

Annexe 1A - Méthodes indirectes

Il existe actuellement cinq variables indicatrices, donnant lieu à cinq types de méthodes indirectes.

1. L'approche monétaire ou de demande de numéraire

Cette approche est certainement celle qui a été la plus utilisée pour mesurer la taille de l'économie souterraine. Cagan (1958), en reliant la demande de monnaie et la pression fiscale, est considéré comme l'initiateur de cette méthode. Dans ce cadre, la hausse excessive de demande de monnaie est attribuée à l'accroissement du fardeau fiscal conduisant les individus vers l'économie souterraine. Cette approche a ensuite été empruntée par Gutmann (1977), sans toutefois recourir à de solides procédures statistiques puisque les estimations obtenues résultent uniquement de l'observation. Feige (1979) adopte une approche très similaire, si ce n'est que l'attention est portée sur le volume des transactions. Ses travaux ont donné naissance à l'approche dite «approche des transactions». Enfin, Tanzi (1980, 1983) offre un fondement économétrique à la méthode monétaire et apporte quelques innovations.

1.1. Approche des transactions

Nous devons cette approche à Feige (1979, 1989, 1996) qui suppose une relation constante, à travers le temps, entre le volume des transactions et le PNB. Cette approche repose sur l'équation quantitative de la monnaie de Fisher. Des hypothèses sont dès lors nécessaires sur la vitesse de circulation de la monnaie et sur la relation entre la valeur des transactions totales et le PNB total (officiel et souterrain). En reliant le PNB total à l'ensemble des transactions, la production souterraine est calculée en soustrayant la part officielle du PNB total. Feige a dû, en outre, postuler une année de référence pour laquelle l'économie souterraine est inexistante, *i.e.* pour laquelle l'activité officielle caractérise la totalité du PNB.

En théorie, cette approche paraît très attrayante, mais les conditions empiriques de son application sont difficiles à remplir, de sorte que les estimations obtenues s'avèrent souvent peu crédibles. Le principal inconvénient de cette approche est qu'elle requiert un nombre considérable de données. En effet, pour relier la totalité des transactions à la production nationale, il faudrait se limiter aux transactions finales sur les biens et services, mais de nombreux échanges correspondent à des transactions intermédiaires, voire purement financières. Il est donc nécessaire de disposer d'une information précise

sur les différents types de transactions. Mais, il est également nécessaire de connaître précisément le volume total des transactions. Or, cela s'avère particulièrement délicat dans le cas des transactions au comptant. Quels que soient les raffinements envisagés, il demeure difficile de déterminer la vitesse de circulation des billets de banques et des dépôts à vue dans l'ensemble des transactions de biens et de services. La vitesse de circulation de l'argent liquide dépend notamment de la durabilité des billets de banque émis, *i.e.* de la qualité du papier d'impression. Pour contourner cette difficulté, Feige a alors estimé la durée de vie des billets par le rapport entre le stock de numéraire et le volume de billets remplacés chaque année. Il a ensuite divisé le nombre de transactions réglées en espèces (estimé à 125 pour chaque billet) par la durée de vie moyenne des billets. Enfin, pour tenir compte de la possibilité que l'amélioration de la qualité des billets de banque ait entraîné une augmentation du nombre de transactions par billet, il a arbitrairement élevé le nombre de transactions par billet, de 125 à 225, ce qui s'est inévitablement traduit par une hausse importante de la taille estimée de l'économie souterraine.

Cette approche a également suscité la critique en raison du nombre conséquent d'estimations requises et des risques de biais qu'elles entraînent, en particulier lorsqu'il s'agit de déterminer la vitesse de circulation des billets de banque et des dépôts à vue.²⁶ Par la suite, plusieurs innovations ont été proposées et ont apporté un fondement économétrique à cette méthode.

1.2. Prolongements économétriques

Le premier apport de Tanzi (1980, 1983) a été d'introduire explicitement un fardeau fiscal dans l'équation de demande de monnaie. Afin d'isoler l'excès de demande de monnaie résultant des transactions souterraines au comptant, tous les facteurs conventionnels, tels que les habitudes de paiements, les taux d'intérêts, les impôts directs et indirects, les réglementations publiques et la complexité du système fiscal, sont pris en compte dans les estimations économétriques. L'équation de régression pour la demande de monnaie, proposée par Tanzi, est la suivante :

$$\ln(C/M_2) = \beta_0 + \beta_1 \ln(1 + TW)_t + \beta_2 \ln(WS/Y)_t \\ + \beta_3 \ln R_t + \beta_4 \ln(Y/N) + u_t,$$

avec $\beta_1 > 0, \beta_2 > 0, \beta_3 < 0$ et $\beta_4 > 0$.

où C/M_2 est le ratio entre l'argent comptant en circulation et les comptes courants et d'épargne ; TW caractérise le taux d'imposition moyen pondéré, servant à approximer les variations de la taille de l'économie souterraine ; WS/Y décrit la proportion des salaires dans le revenu national et permet de capter les changements de comportements de paiements ; R est l'intérêt payé sur les comptes d'épargne, ce qui permet de saisir le coût d'opportunité de la détention d'argent comptant ; et enfin, Y/N représente le

²⁶Pour une critique détaillée de l'approche des transactions, consulter Frey et Pommerehne (1984), Tanzi (1982, 1986), Thomas (1986, 1992, 1999) et Giles (1999).

revenu par tête. La taille et la croissance de l'économie souterraine peut donc être évaluée par la comparaison de l'accroissement de monnaie, lorsque les impôts et les réglementations sont au plus bas niveau, avec l'accroissement monétaire aux niveaux de taxation et de réglementations élevés tels qu'ils sont observés. Ensuite, en supposant une même vitesse de circulation de la monnaie sur le marché noir que sur le marché officiel, la taille de l'économie souterraine peut être calculée et comparée au PIB.

Pour évaluer la taille de l'économie souterraine, cette méthode se sert de l'information véhiculée par les agrégats monétaires et tente d'exploiter le lien qui pourrait exister entre l'économie souterraine et la demande de numéraire. L'avantage principal de cette méthode indirecte est qu'elle ne nécessite pas, de la part des agents économiques (*e.g.* consommateurs, entreprises, travailleurs), de révélation directe de leurs revenus et dépenses dans l'économie souterraine. Cette équation de demande de monnaie a donc été fréquemment utilisée pour estimer la taille de l'économie souterraine. Néanmoins, elle fait l'objet de sévères critiques.²⁷ Son inconvénient majeur est qu'elle repose sur des hypothèses, non seulement invérifiables empiriquement, mais auxquelles les résultats obtenus sont extrêmement sensibles (Lafèche, 1994).

Dans sa forme la plus élémentaire, cette méthode repose sur quatre hypothèses fondamentales. La première suppose que les transactions dans l'économie souterraine sont réglées en argent comptant. Selon la seconde hypothèse, la vitesse de circulation de la monnaie, *i.e.* le nombre de transactions réalisées par unité monétaire durant une période donnée, demeure constante dans le temps.²⁸ La troisième hypothèse est que la taille de l'économie souterraine est nulle pour une année de base particulière. Comme la valeur des transactions réalisées durant une année est simplement le produit de la vitesse de circulation de la monnaie et de la masse monétaire en circulation, ce produit peut être utilisé pour estimer la valeur de l'ensemble des transactions dans l'économie officielle et dans l'économie souterraine au cours d'une année. Il est donc possible de mesurer l'importance des transactions dans l'économie souterraine en prenant la différence entre cette estimation de la valeur totale des transactions dans l'économie et les transactions dans l'économie officielle telles que mesurées dans les Comptes Nationaux. Finalement, la quatrième hypothèse suppose que le rapport entre le volume des transactions et la valeur de la production économique totale est stable, ce qui permet d'évaluer l'importance de la production sur la base du volume des transactions.²⁹

L'hypothèse selon laquelle les transactions dans l'économie souterraine sont réglées en espèces est restrictive. Des études utilisant des données d'enquêtes ont, en effet, pu

²⁷Voir notamment, Feige (1986), Thomas (1986, 1992, 1999), Pozo (1996), Giles (1999) et Bhattacharyya (1999).

²⁸La vitesse de la monnaie correspond au nombre de fois où la monnaie circule dans un pays. Elle est calculée par le rapport entre la production et le stock de monnaie.

²⁹S'il n'y avait pas de production intermédiaire et de transactions financières, le volume des transactions serait égal à la valeur de la production économique. Cependant, les transactions excèdent la production (mesurée par la somme des valeurs ajoutées à chaque étape de la production) lorsqu'un bien ou service est produit en plusieurs étapes et lorsque de nombreuses transactions financières plus ou moins reliées aux activités de production sont réalisées. Or, ces phénomènes sont courants dans une économie moderne.

montrer qu'une certaine proportion des transactions (entre 9 à 15%) sont payées par chèques (*e.g.* Isachsen et Strom, 1980, 1985; Smith, 1982; Ginsburgh *et al.*, 1987a; Fortin *et al.*, 1987). L'hypothèse que seul l'argent comptant est utilisé peut cependant être modifiée en postulant qu'une fraction fixe des transactions souterraines est réglée en chèque.

La deuxième hypothèse est plus problématique puisque les travaux empiriques en économie monétaire démontrent clairement que la vitesse de circulation de la monnaie n'est pas constante dans le temps.³⁰ La raison principale de cette variation est la présence d'innovations financières telle que l'introduction de cartes de débit. Cette innovation a sans doute modifié de façon significative le rôle de l'argent comptant dans les achats de biens et de services par les consommateurs. De tels changements peuvent donc produire un effet imprévisible sur les estimations de la taille de l'économie souterraine basées sur l'approche monétaire. Un autre point faible relatif à la vitesse de circulation de la monnaie est l'hypothèse selon laquelle celle-ci serait identique sur le marché officiel et sur le marché noir. Or, si elle s'avère difficile à appréhender pour l'économie officielle, elle l'est encore davantage pour l'économie souterraine (Klovland, 1984; Hill et Kabir, 1996). Il paraît cependant raisonnable de penser que les innovations financières affectent également les modes de paiements effectués sur le marché noir. En revanche, il est fort probable que leur influence soit plus modérée dans l'économie souterraine, de sorte que la circulation de la monnaie est certainement différente entre les deux marchés.

La méthode d'estimation de la taille de l'économie souterraine fondée sur le produit de l'argent comptant en circulation et de la vitesse de circulation de la monnaie constitue évidemment la version la plus simple de l'approche monétaire. Des méthodes économétriques plus sophistiquées ont été employées en vue d'obtenir des estimations plus crédibles.³¹ Les hypothèses requises pour prédire la vitesse de circulation de la monnaie ne peuvent néanmoins être testées directement. Il s'avère donc difficile de vérifier la validité des estimations de la taille de l'économie souterraine obtenues à l'aide de cette approche.

La troisième hypothèse selon laquelle la taille de l'économie souterraine serait nulle au cours d'une année de base constitue une normalisation arbitraire qui ne modifie pas directement l'estimation des changements d'ampleur de l'économie souterraine. Mais, les estimations en valeur absolue s'avèrent tout aussi arbitraires que cette normalisation. En outre, un aspect insatisfaisant de cette hypothèse est qu'elle présuppose que l'économie souterraine devra nécessairement s'accroître, faute de quoi elle pourrait atteindre une taille négative. Or, rien ne garantit que les estimations fondées sur l'approche monétaire soient positives. Cet aspect constitue l'une des faiblesses propres à cette méthode.

³⁰Se reporter, par exemple à Goodhart (1986), Hall *et al.* (1987), Siklos (1993), ou encore à Bordo *et al.* (1997).

³¹Ainsi, la méthode de Tanzi (1980, 1999) est sans doute plus fiable que celles de Feige (1979) et de Gutmann (1977), puisque certaines des hypothèses sous-jacentes sont testées économétriquement et que la valeur de plusieurs paramètres importants est estimée à l'aide de techniques économétriques au lieu d'être simplement imposée.

Enfin, la quatrième hypothèse supposant la stabilité entre le volume des transactions et la valeur de la production pose également problème. L'explosion récente du volume des transactions sur le marché des changes et la création de nouveaux instruments financiers (fondés sur les hypothèques et les comptes de crédit) suggèrent que les transactions financières se sont accrues plus rapidement que la production. Une mesure de l'économie souterraine fondée sur le volume des transactions aura alors tendance à surestimer la croissance du phénomène. Sur ce point, une critique supplémentaire est apportée aux travaux de Tanzi pour les Etats-Unis. En effet, le dollar américain est une devise servant aux échanges internationaux et, en ce sens, il est détenu pour lui-même, y compris à l'étranger. Il peut dès lors alimenter les marchés noirs tout autant que les devises nationales, mais la masse de dollars en circulation n'est pas représentative des échanges souterrains américains (Blades, 1982; Feige, 1986, 1997). Par ailleurs, la variabilité des estimations fournies par l'approche monétaire est non seulement le résultat de variations véritables de l'ampleur du phénomène, mais également de la variabilité inhérente aux années considérées et de l'instabilité de cette méthode. De nombreux auteurs ont ainsi souligné le manque de stabilité des paramètres estimés (Frey *et al.*, 1984; Thomas, 1986, 1992, 1999). Les estimations sont très sensibles à la modélisation du système fiscal, de même qu'aux valeurs de la vitesse de circulation de la monnaie retenues (Hill et Kabir, 1996). A titre d'exemple pour la seule année 1979, les estimations oscillent entre 10% et 28% pour les Etats-Unis (Feige, 1979; Bhattacharyya *et al.*, 1986); entre 7% et 15% pour le Royaume-Uni (Dilnot et Morris, 1981; Feige, 1981); et 6% et 20% pour le Canada en 1984 (Italiano, 1985; Karoleff *et al.*, 1993). Enfin, il convient de noter que les estimations issues de l'approche monétaire sont généralement nettement plus importantes que celles obtenues à l'aide des comptes nationaux ou des enquêtes.

Afin d'illustrer le manque de fiabilité de l'approche monétaire, notons que si l'économie souterraine représentait réellement 20% du PIB au Canada, elle correspondrait à une dépense annuelle de 13 000\$can. (soit 8 400 euros) par ménage, ce qui semble exagérément élevé. En revanche, si l'économie souterraine ne représente que 3% du PIB, ce qui correspond à un ordre de grandeur obtenu avec l'approche des Comptes Nationaux ou l'approche d'enquête de Fortin *et al.* (1996), son importance est de 2 000\$ par ménage (*i.e.* 1 300 euros), ce qui paraît beaucoup plus raisonnable. Compte tenu des estimations excessivement élevées auxquelles cette méthode conduit, il ne nous semble pas pertinent de nous attarder outre mesure sur les estimations du niveau absolu de l'économie souterraine. Nous nous concentrons plutôt sur les changements récents de la taille de l'économie souterraine tels qu'estimés à l'aide de l'approche monétaire.

Une première remarque, à cet effet, est qu'à peu près toutes les études fondées sur cette approche indiquent une croissance importante de l'économie souterraine. Cela peut paraître quelque peu surprenant puisqu'il existe des raisons de penser que la croissance et le développement économique peuvent réduire l'importance relative de l'économie souterraine telle que définie du point de vue des autorités fiscales ou des Comptes Nationaux. La composition des recettes fiscales des différents pays illustre bien cette possibilité. En effet, les recettes provenant de l'impôt sur le revenu des particuliers sont en général très limitées dans la majorité des pays en développement,

en raison des difficultés de collecte de cette source de revenus. Le développement économique et la modernisation des méthodes de transaction peuvent aussi avoir pour résultat d'augmenter la proportion des activités économiques qui sont mesurées dans les Comptes Nationaux.

Par ailleurs, un certain nombre d'études qui se fondent sur l'approche monétaire ont été publiées récemment au Canada afin d'évaluer l'incidence de l'introduction de la Taxe sur les Produits et Services (TPS) sur la taille relative de l'économie souterraine. Selon l'hypothèse de base de l'analyse, il est plus facile d'éviter de payer les taxes sur les services que sur les biens. Or, l'introduction de la TPS a eu pour effet d'augmenter les taxes sur les services et de réduire les taxes sur certains biens. En conséquence, elle devrait avoir pour effet de favoriser l'expansion de l'économie souterraine. Cette hypothèse a été confirmée partiellement par l'étude de Spiro (1994) qui attribue à la TPS une augmentation de l'ordre de 0,8% dans la taille de l'économie souterraine. Cette conclusion n'est cependant pas partagée par Hill et Kabir (1995) selon lesquels il existe une très large variabilité des résultats, d'après la définition du taux de taxation retenue ainsi que les hypothèses posées sur la vitesse de circulation de la monnaie. Il n'existe donc pas de consensus empirique quant à l'impact de l'introduction de la TPS sur la taille relative de l'économie souterraine.

En conclusion, l'approche monétaire semble suggérer une économie souterraine beaucoup plus considérable que les autres approches. Elle présente l'avantage d'être une méthode indirecte qui ne conduit pas à une sous-estimation de la taille de l'économie souterraine en raison de réticences des agents économiques à révéler leurs transactions réalisées au noir. En revanche, cet aspect indirect de l'approche constitue aussi sa faiblesse dans la mesure où les estimations dépendent de façon critique des hypothèses postulées, notamment sur la stabilité de la vitesse de circulation de la monnaie, qui s'avèrent délicates à vérifier. En outre, la plupart des études ne considère que le fardeau fiscal. Or, celui-ci n'est qu'un aspect parmi tant d'autres des raisons de l'apparition de l'économie souterraine. L'ampleur de la réglementation, les attitudes des individus envers l'Etat, la moralité fiscale en sont les exemples les plus saisissants. Ils sont généralement écartés de l'analyse par manque de données en la matière. Mais, ces facteurs affectent pourtant considérablement les décisions de participation au marché noir et, par conséquent, la taille de l'économie souterraine dans son ensemble.³²

Les estimations obtenues à l'aide des différentes méthodes monétaires indiquent clairement que l'approche des transactions produit les résultats les plus élevés, suivie de l'approche développée par Gutmann et enfin de l'approche de la demande de numéraire

³² Quelques exceptions toutefois sont à souligner. Ainsi, Frey et Weck-Hannemann (1984) introduisent une variable d'«immoralité fiscale» dans leur modèle. Celle-ci a d'ailleurs une influence quantitativement plus grande et statistiquement plus forte que la seule part de fiscalité directe. De même, Pommerehne et Schneider (1985), ainsi que Zilberfarb (1986), confirment l'impact sur l'économie souterraine des considérations morales, des réglementations et, de façon générale, des conditions du marché officiel appréhendés notamment par le taux de salaire minimum. Par ailleurs, il convient de noter qu'un certain nombre de travaux tendent, depuis peu, à se développer autour des notions de morale fiscale (e.g. Beckmann, 2001; Schnellenbach, 2002; Slemrod et Yitzakhi, 2002; Kirchler *et al.*, 2003). A partir d'une démarche expérimentale, ils s'attachent à démontrer le lien entre les comportements de fraude et les perceptions subjectives des individus.

(Schneider et Enste, 2000). La méthode des transactions évalue la taille de l'économie souterraine de 15 à 35% du PIB, entre 1979 et 1990, pour cinq pays de l'OCDE (Allemagne, Canada, Grande-Bretagne, Italie et Etats-Unis). L'approche fondée sur la demande d'argent liquide l'estime de 10 à 30% et l'approche de la demande de numéraire entre 4 et 20% du PIB. Concernant l'évolution du phénomène, cette dernière méthode démontre une nette diminution, tandis que les autres approches monétaires attestent d'une croissance des activités souterraines.

2. L'approche par le taux de participation au marché du travail

Cette approche exploite les statistiques relatives au marché du travail pour estimer la taille de l'économie souterraine. En effet, si le pourcentage de la population active est supposé constant, *ceteris paribus*, alors un déclin de la participation au marché officiel peut être considéré comme un indicateur de l'accroissement des activités souterraines. Une telle approche a été utilisée en Italie (Contini, 1981, 1982) et aux Etats-Unis (O'Neil, 1983). Mais, elle demeure relativement peu employée. La faiblesse de cette approche provient du fait que les différences observées dans le taux de participation au marché du travail officiel peuvent avoir de nombreuses explications. Il se peut, en outre, que les individus cumulent une activité officielle et un emploi dissimulé, et dans ce cas, ils ne peuvent pas être appréhendés par cette méthode. A titre d'exemple, dans notre échantillon, la moitié des travailleurs au noir exercent une activité officielle en parallèle. Par conséquent, les estimations de la taille de l'économie souterraine fondées sur l'approche de l'écart entre la population active mesurée dans les recensements et la population active réelle sont peu crédibles. D'une part, les quelques estimations fondées sur cette approche sont très différentes de celles obtenues par les approches alternatives. Pour certains pays, les valeurs sont plus élevées (*e.g.* Allemagne), pour d'autres, elles se révèlent bien plus faibles (*e.g.* Canada). D'autre part, cette méthode ne permet pas de conclure quant à l'évolution du phénomène (Schneider et Enste, 2000).

3. L'approche des inputs physiques (consommation électrique)

Cette méthode, essentiellement utilisée pour les pays en transition et en voie de développement, repose sur un indicateur physique de l'activité économique -la consommation d'électricité- qui est censé caractériser à la fois l'économie officielle et l'économie souterraine. En effet, une partie de l'électricité est supposée consommée dans les activités souterraines et cette partie est censée isolable de celle qui est consommée officiellement. L'hypothèse sous-jacente est alors que la consommation d'électricité augmente au même taux que le PIB. Mais, l'un des problèmes majeurs réside dans le fait que la demande d'électricité enregistre d'importantes variations au cours du temps et à travers les pays. En outre, la production et la distribution d'électricité sont souvent très sensibles aux progrès techniques du pays considéré. Ainsi, lorsque les progrès techniques sont considérables dans ce secteur d'activités, il y a un risque évident de sous-estimation de la taille de l'économie souterraine. Deux méthodes distinctes

proposent d'estimer l'ampleur des activités souterraines à partir de la consommation d'électricité.

3.1. Méthode de Kaufmann-Kaliberda

La consommation électrique est supposée être le meilleur indicateur de mesure physique de l'activité économique dans son ensemble, *i.e.* officielle et souterraine (Kaufmann et Kaliberda, 1996). L'observation empirique a démontré que l'activité économique et la consommation d'électricité étaient corrélées et que l'élasticité de l'électricité par rapport au PIB était proche de l'unité. En utilisant un indicateur de l'activité économique et en la soustrayant du PIB officiellement mesuré, ces auteurs obtiennent alors une estimation du PIB non officiel. La différence entre la croissance du PIB officiel et celle de la consommation électrique est attribuée à la croissance de l'économie souterraine.

Cette approche est simple et très séduisante, mais de nombreuses critiques lui sont néanmoins adressées. En premier lieu, toutes les activités économiques souterraines n'exigent pas un débit important d'électricité, à l'instar des services personnels, et d'autres sources d'énergie peuvent être utilisées (*e.g.* gaz, fuel, charbon). Par conséquent, une partie seulement de l'activité souterraine peut être captée par cette méthode. En deuxième lieu, le progrès technique est particulièrement diffus dans ce secteur d'activité, de sorte que l'usage de l'électricité est plus efficace que par le passé, aussi bien sur le marché officiel que souterrain. Enfin, l'élasticité de la consommation électrique par rapport au PIB est très différente d'un pays à l'autre et, pour un même pays, varie très fortement dans le temps. Johnson, Kaufmann et Schleifer (1997) ont tenté d'ajuster les élasticités pour ces changements, mais cette approche demeure très instable. Lacko (1996, 1997) soulève de nombreux points négatifs à l'encontre de cette approche et, pour pallier ces critiques, propose une méthode alternative.

3.2 Méthode de Lacko

Dans cette approche, une certaine part de l'économie souterraine est supposée être associée à la consommation d'électricité des ménages incluant la production domestique et toute sorte d'activités de production et services non enregistrés dans les chiffres officiels (Lacko, 1996, 1999). Lorsque cette part de l'économie souterraine associée à la consommation électrique est élevée, le reste de l'économie -qui ne peut être mesuré- est également supposé élevé. Cette approche peut être décrite par les deux équations suivantes :

$$\ln E_i = \alpha_1 \ln C_i + \alpha_2 \ln PR_i + \alpha_3 G_i + \alpha_4 Q_i + \alpha_5 H_i + u_i, \quad (58)$$

avec $\alpha_1 > 0$, $\alpha_3 > 0$, $\alpha_5 > 0$, et $\alpha_2 < 0$, $\alpha_4 < 0$.

$$H_i = \beta_1 T_i + \beta_2 (S_i - T_i) + \beta_3 D_i, \quad (59)$$

avec $\beta_1 > 0$, $\beta_3 > 0$ et $\beta_2 < 0$.

où E_i est la consommation d'électricité par tête dans le pays i , C_i est la consommation réelle par tête hormis la consommation d'électricité en dollars américains, PR_i est le

prix réel de consommation d'un kilowatt-heure d'électricité domestique en dollars US, G_i désigne la fréquence relative du nombre de mois où le chauffage est nécessaire, Q_i est le ratio des sources d'énergie autres que l'électricité dans la consommation des ménages, H_i correspond à l'output par tête de l'économie souterraine, T_i au ratio de la somme du revenu personnel payé, du profit commercial et taxes sur les biens et services sur le PIB, S_i est le ratio des dépenses publiques sur le PIB, et enfin, D_i est la somme du pourcentage de bénéficiaires de transferts sociaux de plus de 14 ans parmi les actifs et du pourcentage des inactifs parmi ces mêmes actifs.

La procédure économétrique consiste à estimer l'équation (58), après avoir remplacé H_i par son équation (59). Les résultats peuvent être utilisés pour classer les pays selon leur consommation d'électricité dans leurs activités non déclarées. Mais, le calcul de la taille réelle de l'économie nécessite de connaître le PIB par unité de production d'électricité dans l'économie souterraine de chaque pays. Or, de telles données ne sont malheureusement pas disponibles. C'est pourquoi, il est nécessaire de recourir aux estimations issues d'approches concurrentes. En effet, l'un des résultats obtenus par ailleurs est appliqué aux autres pays. Lacko, par exemple, se sert de l'estimation de l'économie souterraine aux Etats-Unis, au début des années quatre-vingt-dix, comme base de calcul pour évaluer la taille de l'économie souterraine des autres pays.³³

Certaines des critiques adressées à Kaufmann et Kaliberda restent valables dans ce cadre. En effet, toutes les activités économiques ne requièrent pas l'usage de l'électricité et d'autres sources d'énergie peuvent également être mobilisées. Mais, au-delà de ces critiques générales, la méthode de Lacko présente plusieurs inconvénients. En premier lieu, les activités souterraines ne sont pas liées à la production domestique. En deuxième lieu, il n'est pas évident que le ratio des dépenses publiques puisse être un facteur explicatif des activités souterraines, spécialement dans les pays en transition et dans les pays en développement. Enfin, le choix de l'estimation servant de base de calcul à tous les pays relève de l'arbitraire et, dès lors, est susceptible de poser de sérieux problèmes. Ce type d'approche conduit à des estimations intermédiaires par rapport aux autres méthodes. La taille moyenne de l'économie souterraine, pour cinq pays de l'OCDE (Allemagne, Canada, Grande-Bretagne, Italie et Etats-Unis) est estimée à 12% sur la période 1970-1990 (Schneider et Enste, 2000).

4. L'approche par la modélisation ou méthode MIMIC

Frey et Weck (1983) et Frey et Weck-Hannemann (1984) sont les pionniers en matière d'utilisation des modèles MIMIC dans le contexte de l'économie souterraine. Ils ont appliqué cette méthode à des données issues de vingt-deux pays de l'OCDE. Au préalable, l'idée d'une modélisation simplifiée («soft modeling») a permis d'établir un classement de la taille relative de l'économie souterraine dans différents pays. Cette approche a ensuite été adoptée par Schneider (1997a), Tedds (1998), Giles (1999), Giles, Tedds et Werkneh (2002).

³³L'estimation en question est issue des travaux de Morris (1993) selon lesquels l'économie souterraine américaine équivalait à 10,5% du PIB.

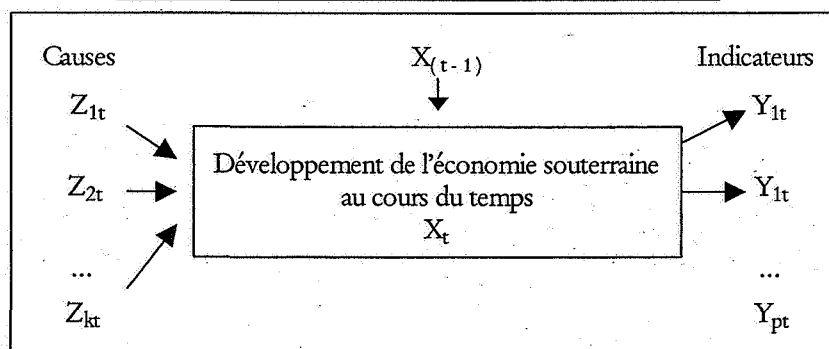
L'ensemble des méthodes analysées jusqu'ici ne considèrent qu'un seul indicateur de tous les effets de l'économie souterraine. Pourtant, ses effets se font ressentir simultanément sur de nombreuses variables, que ce soit sur le marché monétaire, au niveau de la production ou sur le marché du travail. Certaines méthodes ne s'intéressent même qu'à une seule cause de l'apparition de l'économie souterraine. Ainsi, l'approche monétaire, comme nous l'avons souligné précédemment, n'envisage que le fardeau fiscal comme facteur explicatif de l'activité souterraine. A l'inverse, la modélisation de type MIMIC (indicateurs multiples, causes multiples) considère explicitement les nombreuses causes et les nombreux effets de l'économie souterraine. Cette méthode empirique, fondée sur la théorie statistique des variables inobservées, reconnaît le caractère invisible inhérent à l'économie souterraine. Elle relève d'une approche économétrique structurelle et traite la taille de l'économie souterraine comme une variable inobservable latente. Celle-ci est reliée, d'une part à un ensemble d'indicateurs observables qui reflètent les changements dans la taille de l'économie souterraine, et d'autre part, à une combinaison de variables causales observables qui sont considérées comme facteurs explicatifs pertinents de l'activité souterraine. Les modèles MIMIC utilisent donc simultanément l'information contenue dans plusieurs variables indicatrices et plusieurs variables causales pour estimer la taille de l'économie souterraine durant la période visée par l'échantillon.

La procédure d'estimation devient multidimensionnelle. Les coefficients inconnus sont estimés par une série d'équations structurelles dans lesquelles la variable inobservée (*i.e.* la taille de l'économie souterraine) ne peut pas être mesurée directement. Cette variable inobservée est reliée à des indicateurs qui sont observés. Ainsi, la taille de l'économie souterraine, non observée, est reliée à des indicateurs observables de sa propre taille. Ces indicateurs sont censés exercer une influence sur elle et donc capter la dépendance structurelle de l'économie souterraine aux variables susceptibles d'être utiles pour prédire sa taille et sa croissance future.

Le modèle est dynamique et l'interaction, au fil du temps, entre les causes Z_{it} , la taille de l'économie souterraine X_{it} et les indicateurs Y_{it} peut être décrite par le schéma 1A-1 :

Schéma 1A-1 :

Evolution de l'économie souterraine selon le modèle MIMIC



Source : Schneider et Enste (2000).

Trois causes sont habituellement avancées pour expliquer le choix de délaisser un emploi officiel au profit d'une activité souterraine. Il s'agit du fardeau élevé de la fiscalité, du poids accru des réglementations et du déclin de la morale fiscale, ou plus généralement des attitudes négatives des citoyens envers les pouvoirs publics. La principale difficulté étant d'obtenir des données quant au poids de la réglementation, réel ou perçu, et quant à la perception du rôle de l'Etat par les individus. Le choix des variables pertinentes est également une source potentielle de problèmes. Comme toujours en la matière, la littérature empirique n'apporte guère d'enseignements, si ce n'est la présentation de variables significatives, à une date donnée et pour un ensemble de pays donné.

L'évolution de l'économie souterraine est essentiellement appréhendée à travers des indicateurs monétaires, des indicateurs de l'état du marché du travail et de la production marchande. Ainsi, un accroissement du volume de transactions, une diminution de la main-d'œuvre officielle, une réduction de la durée moyenne de travail officiel ou une perte de croissance économique sont interprétés comme autant de signes de l'expansion de l'économie souterraine, du renforcement de la participation au marché noir et du détournement d'inputs de production au détriment du secteur officiel. Bien évidemment, ces postulats sont discutables.

Le modèle MIMIC dynamique sert à retracer l'évolution temporelle de l'économie souterraine, sur une période donnée. Il permet, en effet, de déterminer un indice de séries temporelles du rapport (*production cachée / production mesurée*) pour un ou plusieurs pays. Cet indice permet de classer les pays les uns par rapport aux autres, ou d'ordonner différentes périodes pour un même pays. La série indiciaire peut ensuite être convertie en mesure cardinale grâce à l'utilisation de la moyenne des estimations disponibles, ce qui permet d'obtenir une valeur repère pour la série. Il est également possible d'estimer un modèle de demande de numéraire pour définir cette valeur repère et convertir l'indice en pourcentage du PIB. Plusieurs cheminements chronologiques distincts, suivant différentes combinaisons de variables causales et indicatrices, peuvent être proposées. De la même manière, l'équation de demande de monnaie peut tenir compte de différentes vitesses de circulation dans l'économie officielle et l'économie souterraine. Ce fut notamment le cas dans l'étude de Giles, Tedds et Werkneh (2002) dans laquelle l'équation de demande de numéraire n'est pas assimilée à un input pour déterminer la variation de la taille de l'économie souterraine. Elle ne sert, au contraire, qu'à produire une valeur moyenne de long terme de l'output (*caché / mesuré*), de sorte que le ratio prédit par le modèle peut être utilisé pour estimer l'économie souterraine.

Cette méthode présente l'intérêt de se libérer de certaines hypothèses contraignantes. Elle offre un cadre théorique intéressant et repose sur un solide fondement économétrique permettant de réaliser plusieurs tests statistiques. Néanmoins, Helberger et Knepel (1988) soulignent que les résultats obtenus par cette méthode sont très instables. Non seulement, les coefficients estimés sont très sensibles à la taille de l'échantillon, mais ils varient aussi énormément selon les spécifications retenues du modèle. Ils sont particulièrement affectés par les changements de variables et sont extrêmement sensibles aux diverses pondérations des indices utilisés. Dès lors, il s'avère difficile d'entreprendre une analyse en série temporelle pure avec une telle procédure d'estimation

(Schneider, 1997b). Il est crucial de s'assurer de la stabilité intertemporelle des coefficients et donc de la stationarité des diverses séries chronologiques. L'usage de cette méthode nécessite de porter une attention toute particulière aux propriétés statistiques et à l'interprétation des estimations réalisées. Par ailleurs, l'une des contraintes essentielles demeure le choix des variables pertinentes parmi l'ensemble restreint de celles qui sont accessibles. La fiabilité et le pouvoir explicatif de certaines variables indicatrices peut notamment être mis en doute. Elle requiert, en outre, une information variée et très détaillée, à la fois, du marché du travail, du marché monétaire et de la production marchande. Enfin, son caractère général et automatique empêche toute tentative de compréhension des motivations individuelles à participer à l'économie souterraine.

5. L'approche comptable (écart recettes - dépenses publiques)

Le Système de Comptabilité Nationale a pour objectif de mesurer le Produit Intérieur Brut (PIB). Celui-ci peut être obtenu de deux manières différentes, soit par les dépenses, soit par les revenus.

Selon l'approche par les dépenses, le PIB correspond à la somme de toutes les consommations finales des unités résidentes. L'approche par les revenus, quant à elle, détermine le PIB par la totalité des revenus. Cette dernière approche est indépendante de la précédente puisque les estimateurs des différents types de revenus proviennent de sources distinctes de celles utilisées pour les dépenses. Pourtant, la logique conceptuelle sous-jacente à l'estimation par les revenus est la même que celle utilisée dans l'estimation par les dépenses. En effet, du point de vue des dépenses, la valeur ajoutée est la différence entre le produit total et la consommation intermédiaire. Or, cette valeur ajoutée n'est rien d'autre qu'un estimateur des revenus générés par les différentes activités de production.

L'existence de ces deux approches suggère l'utilisation d'estimateurs différents et la comparaison des résultats obtenus. Un écart constaté entre le revenu mesuré et les dépenses estimées pourrait dès lors fournir une évaluation de la taille de l'économie souterraine. En tout état de cause, la confrontation de données provenant de sources différentes fait partie intégrante de la méthode de compilation des Comptes Nationaux. En outre, cette confrontation peut permettre d'identifier les erreurs et parfois de combler en partie les informations manquantes.

Les estimations du PIB obtenues sur la base des dépenses sont souvent plus élevées que celles qu'on obtient à partir des revenus, ce qui donne lieu à une divergence statistique que les services publics comptables répartissent de façon égale du côté des dépenses et des revenus. Une interprétation possible de cette divergence statistique est que les revenus, tels que mesurés dans les Comptes Nationaux, sous-estiment l'ensemble des revenus puisque ceux-ci découlent de transactions dans l'économie souterraine et ne sont donc pas comptabilisés. En revanche, les dépenses reflètent l'ensemble de celles qui sont effectuées dans l'économie officielle et dans l'économie souterraine. Cette interprétation est cependant incomplète, dans la mesure où la divergence statistique résulte également d'une procédure complexe d'ajustements des revenus et des

dépenses des Comptes Nationaux. Par ailleurs, il importe de souligner que chacune des composantes par poste détaillé du PIB n'est malheureusement pas parfaitement mesuré, qu'il s'agisse du côté dépense ou du côté revenu. La divergence statistique peut dès lors caractériser tout autant les omissions et erreurs perpétrées dans les Comptes Nationaux, ce qui limite la portée de cette approche et conduit à mettre en doute la fiabilité des résultats ainsi obtenus.

La procédure d'estimation de la taille de l'économie souterraine à partir des Comptes Nationaux se fonde sur un processus de réconciliation des revenus et des dépenses par poste détaillé du PIB. La démarche généralement retenue est celle du calcul d'une borne supérieure, encore appelée «analyse de sensibilité». Il s'agit, en fait, d'estimer une limite supérieure à l'importance des activités non observées. Ainsi, pour chaque composante du PIB, est estimé le montant maximum qui pourrait correspondre à des activités non observées. Les résultats obtenus sont ensuite additionnés pour constituer la valeur de la borne supérieure. En raison de la complexité du Système de Comptabilité Nationale, il serait fastidieux de discuter cette approche en détail. Comme dans le cas des approches précédentes, nous nous concentrerons donc sur les principes fondamentaux de cette méthode sans insister outre mesure sur les détails techniques.

Considérons, à titre d'exemple, le cas des cigarettes qui a récemment attiré l'attention des autorités canadiennes à cause des problèmes d'importation clandestine. Ainsi, la part de la production totale de tabac exportée du Canada vers les Etats-Unis s'est rapidement accrue au début des années quatre-vingt-dix. Les cigarettes étaient ensuite réintroduites en contrebande au Canada. En conséquence, la valeur mesurée des ventes domestiques finales de tabac a chuté, au Canada, réduisant la contribution au PIB de la valeur ajoutée (*e.g.* revenus des commerçants) du commerce de tabac (en gros et au détail). En d'autres termes, les données habituellement utilisées pour calculer le PIB du côté des revenus ont sous-estimé de façon importante la taille du commerce du tabac canadien. Il devient dès lors possible de calculer l'importance de l'économie souterraine dans le commerce du tabac en posant des hypothèses simples sur la proportion des exportations de tabac vers les Etats-Unis qui sont ensuite importés en contrebande.³⁴

Notons également que les dépenses rapportées dans l'enquête de l'institut national de statistique canadien (Statistiques Canada) sur les dépenses de consommation des ménages (FAMEX) n'ont pas suivi cette tendance à la baisse au début des années quatre-vingt-dix (Gervais, 1995). Autrement dit, les consommateurs ont continué à déclarer leurs dépenses de consommation dans l'enquête FAMEX, mais ces dépenses étaient effectuées sur des cigarettes obtenues aussi bien légalement qu'en contrebande.³⁵

³⁴Il convient de noter que Statistiques Canada tient compte de la contrebande de tabac dans les Comptes Nationaux depuis juin 1994.

³⁵Mentionnons toutefois que les dépenses de tabac (et d'alcool) sont systématiquement sous-estimées dans les enquêtes de consommation. L'explication habituellement avancée pour expliquer ce phénomène est que le ou les membres du ménage qui consomment du tabac ou de l'alcool en grande quantité sont réticents à avouer, au reste de la famille, l'importance des sommes d'argent qu'ils consacrent à la consommation de ces biens. Néanmoins, ce biais de réponse ne modifie pas l'analyse de la contrebande dans la mesure où il demeure constant dans le temps.

Les conclusions à tirer de cette observation semblent plutôt encourageantes pour l'approche fondée sur la Comptabilité Nationale. Dans ce cas précis où l'économie souterraine a connu un essor considérable, il semble que les estimations de la taille de l'économie souterraine fondées sur cette approche se sont avérées très crédibles. Bien qu'il soit difficile de généraliser sur la base de cette expérience particulière, l'exemple décrit indique qu'il est souvent raisonnable de croire que les dépenses déclarées par les consommateurs dans les enquêtes auprès des ménages comprennent à la fois les dépenses sur le marché officiel et sur le marché noir.³⁶ A l'appui de cette proposition, notons qu'il incombe au vendeur et non à l'acheteur de verser les taxes dues au gouvernement et que l'anonymat est garanti dans la plupart des enquêtes de consommation. Ainsi, l'acheteur ne risque à peu près rien en révélant le montant de ses dépenses effectuées dans l'économie souterraine et semble être disposé à le faire pour une enquête de Statistique Canada. Il est également raisonnable de penser à plus forte raison qu'il fera de même à l'occasion d'une enquête comme la nôtre qui est administrée par une agence de sondage bien connue au Canada plutôt que par des agents du gouvernement.

³⁶Le cas des loyers constitue un autre exemple où des revenus non-déclarés au fisc (revenus de location) sont bien mesurés par des instruments utilisés dans les Comptes Nationaux, comme les enquêtes sur les dépenses de consommation.

Annexe 1B - Taux marginal implicite de taxation

Familles monoparentales

La situation des familles monoparentales diffère considérablement selon le nombre d'enfants et les frais de garde payés. Nous avons choisi de nous concentrer sur le cas d'une famille monoparentale, ayant un enfant de moins de 6 ans et dont les frais de garde s'élèvent à 5 000\$ (environ 3 200 euros). Ces critères correspondent aux conditions d'obtention d'un crédit d'impôt pour frais de garde au Québec.³⁷

La prise en compte des mesures de crédits d'impôts s'avère particulièrement importante pour les familles monoparentales. Comme en témoigne le *tableau 1B-1*, le phénomène de la taxation marginale implicite pose surtout problème pour la zone de revenu comprise entre 30 000 et 35 000\$. Dans ce cas, le taux de taxation atteint le niveau maximal de 85,2%. Il diminue ensuite rapidement et se stabilise dès que le revenu approche de 40 000\$.

Par ailleurs, il importe de souligner que les taux marginaux implicites de taxation peuvent inciter à la reprise d'activité des familles monoparentales, puisqu'ils ne sont que de 15,7%. Ils restent, en outre, limités pour les faibles revenus d'activité, ce qui peut favoriser l'accroissement du temps de travail de cette catégorie de ménages.

Enfin, le *tableau 1B-1* permet également de repérer le moment à partir duquel la famille monoparentale n'est plus bénéficiaire net de l'Etat, mais contribuable net. Cette situation se produit lorsque le revenu atteint 30 000\$.

³⁷ A titre de comparaison, il convient de noter que le système français prévoit la même disposition, si ce n'est que le plafond des frais de garde est porté à 2 300 euros au lieu de 3 200 (*Loi 199 Quater D du Code Général des Impôts*). En revanche, la limite d'âge pour l'enfant est identique à celle du Québec, soit six ans.

Tableau 1B-1 :
Taux marginaux implicites de taxation
(année d'imposition 1999)

Tranche de revenu (en dollars can.)	Contribution nette (en dollars can.)	Taux de contribution nette ^a (en %)	Taux marginal implicite de taxation (en %)	Ensemble des ménages (en %)
0 - 5000\$	11 036	-	15,7	35,9
5000 - 10000\$	11 819	-	27,8	9,1
10000 - 15000\$	10 429	-	27,5	9,1
15000 - 20000\$	9 053	-	60,2	7,1
20000 - 25000\$	6 045	-	43,5	5,9
25000 - 30000\$	3 868	-	79,0	7,0
30000 - 35000\$	-84	0,3	85,2	5,8
35000 - 40000\$	-4 342	12,4	68,3	4,5
40000 - 45000\$	-7 754	19,4	55,1	3,8
45000 - 50000\$	-10 509	23,4	48,4	3,0
50000 - 55000\$	-12 928	25,9	52,9	2,4
55000 - 60000\$	-15 573	28,3	51,0	1,8
60000 - 65000\$	-18 124	30,2	56,9	1,5
65000 - 70000\$	-20 969	32,3	59,1	0,8
70000 - 75000\$	-23 924	34,2	56,8	0,6
75000\$ et plus	-26 764	35,7	52,2	1,8
TOTAL				100,0

a - Impôts nets des transferts en pourcentage du revenu. Le pourcentage est calculé au minimum de la tranche de revenus.

Source : Gouvernement du Québec, 1999.

Le *tableau 1B-2* met en évidence qu'une famille monoparentale ayant un revenu annuel de 30 000\$ est un contribuable net de l'Etat, pour une somme cependant très limitée de 84\$. Son revenu disponible s'établit alors à 29 916\$. Lorsque le revenu de travail atteint 35 000\$, cette famille doit supporter une contribution nette de 4 342\$ et le revenu disponible est égal à 30 658\$. Par rapport à la situation précédente, le revenu disponible a seulement augmenté de 742\$.

Dans cet exemple, l'augmentation de revenu de 30 000 à 35 000\$ s'accompagne d'un accroissement de la contribution nette de 4 285\$, soit 85,2% de l'augmentation de revenu obtenue. Ce taux marginal de taxation implicite est pour l'essentiel imputable, non pas à la baisse des transferts (- 418\$), mais à la hausse des impôts (+3 840\$). Une nuance importante doit toutefois être apportée. En effet, une partie de l'augmentation du prélèvement fiscal est, en fait, liée à la réduction de ce que l'on peut appeler des dispositions socio-fiscales (-2 077\$),³⁸ qui sont des réductions d'impôt appliquées aux revenus les plus bas et de même nature que les transferts sociaux.

³⁸ *i.e.* crédit d'impôt remboursable pour les taxes de ventes du Québec, crédit d'impôt remboursable pour frais de garde, remboursement d'impôts fonciers, etc.

Tableau 1B-2 :

Exemple de calcul du taux marginal implicite

	Revenu de travail		Variation
	30 000	35 000	
A. Revenu de travail	30 000	35 000	5 000
B. Gouvernement du Québec			
- Prélèvements fiscaux :			
. Impôt ²	-3 338	-4 084	-746
. Crédits socio-fiscaux ³	4 477	2 604	-1 873
. Cotisations au RRQ ⁴	-935	-1 113	-178
- Transferts sociaux :			
Allocation familiale	131	131	0
Sous-total	335	-2 462	-2 797
C. Gouvernement du Canada			
- Prélèvements fiscaux :			
. Impôt ¹	-1 493	-2 204	-711
. Crédits socio-fiscaux ⁵	503	299	-204
. Cotisations au RRQ	-765	-893	-128
- Transferts sociaux :			
Prestations fiscales pour enfants	1 336	918	-418
Sous-total	-419	-1 880	-1 461
D. Contribution nette (B+C)	-84	-4 342	-4 258
E. Taux marginal implicite de Taxation (D/A)	-	-	85,2 %
F. Revenu disponible (A+D)	29 916	30 658	742

1- L'enfant est âgé de moins de 6 ans et les frais de garde sont de 5 000\$.

2- Excluant les crédits d'impôt non-remboursables réductibles.

3- Incluant les crédits remboursable pour la TVQ, le remboursement d'impôts fonciers, le crédit d'impôt remboursable pour frais de garde, le crédit d'impôt pour personne vivant seule, la réduction à l'égard des familles.

4- Incluant la cotisation au Fonds de lutte contre la pauvreté par la réinsertion au travail.

5- Incluant les crédits d'impôt non-remboursables réductibles et le crédit d'impôt remboursable pour TPS.

Source : Gouvernement du Québec, 1999.

Il faut ajouter que, dans l'exemple retenu, la famille monoparentale ne bénéficie pas du système de frais de garde à 5\$, mais du crédit d'impôt défini pour ces mêmes frais de garde. Or, cette situation accentue encore les taux de taxation implicite puisque ce soutien est défini en fonction du revenu. C'est pourquoi, le crédit d'impôt en question est, depuis peu, graduellement remplacé par un dispositif universel, nommé programme des services de garde à contribution réduite de 5\$ par jour. Le remplacement d'un programme défini en fonction du revenu par un programme universel a pour conséquence de diminuer le taux marginal implicite de taxation, pour les contribuables concernés.

Le tableau 1B-3 compare la situation des familles monoparentales bénéficiant du système de garde à 5\$ à celle des familles qui accèdent au crédit d'impôt pour frais de garde. L'on observe ainsi que le taux marginal implicite de taxation est passé de 95,3% à 67,2% lorsque le revenu du contribuable a augmenté de 30 000\$ à 31 000\$.

Le nombre de famille monoparentales dont le taux marginal implicite de taxation excède 60%, avoisine 38 500, ce qui représente 24% des familles monoparentales.³⁹ Il

³⁹ La population considérée comprend ici toutes les familles monoparentales ayant un enfant à charge, qu'elles aient ou non des frais de garde et quel que soit l'âge de l'enfant.

est, dès lors, raisonnable de penser que les familles monoparentales soient portées à travailler au noir, bien que des contraintes de temps puissent les empêcher de travailler autant qu'elles le souhaiteraient sur ce marché.

Tableau 1B-3 :
Exemple de calcul du taux marginal implicite

	Famille monoparentale		Personne seule
	Sans garde à 5\$	Avec garde à 5\$	
Taux marginal – régime québécois			
Impôt sur le revenu			
. Table	23,0	23,0	23,0
. Crédit personne vivant seule	3,5	3,5	3,5
. Réduction d'impôt des familles	6,0	6,0	
Crédit pour frais de garde	25,0		
Crédit pour TVQ et RIF ¹	6,0	6,0	6,0
Cotisations au RRQ	3,5	3,5	3,5
<i>Sous-total – Québec</i> ²	67,0	42,0	36,0
Taux marginal – régime fédéral			
	28,3	25,2	28,4
Taux marginal total	95,3	67,2	64,4
1- RIF : remboursement d'impôts fonciers			
2- Excluant le Fonds de lutte contre la pauvreté par la réinsertion au travail			

Source : Gouvernement du Québec, 1999.

Couples sans enfants

Les couples sans enfants ont généralement des taux marginaux implicites de taxation inférieurs à 60%, du moins lorsque les impôts sont supérieurs aux montants reçus en transfert. Ces taux restent relativement modérés lors d'une reprise d'activité. Mais, pour de bas revenus de travail, la taxation marginale implicite peut atteindre jusqu'à 100%. Le *tableau 1B-4* présente les taux marginaux auxquels font face les couples sans enfants, ayant un seul revenu de travail.

Le taux marginal implicite de taxation de cette catégorie de ménage s'échelonne entre 22% et 100%. Un couple sans enfants ayant un revenu de travail de 50 000\$ verse, par exemple, aux gouvernements une contribution nette de 15 665\$. Dans ce cas, une hausse de revenus de 5 000\$ sera à moitié absorbée par la fiscalité fédérale et québécoise, soit 48,5%. A un revenu de travail de 15 000\$, il verse une contribution nette de 188\$ et fait alors face à un taux marginal implicite de 21,6%.

En revanche, lorsqu'un couple est bénéficiaire net de l'Etat, il peut être confronté à un taux de 100%. Ainsi, un ménage gagnant un revenu de travail de 5 000\$ reçoit une aide nette de la part des gouvernements de l'ordre de 8 599\$. Son revenu disponible global est alors de 13 599\$. Mais, si le revenu de travail augmente de 5 000\$, il sera totalement absorbé (le taux marginal implicite est de 100%). L'incitation au travail est, dans ce cas, fortement réduite, voire totalement annihilée. Or, près de 6% des

couples sans enfants sont précisément dans cette situation et 11% environ se situent dans une zone de revenus où le taux implicite dépasse 75%, ce qui correspond à près de 17 000 ménages. Par conséquent, les couples sans enfants sont également susceptibles de participer activement aux activités souterraines et, en particulier, ceux dont les revenus sont les plus faibles. Cette fois encore, ce sont donc les catégories de ménages à bas revenus qui sont touchés par le phénomène. Les couples sans enfants constituent potentiellement une part importante des travailleurs au noir, dans la mesure où leur taux marginal de taxation implicite est le plus élevé.

Tableau 1B-4 :
Taux marginaux implicites de taxation
(année d'imposition 1999)

Tranche de revenu (en dollars can.)	Contribution nette (en dollars can.)	Taux de contribution nette ^a (en %)	Taux marginal implicite de taxation (en %)	Ensemble des ménages (en %)
0 - 5000\$	9 699	-	22,0	44,8
5000 - 10000\$	8 599	-	100,0	5,6
10000 - 15000\$	3 599	-	75,7	5,1
15000 - 20000\$	-188	1,3	21,6	4,9
20000 - 25000\$	-1 265	6,3	38,7	4,6
25000 - 30000\$	-3 201	12,8	51,4	4,9
30000 - 35000\$	-5 768	19,2	59,4	4,7
35000 - 40000\$	-8 737	25,0	49,0	4,3
40000 - 45000\$	-11 187	28,0	44,8	4,2
45000 - 50000\$	-13 426	29,8	44,8	3,4
50000 - 55000\$	-15 665	31,3	48,5	3,1
55000 - 60000\$	-18 091	32,9	49,4	2,5
60000 - 65000\$	-20 561	34,3	51,6	1,9
65000 - 70000\$	-23 141	35,6	51,9	1,5
70000 - 75000\$	-25 738	36,8	52,2	1,0
75000\$ et plus	-28 347	37,8	52,2	3,6
TOTAL				100,0

a - Impôts nets des transferts en pourcentage du revenu. Le pourcentage est calculé au minimum de la tranche de revenus.

Source : Gouvernement du Québec, 1999.

Couples avec enfants

Tout comme pour les familles monoparentales, les taux marginaux implicites de taxation varient en fonction du nombre d'enfants et l'importance des frais de garde payés. Dans ce qui suit, nous considérons le cas d'un couple ayant deux revenus, deux enfants de moins de six ans et des frais de garde de 5 000\$ par enfant.

De façon générale, le taux marginal implicite de taxation n'excède pas 60%. Il s'élève néanmoins à 73,7% pour les revenus compris entre 30 000 et 35 000\$. Cela signifie qu'une augmentation de 5 000\$ du revenu de travail sera prélevé à près de 74%. Le taux de taxation implicite marginal est également important lorsque les revenus

de travail passent de 35 000 à 40 000\$. Dans ce cas, le ménage devient contribuable net de l'Etat. Ces couples, qui exercent chacun une activité professionnelle, ont alors peu d'incitations à accroître leur intensité de travail. Comparativement à une famille monoparentale, le prélèvement opéré sur cette tranche de revenus est plus élevée, ce qui tend à rendre le marché noir plus attractif pour cette catégorie de ménages. Parmi l'ensemble des couples, ayant deux enfants à charge, sans considération du montant des frais de garde, 18% environ font face à un taux implicite dépassant 60%. Cela représente quelque 48 000 ménages susceptibles d'être attirés par une activité dissimulée.

Tableau 1B-5 :
Taux marginaux implicites de taxation
(année d'imposition 1999)

Tranche de revenu (en dollars can.)	Contribution nette (en dollars can.)	Taux de contribution nette ^a (en %)	Taux marginal implicite de taxation (en %)	Ensemble des ménages (en %)
0 - 5000\$	15 467	-	29,8	1,8
5000 - 10000\$	16 956	-	35,3	2,0
10000 - 15000\$	15 191	-	53,3	2,9
15000 - 20000\$	12 525	-	33,5	3,3
20000 - 25000\$	10 850	-	38,8	4,1
25000 - 30000\$	8 911	-	38,8	5,0
30000 - 35000\$	6 972	-	73,7	5,7
35000 - 40000\$	3 288	9,4	69,4	6,0
40000 - 45000\$	-182	0,5	52,7	6,3
45000 - 50000\$	-2 816	6,3	56,1	6,4
50000 - 55000\$	-5 620	11,2	57,1	6,5
55000 - 60000\$	-8 474	15,4	52,0	6,4
60000 - 65000\$	-11 072	18,5	60,5	5,9
65000 - 70000\$	-14 098	21,7	59,0	5,8
70000 - 75000\$	-17 046	24,4	56,9	4,7
75000\$ et plus	-19 893	26,5	45,6	27,5
TOTAL				100,0

a - Impôts nets des transferts en pourcentage du revenu. Le pourcentage est calculé au minimum de la tranche de revenus.

Source : Gouvernement du Québec, 1999.

Personnes âgées de 65 ans et plus

Aux fins de l'analyse, nous avons retenu le cas d'une personne âgée, vivant seule et recevant un revenu de retraite (Régime Enregistré d'Épargne Retraite, par exemple). En effet, la majorité d'entre elles ne perçoit pas ou peu de revenu de travail.

Le tableau 1B-6 révèle que les personnes âgées de 65 ans et plus ont un taux marginal implicite de taxation souvent proche de 60%, notamment lorsqu'elles sont contribuables nets de l'Etat. Ainsi, une personne vivant seule et ayant un revenu de retraite de 25 000\$ verse aux gouvernements une contribution nette de 714\$. Dans ce cas, une hausse de revenu de 5 000\$, au titre de la retraite, sera absorbée à hauteur de 58,2%. A 60 000\$, elle est confrontée à un taux atteignant 60,4%, en raison de la récupération fédérale de la pension vieillesse (15%) à compter d'un revenu de 53 215\$.

Tableau 1B-6:
Taux marginaux implicites de taxation
(année d'imposition 1999)

Tranche de revenu (en dollars can.)	Contribution nette (en dollars can.)	Taux de contribution nette ^a (en %)	Taux marginal implicite de taxation (en %)	Ensemble des ménages (en %)
0 - 5000\$	12 694	-	49,6	79,4
5000 - 10000\$	10 215	-	62,2	9,5
10000 - 15000\$	7 103	-	63,2	4,6
15000 - 20000\$	3 943	-	42,3	2,6
20000 - 25000\$	1 829	-	50,9	1,6
25000 - 30000\$	-714	2,9	58,2	0,9
30000 - 35000\$	-3 623	12,1	53,4	0,5
35000 - 40000\$	-6 292	18,0	52,3	0,3
40000 - 45000\$	-8 905	22,3	51,6	0,2
45000 - 50000\$	-11 486	25,5	55,6	0,1
50000 - 55000\$	-14 268	28,5	57,6	0,1
55000 - 60000\$	-17 148	31,2	59,8	0,0
60000 - 65000\$	-20 137	33,6	60,4	0,0
65000 - 70000\$	-23 155	35,6	60,4	0,0
70000 - 75000\$	-26 172	37,4	60,4	0,0
75000\$ et plus	-29 190	38,9	60,4	0,0
TOTAL				100,0

a - Impôts nets des transferts en pourcentage du revenu. Le pourcentage est calculé au minimum de la tranche de revenus.

Source : Gouvernement du Québec, 1999.

Cependant, lorsqu'une personne âgée est bénéficiaire net de l'Etat, il peut arriver qu'elle soit confrontée à un taux marginal implicite élevé. Tel est le cas d'un pensionné dont le revenu de retraite est de 5 000\$. Il reçoit alors un soutien des gouvernements de 10 215\$ et fait face à un taux marginal implicite de 62,2%, à cause principalement de la réduction du supplément de revenu garanti (50%) par le gouvernement fédéral.

La majorité des personnes âgées se situe dans une zone de revenu où le taux marginal implicite de taxation est inférieur à 60%. Mais, pour 14% d'entre elles, ce taux est supérieur. Cela représente un peu plus de 55 500 ménages pour lesquels l'attrait du marché noir peut s'avérer important. Une activité dissimulée leur permettrait, en effet, de générer des revenus supplémentaires sans toutefois perdre le bénéfice des prestations sociales et sans devoir subir un accroissement de leurs impôts. Une nouvelle fois, ce sont les ménages à faibles revenus qui sont les plus touchés par la taxation marginale implicite et qui seront donc les plus tentés de travailler au noir.

Annexe 1C - Démonstration de l'équiprobabilité de sélection des ménages

Cette annexe a pour but de démontrer qu'à l'intérieur de chaque strate, les probabilités de sélection des ménages sont approximativement égales, que ces ménages soient inscrits ou non dans un annuaire téléphonique. La démonstration est réalisée pour les régions de Montréal et de Québec, où les grappes sont constituées de 30 ménages. Mais, il est facile de la généraliser à n'importe quelle taille de grappe.

En premier lieu, la probabilité qu'un ménage appartienne à une grappe choisie peut s'exprimer de la manière suivante :

$$P(\text{grappe}) = G \times [Ntel(\text{grappe}) / Ntel],$$

où $Ntel(\text{grappe})$ est le nombre de logements résidentiels habités, dont le numéro de téléphone est inscrit dans l'annuaire, parmi les 30 logements précédant et incluant le logement du ménage en question. Par «précédant», on entend correspondant à un déplacement, contraire aux aiguilles d'une montre, à partir du logement du ménage en question. $Ntel$ est le nombre de numéros de téléphone résidentiels inscrits dans les annuaires de la strate; et G est le nombre de grappes choisies dans la strate.

Par ailleurs, la probabilité qu'un logement soit choisi, s'il appartient à une grappe sélectionnée, vaut :

$$P(\text{sélection} / \text{grappe}) = 5/30 = 1/6.$$

Nous supposons, en effet, que 5 ménages sur 30 acceptent de participer à l'enquête. L'hypothèse que le nombre de ménages, choisis au sein d'une même grappe, soit de 5 découle d'un compromis entre d'une part, la nécessité de définir une charge de travail suffisante à un enquêteur se déplaçant en un endroit donné, et d'autre part, l'avantage méthodologique à restreindre le plus possible la taille des grappes.

La probabilité qu'un ménage particulier soit choisi peut donc s'exprimer ainsi :

$$\begin{aligned} P(\textit{selection}) &= P(\textit{grappe}) \times P(\textit{sélection/grappe}), \\ P(\textit{sélection}) &= \frac{G \times [Ntel(\textit{grappe}) / Ntel]}{6}. \end{aligned} \quad (1D-1)$$

Or, le nombre de numéros de téléphone résidentiels d'une région donnée ne représente en fait qu'une fraction F par rapport à l'ensemble des logements résidentiels puisque certains ménages (moins de 3%, en général) n'ont pas le téléphone et d'autres ne sont pas inscrits dans l'annuaire (environ 15 à 20% dans la région de Montréal, par exemple). On peut donc écrire :

$$Ntel = F \times N, \quad (1D-2)$$

où N est le nombre de logements réels dans la strate.

En outre, à l'intérieur d'une même strate, il est raisonnable de supposer que la fraction F est, en général, assez uniforme sur tout le territoire, de sorte que l'approximation suivante :

$$Ntel(\textit{grappe}) = F \times 30, \quad (1D-3)$$

est empiriquement acceptable.

A l'aide des relations (1D-2) et (1D-3), on peut réécrire (1D-1) de la manière suivante :

$$P(\textit{sélection}) = (5 \times G) / N = n/N,$$

où $n = 5 \times G$ et correspond au nombre de sélections dans la strate.

Par conséquent, la probabilité de sélection de chaque ménage inscrit ou non dans un annuaire téléphonique est égale à la fraction d'échantillonnage global n/N . Le seul facteur qui peut rendre les probabilités inégales est la variation, d'une grappe à l'autre à l'intérieur de la même strate, de la proportion de ménages inscrits dans l'annuaire utilisé pour identifier les départs aléatoires.

On peut facilement vérifier que cette relation est maintenue même si l'une des grappes comprend moins de 30 logements résidentiels habités, puisque la procédure fait en sorte que le nombre de logements choisis est toujours égale à 5. La probabilité plus faible de choisir une telle grappe est compensée par une probabilité plus grande de sélectionner chaque ménage de la grappe, si la grappe est choisie. Evidemment, il faut qu'il y ait plus de 4 logements habités dans la grappe ; lorsque ce n'est pas le cas, il faut alors procéder à une pondération appropriée pour être rigoureux.

Annexe 2A - Estimations de première étape

Tableau 2A.1 :
Revenu virtuel hors-travail

<i>Variables</i>	<i>Coefficient</i>	<i>Ecart-type</i>	
Constante	7.8392	3.9911	
Age	0.4493	0.1885	**
Age au carré	-7.5726	1.9746	***
Femme	-2.3862	0.8502	***
Célibataire	2.7955	0.1761	
Conjoint	-0.0154	1.4679	
Monoparent	4.6216	2.3115	**
Années de scolarité	3.8592	1.1665	***
Années de scolarité au carré	1.2574	0.2653	***
Années de scolarité au cube	-6.0392	1.5959	***
Montréal	0.8823	0.9432	
Bas-du-Fleuve	-0.7882	1.1362	
Nombre d'enfants	-0.0798	0.4718	
R ²	0.0470		
Nombre d'observations	4988		

*** Statistiquement significatif à 1%, ** Statistiquement significatif à 5%, * Statistiquement significatif à 10%

Tableau 2A.2 :
Revenu d'activité déclarée

<i>Variables</i>	<i>Coefficient</i>	<i>Ecart-type</i>	
Constante	-6.0419	1.8098	
Age	7.1691	1.3620	***
Age au carré	-1.5051	0.2978	***
Age au cube	8.4015	2.0385	***
Femme	-5.0047	1.7704	***
Célibataire	1.1327	3.7583	
Conjoint	5.3799	3.1282	*
Monoparent	2.0183	4.8666	
Années de scolarité	-1.2142	0.2459	***
Années de scolarité au carré	3.4880	0.5571	***
Années de scolarité au cube	-1.9320	0.3340	***
Montréal	2.3680	1.9639	
Bas-du-Fleuve	-5.2927	2.3659	**
Nombre d'enfants	1.3721	1.0094	
R ²	0.0700		
Nombre d'observations	4988		

*** Statistiquement significatif à 1%, ** Statistiquement significatif à 5%, * Statistiquement significatif à 10%.

Annexe 2B - La variable d'expérience comme instrument

Nous proposons de vérifier si l'expérience acquise sur le marché noir est endogène aux heures de travail non déclarée. A l'instar de Mroz (1987), nous avons estimé une équation de salaires à partir des travailleurs au noir en incluant les variables d'expérience, puis en les excluant. Les salaires prédits à l'aide de ces deux ensemble de paramètres sont ensuite utilisés comme variables exogènes dans une équation d'heures de travail linéaire standard pour tout l'échantillon. Les résultats de ces régressions apparaissent aux tableaux 2B.1 et 2B.2.

Tableau 2B.1 :
Salaire au noir

<i>Variables</i>	<i>Coefficient</i>	<i>Ecart-type</i>		<i>Coefficient</i>	<i>Ecart-type</i>	
Constante	-2.9256	1.1647		-3.0412	1.1606	
Age	0.2065	0.0840	**	0.2049	0.0842	**
Age au carré	-4.2276	1.8797	**	-4.2516	1.8812	**
Age au cube	2.7130	1.2782	**	2.7653	1.2779	**
Femme	-0.0455	0.1597		-0.0220	0.1625	
Montréal	-0.0739	0.1798		-0.0543	0.1801	
Bas-du-Fleuve	-0.2717	0.1754		-0.2866	0.1757	
Années de scolarité	-0.2673	0.3407		-0.3111	0.3475	
Années de scolarité au carré	0.2711	0.2604		0.2858	0.2638	
Expérience	-0.3145	0.1545		-	-	
R ²		0.1430			0.1305	
Ecart-type de régression		5.2245			5.2255	

*** Statistiquement significatif à 1%, ** Statistiquement significatif à 5%, * Statistiquement significatif à 10%.

La variable $\text{Log}(\text{salaire})$ du tableau 2B.2 est le salaire prédit à partir des paramètres du tableau 2B.1. Les colonnes 1 et 3 correspondent aux modèles 1 et 2 du tableau 2B.1, respectivement. Nous remarquons que le paramètre estimé associé aux salaires prédits est très sensible à l'inclusion ou non de l'expérience dans l'équation de salaires. Les résultats ci-dessus doivent être interprétés comme une preuve que la variable d'expérience n'est pas un instrument valable dans l'équation de salaires.

Tableau 2B.2 :

Heures de travail au noir

<i>Variables</i>	<i>Coefficient</i>	<i>Ecart-type</i>		<i>Coefficient</i>	<i>Ecart-type</i>	
Constante	-2.7711	0.2725		-2.8588	0.2683	
Age	2.2152	0.2289	***	2.3025	0.2245	***
Age au carré	-4.3851	0.5181	***	-4.5870	0.5075	***
Age au cube	2.4349	0.3538	***	2.5731	0.3465	***
Femme	-1.1189	0.2235	***	-1.1201	0.2236	***
Nbre d'enfants <6 ans	-5.7490	2.0295	***	-5.8053	2.0288	***
Années de scolarité	2.9162	0.5228	***	2.9607	0.5226	***
Taux marginal d'imposition	1.4148	0.0749	***	1.4151	0.0749	***
Probabilité de détection	-1.1539	0.4402	***	-1.1802	0.4400	***
Pénalités	-0.7960	-1.4590		-0.8014	0.5454	
Proportion dans l'entourage	-8.8864	6.8959		-8.7759	6.9882	
Réaction de l'entourage	4.3795	2.4559	*	4.4885	2.4531	*
Immoralité	3.6515	1.3296	***	3.6526	1.3299	***
Log (salaire) au noir	2.7136	3.9065		0.7748	3.8244	
R ²		0.3040			0.0974	
Ecart-type de régression		1.6706			1.7243	

*** Statistiquement significatif à 1%, ** Statistiquement significatif à 5%, * Statistiquement significatif à 10%.

Annexe 3A - Les fonctions d'offre de travail

Dans le régime 1, les fonctions d'offre de travail sont obtenues en posant :

$$\begin{aligned}
 m_1(h_1, h_2, z) = & \bar{\alpha}_1 + X'_1 \delta_1 + \alpha_3 w_1 + \beta_{11} h_1 + \beta_{12} h_2 \\
 & + \beta_{13} (y^v + 2w_1 h_1 + h_2 E w_2) + \beta_{23} w_1 h_2 \\
 & + \beta_{33} w_1 (y^v + w_1 h_1 + h_2 E w_2) + \varepsilon_1 = 0
 \end{aligned} \tag{3A.1}$$

$$\begin{aligned}
 m_2(h_1, h_2, z) = & \bar{\alpha}_2 + X'_2 \delta_2 + \alpha_3 E w_2 + \beta_{12} h_1 + \beta_{13} h_1 E w_2 \\
 & + \beta_{22} h_2 + \beta_{23} (y^v + w_1 h_1 + 2h_2 E w_2) \\
 & + \beta_{33} [(y^v + w_1 h_1) E w_2 + h_2 E w_2^2] + \varepsilon_2 = 0
 \end{aligned} \tag{3A.2}$$

et en résolvant pour h_1 et h_2 . Les fonctions d'offre de travail associées à ce régime sont les suivantes :

$$h_1 = \frac{\gamma \zeta_1 - \gamma \varepsilon_1 - \beta \zeta_2 + \beta \varepsilon_2}{(\alpha \gamma - \beta^2)} \tag{3A.3}$$

$$h_2 = \frac{-\beta \zeta_1 + \beta \varepsilon_1 + \alpha \zeta_2 - \alpha \varepsilon_2}{(\alpha \gamma - \beta^2)} \tag{3A.4}$$

où

$$\begin{aligned}
 \alpha &= \beta_{11} + 2\beta_{13} w_1 + \beta_{33} w_1^2 \\
 \beta &= \beta_{12} + \beta_{13} E w_2 + \beta_{23} w_1 + \beta_{33} w_1 E w_2 \\
 \gamma &= \beta_{22} + 2\beta_{23} E w_2 + \beta_{33} E w_2^2 \\
 \zeta_1 &= -\bar{\alpha}_1 - X'_1 \delta_1 - \alpha_3 w_1 - \beta_{13} y^v - \beta_{33} y^v w_1 \\
 \zeta_2 &= -\bar{\alpha}_2 - X'_2 \delta_2 - \alpha_3 E w_2 - \beta_{23} y^v - \beta_{33} y^v E w_2
 \end{aligned}$$

De même, en posant $h_2 = 0$ dans (3A.1), nous obtenons :

$$m_1(h_1, 0, z) = \bar{\alpha}_1 + X'_1 \delta_1 + \alpha_3 w_1 + \beta_{11} h_1 + \beta_{13} (y^v + 2w_1 h_1) + \beta_{33} w_1 (y^v + w_1 h_1) + \varepsilon_1 = 0$$

et

$$\begin{aligned}
 m_2(h_1, 0, z) = & \bar{\alpha}_2 + X'_2 \delta_2 + \alpha_3 E w_2 + \beta_{12} h_1 + \beta_{13} h_1 E w_2 + \beta_{23} (y^v + w_1 h_1) \\
 & + \beta_{33} (y^v + w_1 h_1) E w_2 + \varepsilon_2 \leq 0
 \end{aligned}$$

En résolvant pour h_1 , nous obtenons les fonctions d'offre contraintes associées au régime 2 :

$$\bar{h}_1 = \frac{\zeta_1 - \varepsilon_1}{\alpha} \quad (3A.5)$$

Enfin, en posant $h_1 = 0$ dans (3A.2), nous obtenons :

$$m_2(0, h_2, z) = \bar{\alpha}_2 + X'_2 \delta_2 + \alpha_3 Ew_2 + \beta_{22} h_2 + \beta_{23} (y^v + 2h_2 Ew_2) + \beta_{33} (y^v Ew_2 + h_2 Ew_2^2) + \varepsilon_2 = 0$$

et

$$m_1(0, h_2, z) = \bar{\alpha}_1 + X'_1 \delta_1 + \alpha_3 w_1 + \beta_{12} h_2 + \beta_{13} (y^v + h_2 Ew_2) + \beta_{23} w_1 h_2 + \beta_{33} w_1 (y^v + h_2 Ew_2) + \varepsilon_1 \leq 0$$

En résolvant pour h_2 , nous obtenons la fonction d'offre contrainte associée au régime 3 :

$$\bar{h}_2 = \frac{\zeta_2 - \varepsilon_2}{\gamma} \quad (3A.6)$$

Annexe 3B - Les fonctions d'offre de travail sous l'hypothèse de substitution parfaite

Dans le régime 1, les fonctions d'offre de travail sont obtenues en posant :

$$m_1(h_1, h_2, z) = \bar{\gamma}_1 + \gamma_2 w_1 + \delta_{11} h_1 + \delta_{11} h_2 + \delta_{12}(y^v + 2w_1 h_1 + h_2 E w_2) + \delta_{12} w_1 h_2 + \delta_{22} w_1 (y^v + w_1 h_1 + h_2 E w_2) + \mu_1 = 0 \quad (3B.1)$$

$$m_2(h_1, h_2, z) = \bar{\gamma}_1 + \gamma_2 E w_2 + \delta_{11} h_2 + \delta_{11} h_1 + \delta_{12} h_1 E w_2 + \delta_{12}(y^v + w_1 h_1 + 2h_2 E w_2) + \delta_{22} E w_2 (y^v + w_1 h_1 + h_2 E w_2) + \mu_2 = 0 \quad (3B.2)$$

et en résolvant pour h_1 et h_2 . Les fonctions d'offre de travail associées à ce régime sont les suivantes :

$$h_1 = \frac{\gamma \zeta_1 - \gamma \mu_1 - \beta \zeta_2 + \beta \mu_2}{(\alpha \gamma - \beta^2)} \quad (3B.3)$$

$$h_2 = \frac{-\beta \zeta_1 + \beta \mu_1 + \alpha \zeta_2 - \alpha \mu_2}{(\alpha \gamma - \beta^2)} \quad (3B.4)$$

où

$$\begin{aligned} \alpha &= \delta_{11} + \delta_{12} w_1 + \delta_{22} w_1^2 \\ \beta &= \delta_{11} + \delta_{12} E w_2 + \delta_{12} w_1 + \delta_{22} w_1 E w_2 \\ \gamma &= \delta_{11} + \delta_{12} E w_2 + \delta_{22} E w_2^2 \\ \zeta_1 &= -\bar{\gamma}_1 - \gamma_2 w_1 - \delta_{12} y^v - \delta_{12} w_1 y^v \\ \zeta_2 &= -\bar{\gamma}_1 - \gamma_2 E w_2 - \delta_{12} y^v - \delta_{12} y^v E w_2 \end{aligned}$$

De même, en posant $h_2 = 0$ dans (3B.1), nous obtenons :

$$m_1(h_1, 0, z) = \bar{\gamma}_1 + \gamma_2 w_1 + \delta_{11} h_1 + \delta_{12}(y^v + w_1 h_1) + \delta_{22} w_1 (y^v + w_1 h_1) + \mu_1 = 0$$

et

$$m_2(h_1, 0, z) = \bar{\gamma}_1 + \gamma_2 E w_2 + \delta_{11} h_1 + \delta_{12} h_1 E w_2 + \delta_{12}(y^v + w_1 h_1) + \delta_{22} E w_2 (y^v + w_1 h_1) + \mu_2 \leq 0$$

En résolvant pour h_1 , nous obtenons les fonctions d'offre contraintes associées au régime 2 :

$$\bar{h}_1 = \frac{\zeta_1 - \mu_1}{\alpha} \quad (3B.5)$$

Enfin, en posant $h_1 = 0$ dans (3B.2), nous obtenons :

$$m_2(0, h_2, z) = \bar{\gamma}_1 + \gamma_2 Ew_2 + \delta_{11}h_2 + \delta_{12}h_1 Ew_2 + \delta_{12}(y^v + a_2 Ew_2) + \delta_{22}Ew_2(y^v + h_2 Ew_2) + \mu_2 = 0$$

et

$$m_1(0, h_2, z) = \bar{\gamma}_1 + \gamma_2 w_1 + \delta_{11}h_2 + \delta_{12}(y^v + h_2 Ew_2) + \delta_{12}w_1 h_2 + \delta_{22}w_1(y^v + h_2 Ew_2) + \mu_1 \leq 0$$

En résolvant pour h_2 , nous obtenons la fonction d'offre contrainte associée au régime 3 :

$$\bar{h}_2 = \frac{\zeta_2 - \mu_2}{\gamma} \quad (3B.6)$$

Annexe 3C - Les équations de salaire

Dans le **régime 2**, l'individu travaille uniquement sur le marché officiel ($h_1 > 0$ et $h_2 = 0$). Nous observons alors uniquement le salaire obtenu sur ce marché.

$$\begin{aligned}w_1 &= X'_4 \Psi_4 + \varepsilon_4 \\ \varepsilon_5 &= Ew_2 - X'_5 \Psi_5 \\ \text{et } \varepsilon_6 &= Ew_2^2 - X'_6 \Psi_6\end{aligned}\tag{3C.1}$$

A l'inverse, dans le **régime 3**, l'individu travaille uniquement sur le marché noir ($h_1 = 0$ et $h_2 > 0$). Par conséquent, nous observons uniquement le salaire de l'activité souterraine :

$$\begin{aligned}\varepsilon_4 &= w_1 - X'_4 \Psi_4 \\ Ew_2 &= X'_5 \Psi_5 + \varepsilon_5 \\ \text{et } Ew_2^2 &= X'_6 \Psi_6 + \varepsilon_6\end{aligned}\tag{3C.2}$$

Enfin, dans le **régime 4**, l'individu ne travaille sur aucun marché ($h_1 = 0$ et $h_2 = 0$), de sorte que nous obtenons :

$$\begin{aligned}\varepsilon_4 &= w_1 - X'_4 \Psi_4 \\ \varepsilon_5 &= Ew_2 - X'_5 \Psi_5 \\ \text{et } \varepsilon_6 &= Ew_2^2 - X'_6 \Psi_6\end{aligned}\tag{3C.3}$$

Annexe 3D - Les effets de groupe

résultats d'estimations avec douze groupes de référence

PARAMETRES	Coefficients	Ecart-T	PARAMETRES	Coefficients	Ecart-T
α	0.25	-			
β_{11}	-0.5285	0.036	***		
β_{12}	-0.3454	0.018	***		
β_{13}	-0.0976	0.006	***		
β_{22}	-0.2897	0.016	***		
β_{23}	-0.0978	0.019	***		
β_{33}	-0.1548	0.003	***		

MARCHÉ OFFICIEL			MARCHÉ NOIR			
PARAMETRES	Coefficients	Ecart-T	PARAMETRES	Coefficients	Ecart-T	
HEURES OFFICIELLES /1000			HEURES NON-OFFICIELLES /1000			
Nb indiv. dans le ménage	-0.5364	0.073	***	Nb indiv. dans le ménage	-2.4728	0.064
Enfants (<6ans)	-0.0273	0.009	***	Enfants (<6ans)	0.0165	0.005
Célibataire	-0.0084	0.012		Célibataire	-0.0044	0.009
Marié	0.1452	0.032	***	Marié	0.1052	0.016
Famille mono-parentale	0.2054	0.027	***	Famille mono-parentale	0.1029	0.015
Femme	0.1291	0.037	***	Femme	0.0742	0.022
Age	-0.1012	0.019	***	Age	-0.0454	0.018
Age	0.4678	0.135	***	Age	0.1618	0.121
Groupe 1	-0.0160	0.046		Proportion de travailleurs au noir	0.0029	0.001
Groupe 2	0.0482	0.042		Réaction de l'entourage	0.0954	0.014
Groupe 3	-0.0204	0.044		Montreal	-0.0014	0.004
Groupe 4	-0.0743	0.045	*	Groupe 1	0.0080	0.035
Groupe 5	Ref.	-		Groupe 2	0.0297	0.033
Groupe 6	-0.0043	0.041		Groupe 3	0.0015	0.040
Groupe 7	-0.0323	0.040		Groupe 4	0.0479	0.036
Groupe 8	0.0106	0.045		Groupe 5	Réf.	-
Groupe 9	-0.0183	0.056		Groupe 6	0.0034	0.029
Groupe 10	-0.0415	0.047		Groupe 7	-0.0018	0.033
Groupe 11	-0.0848	0.049	*	Groupe 8	0.0135	0.030
Groupe 12	-0.0332	0.047		Groupe 9	0.0050	0.045
				Groupe 10	-0.0192	0.040
				Groupe 11	-0.0198	0.055
				Groupe 12	0.0044	0.041
REVENU BORS-TRAVAIL VIRTUEL /100			SALAIRE ESPERE AU NOIR /100			
Constante	0.2068	0.154		Constante	-2.2219	0.114
Age	0.1860	0.403		Age	4.2760	1.004
Age ²	0.0236	0.030		Age ²	-1.4821	0.307
Années de scolarité	0.2682	0.216	*	Age ³	1.2998	0.290
Années de scolarité ²	-1.4762	0.896	*	Années de scolarité	2.2824	0.335
Age*scolarité	-0.7183	0.437	*	Années de scolarité ²	-2.1556	0.488
Age*scolarité ²	-0.5666	0.189	***	Années de scolarité ³	0.6222	0.169
Femme	-0.0143	0.011		Femme	0.0110	0.017
Montréal	-0.0359	0.011	***	Montréal	-0.0503	0.019
Bas-du-Beuve	-0.0162	0.014		Bas-du-Beuve	-0.0819	0.028
SALAIRE OFFICIEL /100			SALAIRE ESPERE AU NOIR /10000			
Constante	0.1275	0.115		Constante	-3.1213	0.509
Age	1.1370	0.408	**	Age	4.7761	3.932
Age ²	-0.1269	0.045	**	Age ²	-1.7966	1.284
Age ³	0.0671	0.071		Age ³	1.6265	1.165
Années de scolarité	0.0048	0.156		Années de scolarité	3.7190	2.629
Age*scolarité	-0.0381	0.014	**	Années de scolarité ²	-3.3324	2.386
Femme	0.0075	0.014		Années de scolarité ³	1.1080	0.725
Montréal	0.0075	0.014		Femme	0.0027	0.047
Bas-du-Beuve	-0.0212	0.019	*	Montréal	-0.0901	0.048
				Bas-du-Beuve	-0.0965	0.084

Log-Vraisemblance	-8 127.8
Nb d'observations	3 083

*** Statistiquement significatif à 1%
 ** Statistiquement significatif à 5%
 * Statistiquement significatif à 10%

(référence = groupe 5)

où les douze groupes de référence sont les suivants :

Groupes	Régions	Education	Age
1	Montréal	Elevé	≤ 40 ans
2	Montréal	Elevé	>40 ans
3	Montréal	Faible	≤ 40 ans
4	Montréal	Faible	>40 ans
5	Québec	Elevé	≤ 40 ans
6	Québec	Elevé	>40 ans
7	Québec	Faible	≤ 40 ans
8	Québec	Faible	>40 ans
9	Bas-du-fleuve	Elevé	≤ 40 ans
10	Bas-du-fleuve	Elevé	>40 ans
11	Bas-du-fleuve	Faible	≤ 40 ans
12	Bas-du-fleuve	Faible	>40 ans

et la distribution des heures sur le marché soir est :

Groupes	Heures moyennes / an	Heures minimales	Heures maximales	Nbre de travailleurs au noir
1	403	16	3120	66
2	1384	40	4200	13
3	782	36	2400	21
4	1303	144	3600	12
5	384	8	2250	45
6	419	10	1920	15
7	410	12	1320	18
8	854	10	2070	10
9	381	12	1200	18
10	169	30	300	2
11	288	20	540	10
12	410	120	700	2

. résultats d'estimations avec quatre groupes de référence

PARAMETRES			COEFFICIENTS			ECART-TYPE		
α_3		0.25						
β_{11}		-0.5256		0.036		***		
β_{12}		-0.3430		0.018		***		
β_{13}		-0.0976		0.006		***		
β_{22}		-0.2880		0.015		***		
β_{23}		-0.0984		0.017		***		
β_{33}		-0.1549		0.003		***		

MARCHÉ OFFICIEL				MARCHÉ NOIR							
PARAMETRES		COEFFICIENTS		ECART-T		PARAMETRES		COEFFICIENTS		ECART-T	
HEURES OFFICIELLES /1000						HEURES NON-OFFICIELLES /1000					
α_1		0.5406	0.051	***	α_2		-2.4659	0.048	***		
Nb indiv. dans le ménage		0.0278	-0.009	***	Nb indiv. dans le ménage		0.0165	0.005	***		
Enfants (<6ans)		-0.0075	0.012		Enfants (<6ans)		-0.0038	0.008			
Célibataire		0.1391	0.032	***	Célibataire		0.1018	0.016	***		
Marié		0.2006	0.027	***	Marié		0.0998	0.014	***		
Famille mono-parentale		0.1225	-0.037	***	Famille mono-parentale		0.0697	0.020	***		
Femme		-0.1020	0.018	***	Femme		-0.0455	0.018	***		
Age		0.4769	0.131	***	Age		0.1662	0.116			
Groupe 1		-0.0071	0.032		Proportion de travailleurs au noir		0.0030	0.001	**		
Groupe 2		Réf.			Réaction de l'entourage		0.9951	0.013	***		
Groupe 3		-0.0498	0.033		Mortalité		-0.0013	0.004			
Groupe 4		0.0152	0.019		Groupe 1		-0.0056	0.024			
					Groupe 2		Réf.				
					Groupe 3		-0.0146	0.031			
					Groupe 4		0.0148	0.013			
REVENU HORS-TRAVAIL VIRTUEL /100						SALAIRE ESPERE AU NOIR /100					
Constante		0.2098	0.152		Constante		-2.2158	0.107	***		
Age		0.1741	0.401		Age		4.2340	0.935	***		
Age ²		0.0251	0.029		Age ²		-1.4702	0.282	***		
Années de scolarité		0.2661	0.212		Age ³		1.2902	0.264	***		
Années de scolarité ²		-1.4641	0.878	*	Age*scolarité		2.2978	0.307	***		
Age*scolarité		-0.7150	0.426	*	Age*scolarité ²		-2.1843	0.429	***		
Age*scolarité ²		0.5655	0.184	***	Age*scolarité ³		0.6333	0.146	***		
Femme		-0.0147	0.011		Femme		0.0104	0.017			
Montréal		0.0272	0.009	***	Montréal		-0.0552	0.007	***		
Bas-du-fleuve		-0.0050	0.011		Bas-du-fleuve		-0.0781	0.011	***		
SALAIRE OFFICIEL /100						SALAIRE ESPERE AU NOIR ² /10000					
Constante		0.1245	0.114		Constante		-3.1115	0.493	***		
Age		1.1517	0.403	***	Age		4.7159	3.831			
Age ²		-0.1286	0.045	***	Age ²		-1.7793	1.213			
Années de scolarité		0.0672	0.070		Age ³		1.6124	1.086			
Age*scolarité		0.0030	0.153		Age*scolarité		3.7380	2.397			
Femme		-0.0377	0.014	**	Age*scolarité ²		-3.8717	2.163	*		
Montréal		0.0144	0.013		Age*scolarité ³		1.1237	0.653	*		
Bas-du-fleuve		-0.0282	0.016	*	Femme		0.0016	0.046			
					Montréal		-0.0987	0.040	**		
					Bas-du-fleuve		-0.0901	0.070			
LOG-VRAISSEMBLANCE						-8 138.5					
NHD/OBSERVATIONS						3 083					

*** Statistiquement significatif à 1%
 ** Statistiquement significatif à 5%
 * Statistiquement significatif à 10%

(référence = groupe 2)

où les quatre groupes de référence sont :

Groupes	Education	Age
1	Elevé	≤ 40 ans
2	Elevé	> 40 ans
3	Faible	≤ 40 ans
4	Faible	> 40 ans

Annexe 3E - Pentas des fonctions d'offre de travail

Les pentes des fonctions d'offre de travail se définissent par la dérivée de h_i par rapport à ε_i . Ainsi, pour le régime 1 ($h_1 > 0, h_2 > 0$), les pentes sont déterminées de la manière suivante :

$$\begin{aligned} \frac{\partial h_i}{\partial \varepsilon_1} d\varepsilon_1 + \frac{\partial h_i}{\partial \varepsilon_2} d\varepsilon_2 = 0 &\Leftrightarrow \frac{\partial h_i}{\partial \varepsilon_1} + \frac{\partial h_i}{\partial \varepsilon_2} \frac{d\varepsilon_2}{d\varepsilon_1} = 0 \\ &\Leftrightarrow \frac{d\varepsilon_2}{d\varepsilon_1} = -\frac{\partial h_i / \partial \varepsilon_1}{\partial h_i / \partial \varepsilon_2} \end{aligned}$$

Les fonctions d'offres de travail s'écrivent respectivement :

$$\begin{aligned} h_1 = g_1(y, x, \varepsilon_1, \varepsilon_2) &= \frac{\gamma \xi_1 - \beta \xi_2 - \gamma \varepsilon_1 + \beta \varepsilon_2}{\alpha \gamma - \beta^2} \\ \text{et } h_2 = g_2(y, x, \varepsilon_1, \varepsilon_2) &= \frac{-\beta \xi_1 + \alpha \xi_2 + \beta \varepsilon_1 - \alpha \varepsilon_2}{\alpha \gamma - \beta^2} \end{aligned}$$

Par conséquent, la pente de h_1 est donnée par l'expression suivante :

$$\frac{d\varepsilon_2}{d\varepsilon_1} = \frac{\gamma / (\alpha \gamma - \beta^2)}{\beta / (\alpha \gamma - \beta^2)} = \frac{\gamma}{\beta}$$

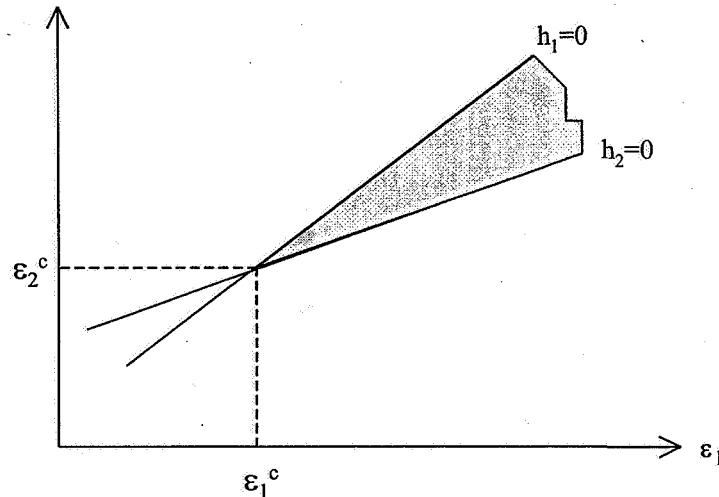
Si l'on décompose les dérivées partielles, $\frac{\partial h_1}{\partial \varepsilon_1} = \frac{-\gamma}{(\alpha \gamma - \beta^2)} > 0$ et $\frac{\partial h_1}{\partial \varepsilon_2} = \frac{\beta}{(\alpha \gamma - \beta^2)} < 0$. Ces dérivées partielles sont de signes opposés, ce qui implique que la pente de h_1 est positive. Celle-ci est comprise sur un intervalle [0.28; 2.33] selon les individus présents dans le premier régime.

Par ailleurs, la pente de h_2 est donnée par :

$$\frac{d\varepsilon_2}{d\varepsilon_1} = \frac{-\beta / (\alpha \gamma - \beta^2)}{-\alpha / (\alpha \gamma - \beta^2)} = \frac{\beta}{\alpha}$$

où, si l'on décompose les dérivées partielles, $\frac{\partial h_2}{\partial \varepsilon_1} = \frac{\beta}{(\alpha \gamma - \beta^2)} < 0$ et $\frac{\partial h_2}{\partial \varepsilon_2} = \frac{-\alpha}{(\alpha \gamma - \beta^2)} > 0$. Ces dérivées partielles sont également de signe opposé, ce qui signifie que la pente de h_2 est positive. Celle-ci s'inscrit dans un intervalle [0.15; 1.02] selon les individus de ce régime. Il convient de noter que la pente de h_2 est toujours plus faible que celle de h_1 . En outre, le signe des dérivées partielles pour h_2 est inversé par rapport à celui de h_1 .

Nous obtenons alors la représentation graphique ci-après des fonctions d'offre de travail sur le marché officiel et le marché non officiel.



La zone grisée correspond à la surface où h_1 et h_2 sont toutes deux positives. Elle nous est donnée par le signe des dérivées partielles de h_1 et h_2 . En effet, $\frac{\partial h_1}{\partial \varepsilon_1} > 0$ et $\frac{\partial h_1}{\partial \varepsilon_2} < 0$, de sorte que la surface pour laquelle h_1 est positif se situe en-dessous de la droite h_1 . À l'inverse, $\frac{\partial h_2}{\partial \varepsilon_1} < 0$ et $\frac{\partial h_2}{\partial \varepsilon_2} > 0$. Ainsi, la surface pour laquelle h_2 est positif se situe au-dessus de la droite h_2 . La zone grisée recoupe ces deux surfaces.

Pour le régime 2 ($h_1 > 0, h_2 = 0$), la pente de h_1 est définie de la manière suivante :

$$\frac{\partial h_1}{\partial \varepsilon_1} = -\frac{1}{\alpha}$$

où $h_1 = g_1(y, x, \varepsilon_1) = \frac{\xi_1 - \varepsilon_1}{\alpha}$. Dans la mesure où α est négatif, la pente de h_1 est positive. Elle est comprise sur l'intervalle [0.13; 1.90].

Enfin, pour le régime 3 ($h_1 = 0, h_2 > 0$), la pente de h_2 s'écrit :

$$\frac{\partial h_2}{\partial \varepsilon_2} = -\frac{1}{\gamma}$$

où $h_2 = g_2(y, x, \varepsilon_2) = \frac{\xi_2 - \varepsilon_2}{\gamma}$. Dans la mesure où γ est négatif, la pente de h_2 est également positive. Celle-ci se situe dans l'intervalle [0.78; 3.57].

Annexe 3F - Modèle exprimé sous forme logarithmique des salaires

Paramètres	Coefficients	Ecart-type
α_1	0.25	-
β_{11}	-0.7746	0.047
β_{12}	-0.5679	0.040
β_{13}	-0.2870	0.006
β_{22}	-0.6018	0.039
β_{23}	-0.5449	0.015
β_{33}	-0.0624	0.002

Marché officiel			Marché noir		
Paramètres	Coefficients	Ecart-type	Paramètres	Coefficients	Ecart-type
Heures officielles / 1000			Heures au noir / 1000		
α_1	1.0348	0.065	α_2	-2.2417	0.072
Nb de pers. dans le ménage	0.0426	0.012	Nb de pers. dans le ménage	0.0495	0.013
Enfants (< 6ans)	0.0018	0.013	Enfants (< 6ans)	0.0064	0.015
Personne seule	0.2026	0.040	Personne seule	0.2492	0.039
Conjoint	0.2770	0.033	Conjoint	0.3012	0.032
Monoparent	0.2302	0.051	Monoparent	0.2638	0.051
Femme	-0.1677	0.021	Femme	-0.2057	0.025
Age	0.6648	0.122	Age	0.8351	0.133
			Proporen	0.0015	0.002
			Réaction de l'entourage	1.0060	0.016
			Moralité	0.0036	0.004
Revenu hors-travail "virtuel" / 100			Salaire espéré au noir / 100		
Constante	0.5276	0.151	Constante	-2.6901	0.070
Age	-0.3231	0.402	Age	8.4360	0.560
Age (au carré)	0.1036	0.031	Age (au carré)	-2.3121	0.151
Années de scolarité	0.2141	0.214	Age (au cube)	1.7564	0.128
Années de scolarité (au carré)	-0.8737	0.895	Années de scolarité	3.7182	0.229
Age*scolarité	-0.6362	0.423	Années de scolarité (au carré)	-3.6796	0.300
Age*(scolarité au carré)	0.3898	0.187	Années de scolarité (au cube)	1.1266	0.105
Femme	-0.0087	0.010	Femme	-0.0298	0.011
Montréal	0.0131	0.010	Montréal	-0.0173	0.007
Bas-du-fleuve	0.0011	0.012	Bas-du-fleuve	-0.0842	0.009
Salaire net officiel / 100			Salaire espéré au noir (au carré) / 10 000		
Constante	0.1440	0.114	Constante	-3.0975	0.476
Age	1.2222	0.427	Age	10.9481	3.791
Age (au carré)	-0.1417	0.049	Age (au carré)	-3.0514	1.124
Années de scolarité	0.0617	0.073	Age (au cube)	2.3320	1.014
Années de scolarité (au carré)	-0.0166	0.160	Années de scolarité	5.8383	1.689
Femme	-0.0412	0.015	Années de scolarité (au carré)	-6.0082	1.824
Montréal	0.0094	0.014	Années de scolarité (au cube)	1.8246	0.607
Bas-du-fleuve	-0.0259	0.019	Femme	-0.0572	0.042
			Montréal	-0.0416	0.042
			Bas-du-fleuve	-0.0979	0.074
Log-Vraisemblance			- 8 914.2		
Nb d'observations			3 083		

Tableau 3F-1 : Estimation de la fonction d'utilité quadratique

Annexe 3F - Modèle exprimé sous forme logarithmique des salaires 305

Matrice de corrélation des termes d'erreur :

	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6
ε_1	0.43539					
ε_2	0.11271	1.49619				
ε_3	0.18284	-0.44881	0.28915			
ε_4	0.04080	-0.44696	-0.16908	0.19965		
ε_5	-0.23833	1.07728	-0.49487	-0.38212	1.02328	
ε_6	-0.36030	1.65764	-0.76296	-0.58685	1.57035	2.42541

où ε_1 correspond à l'équation des heures officielles (h_1), ε_4 correspond à l'équation du salaire net officiel (w_1),
 ε_2 correspond à l'équation des heures au noir (h_2), ε_5 correspond à l'équation du salaire espéré au noir (Ew_2),
 ε_3 correspond à l'équation du revenu virtuel (y^v), ε_6 correspond à l'équation de l'espérance de salaire au noir au carré $E(w_2^2)$.

Tableau 3F-2 : Matrice de corrélation des termes d'erreur

Annexe 4A - Les fonctions d'offre de travail

Dans le régime 1, les fonctions d'offre de travail sont obtenues en posant :

$$\begin{aligned} m_1(h_1, h_2, z) = & \bar{\alpha}_1 + X'_1 \delta_1 + \alpha_3 w_1 + \beta_{11} h_1 + \beta_{12} h_2 \\ & + \beta_{13} (y^v + 2w_1 h_1 + w_2 h_2) + \beta_{23} w_1 h_2 \\ & + \beta_{33} w_1 (y^v + w_1 h_1 + w_2 h_2) + \varepsilon_1 = 0 \end{aligned} \quad (4A.1)$$

$$\begin{aligned} m_2(h_1, h_2, z) = & \bar{\alpha}_2 + X'_2 \delta_2 + \alpha_3 E w_2 + \beta_{12} h_1 + \beta_{13} h_1 w_2 \\ & + \beta_{22} h_2 + \beta_{23} (y^v + w_1 h_1 + 2w_2 h_2) \\ & + \beta_{33} w_2 (y^v + w_1 h_1 + w_2 h_2) + \varepsilon_2 = 0 \end{aligned} \quad (4A.2)$$

et en résolvant pour h_1 et h_2 . Les fonctions d'offre de travail associées à ce régime sont les suivantes :

$$h_1 = \frac{\gamma \zeta_1 - \gamma \varepsilon_1 - \beta \zeta_2 + \beta \varepsilon_2}{(\alpha \gamma - \beta^2)} \quad (4A.3)$$

$$h_2 = \frac{-\beta \zeta_1 + \beta \varepsilon_1 + \alpha \zeta_2 - \alpha \varepsilon_2}{(\alpha \gamma - \beta^2)} \quad (4A.4)$$

où

$$\begin{aligned} \alpha &= \beta_{11} + 2\beta_{13} w_1 + \beta_{33} w_1^2 \\ \beta &= \beta_{12} + \beta_{13} w_2 + \beta_{23} w_1 + \beta_{33} w_1 w_2 \\ \gamma &= \beta_{22} + 2\beta_{23} w_2 + \beta_{33} w_2^2 \\ \zeta_1 &= -\bar{\alpha}_1 - X'_1 \delta_1 - \alpha_3 w_1 - \beta_{13} y^v - \beta_{33} y^v w_1 \\ \zeta_2 &= -\bar{\alpha}_2 - X'_2 \delta_2 - \alpha_3 w_2 - \beta_{23} y^v - \beta_{33} y^v w_2 \end{aligned}$$

De même, en posant $h_2 = 0$ dans (4A.1), nous obtenons :

$$m_1(h_1, 0, w_1, w_2, z) = \bar{\alpha}_1 + X'_1 \delta_1 + \alpha_3 w_1 + \beta_{11} h_1 + \beta_{13} (y^v + 2w_1 h_1) + \beta_{33} w_1 (y^v + w_1 h_1) + \varepsilon_1 = 0$$

et

$$m_2(h_1, 0, w_1, w_2, z) = \bar{\alpha}_2 + X'_2 \delta_2 + \alpha_3 w_2 + \beta_{12} h_1 + \beta_{13} h_1 w_2 + \beta_{23} (y^v + w_1 h_1) + \beta_{33} w_2 (y^v + w_1 h_1) + \varepsilon_2 \leq 0$$

En résolvant pour h_1 , nous obtenons les fonctions d'offre contraintes associées au régime 2 :

$$\bar{h}_1 = \frac{\zeta_1 - \varepsilon_1}{\alpha} \quad (4A.3)$$

Enfin, en posant $h_1 = 0$ dans (4A.2), nous obtenons :

$$m_2(0, h_2, w_1, w_2, z) = \bar{\alpha}_2 + X'_2 \delta_2 + \alpha_3 w_2 + \beta_{22} h_2 + \beta_{23} (y^v + 2w_2 h_2) \\ + \beta_{33} w_2 (y^v + w_2 h_2) + \varepsilon_2 = 0$$

et

$$m_1(0, h_2, w_1, w_2, z) = \bar{\alpha}_1 + X'_1 \delta_1 + \alpha_3 w_1 + \beta_{12} h_2 + \beta_{13} (y^v + w_2 h_2) + \beta_{23} w_1 h_2 \\ + \beta_{33} w_1 (y^v + w_2 h_2) + \varepsilon_1 \leq 0$$

En résolvant pour h_2 , nous obtenons la fonction d'offre contrainte associée au régime 3 :

$$\bar{h}_2 = \frac{\zeta_2 - \varepsilon_2}{\gamma} \quad (4A.4)$$

Annexe 4B - Les fonctions d'offre de travail en l'hypothèse de parfaite substitution

Dans le régime 1, les fonctions d'offre de travail sont obtenues en posant :

$$\begin{aligned} m_1(h_1, h_2, w_1, w_2, z) = & \bar{\gamma}_1 + \gamma_2 w_1 + \delta_{11} h_1 + \delta_{11} h_2 \\ & + \delta_{12}(y^v + w_1 h_1 + w_2 h_2) + \delta_{12} w_1 h_2 \\ & + \delta_{22} w_1 (y^v + w_1 h_1 + w_2 h_2) + \mu_1 = 0 \end{aligned} \quad (4B.1)$$

$$\begin{aligned} m_2(h_1, h_2, w_1, w_2, z) = & \bar{\gamma}_1 + \gamma_2 w_2 + \delta_{11} h_2 + \delta_{11} h_1 \\ & + \delta_{12} h_1 w_2 + \delta_{12}(y^v + w_1 h_1 + w_2 h_2) \\ & + \delta_{22} w_2 (y^v + w_1 h_1 + w_2 h_2) + \mu_2 = 0 \end{aligned} \quad (4B.2)$$

et en résolvant pour h_1 et h_2 . Les fonctions d'offre de travail associées à ce régime sont les suivantes :

$$h_1 = \frac{\gamma \zeta_1 - \gamma \mu_1 - \beta \zeta_2 + \beta \mu_2}{(\alpha \gamma - \beta^2)} \quad (4B.3)$$

$$h_2 = \frac{-\beta \zeta_1 + \beta \mu_1 + \alpha \zeta_2 - \alpha \mu_2}{(\alpha \gamma - \beta^2)} \quad (4B.4)$$

où

$$\begin{aligned} \alpha &= \delta_{11} + \delta_{12} w_1 + \delta_{22} w_1^2 \\ \beta &= \delta_{11} + \delta_{12} w_2 + \delta_{12} w_1 + \delta_{22} w_1 w_2 \\ \gamma &= \delta_{11} + \delta_{12} w_2 + \delta_{22} w_2^2 \\ \zeta_1 &= -\bar{\gamma}_1 - \gamma_2 w_1 - \delta_{12} y^v - \delta_{12} w_1 y^v \\ \zeta_2 &= -\bar{\gamma}_1 - \gamma_2 w_2 - \delta_{12} y^v - \delta_{12} y^v w_2 \end{aligned}$$

De même, en posant $h_2 = 0$ dans (4B.1), nous obtenons :

$$m_1(h_1, 0, w_1, w_2, z) = \bar{\gamma}_1 + \gamma_2 w_1 + \delta_{11} h_1 + \delta_{12}(y^v + w_1 h_1) + \delta_{22} w_1 (y^v + w_1 h_1) + \mu_1 = 0$$

et

$$m_2(h_1, 0, w_1, w_2, z) = \bar{\gamma}_1 + \gamma_2 w_2 + \delta_{11} h_1 + \delta_{12} h_1 w_2 + \delta_{12}(y^v + w_1 h_1) + \delta_{22} w_2 (y^v + w_1 h_1) + \mu_2 \leq 0$$

En résolvant pour h_1 , nous obtenons les fonctions d'offre contraintes associées au régime 2 :

$$\bar{h}_1 = \frac{\zeta_1 - \mu_1}{\alpha} \quad (4B.5)$$

Enfin, en posant $h_1 = 0$ dans (4B.2), nous obtenons :

$$m_2(0, h_2, w_1, w_2, z) = \bar{\gamma}_1 + \gamma_2 w_2 + \delta_{11} h_2 + \delta_{12} h_1 w_2 + \delta_{12} (y^v + a_2 w_2) + \delta_{22} w_2 (y^v + w_2 h_2) + \mu_2 = 0$$

et

$$m_1(0, h_2, w_1, w_2, z) = \bar{\gamma}_1 + \gamma_2 w_1 + \delta_{11} h_2 + \delta_{12} (y^v + w_2 h_2) + \delta_{12} w_1 h_2 + \delta_{22} w_1 (y^v + w_2 h_2) + \mu_1 \leq 0$$

En résolvant pour h_2 , nous obtenons la fonction d'offre contrainte associée au régime 3 :

$$\bar{h}_2 = \frac{\zeta_2 - \mu_2}{\gamma} \quad (4B.6)$$

Annexe 4C - Les équations de salaire

Dans le régime 2, l'individu travaille uniquement sur le marché officiel ($h_1 > 0$ et $h_2 = 0$). Nous observons alors uniquement le salaire obtenu sur ce marché :

$$\begin{aligned} w_1 &= X'_4 \Psi_4 + \varepsilon_4 \\ \text{et } \varepsilon_5 &= w_2 - X'_5 \Psi_5 \end{aligned} \quad (4C.1)$$

A l'inverse, dans le régime 3, l'individu travaille uniquement sur le marché noir ($h_1 = 0$ et $h_2 > 0$). Par conséquent, nous observons uniquement le salaire de l'activité souterraine :

$$\begin{aligned} \varepsilon_4 &= w_1 - X'_4 \Psi_4 \\ \text{et } w_2 &= X'_5 \Psi_5 + \varepsilon_5 \end{aligned} \quad (4C.2)$$

Enfin, dans le régime 4, l'individu ne travaille sur aucun marché ($h_1 = 0$ et $h_2 = 0$), de sorte que nous obtenons :

$$\begin{aligned} \varepsilon_4 &= w_1 - X'_4 \Psi_4 \\ \text{et } \varepsilon_5 &= w_2 - X'_5 \Psi_5 \end{aligned} \quad (4C.3)$$

Annexe 4D - Modèle exprimé sous forme logarithmique des salaires

Paramètres	Coefficients	Ecart-type
α_1	0.25	-
β_{11}	-1.9529	0.021
β_{12}	-0.8488	0.009
β_{13}	-0.5734	0.002
β_{22}	-0.3868	0.009
β_{23}	-0.7932	0.007
β_{33}	-0.0835	0.001

Marché officiel			Marché noir		
Paramètres	Coefficients	Ecart-type	Paramètres	Coefficients	Ecart-type
Heures officielles / 1000			Heures au noir / 1000		
α_1	2.7210	0.052	α_2	-3.0209	0.044
Nb de pers. dans le ménage	0.0849	0.017	Nb de pers. dans le ménage	0.0598	0.008
Enfants (< 6ans)	-0.0856	0.020	Enfants (< 6ans)	-0.0619	0.011
Personne seule	0.1924	0.050	Personne seule	0.1321	0.024
Conjoint	0.1472	0.036	Conjoint	0.0978	0.018
Monoparent	0.2527	0.049	Monoparent	0.1721	0.025
Femme	-0.2855	0.030	Femme	-0.2146	0.023
Age	0.6421	0.136	Age	2.0033	0.100
			Proporen	-0.0005	0.001
			Réaction de l'entourage	0.0023	0.004
			Moralité	0.0021	0.001
Salaires net officiel / 100			Salaires espéré au noir / 100		
Constante	-0.1499	0.114	Constante	-3.8665	0.041
Age	1.6914	0.423	Age	3.1823	0.331
Age (au carré)	-0.2131	0.052	Age (au carré)	-0.7126	0.096
Années de scolarité	-0.0266	0.078	Age (au cube)	0.5761	0.090
Age*scolarité	-0.0059	0.178	Années de scolarité	1.8041	0.120
Femme	-0.0390	0.014	Années de scolarité (au carré)	-3.2955	0.280
Montréal	0.0191	0.015	Années de scolarité (au cube)	1.6946	0.206
Bas-du-fleuve	-0.0031	0.019	Femme	-0.0581	0.010
			Montréal	-0.0577	0.004
			Bas-du-fleuve	-0.0586	0.006
Revenu hors-travail "virtuel" / 100					
Constante	0.5324	0.162			
Age	-0.7748	0.561			
Age (au carré)	0.1161	0.051			
Années de scolarité	-0.1532	0.161			
Années de scolarité (au carré)	0.0592	0.039			
Age*scolarité	0.2244	0.334			
Age*(scolarité au carré)	-0.1185	0.167			
Femme	-0.0051	0.013			
Montréal	0.0043	0.011			
Bas-du-fleuve	0.0082	0.015			
Log-Vraisemblance			-6 889.1		
Nb d'observations			2 934		

Tableau 4D-1 : Estimation de la fonction d'utilité quadratique

Annexe 4D - Modèle exprimé sous forme logarithmique des salaires 312

	ε_1	ε_2	ε_3	ε_4	ε_5
ε_1	2.69346	0.12378	0.34170	-0.14917	-0.99174
ε_2	0.12378	1.99448	-0.22919	0.01252	1.56989
ε_3	0.34170	-0.22919	0.15427	0.02112	-0.34943
ε_4	-0.14917	0.01252	0.02112	0.10349	0.05202
ε_5	-0.99174	1.56989	-0.34943	0.05202	1.68615

où ε_1 correspond à l'équation des heures officielles (h_1),
 ε_2 correspond à l'équation des heures au noir (h_2),
 ε_3 correspond à l'équation du revenu virtuel (y^v),
 ε_4 correspond à l'équation du salaire net officiel (w_1),
 ε_5 correspond à l'équation du salaire espéré au noir (EW_2).

Tableau 4D-2 : Matrice de corrélation des termes d'erreur

Annexe 4E - Pentas des fonctions d'offre de travail

Les pentes des fonctions d'offre de travail se définissent par la dérivée de h_i par rapport à ε_i . Ainsi, pour le régime 1 ($h_1 > 0, h_2 > 0$), les pentes sont déterminées de la manière suivante :

$$\begin{aligned} \frac{\partial h_i}{\partial \varepsilon_1} d\varepsilon_1 + \frac{\partial h_i}{\partial \varepsilon_2} d\varepsilon_2 = 0 &\Leftrightarrow \frac{\partial h_i}{\partial \varepsilon_1} + \frac{\partial h_i}{\partial \varepsilon_2} \frac{d\varepsilon_2}{d\varepsilon_1} = 0 \\ &\Leftrightarrow \frac{d\varepsilon_2}{d\varepsilon_1} = -\frac{\partial h_i / \partial \varepsilon_1}{\partial h_i / \partial \varepsilon_2} \end{aligned}$$

Les fonctions d'offres de travail s'écrivent respectivement :

$$\begin{aligned} h_1 = g_1(y, x, \varepsilon_1, \varepsilon_2) &= \frac{\gamma\xi_1 - \beta\xi_2 - \gamma\varepsilon_1 + \beta\varepsilon_2}{\alpha\gamma - \beta^2} \\ \text{et } h_2 = g_2(y, x, \varepsilon_1, \varepsilon_2) &= \frac{-\beta\xi_1 + \alpha\xi_2 + \beta\varepsilon_1 - \alpha\varepsilon_2}{\alpha\gamma - \beta^2}. \end{aligned}$$

Par conséquent, la pente de h_1 est donnée par l'expression suivante :

$$\frac{d\varepsilon_2}{d\varepsilon_1} = \frac{\gamma / (\alpha\gamma - \beta^2)}{\beta / (\alpha\gamma - \beta^2)} = \frac{\gamma}{\beta}$$

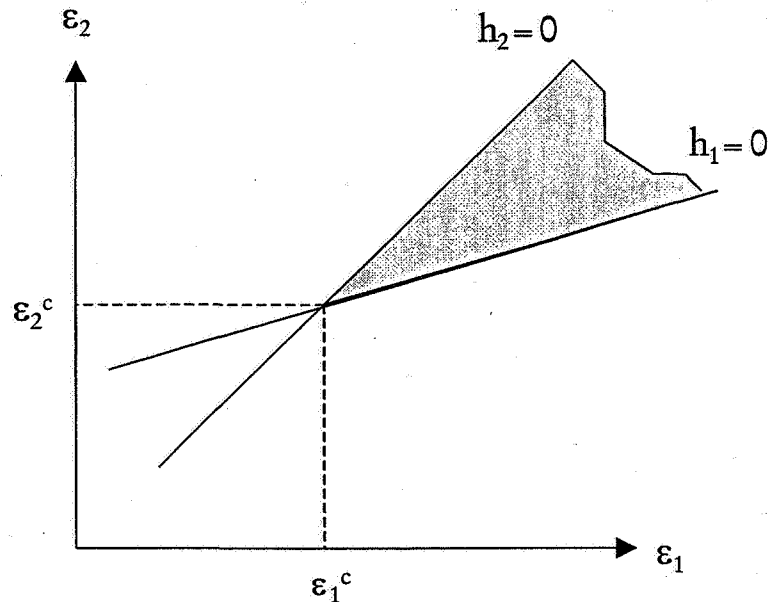
Si l'on décompose les dérivées partielles, $\frac{\partial h_1}{\partial \varepsilon_1} = \frac{-\gamma}{(\alpha\gamma - \beta^2)} > 0$ et $\frac{\partial h_1}{\partial \varepsilon_2} = \frac{\beta}{(\alpha\gamma - \beta^2)} < 0$. Ces dérivées partielles sont de signes opposés, ce qui implique que la pente de h_1 est positive. Celle-ci est comprise sur un intervalle [0.01; 0.45] selon les individus présents dans le premier régime.

Par ailleurs, la pente de h_2 est donnée par :

$$\frac{d\varepsilon_2}{d\varepsilon_1} = \frac{-\beta / (\alpha\gamma - \beta^2)}{-\alpha / (\alpha\gamma - \beta^2)} = \frac{\beta}{\alpha}$$

où, si l'on décompose les dérivées partielles, $\frac{\partial h_2}{\partial \varepsilon_1} = \frac{\beta}{(\alpha\gamma - \beta^2)} < 0$ et $\frac{\partial h_2}{\partial \varepsilon_2} = \frac{-\alpha}{(\alpha\gamma - \beta^2)} > 0$. Ces dérivées partielles sont également de signe opposé, ce qui signifie que la pente de h_2 est positive. Celle-ci s'inscrit dans un intervalle [0.47; 2.73] selon les individus de ce régime. Il convient de noter que la pente de h_2 est toujours plus élevée que celle de h_1 . En outre, le signe des dérivées partielles pour h_2 est inversé par rapport à celui de h_1 .

Nous obtenons alors la représentation graphique ci-après des fonctions d'offre de travail sur le marché officiel et le marché non officiel.



La zone grisée correspond à la surface où h_1 et h_2 sont toutes deux positives. Elle nous est donnée par le signe des dérivées partielles de h_1 et h_2 . En effet, $\frac{\partial h_1}{\partial \varepsilon_1} < 0$ et $\frac{\partial h_1}{\partial \varepsilon_2} > 0$, de sorte que la surface pour laquelle h_1 est positif se situe au-dessus de la droite h_1 . A l'inverse, $\frac{\partial h_2}{\partial \varepsilon_1} > 0$ et $\frac{\partial h_2}{\partial \varepsilon_2} < 0$. Ainsi, la surface pour laquelle h_2 est positif se situe en-dessous de la droite h_2 . La zone grisée recoupe ces deux surfaces.

Pour le régime 2 ($h_1 > 0, h_2 = 0$), la pente de h_1 est définie de la manière suivante :

$$\frac{\partial h_1}{\partial \varepsilon_1} = -\frac{1}{\alpha}$$

où $h_1 = g_1(y, x, \varepsilon_1) = \frac{\xi_1 - \varepsilon_1}{\alpha}$. Dans la mesure où α est négatif, la pente de h_1 est positive. Elle est comprise sur l'intervalle [1.11; 18.47].

Enfin, pour le régime 3 ($h_1 = 0, h_2 > 0$), la pente de h_2 s'écrit :

$$\frac{\partial h_2}{\partial \varepsilon_2} = -\frac{1}{\gamma}$$

où $h_2 = g_2(y, x, \varepsilon_2) = \frac{\xi_2 - \varepsilon_2}{\gamma}$. Dans la mesure où γ est négatif, la pente de h_2 est également positive. Celle-ci se situe dans l'intervalle [31.08; 74.94].

Annexe 5A - Interdépendance des comportements

Nous avons observé que le choix de l'activité souterraine était fortement influencé par l'entourage de l'individu. Lorsque ce dernier estime que le nombre de travailleurs au noir est important, il semble incité à accroître son offre de travail au noir. En conséquence, il importe de tenir compte de l'interdépendance des comportements dans la décision de travailler au noir. Mais, les variables subjectives utilisées pour cela sont susceptibles de générer un biais de dissonance cognitive. En effet, un individu peut être amené à sur-estimer la proportion de travailleurs au noir dans son entourage afin de justifier son emploi non-déclaré. Ces variables subjectives posent, en outre, un problème connu sous le nom de «problème du miroir» (ou *reflection problem*). La difficulté, dans ce cas, est de déterminer si l'individu choisit de travailler au noir parce que les membres de son entourage exercent une activité souterraine ou si cet individu possède, en fait, des caractéristiques exogènes identiques à celles de ses proches.

De ce fait, nous devons poursuivre, de façon plus rigoureuse, l'analyse des comportements interdépendants. Deux types de modélisations sont proposées. La première répond au problème d'endogénéité des effets de voisinage ; la seconde permet, en outre, de contrôler pour le biais potentiel de dissonance cognitive. La fonction de vraisemblance, propre à chacun de ces modèles, est également présentée.

1 Endogénéité des effets de voisinage

A l'instar du *chapitre 3*, 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 :

$$\alpha_1 = \bar{\alpha}_1 + X_1' \delta_1 + \varepsilon_1, \quad (5A.1)$$

$$\alpha_2 = \bar{\alpha}_2 + X_2' \delta_2 + \varepsilon_2. \quad (5A.2)$$

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. Ici cependant, X_2' contient la variable subjective de proportion de travailleurs au noir dans l'entourage (h_2), qui est estimée simultanément et dont la spécification est donnée ci-après (équation 5A.7).

Les équations salariales sont définies de la même manière, soit :

$$y^v = X'_3 \Psi_3 + \varepsilon_3 \quad (5A.3)$$

$$w_1 = X'_4 \Psi_4 + \varepsilon_4 \quad (5A.4)$$

$$Ew_2 = X'_5 \Psi_5 + \varepsilon_5 \quad (5A.5)$$

$$E(w_2^2) = X'_6 \Psi_6 + \varepsilon_6 \quad (5A.6)$$

Afin de contourner le problème d'endogénéité associé à la variable subjective d'évaluation du pourcentage de travailleurs au noir dans l'entourage, nous spécifions la forme linéaire suivante :

$$\widetilde{h}_2 = X'_7 \Psi_7 + \varepsilon_7 \quad (5A.7)$$

On suppose finalement que $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6, \varepsilon_7)$ est distribué $N(0, \Sigma)$. Nous choisissons de procéder aux estimations par la méthode de maximisation de la fonction de vraisemblance à information complète.

Fonction de vraisemblance

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

$$\begin{aligned} L = & \prod_{i \in G_1} f(-\bar{\alpha}_1 - X'_1 \delta_1 - \alpha_3 w_1 - \beta_{11} h_1 - \beta_{12} h_2 - \beta_{13} (y^v + 2w_1 h_1 + h_2 Ew_2) \\ & - \beta_{23} w_1 h_2 - \beta_{33} w_1 (y^v + w_1 h_1 + h_2 Ew_2), \\ & - \bar{\alpha}_2 - X'_2 \delta_2 - \widetilde{h}_2 \delta_2 - \alpha_3 Ew_2 - \beta_{12} h_1 - \beta_{13} h_1 Ew_2 - \beta_{22} h_2 \\ & - \beta_{23} (y^v + w_1 h_1 + 2h_2 Ew_2) - \beta_{33} [(y^v + w_1 h_1) Ew_2 + h_2 Ew_2^2], \\ & y^v - X'_3 \Psi_3, w_1 - X'_4 \Psi_4, Ew_2 - X'_5 \Psi_5, Ew_2^2 - X'_6 \Psi_6, \widetilde{h}_2 - X'_7 \Psi_7) |J_1| \\ & \times \\ & \prod_{i \in G_2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_2^*} f(-\bar{\alpha}_1 - X'_1 \delta_1 - \alpha_3 w_1 - \beta_{11} h_1 - \beta_{13} (y^v + 2w_1 h_1) \\ & - \beta_{33} w_1 (y^v + w_1 h_1), \varepsilon_2, y^v - X'_3 \Psi_3, w_1 - X'_4 \Psi_4, \\ & \varepsilon_5, \varepsilon_6, \widetilde{h}_2 - X'_7 \Psi_7) d\varepsilon_2 d\varepsilon_5 d\varepsilon_6 |J_2| \\ & \times \\ & \prod_{i \in G_3} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_1^*} f(\varepsilon_1, -\bar{\alpha}_2 - X'_2 \delta_2 - \widetilde{h}_2 \delta_2 - \alpha_3 Ew_2 - \beta_{22} h_2 \\ & - \beta_{23} (y^v + 2h_2 Ew_2) - \beta_{33} (y^v Ew_2 + h_2 Ew_2^2), y^v - X'_3 \Psi_3, \varepsilon_4, \\ & Ew_2 - X'_5 \Psi_5, Ew_2^2 - X'_6 \Psi_6, \widetilde{h}_2 - X'_7 \Psi_7) d\varepsilon_1 d\varepsilon_4 |J_3| \\ & \times \\ & \prod_{i \in G_4} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{\varepsilon_1^{**}} \int_{-\infty}^{\varepsilon_2^{**}} f(\varepsilon_1, \varepsilon_2, y^v - X'_3 \Psi_3, \varepsilon_4, \varepsilon_5, \varepsilon_6, \widetilde{h}_2) \\ & d\varepsilon_1 d\varepsilon_2 d\varepsilon_4 d\varepsilon_5 d\varepsilon_6 \end{aligned} \quad (5A.8)$$

où $G_1 = \{h_1 > 0, h_2 > 0\}$, $G_2 = \{h_1 > 0, h_2 = 0\}$,
 $G_3 = \{h_1 = 0, h_2 > 0\}$, $G_4 = \{h_1 = 0, h_2 = 0\}$.

2 Effets de voisinage endogènes et dissonance cognitive

2.1 Modèle théorique

Afin de tenir compte du biais potentiel de dissonance cognitive, la variable \widetilde{h}_2 doit être un paramètre d'optimisation pour l'individu. Ainsi, le programme à résoudre est :

$$\max_{\{h_1, h_2, \widetilde{h}_2\}} EU(h_1, h_2, C, \widetilde{h}_2), \quad (5A.9)$$

sous les contraintes de budget et de non-négativité sur h_1 , h_2 et \widetilde{h}_2 . Une telle formulation permet à la variable h_2 d'entrer de façon non linéaire dans la fonction d'utilité. Les utilités marginales nettes des heures de travail et de la proportion de travailleurs au noir dans l'entourage sont les suivantes :

$$m_1(h_1, h_2, \widetilde{h}_2, z) = EU_1 + w_1 EU_3, \quad (5A.10)$$

$$m_2(h_1, h_2, \widetilde{h}_2, z) = EU_2 + E(U_3 W_2), \quad (5A.11)$$

$$m_3(h_1, h_2, \widetilde{h}_2, z) = EU_4, \quad (5A.12)$$

où z est un vecteur de variables exogènes et les termes EU_1 , EU_2 et EU_4 représentent respectivement l'effet d'une hausse de h_1 , de h_2 et de \widetilde{h}_2 sur l'utilité espérée.

Les équations (5A.10) à (5A.12) permettent de caractériser les conditions d'optimalité du travail propres à chacun des quatre régimes sur les marchés officiel et souterrain. A l'optimum, les conditions de Kuhn-Tucker de premier ordre associées au programme (5A.9) s'écrivent :

$$m_1(h_1, h_2, \widetilde{h}_2, z) \leq 0, \quad (5A.13)$$

$$m_2(h_1, h_2, \widetilde{h}_2, z) \leq 0, \quad (5A.14)$$

$$m_3(h_1, h_2, \widetilde{h}_2, z) \leq 0, \quad (5A.15)$$

$$h_1[m_1(h_1, h_2, \widetilde{h}_2, z)] = 0, \quad (5A.16)$$

$$h_2[m_2(h_1, h_2, \widetilde{h}_2, z)] = 0, \quad (5A.17)$$

$$\widetilde{h}_2[m_3(h_1, h_2, \widetilde{h}_2, z)] = 0. \quad (5A.18)$$

Ces conditions nous permettent de redéfinir les quatre régimes de participation au marché du travail, selon que les contraintes de non-négativité sont serrées ou non.

2.2 Spécification économétrique

Afin d'estimer notre modèle, nous retenons la forme fonctionnelle suivante :

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

où $x = (h_1, h_2, C, \widetilde{h}_2)'$, α est un vecteur de paramètres de dimension 4×1 et β est une matrice de paramètres 4×4 . Dans la mesure où β est définie négative et symétrique, U est strictement concave.

En remplaçant la consommation C par son expression dans chaque état de la nature, nous obtenons :

$$\begin{aligned}
EU(x) = & \alpha_1 h_1 + \alpha_2 h_2 + \alpha_3 (y^v + w_1 h_1 + h_2 Ew_2) + \alpha_4 \widetilde{h}_2 + \frac{1}{2} \beta_{11} h_1^2 \\
& + \beta_{12} h_1 h_2 + \beta_{13} h_1 (y^v + w_1 h_1 + h_2 Ew_2) + \beta_{14} h_1 \widetilde{h}_2 \\
& + \frac{1}{2} \beta_{22} h_2^2 + \beta_{23} h_2 (y^v + w_1 h_1 + h_2 Ew_2) + \beta_{24} h_2 \widetilde{h}_2 \\
& + \frac{1}{2} \beta_{33} E(y^v + w_1 h_1 + w_2 h_2)^2 \\
& + \beta_{34} \widetilde{h}_2 (y^v + w_1 h_1 + h_2 Ew_2) + \frac{1}{2} \beta_{44} (\widetilde{h}_2)^2.
\end{aligned} \tag{5A.19}$$

Les utilités marginales nettes (5A.10) à (5A.12) sont définies par :

$$\begin{aligned}
m_1(h_1, h_2, \widetilde{h}_2, z) = & \alpha_1 + \alpha_3 w_1 + \beta_{11} h_1 + \beta_{12} h_2 \\
& + \beta_{13} (y^v + 2w_1 h_1 + h_2 Ew_2) + \beta_{14} \widetilde{h}_2 \\
& + \beta_{23} w_1 h_2 + \beta_{33} w_1 (y^v + w_1 h_1 + h_2 Ew_2) + \beta_{34} w_1 \widetilde{h}_2,
\end{aligned} \tag{5A.20}$$

$$\begin{aligned}
m_2(h_1, h_2, \widetilde{h}_2, z) = & \alpha_2 + \alpha_3 Ew_2 + \beta_{12} h_1 + \beta_{13} h_1 Ew_2 \\
& + \beta_{22} h_2 + \beta_{23} (y^v + w_1 h_1 + 2h_2 Ew_2) + \beta_{24} \widetilde{h}_2 \\
& + \beta_{33} [(y^v + w_1 h_1) Ew_2 + h_2 Ew_2^2] + \beta_{34} h_2 Ew_2,
\end{aligned} \tag{5A.21}$$

$$\begin{aligned}
m_3(h_1, h_2, \widetilde{h}_2, z) = & \alpha_4 + \beta_{14} h_1 + \beta_{24} h_2 + \beta_{34} (y^v + w_1 h_1 + h_2 Ew_2) \\
& + \beta_{44} \widetilde{h}_2.
\end{aligned} \tag{5A.22}$$

Ces conditions sont globalement satisfaites si la matrice β est définie négative (concavité stricte de $U(x)$).

Structure stochastique

Comme précédemment, les attributs socio-démographiques et les différences non observables entre individus sont introduits dans le modèle de la manière suivante :

$$\alpha_1 = \bar{\alpha}_1 + X_1' \delta_1 + \varepsilon_1, \tag{5A.23}$$

$$\alpha_2 = \bar{\alpha}_2 + X_2' \delta_2 + \varepsilon_2. \tag{5A.24}$$

où $\bar{\alpha}_i$ est une constante, X_i un vecteur de caractéristiques socio-démographiques, δ_i est un vecteur de paramètres à estimer et ε_i caractérise un terme d'erreur, avec $i = 1, 2$.

Mais, pour tenir compte de l'endogénéité de la variable subjective d'évaluation du nombre de travailleurs au noir dans l'entourage et pour contourner le problème de dissonance cognitive, nous retenons la spécification suivante :

$$\alpha_4 = \bar{\alpha}_4 + X_4' \delta_4 + \varepsilon_4. \tag{5A.25}$$

Les équations salariales linéaires sont :

$$y^v = X_3' \Psi_3 + \varepsilon_3, \tag{5A.26}$$

$$w_1 = X_5' \Psi_5 + \varepsilon_5, \tag{5A.27}$$

$$Ew_2 = X_6' \Psi_6 + \varepsilon_6, \tag{5A.28}$$

$$E(w_2^2) = X_7' \Psi_7 + \varepsilon_7. \tag{5A.29}$$

Enfin, nous supposons que $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6, \varepsilon_7)$ est distribué $N(0, \Sigma)$.

Fonction de vraisemblance

La fonction de vraisemblance associée aux quatres 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_{14} \widetilde{h}_2 - \beta_{23} w_1 h_2 - \beta_{33} w_1 (y^v + w_1 h_1 + h_2 Ew_2) - \beta_{34} \widetilde{h}_2 w_1, \\
& - \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_{24} \widetilde{h}_2 \\
& - \beta_{33} [(y^v + w_1 h_1) Ew_2 + h_2 Ew_2^2] - \beta_{34} \widetilde{h}_2 Ew_2, y^v - X'_3 \Psi_3, \\
& - \bar{\alpha}_4 - X'_4 \delta_4 - \beta_{14} h_1 - \beta_{24} h_2 - \beta_{34} (y^v + w_1 h_1 + h_2 Ew_2) - \beta_{44} \widetilde{h}_2 \\
& w_1 - X'_5 \Psi_5, Ew_2 - X'_6 \Psi_6, Ew_2^2 - X'_7 \Psi_7) |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_{14} \widetilde{h}_2 - \beta_{33} w_1 (y^v + w_1 h_1) - \beta_{34} \widetilde{h}_2 w_1, \varepsilon_2, y^v - X'_3 \Psi_3, \\
& - \bar{\alpha}_4 - X'_4 \delta_4 - \beta_{14} h_1 - \beta_{34} (y^v + w_1 h_1) - \beta_{44} \widetilde{h}_2, \\
& w_1 - X'_5 \Psi_5, \varepsilon_6, \varepsilon_7) d\varepsilon_2 d\varepsilon_6 d\varepsilon_7 |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_{24} \widetilde{h}_2 - \beta_{33} (y^v Ew_2 + h_2 Ew_2^2) - \beta_{34} \widetilde{h}_2 Ew_2, y^v - X'_3 \Psi_3, \\
& - \bar{\alpha}_4 - X'_4 \delta_4 - \beta_{24} h_2 - \beta_{34} (y^v + h_2 Ew_2) - \beta_{44} \widetilde{h}_2 \\
& \varepsilon_5, Ew_2 - X'_6 \Psi_6, Ew_2^2 - X'_7 \Psi_7) d\varepsilon_1 d\varepsilon_5 |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, -\bar{\alpha}_4 - X'_4 \delta_4 \\
& - \beta_{34} y^v - \beta_{44} \widetilde{h}_2, \varepsilon_5, \varepsilon_6, \varepsilon_7) d\varepsilon_1 d\varepsilon_2 d\varepsilon_5 d\varepsilon_6 d\varepsilon_7.
\end{aligned} \tag{5A.30}$$

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

et où $|J_1|$, $|J_2|$ et $|J_3|$ représentent la valeur absolue du déterminant de la matrice jacobienne de la transformation des termes inobservés ε_i en termes observés $(h_1, h_2, y^v, \widetilde{h}_2, w_1, Ew_2, Ew_2^2)$.

ε_2^* est tel que $m_2(h_1, 0, y^v, \widetilde{h}_2, w_1, Ew_2, Ew_2^2, X_2, \varepsilon_2^*) = 0$:

$$\begin{aligned}
\varepsilon_2^* = & -\bar{\alpha}_2 - X'_2 \delta_2 - \beta_{12} h_1 - \beta_{23} (y^v + w_1 h_1) - \beta_{24} \widetilde{h}_2 \\
& - (X'_6 \Psi_6) (\alpha_3 + \beta_{13} h_1 + \beta_{33} (y^v + w_1 h_1) + \beta_{34} \widetilde{h}_2) \\
& - \varepsilon_6 (\alpha_3 + \beta_{13} h_1 + \beta_{33} (y^v + w_1 h_1) + \beta_{34} \widetilde{h}_2).
\end{aligned} \tag{5A.31}$$

ε_1^* est tel que $m_1(0, h_2, y^v, \widetilde{h}_2, w_1, Ew_2, Ew_2^2, X_1, \varepsilon_1^*) = 0$. Autrement dit :

$$\begin{aligned} \varepsilon_1^* = & -\bar{\alpha}_1 - X_1' \delta_1 - \beta_{12} h_2 - \beta_{13} (y^v + h_2 Ew_2) - \beta_{14} \widetilde{h}_2 \\ & - (X_5' \Psi_5) (\alpha_3 + \beta_{23} h_2 + \beta_{33} (y + h_2 Ew_2) + \beta_{34} \widetilde{h}_2) \\ & - \varepsilon_5 (\alpha_3 + \beta_{23} h_2 + \beta_{33} (y^v + h_2 Ew_2) + \beta_{34} \widetilde{h}_2). \end{aligned} \quad (5A.32)$$

ε_1^{**} est tel que $m_1(0, 0, y^v, \widetilde{h}_2, w_1, w_2, X_1, \varepsilon_1^{**}) = 0$:

$$\begin{aligned} \varepsilon_1^{**} = & -\bar{\alpha}_1 - X_1' \delta_1 - \beta_{13} y^v - \beta_{14} \widetilde{h}_2 - (X_5' \Psi_5) (\alpha_3 + \beta_{33} y^v + \beta_{34} \widetilde{h}_2) \\ & - \varepsilon_5 (\alpha_3 + \beta_{33} y^v + \beta_{34} \widetilde{h}_2), \end{aligned} \quad (5A.33)$$

et ε_2^{**} est tel que $m_2(0, 0, y^v, \widetilde{h}_2, w_1, w_2, X_2, \varepsilon_2^{**}) = 0$:

$$\begin{aligned} \varepsilon_2^{**} = & -\bar{\alpha}_2 - X_2' \delta_2 - \beta_{23} y^v - \beta_{24} \widetilde{h}_2 - (X_6' \Psi_6) (\alpha_3 + \beta_{33} y^v + \beta_{34} \widetilde{h}_2) \\ & - \varepsilon_6 (\alpha_3 + \beta_{33} y^v + \beta_{34} \widetilde{h}_2), \end{aligned} \quad (5A.34)$$

La méthode du maximum de vraisemblance à information complète a été mise en œuvre afin d'estimer ces deux modèles. La programmation de la fonction de vraisemblance a été réalisée en langage Gauss (sous-routines Optnum et Maxlik), ainsi qu'en langage Fortran.⁴⁰ En dépit de plusieurs mois de persévérance et d'essais de diverses valeurs de départ, aucun de ces modèles n'a malheureusement pu converger.

⁴⁰ Etant donnée la longueur des lignes de programmation, les différents programmes ne sont pas reportés ici, mais ils sont naturellement disponibles à la demande.