C)$ où h_1 et h_2 représentent les heures de travail sur les marchés officiel et souterrain respectivement et où C représente la consommation. Cependant, certains auteurs (*e.g.* Hansson 1985, Baldry 1986, Lacroix 1990) soutiennent que tel n'est pas le cas. En effet, alors que le plaisir de frauder pourrait réduire, pour certains individus, la désutilité marginale de h_2 par rapport à h_1 , l'immoralité associée au travail au noir et l'opprobre social manifesté à son égard peuvent contribuer à accroître la désutilité marginale de h_2 relativement à h_1 . Par conséquent, nous avons retenu une fonction d'utilité plus générale donnée par :

$$U = U(h_1, h_2, C) \quad (3.1)$$

Cette fonction n'impose pas la substitution parfaite entre les deux types d'heures de travail. Elle est, de plus, supposée strictement concave en ses arguments, croissante en C et décroissante en h_1 et h_2 . La concavité provient du fait que l'individu est averse au risque.

Par ailleurs, la nature dissimulée de l'activité au noir implique une limitation de son ampleur. Une fonction de gains non déclarés de type Cobb-Douglas et reliant ceux-ci aux heures non déclarées pourrait être envisagée, tandis que les rémunérations du travail déclaré seraient modélisées par une fonction linéaire (Lemieux *et al.*, 1994). Néanmoins, afin de simplifier notre approche, nous supposons une même forme fonctionnelle linéaire aux gains issus du marché officiel et du marché souterrain.

Le travail au noir est rémunéré au taux W_2 et permet à l'individu d'échapper à l'imposition. Ce dernier encourt, toutefois, le risque d'être contrôlé par les autorités fiscales -ce qui survient avec une probabilité p . Il subit alors un redressement au taux de pénalité θ , avec $\theta > \tau$ où τ est le taux d'impôt de l'individu supposé, pour le moment, proportionnel. A l'instar de Yitzhaki (1974), nous postulons que la pénalité s'applique à l'impôt fraudé et non pas aux revenus dissimulés. Par conséquent, on a $\theta = F\tau = (1 + \lambda)\tau$ où λ est le taux de pénalité sur ce montant d'impôt fraudé. L'incertitude liée à la détection sur le marché noir et l'amende qui y est rattachée rendent la contrainte budgétaire aléatoire, de sorte que celle-ci s'écrit :

$$C \geq y + w_1 h_1 + w_2 h_2, \quad (3.2)$$

où y est le revenu non salarial et w_1 est le salaire officiel net, avec $w_1 = W_1(1 - \tau)$; W_1 étant le salaire brut. Enfin, w_2 est la rémunération nette l'activité souterraine. Comme cette dernière variable est aléatoire, on a :

$$w_2 = \left\{ \begin{array}{ll} W_2 & \text{avec probabilité } (1 - p) \\ W_2(1 - \theta) & \text{avec probabilité } p \end{array} \right\} \quad (3.3)$$

où W_2 est le salaire brut sur le marché noir.

L'individu est supposé maximiser son espérance d'utilité sous cette contrainte, en choisissant ses heures de travail sur les deux marchés. Il doit, à la fois, déterminer sa prestation totale de travail, puis décomposer celle-ci en une prestation sur chacun des marchés. Cette dernière décision est analogue à un choix de portefeuille, dans la mesure où un compromis doit être effectué entre un rendement incertain sur le marché noir et un rendement certain sur le marché officiel. Néanmoins, nous admettons la possibilité pour l'individu de cumuler un emploi officiel et un emploi au noir. Il peut exercer un certain nombre d'heures sur le marché officiel et décider de participer ou non à une activité souterraine.

Une contrainte de non négativité est imposée sur les heures de travail dans chacun des marchés, de sorte que : $h_1 \geq 0$ et $h_2 \geq 0$.

Sous l'hypothèse que les axiomes de von Neuman-Morgenstern sous incertitude sont satisfaits, le programme à résoudre est alors le suivant :

$$\max EU(h_1, h_2, C) \quad (3.4)$$

sous les contraintes de budget et de non négativité sur h_1 et h_2 . E est l'opérateur d'espérance mathématique.

Dans la mesure où la contrainte budgétaire (3.2) est respectée avec égalité stricte (en l'absence de saturation des préférences), il est possible de la substituer dans (3.4) et d'exprimer la fonction d'utilité espérée en termes de h_1 et h_2 et des variables exogènes. Les utilités marginales nettes des heures de travail sont obtenues en dérivant partiellement la fonction d'utilité espérée obtenue par rapport à h_1 et h_2 :

$$m_1(h_1, h_2, z) = EU_1 + w_1 EU_3 \quad (3.5)$$

$$m_2(h_1, h_2, z) = EU_2 + E(U_3 W_2) \quad (3.6)$$

où m_1 et m_2 représentent respectivement les utilités marginales nettes espérées de l'activité déclarée et de l'activité non déclarée, et où $z = (W_1(1 - \tau), W_2, p, \theta, y)$ est le vecteur de variables exogènes. Les termes EU_1 et EU_2 représentent respectivement l'effet d'une hausse de h_1 et h_2 sur l'espérance de l'utilité. Par conséquent, $EU_i < 0$, avec $i = 1, 2$.

Les équations (3.5) et (3.6) nous permettent de caractériser les conditions d'optimalité du travail propres à chacun des quatre régimes sur les marchés officiel et souterrain. Plus précisément, à l'optimum, les conditions de Kuhn-Tucker de premier ordre associées au programme (3.4) peuvent s'écrire de la façon suivante :

$$m_1(h_1, h_2, z) \leq 0 \quad (3.7)$$

$$m_2(h_1, h_2, z) \leq 0 \quad (3.8)$$

$$h_1[m_1(h_1, h_2, z)] = 0 \quad (3.9)$$

$$h_2[m_2(h_1, h_2, z)] = 0 \quad (3.10)$$

Les conditions de second ordre sont :

$$m_{11}(h_1, h_2, z) \leq 0 \quad (3.11)$$

$$m_{22}(h_1, h_2, z) \leq 0 \quad (3.12)$$

$$m_{11}m_{22} - (m_{12})^2 \geq 0 \quad (3.13)$$

Ces conditions (3.7) à (3.10) définissent quatre régimes selon que les contraintes de non négativité sont serrées ou non.

Le **régime 1** caractérise le cas où l'individu exerce une activité sur les deux marchés du travail, *i.e.* $h_1 > 0$ et $h_2 > 0$ (solutions intérieures). Les conditions impliquent que les équations (3.5) et (3.6) sont satisfaites avec égalité stricte :

$$m_1(h_1, h_2, z) = 0 \quad (3.14)$$

$$m_2(h_1, h_2, z) = 0 \quad (3.15)$$

En résolvant ces dernières équations simultanément, nous obtenons les fonctions d'offre de travail non contraintes : $h_1 = h_1(h_2, z)$ et $h_2 = h_2(h_1, z)$.

Dans le **régime 2**, l'individu travaille uniquement sur le marché officiel, *i.e.* $h_1 > 0$ et $h_2 = 0$ (solution intérieure en h_1 et solution de coin en h_2). Par conséquent, nous avons :

$$m_1(h_1, 0, z) = 0 \quad (3.16)$$

$$m_2(h_1, 0, z) \leq 0 \quad (3.17)$$

L'équation (3.16) nous permet de déduire la fonction d'offre de travail officiel contrainte $h_1 = \bar{h}_1(z)$.

Dans le **régime 3**, l'individu travaille uniquement sur le marché noir, *i.e.* $h_1 = 0$ et $h_2 > 0$ (solution intérieure en h_2 et solution de coin en h_1). Dans ce cas, on a :

$$m_1(0, h_2, z) \leq 0 \quad (3.18)$$

$$m_2(0, h_2, z) = 0 \quad (3.19)$$

En résolvant (3.19) pour h_2 , nous obtenons la fonction d'offre de travail au noir contrainte $h_2 = \bar{h}_2(z)$.

Enfin, le **régime 4** fait référence aux individus qui n'exercent aucune activité de travail, *i.e.* $h_1 = 0$ et $h_2 = 0$. Il n'y a aucune solution intérieure. Les inéquations (3.7) et (3.8) deviennent alors :

$$m_1(0, 0, z) \leq 0 \quad (3.20)$$

$$m_2(0, 0, z) \leq 0 \quad (3.21)$$

Le modèle théorique nous permet d'identifier les différentes variables exogènes influençant l'offre de travail sur les deux marchés. Néanmoins, il ne nous permet pas de déterminer quantitativement les effets de ces variables sur l'offre de travail. La stratégie économétrique consiste alors à imposer une forme fonctionnelle à la fonction d'utilité, puis d'estimer les paramètres qui la définissent.

3.3.2 Spécification économétrique

Nous spécifions, en premier lieu, une forme fonctionnelle à la fonction d'utilité et aux équations de salaire et de revenu virtuel hors-travail. Les caractéristiques socio-démographiques et l'hétérogénéité non observable entre individus sont ensuite introduites dans le modèle par des composantes aléatoires.

Une fonction d'utilité quadratique

La spécification d'une forme fonctionnelle appropriée se heurte à la volonté de rendre compte de la réalité, le plus fidèlement possible, tout en répondant à des conditions de faisabilité économétrique. Stern (1986) et Ransom (1987a) ont montré qu'une fonction d'utilité quadratique est suffisamment flexible pour être considérée comme une approximation de second ordre à toute fonction d'utilité. En outre, les fonctions d'utilité marginales qui en découlent sont linéaires dans les paramètres, ce qui simplifie considérablement les estimations.

Afin d'estimer notre modèle et suivant en cela Lacroix et Fortin (1992), nous retenons la forme fonctionnelle suivante :

$$U(x) = \alpha'x + \frac{1}{2}x'\beta x \quad (3.22)$$

où $x = (h_1, h_2, C)'$, α est un vecteur de paramètres de dimension 3×1 et β est une matrice de paramètres 3×3 .

U est strictement concave si β est définie négative et symétrique. Sous une forme vectorielle, les fonctions d'utilité marginales, linéaires en x , sont égales à $\alpha + \beta x$. Elles ne sont pas monotones, en général, de sorte que la consommation peut affecter négativement l'utilité pour certaines valeurs de x et que les heures de travail (h_1 et h_2) peuvent l'affecter positivement. L'espérance de l'utilité en fonction de h_1 , h_2 et z est obtenue en utilisant (3.22) et en remplaçant la consommation C par son expression dans chaque état de la nature :

$$\begin{aligned} EU(x) = & \alpha_1 h_1 + \alpha_2 h_2 + \alpha_3 (y + w_1 h_1 + h_2 Ew_2) + \frac{1}{2} \beta_{11} h_1^2 \\ & + \beta_{12} h_1 h_2 + \beta_{13} h_1 (y + w_1 h_1 + h_2 Ew_2) \\ & + \frac{1}{2} \beta_{22} h_2^2 + \beta_{23} h_2 (y + w_1 h_1 + h_2 Ew_2) \\ & + \frac{1}{2} \beta_{33} E(y + w_1 h_1 + w_2 h_2)^2 \end{aligned} \quad (3.23)$$

où $Ew_2 = W_2(1 - \theta p)$ et $Ew_2^2 = (1 - p)W_2^2 + pW_2^2(1 - \theta)^2 = W_2^2(1 + p\theta^2 - 2p\theta)$.

A partir de (3.23), on obtient :

$$\begin{aligned} m_1(h_1, h_2, z) = & \alpha_1 + \alpha_3 w_1 + \beta_{11} h_1 + \beta_{12} h_2 \\ & + \beta_{13}(y + 2w_1 h_1 + h_2 Ew_2) + \beta_{23} w_1 h_2 \\ & + \beta_{33} w_1 (y + w_1 h_1 + h_2 Ew_2) \end{aligned} \quad (3.24)$$

$$\begin{aligned} m_2(h_1, h_2, z) = & \alpha_2 + \alpha_3 Ew_2 + \beta_{12} h_1 + \beta_{13} h_1 Ew_2 \\ & + \beta_{22} h_2 + \beta_{23}(y + w_1 h_1 + 2h_2 Ew_2) \\ & + \beta_{33} [(y + w_1 h_1) Ew_2 + h_2 Ew_2^2] \end{aligned} \quad (3.25)$$

Les conditions locales de second ordre sont :

$$m_{11} = \beta_{11} + 2\beta_{13} w_1 + \beta_{33} w_1^2 \leq 0 \quad (3.26)$$

$$m_{22} = \beta_{22} + 2\beta_{23} Ew_2 + \beta_{33} Ew_2^2 \leq 0 \quad (3.27)$$

$$\begin{aligned} m_{11} m_{22} - (m_{12})^2 = & (\beta_{11} + 2\beta_{13} w_1 + \beta_{33} w_1^2) (\beta_{22} + 2\beta_{23} Ew_2 + \beta_{33} Ew_2^2) \\ & - (\beta_{12} + \beta_{13} Ew_2 + \beta_{23} w_1 + \beta_{33} w_1 Ew_2)^2 \geq 0 \end{aligned} \quad (3.28)$$

Ces conditions sont globalement satisfaites si la matrice β est définie négative (concavité stricte de $U(x)$). Les conditions globales de second ordre sont les suivantes :

$$\begin{aligned} \beta_{11} & < 0 \\ \beta_{11} \beta_{22} - \beta_{12}^2 & > 0 \end{aligned} \quad (3.29)$$

$$\beta_{11} \beta_{22} \beta_{33} - \beta_{11} \beta_{23}^2 - \beta_{12}^2 \beta_{33} + 2\beta_{12} \beta_{13} \beta_{23} - \beta_{13}^2 \beta_{22} < 0$$

Les fonctions d'offre de travail du régime 1 sont obtenues en égalisant (3.24) et (3.25) à zéro et en résolvant pour h_1 et h_2 (voir équations 3A.1 et 3A.2 de l'annexe 3A). De même, en posant respectivement $h_2 = 0$ dans (3.24) et $h_1 = 0$ dans (3.25), en égalisant (3.24) et (3.25) à zéro et en résolvant pour h_1 et h_2 , nous obtenons les fonctions d'offre associées aux régimes 2 et 3 respectivement (voir les équations 3A.3 et 3A.4 de l'annexe 3A). Ces fonctions admettent un rebroussement dans les salaires puisqu'elles sont non linéaires en w_1 , Ew_2 et Ew_2^2 . Il importe de souligner qu'en raison de la forme quadratique retenue, deux paramètres seulement définissent les caractéristiques du marché noir influençant l'offre de travail au noir, soit Ew_2 et Ew_2^2 . Ainsi, la probabilité d'être contrôlé, le taux d'amende et le salaire brut au noir influencent les offres de travail uniquement à travers ces deux paramètres. On notera en outre que, pour une espérance de salaire au noir donné ($= Ew_2$), la variable Ew_2^2 est une mesure du risque de salaire associé au travail au noir. En effet, la variance de la rémunération au noir est donnée par $Ew_2^2 - (Ew_2)^2$.

Revenu hors-travail virtuel et taux de taxation endogène

L'interaction des systèmes d'imposition et de transferts sociaux donne lieu à une contrainte budgétaire linéaire par segment. En conséquence, le taux marginal d'imposition auquel fait face un individu est endogène. Un modèle comptable de la fiscalité et des transferts québécois a donc été utilisé. Celui-ci calcule le taux marginal effectif de taxation sur le revenu de travail officiel de chaque individu ainsi que ses impôts nets des transferts à partir du revenu de travail, du revenu hors-travail et des caractéristiques démographiques de chaque ménage. Le revenu hors-travail est, quant à lui, calculé à partir des salaires et traitements, des revenus de placement et gains nets en capital, et des diverses prestations.⁹

Le revenu hors-travail virtuel y^v est calculé en linéarisant la contrainte budgétaire de l'individu au point correspondant à son niveau observé des heures de travail sur le marché officiel et sur le marché noir. En généralisant l'équation (6.7) de Blundell et MaCurdy (1999), il s'agit de résoudre :

$$C = W_1(1 - \tau)h_1 + w_2h_2 + y^v = W_1h_1 + w_2h_2 + y - T(.) \text{ pour } y^v.$$

On obtient alors :

$$y^v = y + \tau W_1 h_1 - T \quad (3.30)$$

où y est le revenu hors-travail de l'individu, T définit le montant de l'impôt net des transferts et τ le taux marginal effectif de taxation sur le revenu de travail officiel. Ces deux derniers éléments sont fonction du salaire brut officiel $W_1 h_1$, du revenu hors-travail y et des caractéristiques individuelles.

La *figure 3-1* ci-dessous présente le cas d'un impôt sur le revenu de travail progressif, en présence de programmes sociaux et en supposant l'absence de revenus non déclarés.

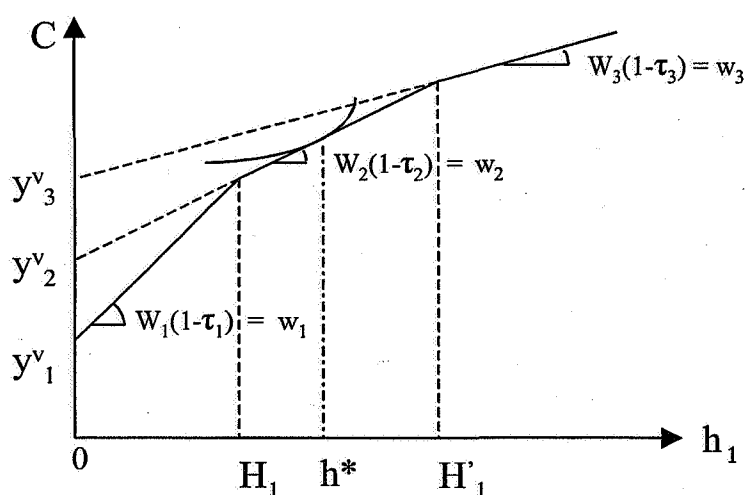
Trois taux d'imposition marginaux τ_1 , τ_2 et τ_3 conduisent à trois salaires marginaux nets w_1 , w_2 et w_3 respectivement. La transition entre les taux d'imposition τ_1 et τ_2 se produit à H_1 heures de travail et celle entre τ_2 et τ_3 à H'_1 heures de travail. Une mesure du revenu hors-travail virtuel est associée à chaque taux d'imposition.

Elle correspond au prolongement de la droite de salaire net sur l'axe des ordonnées. Cela revient à considérer une contrainte budgétaire linéaire, qui est tangente à l'ensemble budgétaire associé au nombre d'heures de travail observées (Hall, 1973). Ainsi,

⁹En raison du caractère statique de notre modèle, on suppose que l'horizon décisionnel de l'individu est d'une année. Les taxes portent sur les revenus présents de l'individu et ce dernier n'anticipe pas les revenus différés associés à certaines cotisations sociales.

nous supposons qu'un travailleur se comporte comme s'il faisait face à un unique taux de salaire et un unique revenu virtuel. En effet, l'individu décide d'offrir un nombre d'heures de travail h^* au taux de salaire net w_2 , étant donnée l'existence de l'impôt progressif. Il devrait également choisir le même nombre d'heures de travail h^* s'il recevait un revenu virtuel hors-travail de y_2^v et pouvait être rémunéré au taux de salaire net de w_2 .

Figure 3-1 :
Impôt progressif sur les revenus de travail



L'effet d'une fiscalité progressive est de créer un ensemble budgétaire convexe et linéaire par segment. Le barème fiscal engendre des coudes caractérisant les différentes tranches d'imposition liées au niveau de revenu imposable de l'individu. Par ailleurs, la présence de programmes de transferts sociaux axés sur le revenu engendre, en général, des non convexités dans l'ensemble budgétaire.

Le principal problème économétrique est, par conséquent, la multiplicité des taux de salaire nets auxquels l'individu fait face lorsqu'il détermine son offre de travail, ainsi que la non convexité de l'ensemble budgétaire. Pour résoudre ce problème, nous adoptons une approche de convexification (locale) de l'ensemble budgétaire (Bourguignon et Magnac, 1990 ; Blundell et MaCurdy, 1999) et nous considérons le taux de salaire marginal net ainsi que le revenu hors-travail virtuel comme variables endogènes. Nous linéarisons la contrainte budgétaire de sorte que le revenu hors-travail doit être remplacé par le revenu virtuel. Afin d'endogénéiser le revenu hors-travail virtuel, nous

spécifions l'équation suivante :

$$y^v = X_3' \Psi_3 + \varepsilon_3 \quad (3.31)$$

où X_3' est un vecteur de variables exogènes. En ce qui concerne les couples, nous avons appliqué différentes règles de partage du revenu hors-travail. La première des règles consiste en une répartition égalitaire du revenu hors-travail virtuel entre les membres du ménage. La seconde règle consiste à attribuer une part du revenu hors-travail proportionnelle aux salaires reçus par chacun des époux. Enfin, la troisième règle repose sur le résultat obtenu par Chiappori, Fortin et Lacroix (2001), à savoir une répartition à hauteur de 70% pour la femme et de 30% pour l'homme. La dernière règle a été retenue dans le présent chapitre. Il convient de noter que les résultats sont en général robustes à ces différentes spécifications.

Les équations de salaire

Nous spécifions les équations salariales linéaires suivantes :

$$w_1 = X_4' \Psi_4 + \varepsilon_4 \quad (3.32)$$

$$Ew_2 = X_5' \Psi_5 + \varepsilon_5 \quad (3.33)$$

$$E(w_2^2) = X_6' \Psi_6 + \varepsilon_6 \quad (3.34)$$

où X_i' est un vecteur de caractéristiques socio-économiques exogènes et ε_i est un terme d'erreur avec $i = 4, 5, 6$.¹⁰ La linéarité de ces équations facilitera l'estimation économétrique du modèle (Moffit, 1984).

Dans la mesure où h_1 et h_2 sont positifs dans le régime 1, les variables de salaire dans les équations (3.32), (3.33) et (3.34) sont observées dans ce régime. En revanche, dans le régime 2, nous observons uniquement le salaire obtenu sur le marché officiel et pour le régime 3, nous observons uniquement le salaire de l'activité au noir. Enfin, dans le régime 4, il importe de tenir compte de ces « données manquantes » dans l'estimation du modèle. Les conditions relatives à ces trois derniers régimes sont reportées à l'annexe 3C.

¹⁰L'équation $E(w_2^2)$ devrait être exprimée en fonction de $E(w_2)$ et d'un terme aléatoire, de sorte que des termes d'interactions devraient apparaître. Cependant, pour plusieurs raisons, nous supposons que ces deux équations de salaire sont indépendantes l'une de l'autre. La première raison tient au fait que la prise en compte d'une telle interaction compliquerait singulièrement la fonction de vraisemblance, et en particulier les régimes 1 et 3. De ce fait, l'estimation du modèle serait plus délicate. Nous serions, en outre, contraints de formuler des hypothèses sur la nature de l'interaction entre les équations afin de procéder à l'estimation d'un tel modèle. Or, l'hypothèse d'une corrélation particulière s'avère tout aussi arbitraire que celle de l'indépendance entre Ew_2 et Ew_2^2 . Enfin, la complexité accrue du modèle pourrait rendre l'interprétation des résultats plus difficile et pourrait finalement n'apporter que peu d'éclairage. C'est pourquoi, nous préférons ne pas tenir compte des termes d'interaction entre Ew_2 et $E(w_2^2)$. En ce sens, l'hypothèse d'indépendance peut être considérée comme un cas particulier d'un modèle plus général [Nous devons ce point à Guy Laroque].

Une structure stochastique

Les attributs socio-démographiques et les différences non observables entre individus sont introduits dans le modèle en décomposant les coefficients α_1 et α_2 en une partie déterministe et une partie aléatoire. A l'instar de Ransom (1987), les différences dans les préférences entre individus sont pris en compte en réécrivant α_1 et α_2 ainsi :

$$\alpha_1 = \bar{\alpha}_1 + X'_1 \delta_1 + \varepsilon_1 \quad (3.35)$$

$$\alpha_2 = \bar{\alpha}_2 + X'_2 \delta_2 + \varepsilon_2 \quad (3.36)$$

où $\bar{\alpha}_i$ est une constante, X'_i un vecteur de caractéristiques socio-démographiques, et δ_i est un vecteur de paramètres à estimer, avec $i = 1, 2$. Les termes aléatoires ε_1 et ε_2 sont introduits pour tenir compte de l'hétérogénéité non observable dans les préférences. On suppose finalement que $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6)$ est distribué $N(0, \Sigma)$.

3.4 Estimation du modèle

Parmi les difficultés liées à l'estimation du modèle figurent l'absence d'information sur les salaires des individus qui ne travaillent pas sur l'un ou l'autre des marchés ainsi que l'endogénéité des taux marginaux d'imposition, du revenu «virtuel» hors-travail et des salaires sur les deux marchés du travail. Nous avons choisi de résoudre ces problèmes en procédant aux estimations par la méthode de maximisation de la fonction de vraisemblance à information complète. Ainsi, nous étudions simultanément les comportements sur les marchés officiel et au noir en tenant compte des problèmes liés aux équations simultanées avec variables endogènes censurées.

Dans une première étape, nous présentons la fonction de vraisemblance associée aux quatre états possibles. Puis, nous développons la procédure d'estimation utilisée.

3.4.1 La fonction de vraisemblance

Dans le régime 1, l'individu travaille sur les deux marchés ($h_1 > 0$ et $h_2 > 0$). En utilisant les équations (3.24) et (3.25), la fonction de densité jointe des heures de travail et des salaires sur les deux marchés est :

$$g_1(h_1, h_2, y^v, w_1, Ew_2, Ew_2^2; X') = f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6) |J_1| \quad (3.37)$$

où X' est le vecteur des X'_i ($i = 1, \dots, 6$), $|J_1|$ représente la valeur absolue du déterminant de la matrice jacobienne de la transformation des termes inobservés $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6)$ en termes observés $(h_1, h_2, y^v, w_1, Ew_2, Ew_2^2)$ et où $f(\cdot)$ est la fonction de densité normale à six dimensions.

La densité de probabilité associée au premier régime se caractérise de la manière suivante :

$$\begin{aligned}
 g_1(h_1, h_2, y^v, w_1, Ew_2, Ew_2^2; X') &= f(\varepsilon_1(h_1, h_2, y^v, w_1, Ew_2; X'_1), \\
 &\varepsilon_2(h_1, h_2, y^v, w_1, Ew_2, Ew_2^2; X'_2), \\
 &\varepsilon_3(y^v; X'_3), \varepsilon_4(w_1, h_1; X'_4), \\
 &\varepsilon_5(Ew_2, h_2; X'_5), \varepsilon_6(Ew_2^2, h_2; X'_6)) \quad |J_1|
 \end{aligned} \tag{3.38}$$

$$\begin{aligned}
 P &= f(-\bar{\alpha}_1 - X'_1\delta_1 - \alpha_3w_1 - \beta_{11}h_1 - \beta_{12}h_2 - \beta_{13}(y^v + 2w_1h_1 + h_2Ew_2) \\
 &\quad - \beta_{23}w_1h_2 - \beta_{33}w_1(y^v + w_1h_1 + h_2Ew_2), \\
 &\quad - \bar{\alpha}_2 - X'_2\delta_2 - \alpha_3Ew_2 - \beta_{12}h_1 - \beta_{13}h_1Ew_2 - \beta_{22}h_2 \\
 &\quad - \beta_{23}(y^v + w_1h_1 + 2h_2Ew_2) - \beta_{33}[(y^v + w_1h_1)Ew_2 + h_2Ew_2^2], \\
 &\quad y^v - X'_3\Psi_3, w_1 - X'_4\Psi_4, Ew_2 - X'_5\Psi_5, Ew_2^2 - X'_6\Psi_6) \quad |J_1|
 \end{aligned}$$

et $|J_1|$ s'écrit :

$$|J_1| = \begin{vmatrix} \Delta_{11} & \Delta_{12} & \Delta_{13} & \Delta_{14} & \Delta_{15} & 0 \\ \Delta_{21} & \Delta_{22} & \Delta_{23} & \Delta_{24} & \Delta_{25} & \Delta_{26} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix} \tag{3.39}$$

où

$$\begin{aligned}
 \Delta_{11} &= -\beta_{11} - 2\beta_{13}w_1 - \beta_{33}w_1^2 \\
 \Delta_{12} &= -\beta_{12} - \beta_{13}Ew_2 - \beta_{23}w_1 - \beta_{33}w_1Ew_2 \\
 \Delta_{13} &= -\beta_{13} - \beta_{33}w_1 \\
 \Delta_{14} &= -\alpha_3 - 2\beta_{13}h_1 - \beta_{23}h_2 - \beta_{33}(y^v + 2w_1h_1 + h_2Ew_2) \\
 \Delta_{15} &= -\beta_{13}h_2 - \beta_{33}w_1h_2 \\
 \Delta_{21} &= -\beta_{12} - \beta_{13}Ew_2 - \beta_{23}w_1 - \beta_{33}w_1Ew_2 \\
 \Delta_{22} &= -\beta_{22} - 2\beta_{23}Ew_2 - \beta_{33}Ew_2^2 \\
 \Delta_{23} &= -\beta_{23} - \beta_{33}Ew_2 \\
 \Delta_{24} &= -\beta_{23}h_1 - \beta_{33}h_1Ew_2 \\
 \Delta_{25} &= -\alpha_3 - \beta_{13}h_1 - 2\beta_{23}h_2 - \beta_{33}(y^v + w_1h_1) \\
 \Delta_{26} &= -\beta_{33}h_2
 \end{aligned}$$

Dans le régime 2, l'individu travaille uniquement sur le marché officiel ($h_1 > 0$ et $h_2 = 0$). La probabilité de n'exercer qu'un emploi déclaré est :

$$P(h_1 > 0, h_2 = 0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_2^*} f(-\bar{\alpha}_1 - X'_1 \delta_1 - \alpha_3 w_1 - \beta_{11} h_1 - \beta_{13}(y^v + 2w_1 h_1) - \beta_{33} w_1 (y^v + w_1 h_1), \varepsilon_2, y^v - X'_3 \Psi_3, w_1 - X'_4 \Psi_4, \varepsilon_5, \varepsilon_6) d\varepsilon_2 d\varepsilon_5 d\varepsilon_6 |J_2| \quad (3.40)$$

La fonction mixte et de densité cumulative est la suivante :

$$g_2(h_1, y^v, w_1; X') = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_2^*} f(h_1, \varepsilon_2, y^v, w_1, \varepsilon_5, \varepsilon_6) d\varepsilon_2 d\varepsilon_5 d\varepsilon_6 |J_2| \quad (3.41)$$

où $|J_2|$ représente la valeur absolue de déterminant de la matrice jacobienne de la transformation des termes inobservés ($\varepsilon_1, \varepsilon_3, \varepsilon_4$) en termes observés (h_1, y^v, w_1). $|J_2|$ s'écrit de la façon suivante :

$$|J_2| = \begin{vmatrix} \Delta_{11} & \Delta_{12} & \Delta_{13} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \quad (3.42)$$

où

$$\begin{aligned} \Delta_{11} &= -\beta_{11} - 2\beta_{13}w_1 - \beta_{33}w_1^2 \\ \Delta_{12} &= -\beta_{13} - \beta_{33}w_1 \\ \Delta_{13} &= -\alpha_3 - 2\beta_{13}h_1 - \beta_{33}(y^v + 2w_1h_1) \end{aligned}$$

et où ε_2^* est tel que $m_2(h_1, 0, y^v, w_1, Ew_2, Ew_2^2; X'_2) = 0$, *i.e.*

$$\varepsilon_2^* = -\bar{\alpha}_2 - X'_2 \delta_2 - (X'_5 \Psi_5)(\alpha_3 + \beta_{13}h_1 + \beta_{33}(y^v + w_1h_1)) - \beta_{12}h_1 - \beta_{23}(y^v + w_1h_1) - \varepsilon_5(\alpha_3 + \beta_{13}h_1 + \beta_{33}(y^v + w_1h_1)) \quad (3.43)$$

Nous constatons par (3.43) que le terme aléatoire ε_2^* est lui-même fonction du terme aléatoire ε_5 , ce qui complique quelque peu la programmation de la vraisemblance. Pour résoudre ce problème numérique, dans le processus d'intégration sur l'intervalle $[-\infty; \varepsilon_2^*]$, nous fixons une valeur de ε_5 et nous évaluons la fonction pour chacune des valeurs de ε_2 , puis nous fixons une nouvelle valeur pour ε_5 et nous réévaluons la fonction pour chaque valeur de ε_2 . Nous réitérons ce mode de calcul sur toutes les valeurs de ε_5 et de ε_2 afin de s'assurer que l'on balaie sur les deux dimensions.

Le régime 3 est le symétrique du deuxième. Il caractérise l'exercice d'un emploi au noir uniquement ($h_1 = 0$ et $h_2 > 0$). La probabilité d'exercer une activité souterraine seulement s'écrit de la façon suivante :

$$P(h_1 = 0, h_2 > 0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_1^*} f(\varepsilon_1, -\bar{\alpha}_2 - X'_2 \delta_2 - \alpha_3 E w_2 - \beta_{22} h_2 - \beta_{23}(y^v + 2h_2 E w_2) - \beta_{33}(y^v E w_2 + h_2 E w_2^2), y^v - X'_3 \Psi_3, \varepsilon_4, E w_2 - X'_5 \Psi_5, E w_2^2 - X'_6 \Psi_6) d\varepsilon_1 d\varepsilon_4 |J_3| \quad (3.44)$$

La fonction mixte et de densité cumulative est donc :

$$g_3(h_2, y^v, E w_2, E w_2^2; X') = \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_1^*} f(\varepsilon_1, h_2, y^v, \varepsilon_4, E w_2, E w_2^2) d\varepsilon_1 d\varepsilon_4 |J_3| \quad (3.45)$$

où $|J_3|$ représente la valeur absolue du déterminant de la matrice jacobienne de la transformation des termes inobservés ($\varepsilon_2, \varepsilon_3, \varepsilon_5, \varepsilon_6$) en termes observés ($h_2, y^v, E w_2, E w_2^2$), i.e.

$$|J_3| = \begin{vmatrix} \Delta_{11} & \Delta_{12} & \Delta_{13} & \Delta_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (3.46)$$

où

$$\begin{aligned} \Delta_{11} &= -\beta_{22} - 2\beta_{23} E w_2 - \beta_{33} E w_2^2 \\ \Delta_{12} &= -\beta_{23} - \beta_{33} E w_2 \\ \Delta_{13} &= -\alpha_3 - 2\beta_{23} h_2 - \beta_{33} y^v \\ \Delta_{14} &= -\beta_{33} h_2 \end{aligned}$$

et ε_1^* est tel que $m_1(0, h_2, y^v, w_1, E w_2, E w_2^2, X'_1) = 0$. Autrement dit :

$$\begin{aligned} \varepsilon_1^* &= -\bar{\alpha}_1 - X'_1 \delta_1 - \beta_{12} h_2 - \beta_{13}(y^v + h_2 E w_2) \\ &\quad - (X'_4 \Psi_2)(\alpha_3 + \beta_{23} h_2 + \beta_{33}(y + h_2 E w_2)) \\ &\quad - \varepsilon_4(\alpha_3 + \beta_{23} h_2 + \beta_{33}(y^v + h_2 E w_2)) \end{aligned} \quad (3.47)$$

Ici encore, nous constatons que ε_1^* est lui-même fonction d'un autre terme aléatoire ε_4 . Le processus d'intégration développé pour le régime 2 est reproduit afin d'évaluer la fonction cumulative. Nous nous garantissons ainsi de couvrir les deux dimensions correspondant aux termes d'erreur.

Dans le régime 4, l'individu ne travaille sur aucun marché ($h_1 = 0$ et $h_2 = 0$). La fonction mixte et de densité cumulative est :

$$g_4(y^v; X') = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_1^{**}} \int_{-\infty}^{\varepsilon_2^{**}} f(\varepsilon_1, \varepsilon_2, y^v, \varepsilon_4, \varepsilon_5, \varepsilon_6) d\varepsilon_1 d\varepsilon_2 d\varepsilon_4 d\varepsilon_5 d\varepsilon_6$$

où ε_1^{**} est tel que $m_1(0, 0, y^v, w_1, Ew_2, Ew_2^2; X'_1) = 0$:

$$\varepsilon_1^{**} = -\bar{\alpha}_1 - X'_1 \delta_1 - \beta_{13} y^v - (X'_4 \Psi_4)(\alpha_3 + \beta_{33} y^v) - \varepsilon_4(\alpha_3 + \beta_{33} y^v) \quad (3.48)$$

et où ε_2^{**} est tel que $m_2(0, 0, y^v, w_1, Ew_2, Ew_2^2; X'_2) = 0$:

$$\varepsilon_2^{**} = -\bar{\alpha}_2 - X'_2 \delta_2 - \beta_{23} y^v - (X'_5 \Psi_5)(\alpha_3 + \beta_{33} y^v) - \varepsilon_5(\alpha_3 + \beta_{33} y^v) \quad (3.49)$$

Une nouvelle fois, les termes ε_1^{**} et ε_2^{**} sont fonction de deux autres termes aléatoires, respectivement ε_4 et ε_5 . Nous reproduisons le même processus d'intégration que pour les régimes 2 et 3 afin d'évaluer la fonction cumulative.

La fonction de vraisemblance associée aux quatre régimes s'écrit :

$$\begin{aligned} L = & \prod_{i \in G_1} f(-\bar{\alpha}_1 - X'_1 \delta_1 - \alpha_3 w_1 - \beta_{11} h_1 - \beta_{12} h_2 - \beta_{13}(y^v + 2w_1 h_1 + h_2 Ew_2) \\ & - \beta_{23} w_1 h_2 - \beta_{33} w_1 (y^v + w_1 h_1 + h_2 Ew_2), -\bar{\alpha}_2 - X'_2 \delta_2 - \alpha_3 Ew_2 \\ & - \beta_{12} h_1 - \beta_{13} h_1 Ew_2 - \beta_{22} h_2 - \beta_{23}(y^v + w_1 h_1 + 2h_2 Ew_2) \\ & - \beta_{33}[(y^v + w_1 h_1)Ew_2 + h_2 Ew_2^2], y^v - X'_3 \Psi_3, w_1 - X'_4 \Psi_4, \\ & Ew_2 - X'_5 \Psi_5, Ew_2^2 - X'_6 \Psi_6) |J_1| \\ & \times \\ & \prod_{i \in G_2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_2^*} f(-\bar{\alpha}_1 - X'_1 \delta_1 - \alpha_3 w_1 - \beta_{11} h_1 - \beta_{13}(y^v + 2w_1 h_1) \\ & - \beta_{33} w_1 (y^v + w_1 h_1), \varepsilon_2, y^v - X'_3 \Psi_3, w_1 - X'_4 \Psi_4, \varepsilon_5, \varepsilon_6) d\varepsilon_2 d\varepsilon_5 d\varepsilon_6 |J_2| \\ & \times \\ & \prod_{i \in G_3} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_1^*} f(\varepsilon_1, -\bar{\alpha}_2 - X'_2 \delta_2 - \alpha_3 Ew_2 - \beta_{22} h_2 - \beta_{23}(y^v + 2h_2 Ew_2) \\ & - \beta_{33}(y^v Ew_2 + h_2 Ew_2^2), y^v - X'_3 \Psi_3, \varepsilon_4, Ew_2 - X'_5 \Psi_5, \\ & Ew_2^2 - X'_6 \Psi_6) d\varepsilon_1 d\varepsilon_4 |J_3| \\ & \times \\ & \prod_{i \in G_4} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_1^{**}} \int_{-\infty}^{\varepsilon_2^{**}} f(\varepsilon_1, \varepsilon_2, y^v - X'_3 \Psi_3, \varepsilon_4, \varepsilon_5, \varepsilon_6) d\varepsilon_1 d\varepsilon_2 d\varepsilon_4 d\varepsilon_5 d\varepsilon_6 \end{aligned} \quad (3.50)$$

$$\begin{aligned} \text{où } G_1 &= \{h_1 > 0, h_2 > 0\}, \\ G_2 &= \{h_1 > 0, h_2 = 0\}, \\ G_3 &= \{h_1 = 0, h_2 > 0\}, \\ G_4 &= \{h_1 = 0, h_2 = 0\}. \end{aligned}$$

Pour les régimes 2, 3 et 4, la fonction de densité jointe peut être réécrite comme le produit d'une densité conditionnelle et d'une densité marginale. Cela permet de réduire le niveau d'intégration. Par ailleurs, le terme aléatoire ε_6 -associé à l'espérance du salaire au noir élevé au carré- est intégré sur l'intervalle $[-\infty; +\infty]$, de sorte que sa densité est égale à 1 et n'affecte donc pas la fonction de vraisemblance. Un niveau d'intégration supplémentaire peut alors être éliminé pour les deuxième et quatrième régimes. Ainsi, l'expression (3.50) devient :

$$\begin{aligned} L = & \prod_{i \in G_1} f(h_1, h_2, y^v, w_1, Ew_2, Ew_2^2; X') |J_1| \\ & \times \\ & \prod_{i \in G_2} \int_{-\infty}^{+\infty} \Phi(\varepsilon_2 | h_1, y^v, w_1, \varepsilon_5; X') f(h_1, y^v, w_1, \varepsilon_5; X) d\varepsilon_5 |J_2| \\ & \times \\ & \prod_{i \in G_3} \int_{-\infty}^{+\infty} \Phi(\varepsilon_1 | h_2, y^v, \varepsilon_4, Ew_2, Ew_2^2; X') f(h_2, y^v, \varepsilon_4, Ew_2, Ew_2^2; X') d\varepsilon_4 |J_3| \\ & \times \\ & \prod_{i \in G_4} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_2^{**}} \Phi(\varepsilon_1 | \varepsilon_2, y^v, \varepsilon_4, \varepsilon_5; X') f(\varepsilon_2, y^v, \varepsilon_4, \varepsilon_5; X') d\varepsilon_2 d\varepsilon_4 d\varepsilon_5 \end{aligned} \quad (3.51)$$

Les densités jointes se résument alors au produit d'une densité cumulative normale univariée et d'une densité normale à quatre dimensions (régimes 2 et 4) ou à cinq dimensions (régime 3).

3.4.2 La procédure d'estimation

Les paramètres de la fonction d'utilité et des équations de salaire et de revenu virtuel sont estimés en maximisant la fonction de vraisemblance (3.51). En pratique, il est plus aisé de maximiser le logarithme naturel de la fonction.¹¹ L'équation (3.51) devient alors :

$$\begin{aligned} l = \ln L = & \sum_1 \ln g_1(h_1, h_2, y^v, w_1, Ew_2, Ew_2^2; X') + \sum_2 \ln g_2(h_1, y^v, w_1; X') \\ & + \sum_3 \ln g_3(h_2, y^v, Ew_2, Ew_2^2; X') + \sum_4 \ln g_4(y^v; X') \end{aligned}$$

¹¹Cela n'affecte en rien les valeurs des paramètres estimés puisque le logarithme est un opérateur monotone.

La procédure utilisée est la méthode du maximum de vraisemblance à information complète. Le calcul de la log-vraisemblance se heurte à l'évaluation d'intégrales multiples définissant l ; celles-ci faisant intervenir une loi normale à cinq dimensions. Un tel calcul d'intégration ne peut être effectué analytiquement. Il requiert une évaluation numérique et relève d'un problème mathématique largement débattu dans la littérature, celui de la quadrature gaussienne.¹² Divers procédés sont envisageables parmi lesquels la méthode des rectangles, celle des trapèzes, la règle de Simpson ou encore les quadratures «Gauss-Legendre», «Gauss-Hermite», «Gauss-Laguerre» ou «Gauss-Chebyshev». La méthode la plus simple pour approximer :

$$G(x) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} g(x) dx,$$

est probablement la règle trapézoïdale suivante :

$$G(x) = \frac{h}{2} [(g_0 + g_N) + h(g_1 + g_2 + \dots + g_{N-1})],$$

où g_j est la fonction évaluée à N points répartis de façon égale sur l'intervalle $[\underline{\varepsilon}, \bar{\varepsilon}]$ et h désigne la longueur de chaque segment, de sorte que $h = (\bar{\varepsilon} - \underline{\varepsilon})/N$ ou $\bar{\varepsilon} = \underline{\varepsilon} + Nh$. Au plan géométrique, cette procédure s'interprète comme la somme des aires des N trapèzes de longueur h et de poids moyen $(g_j + g_{j+1})/2$. Elle consiste en une approximation linéaire par segment. La méthode des rectangles et la règle de Simpson sont des variantes de celle des trapèzes; celles-ci sont qualifiées de formules Newton-Cotes. Le principal problème de ces méthodes provient de la difficulté à obtenir une précision satisfaisante avec un nombre limité de points. Par ailleurs, les formules Newton-Cotes reposent sur le choix arbitraire des différents points d'évaluation, habituellement répartis de façon uniforme sur l'intervalle $[\underline{\varepsilon}, \bar{\varepsilon}]$. A l'inverse, les formules dites «gaussiennes» (*i. e.* Gauss-Legendre, Gauss-Hermite, Gauss-Laguerre) sont construites à partir de segments et de poids choisis simultanément. En ce sens, ces méthodes sont plus efficaces.

De façon générale, les formules gaussiennes consistent à utiliser une approximation de la forme :

$$l = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} P(x)g(x)dx \approx \sum_{j=1}^N p_j g(a_j),$$

où $P(x)$ est considérée comme une fonction de pondération pour intégrer $g(x)$, p_j est le poids de quadrature et a_j est l'abscisse de quadrature. Différents poids et abscisses ont

¹²Pour de plus amples détails, se reporter aux ouvrages mathématiques et économétriques d'Abromowitz et Stegun (1971), de Judd (1998) et de Wooldrige (2001), par exemple.

été dérivés pour de nombreuses fonctions pondérées ; ceux-ci figurent dans les tables numériques usuelles.¹³

Mais, l'intégration numérique par la technique standard de quadrature gaussienne des intégrales de dimension supérieure ou égale à trois requiert un temps de calcul considérable (Train, 2002). Dans notre cas, une difficulté supplémentaire apparaît puisque les intégrales des régimes 2, 3 et 4 sont elles-mêmes exprimées en fonction d'autres intégrales, ce qui nous complique encore davantage la tâche. De ce fait, l'estimation des paramètres par la méthode du maximum de vraisemblance traditionnelle est, à la fois, coûteuse en termes de temps et en termes de programmation.¹⁴ L'estimation de la fonction de vraisemblance a été réalisée à partir de la sous-routine Optimum du logiciel Gauss et la procédure d'intégration numérique a consisté à adapter une formule Gauss-Legendre.

La fonction a été maximisée par rapport à tous ses paramètres à l'aide de l'algorithme BFGS (Broyden-Fletcher-Goldfarb-Shanno).¹⁵ Les valeurs de départ ont été calculées en appliquant l'algorithme itératif Gauss-Seidel (dit «de déplacements successifs») aux formes structurelles du modèle dans le cas de solutions intérieures sur les deux marchés. Il s'agit essentiellement de contraindre certains paramètres à des valeurs arbitraires mais plausibles, et à estimer les autres par des méthodes simples telles que les moindres carrés ordinaires. Les paramètres ainsi obtenus sont fixés à leur tour, et les paramètres initialement contraints sont relâchés. La convergence des paramètres est assurée par répétition successive de la procédure. Comme nous n'utilisons que les observations du premier, deuxième et troisième régimes, ces paramètres sont, en général, biaisés en raison de la sélection de l'échantillon. Ils constituent néanmoins des valeurs de départ utiles dans les circonstances.

Dans la mesure où la fonction d'utilité est définie à une transformation affine près, un des paramètres doit être normalisé. Nous avons choisi de contraindre le paramètre

¹³Se reporter à Press *et al.* (1986, pp.121-125) pour un exposé de la dérivation théorique des poids et abscisses. Par ailleurs, des tables de poids et d'abscisses sont proposées par Abramovitz and Stegun (1971), pour de nombreuses valeurs de N . Enfin, Judd (1998) présente de façon détaillée les différentes techniques d'intégration numérique.

¹⁴A titre indicatif, la convergence du modèle a nécessité trois mois de calculs ininterrompus.

¹⁵Ces algorithmes, qualifiés de correction de second et de troisième rang, appartiennent à la famille des algorithmes Quasi-Newton. Conçus pour éviter le calcul contraignant de la matrice hessienne, ils permettent d'écartier le problème de l'existence d'une matrice non définie négative lorsque l'estimation est relativement loin du maximum. Ces algorithmes, extrêmement efficaces, présentent d'excellentes propriétés de convergence et n'imposent pas la concavité de la fonction de vraisemblance. Néanmoins, l'approximation fournie par ces algorithmes sous-estime systématiquement les écarts-types. La matrice de variance-covariance des estimateurs ne peut donc être utilisée directement et nécessite d'être recalculée à partir de l'algorithme de Newton. Pour plus de détails, voir Judge *et al.* (1980).

α_3 à 0.25.¹⁶ Par le principe d'invariance de la fonction de vraisemblance, les paramètres estimés ne sont pas affectés par le choix de la normalisation. Les coefficients associés aux variables de contrôle doivent être interprétés comme mesurant l'impact de ces variables sur l'utilité marginale de la catégorie de travail où elles apparaissent. L'interaction entre les deux marchés du travail se manifeste par l'intermédiaire du paramètre β_{12} .

3.4.3 Construction des variables

Parmi les variables explicatives, nous distinguons les caractéristiques individuelles objectives et les croyances subjectives relatives à l'économie souterraine. Nous prenons en considération non seulement le sexe, l'âge, le niveau d'éducation, la situation matrimoniale, le nombre d'enfant, les caractéristiques professionnelles, la situation financière des individus, mais également leur sentiment sur l'exercice d'une activité dissimulée et le jugement que, selon eux, la société porte sur ce phénomène. Ces dernières nous permettent d'appréhender partiellement l'effet de voisinage dans le choix de l'activité.

Variables subjectives

L'enquête présente une évaluation par les individus du pourcentage de participants à l'économie souterraine parmi les personnes de leur entourage (PROPOREN). Cette dernière variable est graduée de façon discrète sur l'intervalle [1;11] où 1 correspond à 0% de travailleurs au noir et 11 à 100%. Des questions d'ordre moral ont également été posées. D'une part, l'exercice d'un emploi non déclaré est jugé, selon une échelle de valeur, de très moral à très immoral (IMMORALITE). D'autre part, l'approbation ou la désapprobation des proches à l'exercice d'une activité souterraine est évaluée subjectivement par l'individu à travers la variable «*réaction de l'entourage*».

Les variables subjectives nous permettent de mettre en évidence les processus d'interactions sociales qui sous-tendent les prises de décisions individuelles. Elle reflètent l'interdépendance des comportements au sein d'un groupe d'individus et son impact sur la détermination des préférences de chacun. Néanmoins, le recours à de telles variables soulève un problème d'identification connu sous le nom de problème du miroir¹⁷ (Manski 2000). Un traitement particulier est alors requis. Il serait possible de remédier à ce biais à l'aide de techniques de variables instrumentales. Toutefois, cela signifierait l'introduction de variables supplémentaires à estimer. Or, étant donné la complexité de la fonction de vraisemblable et le nombre limité de travailleurs au noir, cela s'avère proprement irréalisable.

¹⁶Le choix de cette normalisation est fondé sur les travaux de Lacroix et Fortin (1992).

¹⁷«reflection problem».

Taux marginal d'imposition et impôt net des transferts

La structure des programmes de transfert au Québec est telle que le taux marginal effectif d'imposition peut être supérieur en valeur absolue à 100%. Cette variable est, en effet, définie sur l'intervalle $[-133.30; 177]$. Toutefois, pour les individus ayant déclaré une activité souterraine, nous avons fixé une borne inférieure à zéro et substitué les valeurs négatives en conséquence. Ce choix est justifié par le fait que, si $\tau < 0$, le salaire espéré au noir en cas de détection par les autorités fiscales est supérieur au salaire espéré au noir en cas de non détection. Autrement dit, le travailleur aurait intérêt à être repéré par le fisc, ce qui paraît insensé.

Probabilité de détection et taux d'amende

Au cours de l'enquête, il a été demandé aux personnes interrogées d'évaluer la probabilité de détection d'une activité souterraine et le taux de pénalité encouru. Cette information n'est utilisée que pour les travailleurs au noir. Elle permet d'estimer le risque d'une activité souterraine et, par conséquent, l'impact du système fiscal et de répression de la fraude sur l'exercice d'une telle activité. Nous aurions pu exploiter les publications officielles des services fiscaux, et les divers textes de loi relatifs au travail au noir, pour appréhender la probabilité de détection par secteur d'activité et le montant réel des pénalités. Cependant, les décisions individuelles sont probablement davantage fondées sur une évaluation subjective que sur une connaissance précise de ces éléments (Cowell, 1990; Andréoni, Erard et Feinstein, 1998). D'une part, les individus ne peuvent obtenir qu'une connaissance partielle de ces informations dans la mesure où les instances gouvernementales de la plupart des pays développés ne les rendent pas accessibles au grand public (Benjamini *et al.*, 1983). D'autre part, la probabilité de détection étant le produit de la probabilité de contrôle et de la probabilité de détection en cas de contrôle, même si la fréquence des contrôles est connue parfaitement, la probabilité d'être détecté ne l'est pas forcément. Enfin, il est probable que les individus aient une interprétation très subjective des informations diffusées. Il est donc plus judicieux de procéder de cette manière.

Les questions touchant à la probabilité de détection et l'amende sont toutefois de nature à soulever des problèmes d'endogénéité. Cela est dû au fait que les individus ayant travaillé sur le marché noir sont susceptibles de sous-estimer les valeurs de θ et p , de façon à réduire le stress associé à cette activité. Une telle attitude est à la source d'un biais de dissonance cognitive. En effet, selon la théorie de la dissonance cognitive empruntée à Akerlof et Dickens (1982), les individus ont des préférences quant à l'état de la nature et quant à leurs croyances en l'état de la nature. Elle nous enseigne, en

outre, qu'ils peuvent exercer un certain contrôle sur leurs croyances. Ils peuvent, en effet, manipuler leurs croyances en sélectionnant les sources d'information qui confirment leurs désirs. Tout comme les travailleurs d'emplois dangereux sous-estiment le danger inhérent à leur activité, les participants à l'économie souterraine peuvent se convaincre du caractère peu risqué de l'activité non déclarée. Ils rejettent généralement tout sentiment d'insécurité et déclarent de faibles probabilités de sanctions. Par conséquent, le risque de détection et l'amende associée sont endogènes à la participation au marché noir et doivent être traitées comme telles.

Or, la procédure économétrique que nous développons ici tient compte de l'endogénéité de ces paramètres. En effet, nous procédons à l'estimation de l'espérance du salaire (et son espérance au carré) sur le marché noir et non pas à l'estimation du salaire lui-même. La probabilité de détection et l'amende encourue sont alors indirectement endogénéisées et estimées, ce qui écarte ce problème de dissonance cognitive.

La question relative à l'amende à payer pose, en revanche, un autre problème méthodologique. En effet, en posant que la pénalité est proportionnelle à l'impôt impayé, le salaire au noir en cas de détection est donné par $w_2 = W_2(1 - \theta)$ avec :

$$\theta = (1 + \lambda)\tau$$

Dans ce cas, pour que l'amende ait un effet dissuasif, il faut que $\lambda > 0$. Cette variable n'a cependant pas de bornes naturelles, contrairement à la variable de probabilité. Ainsi, sur 1000\$ d'impôt fraudé, certains individus estiment devoir payer jusqu'à 25000\$ d'amende additionnelle aux 1000\$ dus, soit vingt-cinq fois le montant de la fraude. Dans ces conditions, déterminer un intervalle pour le montant de l'amende relève quelque peu de l'arbitraire.

Nous avons alors choisi de redéfinir cette variable de deux manières différentes. En premier lieu, nous avons fixé une borne supérieure équivalente à 100 % de la fraude, de sorte que l'amende maximale soit égale à 1000\$ en plus des 1000\$ impayés. Puis, nous avons envisagé la possibilité de doubler le montant du redressement. Un traitement particulier a alors été effectué sur les valeurs excédant 1000\$. En effet, une fonction cumulative normale centrée-réduite a été créée pour ces valeurs, de sorte que le montant de l'amende est, dans ce cas, défini sur l'intervalle [0 ; 2000]. Après quelques expérimentations, nous avons décidé de retenir cette dernière spécification.

Le tableau 3-1 retrace la moyenne des variables τ , λ , θ et p par régime.

Tableau 3-1 :
Valeurs moyennes

	Régime 1	Régime 2	Régime 3	Régime 4	Total
Taux marginal d'impôt : τ	0.50	0.48	0.21	0.19	0.34
Taux de pénalité : λ	0.52	0.52	0.51	0.56	0.53
Amende : θ	0.76	0.73	0.32	0.30	0.52
(où $\theta = [1 + \lambda] \tau$)					
Probabilité de détection : p	0.23	0.38	0.27	0.40	0.32
Nombre d'observations	104	2893	86	1576	4659 ^c

c - Le nombre d'observations correspond à notre sous-échantillon après retrait des valeurs aberrantes.

A la lumière de ce tableau, apparaissent deux remarques essentielles concernant la manière dont les individus perçoivent leur probabilité de détection, en cas de fraude. D'une part, cette probabilité subjective est jugée, par ceux qui n'exercent pas d'activité souterraine (régime 2 et 4), plus élevée que par ceux qui travaillent au noir (régime 1 et 3). En moyenne, elle est de l'ordre de 40% selon les premiers alors qu'elle n'atteint pas 30% selon les travailleurs au noir. Dès lors, ces évaluations subjectives peuvent contribuer à expliquer le choix de l'activité. En effet, il semble logique qu'un individu ne travaille pas au noir s'il estime encourir un risque important d'être repéré par les autorités fiscales. Cet argument a d'ailleurs été confirmé par les travaux de Sheffrin et Triest (1992). Il reflète, en outre, la nécessité pour les personnes enquêtées de fournir une image cohérente d'elles-mêmes. Ainsi, celles qui déclarent travailler au noir sont tentées de rationaliser leur déclaration en affirmant de faibles probabilités de détection. Les travailleurs au noir rapportent des croyances qui justifient leur activité souterraine, tandis que les autres énoncent des croyances qui justifient leur choix de ne pas travailler au noir. D'autre part, ces évaluations subjectives s'avèrent excessives par rapport à la réalité. Or, depuis les travaux de Kahneman et Tversky (1979) il est largement reconnu que les individus ont tendance à surestimer la probabilité de détection à laquelle ils font face, ce qui explique *in fine* la faible proportion de fraudeurs.¹⁸

Variables de contrôle sur les préférences

La théorie économique ne nous permet malheureusement pas de déterminer les variables exogènes qui influencent l'utilité marginale du travail officiel et du travail au noir. Après plusieurs expérimentations, les variables suivantes ont été incluses dans l'équation de l'utilité marginale nette des heures sur le marché officiel : le nombre de personnes dans le ménage, le nombre d'enfants de moins de six ans, une variable muette pour le sexe ($FEMME=1$), l'âge de la personne et trois variables muettes pour la situation maritale (célibataire, couple et famille monoparentale; la référence étant

¹⁸Se reporter à Andréoni, Erard et Feinstein (1998) pour une recension critique des travaux relatifs à l'évaluation subjective de la probabilité de détection et à sa corrélation avec les valeurs réelles du dispositif de lutte contre la fraude.

constituée par la catégorie *AUTRE* qui se rapporte à une personne vivant chez un membre de sa famille). L'équation de l'utilité marginale nette des heures sur le marché noir contient les mêmes variables auxquelles ont été ajoutées un certain nombre de variables subjectives, *i.e.* l'approbation ou la désapprobation des proches à l'exercice d'une activité souterraine et le pourcentage de participants au marché noir parmi les personnes de leur entourage (*PROPOREN*).

Variables explicatives des équations de salaire et de revenu virtuel

Lorsque l'on estime des fonctions d'offre de travail, le salaire horaire utilisé est souvent obtenu en divisant les revenus de travail annuels par le produit des semaines de travail et des heures hebdomadaires «normales». Toute erreur de mesure sur les heures de travail introduit alors un biais sur les élasticités salariales. Les salaires horaires utilisés ici sont moins susceptibles d'être entachés de ce type de biais puisque les répondants devaient indiquer le revenu moyen durant les séquences où ils ont occupé un emploi au cours de l'année. Les salaires ont ensuite été converti sur une base horaire.

Après plusieurs expérimentations, nous avons retenu dans les trois équations de salaire (salaire sur le marché officiel, salaire espéré sur le marché noir, salaire espéré au carré sur le marché noir) deux variables régionales muettes : la première pour la région de Montréal et la seconde pour la région du Bas-du-Fleuve; la référence étant constituée par la région de Québec. Nous avons également introduit une variable de sexe (*FEMME=1*) et des fonctions polynomiales de l'âge et du nombre d'années de scolarité. Enfin, ces mêmes variables et des termes interactifs entre l'âge et la scolarité ont été inclus dans l'équation du revenu virtuel. Il est facile de vérifier que le choix des variables exogènes retenues dans le système des six équations simultanées fait en sorte que le modèle est suridentifié.

3.5 Résultats économétriques

3.5.1 Fonction d'utilité

Les paramètres de la fonction d'utilité quadratique sont obtenus par maximisation de la fonction de vraisemblance, comme nous l'avons mentionné plus haut. Les résultats de l'estimation sont présentés au tableau 3-2.¹⁹ Les estimateurs des paramètres sont compatibles avec une fonction d'utilité strictement concave puisque les conditions globales de second ordre sont satisfaites.²⁰

¹⁹En raison de problèmes d'ordre numérique lié à la procédure d'intégration, les résultats ne tiennent compte que des trois premiers régimes.

²⁰Se reporter à l'expression (3.29).

En outre, il est important de noter que tous les coefficients associés à la forme quadratique de cette fonction (la matrice β) sont statistiquement significatifs, ce qui implique que la fiscalité et le système de répression de la fraude (probabilité de détection et amendes) ont aussi un impact significatif sur l'offre de travail officiel et au noir. Nos résultats sont donc cohérents avec la théorie du crime, initiée par Becker (1968), selon laquelle les individus averses au risque répondent à l'incertitude des sanctions en réduisant leurs activités criminelles.

Les paramètres associés aux variables de contrôle reflètent les effets de ces variables sur l'utilité marginale des heures de travail sur le marché correspondant.

Les variables de contrôle sur les heures au noir et sur le marché officiel ont le signe généralement attendu. Le statut matrimonial affecte positivement l'utilité marginale du travail sur les deux marchés. *Ceteris paribus*, celle-ci est plus élevée pour les personnes seules, les familles monoparentales et les couples comparativement aux personnes résidant chez un parent ou un proche.

Marché officiel			Marché noir		
Paramètres	Coefficients	Ecart-type	Paramètres	Coefficients	Ecart-type
α_1	0.25	-	α_2	-2.4515	0.043
β_{11}	-0.5257	0.035	Nb de pers. dans le ménage	0.0159	-0.005
β_{12}	-0.3389	0.017	Enfants (< 6ans)	-0.0020	0.007
β_{13}	-0.0987	0.006	Personne seule	0.1037	0.015
β_{22}	-0.2823	0.015	Conjoint	0.1014	-0.014
β_{23}	-0.1077	0.017	Monoparent	0.0719	0.019
β_{33}	-0.1557	0.003	Femme	-0.0461	0.018
			Age	0.1124	0.085
			Préparen	0.0030	-0.001
			Réaction de l'entourage	0.9948	-0.013
			Immoralité	-0.0016	0.004
Heures officielles / 1 000			Heures au noir / 1 000		
α_1	0.5784	0.046	α_2	-2.2059	0.094
Nb de pers. dans le ménage	-0.0271	0.009	Age	4.1547	0.828
Enfants (< 6ans)	-0.0040	0.012	Age (au carré)	-1.4446	0.255
Personne seule	0.1440	0.031	Age (au cube)	1.2676	0.243
Conjoint	0.2062	0.027	Années de scolarité	2.2584	0.265
Monoparent	-0.1286	0.037	Années de scolarité (au carré)	-2.1355	0.357
Femme	-0.1041	0.018	Années de scolarité (au cube)	0.6180	0.121
Age	0.3485	0.088	Femme	0.0098	-0.016
			Montreal	-0.0271	0.007
			Bas-du-Fleuve	-0.0052	0.011
Revenu hors-travail "virtuel" / 100			Salaire espéré au noir / 100		
Constante	0.2154	0.148	Constante	-3.0925	0.478
Age	0.1690	0.393	Age	4.5652	3.671
Age (au carré)	0.0263	0.029	Age (au carré)	-1.7337	1.157
Années de scolarité	0.2454	0.209	Age (au cube)	1.5732	1.030
Années de scolarité (au carré)	-0.1378	0.875	Années de scolarité	3.6846	2.389
Age*scolarité	-0.6099	0.420	Années de scolarité (au carré)	-3.8061	2.154
Age*scolarité (au carré)	0.5244	0.183	Années de scolarité (au cube)	1.1038	0.647
Femme	-0.0152	0.011	Femme	0.0098	0.046
Montreal	0.0271	0.009	Montreal	-0.0988	0.040
Bas-du-Fleuve	-0.0052	0.011	Bas-du-Fleuve	-0.0895	0.069
Salaire net officiel / 100			Salaire espéré au noir (au carré) / 10 000		
Constante	0.1352	0.110	Constante	-3.0925	0.478
Age	1.1191	0.390	Age	4.5652	3.671
Age (au carré)	-0.1284	0.043	Age (au carré)	-1.7337	1.157
Années de scolarité	0.0586	0.067	Age (au cube)	1.5732	1.030
Années de scolarité (au carré)	-0.0303	0.148	Années de scolarité	3.6846	2.389
Femme	-0.0376	0.014	Années de scolarité (au carré)	-3.8061	2.154
Montreal	0.0144	0.013	Années de scolarité (au cube)	1.1038	0.647
Bas-du-Fleuve	-0.0282	0.016	Femme	0.0098	0.046
			Montreal	-0.0988	0.040
			Bas-du-Fleuve	-0.0895	0.069
Log-Vraisemblance		-8 143.6			
Nb d'observations		3 083			

Tableau 3-2 : Estimation de la fonction d'utilité quadratique

Le paramètre associé à la variable du «nombre de personnes dans le ménage» est également positif. Ce coefficient mesure l'impact d'une hausse du nombre de personnes dans le ménage (enfants de plus de 6 ans et adultes) sur l'utilité marginale des heures de travail sur le marché correspondant, pour un statut marital donné. Ainsi, l'utilité marginale du travail officiel et du travail au noir augmente avec la taille du ménage. Ce résultat reflète notamment les contraintes financières auxquelles le ménage fait face lorsque le nombre de personnes au foyer s'accroît. Il confirme, par ailleurs, un des résultats obtenus par Glaetzer et Berger (1983) pour lesquels la composition familiale -définie par le nombre d'individus dans le ménage et les relations qui les lient entre eux- est un déterminant important de l'offre de travail au noir et en particulier du travail au noir domestique. Selon ces auteurs, un changement dans la composition du ménage a une répercussion immédiate sur le volume de la production souterraine, de sorte que cette caractéristique familiale aurait un pouvoir explicatif bien plus puissant que le niveau des revenus, la classe sociale ou encore l'âge.

La variable muette *FEMME* est négative, reflétant une participation au marché du travail plus faible pour les femmes. La variable correspondant au nombre d'enfants en-dessous de 6 ans n'est pas significative. Néanmoins, ce résultat peut être dû au fait que nous contraignons les paramètres à être les mêmes pour les hommes et pour les femmes. Or, il se peut qu'à l'intérieur du ménage, la présence d'enfants en âge préscolaire ait un effet différent sur chacun des conjoints. Elle pourrait affecter uniquement l'intensité au travail des femmes et rester sans effet pour les hommes, ou n'avoir qu'un effet marginal.

La variable *PROPOREN*²¹ affecte positivement l'utilité marginale du travail au noir. Cette variable apporte un premier élément de réponse quant à l'impact du voisinage sur l'offre de travail au noir. Il semblerait que le choix des heures de travail au noir soit influencé par le comportement des proches de l'individu. Les individus noueraient d'importantes relations personnelles sur le marché souterrain, ce qui expliquerait que les travailleurs au noir éprouvent certaines difficultés à retourner sur le marché officiel (Spiro, 1993). Benjamini et Maital (1983) ont été les premiers à mettre en évidence que plus un individu connaît de fraudeurs, plus il est probable qu'il fraude lui-même. La décision de travailler au noir dépend alors du nombre de travailleurs au noir dans l'entourage. Par ailleurs, les membres d'un même groupe social tendent à se comporter de façon similaire. Le groupe auquel l'individu appartient représente alors non seulement un bassin d'emplois potentiels non négligeable, mais génère également davantage de tolérance à l'égard des activités souterraines. Autrement dit, un milieu

²¹Évaluation subjective de la proportion de travailleurs au noir dans l'entourage.

constitué de nombreux travailleurs au noir a tendance à inciter ses membres à exercer le même type d'activités et, par un effet de réseau, conduit à la fois à l'accroissement des opportunités d'emploi sur ce marché et à la levée d'éventuelles réticences à l'égard des activités frauduleuses.

Ce résultat tendrait à confirmer l'hypothèse d'interdépendance des comportements selon laquelle le comportement du groupe affecte celui de ses membres. Les individus interagissent avec leur environnement lors de leurs prises de décision et les interactions sociales parmi les membres du groupe sont de nature à conditionner le choix de l'activité. Cependant, ce résultat mérite quelque peu d'être nuancé. D'une part, il est possible que la variable *PROPOREN* soit entachée d'un biais d'endogénéité. En effet, on peut s'attendre à ce que l'individu qui travaille beaucoup au noir soit porté à déclarer qu'une fraction importante de son entourage travaille au noir, de façon à justifier son comportement (effet de dissonance cognitive). D'autre part, Manski (1993) a montré que la corrélation entre la participation du groupe et celle de l'individu est sujette au problème du «miroir» (*reflection problem*). En d'autres mots, l'individu possède des caractéristiques non observables corrélées avec celles du groupe auquel il s'identifie («Qui se ressemble, s'assemble.»). Dans ce cas, la corrélation observée ne reflète pas un lien de causalité entre le comportement du groupe et celui de l'individu.

Pour vérifier cette éventualité, nous avons poursuivi l'analyse de l'interdépendance des comportements en nous fondant sur l'approche développée par Aronsson *et al.* (1999). Nous avons introduit des variables dichotomiques distinguant différents groupes de référence afin de tenir compte du problème du «miroir». Cependant, le problème fondamental de cette approche provient de la difficulté de cerner avec précision le groupe avec lequel l'individu interagit. En nous fondant sur les critères habituellement retenus pour les décisions d'offre de travail,²² nous avons sélectionné les variables d'âge, de niveau d'éducation et du lieu de résidence des individus pour constituer les groupes de référence. Néanmoins, quelle que soit la spécification -fondée sur les différentes combinaisons possibles des critères choisis- aucun des coefficients associés aux variables de groupe ne s'est révélé statistiquement significatif. En outre, l'introduction des variables de groupe n'a pas affecté le coefficient associé à la variable subjective d'évaluation de la proportion de travailleurs au noir dans l'entourage. Celui-ci est toujours demeuré statistiquement significatif.²³ Deux interprétations sont dès lors possibles. La première est que les critères de sélection pour la constitution des groupes de référence sont inappropriés pour les travailleurs au noir. La seconde est que le problème du «miroir» n'est pas important dans le cas du travail au noir.

²²Voir notamment Grodner *et al.* (2002), Weinberg *et al.* (2001) et Aronsson *et al.* (1999).

²³Les résultats détaillés des estimations des différentes spécifications sont reportés à l'annexe 3D.

De façon similaire, le sentiment d'être approuvé par son entourage dans le choix de l'activité souterraine accroît l'utilité marginale de cet emploi, comme en témoigne le coefficient associé à la variable «*réaction de l'entourage*». Alors qu'un certain opprobre social peut être associé au travail au noir, notamment pour les non participants, une attitude moins négative à l'égard de ces activités peut motiver leur choix. Les préférences individuelles en termes d'heures de travail officiel et au noir sembleraient donc influencées par le comportement et l'approbation des membres du groupe auquel l'individu appartient.

Concernant les équations de salaire, il est intéressant de noter que la plupart des variables explicatives ont le signe intuitivement attendu. Les résultats obtenus sont également conformes à ce qui est observé dans la littérature sur les salaires (Card, 1999). Ainsi, les fonctions de salaire sur le marché officiel et sur le marché noir sont concaves par rapport à l'âge, comme en témoigne le signe positif du paramètre associé à l'âge, et le signe négatif associé à cette variable au carré. On observe aussi une concavité de la relation par rapport au niveau d'éducation. Il existe, sur chacun des deux marchés, une relation positive entre le diplôme et la rémunération, mais la rentabilité de l'éducation est à rendements décroissants. Une nouvelle fois, nous retrouvons un des résultats standards de la littérature du capital humain (*e.g.*, Becker 1993; Card, 1999).

Par ailleurs, des différences salariales défavorables aux femmes apparaissent sur le marché officiel. Le paramètre associé à cette variable est, en effet, négatif, reflétant que celles-ci bénéficient de rémunérations inférieures aux hommes, *ceteris paribus*. Un résultat intéressant est que cette variable n'est pas significative au niveau de l'équation de salaire associé au marché noir. La discrimination en termes de salaire dont souffrent habituellement les femmes semblerait absente sur le marché souterrain. Ce résultat mérite d'être explicité compte tenu du fait que nous ne contrôlons pas pour le type d'activité exercée au noir. Il se pourrait, en effet, que l'absence de significativité de la variable *FEMME* soit le résultat d'une spécialisation plus marquée pour celles-ci sur le marché noir que sur le marché officiel.

Cependant, l'introduction de variables de contrôle pour le type d'emploi est susceptible de soulever des problèmes d'endogénéité et nécessiterait dès lors une correction du biais engendré. Or, l'endogénéisation du choix de l'activité souterraine, dans le cadre d'un modèle à information complète, se traduirait par une complexité accrue de l'estimation et pourrait s'avérer proprement irréalisable. C'est pourquoi, afin de juger de l'impact du secteur d'activité sur le salaire obtenu dans l'emploi souterrain, nous avons mené plusieurs estimations, par moindres carrés ordinaires, en parallèle.

L'introduction de variables dichotomiques de secteur d'activité ne permet nullement de remettre en cause le résultat obtenu. En effet, quel que soit le degré de finesse de la nomenclature des activités professionnelles utilisées dans nos estimations, la variable *FEMME* ne s'est jamais révélée statistiquement significative et son coefficient est resté inchangé. Par conséquent, notre résultat semble robuste aux différentes spécifications, ce qui tend à confirmer l'absence de discrimination salariale défavorable aux femmes sur le marché noir.

Les variables régionales nous permettent d'identifier les zones où les emplois sont les plus rémunérateurs. Ainsi, la région de Montréal est celle qui offre les rémunérations les plus élevées sur le marché officiel alors que, sur le marché noir, les salaires les plus importants sont versés dans la région de Québec. Mais, dans tous les cas, les salaires (officiels et au noir) obtenus dans la région du Bas-du-Fleuve sont plus faibles.

Enfin, la matrice de corrélation des termes d'erreur nous apporte un éclairage quant à la nature des relations entre les différentes équations du modèle. Ainsi, le *tableau 3-3* semble indiquer que les heures de travail officielles sont négativement corrélées au salaire espéré du marché noir (covariance négative entre ε_1 et ε_5). De la même manière, l'équation des heures au noir est corrélée négativement avec le salaire net officiel (covariance entre ε_2 et ε_4). La relation négative entre ε_2 et ε_3 suggère, par ailleurs, que les heures non déclarées diminuent lorsque le revenu hors-travail augmente. Enfin, la covariance entre ε_1 et ε_2 témoigne d'une corrélation positive entre les heures officielles et les heures au noir. Ce résultat tend à réfuter l'hypothèse de substituabilité parfaite des heures de travail.

Tableau 3-3 :
Matrice de corrélation des termes d'erreur

	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6
ε_1	0.17381					
ε_2	0.01452	1.18217				
ε_3	0.05432	-0.26516	0.11294			
ε_4	-0.01230	-0.41073	0.08317	0.18376		
ε_5	-0.04827	1.14037	-0.27354	-0.39779	1.13080	
ε_6	-0.06658	1.69188	-0.40925	-0.58964	1.67535	2.49756
Équation	Heures officielles h_1	Heures au noir h_2	Revenu virtuel y^v	Salaire net officiel $w_1(1-\tau)$	Salaire espéré au noir Ew_2	Salaire espéré au noir au carré $E(w_2^2)$

Afin de conclure sur l'impact de la fiscalité sur l'offre de travail officiel et au noir, il convient d'examiner, dans un premier temps, les élasticités conditionnelles et non conditionnelles aux différents régimes. Puis, dans un second temps, nous analyserons la structure des préférences individuelles. Nous testerons notamment la validité de

l'hypothèse de substitutabilité parfaite des heures de travail sur les marchés officiel et souterrain, de même que celle d'additivité séparable entre la consommation et le nombre total d'heures de travail.

3.5.2 Elasticités

Le calcul des élasticités est fondé sur l'espérance des heures de travail, non conditionnelle à la présence de l'individu dans un régime particulier. En effet, comme nous raisonnons dans le cadre de la méthode du maximum de vraisemblance à information complète, nous devons prendre en compte la possibilité pour l'individu de transiter d'un régime à un autre. Autrement dit, nous devons permettre un ajustement des heures de travail aux variations des différentes variables.

Les élasticités (η_{h_1} et η_{h_2}) des heures de travail sur chaque marché, par rapport aux variables X'_i , sont déterminées de la manière suivante :

$$\eta_{h_1} = \frac{\partial E(h_1)}{\partial X'_i} \times \frac{X'_i}{E(h_1)} \quad (3.52)$$

$$\eta_{h_2} = \frac{\partial E(h_2)}{\partial X'_i} \times \frac{X'_i}{E(h_2)} \quad (3.53)$$

où $E(h_1)$ et $E(h_2)$ représentent respectivement l'espérance non conditionnelle des heures de travail officiel et au noir. Elles sont définies ainsi :

. pour le marché officiel

$$E(h_1) = P_1(h_1 > 0, h_2 > 0).E(h_1|Y, X', h_1 > 0, h_2 > 0) + P_2(h_1 > 0, h_2 = 0).E(h_1|Y, X', h_1 > 0, h_2 = 0) \quad (3.54)$$

. et pour le marché noir

$$E(h_2) = P_1(h_1 > 0, h_2 > 0).E(h_2|Y, X', h_1 > 0, h_2 > 0) + P_3(h_1 = 0, h_2 > 0).E(h_2|Y, X', h_1 = 0, h_2 > 0) \quad (3.55)$$

P_1, P_2 et P_3 désignent respectivement la probabilité de travailler sur les deux marchés (régime 1), sur le marché officiel uniquement (régime 2) et sur le marché noir uniquement (régime 3). $E(h_1|.)$ et $E(h_2|.)$ caractérisent l'espérance conditionnelle des heures de travail sur chaque marché.

Les dérivées analytiques des expressions (3.54) et (3.55) sont :

$$\frac{\partial E(h_1)}{\partial X'_i} = P_1 \frac{\partial E(h_1, h_2|Y, X', h_1 > 0, h_2 > 0)}{\partial X'_i} + E(h_1|Y, X', h_1 > 0, h_2 > 0) \frac{\partial P_1}{\partial X'_i} + P_2 \frac{\partial E(h_1, h_2|Y, X', h_1 > 0, h_2 > 0)}{\partial X'_i} + E(h_1|Y, X', h_1 > 0, h_2 > 0) \frac{\partial P_2}{\partial X'_i} \quad (3.56)$$

$$\begin{aligned} \frac{\partial E(h_2)}{\partial X'_i} = & P_1 \frac{\partial E(h_1, h_2 | Y, X', h_1 > 0, h_2 > 0)}{\partial X'_i} + E(h_2 | Y, X', h_1 > 0, h_2 > 0) \frac{\partial P_1}{\partial X'_i} \\ & + P_3 \frac{\partial E(h_1, h_2 | Y, X', h_1 > 0, h_2 > 0)}{\partial X'_i} + E(h_2 | Y, X', h_1 > 0, h_2 > 0) \frac{\partial P_3}{\partial X'_i} \end{aligned} \quad (3.57)$$

Afin de déterminer les élasticités des heures travaillées sur chaque marché, nous devons donc expliciter chacun des termes des expressions (3.52) à (3.57).

Pour le régime 1 ($h_1 > 0, h_2 > 0$), l'espérance des heures de travail sur le marché officiel, conditionnelle à $h_1 > 0$ et $h_2 > 0$, est définie de la manière suivante :

$$E(h_1 | Y, X', h_1 > 0, h_2 > 0) = \int_0^\infty \int_0^\infty h_1 f(h_1, h_2 | Y, X') dh_1 dh_2$$

où y et x désignent respectivement les variables endogènes et les variables exogènes, et $h_1 = g_1(Y, X', \varepsilon_1, \varepsilon_2) = \frac{\gamma \xi_1 - \beta \xi_2 - \gamma \varepsilon_1 + \beta \varepsilon_2}{\alpha \gamma - \beta^2}$

Les variables $\alpha, \beta, \gamma, \xi_1$ et ξ_2 étant définies, comme précédemment, par :

$$\begin{aligned} \alpha &= \beta_{11} + 2\beta_{13}w_1 + \beta_{33}w_1^2 \\ \beta &= \beta_{12} + \beta_{13}Ew_2 + \beta_{23}w_1 + \beta_{33}w_1Ew_2 \\ \gamma &= \beta_{22} + 2\beta_{23}Ew_2 + \beta_{33}Ew_2^2 \\ \xi_1 &= -\bar{\alpha}_1 - X'_1\delta_1 - \alpha_3w_1 - \beta_{13}y^v - \beta_{33}y^vw_1 \\ \xi_2 &= -\bar{\alpha}_2 - X'_2\delta_2 - \alpha_3Ew_2 - \beta_{23}y^v - \beta_{33}y^vEw_2 \end{aligned}$$

L'espérance conditionnelle s'écrit :

$$E(h_1 | Y, X', h_1 > 0, h_2 > 0) = \frac{\int_{\varepsilon_1^c}^\infty \int_{\varepsilon_2^c(\varepsilon_1)}^{\bar{\varepsilon}_2^c(\varepsilon_1)} g_1 f(\varepsilon_1, \varepsilon_2) d\varepsilon_1 d\varepsilon_2}{1 - F(\varepsilon_1^c, \varepsilon_2^c)}$$

$$\Leftrightarrow E(h_1 | Y, X', h_1 > 0, h_2 > 0) = \frac{\int_{\varepsilon_1^c}^\infty \int_{\varepsilon_2^c(\varepsilon_1)}^{\bar{\varepsilon}_2^c(\varepsilon_1)} g_1 f(\varepsilon_1, \varepsilon_2) d\varepsilon_1 d\varepsilon_2}{\int_{\varepsilon_1^c}^\infty \int_{\varepsilon_2^c(\varepsilon_1)}^{\bar{\varepsilon}_2^c(\varepsilon_1)} f(\varepsilon_1, \varepsilon_2) d\varepsilon_1 d\varepsilon_2}$$

où ε_1^c est obtenu en posant $h_1 = 0$ et $h_2 = 0$, puis en résolvant le système suivant :

$$\begin{cases} \varepsilon_1 = \frac{\gamma \xi_1 - \beta \xi_2 + \beta \varepsilon_2}{\gamma} \\ \varepsilon_2 = \frac{-\beta \xi_1 + \alpha \xi_2 + \beta \varepsilon_1}{\alpha} \end{cases}$$

Les bornes d'intégration sont définies en fonction des pentes respectives de h_1 et h_2 et de leur position relative l'une par rapport à l'autre.²⁴

Par conséquent, $\varepsilon_1^c = \xi_1$.

²⁴Se reporter à l'annexe 3E pour plus de détails.

$\varepsilon_2^c(\varepsilon_1)$ est défini par $h_2 = 0$, de sorte que $\varepsilon_2^c(\varepsilon_1) = \frac{-\beta\xi_1 + \alpha\xi_2}{\alpha} + \frac{\beta}{\alpha}\varepsilon_1$
 et $\bar{\varepsilon}_2^c(\varepsilon_1)$ est défini par $h_1 = 0$, de sorte que $\bar{\varepsilon}_2^c(\varepsilon_1) = \frac{-\gamma\xi_1 + \beta\xi_2}{\beta} + \frac{\gamma}{\beta}\varepsilon_1$.

De la même manière, l'espérance des heures de travail sur le marché noir, conditionnelle à $h_1 > 0$ et $h_2 > 0$, s'écrit de la manière suivante :

$$E(h_2|Y, X', h_1 > 0, h_2 > 0) = \int_0^\infty \int_0^\infty h_2 f(h_1, h_2|Y, X') dh_1 dh_2$$

où $h_2 = g_2(Y, X', \varepsilon_1, \varepsilon_2) = \frac{-\beta\xi_1 + \alpha\xi_2 + \beta\varepsilon_1 - \alpha\varepsilon_2}{\alpha\gamma - \beta^2}$.

L'espérance conditionnelle des heures au noir est donc définie par :

$$\Leftrightarrow E(h_2|Y, X', h_1 > 0, h_2 > 0) = \frac{\int_{\varepsilon_1^c}^\infty \int_{\varepsilon_2^c(\varepsilon_1)}^{\bar{\varepsilon}_2^c(\varepsilon_1)} g_2 f(\varepsilon_1, \varepsilon_2) d\varepsilon_1 d\varepsilon_2}{1 - F(\varepsilon_1^c, \varepsilon_2^c)}$$

$$\Leftrightarrow E(h_2|Y, X', h_1 > 0, h_2 > 0) = \frac{\int_{\varepsilon_1^c}^\infty \int_{\varepsilon_2^c(\varepsilon_1)}^{\bar{\varepsilon}_2^c(\varepsilon_1)} g_2 f(\varepsilon_1, \varepsilon_2) d\varepsilon_1 d\varepsilon_2}{\int_{\varepsilon_1^c}^\infty \int_{\varepsilon_2^c(\varepsilon_1)}^{\bar{\varepsilon}_2^c(\varepsilon_1)} f(\varepsilon_1, \varepsilon_2) d\varepsilon_1 d\varepsilon_2}$$

Pour le régime 2 ($h_1 > 0, h_2 = 0$), l'espérance des heures de travail sur le marché officiel, conditionnelle à $h_1 > 0$ et $h_2 = 0$, s'écrit :

$$E(h_1|Y, X', h_1 > 0, h_2 = 0) = \int_0^\infty h_1 f(h_1|Y, X') dh_1$$

où $h_1 = g_1(Y, X', \varepsilon_1) = \frac{\xi_1 - \varepsilon_1}{\alpha}$.

L'espérance conditionnelle des heures officielles est définie par :

$$E(h_1|Y, X', h_1 > 0, h_2 = 0) = \frac{\int_{\varepsilon_1^c}^\infty g_1 f(\varepsilon_1) d\varepsilon_1}{1 - F(\varepsilon_1^c)}$$

$$\Leftrightarrow E(h_1|Y, X', h_1 > 0, h_2 = 0) = \frac{\int_{\varepsilon_1^c}^\infty g_1 f(\varepsilon_1) d\varepsilon_1}{\int_{\varepsilon_1^c}^\infty f(\varepsilon_1) d\varepsilon_1}$$

$$\Leftrightarrow E(h_1|Y, X', h_1 > 0, h_2 = 0) = \left(\frac{\xi_1}{\alpha} - \frac{1}{\alpha} \right) \frac{\int_{\varepsilon_1^c}^\infty \varepsilon_1 f(\varepsilon_1) d\varepsilon_1}{\int_{\varepsilon_1^c}^\infty f(\varepsilon_1) d\varepsilon_1}$$

où ε_1^c est donnée par $h_1 > 0$, de sorte que $\varepsilon_1^c < -\bar{\alpha}_1 - X'_1 \delta_1 - \alpha_3 w_1 - \beta_{13} y^v - \beta_{33} y^v w_1$.
 Autrement dit, $\varepsilon_1^c < \xi_1$.

Enfin, pour le régime 3 ($h_1 = 0, h_2 > 0$), l'espérance des heures de travail sur le marché noir, conditionnelle à $h_1 = 0$ et $h_2 > 0$, s'écrit de la manière suivante :

$$E(h_2|y, x, h_1 = 0, h_2 > 0) = \int_0^\infty h_2 f(h_2|Y, X') dh_2$$

où $h_2 = g_2(Y, X', \varepsilon_2) = \frac{\xi_2 - \varepsilon_2}{\gamma}$.

L'espérance conditionnelle des heures au noir est définie par :

$$\begin{aligned} \Leftrightarrow E(h_2|Y, X', h_1 = 0, h_2 > 0) &= \frac{\int_{\varepsilon_2^c}^{\infty} g_2 f(\varepsilon_2) d\varepsilon_2}{1 - F(\varepsilon_2^c)} \\ \Leftrightarrow E(h_2|Y, X', h_1 = 0, h_2 > 0) &= \frac{\int_{\varepsilon_2^c}^{\infty} g_2 f(\varepsilon_2) d\varepsilon_2}{\int_{\varepsilon_2^c}^{\infty} f(\varepsilon_2) d\varepsilon_2} \\ \Leftrightarrow E(h_2|Y, X', h_1 = 0, h_2 > 0) &= \left(\frac{\xi_1}{\alpha} - \frac{1}{\alpha} \right) \frac{\int_{\varepsilon_2^c}^{\infty} \varepsilon_2 f(\varepsilon_2) d\varepsilon_2}{\int_{\varepsilon_2^c}^{\infty} f(\varepsilon_2) d\varepsilon_2} \end{aligned}$$

où ε_2^c est déterminé par $h_2 > 0$, de sorte que $\varepsilon_2^c < -\bar{\alpha}_2 - X_2' \delta_2 - \alpha_3 Ew_2 - \beta_{23} y^v - \beta_{33} y^v Ew_2$. Autrement dit, $\varepsilon_2^c < \xi_2$.

Le tableau 3-4 décrit les heures espérées conditionnelles et non conditionnelles moyennes sur chaque marché du travail.

Tableau 3-4 :
Heures espérées moyennes

	Valeurs prédites	Valeurs observées	Nombre d'observations
. Non-Conditionnelles :			
Heures officielles	1,556.18	1,513.49	3 083
Heures au noir	165.36	30.02	
. Conditionnelles :			
REGIME 1 :			
Heures officielles	829.58	1,083.65	104
Heures au noir	1,063.10	343.57	
REGIME 2 :			
Heures officielles	1,593.56	1,570.06	2 893
REGIME 3 :			
Heures au noir	1,669.20	719.58	86

De façon générale, les valeurs prédites des heures officielles sont extrêmement proches de celles qui sont observées dans notre échantillon. En revanche, les heures espérées dans l'économie souterraine sont, en moyenne, moins bien prédites. Les heures au noir conditionnelles ont tendance à être sur-estimées par rapport à leurs valeurs réelles, tandis que les heures espérées non conditionnelles ne représentent qu'un tiers des heures moyennes de l'échantillon. Ce relatif manque de précision sur le marché

noir est, en grande partie, imputable à la faible proportion des travailleurs au noir comparativement aux travailleurs officiels, ce qui rend les prédictions plus délicates. Néanmoins, nos résultats demeurent encourageants.

Les tableaux 3-5 à 3-8 ci-dessous présentent les élasticités non conditionnelles à la participation au marché officiel et/ou au marché noir, ainsi que les élasticités conditionnelles à chacun des quatre régimes.²⁵ L'ensemble des tableaux décrivent, non pas les élasticités évaluées à la moyenne des variables, mais la moyenne des élasticités individuelles. Cette mesure est préférable puisqu'elle permet de mieux tenir compte des caractéristiques individuelles.

Tableau 3-5 :
Elasticités non conditionnelles

	Elasticités	Ecart-type
Heures officielles		
(1) η_{h_1, y^v}	-0.025980	0.001338 ***
(2) $\eta_{h_1, \tau}$	0.022480	0.002508 ***
(3) η_{h_1, W_1}	-0.025243	0.002560 ***
(4) η_{h_1, w_1}	-0.025257	0.002556 ***
(5) η_{h_1, W_2}	0.000053	0.000083
(6) $\eta_{h_1, P}$	-0.000049	0.000018 **
(7) $\eta_{h_1, \theta}$	-0.000003	0.000007
(8) η_{h_1, Ew_2}	0.000041	0.000069
(9) η_{h_1, Ew_2^2}	0.000111	0.000035 ***
Heures au noir		
(10) η_{h_2, y^v}	-0.005644	0.000712 ***
(11) $\eta_{h_2, \tau}$	0.004172	0.001830 *
(12) η_{h_2, W_1}	-0.001032	0.000533 *
(13) η_{h_2, w_1}	-0.001221	0.005019 **
(14) η_{h_2, W_2}	-0.018437	0.008387 *
(15) $\eta_{h_2, P}$	0.003671	0.001354 **
(16) $\eta_{h_2, \theta}$	0.001424	0.000652 *
(17) η_{h_2, Ew_2}	-0.014735	0.007034 *
(18) η_{h_2, Ew_2^2}	-0.005330	0.000265 ***

*** Statistiquement significatif à 1%

** Statistiquement significatif à 5%

* Statistiquement significatif à 10%

L'élasticité des heures de travail par rapport au revenu virtuel (soit η_{h_i, y^v}) décrit l'effet revenu (virtuel). Ce dernier est négatif sur les deux marchés du marché, ce

²⁵ Les élasticités sont calculées pour une variation de 0.0001 point de la variable considérée. Par ailleurs, l'évaluation des différentes intégrales est réalisée à partir de 40 points de calcul.

qui signifie qu'un accroissement du revenu a pour conséquence de diminuer l'intensité du travail. Nous retrouvons, dès lors, un des résultats standards de la littérature de l'offre de travail.²⁶ L'effet revenu est cependant modéré, en particulier pour les heures non déclarées. Les variations du revenu virtuel semblent donc exercer une influence limitée sur le volume horaire de l'activité souterraine. Concernant l'économie officielle, l'effet revenu obtenu est plus important que celui qui est observé empiriquement (*e.g.* Blomquist, Hansson et Brusewitz, 1990; MaCurdy *et al.*, 1990; ou encore Van Soest *et al.*, 1990). En effet, nous obtenons une élasticité-revenu de -0.026 tandis que dans ces modèles, pour lesquels l'existence d'une économie souterraine n'est pas prise en compte, l'élasticité-revenu est proche de -0.001. L'opportunité d'une activité non déclarée semble, par conséquent, accentuer la désincitation au travail officiel lorsque le revenu augmente. Ce résultat suggère alors qu'ignorer le marché noir conduit à un biais à la baisse potentiellement important dans l'évaluation de l'élasticité-revenu des heures de travail officielles.

Le taux marginal d'imposition a, par ailleurs, un impact positif sur les heures de travail au noir. Il convient de noter toutefois que l'offre de travail dissimulé reste, encore une fois, peu élastique. L'élasticité des heures officielles par rapport au taux d'impôt marginal est également positive, de sorte qu'un accroissement du fardeau fiscal entraîne une augmentation du nombre total d'heures travaillées. Ce résultat, quelque peu surprenant pour le marché officiel, pourrait cependant s'expliquer par une courbe d'offre de travail à rebroussement pour laquelle l'effet revenu dominerait l'effet substitution. L'accroissement de la pression fiscale a pour conséquence de diminuer le revenu de l'individu, de sorte que celui-ci choisirait de travailler davantage pour maintenir son niveau de revenu. Ce résultat semble, en outre, être corroboré par les élasticité des heures officielles par rapport aux différentes formes de salaires (lignes 3 et 4). En effet, le *tableau 3-5* indique que les heures officielles diminuent lorsque le salaire brut et le salaire net augmentent. Nous sommes alors typiquement en présence d'une courbe d'offre de travail officiel à rebroussement.²⁷ De la même manière, la relation entre le salaire et le volume horaire de travail sur le marché noir ne semble pas non plus linéaire. Les élasticité des heures au noir par rapport au salaire brut et par rapport au salaire espéré sont négatives, de sorte qu'une augmentation de la rémunération au noir a pour effet de réduire l'activité souterraine. Ce résultat pourrait paraître étonnant, mais il reflète certainement les contraintes de dissimulation auxquelles est confronté tout travailleur au noir. A mesure que les revenus de l'activité souterraine augmentent, il s'avère de plus en plus difficile de ne pas éveiller les soupçons du fisc. Un tel argument

²⁶ Voir notamment Ashenfelter et Card (1999).

²⁷ Un résultat similaire a été obtenu par Flood et MaCurdy (1992), Kaiser *et al.* (1992), et Kuismanen (1997). Dans ces travaux, l'élasticité-salaire est respectivement de -0.025, -0.004 et -0.01.

a été étayé empiriquement par plusieurs études (*e.g.* Lemieux, 1988 ; Lemieux, Fortin et Fréchette, 1994).

Au regard des élasticités croisées, il est intéressant de noter que le volume horaire de l'emploi sur le marché officiel n'est pas affecté par les conditions salariales du marché noir. Les élasticités des heures officielles par rapport au salaire brut au noir et par rapport au salaire espéré dissimulé ne sont pas statistiquement significatives. En revanche, un accroissement du salaire officiel (à la fois brut et net d'impôts) entraîne une diminution des heures travaillées sur le marché noir. Ces résultats sont probablement dus en partie au fait que les heures de travail au noir sont plus flexibles que celles de l'activité officielle. Ils suggèrent, en outre, l'existence d'une certaine substituabilité entre les heures officielles et les heures non déclarées. Néanmoins, l'élasticité des heures au noir par rapport au salaire officiel demeure très faible, ce qui semble indiquer que la substitution n'est qu'imparfaite.²⁸

Il convient de souligner que l'intensité du travail officiel apparaît très peu sensible aux paramètres de répression de la fraude, *i.e.* risque de détection et montant de l'amende. L'élasticité des heures officielles est, d'une part, minime par rapport à la probabilité de contrôle, et d'autre part, celle-ci n'est pas statistiquement significative par rapport au taux de pénalité. Étonnamment, ces paramètres semblent influencer positivement les heures travaillées dans l'économie souterraine. Les élasticités demeurent toutefois assez faibles. De tels résultats méritent d'être nuancés. En effet, il s'agit d'élasticités non conditionnelles au régime dans lequel l'individu se situe. Cette mesure est certes très utile puisqu'elle permet de tenir compte des changements de régimes opérés par les individus à la suite d'une variation marginale des paramètres considérés. Mais, sa pertinence pour le marché noir est limitée par le nombre relativement faible de travailleurs sur ce marché par rapport à l'ensemble de l'échantillon. En conséquence les résultats pour les heures non déclarées sont potentiellement moins satisfaisants. Cette hypothèse semble confirmée par le *tableau 3-4* où les heures espérées au noir non conditionnelles sont largement sous-estimées.

Le *tableau 3-6* présente les élasticités propres au régime 1. Conformément à ce qui est observé dans la littérature (*e.g.* Hausman, 1985), l'effet revenu sur le marché officiel est négatif et de faible ampleur. Nos résultats sont similaires à ceux de Lacroix et Fortin (1992), dans la mesure où ils obtiennent un effet revenu de -0.018 sur le marché officiel et de -0.0045 sur le marché noir. Les lignes 2 et 12 présentent les élasticités par rapport à τ . Un changement dans le taux marginal d'imposition affecte non seulement le salaire officiel net, mais également le taux de pénalité puisque ce dernier est défini

²⁸Nous testerons cette assertion de façon plus rigoureuse au paragraphe 3.5.3.

par $\theta = (1 + \lambda)\tau$. Afin de faciliter la comparaison avec les travaux de Lacroix et Fortin (1992), nous avons calculé l'élasticité, non pas par rapport à τ , mais par rapport à $(1 - \tau)$. De manière plus formelle, ces élasticités sont données par l'équation suivante :

$$\eta_{h_i, (1-\tau)} = \eta_{h_i, W_i} - \eta_{h_i, \theta} [(1 - \tau)/\tau], \quad i = 1, 2.$$

Cette équation montre que $\eta_{h_i, (1-\tau)} \neq \eta_{h_i, W_i}$ même en l'absence d'illusion fiscale.²⁹ Le tableau 3-6 démontre que l'écart entre $\eta_{h_1, (1-\tau)}$ et η_{h_1, W_1} est cependant modéré, dans la mesure où $\eta_{h_i, \theta}$ est faible.

Tableau 3-6 :
Elasticités conditionnelles (régime 1)

	Elasticités	Ecart-type
Heures officielles		
(1) η_{h_1, y^v}	-0.011388	0.000610 ***
(2) $\eta_{h_1, \tau}$	0.031702	0.002751 ***
(3) $\eta_{h_1, (1-\tau)}$	-0.033023	0.002453 ***
(4) η_{h_1, W_1}	-0.031783	0.002536 ***
(5) η_{h_1, w_1}	-0.032228	0.002790 ***
(6) η_{h_1, W_2}	-0.007171	0.001988 ***
(7) $\eta_{h_1, P}$	0.000741	0.000525
(8) $\eta_{h_1, \theta}$	0.000694	0.000265 ***
(9) η_{h_1, Ew_2}	-0.006531	0.001694 ***
(10) η_{h_1, Ew_2^2}	0.003525	0.000172 ***
Heures au noir		
(11) η_{h_2, y^v}	-0.013448	0.000712 ***
(12) $\eta_{h_2, \tau}$	0.020148	0.001830 ***
(13) $\eta_{h_2, (1-\tau)}$	-0.019311	0.003701 ***
(14) η_{h_2, W_1}	-0.005808	0.000533 *
(15) η_{h_2, w_1}	-0.006517	0.005019 **
(16) η_{h_2, W_2}	-0.005975	0.008387 ***
(17) $\eta_{h_2, P}$	0.017185	0.001354 ***
(18) $\eta_{h_2, \theta}$	0.005405	0.000652 ***
(19) η_{h_2, Ew_2}	-0.045146	0.007034 ***
(20) η_{h_2, Ew_2^2}	-0.008666	0.000265 ***

*** Statistiquement significatif à 1%
** Statistiquement significatif à 5%

* Statistiquement significatif à 10%

²⁹ En ce sens, tester l'existence d'illusion fiscale à partir de modèles qui ignorent l'économie souterraine (e.g. Mroz, 1987) pourrait conduire à des résultats biaisés.

L'élasticité des heures au noir par rapport au taux d'impôt marginal (τ) est significativement positive, ce qui signifie qu'une augmentation de la pression fiscale conduit à accroître l'activité souterraine des travailleurs présents sur ce marché. Pour les mêmes raisons exposées précédemment, nous observons des élasticité propres et croisées négatives. Les heures officielles des personnes qui opèrent simultanément sur les deux marchés paraissent, une nouvelle fois, relativement inélastiques aux paramètres de détection de la fraude, tandis qu'ils continuent à affecter positivement les heures au noir. A ce titre, il est nécessaire de rappeler que la probabilité de détection et l'amende encourue reposent sur une évaluation subjective et sont dépendantes de la forme fonctionnelle retenue. Dès lors, il se pourrait que des erreurs de mesure inhérentes aux valeurs subjectives (même indirectement endogénéisées) puissent se produire ou que la spécification retenue ne soit pas appropriée pour rendre compte correctement de l'impact des paramètres fiscaux. Il se pourrait enfin que les individus, anticipant un redressement, choisissent de travailler davantage afin de se couvrir contre les pertes.

Le *tableau 3-7* présente les élasticité propres au régime 2 ; celles-ci correspondant aux élasticité d'offre de travail habituelles. L'effet revenu est semblable à ce qui est observé empiriquement. A titre d'exemple, Lacroix (1990) obtient une élasticité-revenu de -0.029 et Lacroix et Fortin (1992) aboutissent à une valeur de -0.054. En revanche, l'élasticité par rapport au salaire est négative, dans notre cas, alors que pour ces auteurs, elle est positive. Par conséquent, le fait d'estimer simultanément les heures de travail et les salaires sur chaque marché, plutôt que de recourir à une estimation en deux étapes, tend à démontrer que les travailleurs officiels privilégient le loisir lorsque leur emploi leur alloue davantage de revenus.

Tableau 3-7 :
Elasticités conditionnelles (régime 2)

	Elasticités	Ecart-type
Heures officielles		
(1) η_{h_1, y^v}	-0.023658	0.001311 ***
(2) $\eta_{h_1, \tau}$	0.024240	0.002367 ***
(3) η_{h_1, w_1}	-0.026863	0.002468 ***
(4) η_{h_1, w_1}	-0.026863	0.002469 ***

*** Statistiquement significatif à 1%

** Statistiquement significatif à 5%

* Statistiquement significatif à 10%

Enfin, le *tableau 3-8* présente les élasticité propres au régime 3. Nos résultats révèlent une relation positive entre la taux d'imposition marginal et le nombre d'heures de travail réalisées sur le marché noir. Cela tend à confirmer un des résultats standards de la littérature sur l'offre de travail au noir (e.g. Isachsen, Samuelson et Strom, 1983).

L'effet-revenu est identique à celui de Lacroix (1990) et, à l'instar de Lemieux, l'élasticité par rapport au salaire (brut et espéré) sur le marché noir est négative. Cette relation concave serait le résultat de la volonté de ne pas éveiller les soupçons du fisc.

Tableau 3-8 :
Elasticités conditionnelles (régime 3)

	Elasticités	Ecart-type
Heures au noir		
(1) $\eta_{h_2, y}$	-0.011090	0.000184 ***
(2) $\eta_{h_2, \tau}$	0.007185	0.001436 ***
(3) η_{h_2, W_2}	-0.051738	0.008878 ***
(4) $\eta_{h_2, P}$	0.007969	0.001428 ***
(5) $\eta_{h_2, \theta}$	0.003322	0.000685 ***
(6) η_{h_2, Ew_2}	-0.043757	0.007464 ***
(7) η_{h_2, Ew_2^2}	-0.005420	0.000329 ***

*** Statistiquement significatif à 1%
** Statistiquement significatif à 5%

* Statistiquement significatif à 10%

3.5.3 Tests sur les préférences

Notre modèle permet de tester empiriquement la validité de l'hypothèse de substituabilité parfaite entre les heures de travail sur le marché officiel (h_1) et sur le marché noir (h_2). Il nous permet également de tester l'hypothèse de séparabilité additive entre la consommation (C) et le nombre total d'heures de travail ($h_1 + h_2$). Ces hypothèses sont, en effet, très fréquentes dans la littérature sur la fraude fiscale.

La forme générale de la fonction d'utilité quadratique est la suivante :

$$U(h_1, h_2, C) = \alpha_1 h_1 + \alpha_2 h_2 + \alpha_3 C + \frac{1}{2} \beta_{11} h_1^2 + \beta_{12} h_1 h_2 + \beta_{13} h_1 C + \frac{1}{2} \beta_{22} h_2^2 + \beta_{23} h_2 C + \frac{1}{2} \beta_{33} C^2. \quad (3.58)$$

La substitution parfaite entre h_1 et h_2 signifie que (3.58) peut être réécrite de la façon suivante :

$$U(h_1, h_2, C) = \gamma_1 (h_1 + h_2) + \gamma_2 C + \frac{1}{2} \delta_{11} (h_1 + h_2)^2 + \delta_{12} (h_1 + h_2) C + \frac{1}{2} \delta_{22} C^2. \quad (3.59)$$

Dans ce cas, la désutilité marginale de chaque activité est la même. La comparaison de (3.58) et (3.59) nous indique les restrictions imposées sur les paramètres par la formulation de l'hypothèse de substituabilité parfaite :

$$(i) \alpha_1 = \alpha_2; \quad (ii) \beta_{11} = \beta_{22} = 2\beta_{12}; \quad (iii) \beta_{13} = \beta_{23}.$$

Afin de tenir compte des caractéristiques individuelles, nous avons posé $\alpha_1 = \bar{\alpha}_1 + X'_1 \delta_1 + \varepsilon_1$ et $\alpha_2 = \bar{\alpha}_2 + X'_2 \delta_2 + \varepsilon_2$. Or, les variables incluses dans X'_1 ne sont pas forcément les mêmes que celles incluses dans X'_2 . Nous pouvons, dès lors, décomposer α_1 et α_2 ainsi :

$$\begin{aligned}\alpha_1 &= \bar{\alpha}_1 + \bar{X}' \delta_1 + \tilde{X}'_1 \tilde{\delta}_1 + \varepsilon_1, \\ \alpha_2 &= \bar{\alpha}_2 + \bar{X}' \delta_2 + \tilde{X}'_2 \tilde{\delta}_2 + \varepsilon_2,\end{aligned}$$

où \bar{X}' est un vecteur de variables communes aux deux activités, \tilde{X}'_1 et \tilde{X}'_2 sont des vecteurs de variables de contrôle spécifiques aux activités officielles et au noir respectivement. Enfin, $\bar{\alpha}_i$ est une constante (avec $i = 1, 2$). Par conséquent, la condition $\alpha_1 = \alpha_2$ implique : $\bar{\alpha}_1 = \bar{\alpha}_2$, $\bar{\delta}_1 = \bar{\delta}_2$ et $\tilde{\delta}_1 = \tilde{\delta}_2 = 0$. L'ensemble des restrictions découlant de l'hypothèse de substitution parfaite se résume à :

$$\begin{aligned}(i) \quad \bar{\alpha}_1 &= \bar{\alpha}_2; \\ (ii) \quad \beta_{11} &= \beta_{22} = 2\beta_{12}; \\ (iii) \quad \beta_{13} &= \beta_{23}; \\ (iv) \quad \bar{\delta}_1 &= \bar{\delta}_2 \\ (v) \quad \tilde{\delta}_1 &= \tilde{\delta}_2 = 0\end{aligned}$$

Pour tester l'hypothèse de substituabilité parfaite entre h_1 et h_2 , nous avons effectué un test de Wald.³⁰ Brièvement, le test de Wald peut être formulé ainsi. Soit $\hat{\beta}$ l'estimateur du vecteur non contraint β . Les restrictions (i) à (v) peuvent être représentées par $R\hat{\beta} = q$. La statistique W de Wald est alors définie comme :

$$W = (R\hat{\beta} - q)' [Rv(\hat{\beta})R']^{-1} (R\hat{\beta} - q) \quad (3.60)$$

où $V(\hat{\beta})$ est la matrice de variance-covariance de β évaluée à $\hat{\beta}$. On peut montrer que $W \sim \chi^2_{(v)}$ avec $v = \text{rang}(R)$.

La statistique W , calculée à partir des paramètres du *tableau 3-2*, est de $W = 23\,354.8$. Cette valeur doit être comparée à la valeur critique $\chi^2_{0.05}(13) = 22.36$.³¹ Par conséquent, l'hypothèse de substitution parfaite entre les heures de travail est fortement rejetée.

Par ailleurs, l'hypothèse de séparabilité additive revient à imposer $\beta_{12} = \beta_{13} = \beta_{23} = 0$. Un test de Wald a également été effectué et la statistique W nous donne

³⁰ Voir Judge *et al.* (1980) ou Wooldridge (2001) pour plus de détails.

³¹ $\text{Rang}(R) = 13$. En effet, il y a 5 restrictions sur les paramètres β , et $\dim(\bar{\delta}_1) = 7$. Par ailleurs, $\dim(\tilde{\delta}_1) = 0$, $\dim(\tilde{\delta}_2) = 3$, de sorte que nous ne testons qu'une seule restriction additionnelle, *i.e.* $\tilde{\delta}_2 = 0$.

$W = 144.4$. La comparaison de cette dernière avec $\chi_{0.05}^2(3) = 7.82$ nous permet de conclure que nos données ne supportent pas l'hypothèse de séparabilité additive.³²

3.6 Conclusion

Dans ce chapitre, nous avons examiné la relation entre la fiscalité et l'offre individuelle de travail sur les marchés officiel et/ou souterrain. Un rapide survol de la littérature nous a permis de cerner les limites des différentes approches proposées pour l'étude des comportements de travail sur chacun des marchés.

Dès lors, nous avons construit et estimé un modèle d'offre de travail sur les marchés officiel et souterrain en présence de la fiscalité, des programmes sociaux et des politiques de contrôle fiscal du gouvernement. Un modèle théorique et empirique d'offre de travail a été développé dans lequel nous supposons que l'individu maximise son espérance d'utilité. Nous supposons également que ses préférences sont représentées par une fonction d'utilité quadratique. Le choix de cette forme fonctionnelle a été motivé par le fait qu'elle permet d'imposer un minimum de structure aux préférences individuelles. Elle est suffisamment flexible, en effet, pour que la séparabilité additive et la substituabilité parfaite entre les heures de travail officiel et au noir soient des cas particuliers. Elle présente, en outre, l'avantage de faciliter les estimations dans la mesure où les fonctions d'utilités marginales nettes qui en découlent sont linéaires dans les paramètres.

Notre modèle inclut non seulement les heures de travail officiel et au noir, les salaires obtenus sur ces deux marchés, mais également le revenu virtuel hors-travail. De ce fait, il tient compte des non linéarités dues à la fois au système fiscal et aux programmes sociaux. Notre modèle est alors constitué par un système de six équations simultanées avec censure sur les heures de travail dans chaque marché. Les estimations des paramètres de la fonction d'utilité sont obtenus par la méthode du maximum de vraisemblance à information complète. Cette approche économétrique nous permet d'obtenir des estimateurs convergents et asymptotiquement efficaces des coefficients estimés du modèle. Elle nous permet, en outre, d'effectuer des tests d'hypothèses non biaisés sur les paramètres du modèle.

L'information subjective contenue dans la base de données nous a permis de procéder à l'estimation du modèle. La perception subjective du risque associé aux activités souterraines est, en effet, d'autant plus primordiale que l'on doit s'attendre à ce que

³²Les valeurs de la statistique W calculée ne doivent pas surprendre compte tenu du niveau très élevé de significativité des coefficients, au regard de la statistique t de Student.

les individus décident de participer au marché noir, non pas en fonction des risques réels (difficilement appréciables), mais davantage en fonction de l'idée qu'ils s'en font.

A ce propos, il convient de noter que les travailleurs au noir peuvent être amenés à sous-estimer la probabilité de détection et le montant de l'amende de façon à limiter le stress inhérent à leur activité. De ce point de vue, nous sommes confrontés à un problème de dissonance cognitive (Akerlof *et al.*, 1982). En effet, les individus peuvent manipuler leurs croyances pour les mettre en conformité avec leurs désirs. Par conséquent, l'évaluation par l'individu du risque encouru est endogène. Or, l'approche économétrique que nous avons développée tient compte de l'endogénéité des variables subjectives de probabilité de détection et d'amende.

Un autre aspect spécifique au présent chapitre repose sur la prise en compte d'un effet de voisinage. Celui-ci est appréhendé au travers d'une variable subjective d'évaluation de la proportion de travailleurs au noir dans l'entourage de l'individu (*PROPOREN*). Nos résultats tendent à montrer que l'individu interagit avec son environnement lors de ses prises de décision et que les interactions sociales, parmi les membres du groupe auquel il appartient, conditionnent le choix de l'activité souterraine. Autrement dit, l'offre de travail au noir semble être influencée par le type d'activité exercée par l'entourage de l'individu. Néanmoins, cette conclusion doit être nuancée par le fait que la variable *PROPOREN* peut refléter l'existence de caractéristiques non observables similaires entre l'individu et les autres membres du groupe auquel il appartient. C'est pourquoi, dans certaines spécifications, nous avons introduit plusieurs variables de groupes, afin de contrôler les effets corrélés. Or, non seulement les coefficients associés aux variables de groupe ne sont pas significatifs, mais les coefficients associés à la variable subjective ne sont pas affectés par l'introduction des variables de groupe. Cela conforte l'idée selon laquelle le problème du miroir ne semble pas présent sur le marché noir. Par ailleurs, l'approche proposée dans ce chapitre constitue la première tentative d'appréhension des comportements interdépendants sur le marché noir.

Nos résultats témoignent, par ailleurs, d'une certaine similitude entre les marchés officiel et souterrain du travail. En effet, les rémunérations salariales sur chacun des deux marchés augmentent avec l'âge, mais à un taux décroissant. De même, la rentabilité de l'éducation a des rendements décroissants à la fois sur le marché officiel et sur le marché noir. Nous retrouvons ici les résultats standards de la littérature sur le capital humain et les salaires.

En outre, l'utilité marginale du travail sur les deux marchés augmente avec la taille du ménage, de sorte que les contraintes financières incitent tout autant les individus à travailler sur le marché officiel que sur le marché noir.

Nous constatons, en revanche, des divergences non négligeables entre les deux marchés du travail. Ainsi, la discrimination salariale dont semblent souffrir les femmes sur le marché officiel n'est pas significative sur le marché noir. En effet, alors que celles-ci sont moins rémunérées que les hommes sur le marché officiel, il ne semble pas y avoir de différences salariales entre les hommes et les femmes sur le marché noir. De même, la répartition géographique des niveaux de rémunérations est différente selon les deux marchés. Les salaires du marché officiel les plus élevés sont concentrés dans la région de Montréal, alors que ceux du marché noir sont davantage situés dans la région de Québec. Toutefois, le clivage entre les zones urbaines et rurales reste très marqué.

Enfin, nos résultats confirment que les choix d'offre de travail sur les deux marchés sont influencés par les mesures fiscales. Ils montrent ainsi que la participation au marché noir est significativement et positivement corrélée avec le taux d'imposition marginal. Dès lors, l'ampleur du fardeau fiscal semble inciter les individus à travailler au noir. Néanmoins, les élasticités des heures de travail par rapport aux paramètres fiscaux demeurent très faibles. En particulier, les élasticités par rapport aux risques de sanctions (probabilité de détection et amendes) sont dérisoires, ce qui reflète un certain manque de crédibilité des dispositifs de répression. Par conséquent, la crainte de punitions ne peut expliquer qu'une majorité d'individus choisisse d'exercer une activité officielle et de déclarer la totalité de ses revenus. En ce sens, l'analyse des facteurs purement économiques ne permet pas d'appréhender l'offre de travail au noir dans sa globalité. C'est pourquoi, nous poursuivons désormais l'analyse par l'examen des contraintes morales en action.