

**HEC Montréal**

**L'entrepreneuriat chez les immigrants de 2<sup>e</sup> génération : L'ethnicité joue-t-elle  
toujours un rôle ?**

**par**

**Laurent Giguère**

**Sciences de la gestion  
(Option Économie Appliquée)**

*Mémoire présenté en vue de l'obtention  
du grade de maîtrise ès sciences  
(M.Sc.)*

Avril 2017  
© Laurent Giguère, 2017

## Résumé

Ce mémoire se veut une étude d'une des particularités les plus notoires des immigrants dans les pays développés : l'entrepreneuriat. En effet, dans la plupart des pays développés, nombreuses sont les communautés ethniques qui affichent un taux d'entrepreneuriat nettement différent de celui qui prévaut chez les citoyens d'origine. L'objectif principal de ce mémoire est d'analyser ce phénomène en détail et de manière séparée pour les immigrants de 1<sup>ère</sup> et de 2<sup>e</sup> génération.

Pour ce faire, nous isolons, pour chaque génération d'immigrants, l'effet d'appartenir aux différents groupes ethniques sur la probabilité d'être entrepreneur et tentons d'expliquer la provenance de cet effet. Par la suite, nous comparons les résultats obtenus pour chaque génération afin de dresser un portrait de l'évolution de l'entrepreneuriat chez les communautés ethniques entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants. Cette analyse est rendue possible par l'application de différentes méthodes économétriques sur des données de recensement américaines et par l'utilisation d'un modèle théorique servant à interpréter les résultats obtenus grâce à ces méthodes.

Nous arrivons à la conclusion que l'appartenance à un groupe ethnique minoritaire a généralement un effet positif sur la décision des immigrants de se lancer en affaires, un effet que nous attribuons à l'existence de réseaux d'affaires au sein des communautés ethniques. De plus, bien qu'il diminue en importance entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants, cet effet ne disparaît généralement pas à travers le temps. Par conséquent, nous déduisons qu'au-delà de la 1<sup>ère</sup> génération d'immigrants, une part importante des citoyens membres de communautés ethniques continue de se comporter différemment des citoyens de souche dans le marché du travail.

## Table des matières

I. Introduction.....	4
II. Revue de littérature.....	7
II.I. Le choix de devenir entrepreneur : cadre théorique .....	7
II.II. Le choix de devenir entrepreneur : cadre empirique .....	11
II.III. L'entrepreneuriat chez les communautés ethniques .....	13
II.IV. L'entrepreneuriat immigrant à travers les générations .....	16
II.V. La transmission intergénérationnelle des compétences entrepreneuriales....	18
III. Données et Statistiques Descriptives .....	21
III.I Sources .....	21
III.II Méthode d'échantillonnage.....	21
III.III Distribution de l'entrepreneuriat.....	23
III.IV Distribution des caractéristiques sociodémographiques.....	26
III.V Distribution des salaires et des revenus d'entreprise.....	32
IV. Méthodologie .....	37
IV.I L'Effet de l'ethnicité sur le choix sectoriel.....	37
IV.II L'Effet de l'ethnicité sur le choix sectoriel : comparaison entre la 1 <sup>ère</sup> et la 2 <sup>e</sup> génération d'immigrants.....	41
IV.III L'Effet de l'ethnicité sur le salaire et sur le revenu d'entreprise.....	44
IV.IV L'Explication de l'effet de l'ethnicité sur le choix sectoriel : cadre théorique	46
IV.V L'Explication de l'effet de l'ethnicité sur le choix sectoriel : cadre empirique	52
V. Estimations et Résultats .....	60
V.I. Équations de choix sectoriel .....	60
V.II. L'évolution de l'effet de l'ethnicité sur le choix sectoriel entre la 1 <sup>ère</sup> et la 2 <sup>e</sup> génération d'immigrants.....	64
V.III Équations de revenu.....	67
V.IV Les phénomènes à la source de l'effet de l'ethnicité sur le choix sectoriel.....	69
VI. Conclusion .....	76
Bibliographie .....	78
Annexe 1 : Tableaux.....	80
Annexe 2 : Graphiques.....	91
Annexe 3 : Méthode DiNardo-Fortin-Lemieux.....	99

## **Remerciements**

Je tiens à remercier ma directrice de mémoire, Mme Ana Damas de Matos, pour son support et ses nombreux conseils tout au long de la rédaction du mémoire. Je tiens également à remercier Mme Marie Allard, pour ses conseils lors de l'Atelier de Recherche, de même que M. Mario Samano, pour ses réponses à quelques questions. Finalement, je suis très reconnaissant envers les étudiants de la M.Sc. avec qui j'ai pris part à l'Atelier de Recherche pour leurs nombreux commentaires constructifs.



## I. Introduction

Dans la plupart des pays développés, les immigrants de 1<sup>ère</sup> génération se distinguent par leur activité entrepreneuriale élevée. En effet, dans ces pays, les immigrants de première génération affichent un taux d'entrepreneuriat beaucoup plus élevé que celui de la population d'origine, et ce, de manière constante au fil des années. Des écarts tout aussi importants de taux d'entrepreneuriat sont également observés entre les différentes communautés ethniques. Dans certains, ces écarts demeurent largement significatifs même lorsque l'on tient compte des caractéristiques individuelles qui influencent la décision de devenir entrepreneur et qui sont susceptibles d'être distribuées différemment d'une communauté à l'autre, laissant du même coup entrevoir un effet indépendant de l'ethnicité sur la probabilité d'être entrepreneur.

Or, bien que le lien entre l'ethnicité et la probabilité d'être entrepreneur ait fait l'objet d'une littérature considérable en économie et en sociologie, aucune étude n'a jusqu'à maintenant cherché à documenter l'évolution de ce lien entre la première et la deuxième génération d'immigrants. Ainsi, l'objectif de ce mémoire est de combler un vide dans la littérature en répondant à la question suivante : toutes choses étant égales par ailleurs, les immigrants de deuxième génération choisissent-ils de devenir entrepreneur pour les mêmes raisons que les citoyens d'origine du pays d'accueil, ou bien sont-ils toujours influencés par des caractéristiques propres à leur communauté ethnique de la même façon que leurs parents l'étaient?

Étant donné que les études qui rapportent un effet significatif de l'ethnicité utilisent des échantillons composés d'immigrants de différentes générations, il serait surprenant qu'en moyenne, l'influence de l'ethnicité disparaisse au-delà de la première génération d'immigrants. Néanmoins, au sein de chaque communauté, nous n'avons aucune idée du changement de magnitude de cet effet d'une génération à l'autre. De plus, même si l'effet estimé s'avérait être le même, celui-ci pourrait être causé par des facteurs différents. Ainsi, l'étude de ces facteurs nous conduit vers une autre question : dans quelle mesure l'effet de l'ethnicité sur le choix sectoriel est-il attribuable aux mêmes facteurs d'une génération à l'autre ?

L'importance de répondre aux deux questions présentées ci-haut, à savoir le changement de magnitude de l'effet ethnique entre la première et la deuxième génération d'immigrants ainsi que les facteurs qui expliquent ce changement, s'inscrit dans la foulée de l'attention croissante consacrée aux questions d'intégration des communautés ethniques depuis quelques décennies dans les littératures économique et sociologique, mais aussi dans la sphère publique. D'une part, savoir s'il y a disparition ou persistance de l'effet d'appartenir à un groupe ethnique sur le choix sectoriel est une façon d'évaluer, à travers le temps, le niveau d'intégration économique des communautés ethniques au marché du travail. Il s'agit aussi d'une façon de savoir si l'absence de convergence du taux d'entrepreneuriat entre les communautés ethniques est davantage attribuable à une persistance des éléments propres à chaque communauté qu'à une persistance des écarts en matière de caractéristiques socio-économiques telles que l'éducation. D'autre part, le fait d'établir une comparaison entre les facteurs qui expliquent l'effet de l'ethnicité chez la première et chez la deuxième génération nous éclaire sur les raisons pour lesquelles il y a persistance ou disparition de cet effet à travers le temps. De là, il devient possible de concevoir des politiques publiques qui favorisent l'intégration économique ; par exemple en luttant contre la discrimination, ou encore en facilitant l'acquisition de compétences qui sont valorisées dans le marché du travail local.

Notre étude utilise des données américaines tirées du recensement de 1980 ainsi que des suppléments économiques annuels des enquêtes CPS pour la période qui s'étale de 2006 à 2015. Pour chaque groupe ethnique, nous reconstituons des cohortes générationnelles qui servent ensuite à notre analyse économétrique, laquelle se base sur des techniques déjà largement utilisées dans la littérature. Pour les groupes dont les immigrants de 1<sup>ère</sup> génération ont, toutes choses étant égales par ailleurs, plus ou moins de chances d'être entrepreneur que les Américains de souche, nous observons en général une persistance de cet écart chez les immigrants de 2<sup>e</sup> génération, bien que le niveau de persistance varie considérablement d'un groupe ethnique à l'autre. Par ailleurs, même si nous observons des signes de discrimination dans le marché du travail, nos résultats nous laissent croire que l'existence de réseaux d'affaires à l'échelle de la communauté ethnique est le principal facteur responsable de l'effet de l'ethnicité sur la probabilité d'être entrepreneur, peu importe la génération. Par conséquent, nous concluons que la

persistance de l'effet de l'ethnicité entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants résulte d'une persistance de l'avantage économique conféré par ces réseaux sur la même période.

La suite de ce mémoire se profile comme suit. Dans un premier temps, nous présentons dans la section *Revue de Littérature* les articles qui ont inspiré notre réflexion et sur lesquels s'appuie notre méthodologie. Par la suite, nous consacrons une section entière aux données et aux statistiques descriptives que nous jugeons pertinentes pour la poursuite de notre analyse. Nous enchaînons par la suite avec la section *Méthodologie*, dans laquelle nous expliquons en détail les équations que nous estimons et dans laquelle nous présentons le cadre théorique qui sert de référence pour l'interprétation de nos résultats. Finalement, le mémoire se termine par la présentation de nos résultats empiriques, suivie d'une brève conclusion.

## II. Revue de littérature

### II.I. Le choix de devenir entrepreneur : cadre théorique

Dans la littérature économique sur l'entrepreneuriat, la décision de devenir entrepreneur est presque toujours représentée comme un choix rationnel entre un emploi salarié et un emploi non salarié [Borjas (1986), Fairlie & Meyer (1996), Dunn & Holtz-Eakin (2000), Clark & Drinkwater (2000)]. Plus précisément, un individu choisirait de devenir entrepreneur plutôt que salarié si l'espérance de son revenu en tant qu'entrepreneur était plus élevée que l'espérance de son revenu en tant qu'employé qui reçoit un salaire :

$$E(R_E) - E(R_S) > 0 \quad (1)$$

où  $R_E$  représente le revenu provenant de l'entrepreneuriat et  $R_S$  le revenu provenant d'un emploi salarié. Cette équation stipule que deux effets peuvent influencer le choix de devenir entrepreneur : une augmentation du revenu espéré en tant qu'entrepreneur ou une diminution du salaire espéré dans l'économie. Différents facteurs peuvent influencer le revenu qu'un individu peut espérer en tant qu'entrepreneur ou travailleur salarié. Le niveau d'éducation et l'âge (i.e. le nombre d'années d'expérience sur le marché du travail) en sont quelques exemples. Des facteurs non observables et difficilement quantifiables tels que le niveau d'habileté ont également un impact potentiellement considérable sur le revenu qu'un individu peut espérer dans chacun des deux secteurs.

Puisqu'on peut s'attendre à ce que les facteurs énumérés ci-dessus n'aient pas le même effet sur le revenu espéré d'un entrepreneur que sur celui d'un travailleur salarié, il convient généralement de traiter le choix de devenir entrepreneur comme un processus de sélection. Par processus de sélection, on entend le mécanisme par lequel les individus font leur choix en fonction de leurs caractéristiques et du rendement que leur procurent ces caractéristiques en termes de revenu espéré dans chacun des deux secteurs. Le principe de sélection fit sont entrée dans la littérature économique avec l'article de Roy (1951), article dont l'analyse qualitative tente de déterminer comment des travailleurs choisissent entre deux métiers en fonction de leurs habiletés et des

habiletés requises pour exercer chacun des deux métiers. Plus tard, Borjas (1987) formalisa le modèle qualitatif mis de l'avant par Roy dans le contexte du choix d'immigrer ou non vers un autre pays. Puis, dans l'article de Borjas & Bronars (1989), le même modèle fut utilisé, cette fois pour décrire la décision de devenir entrepreneur ou non. Ce modèle suppose tout d'abord que le revenu d'un entrepreneur ou d'un travailleur salarié est fonction d'une série de facteurs observables compris dans la matrice  $X$  et du niveau d'habileté intrinsèque représenté par le terme d'erreur  $\varepsilon$ . Par souci de simplicité, nous écrivons le revenu en niveau plutôt qu'en logarithme, contrairement à ce qui est fait dans l'article. Les deux équations de revenu sont donc les suivantes :

$$R_E = (XB_E + \varepsilon_E) \quad (2)$$

$$R_S = (XB_S + \varepsilon_S) \quad (3)$$

Une fois de plus,  $R_E$  et  $R_S$  représentent le revenu d'un entrepreneur et d'un travailleur salarié respectivement. Par ailleurs, les termes d'erreur  $\varepsilon_E$  et  $\varepsilon_S$  suivent une distribution normale conjointe de moyenne nulle et ont comme variance  $\sigma_E^2$  et  $\sigma_S^2$  respectivement. Suivant la logique de l'équation (1), un individu décide de devenir entrepreneur lorsque son revenu est plus élevé dans l'entrepreneuriat que dans le secteur salarié. En d'autres mots, un travailleur choisit l'entrepreneuriat lorsque la condition suivante est satisfaite :

$$(XB_E + \varepsilon_E) - (XB_S + \varepsilon_S) > 0 \quad (4)$$

Il est à noter qu'en invoquant cette condition, nous délaissions le cadre d'analyse dynamique employé par Borjas & Bronars (1989) en faveur d'un cadre d'analyse statique dans lequel il n'existe aucun coût de déplacement associé à une transition vers l'entrepreneuriat en provenance du secteur salarié. Maintenant, si l'on désire exprimer la probabilité qu'un travailleur choisisse d'être entrepreneur, il suffit de réécrire la condition énoncée ci-haut de la manière suivante :

$$P(y = 1) = P[(\varepsilon_E - \varepsilon_S) > (XB_S - XB_E)] \quad (5)$$

où  $y$  est une variable binaire qui vaut 1 lorsque l'individu est entrepreneur. En remplaçant la différence entre les termes d'erreur  $\varepsilon_E$  et  $\varepsilon_S$  par  $v$  et en standardisant la distribution normale rattachée à cette différence, nous obtenons l'expression suivante :

$$P(y = 1) = P \left[ \frac{v}{\sigma_v} > \frac{(XB_S - XB_E)}{\sigma_v} \right] \quad (6)$$

À partir de cette expression, la manière la plus simple d'illustrer le processus de sélection est de définir l'espérance du revenu d'un entrepreneur ou d'un travailleur salarié lorsque celui-ci se trouve dans le sous-ensemble de la population auquel il est confiné (i.e. tel qu'observé dans les données). Par exemple, dans le cas d'un entrepreneur, nous voudrions exprimer son revenu conditionnellement au fait qu'il ait choisi d'être entrepreneur et conditionnellement à l'ensemble de ses caractéristiques observables. L'équation suivante résume cette situation :

$$E(w_E|X, y = 1) = XB_E + E(\varepsilon_E|X, y = 1) \quad (7)$$

Puis, si nous réutilisons la condition d'inégalité comprise dans l'équation (6), nous pouvons réécrire l'équation ci-dessus de la manière suivante :

$$E(w_E|X, y = 1) = XB_E + E \left( \varepsilon_E \left| \frac{v}{\sigma_v} > \frac{(XB_S - XB_E)}{\sigma_v} \right. \right) \quad (8)$$

Ainsi, le revenu d'un entrepreneur tel qu'observé dans les données dépend de l'espérance conditionnelle de son revenu d'entrepreneuriat étant donné ses caractéristiques observables ( $XB_E$ ) et de l'espérance de son niveau d'habileté étant donné le fait qu'il a choisi d'être entrepreneur. Suivant le même raisonnement, le revenu d'un travailleur salarié lorsque l'on tient compte du processus de sélection s'écrit comme suit :

$$E(w_S|X, y = 0) = XB_S + E \left( \varepsilon_S \left| \frac{v}{\sigma_v} < \frac{(XB_S - XB_E)}{\sigma_v} \right. \right) \quad (9)$$

La raison pour laquelle ce modèle ne s'intéresse qu'aux revenus qui sont conditionnels aux caractéristiques observables est qu'il est possible de tenir compte de l'assignation non aléatoire relativement à ces caractéristiques en incluant les variables de contrôle appropriées dans nos régressions. Finalement, grâce à une série de manipulations algébriques, il est possible de réécrire les équations (8) et (9) de la manière suivante<sup>1</sup> :

---

<sup>1</sup> Ces deux équations sont basées sur les équations auxquelles arrivent Borjas & Bronars (1989).

$$E(w_E|X, y = 1) = XB_E + \frac{\sigma_E \sigma_S}{\sigma_v} \left( \frac{\sigma_E}{\sigma_S} - \rho \right) \left( \frac{\phi \left( \frac{XB_S - XB_E}{\sigma_v} \right)}{1 - \Phi \left( \frac{XB_S - XB_E}{\sigma_v} \right)} \right) \quad (10)$$

$$E(w_S|X, y = 0) = XB_S + \frac{\sigma_E \sigma_S}{\sigma_v} \left( \rho - \frac{\sigma_S}{\sigma_E} \right) \left( \frac{-\phi \left( \frac{XB_S - XB_E}{\sigma_v} \right)}{\Phi \left( \frac{XB_S - XB_E}{\sigma_v} \right)} \right) \quad (11)$$

où  $\rho$  correspond au coefficient de corrélation entre les termes d'erreur  $\varepsilon_E$  et  $\varepsilon_S$  ( $\rho = \frac{\sigma_{ES}}{\sigma_E \sigma_S}$ ) et où le rapport entre la fonction de densité d'une loi normale ( $\phi$ ) et sa fonction de répartition ( $\Phi$ ), toutes deux évaluées à  $\left( \frac{XB_S - XB_E}{\sigma_v} \right)$ , correspond au ratio de Mills inversé. Sachant que la sélection est représentée par le deuxième terme dans chaque équation, on constate que deux paramètres,  $\rho$  et  $\frac{\sigma_E}{\sigma_S}$ , déterminent à eux seuls la direction du processus de sélection. En d'autres mots, ces deux paramètres nous permettent de savoir s'il y a sélection positive ou négative dans chacun des deux secteurs. Premièrement, pour que le signe associé à la sélection diffère d'un secteur à l'autre, le terme  $\rho$  doit absolument être positif. Cela revient à dire que les habiletés requises dans les deux secteurs doivent être sensiblement les mêmes de façon à ce que la corrélation entre le niveau de succès dans l'entrepreneuriat et le niveau de succès dans le travail salarié soit positive. Autrement, les meilleurs entrepreneurs deviendraient entrepreneurs et les meilleurs travailleurs salariés deviendraient travailleurs salariés, ce qui entraînerait une sélection positive dans les deux secteurs. Deuxièmement, du moment que  $\rho$  est *suffisamment* positif<sup>2</sup>, la sélection positive aura lieu dans le secteur où la variance conditionnelle du revenu est la plus élevée. En d'autres mots, les individus les plus habiles choisiront d'évoluer dans le secteur où le rendement sur l'habileté est le plus haut, pourvu que l'habileté soit suffisamment transférable d'un secteur à l'autre. Nous résumons les deux conditions décrites ci-haut dans le tableau suivant :

---

<sup>2</sup> Le terme *suffisamment* signifie que  $\rho > \min \left( \frac{\sigma_E}{\sigma_S}, \frac{\sigma_S}{\sigma_E} \right)$ . Pour plus d'informations, voir Borjas (1987).

Signe de la Sélection	$\rho > 0$ (suffisamment)		$\rho < 0$ ou $\rho > 0$ (non suffisamment)
	$\frac{\sigma_E}{\sigma_S} > \rho$	$\frac{\sigma_E}{\sigma_S} < \rho$	
Entrepreneuriat	(+)	(-)	(+)
Travail Salarié	(-)	(+)	(+)

L'importance de comprendre le choix sectoriel comme un processus de sélection relativement au niveau d'habileté du travailleur est intrinsèquement liée à la deuxième partie de notre analyse, où nous tentons d'expliquer la provenance de l'effet de l'ethnicité sur la propension à devenir entrepreneur. En effet, qu'il s'agisse de la discrimination dans le marché du travail et dans le marché des produits, de l'impact des réseaux d'affaires et des enclaves ethniques ou encore de l'impact de simples facteurs culturels, ces mécanismes impliquent potentiellement tous une répartition inégale des individus entre les deux secteurs de l'emploi en ce qui a trait à leur niveau d'habileté. Dans la section *Méthodologie*, nous employons le modèle présenté ci-haut afin d'établir un lien entre ces différents mécanismes, la sélection des individus et l'effet de l'ethnicité sur le salaire et sur le revenu d'entreprise. En procédant ainsi, il devient alors possible de comprendre quels sont les mécanismes à l'origine de l'effet de l'ethnicité sur la probabilité de devenir entrepreneur.

## II.II. Le choix de devenir entrepreneur : cadre empirique

Bien que le problème de maximisation du revenu et le modèle de sélection qui en découle constituent la principale représentation formelle du choix de devenir entrepreneur, très peu d'études empiriques dans la littérature testent directement la relation impliquée par cette représentation. Comme principale raison, on note le fait que la plupart des études qui souhaitent mieux comprendre quels sont les déterminants de l'entrepreneuriat n'ont pas besoin de connaître le mécanisme via lequel les déterminants présumés ont un effet sur la propension à devenir entrepreneur. En effet, outre Clark & Drinkwater (2000) qui, dans leur modèle, incluent à la fois le différentiel de revenu espéré et l'ensemble des caractéristiques sociodémographiques, la plupart des études qui portent sur les déterminants de l'entrepreneuriat utilisent une spécification



empirique qui fait abstraction du processus de choix en fonction du revenu espéré. En d'autres termes, cette spécification ne comprend que les caractéristiques individuelles sociodémographiques, indépendamment du fait qu'elles impactent ou non le choix sectoriel par l'entremise de leur effet sur le revenu dans chaque secteur. L'équation suivante illustre la forme générale de cette spécification :

$$P_i = X_i B + \varepsilon_i \quad (12)$$

où  $P_i$  est une variable continue représentant la propension à devenir entrepreneur,  $X_i$  une matrice contenant un ensemble de variables sociodémographiques qui ont un impact sur la propension à devenir entrepreneur et  $\varepsilon_i$  un terme d'erreur exogène. De manière implicite, la spécification ci-haut suppose que les variables explicatives contenues dans  $X_i$  peuvent influencer la propension à devenir entrepreneur à la fois par l'entremise de leur effet sur le différentiel sectoriel de revenu<sup>3</sup> et de leur effet sur une quelconque préférence pour l'entrepreneuriat. Pour voir cela, représentons le différentiel de revenu par  $D$  et la préférence pour l'entrepreneuriat par  $Z$ . Nous pouvons alors réécrire l'équation (12) de la manière suivante :

$$P_i(D(X_i), Z(X_i)) = X_i B + \varepsilon_i \quad (13)$$

Un cadre analytique similaire à l'équation ci-dessus est présenté de manière explicite dans Dunn & Holtz-Eakin (2000). Toutefois, ce genre de cadre analytique est implicitement utilisé dans la majorité des articles qui étudient les déterminants de l'entrepreneuriat, notamment dans Fairlie & Meyer (1996). Partant de l'équation ci-haut, pour qu'un individu décide de devenir entrepreneur, il faut que sa propension à le devenir soit positive. Ainsi, l'équation estimée doit prendre la forme suivante :

$$E_i = Prob(P_i > 0) = Prob(\varepsilon_i > -X_i B) \quad (14)$$

---

<sup>3</sup> Nous entendons par différentiel sectoriel de revenu la différence entre le revenu espéré en tant qu'entrepreneur et en tant que travailleur salarié.

où  $E_i$  est une variable dichotomique qui vaut 1 si l'individu est entrepreneur et 0 s'il est salarié. Cette équation constitue la base des modèles économétriques que l'on retrouve dans la littérature sur les déterminants de l'entrepreneuriat, et donc la base de notre méthodologie empirique. Bien que chaque étude utilise des variables qui diffèrent légèrement les unes des autres, plusieurs prédicteurs de la décision d'être entrepreneur sont systématiquement inclus dans la matrice  $X_i$  des modèles étant donné leur pouvoir explicatif élevé ou potentiellement élevé. C'est le cas notamment de variables sociodémographiques telles que l'âge, le statut marital, le nombre d'enfants, l'éducation, le lieu de résidence géographique et, bien sûr, l'ethnicité.

### II.III. L'entrepreneuriat chez les communautés ethniques

Dans leur article, Fairlie & Meyer (1996) explorent les écarts de taux d'entrepreneuriat observés entre les différents groupes ethniques aux États-Unis. Dans un premier temps, ils répertorient ces écarts et tentent de mesurer à quel point ceux-ci sont attribuables aux changements dans la distribution des caractéristiques individuelles d'un groupe à l'autre. Pour ce faire, ils estiment des modèles Probit contenant l'ensemble des variables sociodémographiques susceptibles d'avoir un effet sur la propension à devenir entrepreneur ainsi qu'un vecteur de variables dichotomiques servant à identifier le groupe ethnique auquel appartient chaque individu dans l'échantillon. Puisque le groupe de référence est constitué des individus de descendance américaine, les coefficients associés à ces variables dichotomiques représentent la déviation de la probabilité d'être entrepreneur par rapport à la probabilité associée aux Américains de souche une fois toutes les autres variables prises en compte.

Les données utilisées dans l'échantillon sont celles du recensement américain de 1990. Pour construire leur définition de l'ethnicité, les auteurs utilisent à la fois les réponses des questions sur la descendance, la race ainsi que l'appartenance à un groupe ethnique hispanique. Cette définition de l'ethnicité contraste avec notre méthode, qui utilise comme variable uniquement le pays de naissance des parents pour définir le groupe ethnique auquel appartient chaque individu.

Les résultats des régressions Probit montrent qu'une partie des écarts de taux d'entrepreneuriat sont attribuables aux caractéristiques individuelles qui sont distribuées différemment d'un groupe à l'autre, car les effets fixes associés à chacun des groupes donnent lieu à des écarts plus faibles que ceux réellement observés. Cependant, étant donné qu'une bonne partie des variables dichotomiques pour l'ethnicité affichent des coefficients statistiquement significatifs, les écarts entre les différents groupes demeurent significatifs.

Dans un deuxième temps, les auteurs cherchent à expliquer pourquoi, même après avoir contrôlé pour les caractéristiques individuelles observables, des écarts entre les groupes ethniques subsistent. Pour y parvenir, ils utilisent les coefficients des variables dichotomiques associées à l'ethnicité qu'ils ont estimés dans les régressions Probit. Ces coefficients sont ensuite utilisés dans des régressions afin de tester différentes théories qui pourraient justifier le fait que l'ethnicité a un effet indépendant sur le taux d'entrepreneuriat.

Parmi ces théories, les auteurs testent l'hypothèse selon laquelle l'effet propre à chaque groupe ethnique serait corrélé avec le différentiel de revenu entre l'entrepreneuriat et le travail salarié pour chaque groupe. Pour tester cette hypothèse, les auteurs devraient normalement estimer un modèle dans lequel l'effet de l'ethnicité (le coefficient ethnique) est régressé sur une mesure ajustée du différentiel de revenu entre l'entrepreneuriat et le travail salarié, tel qu'exprimé dans l'équation suivante :

$$\hat{\alpha}_j = \delta + \beta(\hat{w}_j^E - \hat{w}_j^S) + \varepsilon_j \quad (15)$$

où  $\hat{\alpha}_j$  est le coefficient de l'effet du groupe ethnique j sur la propension à devenir entrepreneur dans le modèle probit et où  $\hat{w}_j^E$  et  $\hat{w}_j^S$  sont les revenus prédits pour l'entrepreneuriat et le travail salarié respectivement, lesquels ont été calculés à partir d'équations servant à expliquer le revenu. Les équations pour le revenu contiennent l'ensemble des variables explicatives généralement utilisées dans la littérature telles que l'âge, le niveau d'éducation et le statut familial ainsi qu'un vecteur de variables dichotomiques identifiant le groupe ethnique. Pour cette raison, dans l'équation ci-haut, les revenus prédits ne diffèrent d'une communauté à l'autre que de par l'effet indépendant de l'ethnicité sur chaque type de revenu. Cette particularité permet aux auteurs

d'estimer l'importance du différentiel de revenu ethnique ajusté de manière beaucoup plus simple, soit en remplaçant les revenus prédits par les coefficients ethniques qui proviennent des régressions expliquant le revenu, tel qu'illustré dans l'équation ci-dessous :

$$\hat{\alpha}_j = \delta + \beta(\hat{\gamma}_j^E - \hat{\gamma}_j^S) + \varepsilon_j \quad (16)$$

où  $\hat{\gamma}_j^E$  et  $\hat{\gamma}_j^S$  sont les coefficients de l'effet du groupe ethnique  $j$  sur le revenu d'entrepreneuriat et sur le revenu de travail salarié respectivement. Bien que dans les équations de revenu, ces deux coefficients ne soient par construction pas corrélés avec des caractéristiques individuelles observables, ils demeurent potentiellement biaisés en raison du processus de sélection relativement à l'habileté (caractéristiques individuelles non observables) qui est inhérent à la décision de devenir entrepreneur et dont nous avons parlé plus tôt. Pour remédier à ce problème, les auteurs utilisent la méthode de Heckman à deux étapes pour estimer leur modèle, méthode que nous préférons ne pas utiliser en raison de la fragilité d'une des conditions nécessaires à son application<sup>4</sup>. Ils obtiennent comme résultat que le différentiel de revenu ethnique ajusté possède un pouvoir explicatif important sur la tendance naturelle d'un groupe ethnique à produire des entrepreneurs. De plus, en vérifiant s'il y a corrélation entre l'effet de l'ethnicité sur le salaire et l'effet de l'ethnicité sur la propension à être entrepreneur, ils déterminent si oui ou non l'effet trouvé pourrait être attribuable à de la discrimination dans le marché du travail. Leur conclusion est que la discrimination ne joue aucun rôle et que l'effet est plutôt attribuable aux retours exceptionnels que procure l'entrepreneuriat chez certains groupes, sans pour autant expliquer d'où émanent ces retours exceptionnels.

Dans ce mémoire, nous reproduisons partiellement la méthode empirique employée par Fairlie & Meyer (1996) que nous avons décrite ci-haut. Toutefois, plutôt que d'estimer les régressions sur un seul échantillon composé d'individus appartenant à de multiples générations d'immigrants, nous appliquons la méthode séparément pour la première et la deuxième génération

---

<sup>4</sup> Dans notre contexte, l'utilisation du Heckman à deux étapes exigerait que le modèle de choix sectoriel comporte au minimum une variable explicative qui n'a aucun impact sur le revenu d'un entrepreneur, ce qui est difficilement envisageable.

d'immigrants. Nous essayons ensuite de voir dans quelle mesure nos résultats sont comparables à ceux obtenus par les deux auteurs.

De plus, contrairement à ce qui est fait dans l'article, nous ne tentons pas de corriger le biais de sélection au moyen d'une méthode économétrique telle que le modèle de Heckman à deux étapes. Plutôt, nous employons un modèle de sélection très similaire à celui présenté dans l'article de Borjas & Bornars (1989) afin de prédire l'effet sur le revenu d'entreprise et sur le salaire des divers facteurs susceptibles d'entraîner une sélection des individus en fonction de leur niveau d'habileté. Les prédictions que nous obtenons nous permettent ensuite d'identifier les facteurs responsables de l'effet de l'ethnicité sur la propension à devenir entrepreneur à partir des estimés de l'effet des différents groupes ethniques sur le revenu d'entreprise et sur le salaire.

#### II.IV. L'entrepreneuriat immigrant à travers les générations

Dans leur article, Hou & al. (2013) souhaitent vérifier si les immigrants de deuxième génération sont plus ou moins susceptibles de devenir entrepreneur que leurs parents. Pour y parvenir, ils estiment des régressions Probit sur trois échantillons distincts : les immigrants de 1<sup>ère</sup> génération, les enfants de ces derniers (les immigrants de 2<sup>e</sup> génération) ainsi qu'un échantillon réunissant ces deux groupes. La plupart des variables incluses dans le modèle sont les mêmes que celles que l'on retrouve généralement dans la littérature, à savoir des variables sociodémographiques telles que l'éducation, le statut marital et le nombre d'enfants ainsi qu'un vecteur de variables dichotomiques qui servent à identifier la région de résidence et le groupe ethnique. Les auteurs utilisent ensuite les estimés des coefficients associés à ces variables pour calculer la probabilité moyenne d'être entrepreneur chez les immigrants de 1<sup>ère</sup> génération et chez les immigrants de 2<sup>e</sup> génération. Ils calculent ensuite la différence entre ces deux moyennes et appliquent une variante de la décomposition d'Oaxaca-Blinder sur cette différence afin d'établir une distinction entre la portion attribuable à la distribution inégale des caractéristiques sociodémographiques et la portion attribuable aux différentes valeurs que prennent les coefficients des régressions (i.e. les retours sur les caractéristiques démographiques).

L'étude portant sur le Canada, les auteurs utilisent des données en provenance du recensement canadien de 1981 et du recensement canadien de 2006. À partir de ces données, ils bâtissent des cohortes synthétiques pour chacune des générations précédemment mentionnées en restreignant, dépendamment de la génération, l'âge et le lieu de naissance ou encore l'âge et le lieu de naissance des parents. Cette méthode est d'ailleurs celle que nous utilisons pour bâtir les cohortes synthétiques<sup>5</sup> dont nous avons besoin pour effectuer notre analyse intergénérationnelle. Il s'agit également de la même méthode que celle qui est généralement utilisée dans la littérature sur la mobilité sociale des communautés ethniques, les articles de Card & al. (1998) et Aydemir & al. (2009) étant des références en la matière.

Les auteurs obtiennent comme résultat que la diminution du taux d'entrepreneuriat observée entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants est surtout attribuable à l'évolution de la distribution des caractéristiques sociodémographiques. Plus précisément, les résultats de la décomposition employée par les auteurs indiquent que l'augmentation du niveau d'éducation moyen, la diminution du taux de mariage et la diminution du nombre d'enfants moyen auraient tous contribué à faire baisser le taux d'entrepreneuriat entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. Cela n'empêche pas la partie de la décomposition représentant l'effet du changement des coefficients, c'est-à-dire le changement du taux d'entrepreneuriat une fois les caractéristiques sociodémographiques prises en compte, d'afficher une valeur positive. Ainsi, même si le taux d'entrepreneuriat des immigrants de deuxième génération est plus faible en 2006 qu'il ne l'était pour leurs parents en 1981, les immigrants de 2<sup>e</sup> génération ne seraient pas moins enclins à devenir entrepreneur que leurs parents, toutes choses étant égales par ailleurs.

Il est à noter toutefois que les auteurs interprètent leurs résultats sans prêter attention à l'évolution des coefficients associés aux groupes ethniques. En d'autres termes, ces derniers ne s'intéressent pas à l'effet de l'ethnicité sur la propension à devenir entrepreneur et les résultats de leur étude ne permettent donc pas de fournir une explication au fait que les taux d'entrepreneuriat des différentes communautés ethniques ne convergent pas à travers le temps. De plus, le fait de ne pas s'intéresser à l'effet de l'ethnicité ne permet pas d'explorer une partie importante des facteurs responsables de l'impact du changement de génération sur la propension naturelle à devenir

---

<sup>5</sup> Nous expliquons en détail cette méthode dans la section *Données et Statistiques Descriptives*.

entrepreneur. Par exemple, si les immigrants de 2<sup>e</sup> génération sont naturellement plus enclins que les immigrants de 1<sup>ère</sup> génération à devenir entrepreneur, est-ce en partie parce qu'ils sont autant discriminés que leurs parents sur le marché du travail, ou encore est-ce parce qu'ils sont moins discriminés que ces derniers sur le marché des produits et services ? Bien qu'il serait possible d'estimer le niveau de discrimination dont sont victimes l'ensemble des immigrants dans chaque génération, notamment à partir de décompositions d'Oaxaca-Blinder sur des équations de revenu estimées de manière séparée pour les immigrants et pour les citoyens d'origine, les résultats obtenus risqueraient selon nous de ne pas être très révélateurs puisque nous supposons que les immigrants subissent les effets de la discrimination non pas seulement en raison de leur statut d'immigrant, mais bien surtout en raison de leur ethnicité. Outre la discrimination, il existe d'autres facteurs qui peuvent expliquer l'évolution intergénérationnelle de la propension naturelle à devenir entrepreneur et qui sont propres à chaque communauté ethnique. C'est le cas notamment de l'existence d'enclaves ethniques (Edin & al., 2003, Borjas, 1986), de réseaux d'affaires propres à chaque communauté (Mandorff & Kerr, 2015, Edin & al., 2003) et de facteurs purement culturels. Cependant, même si les auteurs décidaient de prêter attention à la valeur des coefficients ethniques, le potentiel d'analyse resterait limité étant donné que les échantillons sur lesquels ils effectuent les régressions Probit excluent les Canadiens de souche comme groupe de référence.

Ainsi, contrairement à ce qui est fait dans cet article, nous portons une attention toute particulière sur l'évolution de l'effet partiel de l'ethnicité sur la probabilité d'être entrepreneur entre la 1<sup>ère</sup> génération et la 2<sup>e</sup> génération d'immigrants. En procédant ainsi, nous étudions non seulement les raisons pour lesquelles les taux d'entrepreneuriat des différentes communautés ethniques ne convergent pas à travers le temps, mais nous arrivons également à mieux saisir pour quelles raisons, de manière agrégée, les immigrants sont plus ou moins enclins à se lancer en affaires au fil des générations.

#### II.V. La transmission intergénérationnelle des compétences entrepreneuriales

Dans leur étude, Dunn & Holtz-Eakin (2000) tentent de trouver une explication au fait que les enfants de parents entrepreneurs affichent une plus grande propension à devenir eux aussi

entrepreneurs. Leur analyse explore deux mécanismes possibles, le premier étant qu'en transférant une partie de leur avoir financier à leurs enfants, les parents entrepreneurs aident ces derniers à se lancer en affaires en réduisant les barrières à l'emprunt. Le deuxième mécanisme, quant à lui, stipule qu'en transférant à leurs enfants une partie de leur expertise managériale et entrepreneuriale, les parents entrepreneurs confèrent à leurs enfants un avantage compétitif en ce qui concerne le démarrage d'entreprise.

La méthode qu'ils utilisent est celle de régressions Logit estimées sur des données de panel tirées du NLS (*National Longitudinal Surveys of Labor Market Experience*). Le modèle estimé est de la même nature que la plupart des modèles utilisés dans la littérature, la seule différence étant que la variable dépendante observée correspond à la transition vers l'entrepreneuriat à partir d'un emploi salarié entre la période  $t$  et  $t+1$  (codée 1 s'il y a transition et 0 s'il y a absence de transition) plutôt que de correspondre à une observation statique du statut d'entrepreneur. Cette spécification est utilisée afin d'éliminer la relation de simultanéité entre le statut d'entrepreneur au temps  $t$  et la valeur des actifs financiers détenus au temps  $t$ . En plus de comprendre les variables sociodémographiques généralement utilisées dans la littérature, le modèle inclut comme variables explicatives l'avoir financier de l'individu interrogé et des parents de l'individu interrogé ainsi qu'une série de variables dichotomiques indiquant si l'un des deux parents, le père, la mère ou encore les deux parents étaient entrepreneurs. Ces variables sont incluses afin de valider ou d'invalider les mécanismes hypothétiques mentionnés dans le dernier paragraphe.

Les auteurs obtiennent comme résultat que parmi toutes les nouvelles variables qu'ils ont incluses dans le but de tester leurs hypothèses, le statut d'entrepreneur des parents est de loin celle qui affiche le plus grand pouvoir explicatif. Puis, en ajoutant aux régressions de base des variables faisant interagir le statut d'entrepreneur avec différentes mesures du niveau de succès de l'entreprise détenue, ils concluent que l'effet d'avoir un parent entrepreneur sur la propension à devenir soi-même entrepreneur est plus grand lorsque l'entreprise appartenant au parent concerné connaît du succès. Pour les auteurs, cela signifie que la corrélation entre l'entrepreneuriat parental et la propension des enfants à devenir entrepreneur dépend de la transmission de compétences managériales, et non pas de la transmission de simples préférences pour l'entrepreneuriat. De



plus, les auteurs s'appuient sur des statistiques descriptives pour montrer que cette corrélation n'est pas entièrement attribuable au lègue d'entreprises familiales.

L'importance de cet article s'inscrit dans la nécessité de tenir compte de la possibilité que la persistance de l'effet de l'ethnicité au-delà de la première génération d'immigrants soit en partie biaisée à la hausse en raison du mécanisme de transmission intergénérationnelle de la propension à devenir entrepreneur qui s'opère au sein d'une même famille. En effet, suivant la logique de ce mécanisme, un groupe ethnique dont la prévalence de l'entrepreneuriat est grande au sein de la première génération d'immigrants aura naturellement tendance à afficher une proportion plus élevée d'entrepreneurs dans sa deuxième génération, toutes choses étant égales par ailleurs. Par conséquent, nous incluons dans les équations de choix sectoriel une variable qui approxime la probabilité d'avoir un père entrepreneur. En procédant ainsi, nous tentons d'établir une séparation entre le facteur de transmission intergénérationnelle des compétences managériales et les facteurs culturels, discriminatoires et autres qui expliquent l'effet de chaque groupe ethnique sur la propension des immigrants de deuxième génération à devenir entrepreneur. Nous sommes contraints d'utiliser une probabilité à défaut d'avoir accès aux données sur le type de travail exercé par les parents. Les détails entourant la façon dont est calculée cette probabilité sont présentés dans la section V.

### III. Données et Statistiques Descriptives

#### III.I Sources

Les données utilisées dans ce travail de recherche sont tirées de deux sources : le recensement américain de 1980 (*1980 United States Census*) ainsi que les enquêtes démographiques courantes (*Current Population Survey*) conduites par le Bureau Américain du Recensement durant la période qui s'est étalée de 2006 à 2015. Les données du recensement de 1980 proviennent de son plus gros échantillon, soit celui couvrant 5% de la population américaine (*1980 5% State Sample*). Nous utilisons ces données pour étudier la 1<sup>ère</sup> génération d'immigrants. Pour la 2<sup>e</sup> génération d'immigrants, nous mettons en commun les dix suppléments économiques et sociaux annuels (*Annual Social and Economic Supplement*) dérivés des enquêtes démographiques courantes effectuées de 2006 à 2015. L'objectif de mettre en commun plusieurs suppléments économiques et sociaux annuels est d'obtenir un échantillon de taille suffisamment grande pour que nous puissions étudier la 2<sup>e</sup> génération d'immigrants avec autant de précision que la 1<sup>ère</sup> génération d'immigrants. Cette méthode est entre autres celle qu'utilisent Card & al. (2000)<sup>6</sup>. À défaut d'établir un lien direct entre les parents de la 1<sup>ère</sup> génération d'immigrants et leurs enfants, les données du bureau de recensement possèdent l'avantage de nous permettre de travailler avec de grands échantillons qui augmentent la précision des estimations auxquelles nous procédons.

#### III.II Méthode d'échantillonnage

L'absence de lien direct entre les immigrants de 1<sup>ère</sup> génération et les enfants de ces derniers nous force à employer la méthode des cohortes synthétiques. Cette méthode, largement utilisée dans la littérature sur la mobilité sociale des immigrants [Card & al., 2000 et Aydemir & al., 2009], consiste à bâtir deux échantillons distincts, soit un pour la 1<sup>ère</sup> génération dans lequel les individus sont contraints d'être parents et un pour la 2<sup>e</sup> génération, où chaque individu, de par ses

---

<sup>6</sup> Tout comme Card & al. (2000), nous ne tenons pas compte du fait que certains ménages sont interrogés chaque année pour plusieurs années consécutives. Selon les auteurs, les écarts-types que nous calculons à partir des données CPS sont alors sous-estimés par une marge d'environ 17%.

caractéristiques, doit être l'enfant plausible d'un de ces derniers. En raison de la très faible proportion d'entrepreneurs de sexe féminin dans la 1<sup>ère</sup> génération d'immigrants sondée en 1980 et de la faiblesse relative des mécanismes de transmission de type père-fille recensée dans la littérature (Aydemir & al., 2009), nos cohortes synthétiques sont restreintes de façon à ce que notre analyse comparative ne porte que sur les liens de type père-fils.

Pour la 1<sup>ère</sup> génération, notre échantillon est constitué de 592 374 individus questionnés dans le cadre du recensement de 1980. Tous les individus de l'échantillon sont des hommes nés à l'étranger, qui habitent avec une conjointe née dans le même pays qu'eux et dont l'enfant le plus vieux est âgé entre 0 et 15 ans. Nous avons choisi l'intervalle allant de 0 à 15 ans afin de nous assurer que les enfants les plus vieux n'approchent pas trop l'âge de la retraite dans l'échantillon correspondant aux individus de 2<sup>e</sup> génération. De plus, afin de stabiliser les estimés obtenus, nous restreignons l'échantillon aux hommes âgés entre 25 et 65 ans, soit une tranche d'âge qui s'apparente à la période d'activité moyenne chez un individu de sexe masculin. Pour la 2<sup>e</sup> génération, notre échantillon est composé de 239 101 individus de sexe masculin, tous âgés de 26 à 50 ans afin de pouvoir être considérés comme des enfants plausibles des hommes de la 1<sup>ère</sup> génération. De plus, les parents de ces individus doivent tous deux être nés dans le même pays. Par ailleurs, certaines des restrictions que nous imposons sont communes aux deux générations. Premièrement, tel qu'il est d'usage dans la littérature sur l'entrepreneuriat, nous éliminons des deux échantillons tous les individus qui travaillent dans l'industrie agricole. Deuxièmement, nous éliminons tous les individus qui se disent entrepreneur et dont le revenu d'entreprise déclaré est négatif, de même que tous les individus qui se disent travailleur salarié et dont le salaire déclaré est négatif. Finalement, nous ne gardons que les individus qui ont un emploi et éliminons tous les individus appartenant à des groupes ethniques pour lesquels il y a moins de 30 entrepreneurs dans l'une ou l'autre des deux générations. En procédant ainsi, nous évitons d'obtenir des estimés potentiellement biaisés de l'effet de certains groupes ethniques sur la décision de devenir entrepreneur.

La manière dont nous bâtissons les cohortes synthétiques nous amène à définir l'origine ethnique des immigrants de 2<sup>e</sup> génération de manière beaucoup plus restrictive que ce qui est généralement fait dans la littérature. En effet, alors que la plupart des études définissent l'origine ethnique d'un

immigrant de 2<sup>e</sup> génération à partir du pays de naissance d'un seul de ses parents (Aydemir & al., 2009) ou encore à partir de la réponse que celui-ci fournit à la question sur l'ascendance (Fairlie & Meyer, 1996), nous optons plutôt pour une méthode qui exige que les deux parents soient nés dans le même pays, de manière à exclure tous les individus dont l'origine ethnique peut être qualifiée de mixte. De la même façon, le groupe de référence que nous utilisons pour la 2<sup>e</sup> génération de travailleurs correspond à l'ensemble des individus dont les deux parents sont nés aux États-Unis. Ce groupe de travailleurs correspond à vrai dire à l'ensemble des hommes de 3<sup>e</sup> génération ou plus présents dans les données de recensement des années 2006 à 2015 et qui satisfont les critères d'échantillonnage énoncés plus haut.

Il existe deux principaux avantages liés à cette méthode. Dans un premier temps, les résultats de notre analyse ne risquent pas d'être influencés par les caractéristiques d'individus ayant potentiellement mieux intégré le marché du travail américain en raison du fait qu'un de leurs parents était né aux États-Unis. Dans un deuxième temps, les groupes ethniques que nous définissons dans notre analyse ne risquent pas de contenir des membres qui les représentent de manière imparfaite en raison de leur origine mixte. Il se peut que les études qui définissent l'ethnicité de manière peu rigoureuse le fassent afin de ne pas éliminer trop d'observations de leur échantillon. Or, nous préférons réduire la précision de certaines de nos estimations plutôt que d'obtenir des estimations à coup sûr biaisées, dont l'interprétation est difficile à effectuer.

### III.III Distribution de l'entrepreneuriat

Puisqu'un des objectifs de ce mémoire est de mesurer l'évolution de l'effet d'appartenir à un groupe ethnique sur la probabilité d'être entrepreneur entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants, il peut être pertinent de débiter notre analyse en observant l'évolution des taux d'entrepreneuriat associés aux différents groupes ethniques sur cette même période. Pour ce faire, nous calculons le taux d'entrepreneuriat associé à chaque groupe en divisant le nombre total de travailleurs autonomes par le nombre total d'individus qui occupent un emploi dans le groupe en question. Cette méthode de calcul correspond à la méthode conventionnellement utilisée dans la

littérature scientifique, ainsi que dans les publications émises par des organisations d'études économiques telles que l'OCDE ou la Banque Mondiale. Également, il est important de mentionner que comme pour toutes les statistiques descriptives que nous calculons dans les sous-sections qui suivent, nous faisons l'usage des poids d'échantillonnage fournis par le Bureau Américain du Recensement afin d'obtenir des moyennes représentatives de la population d'intérêt. Ces poids varient d'un individu à l'autre et correspondent à l'inverse de la probabilité qu'un individu soit inclus dans l'échantillon.

Les deux premières colonnes du tableau 1 contiennent les taux d'entrepreneuriat observés dans chaque groupe ethnique lors de la 1<sup>ère</sup> et de la 2<sup>e</sup> génération, respectivement. En accord avec ce qui est relaté dans la littérature, nous y remarquons que les taux d'entrepreneuriat diffèrent énormément d'un groupe à l'autre, et ce, à la fois au sein de la 1<sup>ère</sup> et de la 2<sup>e</sup> génération. Cette constatation soulève deux questions. Dans un premier temps, nous aimerions savoir si la dispersion du taux d'entrepreneuriat est davantage élevée dans la 1<sup>ère</sup> génération que dans la 2<sup>e</sup> génération. Dans un deuxième temps, dans l'éventualité où nous observerions une atténuation de la dispersion entre les deux générations, nous aimerions savoir si cette atténuation traduit une convergence vers le taux d'entrepreneuriat qui prévaut dans la population native.

Les graphiques 1 et 2 nous aident à répondre à ces questions. Le graphique 1 illustre simplement le taux d'entrepreneuriat dans la 2<sup>e</sup> génération en fonction du taux d'entrepreneuriat dans la 1<sup>ère</sup> génération. La droite pointillée qui le traverse constitue sa bissectrice, qui représente le tracé que suivraient l'ensemble de ses points si tous les groupes ethniques affichaient le même taux d'entrepreneuriat dans chacune des générations. Il est à noter que tous les groupes sont représentés dans le graphique, excepté les natifs. La conclusion que nous tirons de ce graphique est qu'il semble y avoir une certaine diminution de la dispersion du taux d'entrepreneuriat entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. En d'autres mots, un phénomène de régression vers la moyenne semble se produire. Une manière facile de s'en convaincre est d'observer la position des points par rapport à la bissectrice. En procédant à l'exercice, on remarque que les points situés sur la portion droite de l'axe des abscisses ont tendance à se retrouver sous la bissectrice, alors que les points situés plus à gauche ont tendance à se retrouver au-dessus de la bissectrice. De manière plus

simple, on remarque aussi que la dispersion verticale semble plus petite que la dispersion horizontale des points.

Pour répondre à la deuxième question, il faut se tourner vers le graphique 2. Ce graphique illustre l'évolution intergénérationnelle de l'écart entre le taux d'entrepreneuriat de chaque groupe ethnique et le taux d'entrepreneuriat des natifs (nous nommons cet écart *dévi*ation sur le graphique). La droite pointillée qui scinde le graphique en deux représente encore une fois la bissectrice alors que l'autre droite représente la droite des moindres carrés ordinaires qui passe par le nuage de points présent sur le graphique. Très vite, on remarque que le nuage de points possède exactement la même forme que sur le graphique 1, ce qui est tout à fait normal étant donné que le fait de soustraire une constante à une variable  $x$  et à une variable  $y$  ne change en rien la covariance entre ces deux variables, de même que la variance de chacune d'entre elles. Dans ce contexte, le coefficient de la droite des moindres carrés ordinaires illustrée sur le graphique 2 valide l'existence du phénomène de régression vers la moyenne étant donné que celui-ci est inférieur à 1 (0.49).

Maintenant, pour savoir si cette régression vers la moyenne se produit autour du taux d'entrepreneuriat qui prévaut chez les natifs, il faut porter notre attention non pas sur la pente de cette droite, mais bien sur son ordonnée à l'origine. Si cette dernière possède une valeur située aux alentours de 0, nous pouvons conclure à une régression parfaite vers le taux observé chez les natifs, car tous les taux d'entrepreneuriat seraient alors moins éloignés du taux observé chez les natifs dans la 2<sup>e</sup> génération que dans la 1<sup>ère</sup> génération. Par opposition, si l'ordonnée à l'origine se trouve au-dessus ou en-dessous de 0, nous pouvons conclure à une homogénéisation des taux d'entrepreneuriat, mais à un niveau ou bien inférieur, ou bien supérieur au taux d'entrepreneuriat des natifs. Sur le graphique 2, nous voyons que l'ordonnée à l'origine de la droite est située tout juste au-dessus de 0, plus exactement à un taux de 1,49%. Ainsi, bien que nous ne puissions pas conclure à une convergence parfaite vers le taux d'entrepreneuriat de la population native, nous observons tout de même une tendance forte qui va dans cette direction.

Cette tendance est en quelque sorte reflétée sur le graphique 3, où nous voyons que l'écart entre le taux d'entrepreneuriat observé chez l'ensemble des immigrants et celui observé chez les natifs

est plus petit dans la 2<sup>e</sup> génération que dans la 1<sup>ère</sup> génération. Il est à noter que l'écart négatif entre les immigrants et les natifs observé sur ce graphique, alors même que l'ordonnée à l'origine de la droite des moindres carrés ordinaires qui traverse le graphique 2 est supérieure à 0, est attribuable au fait que les groupes ethniques dont le taux d'entrepreneuriat est le plus faible possèdent un poids démographique en moyenne plus important. La question qui demeure et qui fait l'objet d'une partie importante de notre analyse économétrique est de savoir si la convergence partielle vers le taux d'entrepreneuriat qui caractérise la population native est le fruit d'une convergence des distributions des caractéristiques individuelles qui influencent la décision de devenir entrepreneur, ou s'il s'agit plutôt d'une conséquence de l'atténuation de facteurs propres à chaque groupe ethnique.

#### III.IV Distribution des caractéristiques sociodémographiques

Nous refaisons dans cette section le même type d'analyse que pour le taux d'entrepreneuriat, car tel que nous venons de le mentionner dans le dernier paragraphe, il se pourrait que la convergence partielle des taux d'entrepreneuriat ethniques vers le taux de la population d'origine résulte en partie d'une convergence en ce qui concerne la distribution des caractéristiques individuelles. Nous portons notre attention sur trois variables individuelles à caractère sociodémographique : le niveau d'éducation, le nombre d'enfants ainsi que le statut marital. Notre choix est motivé par le fait que dans la littérature, ces variables sont constamment incluses dans les modèles régressifs et qu'elles sont constamment parmi celles qui expliquent le mieux la décision de devenir entrepreneur (Hou & al. , 2013, Holtz-Eakin, 2000, Fairlie & Meyer, 1996).

Nous débutons notre analyse avec le niveau d'éducation, que nous définissons en termes de nombres d'années. Nous calculons ce nombre d'années en soustrayant 5 années (nous prenons pour acquis que l'individu moyen débute son éducation à 5 ans) à l'âge auquel l'individu complète son éducation. Quant à cet âge, il est standardisé pour tous les individus en fonction du type de diplôme obtenu. Pour chaque groupe ethnique, les niveaux d'éducation moyens dans la 1<sup>ère</sup> et la 2<sup>e</sup> génération sont affichés dans les colonnes 3 et 4 du tableau 1, respectivement. Une fois de plus, on constate qu'à la fois dans la 1<sup>ère</sup> et dans la 2<sup>e</sup> génération, il existe des disparités

importantes entre les groupes ethniques. Certains des écarts observés dans la 1<sup>ère</sup> génération sont particulièrement saisissants, tel que celui entre les Indiens, qui possèdent en moyenne 19,1 années d'éducation dans la 1<sup>ère</sup> génération, et la communauté portugaise, dont les membres de la 1<sup>ère</sup> génération n'ont que 7,1 années d'éducation en moyenne. Évidemment, ces statistiques datent de 1980, et les écarts que nous observerions chez les immigrants de 1<sup>ère</sup> génération de l'époque actuelle seraient probablement beaucoup moins élevés.

Pareillement à ce que le graphique 1 faisait pour le taux d'entrepreneuriat, le graphique 4 illustre l'évolution du nombre d'années d'éducation moyen pour chaque groupe ethnique entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. En y comparant les étalements vertical et horizontal des points, on remarque que la dispersion du niveau moyen d'éducation semble beaucoup plus élevée dans la 1<sup>ère</sup> génération que dans la 2<sup>e</sup> génération. Ainsi, comme c'était le cas pour le taux d'entrepreneuriat, nous sommes en présence d'un processus de régression vers la moyenne. Cependant, contrairement à ce que nous observions avec le taux d'entrepreneuriat, la moyenne du niveau d'éducation se déplace de manière considérable d'une génération à l'autre, comme l'illustre le fait que la majorité des points se situent nettement au-dessus de la bissectrice du graphique. Les points les plus éloignés de la bissectrice illustrent la force de la hausse du niveau d'éducation qu'ont connu certaines communautés qui, lors de la 1<sup>ère</sup> génération, étaient situées dans le bas de la distribution. C'est notamment le cas pour les communautés portugaise et italienne, qui avec un bond du nombre d'années moyen d'éducation de 7,2 années et 6,1 années respectivement, ont subi un véritable rattrapage éducationnel ayant même permis aux hommes de leur 2<sup>e</sup> génération de se hisser au-dessus de la moyenne pour l'ensemble des immigrants, chiffrée à 14,02 années. La hausse du niveau d'éducation moyen chez la majorité des groupes ethniques est en partie reflétée dans le graphique 6, où l'on voit très bien que le niveau d'éducation moyen de la population immigrante augmente entre la 1<sup>ère</sup> génération et la 2<sup>e</sup> génération. La même tendance s'opère chez la population native, qui voit son niveau d'éducation augmenter dans des proportions similaires.

La question est maintenant de savoir si la régression vers la moyenne du niveau d'éducation entre les groupes ethniques converge vers le niveau d'éducation moyen de la population native dans la 2<sup>e</sup> génération. Pour cela, il faut entre autres que la hausse généralisée du niveau d'éducation qu'ont subie les immigrants entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération parvienne à suivre la hausse qu'ont



connue les natifs sur cette même période. Le graphique 5 nous permet de répondre à cette question en illustrant l'évolution intergénérationnelle de l'écart entre le niveau d'éducation de chaque groupe ethnique et celui de la population native. Encore une fois, nous nommons cet écart *déviatio*n sur le graphique. Nous constatons sur ce graphique que l'ordonnée à l'origine est située très près de 0, à 0,84 années plus précisément. Ceci nous indique que la convergence s'effectue bel et bien *vers* le niveau d'entrepreneuriat qui prévaut dans la population native. De plus, le coefficient de 0,32 associé à la droite des moindres carrés ordinaires qui passe par le nuage de points témoigne de la force - bien plus élevée que celle observée dans le contexte du taux d'entrepreneuriat - du processus de régression vers la moyenne dont il était question dans le dernier paragraphe. Ainsi, grâce à ces deux éléments, il est possible de dire que le niveau d'éducation des différentes communautés ethniques converge fortement vers celui des natifs entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération, et ce, dans des proportions supérieures à la convergence observée pour le taux d'entrepreneuriat.

À la lumière des constatations que nous avons faites sur l'évolution de la distribution du niveau d'éducation entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération, il ne serait pas surprenant que l'éducation soit une variable qui explique une partie importante de la convergence partielle des taux d'entrepreneuriat dont nous avons discuté dans la dernière sous-section. Cependant, ce pouvoir explicatif est conditionnel au fait que l'effet de l'éducation sur la propension à devenir entrepreneur soit sensiblement le même dans la 1<sup>ère</sup> et dans la 2<sup>e</sup> génération, ou du moins qu'il soit du même signe. Alors que toutes les études relatent un effet nettement positif de l'éducation sur la propension qui nous intéresse, il faut se rappeler que presque aucune de ces études ne fait la distinction entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération dans leurs échantillons. Ainsi, il faut attendre les résultats de nos régressions à la section V pour savoir si la convergence des niveaux d'éducation explique bel et bien une partie de la convergence des taux d'entrepreneuriat que nous observons. Entre temps, il est pertinent de noter que l'étude de Hou & al. (2013), qui porte sur le Canada, conclut que, chez les hommes, la hausse du niveau d'éducation des immigrants entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération a eu un impact considérable sur la diminution de l'écart de taux d'entrepreneuriat entre les immigrants et les natifs sur cette même période.

Nous poursuivons notre analyse avec comme variable le nombre d'enfants que possède chaque individu. Les différentes valeurs que prend cette variable en fonction du groupe ethnique et de la génération sont affichées dans les colonnes 7 et 8 du tableau 1. Une fois de plus, on constate la présence de disparités entre les groupes ethniques, à la fois dans la 1<sup>ère</sup> et la 2<sup>e</sup> génération. Or, ces disparités semblent en moyenne de taille inférieure aux disparités que nous observions dans le cas de l'éducation et du taux d'entrepreneuriat, ce qui est probablement lié à la nature même de la variable. Dans tous les cas, l'élément clé demeure la taille relative des disparités d'une génération à l'autre.

Le graphique 7 nous éclaire à cet effet en nous montrant l'évolution intergénérationnelle du nombre d'enfants moyen pour chaque groupe ethnique. On y constate que la dispersion du nombre d'enfants sur l'axe correspondant à la 1<sup>ère</sup> génération semble plus élevée que sur l'axe correspondant à la 2<sup>e</sup> génération, et pour cette raison, nous pouvons affirmer qu'il existe un certain processus de régression vers la moyenne entre les deux générations. Toutefois, la force de ce processus ne semble pas aussi élevée que dans le cas du niveau d'éducation ou encore du taux d'entrepreneuriat. L'aspect le plus flagrant de ce graphique est de loin la diminution du nombre d'enfants moyen chez les immigrants entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. En effet, nous voyons que tous les points du graphique sont situés bien en-dessous de la bissectrice. Le graphique 9 montre que cette tendance n'est pas unique aux immigrants et qu'elle affecte également la population native dans des proportions très similaires. À vrai dire, cette tendance est le reflet de la diminution importante du taux de natalité aux États-Unis depuis plusieurs décennies<sup>7</sup>.

Afin de savoir si la régression vers la moyenne donne lieu à une convergence entre le nombre d'enfants moyen associé à chaque groupe ethnique et le nombre d'enfants moyen dans la population native, nous nous tournons vers le graphique 8. Le coefficient de la droite des moindres carrés ordinaires représentée sur ce graphique est de 0,48, ce qui, a priori, illustre une régression vers la moyenne du même ordre de grandeur que celle observée dans le cas du taux d'entrepreneuriat. Cependant, lorsque nous observons le graphique de plus près, nous remarquons que la pente de cette droite semble fortement tirée vers le haut par les deux points situés le plus à

---

<sup>7</sup> Entre 1960 et 2006, le taux de naissance brut aux États-Unis est passé de 23,7 enfants par 1000 personnes à 14,3 enfants par 1000 personnes. Source : Banque Mondiale, 2016.

droite sur l'axe des abscisses. Si l'on retire de l'échantillon les deux groupes ethniques qui correspondent à ces points, la pente de la droite fait un bond considérable et atteint 0,78, valeur à partir de laquelle nous ne concluons qu'à une faible régression vers la moyenne. Toutefois, peu importe l'échantillon utilisé, le nombre d'enfants moyen dans les différents groupes ethniques semble vouloir converger vers le nombre d'enfants moyen qui caractérise la population native dans la 2<sup>e</sup> génération, les ordonnées à l'origine étant de -0,042 et de -0,019 enfants lorsque nous gardons et lorsque nous enlevons les données aberrantes, respectivement.

Ainsi, il est possible d'affirmer qu'entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération, le nombre d'enfants moyen par groupe ethnique converge vers le nombre d'enfants moyen observé dans la population native, quoique de manière faible. En effet, la magnitude de cette convergence n'est pas équivalente à celle de la convergence des taux d'entrepreneuriat qu'elle serait sensée expliquer, et encore moins à celle de la convergence que nous observions dans le cas du niveau d'éducation.

En ce qui a trait au statut marital, la manière dont nous parvenons à comparer les groupes ethniques par rapport à cette variable est en calculant le taux de mariage (chez la population masculine) propre à chaque groupe ethnique. Pour ce faire, nous divisons tout simplement le nombre d'individus mariés par le nombre total d'individus appartenant à chaque groupe ethnique. Bien entendu, nous faisons un calcul séparé pour les hommes de la 1<sup>ère</sup> et de la 2<sup>e</sup> génération. Tel que nous pouvons le voir dans le tableau 1, le taux de mariage dans la 1<sup>ère</sup> génération est de 100% pour la majorité des groupes ethniques. En effet, seuls les natifs, les Mexicains, les Cubains et les Antillais affichent un taux légèrement inférieur à 100%, une situation probablement attribuable au fait qu'il s'agit des sous-échantillons les plus grands dont nous disposons. Bien que cette situation puisse paraître extrême, elle n'entre pas en conflit avec ce que l'on retrouve dans la littérature. À titre d'exemple, dans leur article, Hou & al. (2013), qui emploient la méthode des cohortes synthétiques sur des données canadiennes, estiment à 98,6% le taux de mariage chez les pères de famille considérés comme des immigrants de 1<sup>ère</sup> génération en 1981. De plus, il existe deux facteurs pouvant faire en sorte que les taux d'entrepreneuriat que nous estimons sont encore plus élevés que ceux qui sont rapportés dans certains des articles de référence. En premier lieu, puisque nous décomposons la population d'immigrants de 1<sup>ère</sup> génération en plusieurs groupes ethniques et que les sous-échantillons rattachés à ces groupes sont parfois de taille assez petite, il

n'est pas surprenant que beaucoup de groupes affichent un taux de mariage de 100%, considérant que chaque père de famille a beaucoup plus de chances d'être marié que de faire partie d'une union libre. En deuxième lieu, contrairement aux autres études qui font usage de la méthode des cohortes synthétiques, notre échantillon est restreint aux pères de famille dont l'épouse vient du même pays. Par conséquent, notre échantillon risque de représenter un segment plus conservateur de la population dans lequel les individus sont potentiellement davantage susceptibles d'être mariés.

Étant donné que les taux de mariage sont presque tous les mêmes dans la 1<sup>ère</sup> génération et qu'ils varient dans la 2<sup>e</sup> génération, ils ne peuvent faire autrement que de diverger entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. Or, bien qu'ils suivent un processus de divergence, ils subissent néanmoins tous une diminution importante. Il est à noter que même si, sur la période étudiée, la baisse généralisée du taux de mariage aux États-Unis ne s'était pas produite, nous observerions tout de même une diminution importante des taux de mariage étant donné que les hommes présents dans l'échantillon de 2<sup>e</sup> génération n'ont pas tous des enfants et ne vivent pas tous avec une conjointe, contrairement à leur père. Le graphique 10 illustre le taux de mariage associé à chaque groupe ethnique dans la 2<sup>e</sup> génération. La ligne verticale qui traverse le graphique représente le taux de mariage en vigueur dans la population native. On remarque que la distance moyenne des points par rapport à la ligne verticale est plus ou moins la même de part et d'autre de celle-ci. Cependant, la majorité des groupes ethniques semblent posséder un taux de mariage inférieur à celui des natifs. Puisque la littérature recense généralement un effet positif du fait d'être marié sur la propension à devenir entrepreneur, cela signifie que dans la 2<sup>e</sup> génération, beaucoup de groupes ethniques devraient voir leur taux d'entrepreneuriat être négativement influencé par rapport à celui des natifs en raison de leur taux de mariage. Ce phénomène tend à agrandir plutôt qu'à rétrécir les écarts de taux d'entrepreneuriat entre les groupes ethniques et les natifs sur la période s'étalant de la 1<sup>ère</sup> à la 2<sup>e</sup> génération, à moins que les groupes dont le taux d'entrepreneuriat était le plus élevé dans la 1<sup>ère</sup> génération soient en moyenne ceux dont le taux de mariage est le plus faible dans la 2<sup>e</sup> génération, et vice-versa. N'ayant trouvé aucune évidence plausible de l'existence de ce phénomène, nous rejetons cette hypothèse et concluons plutôt que la divergence des taux de mariage entre les deux générations a surtout contribué à accroître les écarts de taux d'entrepreneuriat entre les natifs et les différents groupes ethniques. Dans ce

contexte, il devient de plus en plus probable qu'au-delà de l'évolution de facteurs sociodémographiques tels que le niveau d'éducation et le nombre d'enfants, l'évolution de facteurs propres à chaque communauté ait un effet sur la convergence des taux d'entrepreneuriat, sans quoi la convergence partielle que nous observons entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération n'aurait probablement pas lieu. En d'autres mots, l'existence de l'effet de l'ethnicité, de même que son atténuation entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération, semble fort probable. Tel que précédemment annoncé, il s'agit de la question centrale de ce mémoire, que nous analysons en détail à partir de la section IV.

### III.V Distribution des salaires et des revenus d'entreprise

Nous calculons le salaire et le revenu d'entreprise des individus faisant partie de nos échantillons directement à partir des données tirées du recensement de 1980 et des enquêtes démographiques courantes (CPS). Grâce à la compatibilité des données de recensement avec les données CPS, nous utilisons les mêmes variables à la fois pour la 1<sup>ère</sup> génération et la 2<sup>e</sup> génération d'immigrants, soit le revenu d'emploi (salaire) et le revenu d'entreprise (dividendes). Pour les individus qui se déclarent salariés, nous considérons que leur revenu est tout simplement égal au revenu d'emploi qui apparaît dans les données. Par opposition, pour les individus qui disent travailler à leur propre compte, nous sommes contraints d'additionner le revenu d'emploi et le revenu d'entreprise, car dans de nombreux cas, les entrepreneurs dérivent une partie ou encore la totalité de leur revenu d'un salaire qu'ils se versent eux-mêmes. Il est à noter qu'afin de rendre toute comparaison intergénérationnelle possible, nous avons converti tous les revenus présents dans nos données en dollars de 2006 à l'aide des données sur l'inflation fournies par le *Bureau of Labor Statistics*. Le tableau 2 affiche le salaire moyen et le revenu d'entreprise moyen en fonction du groupe ethnique et de la génération.

L'objectif de cette sous-section est double. Dans un premier temps, nous comparons la distribution du revenu d'entreprise avec la distribution du salaire pour chacun des quatre grands groupes sur lesquels notre étude porte: les immigrants de 1<sup>ère</sup> génération, les natifs de 1<sup>ère</sup> génération, les immigrants de 2<sup>e</sup> génération et les natifs de 2<sup>e</sup> génération. L'objectif de cette

comparaison est de déterminer la valeur du rapport entre la variance de l'habileté dans l'entrepreneuriat et la variance de l'habileté dans le secteur salarié. Dans le modèle de sélection présenté à la section II.I, la valeur de ce paramètre est représentée par l'expression  $\frac{\sigma_E}{\sigma_S}$ . L'importance d'accoler une valeur à ce paramètre est directement liée au modèle de sélection que nous utilisons dans notre méthodologie pour tenter d'expliquer la provenance de l'effet de l'ethnicité sur la propension à devenir entrepreneur. En effet, tel que nous le verrons, avoir une idée de la valeur du rapport des variances est crucial afin de poser les bonnes conditions initiales du modèle. Dans un deuxième temps, nous analysons comment le différentiel de revenu entre l'entrepreneuriat et le travail salarié évolue par rapport au taux d'entrepreneuriat entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. En d'autres mots, nous souhaitons voir si, pour chacune des deux générations, le taux d'entrepreneuriat associé à chaque groupe ethnique est corrélé avec le différentiel de revenu qui caractérise chaque groupe. Cette vérification empirique nous permet ensuite de savoir s'il existe un lien potentiel entre l'effet d'appartenir à un groupe ethnique sur la propension à devenir entrepreneur et le différentiel de revenu associé au groupe ethnique en question, et donc de mieux orienter nos démarches empiriques.

Pour comparer la variance de l'habileté dans l'entrepreneuriat avec la variance de l'habileté dans le secteur salarié, nous commençons par comparer les distributions du revenu associées à chacun des deux secteurs. Le graphique 11 accomplit cette tâche. On y constate que pour tous les échantillons dont nous disposons, la variance du revenu d'entreprise est largement supérieure à la variance du salaire. En d'autres mots, à la fois pour les natifs et les immigrants, l'écart entre la variance du revenu d'entreprise et la variance du salaire est nettement positif, et ce, peu importe la génération. Or, il faut faire attention de ne pas conclure que les écarts de variance illustrés sur ce graphique sont uniquement le produit d'écarts en ce qui concerne la variance de l'habileté. À vrai dire, la variance du revenu dans chacun des quatre échantillons est le produit de trois facteurs distincts : la variance des caractéristiques sociodémographiques, les rendements sur les caractéristiques sociodémographiques et la variance de l'habileté<sup>8</sup>. Notre objectif est bien entendu d'isoler la composante reliée exclusivement à la variance de l'habileté. Pour ce faire, nous avons effectué une décomposition de la variance, tel que dans Lemieux & al. (2009). Cette méthode

---

<sup>8</sup> Il est possible de décomposer la variance de l'habileté en deux composantes, soit le rendement sur l'habileté et la variance de l'habileté en tant que telle. Or, par souci de simplicité, nous ignorons ce fait.

permet de décomposer la variance de la variable dépendante de tout modèle en sa partie expliquée, c'est-à-dire la variance des variables explicatives et des coefficients associés à ces variables, et en sa partie non expliquée, c'est-à-dire la variance du terme d'erreur du modèle. Dans le cas de notre modèle, la variance du terme d'erreur correspond à la variance de l'habileté et constitue donc notre paramètre d'intérêt.

Le tableau 3 contient les résultats de cette décomposition pour chaque groupe que nous étudions, soit les natifs de 1<sup>ère</sup> génération, les immigrants de 1<sup>ère</sup> génération, les natifs de 2<sup>e</sup> génération et les immigrants de 2<sup>e</sup> génération. Pour chacun de ces groupes, la variance du revenu d'entreprise et du salaire  $y$  est décomposée. Les modèles économétriques que nous avons utilisés correspondent aux équations de revenu présentées à la section IV.III. Pour obtenir la portion expliquée de la variance, nous avons calculé la variance des valeurs prédites par les équations de revenu, alors que pour obtenir la portion non expliquée, nous avons calculé la variance des résidus du modèle. Les colonnes 1 et 4 illustrent ces résultats pour chacun des groupes. Le constat est assez clair : peu importe le groupe étudié, la portion non expliquée de la variance du revenu d'entreprise est nettement plus élevée que la portion non expliquée de la variance du salaire. En d'autres termes, la variance de l'habileté est nettement plus élevée chez les entrepreneurs que chez les travailleurs salariés.

Or, il faut faire attention de ne pas conclure trop rapidement étant donné que les échantillons dont nous disposons sont le résultat d'un processus de sélection non aléatoire des travailleurs au sein de l'entrepreneuriat et du salariat. Pour remédier à ce problème, nous prenons soin de bâtir un échantillon contrefactuel d'entrepreneurs, c'est-à-dire un échantillon d'entrepreneurs dont le revenu correspond au revenu qu'ils auraient obtenu s'ils avaient été payés comme des travailleurs salariés, conditionnellement à leurs caractéristiques observables. Nous bâtissons cet échantillon au moyen de la méthode de DiNardo-Fortin-Lemieux, que nous expliquons dans l'annexe 3. En comparant la portion non expliquée de la variance du revenu dans l'échantillon contrefactuel à celle dans l'échantillon d'entrepreneurs, nous sommes en mesure de savoir si l'entrepreneuriat offre bel et bien un rendement sur le niveau d'habileté plus élevé que le travail salarié. Les colonnes 2 et 5 contiennent les résultats de la décomposition de variance pour les échantillons contrefactuels d'entrepreneurs et les colonnes 3 et 6 contiennent la différence entre la variance

observée dans l'échantillon réel et celle observée dans l'échantillon contrefactuel d'entrepreneurs. Puisque chez tous les groupes étudiés, cette différence est fortement positive pour la portion non expliquée de la variance, nous concluons que la variance de l'habileté est nettement plus élevée chez les entrepreneurs que chez les travailleurs salariés. Autrement dit, le rapport  $\frac{\sigma_E}{\sigma_S}$  est supérieur à 1.

En ce qui concerne l'évolution intergénérationnelle du différentiel de revenu associé à chaque groupe ethnique, le graphique 12 illustre le différentiel de revenu observé dans la 2<sup>e</sup> génération en fonction du différentiel de revenu qui prévalait dans la 1<sup>ère</sup> génération. On y remarque plusieurs choses. Premièrement, la presque totalité des groupes voient leur différentiel de revenu diminuer entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. En d'autres mots, l'écart entre le revenu d'entrepreneuriat moyen et le salaire moyen diminue pour presque tous les groupes. Deuxièmement, cette tendance à la baisse coïncide avec le fait que dans la 1<sup>ère</sup> génération, presque tous les groupes affichaient un écart positif entre le revenu d'entrepreneuriat moyen et le salaire moyen, alors que dans la 2<sup>e</sup> génération, beaucoup de groupes voient cet écart tomber dans le négatif. Alors qu'il serait intéressant de connaître les phénomènes à l'origine de cette tendance, cela nécessiterait une analyse très approfondie qui dépasse les objectifs de ce mémoire. De plus, puisque presque tous les groupes, y compris les natifs (graphique 14), sont affectés de façon similaire par cette tendance, son analyse ne nous permettrait pas de mieux comprendre les causes de l'effet de l'ethnicité sur la probabilité d'être entrepreneur. Finalement, alors que le taux d'entrepreneuriat, tout comme plusieurs variables sociodémographiques, est soumis à un processus de convergence partielle entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération, le graphique 12 ne montre aucun signe de convergence. Cette absence de convergence est également reflétée sur le graphique 13, où il semble n'y avoir aucun lien entre la variation du taux d'entrepreneuriat et la variation du différentiel de revenu sectoriel entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. Autrement dit, à la lumière de ce graphique, le lien entre le taux d'entrepreneuriat d'un groupe ethnique et son différentiel de revenu sectoriel serait inexistant.

Bien que les constatations faites sur le graphique 13 puissent surprendre, elles ne sont pas anormales du moment qu'on comprend le choix sectoriel comme un processus de sélection, tel que présenté à la section II.I. En effet, lorsqu'un groupe ethnique voit son taux d'entrepreneuriat



changer, cela signifie que la différence entre le revenu d'entrepreneuriat espéré et le salaire espéré devient ou bien négative, ou bien positive pour une plus grande proportion de sa population de travailleurs. La question est de savoir qui sont les individus pour lesquels cette différence passe du négatif au positif, ou l'inverse. Ce que nous savons, c'est qu'il y a de fortes chances que ces individus ne possèdent en moyenne pas les mêmes caractéristiques que le reste de la population. Par exemple, si pour une raison quelconque, le salaire moyen d'un groupe augmentait fortement par rapport à son revenu d'entrepreneuriat moyen, et ce, pour l'ensemble des individus, on s'attendrait à ce qu'une plus grande proportion d'individus œuvre dans le secteur salarié, et donc à ce qu'il y ait une diminution du taux d'entrepreneuriat. Or, les individus qui quitteraient l'entrepreneuriat pour devenir travailleurs salariés sont sans doute ceux qui, en moyenne, seraient le plus positivement impactés par l'augmentation du revenu moyen. Il s'agirait donc probablement des entrepreneurs les moins qualifiés et les moins doués. Par conséquent, même si l'augmentation du salaire moyen diminue de façon automatique le différentiel de revenu, la sélection négative des nouveaux travailleurs salariés a le potentiel de contrecarrer cet effet en haussant le revenu moyen des entrepreneurs. L'effet net sera donc déterminé par la taille de l'augmentation du salaire moyen, de même que par d'autres paramètres tels que la variance du salaire et la variance du revenu moyen.

Cet exemple fictif, conjointement avec les résultats du graphique 13, nous montre qu'un changement dans la distribution des revenus d'entreprise ou dans la distribution des salaires peut avoir de multiples impacts sur le différentiel de revenu sectoriel. Par conséquent, lorsque plus loin dans ce mémoire, nous tentons d'identifier les mécanismes responsables de l'effet de l'ethnicité sur la probabilité d'être entrepreneur, nous n'analysons pas l'effet de l'ethnicité sur le différentiel de revenu, qui est très souvent ambigu, mais plutôt son effet sur le revenu d'entreprise et sur le salaire de manière séparée. Cette manière de procéder est également justifiée par les prédictions du modèle théorique que nous présentons dans la sous-section V.IV.

## IV. Méthodologie

### IV.I L'Effet de l'ethnicité sur le choix sectoriel

Pour estimer l'effet d'appartenir à un groupe ethnique sur la probabilité d'être entrepreneur, nous estimons dans cette sous-section une série d'équations de choix sectoriel, c'est-à-dire une série d'équations qui expliquent le choix de devenir entrepreneur ou non. Plus précisément, nous estimons :

$$y_i = \alpha + x_i' \beta + \sum_{j=1}^{22} \delta_j D_j + \varepsilon_i \quad (17)$$

où  $y_i$  est une variable dichotomique qui vaut 1 si l'individu  $i$  est entrepreneur,  $\alpha$  est une constante,  $x$  est un vecteur de variables sociodémographiques,  $\beta$  est le vecteur de coefficients associé à ces variables,  $D_j$  est une variable dichotomique qui vaut 1 si l'individu appartient au groupe ethnique  $j$ ,  $\delta_j$  est le coefficient associé à cette variable et où  $\varepsilon_i$  est un terme d'erreur propre à chaque individu. Nous avons choisi comme groupe de référence les Américains d'origine, de manière à ce que nous puissions interpréter les coefficients ethniques comme des déviations par rapport à la population native. Nous utilisons deux méthodes différentes pour estimer l'équation décrite ci-haut: la régression des moindres carrés ordinaires (MCO) et le modèle Probit, estimé par maximum de vraisemblance. Alors que contrairement au Probit, les moindres carrés ordinaires n'exigent pas que le terme d'erreur du modèle suive une loi normale centrée en 0 afin de produire des estimés convergents et non biaisés, le Probit a l'avantage de forcer les probabilités prédites par le modèle à se trouver à l'intérieur de l'intervalle  $[0,1]$ . Or, même si l'hypothèse de normalité du terme d'erreur n'est généralement pas violée en présence de gros échantillons, nous employons les deux méthodes afin de garantir un certain niveau de robustesse de nos résultats.

Dans un premier temps, lorsque nous estimons les équations de choix sectoriel par MCO, nous estimons dans les faits un modèle de probabilité linéaire (MPL) en raison de la nature dichotomique de la variable dépendante  $y$ . Dans notre contexte, un modèle de probabilité linéaire implique que la probabilité associée à la réalisation de notre variable dichotomique  $y$ , soit

le fait d'être entrepreneur, est fonction linéaire des paramètres présents dans l'équation (17). Nous pouvons exprimer cette relation de la manière suivante :

$$P(y_i = 1|x_i, D_j) = E(y_i|x_i, D_j) = \alpha + x_i'\beta + \sum_{j=1}^{22} \delta_j D_j \quad (18)$$

Par conséquent, les coefficients qui nous intéressent, soit les coefficients  $\delta_j$  associés aux variables dichotomiques  $D_j$  qui servent à identifier les 22 groupes ethniques de notre analyse, représentent l'effet partiel du fait d'appartenir au groupe ethnique  $j$  sur la probabilité d'être entrepreneur :

$$\frac{\partial P(y_i = 1|x_i, D_j)}{\partial D_j} = \delta_j \quad (19)$$

Pour obtenir des estimés convergents et non biaisés des coefficients du modèle, nous émettons l'hypothèse que  $E(\varepsilon_i|D_j)=0$ . Cette hypothèse suppose que nous avons inclus dans le modèle toutes les variables qui influencent la décision de devenir entrepreneur et qui sont corrélées avec les variables dichotomiques  $D_j$ . Ainsi, dans l'éventualité où quelques-unes des variables dichotomiques  $D_j$  seraient corrélées avec d'autres variables sociodémographiques non incluses dans le modèle, nous prendrions pour acquis que ces variables ont un effet partiel nul sur la probabilité de devenir entrepreneur.

Afin que nous puissions faire de l'inférence statistique à partir des estimés de nos coefficients, la méthode des MCO exige que les termes d'erreur  $\varepsilon$  de notre modèle soient homoscédastiques et qu'ils soient distribués indépendamment les uns des autres selon une loi normale centrée en 0 et de variance  $\sigma_\varepsilon^2$ . Tel que précédemment mentionné, l'hypothèse de normalité asymptotique du terme d'erreur est généralement respectée en présence de gros échantillons. Cependant, il n'en est pas de même pour l'hypothèse d'homoscédasticité, car la nature même d'un modèle en probabilité linéaire force le terme d'erreur de ce dernier à être hétéroscédastique. Pour cette raison, nous devons utiliser des écarts-types qui sont robustes à l'hétéroscédasticité lorsque nous estimons les équations de choix sectoriel à partir des MCO. Puisque les variables dichotomiques qui identifient les différents groupes ethniques captent déjà la composante du terme d'erreur qui demeure fixe au sein de chaque groupe, nous employons simplement des écarts-types de White.

Ces écarts-types sont calculés à partir d'une matrice de variance-covariance qui permet à chaque terme d'erreur de l'échantillon d'afficher une valeur différente. Plus précisément, ils permettent aux termes de la diagonale de la matrice de variance-covariance de différer les uns des autres.

Dans un second temps, lorsque nous utilisons un modèle Probit pour estimer les équations de choix sectoriel, la probabilité que la variable dépendante dichotomique se réalise (la probabilité que  $y$  vaille 1) est liée aux variables explicatives du modèle par l'entremise d'une variable latente plutôt que de leur être directement liée, comme c'était le cas dans le modèle de probabilité linéaire. Cette variable latente est fonction linéaire des variables explicatives du modèle. Dans le contexte du modèle de choix sectoriel, cette variable latente correspond à la propension à devenir entrepreneur. Pour établir un lien entre la variable latente et la variable dichotomique, le modèle Probit suppose que l'individu choisira de devenir entrepreneur si sa propension à le devenir est positive. L'objectif du modèle Probit est donc d'estimer la probabilité que cette propension soit positive, tel qu'illustré par l'équation suivante :

$$P(y_i = 1 | x_i, D_j) = P\left(\alpha + x_i' \beta + \sum_{j=1}^{22} \delta_j D_j > -\varepsilon_i\right) \quad (20)$$

Puisqu'un modèle Probit émet l'hypothèse que le terme d'erreur suit une loi normale centrée réduite ( $\mathcal{N}(0,1)$ ), cela revient à dire que :

$$P(y_i = 1 | x_i, D_j) = \Phi\left(\alpha + x_i' \beta + \sum_{j=1}^{22} \delta_j D_j\right) \quad (21)$$

Finalement, en prenant la dérivée partielle de cette expression, nous obtenons comme résultat que l'effet partiel du fait d'appartenir au groupe  $j$  sur la probabilité d'être entrepreneur dépend des valeurs que prennent les autres variables explicatives :

$$\frac{\partial P(y_i = 1 | x_i, D_j)}{\partial D_j} = \delta_j \phi\left(\alpha + x_i' \beta + \sum_{j=1}^{22} \delta_j D_j\right) \quad (22)$$

où  $\phi$  représente la fonction de densité d'une distribution normale centrée réduite. Ainsi, si l'on s'intéresse à l'effet partiel du fait d'appartenir au groupe  $j$  non pas sur la propension à devenir entrepreneur, mais bien sur la probabilité d'être entrepreneur, il est nécessaire de choisir en quel point de la fonction de densité cumulative  $\Phi$  cet effet sera évalué. Toutefois, puisque seul le signe de l'effet des groupes ethniques sur le revenu nous intéresse, et puisque la multiplication du coefficient  $\delta_j$  par une constante positive, en l'occurrence une probabilité, ne change ni son signe, ni sa corrélation avec les autres variables, nous ne calculons pas d'effets marginaux partiels pour effectuer notre analyse. De plus, en procédant ainsi, nous évitons que le niveau de significativité des variables ethniques soit affecté par le niveau de significativité des autres variables explicatives. Malgré tout, afin de respecter les usages de la littérature, nous avons choisi de présenter les résultats de nos régressions Probit sous forme d'effets marginaux moyens dans le tableau 5, en complément au tableau 4.

Il est également important de mentionner qu'en plus de satisfaire la condition d'exogénéité et de normalité asymptotique du terme d'erreur, un modèle Probit doit absolument être homoscédastique afin de produire des estimés convergents et non biaisés des coefficients d'intérêt. Puisqu'il est difficile de tester la présence d'hétéroscédasticité dans le cas d'un modèle Probit (les tests de Breusch-Pagan et de White sont non applicables), nous émettons simplement l'hypothèse que le modèle que nous estimons est homoscédastique.

Finalement, peu importe la méthode d'estimation employée, lorsque nous étudions l'échantillon qui correspond aux immigrants de 2<sup>e</sup> génération, nous utilisons les poids d'échantillonnage pour nous assurer que le sur-échantillonnage ou le sous-échantillonnage par rapport à certaines variables non incluses dans le modèle (par exemple, les états et les villes) ne biaise pas les coefficients estimés. Dans les faits, nous estimons donc un modèle transformé dans lequel chaque observation est multipliée par l'inverse de sa probabilité d'échantillonnage :

$$\frac{y_i}{Prob_i} = \frac{1}{Prob_i} \left( \alpha + x_i' \beta + \sum_{j=1}^{22} \delta_j D_j + \varepsilon_i \right) \quad (23)$$

Or, puisque pour chaque observation, chaque composante du modèle est multipliée par la même constante, les effets partiels que nous estimons conservent la même interprétation que dans un modèle non pondéré.

#### IV.II L'Effet de l'ethnicité sur le choix sectoriel : comparaison entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants

L'objectif d'estimer des équations de choix sectoriel de manière séparée pour la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants est de comparer la magnitude et le signe des coefficients ethniques d'une génération à l'autre. Pour ce faire, nous employons deux méthodes différentes.

Dans un premier temps, nous régressons les coefficients ethniques estimés de la 2<sup>e</sup> génération sur ceux de la 1<sup>ère</sup> génération. Nous effectuons deux régressions séparées, soit une pour les coefficients issus des modèles de probabilité linéaire et l'autre pour les coefficients issus des modèles Probit. Tel que mentionné dans la section précédente, les coefficients issus des modèles Probit ne représentent pas l'effet de l'ethnicité sur la probabilité d'être entrepreneur, mais plutôt l'effet de l'ethnicité sur la propension à être entrepreneur. Alors qu'il serait possible de calculer l'effet sur la probabilité d'être entrepreneur à partir de l'équation (22), il serait inutile de le faire, car les corrélations estimées ne changeraient pas si tous les coefficients faisant partie de chaque génération étaient multipliés par la même constante ; dans ce cas-ci la fonction de densité évaluée en un point précis. Par conséquent, à la fois pour les coefficients issus des modèles de probabilité linéaire et ceux issus des modèles Probit, nous estimons la régression suivante:

$$\widehat{\delta}_j^2 = \alpha + \beta \widehat{\delta}_j^1 + \varepsilon_j \quad (24)$$

où  $\widehat{\delta}_j^2$  et  $\widehat{\delta}_j^1$  représentent les coefficients ethniques estimés de la 2<sup>e</sup> et de la 1<sup>ère</sup> génération respectivement,  $\alpha$  représente une constante et où  $\varepsilon_j$  représente le terme d'erreur de la régression. Nous estimons la régression par moindres carrés ordinaires et utilisons les écart-types de White afin de tenir compte de toute forme d'hétéroscédasticité. En premier lieu, nous nous intéressons au signe du coefficient  $\beta$  ainsi qu'à son niveau de significativité, de façon à savoir s'il existe une corrélation (i.e. une relation de dépendance linéaire) entre les coefficients de chacune des deux

générations, et si oui, quel est le signe de cette corrélation. En second lieu, nous désirons savoir si le coefficient  $\beta$  est inférieur à 1 en valeur absolue, auquel cas nous concluons à une convergence (en moyenne) de l'effet d'appartenir à un groupe ethnique sur la probabilité d'être entrepreneur, ou s'il est plutôt supérieur à 1 en valeur absolue, auquel cas nous concluons à une divergence (en moyenne) de l'effet d'appartenir à un groupe ethnique d'une génération à l'autre. Finalement, tel que nous l'avons fait pour le taux d'entrepreneuriat et diverses variables sociodémographiques dans la section III, nous portons notre attention sur la valeur associée à la constante  $\alpha$ . Puisque les coefficients ethniques représentent des déviations par rapport à la probabilité moyenne ou à la propension moyenne en vigueur chez les natifs (à caractéristiques égales), une ordonnée à l'origine située à proximité de 0 signifierait qu'il y a ou bien atténuation de l'effet de l'ethnicité sur la probabilité d'être entrepreneur, ou bien accroissement de cet effet, dépendamment de la valeur de  $\beta$ .

Dans un deuxième temps, pour les groupes ethniques dont l'effet sur le choix sectoriel est significatif au minimum dans l'une ou l'autre des deux générations, nous calculons simplement, en valeur absolue, la différence entre l'estimé du coefficient dans la 2<sup>e</sup> génération ( $\delta_j^2$ ) et l'estimé du coefficient dans la 1<sup>ère</sup> génération ( $\delta_j^1$ ). Nous calculons cette différence en pourcentage par rapport à la valeur du coefficient estimé dans la 1<sup>ère</sup> génération. De cette façon, en faisant la moyenne des différences, nous sommes en mesure de voir si les coefficients estimés à partir des régressions décrites dans le dernier paragraphe sont représentatifs de la magnitude du changement subi par les groupes qui ont véritablement un effet sur la probabilité d'être entrepreneur dans au minimum l'une ou l'autre des deux générations.

Or, avant de procéder à tous ces calculs, il nous faut vérifier si les coefficients diffèrent de manière significative entre les deux générations. Pour ce faire, nous devrions normalement utiliser un test d'hypothèse qui nous permet d'inférer sur la valeur de la différence entre les coefficients associés à chacune des deux populations d'intérêt (les immigrants de 1<sup>ère</sup> génération et les immigrants de 2<sup>e</sup> génération). Cependant, puisque le modèle de choix sectoriel que nous estimons n'est pas exactement le même d'une génération à l'autre, et puisque les deux modèles estimés ne sont pas « imbriqués »<sup>9</sup> l'un dans l'autre, aucun test standard ne permet de tester la

---

<sup>9</sup> Par « imbriqué », nous entendons le fait que les variables incluses dans l'un des deux modèles constituent un sous-ensemble des variables incluses dans l'autre modèle.

différence entre les deux coefficients. Par conséquent, conformément aux recommandations de Schenker & Gentleman (2001), nous utilisons les intervalles de confiance associés aux coefficients estimés dans chacune des deux générations afin d'inférer sur la différence entre ces derniers. Cette méthode est valide lorsque les deux estimateurs comparés sont convergents, asymptotiquement normaux et asymptotiquement indépendants. Dans le cas de la convergence et de la normalité asymptotique, nous avons déjà pris pour acquis que l'estimateur des MCO (MPL) et que l'estimateur du maximum de vraisemblance (Probit) satisfont ces deux conditions étant donné les hypothèses que nous avons validées ou émises dans la dernière sous-section. Quant à la condition d'indépendance asymptotique, nous présumons que la covariance entre les termes d'erreur des modèles associés aux deux générations est asymptotiquement nulle, entre autres parce que nous contrôlons pour la probabilité d'avoir un père entrepreneur chez les immigrants de 2<sup>e</sup> génération<sup>10</sup>. Nous calculons les intervalles de confiance selon la formule suivante :

$$[Min(\delta_j^1)_{95\%}, Max(\delta_j^1)_{95\%}] = [\widehat{\delta}_j^1 - t_{97,5\%} \widehat{SE}_j^1, \widehat{\delta}_j^1 + t_{97,5\%} \widehat{SE}_j^1] \quad (25)$$

$$[Min(\delta_j^2)_{95\%}, Max(\delta_j^2)_{95\%}] = [\widehat{\delta}_j^2 - t_{97,5\%} \widehat{SE}_j^2, \widehat{\delta}_j^2 + t_{97,5\%} \widehat{SE}_j^2] \quad (26)$$

où  $\widehat{SE}_1$  représente l'écart-type de la distribution de l'estimateur du coefficient  $\delta_j^1$ ,  $\widehat{SE}_2$  représente l'écart-type de la distribution de l'estimateur du coefficient  $\delta_j^2$  et où  $t_{97,5\%}$  représente la statistique de Student associée à un seuil de signification unilatéral de 2,5%. Ainsi, à l'aide de ces formules, nous calculons des intervalles de confiance dont le niveau de certitude est de 95%. Autrement dit, avec ces intervalles, nous sommes sûrs à 95% que la valeur du coefficient dans la population se trouve à l'intérieur des bornes calculées. Par conséquent, si pour un groupe ethnique  $j$ , les intervalles de confiance associés à chacune des deux générations ne se recoupent pas, nous sommes sûrs à 95% que le coefficient du groupe en question diffère d'une génération à l'autre.

Évidemment, cette manière d'inférer sur la différence entre les coefficients n'est pas parfaite ; elle serait autrement toujours préférée aux tests conventionnels vu sa simplicité. En effet, Schenker & Gentleman (2001) recommandent d'utiliser les intervalles de confiance lorsqu'aucun

---

<sup>10</sup> Voir la sous-section V.I.



test statistique convergent ne peut être utilisé, mais reconnaît qu'il s'agit d'une méthode qui augmente la probabilité de commettre une erreur de type II. En d'autres termes, lorsque nous comparons des intervalles de confiance au seuil de confiance de 95%, nous avons plus de chances de ne pas rejeter l'hypothèse nulle alors que l'hypothèse alternative est vraie que lorsque nous employons un test valide au même seuil de confiance. En contrepartie, le risque de commettre une erreur de type I, soit de rejeter l'hypothèse nulle alors que celle-ci est vraie, est en réalité inférieur à 5% lorsque nous utilisons des intervalles de confiance au seuil de confiance de 95%. En gros, il s'agit donc d'une méthode plus conservatrice que l'utilisation d'un test d'hypothèse. Ainsi, pour les groupes ethniques dont les intervalles ne se recoupent pas, nous aurons un haut niveau de certitude que leurs coefficients diffèrent d'une génération à l'autre. Cependant, pour les groupes dont les intervalles ne se recoupent que légèrement, nous devons exercer un certain jugement arbitraire pour déterminer si les coefficients diffèrent bel et bien. Pour savoir si les intervalles se recoupent, et pour connaître la taille du recouvrement, nous employons une des deux formules suivantes, dépendamment du contexte :

$$\text{Min}(\delta_j^1)_{95\%} - \text{Max}(\delta_j^2)_{95\%} \text{ lorsque } \widehat{\delta}_j^2 < \widehat{\delta}_j^1 \quad (27)$$

$$\text{Min}(\delta_j^2)_{95\%} - \text{Max}(\delta_j^1)_{95\%} \text{ lorsque } \widehat{\delta}_j^2 > \widehat{\delta}_j^1 \quad (28)$$

Lorsque le résultat de la formule utilisée est négatif, nous concluons à un recouvrement. De plus, la valeur du résultat correspond à la taille du recouvrement.

#### IV.III L'Effet de l'ethnicité sur le salaire et sur le revenu d'entreprise

Afin d'estimer l'effet d'appartenir à un groupe ethnique sur le revenu dans chaque secteur, nous devons d'abord et avant tout estimer une série d'équations de revenu de la forme suivante :

$$\log(R_i) = \alpha + x_i'\beta + \sum_{j=1}^{22} \gamma_j D_j + \varepsilon_i \quad (29)$$

où  $\log(R_i)$  est le log du revenu de l'individu  $i$  en dollars<sup>11</sup>,  $\alpha$  est une constante,  $x$  est un vecteur de variables sociodémographiques,  $\beta$  est le vecteur de coefficients associé à ces variables,  $D_j$  est une variable dichotomique qui vaut 1 lorsque l'individu  $i$  appartient au groupe ethnique  $j$ ,  $\gamma_j$  est le coefficient associé à cette variable et où  $\varepsilon_i$  est un terme d'erreur propre à chaque individu. Comme c'était le cas pour les équations de choix sectoriel, le groupe de référence correspond aux Américains d'origine. Tel qu'il est d'usage dans la littérature, nous utilisons le log du revenu plutôt que le revenu comme variable dépendante puisque nous tenons à ce que nos variables explicatives aient un effet partiel constant en pourcentage sur le revenu. Nous estimons au total quatre équations de revenu, soit une équation par génération pour chacune des deux catégories de travailleurs auxquelles nous nous intéressons: les entrepreneurs et les travailleurs salariés. Nous estimons l'ensemble de ces équations par moindres carrés ordinaires et utilisons des écarts-types de White afin de tenir compte de toute forme potentielle d'hétéroscédasticité engendrée par les données elles même. De cette façon, nous sommes certains de pouvoir faire de l'inférence à partir des estimés issus des quatre équations de revenu.

Puisque nous nous intéressons à l'effet de l'ethnicité sur le revenu des entrepreneurs et des travailleurs salariés, nous souhaiterions normalement obtenir des estimés convergents et non biaisés des coefficients associés aux variables dichotomiques qui représentent les différents groupes ethniques. Pour cela, nous avons inclus dans chacune des équations l'ensemble des variables susceptibles d'influencer le revenu des travailleurs et potentiellement corrélées avec les variables  $D_j$ . Or, alors que l'inclusion de toutes les variables explicatives pertinentes permet d'éliminer le biais de variable omise de notre estimation, cette pratique ne permet pas d'éliminer le biais d'auto sélection. Pour bien comprendre pourquoi, il faut avant tout se rappeler que les travailleurs salariés et les entrepreneurs ne constituent tous les deux pas un échantillon aléatoire de la population de travailleurs. Bien que nous contrôlions pour l'assignation non aléatoire des travailleurs par rapport aux caractéristiques observables telles que l'éducation, nous ne sommes

---

<sup>11</sup> Pour la 1<sup>ère</sup> génération, nous utilisons le revenu en dollars de 1980. Pour la 2<sup>e</sup> génération, nous utilisons comme unité de compte le dollar de 2006.

pas en mesure de contrôler pour l'assignation non aléatoire par rapport aux caractéristiques non observables (i.e. l'habileté). Conséquemment à l'analyse que nous avons faite dans la sous-section III.V, nous avons de bonnes raisons de croire que les individus qui choisissent de devenir entrepreneur possèdent en moyenne un niveau d'habileté supérieur au travailleur moyen, et qu'il s'agit du contraire pour les individus qui décident d'occuper un emploi salarié. Ainsi, si certaines de nos variables explicatives sont corrélées avec le niveau d'habileté, les estimés des coefficients associés à ces variables seront biaisés. À titre d'exemple, si l'éducation est positivement corrélée avec le niveau d'habileté, le coefficient associé à l'éducation risque d'être biaisé à la hausse chez les entrepreneurs, et biaisé à la baisse chez les travailleurs salariés. Nous appliquons un raisonnement similaire aux coefficients qui modélisent l'effet partiel d'appartenir aux différents groupes ethniques. À vrai dire, pour chaque groupe ethnique, ces coefficients captent, entre autres choses, le niveau d'habileté moyen des entrepreneurs et des travailleurs salariés comme conséquence de l'auto-sélection de ces derniers dans chacun des secteurs. Pour des raisons expliquées dans les deux prochaines sous-sections, nous émettons l'hypothèse que cette auto-sélection diffère d'un groupe ethnique à l'autre. De fait, nous concluons que  $E(\varepsilon_i|D_j) \neq 0$ . Or, dans le contexte de notre analyse, ceci n'est pas un problème, car tel que nous le verrons dans les deux prochaines sous-sections, nous souhaitons mieux comprendre d'où provient l'effet de l'ethnicité sur la probabilité d'être entrepreneur à partir de l'auto-sélection par rapport à l'habileté des travailleurs salariés et des entrepreneurs de chaque groupe ethnique. Par conséquent, nous ne tentons pas d'éliminer le biais de sélection à partir d'une quelconque méthode économétrique, car autrement, nous ne serions plus en mesure de conduire l'analyse que nous souhaitons conduire.

#### IV.IV L'Explication de l'effet de l'ethnicité sur le choix sectoriel : cadre théorique

Nous savons que différents phénomènes sociaux peuvent être à l'origine de l'effet de l'ethnicité sur le choix sectoriel. Parmi ceux-ci, nous avons mentionné la discrimination sur la base de l'ethnicité dans l'un ou l'autre des secteurs, de même que l'existence de phénomènes plus positifs tels que le développement de réseaux d'affaires au sein des communautés. L'hypothèse selon laquelle les individus choisissent le secteur dans lequel ils travaillent en fonction de la différence entre le revenu qu'ils peuvent espérer gagner en tant qu'entrepreneur et le revenu qu'ils peuvent

espérer gagner en tant que travailleur salarié nous laisse croire que si l'un de ces phénomènes se produisait bel et bien, son impact sur le choix sectoriel des immigrants s'effectuerait par l'entremise de son influence sur le salaire et sur le revenu d'entreprise espérés, ou du moins en partie. Ainsi, dans le but de détecter quels phénomènes sont à la source de l'effet de l'ethnicité sur le choix sectoriel, nous tentons de développer un cadre analytique formel qui nous permet de prévoir l'impact de différents phénomènes sociaux sur le salaire et le revenu d'entreprise que les membres de différentes communautés ethniques peuvent espérer toucher. Pour ce faire, nous nous appuyons sur le modèle de sélection classique adapté pour le choix sectoriel (Borjas & Bronars, 1989) que nous avons présenté dans la sous-section II.I. Nous débutons notre analyse en écrivant les équations qui définissent l'espérance du revenu dans chaque secteur conditionnellement au fait que nous nous trouvions dans le secteur en question:

$$E(w_E|X, y = 1) = XB_E + E(\varepsilon_E|X, y = 1) \quad (30)$$

$$E(w_S|X, y = 0) = XB_S + E(\varepsilon_S|X, y = 0) \quad (31)$$

Selon ces deux équations, deux éléments peuvent faire en sorte que le revenu conditionnel espéré diffère d'un secteur à l'autre : les retours sur les caractéristiques observables et l'espérance conditionnelle du terme d'erreur. En réécrivant les espérances conditionnelles des termes d'erreur, communément appelées *termes de sélection*, en fonction des paramètres du modèle de sélection présenté dans la sous-section II.I, nous obtenons les deux équations suivantes :

$$E(w_E|X, y = 1) = XB_E + \frac{\sigma_E \sigma_S}{\sigma_v} \left( \frac{\sigma_E}{\sigma_S} - \rho \right) \left( \frac{\phi \left( \frac{XB_S - XB_E}{\sigma_v} \right)}{1 - \Phi \left( \frac{XB_S - XB_E}{\sigma_v} \right)} \right) \quad (32)$$

$$E(w_S|X, y = 0) = XB_S + \frac{\sigma_E \sigma_S}{\sigma_v} \left( \rho - \frac{\sigma_S}{\sigma_E} \right) \left( \frac{-\phi \left( \frac{XB_S - XB_E}{\sigma_v} \right)}{\Phi \left( \frac{XB_S - XB_E}{\sigma_v} \right)} \right) \quad (33)$$

Nous ne prenons pas le temps de définir tous les paramètres présents dans ces deux équations étant donné que nous avons déjà accompli cet exercice dans la sous-section II.I. L'objectif est maintenant de savoir lesquels de ces paramètres sont susceptibles d'être affectés par les différents phénomènes de nature sociale que nous avons mentionnés plus haut. Dans le cas de la discrimination dans l'un ou l'autre des secteurs, il serait peu logique de croire que celle-ci engendre un changement du rapport des variances. En effet, généralement, lorsqu'un groupe est discriminé, l'ensemble de sa population s'en trouve négativement affectée. Or, dans le contexte du modèle de sélection que nous employons, l'habileté est distribuée selon une loi normale centrée en 0. Par conséquent, si nous supposons, à titre d'exemple, que la discrimination dans le marché du travail (i.e. la discrimination à l'endroit des travailleurs salariés) engendre une compression de la distribution des salaires, cela signifierait que les travailleurs les moins habiles s'en tirent mieux lorsqu'ils sont discriminés que lorsqu'ils ne le sont pas. Pour cette raison, nous écartons la possibilité d'un changement du rapport des variances. Plutôt, nous émettons l'hypothèse que la discrimination diminue le rendement sur les caractéristiques observables dans le secteur concerné. Cette hypothèse est réaliste compte tenu du fait que dans la littérature économique, la discrimination est souvent mesurée en comparant les rendements sur les caractéristiques observables (par exemple, l'éducation) du groupe discriminé avec ceux du groupe de référence. Nous supposons l'existence de deux types de discrimination pouvant influencer la propension à devenir entrepreneur : la discrimination dans le marché du travail (Fairlie & Meyer, 1996, Clark & Drinkwater, 2000) et la discrimination dans le marché des produits (Borjas & Bronars, 1989).

Dans le cas de la discrimination dans le marché du travail, nous pouvons nous attendre à ce que les rendements sur les caractéristiques observables diminuent pour les travailleurs salariés qui sont discriminés. En d'autres mots, nous supposons que la discrimination dans le marché du travail engendre une baisse de la valeur du coefficient  $B_s$ .<sup>12</sup> Empiriquement, cette baisse du coefficient  $B_s$  est une baisse par rapport au coefficient estimé sur l'ensemble de la population que nous étudions, laquelle est à grande majorité composée d'Américains de souche qui, normalement, ne devraient pas être victimes de discrimination. De manière générale, l'effet d'un

---

<sup>12</sup> Afin de simplifier l'analyse, nous supposons que le modèle de sélection est un modèle linéaire simple, autrement dit que le terme  $X$  représente une seule variable, par exemple l'éducation. Dans ce contexte, le terme  $B$  représente un scalaire plutôt qu'un vecteur.

changement de  $B_S$  sur l'espérance du revenu conditionnel dans chaque secteur peut être représenté par les deux équations suivantes :

$$\frac{\partial E(w_E|X, y = 1)}{\partial B_S} = \frac{\sigma_E \sigma_S}{\sigma_v} \left( \frac{\sigma_E}{\sigma_S} - \rho \right) \frac{\partial \lambda_E(z)}{\partial z} \frac{\partial z}{\partial B_S} \quad (34)$$

$$\frac{\partial E(w_S|X, y = 0)}{\partial B_S} = X + \frac{\sigma_E \sigma_S}{\sigma_v} \left( \rho - \frac{\sigma_S}{\sigma_E} \right) \frac{\partial \lambda_S(z)}{\partial z} \frac{\partial z}{\partial B_S} \quad (35)$$

où  $z$  représente le terme  $\left( \frac{XB_S - XB_E}{\sigma_v} \right)$ ,  $\lambda_E$  représente le ratio de Mills inversé présent dans l'équation du revenu d'entreprise  $\left( \left( \frac{\phi(z)}{1 - \Phi(z)} \right) \right)$  et où  $\lambda_S$  représente le ratio de Mills inversé présent dans l'équation du salaire  $\left( \left( \frac{-\phi(z)}{\Phi(z)} \right) \right)$ . À la lumière de ces deux expressions, pour savoir si l'effet du changement de  $B_S$  sur le revenu conditionnel espéré dans chacun des secteurs est positif ou négatif, il est primordial de connaître le signe des dérivées  $\frac{\partial \lambda_E(z)}{\partial z}$  et  $\frac{\partial \lambda_S(z)}{\partial z}$ . Il s'agit d'une tâche simple, car par définition, nous savons que les termes  $\lambda_E(z)$  et  $\lambda_S(z)$  représentent l'espérance de la variable  $\frac{\varepsilon_E - \varepsilon_S}{\sigma_v}$  conditionnellement au fait que cette variable soit supérieure et inférieure à  $z$ , respectivement. Plus précisément, nous savons que :

$$\lambda_E(z) = E \left( \frac{\varepsilon_E - \varepsilon_S}{\sigma_v} \mid \frac{\varepsilon_E - \varepsilon_S}{\sigma_v} > z \right) \quad (36)$$

$$\lambda_S(z) = E \left( \frac{\varepsilon_E - \varepsilon_S}{\sigma_v} \mid \frac{\varepsilon_E - \varepsilon_S}{\sigma_v} < z \right) \quad (37)$$

Intuitivement, nous devinons qu'une hausse de la valeur de  $z$  entraînera une hausse de l'espérance de  $\frac{\varepsilon_E - \varepsilon_S}{\sigma_v}$  dans les deux cas, et nous en déduisons donc que les dérivées  $\frac{\partial \lambda_E(z)}{\partial z}$  et  $\frac{\partial \lambda_S(z)}{\partial z}$  sont toutes les deux positives.<sup>13</sup> Par conséquent, du moment que  $\rho$  est suffisamment positif<sup>14</sup> (i.e. que

---

<sup>13</sup> Cette façon d'analyser la dérivée du ratio de Mills inversé est aussi employée par Borjas (1987).

l'habileté est suffisamment transférable entre l'entrepreneuriat et le salariat), nous sommes certains que la dérivée de l'espérance conditionnelle du revenu d'entreprise par rapport à  $B_S$  et que la dérivée de l'espérance conditionnelle du salaire par rapport à  $B_S$  sont toutes les deux positives. De manière générale, nous sommes donc sûrs que la variation de l'espérance conditionnelle du salaire sera du même signe que la variation de l'espérance conditionnelle du revenu d'entreprise à la suite d'un quelconque changement de  $z$ . Il reste la question de savoir laquelle des deux espérances varie le plus en valeur absolue. Plus précisément, nous aimerions savoir quand est-ce que la condition illustrée ci-dessous est valide :

$$\frac{\partial E(w_E|X, y = 1)}{\partial B_S} - \frac{\partial E(w_S|X, y = 0)}{\partial B_S} > 0 \quad (38)$$

Cependant, il est difficile de répondre à cette question, car la valeur de la différence entre les deux dérivées dépend simultanément de deux variables : le rapport des variances  $\frac{\sigma_E}{\sigma_S}$  et la variable  $z$ . De fait, nous nous contentons pour l'instant de dériver les deux conditions suivantes :

1. Le revenu d'entreprise moyen observé chez un groupe discriminé est, toutes choses étant égales par ailleurs, inférieur à celui observé chez groupe non discriminé. Ainsi, l'effet partiel d'appartenir à un groupe discriminé sur l'espérance conditionnelle du revenu d'entreprise est négatif. Autrement dit, lorsque  $B_S$  diminue:

$$\frac{\partial E(w_E|X, y = 1)}{\partial B_S} dB_S < 0 \quad (39)$$

2. Le salaire moyen observé chez un groupe discriminé est, toutes choses étant égales par ailleurs, inférieur à celui observé chez groupe non discriminé. Ainsi, l'effet partiel d'appartenir à un groupe discriminé sur l'espérance conditionnelle du salaire est négatif. Autrement dit, lorsque  $B_S$  diminue:

---

<sup>14</sup> Consulter la sous-section II.I pour plus de détails à cet effet.

$$\frac{\partial E(w_S|X, y = 0)}{\partial B_S} dB_S < 0 \quad (40)$$

Ainsi, l'impact de la discrimination dans le marché du travail sur le niveau d'habileté moyen est négatif dans les deux secteurs. L'intuition derrière ce résultat est plutôt simple. Lorsqu'un groupe ethnique est victime de discrimination dans le marché du travail, ce groupe aura tendance à perdre ses meilleurs travailleurs salariés au profit de l'entrepreneuriat. Conséquemment, le niveau d'habileté moyen de ses travailleurs salariés diminuera. Les nouveaux entrepreneurs seront aussi en moyenne moins habiles que les entrepreneurs déjà présents, étant donné que les individus les plus habiles avaient déjà choisi d'être entrepreneur en raison du fait que l'entrepreneuriat offre des rendements plus élevés sur l'habileté ( $\sigma_E > \sigma_S$ ) et en raison du fait que l'habileté est suffisamment transférable d'un secteur à l'autre ( $\rho$  est suffisamment positif).

Puisque la discrimination dans le marché du travail impacte le différentiel de revenu sectoriel, on s'attend à ce qu'elle impacte aussi la propension à devenir entrepreneur des membres des groupes discriminés. À la lumière du raisonnement que nous avons exposé dans le dernier paragraphe, nous présumons qu'un groupe discriminé dans le marché du travail affichera un taux d'entrepreneuriat plus élevé qu'un groupe non discriminé. Pour en être certain, il convient d'abord de réécrire la condition nécessaire pour qu'un individu décide de devenir entrepreneur :

$$w_E - w_S > 0 \quad (41)$$

Dans un contexte où le revenu dans chaque secteur dépend de caractéristiques observables, cette équation prend la forme :

$$(XB_E + \varepsilon_E) - (XB_S + \varepsilon_S) > 0 \quad (42)$$

Ainsi, la probabilité qu'un individu devienne entrepreneur s'écrit comme suit :

$$P(y = 1|X) = P(\varepsilon_E - \varepsilon_S > XB_S - XB_E) \quad (43)$$



Puisque les termes d'erreur  $\varepsilon_E$  et  $\varepsilon_S$  sont chacun distribués selon une loi normale centrée en 0, cette probabilité peut être reformulée ainsi :

$$P(y = 1|X) = 1 - \Phi\left(\frac{(XB_S - XB_E)}{\sigma_v}\right) \quad (44)$$

ou encore

$$P(y = 1|X) = 1 - \Phi(z) \quad (45)$$

où  $\sigma_v$  représente la variance de la différence  $\varepsilon_E - \varepsilon_S$  et où  $\Phi$  représente la fonction de répartition d'une loi normale centrée réduite. Puisque l'impact d'une diminution de  $B_S$  sur le terme  $z$  est négatif, nous devrions observer, conformément à nos attentes, une troisième relation chez les groupes ethniques victimes de discrimination :

**3.** Un groupe discriminé affiche, toutes choses étant égales par ailleurs, un taux d'entrepreneuriat supérieur au taux observé chez un groupe non discriminé. Ainsi, l'effet partiel d'appartenir à un groupe discriminé sur la probabilité conditionnelle d'être entrepreneur est positif. Autrement dit, lorsque  $B_S$  diminue :

$$\frac{\partial P(y = 1|X)}{\partial B_S} dB_S > 0 \quad (46)$$

#### IV.V L'Explication de l'effet de l'ethnicité sur le choix sectoriel : cadre empirique

Pour que les trois relations que nous venons de présenter puissent être reproduites à partir des estimés tirés des équations de revenu et de choix sectoriel, il faudrait normalement que les différentielles partielles rattachées à chacune de ces relations constituent *l'équivalent direct* des coefficients ethniques présents dans ces équations. En d'autres mots, les trois conditions suivantes devraient être satisfaites :

$$\delta_j = \frac{\partial P(y = 1|X, D_j)}{\partial D_j} = \frac{\partial P(y = 1|X)}{\partial B_s} dB_s \quad (47)$$

$$\gamma_j^E = \frac{\partial E(w_E|X, D_j, y = 1)}{\partial D_j} = \frac{\partial E(w_E|X, y = 1)}{\partial B_s} dB_s \quad (48)$$

$$\gamma_j^S = \frac{\partial E(w_S|X, D_j, y = 0)}{\partial D_j} = \frac{\partial E(w_S|X, y = 0)}{\partial B_s} dB_s \quad (49)$$

Or, nous verrons que les coefficients estimés ne constituent pas des équivalents directs, mais que dans bien des cas, ils permettent néanmoins d'arriver aux mêmes conclusions que dans le cadre d'analyse théorique. Pour débiter, dans le cas des coefficients ethniques tirés des équations de choix sectoriel, nous savons que la probabilité qu'un individu soit entrepreneur est donnée par l'équation suivante lorsque nous employons le modèle de probabilité linéaire (MPL):

$$P(y = 1|X, D_j) = X\beta + \sum_{j=1}^{22} \delta_j D_j \quad (50)$$

Puisque toutes les variables sociodémographiques pour lesquelles nous contrôlons dans les équations de revenu sont incluses dans les équations de choix sectoriel (i.e. les X des équations de revenu sont un sous-ensemble des X des équations de choix sectoriel), nous pouvons affirmer que l'effet d'appartenir au groupe ethnique j sur la probabilité de devenir entrepreneur correspond bel et bien à l'effet de la discrimination dont est victime le groupe j sur la probabilité d'être entrepreneur. En d'autres mots, nous sommes en présence de la relation suivante :

$$\delta_j = \frac{\partial P(y = 1|X, D_j)}{\partial D_j} = \frac{\partial P(y = 1|X)}{\partial B_s} dB_s \quad (51)$$

En ce qui concerne les coefficients ethniques tirés des équations de revenu, nous savons que, conformément à l'équation (31), l'effet de la discrimination dans le marché du travail sur l'espérance conditionnelle du salaire est représenté par l'équation suivante:

$$\frac{\partial E(w_S|X, y = 0)}{\partial B_S} dB_S = \left( X + \frac{\partial E(\varepsilon_S|X, y = 0)}{\partial B_S} \right) dB_S \quad (52)$$

Au même moment, nous savons que les équations de salaire estimées par MCO fournissent des fonctions d'espérance conditionnelle ayant la forme suivante:

$$E(w_S|X, D_j, y = 0) = XB_S + \sum_{j=1}^{22} \gamma_j^S D_j \quad (53)$$

Suivant ces deux expressions, chaque coefficient estimé  $\gamma_j^S$  devrait correspondre à l'expression ci-dessous :

$$\gamma_j^S = \left( X + \frac{\partial E(\varepsilon_S|X, y = 0)}{\partial B_S} \right) dB_S \quad (54)$$

Or, dans les équations de revenu, les coefficients  $\gamma_j^S$  représentent l'effet fixe d'appartenir au groupe j, et cet effet n'est pas conditionnel à la valeur de  $X$ . Par conséquent, ces coefficients captent en réalité l'effet partiel d'appartenir au groupe j sur l'espérance *non conditionnelle par rapport à X* du salaire, tel qu'illustré par l'équation suivante <sup>15</sup>:

$$\gamma_j^S = \left( \overline{X_j^S} + E_j^S \left( \frac{\partial E(\varepsilon_S|X, y = 0)}{\partial B_S} \right) \right) dB_S \quad (55)$$

où  $\overline{X_j^S}$  représente la moyenne de la variable  $X$  chez les travailleurs salariés du groupe j et où  $E_j^S$  représente l'espérance non conditionnelle de la fonction située entre parenthèse chez les travailleurs salariés du groupe j. En remplaçant la dérivée de l'espérance du terme d'erreur par son expression exacte, nous obtenons l'équation suivante :

---

<sup>15</sup> Nous utilisons la propriété  $E(ax)=aE(x)$  pour arriver à ce résultat.

$$\gamma_j^S = \left( \overline{X_j^S} + E_j^S \left( \frac{\sigma_E \sigma_S}{\sigma_v} \left( \rho - \frac{\sigma_S}{\sigma_E} \right) \frac{\partial \lambda_S(z)}{\partial z} \frac{\partial z}{\partial B_S} \right) \right) dB_S \quad (56)$$

Cette expression est équivalente à l'expression à suivante :

$$\gamma_j^S = \left( \overline{X_j^S} + \frac{\sigma_E \sigma_S}{\sigma_v} \left( \rho - \frac{\sigma_S}{\sigma_E} \right) E_j^S \left( \frac{\partial \lambda_S(z)}{\partial z} \frac{\partial z}{\partial B_S} \right) \right) dB_S \quad (57)$$

De manière analogue, nous obtenons comme expression pour l'effet partiel d'appartenir au groupe j sur l'espérance conditionnelle du revenu d'entreprise :

$$\gamma_j^E = \left( \frac{\sigma_E \sigma_S}{\sigma_v} \left( \frac{\sigma_E}{\sigma_S} - \rho \right) E_j^E \left( \frac{\partial \lambda_E(z)}{\partial z} \frac{\partial z}{\partial B_S} \right) \right) dB_S \quad (58)$$

À la lumière de ces deux expressions, nous constatons que lorsque le coefficient  $B_S$  diminue (i.e. lorsque  $dB_S$  est négatif), les deux coefficients prennent automatiquement une valeur négative, car pourvu que  $\rho$  soit *suffisamment positif*, tous les termes situés à l'intérieur de chacune des deux parenthèses sont strictement positifs. Ainsi, les relations 2 et 3 que nous devrions observer lorsqu'un groupe est victime de discrimination dans le marché du travail peuvent être reproduites à partir des estimés tirés des équations de revenu.

Pour résumer le tout, lorsqu'un groupe ethnique est victime de discrimination dans le marché du travail, nous devrions obtenir les résultats empiriques suivants :  $\widehat{\delta}_j > 0$ ,  $\widehat{\gamma}_j^S < 0$  et  $\widehat{\gamma}_j^E < 0$ . En d'autres mots, l'effet partiel d'appartenir à ce groupe sur la probabilité d'être entrepreneur devrait être positif, alors que les effets partiels sur le salaire et sur le revenu d'entreprise devraient tous deux être négatifs. Évidemment, comme nous l'avons mentionné plus tôt, d'autres phénomènes sociaux peuvent influencer le taux d'entrepreneuriat chez certains groupes ethniques par l'entremise de leur effet sur le revenu et sur le salaire d'entreprise. C'est le cas notamment de la discrimination dans le marché des produits (Borjas & Bronars, 1989) et des réseaux d'affaires propres à chaque communauté (Mandorff & Kerr, 2015) et (Edin & al., 2003). Bien entendu, nous pouvons appliquer le même cadre analytique à ces deux phénomènes sociaux.

Concernant la discrimination dans le marché des produits, plutôt que d'engendrer une baisse du coefficient  $B_S$ , celle-ci engendre une baisse du coefficient  $B_E$ , lequel représente le rendement des caractéristiques individuelles en termes de revenu d'entreprise. Pour connaître l'effet de cette baisse sur l'espérance conditionnelle du revenu dans chaque secteur, il convient de se référer aux équations (34) et (35). Nous savons que les dérivées  $\frac{\partial \lambda_E(z)}{\partial z}$  et  $\frac{\partial \lambda_S(z)}{\partial z}$  sont toutes les deux positives. Quant aux termes  $\frac{\partial z}{\partial B_E}$  et  $dB_E$ , ils sont tous les deux négatifs. Ainsi, l'effet de la discrimination dans le marché des produits sur la sélection, ou encore sur le niveau d'habileté moyen, est positif dans les deux secteurs. Ceci n'est pas surprenant : les entrepreneurs les moins habiles, qui sont pour la plupart plus habiles que les travailleurs salariés déjà actifs, quittent l'entrepreneuriat pour intégrer le travail salarié, entraînant du même coup une hausse du niveau d'habileté moyen dans chacun des deux secteurs<sup>16</sup>. Pour le salariat, la hausse du niveau d'habileté moyen constitue l'unique impact sur le revenu et fait ainsi en sorte que le salaire conditionnel espéré augmente. Ainsi, nous devrions obtenir comme résultat empirique  $\widehat{\gamma}_j^S > 0$ . Quant à l'espérance conditionnelle du revenu d'entreprise, l'effet de la hausse du niveau d'habileté est contrecarré par la baisse des rendements des caractéristiques individuelles. Dit autrement, l'effet de la sélection est contrecarré par « l'effet prix », qui dépend de  $X$ . En effet, lorsque nous ne contrôlons pas pour l'hétérogénéité des rendements entre les groupes ethniques, les coefficients ethniques que nous estimons dans l'équation de revenu d'entreprise devraient être représentés par l'expression suivante :

$$\gamma_j^E = \left( \overline{X}_j^E + E_j^E \left( \frac{\sigma_E \sigma_S}{\sigma_v} \left( \rho - \frac{\sigma_S}{\sigma_E} \right) \frac{\partial \lambda_E(z)}{\partial z} \frac{\partial z}{\partial B_E} \right) \right) dB_E \quad (59)$$

où  $\overline{X}_j^E$  représente la moyenne de la variable  $X$  chez les entrepreneurs du groupe  $j$  et où  $E_j^E$  représente l'espérance non conditionnelle de la fonction située entre parenthèse chez les entrepreneurs du groupe  $j$ . L'équation ci-haut est analogue à l'équation (57) et nous montre bien qu'il est difficile de connaître l'impact de la discrimination des entrepreneurs sur le revenu

---

<sup>16</sup> Il est important de rappeler que nous faisons ici référence à l'habileté conditionnelle aux caractéristiques observables.

d'entreprise de ces derniers, étant donné que le premier terme de la parenthèse est strictement positif alors que le deuxième terme est strictement négatif. Cependant, si nous incluons des termes d'interaction dans l'équation de revenu d'entreprise que nous estimons, nous pourrions conclure quant à l'effet de la baisse de  $B_E$  sur le revenu moyen observé chez les entrepreneurs. En effet, dans ce contexte, la fonction d'espérance conditionnelle découlant de l'équation de revenu d'entreprise que nous estimons prendrait la forme suivante :

$$E(w_E | X, D_j, y = 1) = XB_E + X \sum_{j=1}^{22} (B_E^j - B_E) D_j + \sum_{j=1}^{22} \gamma_j^E D_j \quad (60)$$

Dans cette équation, on voit bien que la valeur du coefficient  $\gamma_j^E$  ne serait pas influencée par la distribution des caractéristiques sociodémographiques au sein du groupe discriminé grâce au coefficient  $(B_E^j - B_E)$  qui viendrait capter « l'effet prix » de la baisse de  $B_E$ . En d'autres mots, la valeur du coefficient  $\gamma_j^E$  serait plutôt donnée par l'expression suivante :

$$\gamma_j^E = \left( \frac{\sigma_E \sigma_S}{\sigma_v} \left( \frac{\sigma_E}{\sigma_S} - \rho \right) E_j^E \left( \frac{\partial \lambda_E(z)}{\partial z} \frac{\partial z}{\partial B_E} \right) \right) dB_E \quad (61)$$

Toutefois, nous devons nous contenter d'estimer des coefficients  $\gamma_j^E$  dont le signe est ambigu en présence de discrimination dans le marché des produits, car l'inclusion de termes d'interaction pour chacun des vingt-deux groupes ethniques que nous étudions réduirait de beaucoup trop la puissance statistique de nos estimateurs. Pour finir, nous aimerions savoir quel sera l'impact de la baisse du coefficient  $B_E$  sur les coefficients estimés dans l'équation de choix sectoriel. Puisque pour tout  $X$ , la diminution de la valeur de  $B_E$  entraîne une hausse de  $z$ , le taux d'entrepreneuriat d'un groupe discriminé dans le marché des produits diminuera, toutes choses étant égales par ailleurs. Ce résultat est logique : pour n'importe quel niveau de  $X$ , moins d'individus auront le niveau d'habileté nécessaire pour demeurer entrepreneur. L'effet partiel d'appartenir au groupe  $j$  sur la probabilité d'être entrepreneur, lorsque les entrepreneurs de ce groupe sont victimes de discrimination, devrait alors être négatif ( $\widehat{\delta}_j < 0$ ).

Pour récapituler, la discrimination des entrepreneurs par l'entremise du marché des produits devrait faire augmenter le salaire moyen observé chez les travailleurs salariés faisant partie du groupe discriminé ( $\widehat{\gamma}_j^s > 0$ ) et faire diminuer le taux d'entrepreneuriat observé chez ce groupe ( $\widehat{\delta}_j < 0$ ). Quant à son effet sur le revenu d'entreprise moyen observé, celui-ci est ambigu, à moins d'inclure des termes d'interaction dans les équations estimés, chose que nous ne sommes malheureusement pas en mesure de faire.

En ce qui a trait aux réseaux d'affaires propres à chaque communauté<sup>17</sup>, nous présumons que leur existence entraîne une hausse des rendements sur les caractéristiques observables dans le domaine de l'entrepreneuriat, en raison de la meilleure performance des entreprises qui œuvrent au sein de ces derniers (Mandorff & Kerr, 2015). Évidemment, il est vrai que ces réseaux peuvent également affecter de manière positive les salaires des individus qui travaillent pour des entreprises qui en font partie (Mandorff & Kerr, 2015, Edin & al., 2003). Ainsi, plutôt que d'engendrer seulement une hausse des rendements chez les entrepreneurs, nous présumerons que ces réseaux engendrent une hausse *relative* des rendements dans le secteur de l'entrepreneuriat ( $\Delta(B_E - B_S) > 0$ ), car il est logique de croire que la plus-value des activités entrepreneuriales revient majoritairement à l'entrepreneur, et que celle-ci est reflétée par des retours plus élevés sur les attributs individuels. L'effet d'une hausse relative des rendements chez les entrepreneurs entraîne, tout comme la discrimination dans le marché du travail, une baisse de la différence conditionnelle espérée entre le salaire et le revenu d'entreprise (une baisse de  $z$ ). Ainsi, l'impact sur la sélection dans chacun des secteurs sera négatif. Toutefois, l'impact net sur le salaire et le revenu d'entreprise espérés sera ambigu en raison de « l'effet prix » qui est positif dans chaque secteur. En effet, si nous revenons au cadre théorique, l'impact des réseaux d'affaires sur le revenu dans chacun des secteurs est illustré par les deux équations suivantes:

$$dE(w_E|X, y = 1) = XdB_E + \frac{\sigma_E\sigma_S}{\sigma_v} \left( \frac{\sigma_E}{\sigma_S} - \rho \right) \frac{\partial \lambda_E(z)}{\partial z} \frac{\partial z}{\partial (B_S - B_E)} d(B_S - B_E) \quad (62)$$

---

<sup>17</sup> Par réseaux d'affaires, nous entendons également les enclaves, étant donné que l'impact de ces dernières sur le revenu des immigrants s'effectue majoritairement par l'entremise des réseaux d'affaires.

$$dE(w_S|X, y = 0) = XdB_S + \frac{\sigma_E\sigma_S}{\sigma_v} \left( \rho - \frac{\sigma_S}{\sigma_E} \right) \frac{\partial \lambda_S(z)}{\partial z} \frac{\partial z}{\partial (B_S - B_E)} d(B_S - B_E) \quad (63)$$

Par conséquent, sans refaire toute la démarche que nous avons faite pour la discrimination dans le marché du travail et la discrimination dans le marché des produits, nous savons d'avance que nous ne pouvons pas conclure quant au signe des coefficients estimés  $\widehat{\gamma}_j^E$  et  $\widehat{\gamma}_j^S$ , à moins que nous ne contrôlions pour l'hétérogénéité des rendements dans chaque équation de revenu, auquel cas nous devrions obtenir des coefficients négatifs pour chaque secteur. Enfin, puisque la valeur du terme  $z$  diminue conséquemment au fait que  $\Delta(B_E - B_S) > 0$ , l'effet partiel de l'ethnicité sur la probabilité d'être entrepreneur devrait augmenter en présence de réseaux d'affaires ( $\widehat{\delta}_j < 0$ ), tel que prédit par l'équation (45).

Une synthèse des résultats prédits par le cadre d'analyse que nous venons de présenter est disponible dans le tableau ci-dessous. Nous y ferons fréquemment référence dans la prochaine section, où nous déterminerons entre autres combien de groupes ethniques satisfont l'ensemble des critères inhérents à chacun des phénomènes qui s'y trouvent.

Coefficients	Phénomène Social		
	Discrimination (Marché du travail)	Discrimination (Marché des produits)	Réseaux d'affaires (industries, enclaves)
*Lorsque les termes d'interaction sont inclus			
$\delta_j$	>0	<0	>0
$\gamma_j^E$	<0	Ambigu (>0*)	Ambigu (<0*)
$\gamma_j^S$	<0	>0	Ambigu (<0*)



## V. Estimations et Résultats

### V.I. Équations de choix sectoriel

Les résultats des équations de choix sectoriel sont présentés dans le tableau 4. Ce tableau comporte quatre régressions, soit une régression des moindres carrés ordinaires (MPL) et une régression de type Probit pour chacune des deux générations. Pour chaque régression, on retrouve les estimés des coefficients associés aux variables dichotomiques qui représentent les différents groupes ethniques, de même que les estimés des coefficients associés aux différentes variables sociodémographiques. Ces variables sont l'âge, l'âge au carré (dont la présence modélise la non linéarité de l'effet de l'âge), le nombre d'années d'éducation, le nombre d'enfants, le statut marital, le fait d'être propriétaire d'un logement ainsi que le fait d'être victime d'un handicap physique ou mental quelconque. Alors que nous avons déjà amplement discuté de la pertinence des variables que sont le nombre d'années d'éducation, le nombre d'enfants et le statut marital dans la section III, il est important de mentionner que les variables liées au fait d'être propriétaire et au fait d'être aux prises avec un handicap sont largement utilisées dans la littérature, notamment dans Fairlie & Meyer (1996) et Clark & Drinkwater (2000). Pour cette raison, nous avons cru bon de les inclure pour éviter tout biais de variable omise. Toujours en observant le tableau 4, on remarque que les variables incluses dans le modèle sont les mêmes pour chaque génération, hormis la variable qui représente la probabilité d'avoir un père entrepreneur conditionnellement au groupe ethnique d'appartenance et à l'état de résidence. Cette variable est uniquement incluse pour les individus de la 2<sup>e</sup> génération et correspond, de manière plus précise, au taux d'entrepreneuriat chez les pères plausibles (i.e. les individus faisant partie de l'échantillon de la 1<sup>ère</sup> génération) qui appartenaient au même groupe ethnique que l'individu de 2<sup>e</sup> génération et qui résidaient dans le même état que ce dernier au moment de la conduite du recensement. Nous avons choisi l'état comme unité géographique pour deux raisons : d'une part, il s'agit de l'unité géographique la plus petite à laquelle nous avons accès dans l'échantillon du recensement que nous utilisons, et, d'autre part, les chances que les immigrants de 2<sup>e</sup> génération choisissent de résider dans le même état que leurs parents sont plutôt élevées lorsque leur communauté ethnique

a une forte présente dans cet état<sup>18</sup>. L'inclusion de cette variable permet d'atténuer le biais auquel les estimés des coefficients ethniques sont potentiellement assujettis en raison de la transmission intergénérationnelle des compétences entrepreneuriales dont nous avons discuté dans la section II.IV. La raison pour laquelle nous n'avons pas utilisé une variable semblable pour les individus de la 1<sup>ère</sup> génération est qu'il était impossible de trouver une variable qui était indicatrice de la probabilité d'avoir un père entrepreneur et qui n'était pas parfaitement corrélée avec le groupe ethnique d'appartenance. Néanmoins, l'omission de cette variable n'est selon nous pas trop problématique, car nous présumons que la transmission des compétences entrepreneuriales est beaucoup moins forte dans un contexte où les individus sont forcés de rompre avec leur pays d'origine. Pour appuyer cette hypothèse, nous nous basons sur les résultats de Fairlie & Meyer (1996), qui ne trouvent aucun lien entre le taux d'entrepreneuriat des communautés ethniques aux États-Unis et le taux d'entrepreneuriat dans leur pays d'origine.

Par ailleurs, chaque régression comporte des variables dichotomiques permettant de contrôler pour l'effet de la région de résidence. Pour la 1<sup>ère</sup> génération, une série de variables dichotomiques permettent de contrôler pour la cohorte d'immigration, c'est-à-dire la période lors de laquelle les immigrants de chaque pays sont arrivés. Chaque cohorte correspond à une période de cinq ans et l'inclusion des variables dichotomiques leur étant associées est surtout justifiée par le besoin de contrôler pour l'impact des conditions économiques aux États-Unis à l'arrivée des immigrants. Pour la 2<sup>e</sup> génération, une série de variables dichotomiques permettent de tenir compte de l'impact de la fluctuation annuelle de l'état de santé du marché du travail américain. Ainsi, les années 2007 à 2015 sont chacune associées à une variable dichotomique. Finalement, nous n'incluons pas de variables de contrôle pour l'industrie, car celles-ci pourraient engendrer un biais dans l'estimation des coefficients rattachés aux variables ethniques. En effet, selon Angrist & Pischke (2009), l'inclusion d'une variable de contrôle qui n'est pas prédéterminée par rapport à la variable d'intérêt peut entraîner un biais similaire au biais de sélection dans l'estimation du coefficient associé à cette variable. Dans le contexte de notre modèle, on ne peut affirmer que l'industrie dans lequel un individu travaille est prédéterminée par rapport au groupe ethnique auquel cet individu appartient. À vrai dire, dans de nombreux cas, l'industrie dans

---

<sup>18</sup> Voir Ellis & Goodwin-White (2006).

laquelle un immigrant travaille est fortement influencée par sa communauté d'appartenance, et ce, davantage encore lorsqu'un individu est entrepreneur (Mandorff & Kerr, 2015).

Le tableau 4 nous fournit également des informations sur la valeur des coefficients ainsi que sur leur niveau de significativité. On y remarque que pratiquement toutes les variables sociodémographiques ont un effet fortement significatif sur la probabilité d'être entrepreneur, et ce, à la fois dans la 1<sup>ère</sup> et dans la 2<sup>e</sup> génération. La seule exception à cette règle est le statut marital, dont le coefficient n'est pas significatif dans chacune des deux générations. Par conséquent, le fait que les taux de mariage soient beaucoup plus hétérogènes dans la 2<sup>e</sup> génération que dans la 1<sup>ère</sup> génération n'entraîne probablement pas une divergence importante des taux d'entrepreneuriat, contrairement à ce que nous présumions dans la section III.IV. De plus, on remarque aussi qu'alors que l'impact de l'éducation est positif pour les individus de la 1<sup>ère</sup> génération, celui-ci est négatif et de magnitude beaucoup plus faible pour les individus de la 2<sup>e</sup> génération. Cette particularité nous amène à conclure que la convergence des niveaux d'éducation n'est peut-être pas aussi importante que ce que la littérature suggère pour expliquer la convergence partielle des taux d'entrepreneuriat.

Compte tenu des deux constats que nous venons de faire, il est d'autant plus probable que les effets fixes associés aux différents groupes ethniques soient liés à l'évolution des taux d'entrepreneuriat observée entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. Or, pour en être sûr, il convient bien entendu de jeter un coup d'œil aux estimés des coefficients des différentes variables ethniques, eux aussi disponibles dans le tableau 4. On remarque tout d'abord une disparité entre le nombre de groupes ethniques qui affichent des coefficients significatifs dans la 1<sup>ère</sup> et dans la 2<sup>e</sup> génération d'immigrants. Alors que neuf groupes ethniques affichent des coefficients significatifs dans la 1<sup>ère</sup> génération, ce chiffre est de quatorze dans la 2<sup>e</sup> génération. Il est vrai que cette disparité peut être partiellement attribuable au fait que les écarts-types estimés à partir des données de l'échantillon de 2<sup>e</sup> génération (i.e. les données CPS) sont sous-estimés, tel que nous l'avons déjà mentionné dans la sous-section III.II. En effet, même si les coefficients associés aux communautés chinoise et mexicaine sont uniquement significatifs dans la 2<sup>e</sup> génération, les estimés de ces derniers sont plus petits dans la 2<sup>e</sup> génération que dans la 1<sup>ère</sup> génération. Or, dans le cas des communautés latino-américaine, britannique, portugaise et polonaise, il semble bel et

bien y avoir un accroissement de l'effet de l'ethnicité entre les deux générations, un résultat quelque peu contre-intuitif que nous analysons en détail dans la sous-section V.IV.

Quant au fait que seulement neuf groupes ethniques affichent un coefficient significatif dans la 1<sup>ère</sup> génération, ce résultat contredit les résultats de Fairlie & Meyer (1996), qui relevaient un effet significatif de l'ethnicité pour la majorité des communautés ethniques. Nous pourrions croire que la taille très petite des sous-échantillons associés à certains groupes ethniques a, dans certains cas, empêché de conclure à un effet significatif. En effet, il est vrai que pour les groupes dont l'effet partiel dans la population est faible, il peut être difficile de conclure à un effet significatif à partir de coefficients estimés sur de petits échantillons. Cette hypothèse est d'autant plus plausible que les écarts-types varient peu d'un groupe à l'autre. Cependant, il faut tout de même noter que plusieurs des groupes ethniques dont l'effet est considéré comme non significatif affichent des coefficients estimés très faibles. Puisque nous ne sommes pas certains que tous les niveaux de significativité que nous trouvons sont valides, nous conservons l'ensemble de nos estimés pour débiter l'analyse de l'évolution intergénérationnelle de l'effet de l'ethnicité, peu importe le niveau de significativité du coefficient auquel ils sont liés. Ensuite, nous écartons tous les groupes ethniques dont les coefficients ne sont pas significatifs dans la 1<sup>ère</sup> génération et comparons les résultats nouvellement obtenus avec les résultats que nous aurons préalablement obtenus au moyen de tous les estimés.

Par ailleurs, on constate, toujours grâce au tableau 4, que parmi les neuf groupes qui affichaient un coefficient significatif dans la 1<sup>ère</sup> génération, huit d'entre eux affichaient un coefficient positif. Une façon d'interpréter ce fait est que lorsque l'effet de l'ethnicité est bel et bien présent, celui-ci a tendance à être positif. En d'autres mots, les groupes ethniques pour lesquels la communauté a un effet sur la propension à devenir entrepreneur ont fortement tendance à afficher un taux d'entrepreneuriat supérieur à celui des natifs, toutes choses étant égales par ailleurs. Ces groupes ethniques sont les Cubains, les Grecs, les Italiens, les Allemands, les Coréens, les Iraniens, les Israélo-Palestiniens et les Libanais. Certains des résultats obtenus pour ces groupes sont d'ailleurs impressionnants : selon les estimés MPL, un homme israélo-palestinien de la 1<sup>ère</sup> génération aurait 28,7% plus de chances d'être entrepreneur qu'un natif de la même génération possédant les mêmes caractéristiques observables. Pour les hommes grecs, libanais et iraniens, ce

pourcentage se situe également au-dessus de la barre des 20%. Outre ces chiffres, on remarque aussi dans le tableau 4 que le seul groupe à afficher un effet partiel négatif dans la 1<sup>ère</sup> génération est celui des Philippins.

## V.II. L'évolution de l'effet de l'ethnicité sur le choix sectoriel entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants

Les graphiques 15 et 16 nous montrent ce qui se produit lorsque nous régressons l'ensemble des estimés des coefficients ethniques de la 2<sup>e</sup> génération sur l'ensemble des estimés des coefficients ethniques de la 1<sup>ère</sup> génération. Peu importe le type d'estimé (MPL ou Probit), on remarque qu'en moyenne, l'effet partiel de l'ethnicité semble plus important (en valeur absolue) dans la 1<sup>ère</sup> génération que dans la 2<sup>e</sup> génération. En effet, la majorité des points situés sur la partie droite du graphique se retrouvent sous sa bissectrice, alors que la majorité des points situés sur la partie gauche du graphique se retrouvent au-dessus de cette dernière. La pente de la droite des moindres carrés ordinaires qui traverse chacun des graphiques valide l'existence du phénomène de convergence, étant donné que celle-ci est nettement inférieure à 1 dans chacun des cas. Le tableau 6 affiche justement la valeur du coefficient de la droite des moindres carrés ordinaires qui apparaît sur chacun des deux graphiques. Sur la base des estimés du modèle de probabilité linéaire, on conclut qu'en moyenne, l'effet partiel de l'appartenance au groupe ethnique diminue, en valeur absolue, d'environ 63% entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants. Si l'on se fie aux estimés du modèle Probit, cette diminution moyenne est encore plus élevée, à environ 67%. De plus, la valeur de l'ordonnée à l'origine de chacune des pentes et le fait qu'elles ne soient toutes les deux pas statistiquement significatives confirment qu'il s'agit bien d'une convergence vers 0, un résultat qui nous fait dire que l'effet de l'ethnicité subit une atténuation, en moyenne, entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants. Toutefois, il faut se rappeler que nous ne tenons pas compte de l'incertitude liée aux estimés des coefficients ethniques lorsque nous utilisons la méthode des régressions linéaires, et que plusieurs des estimés inclus dans les régressions sont non significatifs.

Par opposition, lorsque nous conservons uniquement les groupes ethniques dont le coefficient est significatif au minimum dans l'une ou l'autre des deux générations, nous comparons, un groupe ethnique à la fois, la valeur des estimés des coefficients d'une génération à l'autre en tenant compte des écarts-types qui leur sont associés. Tous les groupes qui satisfont le critère énoncé plus haut sont présents dans le tableau 7. Les quatre premières colonnes du tableau reprennent tout simplement les estimés des coefficients ethniques de l'équation de choix sectoriel (tableau 4), alors que les colonnes 5 et 6 affichent, pour chaque groupe ethnique, la différence entre le coefficient estimé dans 2<sup>e</sup> génération et le coefficient estimé dans la 1<sup>ère</sup> génération (en %). Afin de rendre l'analyse plus simple, nous commençons par analyser uniquement les groupes dont le coefficient est significatif dans la 1<sup>ère</sup> génération (indépendamment de si leur coefficient est aussi significatif dans la 2<sup>e</sup> génération). Pour savoir si le coefficient de chacun de ces groupes est plus élevé dans la 1<sup>ère</sup> génération que dans la 2<sup>e</sup> génération, nous comparons l'intervalle de confiance associé au coefficient de la 1<sup>ère</sup> génération avec l'intervalle de confiance associé au coefficient de la 2<sup>e</sup> génération. Tel que mentionné dans la sous-section IV.II, ces intervalles sont calculés pour un seuil de confiance de 95%. En comparant les bornes inférieures et supérieures de ces intervalles, nous sommes en mesure de savoir si ces derniers se recoupent. S'ils se recoupent, nous ne pouvons pas conclure que les coefficients diffèrent d'une génération à l'autre alors que s'ils ne se recoupent pas, nous pouvons conclure qu'ils diffèrent. Les deux dernières colonnes du tableau 7 nous renseignent à ce sujet. Les valeurs qui figurent dans ces deux colonnes proviennent des deux dernières formules que nous avons présentées dans la sous-section IV.II (Équations (27) et (28)). Les valeurs négatives signifient que les intervalles se recoupent alors que les valeurs positives signifient qu'ils ne se recoupent pas. Ainsi, sur la base des estimés MPL (MCO), quatre groupes voient leurs coefficients différer de manière significative entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération : les Grecs, les Israélo-Palestiniens, les Coréens et les Allemands. Si l'on se fie plutôt aux estimés Probit, seulement trois groupes appartiennent à cette catégorie : les Grecs, les Israélo-Palestiniens et les Coréens. Or, puisque la méthode des intervalles de confiance est une méthode conservatrice, nous prendrons pour acquis que l'absence de recoupement dans un cas sur deux est suffisante pour conclure que les coefficients de la communauté allemande diffèrent de manière significative. Par conséquent, parmi les neuf groupes dont l'effet sur le choix sectoriel est significatif dans la 1<sup>ère</sup> génération, nous pouvons conclure que quatre d'entre eux voient leur effet changer entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. Dans tous les cas, il s'agit d'une atténuation, dans le

sens où le coefficient associé à la 2<sup>e</sup> génération est plus petit, en valeur absolue, que le coefficient associé à la 1<sup>ère</sup> génération. De plus, pour trois des quatre groupes, le coefficient conserve le même signe d'une génération à l'autre. L'unique exception à cette règle est l'Allemagne, dont le coefficient passe d'une valeur positive à une valeur légèrement négative. Cette dernière valeur n'est d'ailleurs pas significative et suggère donc la disparition totale de l'effet. Lorsque nous calculons la diminution moyenne de l'effet de l'ethnicité, en valeur absolue, chez les quatre groupes dont l'effet diminue de manière significative entre les deux générations, nous obtenons une diminution moyenne d'environ 64% selon les estimés MPL et d'environ 60% selon les estimés Probit, ce qui est considérable. De plus, ces chiffres sont très similaires aux résultats obtenus à l'aide des régressions linéaires, ce qui porte à croire que les pentes de ces dernières étaient fortement influencées par l'atténuation de l'effet de l'ethnicité chez les communautés grecque, coréenne, israélo-palestinienne et allemande.

Toutefois, il faut se rappeler que pour les cinq autres groupes ethniques dont l'effet est significatif dans la 1<sup>ère</sup> génération (Cuba, l'Italie, l'Iran, le Liban et les Philippines), nous ne pouvons pas conclure que l'effet de l'ethnicité diminue entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants. Également, il est important de se rappeler que les diminutions calculées (64% et 60%) ne représentent pas les écarts réels entre les deux populations d'immigrants (la 1<sup>ère</sup> et la 2<sup>e</sup> génération), mais bien uniquement les écarts estimés. Ainsi, à la lumière de nos résultats, la seule chose dont nous pouvons être sûrs est que la diminution de l'effet de l'ethnicité entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération est supérieur à 0%, en valeur absolue, pour quatre groupes ethniques parmi les neuf qui ont un effet significatif sur le choix sectoriel des immigrants de 1<sup>ère</sup> génération.

Finalement, pour les groupes ethniques dont l'effet sur la probabilité d'être entrepreneur est seulement significatif chez les immigrants de 2<sup>e</sup> génération, nous nous attendons à ce que la différence entre les coefficients associés à chacune des deux générations soit significative. Toutefois, pour en être sûr, il faut encore une fois comparer les intervalles de confiance des coefficients de la 1<sup>ère</sup> génération avec ceux des coefficients de la 2<sup>e</sup> génération. Tel que nous pouvons le voir dans le tableau 7, au seuil de confiance de 95%, seule la Pologne affiche une différence significative. Ce résultat complique notre interprétation des résultats, car il devient alors difficile de déterminer si les cinq autres groupes ethniques faisant partie de cette catégorie, à

savoir le Mexique, l'Amérique latine, le Royaume-Uni, le Portugal et la Chine, ont véritablement un effet sur la probabilité qu'ont les immigrants de 2<sup>e</sup> génération d'être entrepreneur. Or, si ces groupes n'affichent pas une différence significative, il semble que ce soit en raison du fait que contrairement à la Pologne, la significativité de leurs coefficients dans la 2<sup>e</sup> génération est davantage attribuable à la petite taille relative des écarts-types associés à ces coefficients qu'à la magnitude de ces derniers. En d'autres mots, la diminution de la taille relative des écarts-types entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération contribue davantage à l'augmentation de la significativité des coefficients que la hausse, en valeur absolue, des coefficients eux-mêmes. Ainsi, pour éviter d'aboutir à des conclusions erronées en raison de la sous-estimation des écarts-types calculés à partir des données CPS, nous prenons dorénavant pour acquis que la communauté polonaise est la seule à avoir un effet sur la probabilité d'être entrepreneur uniquement chez les immigrants de 2<sup>e</sup> génération. Cette mesure fait du même coup passer le nombre de groupes ethniques qui ont un effet significatif sur la probabilité d'être entrepreneur dans la 2<sup>e</sup> génération de quatorze à neuf.

### V.III Équations de revenu

Le tableau 8 présente les résultats des équations de revenu. Il est à noter que les équations de revenu comprennent les mêmes variables dichotomiques de contrôle (années, cohortes et régions) et les mêmes variables sociodémographiques que les équations de choix sectoriel, à l'exception de la variable représentant le fait d'être propriétaire d'un logement, qui est généralement exclue des équations de revenu dans la littérature. Le tableau 8 comporte quatre régressions, soit une équation qui explique le salaire et une équation qui explique le revenu d'entreprise pour chacune des deux générations d'immigrants. Tel qu'expliqué dans la sous-section IV.III, chaque équation est estimée par moindres carrés ordinaires. Dans chacune des quatre populations étudiées (entrepreneurs de la 1<sup>ère</sup> génération, travailleurs salariés de la 1<sup>ère</sup> génération, entrepreneurs salariés de la 2<sup>e</sup> génération et travailleurs salariés de la 2<sup>e</sup> génération), les coefficients associés aux variables sociodémographiques de contrôle sont tous significatifs au minimum au seuil de 5% ; la plupart d'entre eux l'étant au seuil de 0,1%. Un autre fait intéressant est que les rendements sur le niveau d'éducation semblent augmenter de manière considérable entre les deux générations, et ce, à la fois pour les entrepreneurs et les travailleurs salariés. Toutefois,



l'augmentation est plus importante pour les travailleurs salariés. Puisque nous croyons qu'il existe une corrélation positive entre l'éducation et le niveau d'habileté, il est probable que la hausse du coefficient lié à l'éducation reflète non seulement une augmentation des rendements sur l'éducation, mais également l'augmentation des rendements sur l'habileté dans le secteur salarié entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération, dont nous avons soupçonné l'existence dans la sous-section III.V.

Quant à l'effet d'appartenir aux différents groupes ethniques sur le revenu, on remarque rapidement que quelque soit la génération, beaucoup plus de groupes affichent des coefficients significatifs dans l'équation de salaire que dans l'équation de revenu d'entreprise. Puisque l'échantillon de 1<sup>ère</sup> génération comporte 525 179 travailleurs salariés et 67 195 entrepreneurs et que l'échantillon de 2<sup>e</sup> génération comporte 212 066 travailleurs salariés et 27 035 entrepreneurs, l'explication la plus probable à ce phénomène est que les sous-échantillons associés aux groupes ethniques sont beaucoup plus petits pour les entrepreneurs que pour les travailleurs salariés. Tel que nous pouvons l'observer dans le tableau 8, le résultat de cette disparité est qu'à effet estimé égal, les écarts-types associés aux estimés des entrepreneurs semblent en général beaucoup plus élevés que ceux associés aux estimés des travailleurs salariés, empêchant du même coup de trouver une relation statistiquement significative. Évidemment, nous ne pouvons pas être certains qu'en présence d'un échantillon d'entrepreneurs plus volumineux, nous trouverions un plus grand nombre de coefficients ethniques significatifs. Ce que nous affirmons, c'est que la petite taille des sous-échantillons diminue de manière considérable la puissance statistique de nos tests; dans ce cas-ci notre capacité à détecter l'effet de l'appartenance à un groupe ethnique sur le revenu d'un entrepreneur lorsque cet effet existe bel et bien. Dans le but de contrer ce problème (même s'il s'agit d'une méthode imparfaite), nous faisons abstraction du niveau de significativité des coefficients des équations de revenu d'entreprise lorsque nous analysons ces derniers en détail dans la prochaine sous-section.

Par ailleurs, on remarque également qu'un nombre élevé de groupes ethniques affiche un coefficient estimé statistiquement négatif dans l'équation de salaire, et ce, à la fois dans la 1<sup>ère</sup> et dans la 2<sup>e</sup> génération. De manière purement intuitive, ce résultat nous porte à croire que les immigrants sont victimes de discrimination dans le marché du travail. Or, nous avons vu dans la

sous-section IV.V que la discrimination dans le marché du travail implique normalement un processus de sélection, et que par conséquent, lorsqu'un groupe ethnique est victime de ce type de discrimination, son coefficient devrait être négatif à la fois dans l'équation de salaire et dans l'équation de revenu d'entreprise. Même si nous ne remarquons pas cette tendance dans le tableau 8, il est difficile de mettre de côté l'hypothèse de l'existence d'un certain niveau de discrimination sur la base de l'ethnicité dans le marché du travail, car lorsqu'on observe de plus près les coefficients ethniques chez les travailleurs salariés de la 1<sup>ère</sup> génération, on remarque que les groupes dont le coefficient est significativement inférieur à 0 sont tous des groupes d'origine non européenne, sauf un. En effet, les Mexicains, les Cubains, les Antillais, les Latino-Américains, les Chinois, les Coréens, les Philippins, les Indiens, les Iraniens, les Israélo-Palestiniens et les Africains affichent tous un salaire moyen significativement inférieur à celui des Américains de souche, toutes choses étant égales par ailleurs. Les seuls immigrants de 1<sup>ère</sup> génération d'origine européenne à se retrouver dans la même catégorie sont les Grecs. On remarque une situation similaire chez les immigrants de 2<sup>e</sup> génération, avec les Mexicains, les Cubains, les Antillais, les Latino-Américains, les Chinois, les Coréens, les Philippins et les Africains qui affichent toujours un salaire moyen significativement inférieur à celui des natifs. À la lumière de ces constats, il se pourrait que les prédictions du modèle de sélection que nous avons présentées dans les sous-sections IV.IV et IV.V soient erronées, notamment en raison de l'absence de tout processus de sélection significatif par rapport à l'habileté. En d'autres mots, certains groupes ethniques seraient fortement discriminés dans le marché du travail, mais cette discrimination n'aurait aucun impact sur le taux d'entrepreneuriat, et donc aucun impact sur le niveau d'habileté moyen dans chaque secteur. Nous explorons cette piste dans la prochaine sous-section en comparant les coefficients ethniques des équations de salaire et de revenu d'entreprise avec les coefficients ethniques des équations de choix sectoriel.

#### V.IV Les phénomènes à la source de l'effet de l'ethnicité sur le choix sectoriel

Notre objectif est maintenant de comprendre quels sont les phénomènes sociaux qui font en sorte que les membres de certains groupes ont, toutes choses étant égales par ailleurs, plus ou moins de chances de devenir entrepreneur que les Américains de souche. Pour ce faire, nous comparons les

estimés que nous avons obtenus pour chaque groupe ethnique aux prédictions que nous avons faites au moyen du modèle de sélection dans la sous-section IV.V. Nous effectuons cet exercice de manière séparée pour la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants. Ensuite, nous comparons les deux générations entre elles afin de voir si, pour chaque groupe ethnique, le même phénomène persiste d'une génération à l'autre. Concrètement parlant, nous employons une méthode d'analyse qui est très simple : pour chaque groupe ethnique, nous regardons si les signes des coefficients estimés correspondent à un des trois phénomènes sociaux que nous avons mis de l'avant dans la sous-section IV.V. Par exemple, un groupe victime de discrimination dans le marché du travail devrait normalement afficher un coefficient positif dans l'équation de choix sectoriel, mais négatif dans chacune des deux équations de revenu (salaire et revenu d'entreprise).

Le résultat de cette analyse est visible dans le tableau 9. Il est à noter que le tableau 9 ne contient que les résultats pour les groupes dont le taux d'entrepreneuriat chez les immigrants est, toutes choses étant égales par ailleurs, significativement différent de celui chez les natifs dans au moins une des deux générations. Nous avons choisi de nous limiter à ces groupes pour la raison bien simple que pour les autres groupes, nous ne pouvons pas conclure que l'ethnicité a véritablement un effet sur la probabilité d'être entrepreneur. La première chose que nous remarquons dans ce tableau est que la catégorie «Réseaux d'affaires» est celle à laquelle le plus grand nombre de groupes ethniques correspondent, peu importe la génération d'immigrants. En effet, dans la 1<sup>ère</sup> génération, pas moins de 7 groupes sur 9 appartiennent à cette catégorie, alors que dans la 2<sup>e</sup> génération, ce sont toujours 6 groupes sur 9 qui y appartiennent. Pour cette raison, nous pourrions déjà être tentés de conclure que le développement de réseaux d'affaires au sein des communautés ethniques est la principale raison pour laquelle certains groupes ethniques affichent un taux d'entrepreneuriat supérieur aux Américains de souche, même lorsque nous tenons compte de la distribution inégale des caractéristiques sociodémographiques d'une population à l'autre. Cependant, afin de garantir la robustesse de cette affirmation, il convient de vérifier si les groupes ethniques qui bénéficient des réseaux d'affaires dans la 1<sup>ère</sup> génération en bénéficient toujours dans la 2<sup>e</sup> génération. En procédant à cet exercice, nous constatons que 5 des 7 groupes pour lesquels l'existence de réseaux d'affaires était soupçonnée dans la 1<sup>ère</sup> génération font toujours l'objet des mêmes soupçons dans la 2<sup>e</sup> génération. Ainsi, dans la majorité des cas, le phénomène persiste entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération.

Quant aux 2 groupes pour lesquels l'effet des réseaux d'affaires semble disparaître d'une génération à l'autre, que leur arrive-t-il ? Dans le cas de la communauté allemande, ceci s'explique tout simplement par le fait que son effet sur la probabilité d'être entrepreneur disparaît entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. Or, dans le cas de la communauté cubaine, il semble que le phénomène social responsable de l'effet sur la probabilité d'être entrepreneur passe des réseaux d'affaires à la discrimination dans le marché du travail, selon les résultats du tableau 9. Pour bien comprendre pourquoi nous arrivons à ce résultat contre-intuitif, il est utile de se référer au tableau que nous avons présenté à la toute fin de la sous-section IV.V. On y voit que sur la base des signes des coefficients, tout groupe qui affiche les caractéristiques d'un groupe discriminé dans le marché du travail affiche également les caractéristiques d'un groupe au sein duquel se développent des réseaux d'affaires. Suivant cette logique, il se pourrait donc que la communauté cubaine soit dans les faits toujours sous l'influence des réseaux d'affaires. Dans ce contexte, le changement de catégorie voudrait simplement dire que dans la 2<sup>e</sup> génération, l'effet sélection (négatif) sur le revenu d'entreprise et sur le salaire l'emporte sur « l'effet prix » (positif), contrairement à dans la 1<sup>ère</sup> génération.

Or, il est également possible que la communauté cubaine soit bel et bien victime de discrimination dans le marché du travail. Pour comprendre pourquoi, nous devons tenir compte d'une réalité dont nous avons fait mention dans la dernière sous-section : la ségrégation de l'effet de l'ethnicité sur le salaire sur la base de la *race*. En effet, nous avons vu dans la dernière sous-section que les groupes ethniques dont l'effet sur le salaire était significativement négatif dans la 1<sup>ère</sup> génération sont tous d'origine non européenne, sauf un, soit les Grecs. Pour cette raison, nous avons conclu qu'il serait surprenant que la discrimination sur la base de l'ethnicité ne joue aucun rôle lorsque vient le temps d'expliquer pourquoi, toutes choses étant égales par ailleurs, les travailleurs salariés appartenant à certains groupes ont un salaire moyen plus faible que celui des natifs. Cette hypothèse est d'autant plus plausible que parmi ces groupes, beaucoup n'affichent pas, toutes choses étant égales par ailleurs, un taux d'entrepreneuriat supérieur au taux d'entrepreneuriat des natifs, ce qui signifie que la sélection négative des entrepreneurs en raison de la présence de réseaux d'affaires ne peut en principe pas expliquer le signe négatif du coefficient ethnique dans l'équation de salaire. C'est notamment le cas pour les Mexicains, les

Antillais, les Latino-Américains, les Philippins, les Chinois et les Indiens. De fait, nous concluons que ces groupes sont victimes de discrimination dans le marché du travail, mais qu'en moyenne, cette discrimination n'incite pas, pour des raisons que nous ignorons, les immigrants de ces groupes à devenir entrepreneur. Pour revenir au cas des Cubains, étant donné que ces derniers constituent une minorité visible au même titre que les groupes mentionnés plus haut, il serait surprenant qu'ils ne soient victimes d'aucune forme de discrimination dans le marché du travail. Par conséquent, nous émettons l'hypothèse que la présence de réseaux d'affaires au sein de cette communauté aurait permis, par l'entremise de son effet haussier sur les rendements dans les deux secteurs ( $B_E$  et  $B_S$ ), de masquer les effets négatifs de la discrimination dans le marché du travail sur le salaire (via la sélection et l'effet prix) et sur le revenu d'entreprise (via la sélection) lors de la 1<sup>ère</sup> génération. Autrement dit, les réseaux d'affaires ethniques auraient permis aux travailleurs salariés cubains d'obtenir des emplois à l'abri de la discrimination et aux entrepreneurs d'obtenir des rendements élevés sur leurs activités en dépit de leur faible niveau de talent en affaires (dans certains cas). Or, une diminution de la taille ou de la performance des réseaux d'affaires entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération effacerait l'effet compensatoire que procurent ces derniers. Ainsi, s'il fallait qu'un certain niveau de discrimination dans le marché du travail subsiste entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération, la diminution de la taille ou de la performance des réseaux d'affaires pourrait expliquer ce que nous observons dans le cas des Cubains.

En observant les coefficients associés à plusieurs minorités visibles dans l'équation de salaire de la 2<sup>e</sup> génération, nous constatons que dans le cas des Mexicains, des Antillais, des Latino-Américains, des Chinois, des Philippins et des Africains, les coefficients sont, contrairement aux coefficients des autres groupes, fortement négatifs. Puisqu'il ne s'agit pas de groupes dont les membres sont plus susceptibles d'être entrepreneurs que les natifs (toutes choses étant égales par ailleurs), nous pouvons présumer que les signes de ces coefficients ne sont pas le résultat d'un processus de sélection négative, et qu'ils témoignent encore une fois d'un certain niveau de discrimination sur la base de l'ethnicité. Par conséquent, nous pouvons conclure qu'il y a de fortes chances que les travailleurs salariés cubains de la 2<sup>e</sup> génération soient victimes d'un certain niveau de discrimination. De la même façon, il y a de fortes chances que les travailleurs salariés iraniens, israélo-palestiniens et libanais de la 2<sup>e</sup> génération soient eux-aussi victimes de discrimination. Si ces groupes ne subissent aucun changement entre les deux générations dans le

tableau 9, nous croyons que c'est parce que la taille et la performance de leurs réseaux d'affaires demeurent suffisamment élevées dans la 2<sup>e</sup> génération pour qu'ils puissent contrer l'effet négatif de la discrimination dans le marché du travail sur le revenu moyen des entrepreneurs et des travailleurs salariés. En effet, selon le modèle MPL, les immigrants de 2<sup>e</sup> génération iraniens, israélo-palestiniens et libanais affichent, dans l'ordre respectif, un taux d'entrepreneuriat 14%, 11% et 9% plus élevé que celui des natifs, toutes choses étant égales par ailleurs. Par opposition, ce chiffre n'est que de 5% pour la communauté cubaine.

Le raisonnement que nous avons développé dans les derniers paragraphes nous permet aussi de mieux comprendre la situation de la communauté coréenne. Le tableau 9 suggère que celle-ci possède, toutes choses étant égales par ailleurs, un taux d'entrepreneuriat plus élevé que celui des natifs en raison du fait que ses travailleurs salariés sont discriminés, et ce, à la fois dans 1<sup>ère</sup> et dans la 2<sup>e</sup> génération. Or, à la lumière de l'analyse que nous venons de faire, cela ne voudrait pas dire qu'il n'existe pas de réseaux d'affaires propres à la communauté coréenne, bien au contraire. En effet, compte tenu du nombre élevé de groupes ethniques qui, à caractéristiques égales, affichent un salaire moyen inférieur à celui des Américains de souche et qui n'affichent pas un taux d'entrepreneuriat supérieur à ces derniers, nous avons de bonnes raisons de croire que la discrimination dans le marché du travail n'entraîne généralement pas à elle seule une hausse significative du taux d'entrepreneuriat chez les communautés affectées. Suivant cette logique, le taux d'entrepreneuriat élevé observé chez les immigrants coréens de 1<sup>ère</sup> génération et de 2<sup>e</sup> génération serait donc surtout attribuable à la présence de réseaux d'affaires. Au même moment, le signe négatif du coefficient associé à la Corée dans l'équation de salaire et dans l'équation de revenu d'entreprise serait attribuable au fait que l'impact positif des réseaux d'affaires (leur effet prix) ne compense pas complètement l'effet négatif engendré par la discrimination dans le marché du travail (via la sélection et le prix).

Quant à la communauté polonaise, le tableau 9 suggère que celle-ci a un effet sur la probabilité d'être entrepreneur à partir de la 2<sup>e</sup> génération, et que cet effet résulte de l'émergence de réseaux d'affaires entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération. Comme de fait, on remarque que dans l'équation de salaire de la 2<sup>e</sup> génération, le coefficient associé à la Pologne n'est pas négatif. Ces résultats appuient notre hypothèse selon laquelle les minorités visibles seraient victimes de discrimination

dans le marché du travail alors que les groupes ethniques d'origine européenne ne le seraient pas. De plus, le fait qu'il faille attendre la 2<sup>e</sup> génération pour voir l'effet de l'ethnicité se manifester valide en quelque sorte l'hypothèse de l'existence de réseaux d'affaires. En effet, on imagine mal la discrimination ou les facteurs culturels se manifester uniquement chez les immigrants de 2<sup>e</sup> génération, tandis qu'il paraît tout à fait naturel qu'un réseau de commerces ou qu'une grappe industrielle puisse nécessiter une génération complète pour voir le jour.

Finalement, en ce qui concerne la communauté philippine, nous savons que celle-ci est la seule à afficher un taux d'entrepreneuriat inférieur à celui des Américains de souche, toutes choses étant égales par ailleurs. Or, à la lumière des coefficients estimés dans les équations de revenu, cette particularité n'est ni dans la 1<sup>ère</sup> ni dans la 2<sup>e</sup> génération attribuable à de la discrimination dans le marché de produits. Comme il s'agit du seul groupe ethnique à se retrouver dans cette situation, il est très probable que des facteurs culturels ou autres propres à cette communauté expliquent la plus faible propension de ses membres à devenir entrepreneur.

Il est à noter que lorsque nous attribuons l'effet positif de l'ethnicité sur le taux d'entrepreneuriat (i.e. sur la propension à devenir entrepreneur) à l'existence de réseaux d'affaires au sein des groupes ethniques concernés, nous ne pouvons en principe pas être certains de la validité de notre conclusion. En effet, lorsque l'effet d'un groupe ethnique sur le taux d'entrepreneuriat est positif, nous attribuons par défaut ce dernier à l'existence de réseaux d'affaires dès que les critères associés à la discrimination dans le marché du travail ne sont pas satisfaits. Pour cette raison, certains pourraient croire que l'effet positif de l'ethnicité est majoritairement le produit de facteurs culturels qui n'ont rien à voir avec l'existence d'enclaves, de grappes industrielles ou de toute autre forme de réseau d'affaires à caractère ethnique. Cependant, nous pensons que cette hypothèse est très peu plausible pour deux principales raisons. Dans un premier temps, les groupes dont l'effet sur la propension à devenir entrepreneur est significativement positif ne partagent pas d'héritage culturel commun. En effet, les cinq groupes pour lesquels l'effet est le plus important, c'est-à-dire les Israélo-Palestiniens, les Grecs, les Libanais, les Iraniens et les Coréens, ne partagent ni religion commune, ni bagage ethnoculturel a priori commun. Dans un deuxième temps, il serait selon nous très surprenant que les membres de certains groupes aient jusqu'à 23%, 25% et 29% plus de chances d'être entrepreneur que les Américains de souche sans

que ces probabilités ne soient de quelque manière que ce soit liées à certains facteurs économiques.



## VI. Conclusion

Dans ce mémoire, nous avons étudié l'effet de l'ethnicité sur la probabilité d'être entrepreneur, son évolution entre la 1<sup>ère</sup> et la 2<sup>e</sup> génération d'immigrants, ainsi que les différentes causes possibles de cet effet. Nous sommes arrivés à la conclusion que tous les groupes ethniques n'ont pas un effet sur la probabilité d'être entrepreneur, et que lorsque cet effet est présent, celui-ci est généralement positif. En effet, parmi les vingt-deux groupes ethniques que nous avons étudiés, neuf d'entre eux ont un effet sur la probabilité d'être entrepreneur chez les immigrants de 1<sup>ère</sup> génération, et dans huit cas sur neuf, cet effet est positif. Ces résultats contredisent les résultats obtenus par Fairlie & Meyer (1996), qui relevaient dans leur étude un effet significatif pour la quasi-totalité des groupes ethniques qu'ils étudiaient. Tel que nous l'avons déjà mentionné, il est possible que la non significativité de certains des coefficients estimés soit reliée à la petite taille des sous-échantillons associés à certains groupes ethniques. Toutefois, dans plusieurs cas, la faiblesse des coefficients estimés peut justifier à elle seule l'absence de significativité.

Par ailleurs, parmi les neuf groupes qui ont un effet sur le choix sectoriel des immigrants de 1<sup>ère</sup> génération, huit d'entre eux voient leur effet persister jusqu'à la 2<sup>e</sup> génération. De manière plus précise, quatre groupes voient leur effet diminuer mais demeurer significatif, quatre groupes voient leur effet rester approximativement le même, et un seul groupe voit son effet disparaître. Par conséquent, il semble qu'en général, les immigrants de 2<sup>e</sup> génération continuent de se comporter différemment des Américains de souche dans le marché du travail, du moins en ce qui a trait au choix sectoriel.

Finalement, nous concluons que pour les groupes ethniques dont l'effet sur la probabilité d'être entrepreneur est positif, la prolifération de réseaux d'affaires au sein des communautés est la cause la plus probable de l'effet observé, et ce, à la fois pour les immigrants de 1<sup>ère</sup> et de 2<sup>e</sup> génération. Alors que les résultats que nous avons obtenus suggèrent également que beaucoup de groupes ethniques sont victimes de discrimination dans le marché du travail, rien ne nous indique que cette discrimination est une cause directe des taux d'entrepreneuriat anormalement élevés qui caractérisent certaines communautés. Ces conclusions concordent avec celles de Fairlie & Meyer

(1996), qui affirment que l'effet de l'ethnicité sur la probabilité d'être entrepreneur est lié non pas à la discrimination des travailleurs salariés, mais bien à l'existence de retours exceptionnels sur l'entrepreneuriat chez certains groupes. Toutefois, elles contredisent celles de Clark & Drinkwater (2000), dont l'étude suggère que la discrimination salariale a un impact positif sur la propension à devenir entrepreneur. Quant au seul groupe dont l'effet sur la probabilité d'être entrepreneur est négatif, l'absence de tout lien avec les phénomènes sociaux que nous avons identifiés comme étant des causes potentielles nous a amenés à conclure que cet effet est attribuable à des facteurs culturels propres à ce groupe.

Parmi les limitations de ce mémoire, il est important de rappeler que notre analyse ne vise que les travailleurs de sexe masculin, et que par conséquent, nous ne pouvons pas généraliser à l'ensemble de la population de travailleurs immigrants nos résultats. De plus, la faible précision de certaines de nos estimations fait en sorte que les conclusions auxquelles nous sommes arrivés ne sont pas dépourvues de tout jugement arbitraire, notamment en ce qui concerne la significativité de certains coefficients.

Enfin, nous terminons en proposant deux avenues de recherche pouvant découler de ce mémoire. Dans un premier temps, il pourrait être intéressant de valider l'existence des réseaux d'affaires au moyen de données détaillées sur les industries dans lesquelles œuvrent les entrepreneurs et les travailleurs salariés, et d'identifier le type d'industrie dans lequel les réseaux prolifèrent le mieux. Dans un deuxième temps, puisque nous soupçonnons que, parmi les groupes ethniques dont l'effet sur la probabilité d'être entrepreneur est attribué aux réseaux d'affaires, plusieurs sont victimes de discrimination dans le marché du travail, il pourrait être intéressant de vérifier s'il existe quelque lien que ce soit entre le niveau de discrimination des travailleurs salariés d'un groupe ethnique et la tendance des membres de ce groupe à se lancer en affaires tout en se concentrant dans certaines industries.

## Bibliographie

Angrist, Joshua David et Jörn-Steffen Pischke. 2009. « *Mostly harmless econometrics: An empiricist's companion.* » Princeton: Princeton University Press.

Aydemir, Abdurrahman, Wen-Hao Chen et Miles Corak. 2009. « Intergenerational earnings mobility among the children of canadian immigrants. » *The Review of Economics and Statistics* 91 (2): 377-97.

Borjas, George J. et Stephen G. Bronars. 1989. « Consumer discrimination and self-employment. » *Journal of Political Economy* 97 (3): 581-605.

Borjas, George J. 1987. « Self-selection and the earnings of immigrants. » *The American Economic Review* 77 (4): 531-53.

Borjas, George J. 1986. « The self-employment experience of immigrants. » *The Journal of Human Resources*, 21 (4): 485-506.

Card, David, John DiNardo et Eugena Estes. 1998. « The more things change: Immigrants and the children of immigrants in the 1940s, the 1970s, and the 1990s. » National Bureau of Economic Research, No. w6519.

Clark, Kenneth et Stephen Drinkwater. 2000. « Pushed out or pulled in? Self-employment among ethnic minorities in england and wales. » *Labour Economics* 7 (5): 603-28.

Dunn, Thomas et Douglas Holtz-Eakin. 2000. « Financial capital, human capital, and the transition to Self-Employment: Evidence from intergenerational links. » *Journal of Labor Economics* 18 (2): 282-305.

Edin, Per-Anders, Peter Fredriksson et Olof Åslund. 2003. « Ethnic enclaves and the economic success of immigrants: Evidence from a natural experiment. » *The Quarterly Journal of Economics* 118 (1): 329-57.

Ellis, Mark et Jamie Goodwin-White. 2006. « 1.5 generation internal migration in the U.S.: Dispersion from states of immigration? » *The International Migration Review* 40 (4): 899-926.

Fairlie, Robert W. et Bruce D. Meyer. 1996. « Ethnic and racial self-employment differences and possible explanations. » *The Journal of Human Resources* 31 (4): 757-93.

Hou, Feng, Teresa Abada et Yuqian Lu. 2013. « Bosses of their own: Are the children of immigrants more likely to be self-employed than their parents? » *Journal of Ethnic and Migration Studies* 39 (8): 1319-35.

Kerr, William R. et Martin Mandorff. 2015. « Social Networks, Ethnicity and Entrepreneurship » National Bureau of Economic Research, No. 21597.

Lemieux, Thomas, W. Bentley Macleod et Daniel Parent. 2009. « Performance pay and wage inequality. » *The Quarterly Journal of Economics* 124 (1): 1-49.

Roy, A. D. 1951. « Some thoughts on the distribution of earnings. » *Oxford Economic Papers* 3 (2): 135-46.

Schenker, Nathaniel et Jane F. Gentleman. 2001. « On judging the significance of differences by examining the overlap between confidence intervals. » *The American Statistician* 55 (3): 182-6.

## Annexe 1 : Tableaux

**Tableau 1 : Taux d'entrepreneuriat et moyennes de différentes variables sociodémographiques**

	Taux d'entrepreneuriat (%)		Éducation (années)		Taux de Mariage (%)		Nombre d'enfants	
	1ère gén.	2e gén.	1ère gén.	2e gén.	1ère gén.	2e gén.	1ère gén.	2e gén.
Natifs	11.3	10.7	14.6	15.1	99.9	62.2	2.0	1.1
Canada	16.3	12.0	14.2	16.7	100	68.0	2.0	1.1
Mexique	4.8	8.0	8.2	11.4	99.8	65.2	2.9	1.5
Cuba	20.5	16.5	13.0	14.8	99.9	66.8	1.8	1.0
Antilles	5.7	9.1	12.8	14.2	99.9	56.1	2.1	1.1
Amérique du Sud	11.0	13.8	13.8	14.4	100	61.8	2.1	1.0
Royaume-Uni	11.4	8.8	16.5	16.5	100	74.8	1.9	1.0
Irlande	10.2	13.7	12.8	16.4	100	64.4	2.7	1.1
Grèce	35.3	25.4	11.2	16.0	100	52.4	2.1	0.8
Italie	20.4	20.5	9.7	15.8	100	69.8	2.2	1.2
Portugal	5.7	16.2	7.1	14.3	100	64.5	1.9	1.0
Allemagne	20.9	11.3	14.6	17.0	100	70.1	1.8	1.1
Pologne	7.1	19.4	12.4	15.5	100	68.5	1.8	1.0
Yougoslavie	11.9	13.5	11.7	14.7	100	73.8	2.0	1.3
Chine	16.4	9.6	16.4	17.2	100	69.9	1.9	0.9
Japon	8.7	9.1	17.1	17.3	100	75.3	1.8	0.7
Corée	29.7	19.7	17.0	17.4	100	66.5	2.0	1.0
Philippines	6.8	6.8	16.6	15.8	100	63.5	2.1	1.1
Inde	11.5	10.2	19.1	18.1	100	81.7	1.8	1.1
Iran	34.9	28.6	18.2	16.7	100	61.7	1.8	0.8
Israël & Palestine	39.2	25.8	15.6	16.3	100	66.9	2.4	1.6
Liban	33.0	22.7	12.8	15.5	100	77.0	2.3	1.5
Afrique	15.2	10.0	18.3	15.7	100	62.6	2.0	1.1

\*Consulter la section III.IV afin de comprendre pourquoi certains groupes ethniques affichent un taux de mariage de 100% alors que d'autres non chez les immigrants de 1<sup>ère</sup> génération

**Tableau 2: Revenu moyen par groupe (k\$ de 2006)**

	1ère génération		2e génération	
	Entreprise	Salariat	Entreprise	Salariat
Natifs	62.6	47.3	60.6	50.3
Canada	83.3	60.1	87.6	70.9
Mexique	42.1	30.4	31.5	29.4
Cuba	56.7	40.6	49.2	41.8
Antilles	46.5	32.9	34.7	36.9
Amérique du Sud	57.9	38.2	44.9	41.1
Royaume-Uni	81.0	65.7	69.6	80.7
Irlande	59.2	50.6	64.7	78.1
Grèce	49.2	38.6	57.3	67.2
Italie	43.3	40.9	68.1	58.6
Portugal	50.0	34.2	57.6	52.1
Allemagne	68.3	63.1	76.6	70.0
Pologne	47.0	45.5	48.3	53.6
Yougoslavie	50.6	46.8	45.2	54.1
Chine	61.0	45.7	58.7	58.1
Japon	60.1	66.3	53.3	69.4
Corée	61.0	44.8	62.1	60.2
Philippines	100.1	39.9	75.9	47.5
Inde	94.2	54.0	77.7	69.6
Iran	88.5	46.8	115.9	59.6
Israël & Palestine	61.8	41.8	49.2	74.7
Liban	52.7	45.6	59.2	52.2
Afrique	77.0	46.5	52.0	43.7

**Tableau 3: Décomposition de la variance des revenus**

---

	<b>Nat. 1ère</b>	<b>Nat. 1ère (Ctf)</b>	<b>Nat. 1ère (Eff)</b>	<b>Nat. 2e</b>	<b>Nat. 2e (Ctf)</b>	<b>Nat. 2e (Eff)</b>
Var (XB) - Entrepreneurs	0.105	0.038	0.067	0.141	0.114	0.026
Var (XB) - Salariés	0.039	0.039	0.000	0.112	0.112	0.000
Var (e) - Entrepreneurs	0.566	0.233	0.332	0.758	0.334	0.424
Var (e) - Salariés	0.233	0.233	0.000	0.334	0.334	0.000
	<b>Imm. 1ère</b>	<b>Imm. 1ère (Ctf)</b>	<b>Imm. 1ère (Eff)</b>	<b>Imm. 2e</b>	<b>Imm. 2e (Ctf)</b>	<b>Imm. 2e (Eff)</b>
Var (XB) - Entrepreneurs	0.134	0.099	0.035	0.190	0.187	0.002
Var (XB) - Salariés	0.088	0.088	0.000	0.177	0.177	0.000
Var (e) - Entrepreneurs	0.566	0.281	0.285	0.709	0.365	0.344
Var (e) - Salariés	0.279	0.279	0.000	0.365	0.365	0.000

---

**Tableau 4: Déterminants du choix de devenir entrepreneur (coefficients)**

	1ère génération		2e génération	
	MCO	Probit	MCO	Probit
<b>Variabes ethniques</b>				
Canada	0.045 (0.029)	0.190 (0.140)	0.002 (0.016)	-0.007 (0.080)
Mexique	-0.022 (0.024)	-0.270* (0.120)	-0.013*** (0.003)	-0.095*** (0.020)
Cuba	0.085** (0.026)	0.340** (0.130)	0.046*** (0.012)	0.210*** (0.051)
Antilles	-0.029 (0.025)	-0.260* (0.130)	0.005 (0.007)	0.015 (0.041)
Amérique du Sud	0.013 (0.025)	0.034 (0.130)	0.042*** (0.007)	0.220*** (0.033)
Royaume-Uni	-0.013 (0.027)	-0.091 (0.140)	-0.028* (0.012)	-0.160* (0.077)
Irlande	-0.011 (0.030)	-0.091 (0.160)	0.025 (0.027)	0.100 (0.130)
Grèce	0.250*** (0.029)	0.900*** (0.130)	0.110** (0.035)	0.410*** (0.110)
Italie	0.120*** (0.026)	0.490** (0.120)	0.076*** (0.020)	0.300*** (0.073)
Portugal	0.002 (0.025)	-0.110 (0.140)	0.065* (0.026)	0.310** (0.110)
Allemagne	0.082** (0.030)	0.310* (0.140)	-0.019 (0.018)	-0.110 (0.099)
Pologne	-0.028 (0.027)	-0.210 (0.150)	0.110*** (0.021)	0.490*** (0.080)
Yougoslavie	0.023 (0.028)	0.088 (0.140)	0.035 (0.022)	0.160 (0.100)
Chine	0.035 (0.025)	0.130 (0.120)	-0.017* (0.007)	-0.100* (0.041)
Japon	-0.010 (0.027)	-0.072 (0.140)	-0.015 (0.025)	-0.089 (0.170)
Corée	0.180*** (0.027)	0.630*** (0.130)	0.066*** (0.013)	0.260*** (0.050)
Philippines	-0.062** (0.024)	-0.400** (0.130)	-0.036*** (0.007)	-0.240*** (0.052)
Inde	-0.016	-0.100	0.009	0.043



	(0.025)	(0.130)	(0.005)	(0.031)
Iran	0.220 <sup>***</sup> (0.042)	0.740 <sup>***</sup> (0.150)	0.140 <sup>***</sup> (0.029)	0.490 <sup>***</sup> (0.091)
Israël & Palestine	0.290 <sup>***</sup> (0.045)	0.990 <sup>***</sup> (0.160)	0.110 <sup>**</sup> (0.041)	0.390 <sup>**</sup> (0.130)
Liban	0.230 <sup>***</sup> (0.055)	0.860 <sup>***</sup> (0.190)	0.091 <sup>*</sup> (0.037)	0.330 <sup>*</sup> (0.130)
Afrique	0.026 (0.029)	0.110 (0.140)	0.004 (0.007)	0.012 (0.042)
<b>Variables sociodémographiques</b>				
Éducation (années)	0.006 <sup>***</sup> (0.000)	0.030 <sup>***</sup> (0.001)	-0.001 <sup>**</sup> (0.000)	-0.006 <sup>***</sup> (0.002)
Âge	0.008 <sup>***</sup> (0.001)	0.052 <sup>***</sup> (0.003)	0.007 <sup>***</sup> (0.001)	0.072 <sup>***</sup> (0.008)
Âge <sup>2</sup>	-0.006 <sup>***</sup> (0.001)	-0.044 <sup>***</sup> (0.003)	-0.004 <sup>*</sup> (0.002)	-0.066 <sup>***</sup> (0.010)
Marié	-0.016 (0.011)	-0.086 (0.060)	0.001 (0.002)	0.007 (0.011)
Enfants (Nombre)	0.005 <sup>***</sup> (0.000)	0.023 <sup>***</sup> (0.002)	0.007 <sup>***</sup> (0.001)	0.038 <sup>***</sup> (0.004)
Handicap	0.013 <sup>***</sup> (0.002)	0.071 <sup>***</sup> (0.011)	0.033 <sup>***</sup> (0.008)	0.170 <sup>***</sup> (0.035)
Propriétaire	0.028 <sup>***</sup> (0.001)	0.170 <sup>***</sup> (0.006)	0.023 <sup>***</sup> (0.002)	0.140 <sup>***</sup> (0.011)
Taux d'entrepreneuriat (pères du même état)			0.170 <sup>***</sup> (0.017)	0.870 <sup>***</sup> (0.082)
Années	Non	Non	Oui	Oui
Cohortes	Oui	Oui	Non	Non
Régions	Oui	Oui	Oui	Oui
Constante	-0.210 <sup>***</sup> (0.014)	-3.100 <sup>***</sup> (0.079)	-0.140 <sup>***</sup> (0.024)	-3.280 <sup>***</sup> (0.150)
Observations	592374	592374	239101	239101

Écart-types entre parenthèses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Tableau 5: Déterminants du choix de devenir entrepreneur (effets marginaux moyens)**

	1ère génération		2e génération	
	MCO	Probit	MCO	Probit
<b>Variabes ethniques</b>				
Canada	0.045 (0.029)	0.035 (0.025)	0.002 (0.016)	-0.005 (0.014)
Mexique	-0.022 (0.024)	-0.049* (0.023)	-0.013*** (0.003)	-0.016*** (0.004)
Cuba	0.085** (0.026)	0.064** (0.023)	0.046*** (0.012)	0.038*** (0.009)
Antilles	-0.029 (0.025)	-0.047 (0.024)	0.005 (0.007)	0.004 (0.007)
Amérique du Sud	0.013 (0.025)	0.007 (0.024)	0.042*** (0.007)	0.038*** (0.006)
Royaume-Uni	-0.013 (0.030)	-0.018 (0.026)	-0.028* (0.012)	-0.030* (0.014)
Irlande	-0.011 (0.029)	-0.014 (0.029)	0.025 (0.027)	0.019 (0.022)
Grèce	0.250*** (0.029)	0.170*** (0.024)	0.110** (0.035)	0.071*** (0.020)
Italie	0.120*** (0.026)	0.092*** (0.023)	0.076*** (0.020)	0.051*** (0.013)
Portugal	0.002 (0.025)	-0.020 (0.026)	0.065* (0.026)	0.056** (0.020)
Allemagne	0.082** (0.030)	0.057* (0.025)	-0.019 (0.018)	-0.020 (0.017)
Pologne	-0.028 (0.027)	-0.039 (0.028)	0.110*** (0.021)	0.087*** (0.014)
Yougoslavie	0.023 (0.028)	0.016 (0.027)	0.035 (0.022)	0.031 (0.018)
Chine	0.035 (0.025)	0.027 (0.023)	-0.017* (0.007)	-0.017* (0.007)
Japon	-0.010 (0.027)	-0.018 (0.027)	-0.015 (0.025)	-0.019 (0.030)
Corée	0.180*** (0.027)	0.120*** (0.023)	0.066*** (0.013)	0.049*** (0.009)
Philippines	-0.062** (0.024)	-0.071** (0.023)	-0.036*** (0.007)	-0.044*** (0.009)
Inde	-0.016	-0.011	0.009	0.008

	(0.025)	(0.023)	(0.005)	(0.006)
Iran	0.220 <sup>***</sup> (0.042)	0.130 <sup>***</sup> (0.028)	0.140 <sup>***</sup> (0.029)	0.087 <sup>***</sup> (0.016)
Israël & Palestine	0.290 <sup>***</sup> (0.045)	0.180 <sup>***</sup> (0.029)	0.110 <sup>**</sup> (0.041)	0.075 <sup>**</sup> (0.023)
Liban	0.230 <sup>***</sup> (0.055)	0.160 <sup>***</sup> (0.034)	0.091 <sup>*</sup> (0.037)	0.060 <sup>*</sup> (0.023)
Afrique	0.026 (0.029)	0.024 (0.026)	0.004 (0.007)	0.001 (0.007)

### Variables sociodémographiques

Éducation (années)	0.006 <sup>***</sup> (0.001)	0.006 <sup>***</sup> (0.000)	-0.001 <sup>**</sup> (0.000)	-0.001 <sup>***</sup> (0.000)
Âge	0.008 <sup>***</sup> (0.001)	0.010 <sup>***</sup> (0.001)	0.007 <sup>***</sup> (0.001)	0.013 <sup>***</sup> (0.001)
Âge <sup>2</sup>	-0.006 <sup>***</sup> (0.001)	-0.008 <sup>***</sup> (0.001)	-0.004 <sup>*</sup> (0.002)	-0.012 <sup>***</sup> (0.002)
Marié	-0.016 (0.011)	-0.011 (0.011)	0.001 (0.002)	0.002 (0.002)
Enfants (Nombre)	0.005 <sup>***</sup> (0.000)	0.004 <sup>***</sup> (0.000)	0.007 <sup>***</sup> (0.001)	0.007 <sup>***</sup> (0.001)
Handicap	0.013 <sup>***</sup> (0.002)	0.014 <sup>***</sup> (0.002)	0.033 <sup>***</sup> (0.008)	0.025 <sup>***</sup> (0.006)
Propriétaire	0.028 <sup>***</sup> (0.001)	0.033 <sup>***</sup> (0.001)	0.023 <sup>***</sup> (0.002)	0.025 <sup>***</sup> (0.002)
Taux d'entrepreneuriat (pères du même état)			0.170 <sup>***</sup> (0.017)	0.150 <sup>***</sup> (0.014)
Années	Non	Non	Oui	Oui
Cohortes	Oui	Oui	Non	Non
Régions	Oui	Oui	Oui	Oui
Constante	-0.210 <sup>***</sup> (0.014)	-	-0.140 <sup>***</sup> (0.024)	-
Observations	592374	592374	239101	239101

Écart-types entre parenthèses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Tableau 6: Dépendance intergénérationnelle**

	Estimés	
	MCO	Probit
Beta	0.37*** (0.06)	0.33*** (0.08)
Constante	0.01 (0.01)	0.07 (0.05)

Écart-types entre parenthèses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ **Tableau 7: Coefficients ethniques significatifs et différences intergénérationnelles**

	1ère génération		2e génération		Différence MCO (%)	Différence Probit (%)	Recoupement MCO (95%)	Recoupement Probit (95%)
	MCO	Probit	MCO	Probit				
<b>Effet significatif dans la 1ère génération</b>								
Cuba	0.085** (0.026)	0.340** (0.130)	0.046*** (0.012)	0.210*** (0.051)	-45.67	-36.06	-0.04	-0.23
Grèce	0.250*** (0.029)	0.900*** (0.130)	0.110** (0.035)	0.410*** (0.110)	-56.49	-54.62	0.02	0.02
Italie	0.120*** (0.026)	0.490*** (0.120)	0.076*** (0.020)	0.300*** (0.073)	-34.17	-38.39	-0.05	-0.19
Allemagne	0.082** (0.030)	0.310* (0.140)	-0.019 (0.018)	-0.110 (0.099)	-122.95	-135.48	0.01	-0.05
Corée	0.180*** (0.027)	0.630*** (0.130)	0.066*** (0.013)	0.260*** (0.050)	-62.55	-58.33	0.03	0.02
Philippines	-0.062** (0.024)	-0.400** (0.130)	-0.036*** (0.007)	-0.240*** (0.052)	-42.52	-40.76	-0.04	-0.19
Iran	0.220*** (0.042)	0.740*** (0.150)	0.140*** (0.029)	0.490*** (0.091)	-35.22	-32.99	-0.06	-0.23
Israël & Palestine	0.290*** (0.045)	0.990*** (0.160)	0.110** (0.041)	0.390** (0.130)	-61.60	-60.73	0.01	0.03
Liban	0.230*** (0.055)	0.860*** (0.190)	0.091* (0.037)	0.330* (0.130)	-61.04	-61.14	-0.04	-0.10
<b>Effet significatif dans la 2e génération seulement</b>								
Portugal	0.002 (0.025)	-0.110 (0.140)	0.065* (0.026)	0.310** (0.110)	2850.04	390.62	-0.04	-0.07
Pologne	-0.028	-0.210	0.110***	0.490***	475.31	338.44	0.04	0.24

	(0.027)	(0.150)	(0.021)	(0.080)				
Mexique	-0.022 (0.024)	-0.270* (0.120)	-0.013*** (0.003)	-0.095*** (0.020)	-38.26	-64.29	-0.05	-0.10
Amérique du Sud	0.013 (0.025)	0.034 (0.130)	0.042*** (0.007)	0.220*** (0.033)	229.08	538.28	-0.03	-0.14
Royaume- Uni	-0.013 (0.027)	-0.091 (0.140)	-0.028* (0.012)	-0.160* (0.077)	116.36	77.54	-0.09	-0.50
Chine	0.035 (0.025)	0.130 (0.120)	-0.017* (0.007)	-0.100* (0.041)	-149.07	-178.30	-0.11	-0.55

Écart-types entre parenthèses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Tableau 8: Équations de revenu**

	1ère génération		2e génération	
	Revenu d'entreprise	Salaire	Revenu d'entreprise	Salaire
<b>Variabes ethniques</b>				
Canada	0.500*** (0.150)	0.170*** (0.045)	0.250* (0.130)	0.140*** (0.035)
Mexique	0.086 (0.140)	-0.190*** (0.038)	-0.220*** (0.029)	-0.190*** (0.006)
Cuba	0.130 (0.140)	-0.190*** (0.041)	-0.130 (0.068)	-0.200*** (0.025)
Antilles	-0.160 (0.150)	-0.350*** (0.040)	-0.410*** (0.070)	-0.220*** (0.018)
Amérique du Sud	-0.008 (0.140)	-0.260*** (0.040)	-0.150** (0.051)	-0.160*** (0.014)
Royaume-Uni	0.350* (0.160)	0.230*** (0.044)	0.057 (0.110)	0.260*** (0.034)
Irlande	0.140 (0.190)	-0.023 (0.051)	0.240 (0.130)	0.240*** (0.052)
Grèce	0.170 (0.140)	-0.230*** (0.046)	0.160 (0.110)	0.054 (0.083)
Italie	0.150 (0.140)	-0.047 (0.039)	-0.003 (0.100)	-0.001 (0.037)
Portugal	0.420* (0.180)	-0.008 (0.041)	0.290* (0.120)	0.110* (0.046)
Allemagne	0.300	0.180***	0.019	0.120**

	(0.150)	(0.044)	(0.120)	(0.037)
Pologne	-0.008 (0.200)	-0.047 (0.046)	-0.008 (0.073)	0.009 (0.039)
Yougoslavie	0.039 (0.170)	0.005 (0.045)	-0.110 (0.110)	-0.026 (0.054)
Chine	-0.100 (0.140)	-0.260*** (0.040)	-0.068 (0.062)	-0.140*** (0.016)
Japon	0.160 (0.160)	0.260*** (0.047)	-0.240 (0.190)	0.081 (0.047)
Corée	-0.050 (0.140)	-0.190*** (0.043)	-0.100 (0.063)	-0.088** (0.027)
Philippines	0.180 (0.150)	-0.280*** (0.039)	-0.064 (0.120)	-0.130*** (0.017)
Inde	0.230 (0.140)	-0.110** (0.040)	-0.018 (0.044)	0.008 (0.012)
Iran	0.250 (0.160)	-0.400*** (0.084)	0.450*** (0.120)	-0.055 (0.055)
Israël & Palestine	0.220 (0.170)	-0.260*** (0.067)	-0.092 (0.120)	0.098 (0.089)
Liban	-0.017 (0.200)	0.008 (0.080)	-0.024 (0.170)	-0.120 (0.087)
Afrique	0.120 (0.170)	-0.260*** (0.049)	-0.200*** (0.060)	-0.270*** (0.020)

### Variables sociodémographiques

Éducation (années)	0.070*** (0.001)	0.042*** (0.000)	0.100*** (0.003)	0.087*** (0.001)
Âge	0.095*** (0.004)	0.078*** (0.001)	0.037** (0.014)	0.068*** (0.003)
Âge <sup>2</sup>	-0.100*** (0.005)	-0.085*** (0.001)	-0.037* (0.018)	-0.070*** (0.004)
Marié	0.210* (0.087)	0.180*** (0.020)	0.230*** (0.018)	0.230*** (0.004)
Enfants (nombre)	0.037*** (0.003)	0.006*** (0.001)	0.042*** (0.006)	0.035*** (0.002)
Handicap	-0.320*** (0.016)	-0.230*** (0.004)	-0.660*** (0.059)	-0.440*** (0.022)
Années	Non	Non	Oui	Oui
Cohortes	Oui	Oui	Non	Non

Régions	Oui	Oui	Oui	Oui
Constante	6.420*** (0.110)	7.320*** (0.026)	7.890*** (0.280)	7.630*** (0.055)
Observations	67195	525179	27035	212066
$R^2$	0.158	0.154	0.180	0.285

Écart-types entre parenthèses

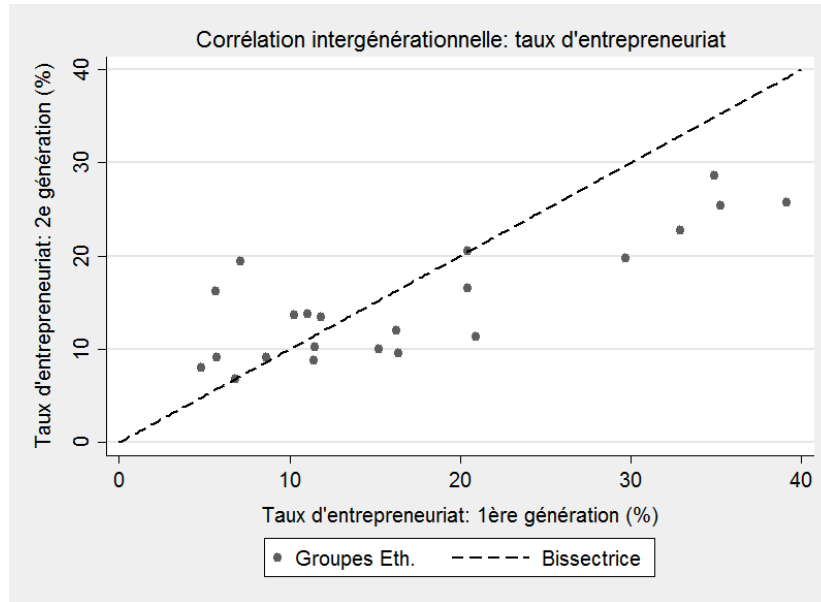
\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Tableau 9 : Phénomènes sociaux à l'origine de l'effet de l'ethnicité sur le choix sectoriel**

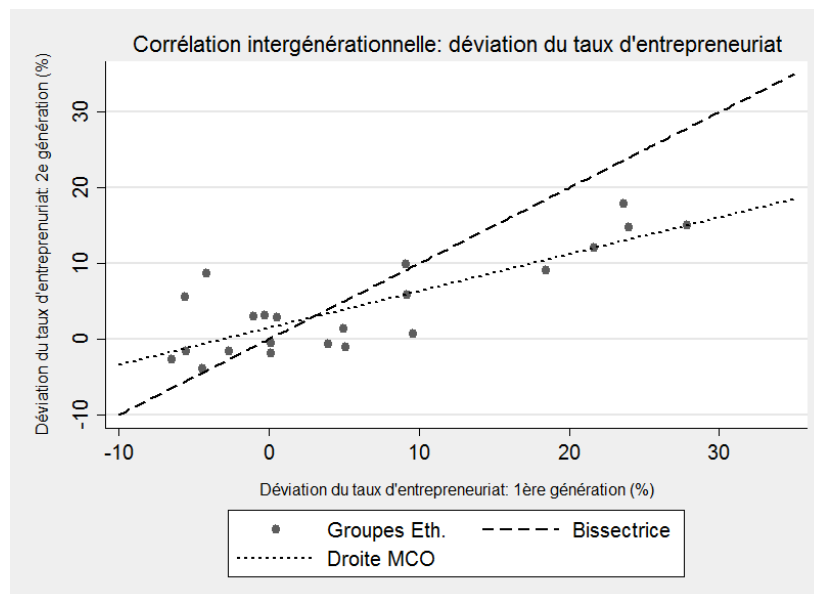
Groupes ethniques	1 <sup>ère</sup> génération			2 <sup>e</sup> génération		
	Discrimination (Marché du travail)	Discrimination (Marché des produits)	Réseaux d'affaires	Discrimination (Marché du travail)	Discrimination (Marché des produits)	Réseaux d'affaires
Cuba			√	√		
Grèce			√			√
Italie			√			√
Allemagne			√	Pas d'effet		
Pologne	Pas d'effet					√
Corée	√			√		
Philippines						
Iran			√			√
Israël & Pal.			√			√
Liban			√			√
<b>TOTAL</b>	<b>1</b>	<b>0</b>	<b>7</b>	<b>2</b>	<b>1</b>	<b>6</b>

## Annexe 2 : Graphiques

### Graphique 1

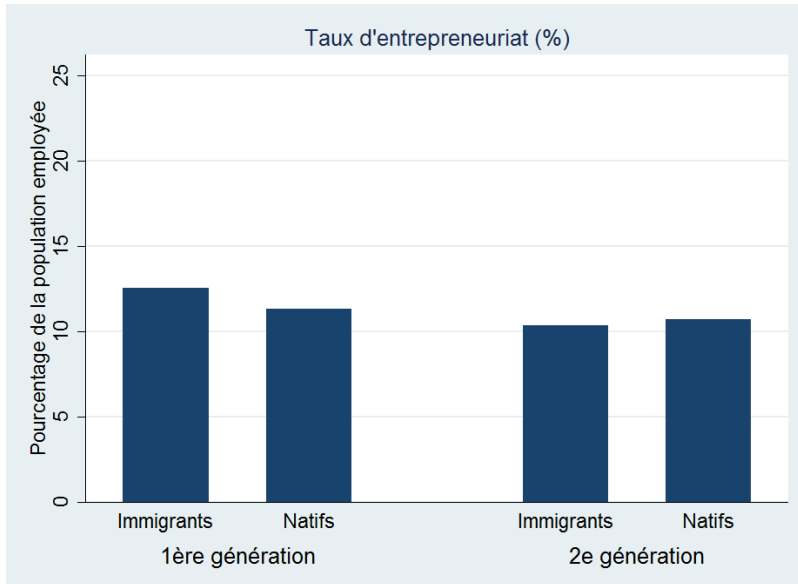


### Graphique 2

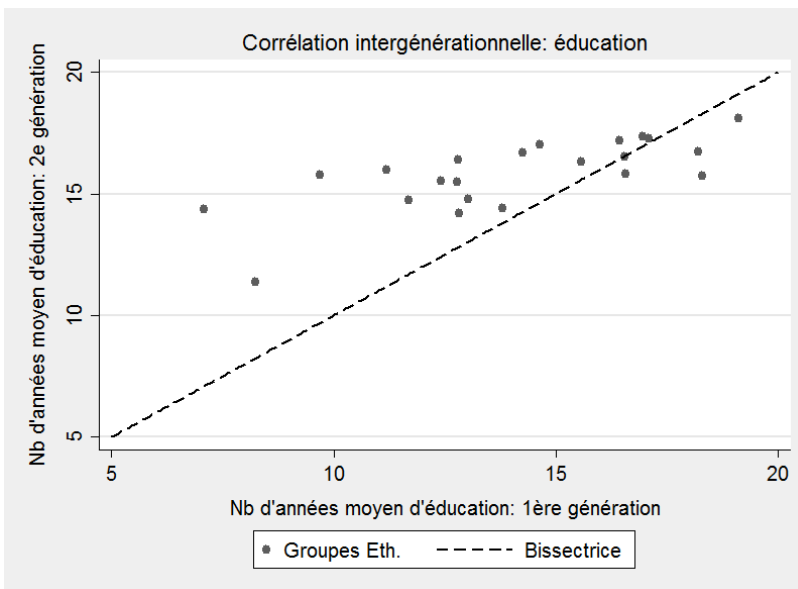




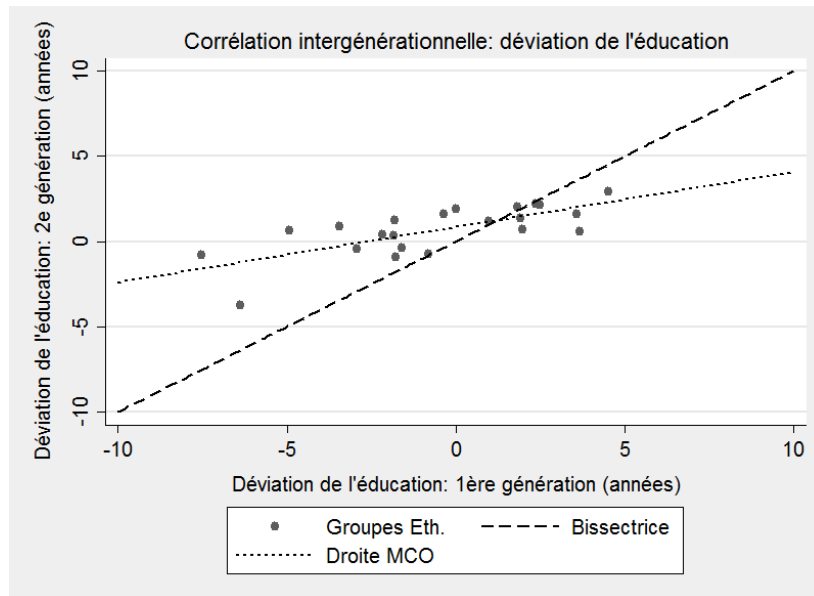
### Graphique 3



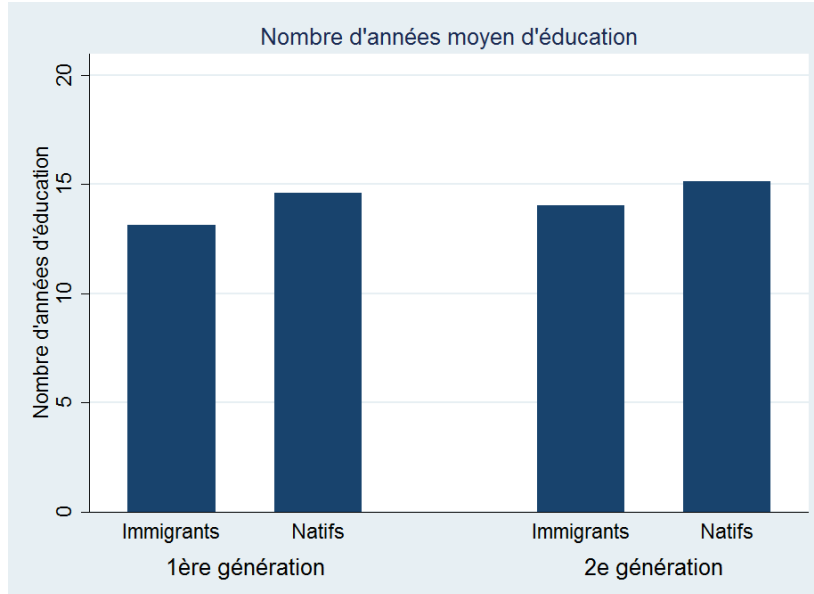
### Graphique 4



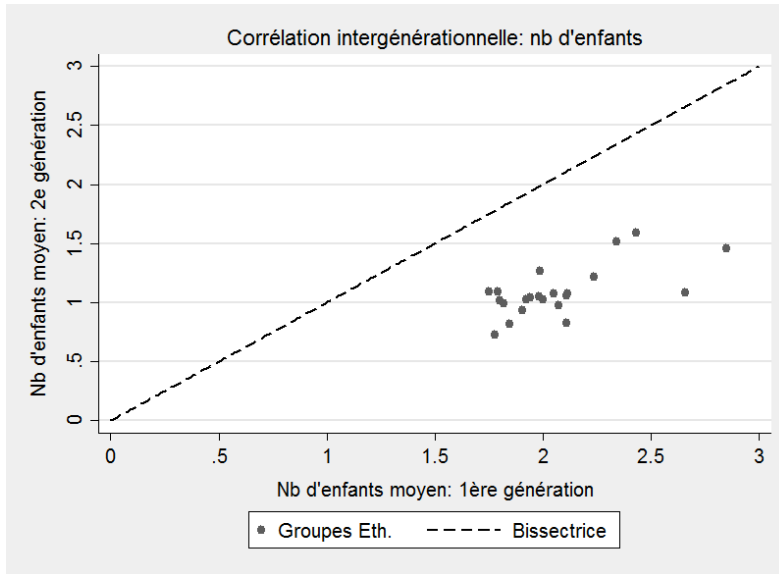
Graphique 5



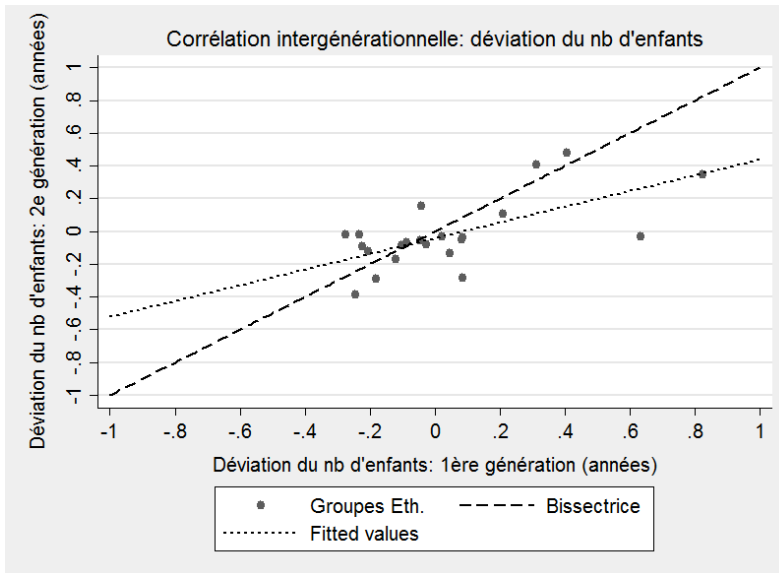
Graphique 6



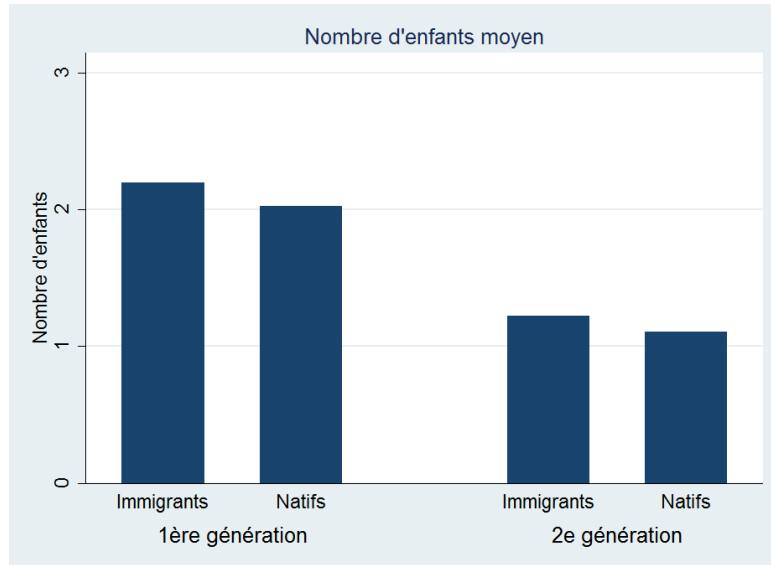
### Graphique 7



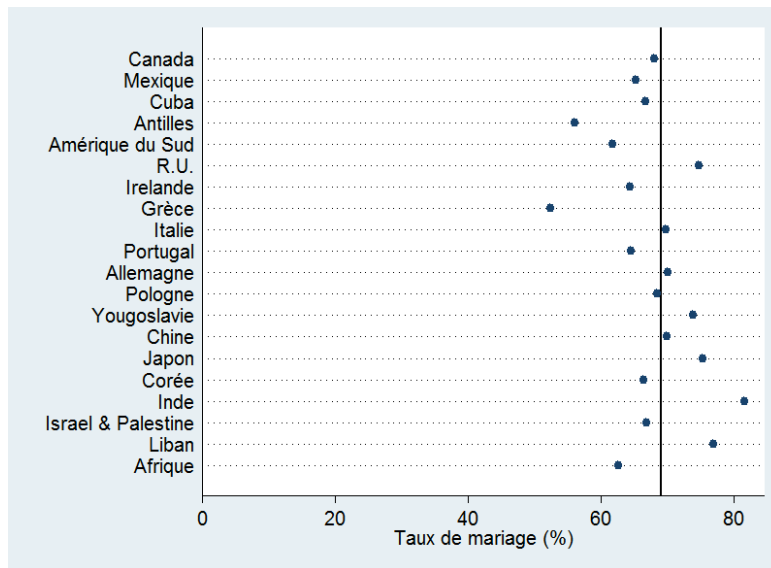
### Graphique 8



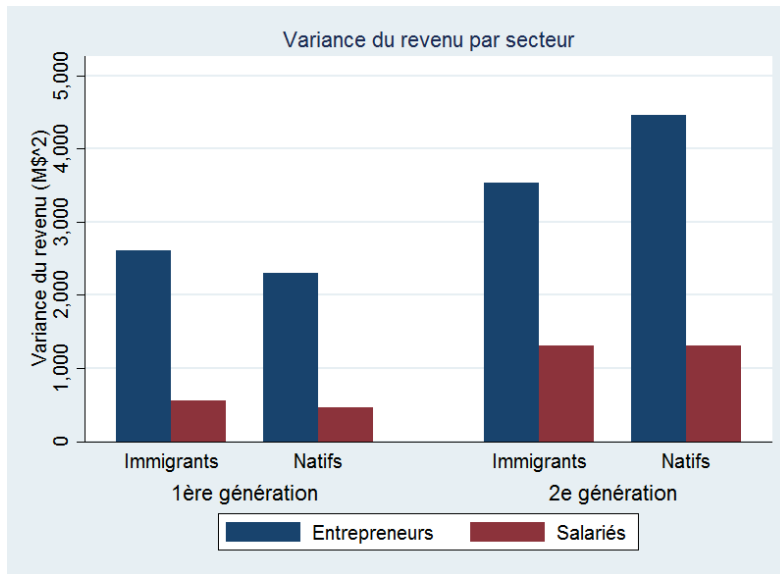
### Graphique 9



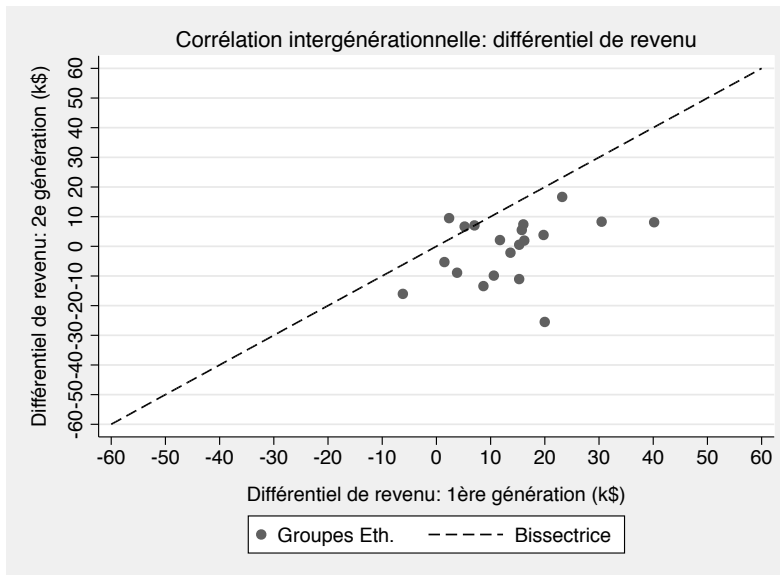
### Graphique 10



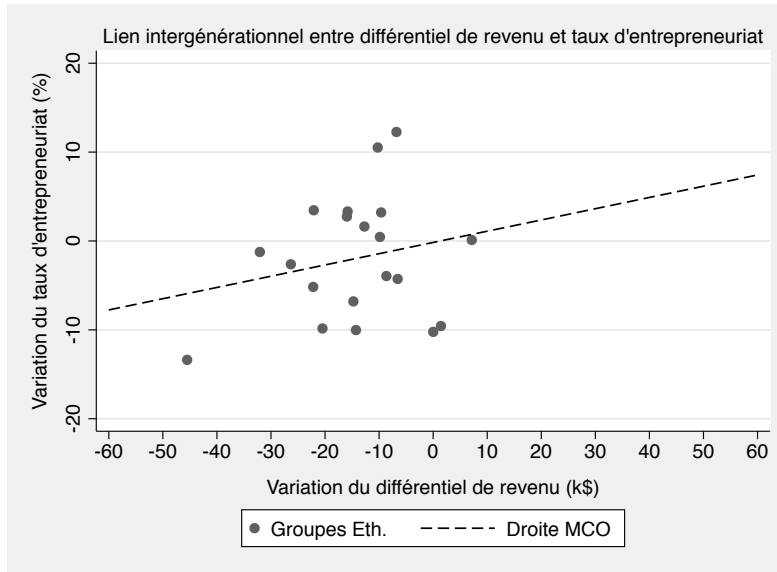
### Graphique 11



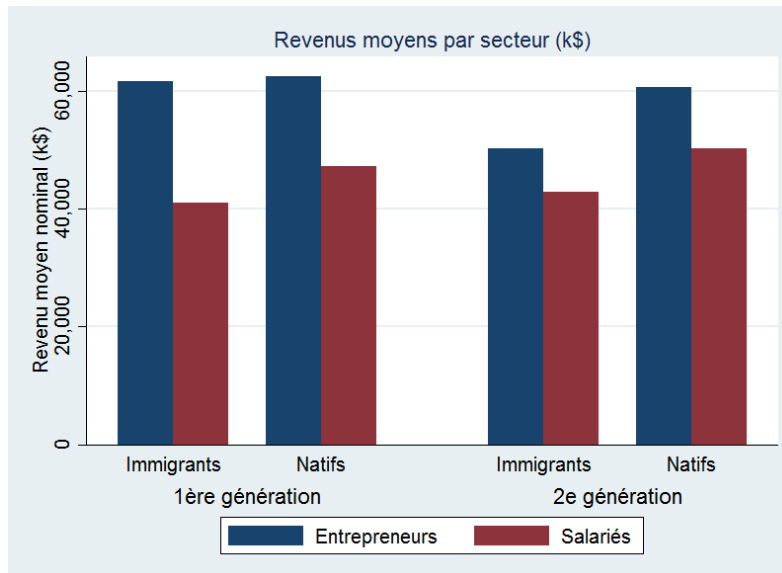
### Graphique 12



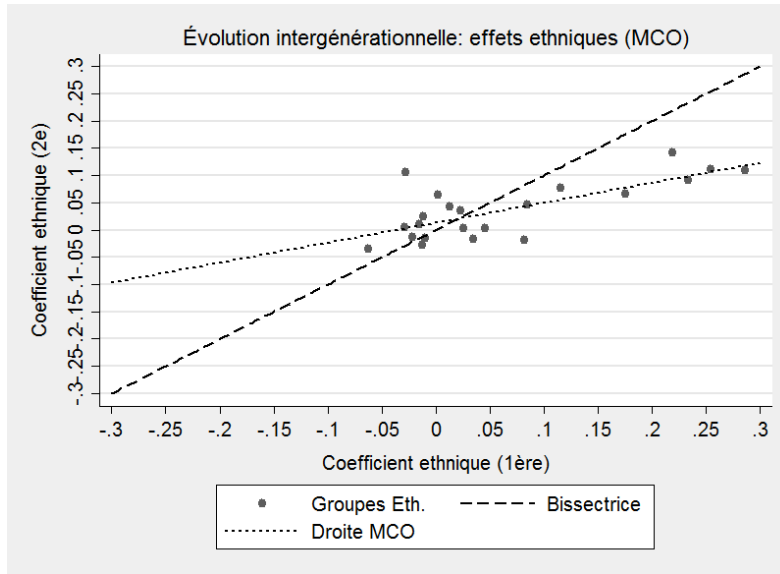
Graphique 13



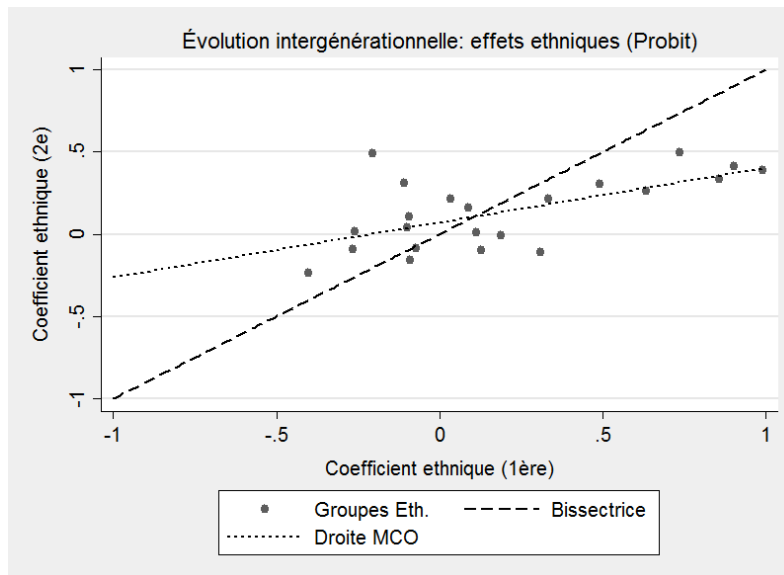
Graphique 14



Graphique 15



Graphique 16



### Annexe 3 : Méthode DiNardo-Fortin-Lemieux

Posons la variable dichotomique  $E$ , laquelle vaut 1 si l'individu est un entrepreneur et 0 si celui-ci est un travailleur salarié. Posons également la variable  $w$ , qui représente le revenu engrangé par le travailleur, peu importe le secteur d'activité. Par conséquent, la distribution contrefactuelle du revenu des entrepreneurs, c'est-à-dire la distribution de leur revenu si ces derniers avaient été payés comme des travailleurs salariés, est représentée par l'expression suivante :

$$g_{E=0}^C(w|E = 1) = \int f(w|X, E = 0)h(X|E = 1)dx$$

où  $f(w|X, E = 0)$  correspond à la densité conditionnelle des salaires et où  $h(X|E = 1)$  correspond à la distribution des caractéristiques observables chez les entrepreneurs. Puisque nous observons la densité conditionnelle des salaires uniquement dans l'échantillon de travailleurs salariés, il nous faut bâtir la distribution contrefactuelle de manière synthétique. Pour ce faire, nous effectuons une pondération de l'échantillon de travailleurs salariés en fonction de la probabilité conditionnelle que chaque individu soit entrepreneur. De manière plus précise, chaque travailleur salarié se voit attribuer le poids suivant :

$$\theta(X) = \frac{Prob(E = 1|X)}{Prob(E = 0|X)}$$

Chaque poids est calculé directement à partir du Probit qui explique le choix sectoriel des travailleurs (voir section IV.I), car à partir du moment où nous sommes capables de prédire la probabilité qu'un individu soit entrepreneur, nous sommes également capables de prédire la probabilité que celui-ci soit travailleur salarié ( $1 - Prob(E = 1|X)$ ). Grâce à l'attribution des poids, nous pouvons illustrer la distribution contrefactuelle du revenu des entrepreneurs de la manière suivante :



$$g_{E=0}^C(w|E = 1) = \int f(w|X, E = 0)\theta(X)h(X|E = 0)dx$$

Finalemment, afin d'effectuer la décomposition de variance sur l'échantillon contrefactuel d'entrepreneurs, nous estimons une régression linéaire (voir section IV.III) sur l'échantillon de travailleurs salariés pondéré qui est représenté par l'expression ci-haut. Pour estimer cette régression, nous employons la méthode des moindres carrés ordinaires pondérés et utilisons comme poids  $\theta(X)$ .