

Table des matières

Résumé	iii
Table des matières	iv
Liste des tableaux	v
Liste des figures	vi
Remerciements	ix
Introduction	1
1 Revue de la littérature	3
1.1 Le niveau de scolarité	3
1.2 Le lieu de formation	4
1.3 L’ethnicité	5
1.4 L’âge et l’expérience	6
1.5 Le domaine de la formation et de l’emploi	7
2 Méthodologie	10
2.1 L’appariement multiple	10
2.2 La régression par quantile	13
3 Base de données	16
3.1 Source	16
3.2 Analyse descriptive	16
4 Résultats	27
4.1 Présentation des résultats	27
4.2 Discussion des résultats	32
A Équations	36
A.1 Les conditions du premier moment	36
A.2 Détermination des quantiles	38
Bibliographie	40

Liste des tableaux

3.1	Caractéristiques observables des immigrants	18
3.2	Caractéristiques des immigrants en fonction de la grille de sélection et de l'âge	19
3.3	Fréquences observées auprès des immigrants selon le genre et les régions de provenance	20
3.4	Fréquences des qualifications des immigrants insérés dans le premier emploi	23
3.5	Fréquences des immigrants selon les difficultés rencontrées et les domaines d'études	24
3.6	Caractéristiques des immigrants selon le lieu d'obtention des diplômes	24
3.7	Caractéristiques linguistiques des immigrants avant l'arrivée	25
3.8	Niveaux linguistiques des immigrants formés avant l'arrivée et difficultés rencontrées lors de la recherche d'emploi au Québec	25
3.9	Fréquences des participants des formations professionnelles	26
4.1	Effets estimés des comparaisons inter-niveaux de formations formelles pour la durée du chômage	28
4.2	Résultats du test d'égalité des effets pour la durée du chômage	29
4.3	Effets estimés des comparaisons inter-niveaux des formations formelles pour le revenu	30
4.4	Résultats du test d'égalité des effets pour le revenu	32

Liste des figures

3.1	Niveau de scolarité des immigrants selon leurs régions de provenance	19
3.2	Distribution du logarithme des salaires horaires des immigrants selon l'ethnicité	21
3.3	Répartition du logarithme des salaires horaires des immigrants selon le genre .	22
4.1	Intervalles de confiance des comparaisons inter-niveaux pour la durée du chômage	29
4.2	Intervalles de confiance des comparaisons inter-niveaux pour le revenu	31

À ma mère bien aimée.

C'est là que je puisais comme la
plante dans le sol les suc
nourriciers de ma jeune
intelligence.

Johann Goethe

Remerciements

Mes remerciements les plus ineffables sont adressés à mon Père qui demeure un soutien pour moi dans mes moindres entreprises. Je tiens à remercier particulièrement mon directeur de recherche Monsieur Guy Lacroix qui n'a ménagé aucun effort à mettre à ma disposition, toutes les informations dont j'avais besoin pour la réalisation de mon projet de recherche. Je voulais également exprimer toute ma gratitude à mon codirecteur Monsieur Vincent Boucher pour son aide financière en collaboration avec mon directeur. Ces remerciements ne seraient pas complets sans souligner le soutien moral et financier de mes parents et de mes proches.

Introduction

Le vieillissement croissant de la population canadienne a conduit le gouvernement vers l'adoption d'une politique d'immigration. De nos jours, le Canada fait partie des dix premiers pays au monde qui accueillent le plus d'immigrants. Pour ces derniers, la motivation première de l'immigration demeure l'amélioration de leur bien-être par l'obtention d'un emploi qualifié. Une fois dans les provinces canadiennes, les immigrants tout comme les natifs rencontrent des barrières face à l'emploi, nonobstant leurs compétences. Ainsi, les premiers jours des immigrants sont marqués par un chômage dû à ces barrières, mais encore, à cause de leur faible connaissance du marché du travail canadien. Partant de ces faits, la plupart d'entre eux décident de suivre des formations formelles afin de faciliter leur insertion en emploi. Il s'agit principalement des formations linguistiques, professionnelles et académiques.

Cependant, la majorité des formations formelles sont rarement ciblées vers les immigrants et ne tiennent pas compte de leurs caractéristiques observables. Par ailleurs, il existe des immigrants qui n'ont suivi aucune formation avant d'obtenir leur premier emploi. C'est pourquoi cette étude cherche à évaluer l'impact des formations formelles sur le revenu et la durée du chômage avant l'obtention du premier emploi des immigrants qualifiés de la belle province du Canada. Ainsi, les résultats de l'étude permettront de conclure sur la pertinence de la participation des immigrants aux formations formelles.

À l'aide d'une base de données issue d'une enquête réalisée auprès des immigrants, nous estimerons essentiellement sur le revenu et la durée du chômage des immigrants, l'effet causal de la participation aux formations linguistiques, professionnelles et académiques. Il convient de noter que l'objectif principal de cette étude est de vérifier l'adéquation entre ces formations formelles et la situation économique des immigrants en fonction de leurs caractéristiques observées. En outre, cette étude cherche à confirmer ou infirmer, dans le contexte québécois, la pertinence de la participation des immigrants aux formations formelles.

Pour atteindre ces objectifs, la méthode d'appariement et la méthode de régression par quantile seront utilisées. Vu qu'il n'existe aucune règle d'assignation claire à la participation aux formations formelles, la méthode d'appariement basée sur des scores de propension généralisés s'impose dans la méthodologie. De plus, nos données sont majoritairement des données issues d'une enquête et non des données administratives. La seconde méthode se basera sur des quan-

tiles estimés, à partir du modèle général pondéré, pour estimer l'effet causal des traitements d'un quantile donné sur les quantiles des immigrants traités.

Une fois cette étude réalisée, les résultats pourront permettre au gouvernement du Québec de cibler les formations formelles vers les immigrants en tenant compte de leurs caractéristiques observées. De plus, l'implantation du ciblage réduirait énormément les aides financières gouvernementales du Québec vers les participants des formations, mais aussi, les coûts d'investissements vers les différentes formations formelles. Par ailleurs, l'intérêt d'analyser l'impact causal des formations formelles par quantile est de permettre aux décideurs économiques, privés et publics, de cibler certaines formations formelles vers les immigrants défavorisés économiquement.

Dans les lignes à suivre, une revue de la littérature des études empiriques, une analyse descriptive de la base de données et la méthodologie utilisée seront présentées. Ensuite, les résultats obtenus suite aux différentes estimations seront analysés. Enfin, la conclusion de l'étude sera marquée par des recommandations en fonction des résultats obtenus.

Chapitre 1

Revue de la littérature

Dans ce chapitre, nous présenterons les résultats empiriques des travaux de recherche qui se sont intéressés à l'évaluation d'impact des formations formelles dans le monde. Peu importe les formations formelles, nos recherches révèlent que leurs effets varient en fonction du niveau de scolarité, du lieu de formation, de l'ethnicité, de l'âge et de l'expérience, puis du domaine de la formation et de l'emploi. Dans les sections à suivre, il sera donc présenté les résultats de plusieurs études qui étayaient ces affirmations.

1.1 Le niveau de scolarité

La quasi-totalité des études aboutissent à des résultats qui divergent selon le niveau d'étude des individus. Dans cette section, il sera donc présenté trois (3) études qui illustrent un tel résultat.

Kang et Bishop (1989) [17] partent des données longitudinales de recensement des années 1980 et 1982 aux États-Unis, pour montrer que les personnes qui ont combiné au moins une formation professionnelle de 4 ans à une formation académique de 8 ans ont en moyenne un revenu espéré plus élevé que ceux qui ont un niveau d'étude scolaire inférieur et qui ont néanmoins suivi la même formation professionnelle pendant 4 ans. Ainsi, les hommes et femmes qui ont complété au moins 8 années du secondaire et qui ont suivi une formation professionnelle de 4 ans gagnent en moyenne respectivement 944 \$ et 624 \$ par an de plus que ceux qui n'ont pas complété les 8 années du secondaire. Il convient de noter que la base de données contient des informations sur 12 000 personnes ayant un niveau d'éducation inférieur ou égal au secondaire.

Ensuite, Chiswick et Miller (1995) [7] ont évalué l'effet de la maîtrise de la langue de destination des immigrants sur leurs revenus à partir de plusieurs bases de données qui contiennent des informations sur les immigrants interviewés lors du recensement de 1981 et 1986 en Australie, de 1980 aux États-Unis, de 1981 au Canada et de 1983 en Israël. Les bases de données contiennent essentiellement des informations sur les hommes âgés de 25 à 64 ans qui étaient

insérés en emploi lors des enquêtes. Le but de l'utilisation de plusieurs bases de données est de pouvoir démontrer la robustesse du modèle utilisé pour l'évaluation d'impact. Il convient de noter que la base de données australiennes de 1981 est plus riche en nombre d'individus que celle de 1986. Les résultats révèlent qu'une année additionnelle de scolarité augmente le revenu moyen de 6 % en Australie. Mais, après avoir effectué leur régression, ils trouvent qu'une année additionnelle de scolarité entraîne un gain additionnel de 8 % sur le revenu annuel des immigrants formés qui maîtrisaient l'anglais alors que ce gain est de 2 % sur le revenu annuel des immigrants formés qui ne maîtrisaient pas l'anglais.

Ferrer et Ridell (2001) [10] montrent qu'au Canada, les formations académiques augmentent environ de 20 % le revenu des immigrants titulaires d'un baccalauréat comparativement à ceux titulaires d'un diplôme du secondaire. En fait, les auteurs ont voulu évaluer l'impact des formations académiques sur les revenus de différentes cohortes d'immigrants à partir des données du recensement canadien des années 1981, 1986, 1991, 1996.

1.2 Le lieu de formation

Une lecture succincte de la littérature révèle que l'impact des formations formelles sur le revenu et la durée du chômage des participants varie selon le lieu de formation. Cependant, les référentiels des lieux de formation diffèrent selon les études. Il convient de noter que cette section de l'étude ne distinguera pas les articles selon les différents référentiels abordés dans les travaux empiriques. À ce titre, trois (3) articles seront présentés pour illustrer ce résultat.

De prime abord, les résultats linguistiques de Chiswick et Miller (1995) [7] sont positifs et significatifs sur les revenus annuels des immigrants quel que soit le pays sélectionné. En effet, ils trouvent que la maîtrise de la langue de destination augmente en moyenne de 5 % les revenus annuels des immigrants formés de 1981 en Australie et de 8 % ceux de 1986 dans le même pays. Toutefois, cet impact est plus grand en Israël, au Canada et aux États-Unis, soit respectivement 11 %, 12 % et 17 % pour les immigrants formés dans les langues nationales.

Cameron et al. (2001) [4] utilisent une base de données issue des enquêtes sur la fécondité mondiale entre 1975 et 1976, pour évaluer l'impact de l'éducation sur le taux de participation au marché de l'emploi, dans cinq pays de l'Asie à savoir, la Corée, la Thaïlande, le Sri Lanka, l'Indonésie et les Philippines. À caractéristiques identiques, les résultats montrent que l'impact du secondaire sur le taux de participation au marché de l'emploi des femmes de la Corée (5,5 %) et du Sri Lanka (9,3 %), est faiblement positif comparativement à l'effet observé chez les femmes qui résident en Thaïlande (57,3 %), en Indonésie (45,5 %) et aux Philippines (27,5 %).

Bellakhal et Mahjoub (2013) [2] font l'évaluation des formations professionnelles dispensées en Tunisie à partir d'une base de données constituée par le ministère de la formation pro-

fessionnelle tunisienne en 2001. Cette base contient des informations sur 462 personnes qui ont participé à différentes formations professionnelles et 418 personnes qui ne l'ont pas fait entre 1996 et 1998. Aux moyens d'un modèle de double sélection et d'une estimation par maximum de vraisemblance, les auteurs montrent que l'impact des formations professionnelles sur l'emploi et le revenu diffère selon la localisation des formations. Il se trouve que l'effet des formations professionnelles dispensées dans les grandes villes comparativement à celles des zones rurales est positif et significatif sur la probabilité d'être en emploi, soit 0,351. Bien que cet effet soit non significatif pour les personnes formées dans les villes intermédiaires, il est positif et significatif sur leur salaire moyen et s'établit à 0,155.

1.3 L'ethnicité

Les formations formelles sont largement embrassées par les individus à minorités ethniques qui se retrouvent dans un milieu où la nécessité de s'intégrer aux habitudes du terroir s'impose. Bien que leurs causes de participation aux formations formelles soient diverses, les lignes à suivre rapporteront essentiellement les résultats de trois (3) études qui aboutissent à des effets des formations formelles qui divergent selon l'ethnicité.

Tout d'abord, Jaeger et Page (1996) [16] utilisent une base de données qui contient des informations sur la population active américaine de 1991 et 1992. Les mêmes citoyens, questionnés sur leurs nombres d'années d'éducation en 1991 par le bureau du recensement américain, ont été questionnés sur leurs diplômes obtenus en 1992. Les résultats de l'étude révèlent que l'impact des formations académiques sur le salaire horaire des hommes noirs est significatif et plus élevé que celui observé chez les hommes caucasiens lorsqu'ils prennent en compte uniquement les diplômés de la maîtrise. Ainsi, les hommes noirs titulaires d'une maîtrise gagnent 65 % de plus que les hommes noirs titulaires d'un baccalauréat universitaire alors que les hommes caucasiens titulaires d'une maîtrise gagnent 5,5 % de plus que les hommes caucasiens titulaires d'un baccalauréat.

En outre, Leslie et Lindley (2001) [20] ont évalué l'impact des formations linguistiques sur le revenu et l'emploi en Angleterre à partir de deux modèles Tobit. Pour ce faire, les auteurs utilisent une base de données obtenue à partir de quatre enquêtes nationales réalisées auprès des minorités ethniques de 1966. Les auteurs ont voulu voir quel serait l'effet du langage sur l'emploi de la population si toutes les minorités ethniques comprenaient l'anglais. Après avoir postulé l'hypothèse de non-discrimination ethnique à l'emploi, les résultats montrent que la différence d'impact du langage sur l'emploi des hommes caucasiens et sur l'emploi des minorités ethniques est de 0,047 en faveur des hommes caucasiens. Cette différence monte à 0,086 lorsque l'hypothèse de non-discrimination ethnique est relâchée.

Dustmann et Fabbri (2003) [9] reposent leur étude sur l'exploitation de deux bases de données qui contiennent des informations sur les minorités ethniques présentes en Angleterre de 1993

à 1995. Il s'agit des personnes âgées de 16 ans et plus recensées d'une part, lors des enquêtes nationales sur les minorités ethniques réalisées de 1993 à 1994 et d'autre part, lors des enquêtes sur la vie familiale et professionnelle des minorités ethniques de 1994 à 1995. Ils montrent que la connaissance de l'anglais augmente significativement les salaires hebdomadaires des Noirs caraïbéens de 28 % comparativement à leurs homologues qui ne connaissent pas la langue. Par contre, cet impact est de 22 % pour les Afro-Asiatiques et de 16 % pour les Indiens. Ils trouvent que le fait de parler couramment l'anglais augmente environ de 15 % la probabilité d'être en emploi des minorités ethniques sélectionnées à partir des enquêtes nationales. Mais, lorsqu'ils séparent les aptitudes "écrire" et "parler" de la langue anglaise, ils trouvent que savoir écrire en anglais augmente environ de 13 % la probabilité d'être en emploi alors que savoir parler la langue l'augmente seulement de 5 %. Il en ressort donc que l'impact des formations linguistiques sur le revenu et l'emploi des minorités ethniques varie en fonction de l'aptitude développée lors de la formation.

1.4 L'âge et l'expérience

Plusieurs études font ressortir que l'impact des formations formelles sur le revenu et la durée du chômage des participants dépend de l'expérience et de l'âge des individus. Mais dans les lignes qui suivent, il sera présenté quatre (4) études qui soulignent la pertinence de l'âge et de l'expérience.

En premier lieu, Krueger et Rouse (1994) [19] sont partis de données d'enquêtes et administratives pour analyser l'impact des formations académiques offertes en milieu de travail sur le revenu des travailleurs. Dispensée au sein d'une entreprise de service, cette formation avait pour objectif principal d'augmenter les rendements des travailleurs. Ils trouvent que l'impact de la formation sur les revenus des travailleurs chevronnés et moins éduqués de l'entreprise de service est plus élevé que celui observé chez les travailleurs moins expérimentés et plus éduqués. Par conséquent, l'impact des formations académiques sur les revenus des travailleurs est faible lorsque les travailleurs disposent déjà des qualifications fournies par ladite formation.

En outre, Gonzalez (2000) [14] a évalué l'impact des aptitudes linguistiques sur le revenu des immigrants présents aux États-Unis en 1992. À l'aide d'une base de données qui contient des informations démographiques et des informations liées au marché de l'emploi américain, il conclut que l'impact des formations linguistiques sur le revenu des immigrants varie en fonction de l'âge d'entrée des immigrants aux États-Unis. Ainsi, les résultats montrent que la probabilité d'exceller dans n'importe quelle aptitude en anglais est réduite de 10 % pour les immigrants qui sont venus aux États-Unis au-delà de 30 ans d'âge comparativement à ceux qui ont immigré à l'âge de 20 ans.

Ensuite, Flores-Lagunes et Light (2004) [12] ont évalué l'impact des formations académiques sur les salaires horaires. En effet, leur étude repose sur 7 113 Américains (hommes et femmes)

caucasiens des États-Unis présents dans leur base de données longitudinales. Ils trouvent des résultats différents lorsqu'ils utilisent l'expérience potentielle et l'expérience actuelle des Américains caucasiens de la base de données. Ainsi, ils s'aperçoivent que l'expérience potentielle surestime l'effet moyen du nombre d'années d'éducation sur le revenu des diplômés, comparativement aux résultats obtenus avec l'expérience actuelle. Cependant, l'expérience potentielle sous-estime l'effet moyen des diplômes obtenus sur le revenu des personnes formées comparativement aux résultats observés avec l'expérience actuelle. En effet, les auteurs ont utilisé l'âge des individus comme mesure de l'expérience potentielle et le nombre d'heures de travail cumulé (depuis le 20^{ième} anniversaire des individus à la date notifiée pour les salaires) comme mesure de l'expérience actuelle.

De surcroît, Hanushek et al. (2011) [15] se servent d'une base de données issue de l'enquête internationale sur l'alphabétisation des adultes, pour évaluer l'impact des formations professionnelles et académiques sur le revenu et l'emploi suivant le cycle de vie des individus. Les données constituées regroupent des informations sur des personnes, âgées de 15 à 65 ans, de la Nouvelle Zélande, du Chili, des États-Unis et de 15 pays européens. Les résultats révèlent qu'à l'âge de 16 ans, les individus qui ont reçu une formation professionnelle ont une probabilité de 7 % de plus d'obtenir un emploi que ceux qui ont reçu une formation académique. Cependant, cette différence varie en fonction de l'âge si bien qu'elle baisse de 2 % à chaque 10 ans. Par contre à 50 ans, l'effet contraire est observé. Ces résultats sont vérifiés quelle que soit la méthode d'estimation employée par les auteurs.

Kilolo-Malambwe (2013) [18] utilise les données de l'Enquête québécoise sur les conditions de travail, d'emploi et de SST (EQCOTESST) qui contient des informations sur 5 071 personnes âgées de 15 à 64 ans entre 2007 et 2008, pour montrer que le fait d'être surqualifié ou sous-qualifié a des répercussions sur la rémunération des travailleurs. En effet, ses résultats montrent que les travailleurs surqualifiés gagnent 13 % de moins que les travailleurs qualifiés alors que les travailleurs sous-qualifiés gagnent 14 % de plus que les travailleurs qui ne le sont pas. Ainsi, les travailleurs surqualifiés occupent des emplois qui exigent des qualifications inférieures à celles qu'ils détiennent. Par conséquent, un travailleur surqualifié gagne moins qu'un travailleur qualifié lorsque les deux occupent le même emploi.

1.5 Le domaine de la formation et de l'emploi

Dans le cadre de la présente revue de la littérature, les travaux empiriques abordés font ressortir que l'impact des formations académiques sur le revenu et la durée du chômage des personnes formées varie en fonction de leurs domaines d'étude. Par ailleurs, il arrive que certains domaines d'emploi aient moins recours à une connaissance approfondie d'une langue que d'autres si bien que l'impact des formations linguistiques sur le revenu et l'emploi varie selon les domaines d'emploi. Dans les paragraphes à suivre, il sera présenté trois (3) études

qui illustrent ces affirmations.

Tout d'abord, Neuman et Ziderman (1991) [21] utilisent 20 % des données du recensement de la population d'Israël de 1983 pour effectuer une étude d'impact de différentes formations professionnelles sur le revenu mensuel des Juifs travaillant à temps plein. En fait, la base de données contient des informations sur 14 000 individus répartis en deux groupes d'hommes âgés de 25 à 49 ans. Le premier est celui des hommes ayant terminé une formation professionnelle et l'autre groupe celui des hommes ayant terminé une formation académique générale. Ces auteurs trouvent que les Juifs qui ont suivi une formation professionnelle en électricité, en électronique, en mécanique automobile et qui sont employés dans leur domaine, gagnent respectivement 0,05, 0,18 et 0,04 de plus que les Juifs formés en agriculture et insérés en emploi lié directement à l'agriculture. Par contre, les personnes formées en gestion d'hôtellerie et en couture, puis insérées dans un emploi en lien avec leurs formations gagnent respectivement 0,02 et 0,05 de moins que les personnes formées en agriculture et insérées dans un emploi en lien avec leur formation.

Ensuite, Battu et al. (1999) [1] fondent leur étude sur une base de données constituée à l'issue d'une enquête réalisée en 1996 par l'université de Birmingham. Cette base contient des données de panel sur des Anglais diplômés de deux cohortes différentes à savoir celle de l'année académique de 1985, puis celle de 1990 au Royaume-Uni. Ils montrent que les hommes diplômés en médecine, en mathématique et en ingénierie ont plus facilement accès à un emploi qui correspond à leurs qualifications, comparativement aux diplômés des autres domaines tels que le droit et les sciences sociales. De plus, cet effet croît dans le temps. Ainsi en 1986, l'impact des formations académiques sur l'emploi est significativement positif pour les diplômés de 1985 en médecine (1,28), en mathématique (0,57), puis en ingénierie (0,94), comparativement aux diplômés en droit (-0,32) et en sciences sociales (-0,23). Néanmoins, les résultats obtenus pour ces deux dernières filières ne sont pas significatifs. Six ans plus tard, les auteurs constatent que l'impact de la formation sur l'emploi passe de 1,28 à 1,23 pour les diplômés en médecine, puis de 0,58 à 0,69 pour les diplômés en mathématique, et demeure constant pour les diplômés en ingénierie.

Berman et al. (2003) [3] utilisent une base de données qui contient des informations sur les travailleurs masculins en Israël de 1994. La base de données contient surtout des informations sur des immigrants en provenance de l'ancienne Union soviétique, car la grande migration des Russes vers Israël prenait place 5 ans plus tôt. Les résultats de l'étude montrent que l'apprentissage de l'hébreu augmente en moyenne le revenu mensuel des techniciens Russes en informatique de 11,2 % comparativement à leurs homologues qui n'ont aucune aptitude en hébreu. Par contre, la même formation augmente en moyenne le revenu mensuel des programmeurs Russes en informatique de 8,1 %, puis environ de 3 % le revenu mensuel des travailleurs Russes de la construction et des stations d'essence comparativement à leurs homologues qui n'ont suivi aucune formation en hébreu. Pour pallier aux erreurs de mesure occasionnées par

la corrélation existante entre la durée de séjours dans un pays et l'apprentissage d'une langue, les auteurs ont spécifié un autre modèle de différence qui leur permet d'utiliser le revenu et le niveau linguistique des immigrants avant leur insertion en emploi ainsi que les mêmes variables observées au cours de l'enquête. Ils constatent que les résultats ne changent guère pour les techniciens et les programmeurs Russes en informatique, mais que l'effet de l'hébreu observé chez les travailleurs Russes de la construction et des stations d'essence est nul voire négatif. L'étude conclut donc que les Russes qui maîtrisent plus la langue hébreu se sont dirigés vers des emplois qui nécessitent l'usage de la langue.

Chapitre 2

Méthodologie

Dans ce chapitre, il sera présenté premièrement la méthode d'appariement et ensuite la régression par quantile.

2.1 L'appariement multiple

Cette section de la méthodologie est basée sur les travaux de Rosenbaum et Rubin (1983) [22], de Dehejia et Whaba (2002) [8], de Cattaneo (2010) [5], de Gertler et al. (2011) [13], puis de Cattaneo et al. (2013) [6] pour présenter la méthode d'appariement multiple basée sur les scores de propension généralisés. En fait, le score généralisé est un scalaire qui résume toutes les caractéristiques observables des individus en fonction de leur participation à un traitement donné. Sur la base de ce score, un participant d'un traitement quelconque peut être apparié à son contrefactuel qui n'est rien d'autre qu'un individu du groupe de contrôle ayant le même score que lui. Après avoir apparié les contrefactuels aux participants des traitements, les estimateurs des effets causaux sur les personnes traitées sont obtenus en faisant la différence moyenne, entre les variables d'intérêts des individus traités (Y_1) et celles de leurs contrefactuels (Y_0).

$$\hat{\Delta}_{ATT} = E(Y_1|T = t) - E(Y_0|T = t); \text{ avec } t \text{ un traitement donné.} \quad (2.1)$$

La méthode d'appariement multiple repose sur deux hypothèses fondamentales à savoir :

- l'hypothèse d'indépendance conditionnelle (CIA¹);
- l'hypothèse du support commun.

Bien que l'hypothèse d'indépendance conditionnelle ne puisse être vérifiée et que celle du support commun s'impose, cette méthode possède des avantages d'utilisation et des tests

1. Le sigle de cette hypothèse nous provient de son sens en anglais « Conditional Independence Assumption » (CIA).

de vérification de la robustesse des résultats. Toutefois, nous prendrons le temps de bien comprendre les hypothèses de cette méthode avant de présenter le calcul des effets estimés du traitement multiple. Enfin, les limites et les avantages de la méthode d'appariement seront exposés.

2.1.1 L'hypothèse d'indépendance conditionnelle (CIA)

L'hypothèse d'indépendance conditionnelle se résume par l'équation suivante :

$$Y_0 \perp T|X \Rightarrow E[Y_0|T = t, X] = E[Y_0|T = 0, X]; t = 1, 2, 3, \dots \quad (2.2)$$

Pour un vecteur de caractéristiques X donné, la CIA stipule qu'en moyenne la variable d'intérêt Y observée pour les non-participants aux formations doit correspondre à la variable d'intérêt Y qui aurait été observée pour les participants s'ils n'avaient pas été formés. C'est de la vraisemblance de cette hypothèse que dépend la robustesse de l'effet causal estimé par la méthode d'appariement.

2.1.2 L'hypothèse du support commun

Cette hypothèse stipule qu'il existe un groupe de comparaison pour les personnes traitées. C'est-à-dire que tout le monde possède une probabilité non nulle et strictement inférieure à 1 ($0 < r(t, \mathbf{X}) < 1$) de choisir un traitement. Par conséquent, un contrefactuel peut être trouvé pour toutes les personnes traitées. La méthode d'appariement impose le support commun des probabilités d'être traitées suivant les caractéristiques observables des deux groupes. De même, une personne traitée et son contrefactuel doivent avoir les mêmes scores généralisés, car la construction du groupe de comparaison est basée sur la proximité des scores.

2.1.3 Estimation des effets du traitement multiple

En partant du support commun et du CIA, il existe deux procédures d'identification et d'Estimation semi-paramétriques (ESP) des effets du traitement multiple $\hat{\Delta}_{(ATT_{ESP})}$. La première procédure d'estimation part d'une fonction influente partielle $\phi_{(ipw,t)}$ qui renferme des probabilités inverses pondérées, pour construire les estimateurs des effets moyens $\hat{\Delta}_{ATT_{(ipw,t)}}$. En fait, les probabilités inverses pondérées ne sont que l'inverse des scores généralisés des individus. La deuxième procédure d'estimation s'appuie sur une fonction influente complète ($\phi_{(eif,t)}$) pour construire les estimateurs des effets moyens $\hat{\Delta}_{ATT_{(eif,t)}}$. Ainsi, la fonction influente complète ($\phi_{(eif,t)}$) est une combinaison entre les probabilités inverses pondérées et une fonction non paramétrique $e_t(X_i, \mu_t)$. Cette dernière n'est rien d'autre que la différence moyenne entre la somme et la moyenne des Y des participants au traitement t , étant données leurs caractéristiques observables.

En outre, les estimateurs obtenus par ces deux procédures d'estimation sont semi-paramétriques, efficaces, convergents et normalement distribués. Ils sont asymptotiquement équivalents sous

la forme non paramétrique. La validité de ces estimateurs est sujette à la vérification des conditions du premier moment² pour chaque traitement t . Ces conditions stipulent qu'en moyenne, les fonctions influentes partielle ($\phi_{(ipw,t)}$) et complète ($\phi_{(eif,t)}$) doivent sommer à 0 lorsque tous les individus ayant participé au traitement t sont pris en compte.

Dans les équations suivantes, $\mathbf{D}(t) = 0$ si $t = 0$ désignera le groupe de contrôle et $\mathbf{D}(t) = 1$ si $t \neq 0$ désignera le groupe des participants, puis $\hat{Y}_i(t)$ sera la valeur prédite issue de la régression des Y des participants au traitement t sur leurs caractéristiques observables X . Les estimateurs $\hat{\Delta}_{(ATT_{ESP})}$ sont donc obtenus comme suit :

$$\hat{\Delta}_{ATT_{(ipw,t)}} = \left(\sum_{i=1}^n \frac{\mathbf{D}_i(t)}{\hat{r}(t, \mathbf{X})} \right)^{-1} \sum_{i=1}^n \frac{\mathbf{D}_i(t) Y_i}{\hat{r}(t, \mathbf{X})}; \quad (2.3)$$

$$\hat{\Delta}_{ATT_{(eif,t)}} = \frac{1}{n} \sum_{i=1}^n \left(\frac{\mathbf{D}_i(t) Y_i}{r(t, \mathbf{X})} - \left[\frac{\mathbf{D}_i(t)}{r(t, \mathbf{X})} - 1 \right] \hat{Y}_i(t) \right). \quad (2.4)$$

Dans l'équation (2.3), l'estimateur $\hat{\Delta}_{ATT_{(ipw,t)}}$ est obtenu en faisant le rapport entre la pondération de la somme des Y des participants au traitement t et la somme des poids individuels des participants dudit traitement. Comme le montre l'équation (2.4), l'estimateur $\hat{\Delta}_{ATT_{(eif,t)}}$ est une moyenne de la somme des résidus (issus de la régression donnant $\hat{Y}_i(t)$) et de la valeur prédite $\hat{Y}_i(t)$ des participants au traitement t . Ce dernier estimateur est recommandé par Cattaneo et al. (2013) [6] et sera utilisé dans cette étude car il est perçu comme la version non paramétrique des estimateurs "doublement robustes"³. Ce qui n'est pas le cas de l'estimateur $\hat{\Delta}_{ATT_{(ipw,t)}}$. Les écarts-types de ces estimateurs peuvent être obtenus à partir de z simulations comme suit :

$$\hat{\sigma}_{(ATT_{ESP})}^2 = \frac{\sum_{i=1}^z \left(\hat{ATT}_{ESP_i} - \bar{ATT}_{ESP} \right)^2}{z - 1}; \text{ avec } \bar{ATT}_{ESP} = \sum_{i=1}^z \hat{ATT}_{ESP_i}. \quad (2.5)$$

2.1.4 Limites et avantages

D'entrée de jeu, l'une des limites majeures de l'appariement est qu'il repose sur une hypothèse forte qui n'est pas vérifiable. Sur la base de la CIA, il est impossible de prouver qu'il n'existe pas des caractéristiques non observées qui pourraient influencer la participation et biaiser les résultats. Toutefois, cette hypothèse devient vraisemblable si l'historique des Y est intégré dans le vecteur des caractéristiques observées des individus. En fait, tout ce qui n'est pas observable serait contenu dans la variable retardée, conditionnée sur le traitement. De même, la CIA est plus vraisemblable lorsque la matrice \mathbf{X} renferme plus de variables explicatives.

2. Voir la sous-section A.1.1 en Annexe A.

3. Cet estimateur est doublement robuste lorsqu'il est sous sa forme paramétrique.

En outre, l'hypothèse du support commun peut conduire l'estimateur d'appariement à être local. En effet, la taille de l'échantillon d'étude diminue lorsque les caractéristiques observées des groupes de traitement ne sont pas en moyenne semblables à celles du groupe de contrôle. Ce qui ne permettra pas d'estimer l'effet causal d'un traitement donné sur les Y des personnes prises au hasard dans la population («Average Treatment Effect» ATE). Néanmoins, après le support commun, la méthode d'appariement permet d'estimer sans biais l'effet causal des traitements sur les personnes traitées («Average Treatment on the Treated» ATT). Enfin, l'avantage principal de l'appariement réside dans la possibilité de reproduire une expérience randomisée avec des données non randomisées, sans spécifier une hypothèse sur la distribution des termes d'erreurs.

2.2 La régression par quantile

Après avoir obtenu les effets moyens du traitement multiple sur les Y des participants, il est nécessaire d'évaluer les effets quantiles du traitement multiple sur les Y participants. En effet, le quantile n'est rien d'autre que l'inverse de la distribution cumulative des Y de la population. Dans les lignes qui suivront, nous allons nous inspirer des travaux de Firpo (2007) [11], de Cattaneo (2010) [5], puis de Cattaneo et al. (2013) [6], pour présenter premièrement les hypothèses cruciales de la méthode, ensuite le calcul des estimateurs quantiles des effets du traitement multiple, enfin les limites et les avantages de la méthode.

2.2.1 Les hypothèses

Hormis l'hypothèse d'indépendance conditionnelle, la régression par quantile s'appuie sur l'hypothèse d'invariance des rangs pour faire des interprétations individuelles. Étant donnée une proportion d'individus de l'échantillon d'étude τ , cette hypothèse stipule qu'une personne traitée ne change pas de quantile après le traitement, c'est-à-dire qu'une personne du groupe de contrôle dont le Y se retrouve dans un certain quantile se trouve dans le même quantile de la distribution des Y avec traitement.

$$\text{Si } i \text{ est tel que } Y_{0i} < Q_{Y_{0i}}(\tau), \text{ alors } Y_{1i} < Q_{Y_{1i}}(\tau). \quad (2.6)$$

Sous ces hypothèses, l'estimation des effets quantiles du traitement multiple sur les personnes traitées (QTET⁴) est possible. En effet, l'estimateur QTET n'est rien d'autre que la différence des quantiles entre une personne traitée et son contrefactuel, étant donnée τ et un traitement t .

4. Quantile Treatment Effect on the Treated.

2.2.2 Calcul de l'estimateur QTET et de sa variance estimée

L'estimation des effets quantiles du traitement nécessite au préalable l'estimation et l'identification des quantiles conditionnels au traitement t . D'une part, ces quantiles seront estimés à partir de la minimisation de la somme des valeurs absolues des résidus selon le modèle général pondéré⁵, car ce dernier tient compte des effets de sélection en affectant un poids à chaque résidu. D'autre part, l'identification des quantiles conditionnels s'effectuera grâce aux méthodes semi-paramétriques⁶ proposées par Firpo (2007) [11].

Après l'identification et l'estimation des quantiles conditionnels, l'évaluation des effets quantiles du traitement multiple se fait par le biais des estimateurs proposés par Cattaneo (2010) [5]. Soient $\hat{Q}_{TET(ipw,t)}(\tau)$ et $\hat{Q}_{TET(eif,t)}(\tau)$ les estimateurs quantiles semi-paramétriques obtenus respectivement à partir des fonctions influentes partielle $\phi_{(ipw,t)}$ et complète $\phi_{(eif,t)}$. Tout comme les estimateurs moyens, la validité des estimateurs quantiles passe par la vérification des conditions du premier moment⁷ pour chaque τ^e quantile. Une fois ces conditions respectées, les estimateurs quantiles semi-paramétriques⁸ peuvent donc être utilisés pour mesurer l'effet causal de la participation à un traitement sur le τ^e quantile des personnes traitées. Ainsi, l'effet causal et sa variance sont obtenus de la manière suivante :

$$\hat{Q}_{TET_{T=t}}(\tau) = \hat{Q}_{Y_1}(\tau|T=t) - \hat{Q}_{Y_0}(\tau|T=t) ; \quad (2.7)$$

$$\hat{V}_{(\tau|T=t)} = \frac{1}{n} \sum_{i=1}^N \left(\hat{\phi}_{ESP,t} - \hat{\theta}_{t,i} \right)^2 . \quad (2.8)$$

$\hat{\phi}_{ESP,t}$ étant la fonction influente estimée, partielle $\phi_{(ipw,t)}$ ou complète $\phi_{(eif,t)}$; $\hat{\theta}_{t,i}$ étant l'effet estimé des scores généralisés sur la fonction influente retenue.

2.2.3 Limites et avantages

L'une des limites de cette méthode est que les quantiles conditionnels ne correspondent pas en moyenne aux quantiles de la population. En fait, la loi des espérances itérées est inapplicable sur les quantiles si bien que la moyenne des quantiles conditionnés diffère du quantile moyen de la population. Par ailleurs, il est impossible de faire des interprétations individuelles sans la contrainte imposée par l'hypothèse d'invariance des rangs.

Néanmoins, le principal avantage de la régression par quantile est qu'elle ne nécessite aucune hypothèse sur la distribution des termes d'erreurs, aussi longtemps que l'inférence n'est pas prise en compte. De plus, cette méthode permet de voir l'effet des traitements sur la constante

5. Voir la sous-section A.2.1 en Annexe A.

6. Voir la sous-section A.2.2 en Annexe A.

7. Voir la sous-section A.1.2 en Annexe A.

8. Voir la sous-section A.2.3 en Annexe A.

et la pente des distributions cumulatives des Y des participants. Enfin, les estimateurs des effets du traitement multiple par quantile sont moins sensibles aux valeurs extrêmes et aberrantes.

Chapitre 3

Base de données

Dans ce chapitre, il sera présenté la source et l'analyse descriptive de la base de données utilisée dans le cadre de cette étude.

3.1 Source

Dans le cadre de cette étude, le Ministère de l'Emploi, de la Solidarité Sociale et de la Famille du Québec en collaboration avec le Ministère de l'Immigration et des Communautés Culturelles (MICC) ont mis à notre disposition une base de données qui contient des informations sur des immigrants de la classe des travailleurs qualifiés. En effet, cette base de données est issue principalement des données d'enquête réalisée en 2011 auprès de 90 000 immigrants qui sont venus au Québec entre le 1^{er} janvier 2002 et le 31 décembre 2009. Ainsi, un échantillon aléatoire de 10 000 immigrants a été tiré de la population cible. Ensuite, les informations des personnes recensées par le MICC ont été fusionnées avec certaines données administratives de la Régie de l'assurance-maladie du Québec (RAMQ). Il en ressort que seulement 7 401 immigrants résidaient à une adresse valide au Québec. Par ailleurs, la liste des noms et des adresses des immigrants a été remise à la firme Léger Marketing afin qu'elle se charge du sondage auprès des immigrants. Enfin, la base de données contient des informations sur 3 009 immigrants qui ont répondu aux questions de l'enquête, avec 2 607 immigrants sélectionnés à partir de la grille de 1996 et 402 immigrants sélectionnés à partir de la grille de 2006.

3.2 Analyse descriptive

L'analyse descriptive s'effectuera selon deux volets à savoir, les caractéristiques de l'échantillon et les caractéristiques des formations formelles. Ainsi, l'analyse de l'échantillon tiendra compte de la grille de sélection, des régions de provenance et du revenu des immigrants pour faire ressortir les différences observées au niveau des critères de sélection et des lieux de provenance des immigrants. Enfin, les caractéristiques des immigrants formés seront présentées en fonction

de la typologie des formations.

3.2.1 Caractéristiques de l'échantillon

Cette sous-section analyse les caractéristiques observables des immigrants échantillonnés dans la base de données. De ce fait, une analyse succincte des immigrants sélectionnés, selon les grilles de sélection des années 1996 et 2006, débutera cette sous-section. Ensuite, les statistiques descriptives observées auprès des immigrants seront présentées selon leurs régions de provenance et leurs revenus du premier emploi.

Grille de sélection

En 2006, le gouvernement québécois a modifié les critères de la grille de sélection de 1996, en accordant plus de poids au niveau de scolarité acquis dans les pays d'origine et à l'autonomie financière¹ des immigrants. Le tableau 3.1 présente un bref résumé des caractéristiques observables des immigrants de l'échantillon. Ce tableau montre que 81,4 % des immigrants sélectionnés en vertu de la grille de 1996 sont des diplômés universitaires comparativement à 80,7 % pour la grille de 2006. De même, le pourcentage d'immigrants ayant de l'expérience en lien avec leurs diplômes, varie très peu entre les grilles de sélection, c'est-à-dire 84 % et 82,1 % respectivement. Au niveau de l'âge, les adultes (tous les immigrants âgés de 30 ans et plus) de la grille de 2006 sont relativement moins âgés que ceux de 1996. En moyenne, le tableau 3.1 révèle que les adultes de 2006 sont âgés de 34 ans comparativement à 36 ans pour les adultes de 1996. Cependant, il convient de faire ressortir que l'âge moyen des jeunes est de 27 ans quelle que soit la grille de sélection. Concernant la moyenne du salaire horaire des immigrants ayant eu un premier emploi une fois au Québec, une différence de 1 \$ est constatée entre les immigrants des deux grilles. C'est à dire que les immigrants de la grille de 1996 gagnent en moyenne 15 \$ ($e^{2,7}$) et ceux de 2006 gagnent en moyenne 16 \$ ($e^{2,8}$). Les résultats des tests d'obtention de la résidence permanente révèlent qu'en moyenne tous les immigrants comprennent le français. En d'autres termes, les immigrants de 1996 et 2006 totalisent en moyenne 14/16 points en français. Une différence est quand même observée au niveau des adultes qui totalisent en moyenne des résultats (13/16) plus bas que ceux observés auprès des jeunes (14/16). Pour ce qui est des points obtenus en anglais, les adultes totalisent en moyenne 4/6 peu importe la grille alors que les jeunes immigrants de 2006 ont de meilleurs scores au test d'anglais que les jeunes de 1996, soit respectivement 5/6 et 4/6.

Le tableau 3.1 révèle également que, plus de la moitié des immigrants sélectionnés en vertu de la grille de 2006 n'ont pas participé aux formations formelles, soit 56,6 %. De plus, le pourcentage des jeunes ayant participé aux formations académiques est plus élevé que celui des adultes. En effet, les jeunes ont eu respectivement un taux de participation de 31,1 % et

1. Les immigrants doivent certifier par voie contractuelle qu'ils sont capables de subvenir non seulement à leurs besoins personnels, mais également à ceux de leurs familles pour une période minimale de trois mois au cours desquels ils ne bénéficient d'aucune aide sociale.

Tableau 3.1 – Caractéristiques observables des immigrants

	Jeunes		Adultes		Sans distinction	
	1996	2006	1996	2006	1996	2006
Scolarité (%)						
< Secondaire	0.3	-	0.4	0.5	0.3	0.5
Secondaire	1.9	2.7	1.2	1.8	1.7	1.9
Post-secondaire	17.6	8.2	14.1	17.6	16.6	16.9
Universitaire	80.3	89.1	84.3	80.1	81.4	80.7
Total	100	100	100	100	100	100
Points en Français (moy)	14/16	14/16	13/16	13/16	14/16	14/16
Points en Anglais (moy)	4/6	5/6	4/6	4/6	4/6	4/6
Log(salaire horaire) (moy)	2.8	2.8	2.7	2.8	2.7	2.8
Âge (moy)	27	27	36	34	34	32
Expérience (%)						
Non	23.5	24.5	13.6	14.1	16	17.9
Oui	76.5	75.5	86.4	85.9	84	82.1
Total	100	100	100	100	100	100
Formations Formelles (%)						
Sans formation	42.1	62.3	37	53.4	38.3	56.6
Linguistique	10	8.5	12.5	18.3	11.9	14.8
Professionnelle	16.8	11.3	28.3	16.2	25.5	14.5
Académique	31.1	18	22.2	12	24.4	14.1
Total	100	100	100	100	100	100

18 % aux formations académiques dans les grilles de 1996 et 2006 alors que les adultes de ces mêmes grilles ont participé respectivement à hauteur de 22,2 % et 12 %. En général, les taux de participation des immigrants aux formations formelles ont diminué en 2006 comparativement à ceux de 1996. Dans la grille de 1996, le taux de participation des jeunes aux formations formelles est de 57,9 % et s'établit à 37,7 % pour la grille de 2006. Le même fait est observé au niveau des adultes pour qui le taux de participation aux formations formelles dans la grille de 1996 est de 63 % alors qu'il s'établit à 46,6 % dans la grille de 2006.

Étant donné le niveau de scolarité, il ressort du tableau 3.2 que les diplômés sont majoritairement des adultes quelle que soit la grille de sélection. Par exemple, 75,5 % et 75 % des diplômés post-secondaires de la grille de 1996 et de celle de 2006 sont en moyenne des adultes alors que 24,5 % et 25 % sont respectivement des jeunes dans les deux grilles. Une exception est néanmoins constatée au niveau des diplômés du secondaire de 2006. La proportion des jeunes diplômés du secondaire de la grille de 2006 est plus élevée que celle des adultes de la même grille. En fait, 57,1 % des diplômés secondaires de la grille de 2006 sont des jeunes alors que 42,9 % des diplômés de la grille sont des adultes. En somme, les immigrants sélectionnés en vertu de la grille de 2006 sont relativement jeunes et plus scolarisés que ceux sélectionnés en vertu de la grille de 1996.

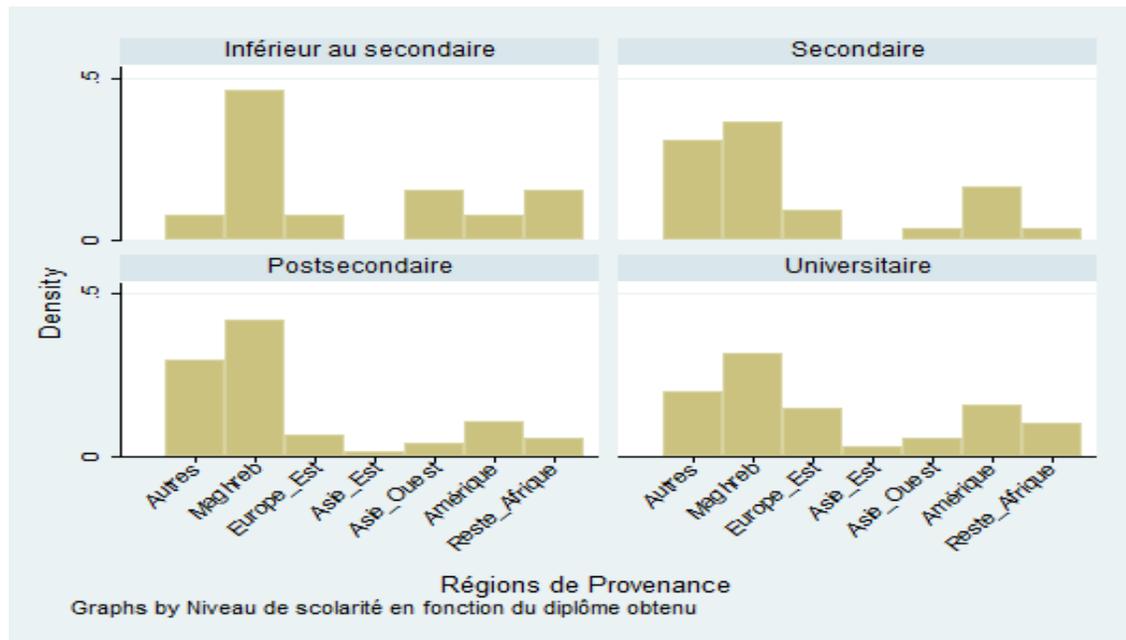
Tableau 3.2 – Caractéristiques des immigrants en fonction de la grille de sélection et de l'âge

Grille		Scolarité (%)			Expérience (%)			Formations formelles (%)			
		< Sec.	Sec.	> Sec.	Univ.	Non	Oui	Sans f.	Lingu.	Profes.	Acad.
1996	Jeunes	16.7	25	24.5	24.5	35.9	22.3	27.3	39.3	12.3	27
	Adultes	83.3	75	75.5	75.5	64.1	77.7	72.7	60.7	87.7	73
	Total	100	100	100	100	100	100	100	100	100	100
2006	Jeunes	0	57.1	25	37.9	50	33.6	39.3	23.1	26.3	39.4
	Adultes	100	42.9	75	62.1	50	66.4	60.7	76.9	73.7	60.6
	Total	100	100	100	100	100	100	100	100	100	100

Régions de provenance

L'objectif de cette partie est de faire ressortir les statistiques descriptives des immigrants en fonction de leurs régions de provenance. Pour ce faire, le graphique 3.1 et le tableau 3.3 présentent respectivement le niveau de scolarité des immigrants en fonction de leurs régions de provenance puis les fréquences observées auprès des immigrants selon le genre.

FIGURE 3.1 – Niveau de scolarité des immigrants selon leurs régions de provenance



En partant du graphique 3.1, il ressort clairement que les Maghrébins sont les plus scolarisés de la base de données. En fait, tous les diplômes obtenus sont marqués par une forte fréquence des Maghrébins. Après, viennent les immigrants en provenance des autres régions² ou des régions non spécifiées dans la base de données, puis les Américains en troisième position. Toutefois,

2. Par exemple l'Océanie, l'Europe de l'Ouest, l'Asie du Sud, etc...

lorsque le diplôme obtenu est inférieur au secondaire, les immigrants en provenance des Autres régions et les Américains sont supplantés par les Asiatiques de l'ouest et le reste de l'Afrique.

Tableau 3.3 – Fréquences observées auprès des immigrants selon le genre et les régions de provenance

Régions	Expérience		Autonomie fin.		Séjours		Emp_enquête	
	Non	Oui	Non	Oui	Non	Oui	Non	Oui
Reste de l'Afrique								
Femmes	23	42	10	55	32	33	20	45
Hommes	41	168	46	163	143	66	50	159
Amérique								
Femmes	18	148	55	111	99	67	46	120
Hommes	25	261	78	208	197	89	75	211
Asie Ouest								
Femmes	9	57	8	58	46	20	17	49
Hommes	11	84	14	81	54	41	24	71
Asie Est								
Femmes	7	29	15	21	25	11	14	22
Hommes	5	32	11	26	22	15	14	23
Europe Est								
Femmes	25	120	40	105	132	13	42	103
Hommes	36	214	112	138	236	14	44	206
Maghreb								
Femmes	55	226	97	184	245	36	97	184
Hommes	97	629	269	457	644	82	225	501
Autres								
Femmes	58	193	96	155	29	222	36	215
Hommes	80	316	147	249	49	347	44	352
Total								
Femmes	195	815	321	689	608	402	272	738
Hommes	295	1704	677	1322	1345	654	476	1523
Nombre d'obs.	490	2519	998	2011	1953	1056	748	2261

Le tableau 3.3 révèle que les Maghrébins ont plus d'expérience en lien avec leur diplôme avant l'obtention de leur résidence permanente. En fait, sur les 2 519 immigrants qui ont attesté avoir eu une expérience en lien avec leurs diplômes, environ 34 % (855/2519) d'entre eux venaient du Maghreb. De surcroît, sur les 2 011 immigrants autonomes financièrement, environ 34 % (641/2011) étaient en provenance du Maghreb. En outre, lors de la réalisation de l'enquête, environ 33 % des hommes en emploi étaient des Maghrébins. Concernant les séjours préalables au Québec, les Maghrébins en ont fait moins comparativement aux immigrants en provenance des régions inconnues et aux américains. Contrairement aux hommes, les femmes du Maghreb étaient moins en emploi au cours de l'enquête [25 % (184/738)] comparativement aux femmes en provenance des régions non spécifiées [29 % (215/738)]. Il convient de faire ressortir que les immigrants de l'Asie de l'Est sont minoritaires dans la base de données. En effet, sur les 3 009 immigrants, il n'y a que 73 Asiatiques de l'Est présents dans la base de données.

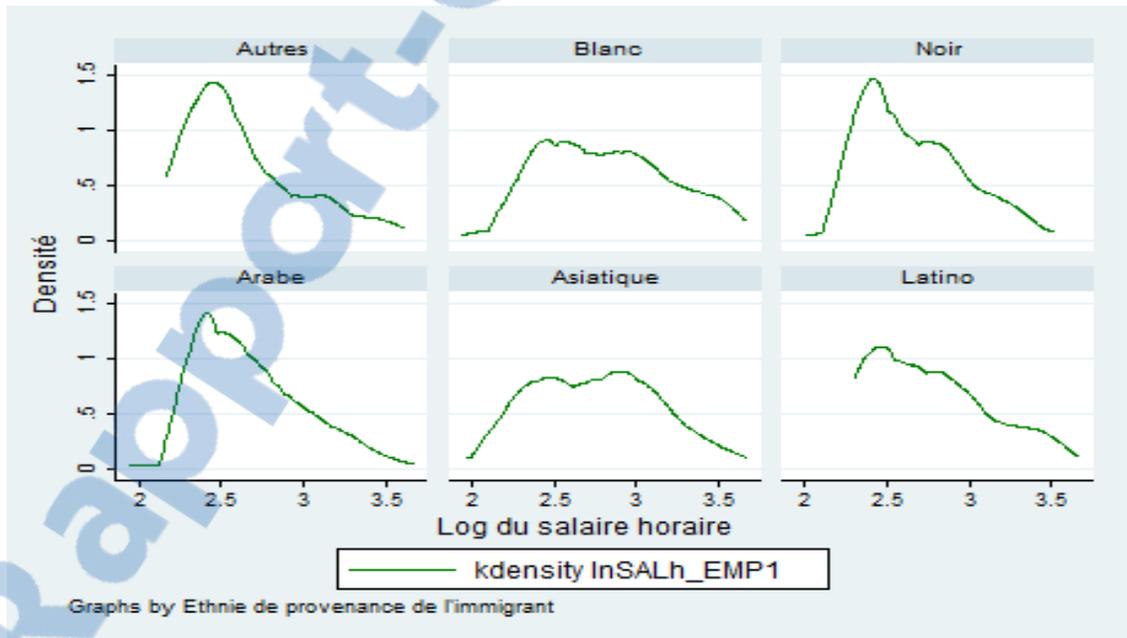
Revenus des immigrants

Dans la base de données, les salaires des immigrants ont été collectés sur une base horaire, journalière, hebdomadaire, bihebdomadaire, mensuelle, bimensuelle ou annuelle. Dans un premier temps tous les salaires ont été convertis sur une base horaire en considérant que :

- une année de travail s'étend sur 50 semaines ;
- un mois de travail s'étend sur 4 semaines ;
- une semaine de travail s'étend sur 5 jours ;
- une journée de travail s'étend sur 7 heures.

Une fois les salaires horaires obtenus, ils ont été minorés par le salaire horaire minimum (7 \$) au taux général du 1^{er} février 2001 fixé par la Commission des normes, de l'équité, de la santé et de la sécurité du travail. Le graphique 3.2 présente les fonctions de distribution du logarithme des salaires horaires du premier emploi des immigrants en fonction de leurs appartenances ethniques. Il en ressort qu'une forte proportion des Caucasiens et des Asiatiques de la base de données ont les revenus les plus élevés alors que les revenus les plus faibles sont localisés parmi toutes les minorités ethniques hormis les Latino-Américains et les immigrants des Autres régions.

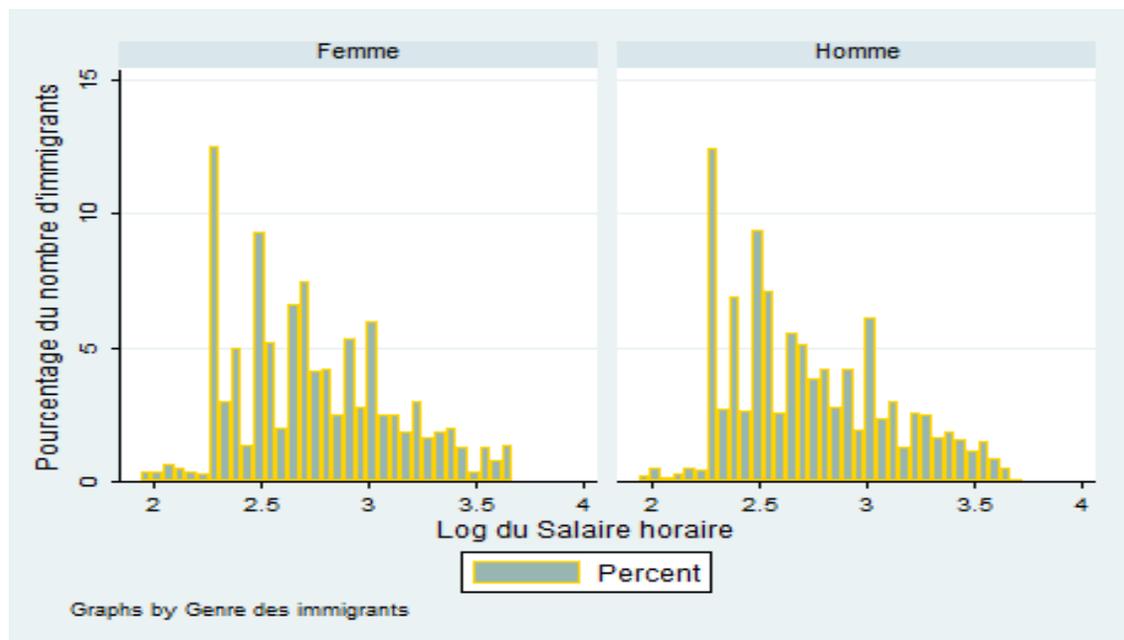
FIGURE 3.2 – Distribution du logarithme des salaires horaires des immigrants selon l'ethnicité



Le graphique 3.3 illustre les pourcentages du nombre d'immigrants en fonction de leurs salaires horaires en logarithme. Quel que soit le genre, il est observé que 12,5 % des hommes et des femmes gagnent environ 9 \$ ($e^{2,25}$) par heure au cours de leur premier emploi. Le gra-

phique 3.3 montre également qu'il n'y a pas de discrimination de revenu selon le genre parmi les immigrants qui ont obtenu un premier emploi, peu importe les intervalles de revenu.

FIGURE 3.3 – Répartition du logarithme des salaires horaires des immigrants selon le genre



3.2.2 Caractéristiques des formations formelles

Les formations formelles sont regroupées au sein d'une variable catégorique de 4 niveaux, le groupe de contrôle y compris. En fait, ces formations doivent mener à l'obtention de crédits, d'un diplôme ou d'une attestation délivrée par une institution d'enseignement, un ministère ou un organisme gouvernemental. Par ailleurs, la variable « Formations formelles » a été construite en tenant compte uniquement des immigrants qui ont strictement participé aux formations durant leurs périodes d'inactivités. Tous les immigrants qui cumulaient leur premier emploi avec une formation formelle ont été abandonnés pour éviter les problèmes de biais dans l'étude. En effet, les résultats de l'étude auraient compris l'impact de leur expérience acquise au cours du premier emploi si leur retrait n'avait pas été fait. Par ailleurs, des erreurs liées à l'entrée des dates d'arrivée, des dates de début des formations formelles, des dates de fin des formations formelles ainsi que des dates de début du premier emploi ont été constatées et corrigées grâce à la disponibilité du questionnaire d'enquête. Ces corrections ont été faites de manière à ne pas changer les dates de formation et les dates du premier emploi des personnes qui jumelaient la formation et l'emploi. Par conséquent, la variable « Formations formelles » regroupe 2 481 immigrants dont 1 004 d'entre eux n'ont suivi aucune formation et 1 477 parmi eux en ont suivi. Parmi les 528 immigrants absents de la variable, 41 immigrants ont suivi d'autres types de formations. Le reste cumulait soit un premier emploi et une formation formelle (135 immigrants) ou soit n'a pas connu de chômage avant l'obtention d'un

premier emploi (352 immigrants). En fait, une période d'inactivité correspond à un minimum de 14 jours sans emploi. Quant à la variable³ « Durée du chômage avant l'obtention du premier emploi », elle n'inclut pas la durée de la formation des personnes formées. En effet, la durée du chômage est le nombre de jours à temps plein passés à chercher le premier emploi.

Tableau 3.4 – Fréquences des qualifications des immigrants insérés dans le premier emploi

Formation formelles	Qualification du 1 ^{er} emploi			Total
	Surqualifié	Qualifié	Sous-qualifié	
Sans formation	422 (45.33 %)	470 (50.48 %)	39 (4.19 %)	931 (100 %)
Linguistique	187 (62.75 %)	91 (30.54 %)	20 (6.71 %)	298 (100 %)
Professionnelle	365 (61.66 %)	182 (30.74 %)	45 (7.6 %)	592 (100 %)
Académique	363 (63.68 %)	189 (33.16 %)	18 (3.16 %)	570 (100 %)
Total	1337 (55.92 %)	932 (38.98 %)	122 (5.1 %)	2391 (100 %)

Le tableau 3.4 montre les fréquences des qualifications des immigrants insérés dans un premier emploi à l'arrivée. Plus de la moitié des participants aux formations linguistiques, professionnelles et académiques étaient surqualifiés pour le premier emploi occupé, avec des proportions respectives de 62,75 %, 61,66 % et 63,68 %. Cependant, seulement 45,33 % des non-participants aux formations formelles étaient surqualifiés pour leur premier emploi. L'effet contraire est observé concernant les immigrants correctement qualifiés pour le premier emploi. En fait, la moitié (50,48 %) des non-participants étaient correctement qualifiés pour leur premier emploi comparativement à 30 % environ pour les participants de chaque formation formelle.

Les formations académiques

Le tableau 3.5 présente les difficultés rencontrées par les immigrants qui ont reçu une formation académique en fonction de leur domaine de diplomation. La variable « Génie » renferme tous les domaines d'études rattachés au Génie, en plus de l'agriculture et de l'architecture. Dans le domaine de la santé, 81,5 % des diplômés ont fait face aux difficultés entraînées par la réglementation sanitaire québécoise. D'une part, 22,2 % des diplômés en santé ont eu des problèmes de reconnaissance de leurs diplômes et d'autre part, 14,8 % des diplômés n'ont pas

3. Stata est le logiciel statistique qui a été utilisé pour le travail économétrique. Ce logiciel prend automatiquement comme date de référence le 1^{er} janvier 1960 pour affecter des nombres aux dates qui sont dans la base de données. Ainsi, pour un immigrant qui aurait obtenu un premier emploi le 12 décembre 2004, Stata affectera un nombre à cette date qui serait le nombre de jours séparant le 12 décembre 2004 au 1^{er} janvier 1960. Ce procédé ne modifie en rien les résultats et c'est ainsi que Stata arrive à faire des opérations mathématiques avec les dates.

Tableau 3.5 – Fréquences des immigrants selon les difficultés rencontrées et les domaines d'études

Formation formelle = Académique	Domaines d'obtention des diplômes						
	Santé	Génie	Gest&com	Sc_hum&so	Droit	Arts&tour	Total
Profession régie (%)							
Non	18.5	45.8	66.9	59.2	26.7	73.9	52.1
Oui	81.5	54.2	33.1	40.8	73.3	26.1	47.9
Total	100	100	100	100	100	100	100
Diplôme non reconnu (%)							
Non	77.8	84	85.5	86.7	86.7	91.3	84.9
Oui	22.2	16	14.5	13.3	13.3	8.7	15.1
Total	100	100	100	100	100	100	100
Manque d'expérience (%)							
Non	85.2	76	66.9	75.5	73.3	65.2	73.8
Oui	14.8	24	33.1	24.5	26.7	34.8	26.2
Total	100	100	100	100	100	100	100

pu obtenir leur premier emploi à cause de leur manque d'expérience québécoise souligné par les employeurs. Mais encore, une forte proportion des immigrants diplômés en droit a fait face à des problèmes de reconnaissance de diplôme, soit 73,3 %. Les domaines d'étude qui sont sujets à plus de difficultés liées au manque d'expérience au Québec sont premièrement l'art et le tourisme (34,8 %) puis la gestion et la communication (33,1 %). Il en résulte donc que les barrières à l'insertion au premier emploi varient selon les domaines d'étude.

Tableau 3.6 – Caractéristiques des immigrants selon le lieu d'obtention des diplômes

Formation formelle = Académique	Lieu d'obtention des diplômes				
	Canada	Europe	USA	Autres	Total
Âge des immigrants (moy)	28	32	35	33	32
Durée chômage avant 1 ^{er} emploi (jours)	1311	1503	1338	1500	1484
Log du Salaire horaire (moy)	2.8	2.8	2.7	2.7	2.7

Par ailleurs, le tableau 3.6 présente certaines caractéristiques des immigrants formés sur le plan académique selon le lieu d'obtention de leurs diplômes. Au premier abord, il appert que l'âge moyen des diplômés Américains qui ont participé aux formations académiques est de 35 ans, soit 3 ans de plus que la moyenne d'âge de l'ensemble des participants des formations académiques (32 ans). En moyenne, les nombres de jours de chômage des immigrants formés au Canada et aux États-Unis sont inférieurs à la moyenne de l'ensemble des participants aux formations académiques. Il se trouve que les immigrants formés dans les deux pays mettent en moyenne 3 ans pour obtenir un premier emploi. Enfin, les diplômés Canadiens et Européens qui ont reçu les formations académiques gagnent en moyenne 1 \$ de plus que les immigrants

diplômés des autres pays qui ont reçu le même type de formation.

Les formations linguistiques

Les formations linguistiques étudiées dans ce mémoire ne tiennent pas compte de la typologie des langues. Elle regroupe donc tous les immigrants ayant suivi des cours de perfectionnement d'une langue parlée au Québec. De ce fait, le tableau 3.7 se base sur l'autoévaluation des immigrants, sur leur connaissance des langues française et anglaise, pour faire ressortir les caractéristiques des personnes formées en langue. Il en ressort que les immigrants qui ont reçu les formations linguistiques pensent avoir au moins un niveau moyen dans les langues française et anglaise. En revanche, très peu de ces immigrants pensent avoir des problèmes d'écriture de la langue française, soit 2 %. La question émanant de ces statistiques est : quelle est donc la raison qui les a poussés à suivre des cours de langue ?

Tableau 3.7 – Caractéristiques linguistiques des immigrants avant l'arrivée

Cours de langue	Très faible	Faible	Niveau			Total
			Moyen	Bon	Très bon	
Comprendre (%)						
Français	-	1.6	10.5	38.8	49	100
Anglais	2.6	12.2	25.3	29.3	30.6	100
Lecture (%)						
Français	-	1	4.6	35.9	58.6	100
Anglais	2.3	9.5	17.8	38.5	31.9	100
Parler (%)						
Français	-	1.6	15.5	44.4	38.5	100
Anglais	4.6	17.1	28.9	26.0	23.4	100
Écrire (%)						
Français	2	6.3	17.1	39.1	35.5	100
Anglais	5.3	13.5	26	35.2	20.1	100

Tableau 3.8 – Niveaux linguistiques des immigrants formés avant l'arrivée et difficultés rencontrées lors de la recherche d'emploi au Québec

Cours de langue	Tests avant l'arrivée (moy)		Insuffisance linguistique (%)		
	Français	Anglais	Français	Anglais	Bilinguisme
Sans formation	14/16	4/6	18.6	21.5	28.2
Linguistique	11/16	4/6	46.5	22.2	21.8
Professionnelle	14/16	3/6	19.4	28.5	33.1
Académique	14/16	4/6	15.5	27.8	16.9
Total	14/16	4/6	100	100	100

Le tableau 3.8 nous renseigne sur les résultats des tests linguistiques effectués par les immigrants avant leur arrivée au Québec. Comparativement aux immigrants des autres formations

formelles, il appert que les participants aux formations linguistiques ont eu le score le plus faible au test de français, soit 11/16. De surcroît, 46,5 % des immigrants formés en langue ont rencontré des barrières à l’insertion au premier emploi dues à une insuffisance de la connaissance de la langue française. En somme, le tableau 3.8 montre que les immigrants qui ont participé aux formations linguistiques avaient des problèmes en français et un peu moins en anglais.

Les formations professionnelles

Le tableau 3.9 révèle que les immigrants qui ont rencontré des difficultés de reconnaissance de diplômes ou qui n’avaient pas d’expérience au Québec se sont beaucoup plus dirigés vers les formations professionnelles. Par exemple, sur 1 127 immigrants qui ont rencontré des problèmes dus à leurs professions qui sont régies par un ordre professionnel, 26,1 % d’entre eux se sont dirigés vers les formations professionnelles en vue d’obtenir un emploi. Cette proportion observée est la plus élevée des trois formations formelles.

Tableau 3.9 – Fréquences des participants des formations professionnelles

Formation formelle = Professionnelle	Profession régie		Manque d’expérience		Niveau de scolarité			
	Non	Oui	Non	Oui	<Sec.	Sec.	>Sec.	Univ.
Sans formation (%)	598 44.2	406 36	833 44.5	171 28.1	8 80	24 50	172 41.1	800 39.9
Linguistique (%)	148 10.9	156 13.8	202 10.8	102 16.7	- -	2 4.2	28 6.7	274 13.7
Professionnelle (%)	305 22.5	294 26.1	411 22	188 30.9	- -	15 31.3	165 39.4	419 20.9
Académique (%)	303 22.4	271 24	426 22.8	148 24.3	2 20	7 14.6	54 12.9	511 25.5
Total	1354 100	1127 100	1872 100	609 100	10 100	48 100	419 100	2004 100

En outre, 188 immigrants parmi les 609 immigrants qui n’avaient pas d’expérience au Québec, ont également choisi de suivre une formation professionnelle. Enfin, le tableau 3.9 nous montre que les immigrants des formations professionnelles sont en majorité des diplômés universitaires et postsecondaires. En effet, ces derniers ont participé dans des proportions respectives de 39,4 % et 20,9 % à une formation professionnelle. Il convient de noter que la participation des diplômés postsecondaires représente environ 67 % (165/247) de tous les 247 immigrants de niveau postsecondaires qui ont participé aux formations formelles. De même, les formations professionnelles sont marquées par une proportion non négligeable des 1 204 diplômés universitaires qui ont participé aux formations formelles. En effet, environ 35 % (419/1204) de ces diplômés ont suivi une formation professionnelle.

Chapitre 4

Résultats

4.1 Présentation des résultats

Dans cette section, il sera présenté premièrement les résultats de la durée du chômage avant l'obtention du premier emploi et ensuite les résultats concernant le revenu des immigrants.

4.1.1 La durée du chômage

Le tableau 4.1 présente les résultats de l'appariement multiple et de la régression par quantile pour la durée du chômage. Il peut y être vu que les effets moyens de la participation aux formations formelles sont tous significatifs. Après avoir postulé l'hypothèse du support commun, le nombre d'observations retenu dans l'appariement est de 2 480 immigrants. En moyenne, il ressort des résultats que les formations formelles baissent la durée du chômage des participants comparativement aux non-participants. Ainsi, la participation aux formations linguistiques baisse de 580 jours en moyenne la durée de chômage des participants comparativement aux non-participants. Cet effet est moins élevé dans les deux autres formations formelles avec l'effet le plus faible observé au niveau des formations académiques. En fait, les formations professionnelles baissent de 422 jours en moyenne la durée du chômage des participants alors que celui des participants aux formations académiques est de 304 jours.

Le graphique 4.1 illustre bien la démarcation observée entre l'effet moyen des formations linguistiques et les effets moyens des autres formations. Si les participants des formations linguistiques avaient reçu une formation académique, cette dernière aurait réduit leur baisse du nombre de jours de chômage de 276 jours en moyenne. Cette réduction aurait été de 158 jours si les participants des formations linguistiques avaient reçu une formation professionnelle et de 118 jours si les participants des formations professionnelles avaient reçu une formation académique.

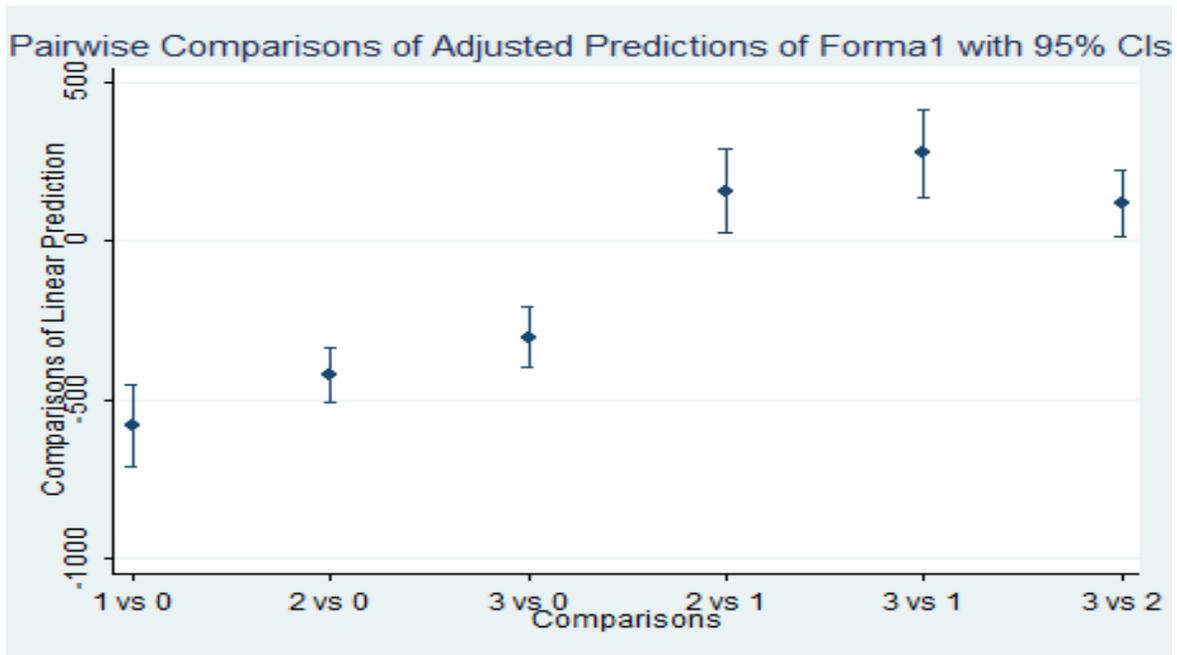
Les résultats de la régression par quantile révèlent que les formations formelles réduisent plus le nombre de jours de chômage de la médiane des distributions que celui de la moyenne

Tableau 4.1 – Effets estimés des comparaisons inter-niveaux de formations formelles pour la durée du chômage

Durée du chômage		Delta-method	Unadjusted
	Contrast	Std. Err.	[95% Conf. Interval]
mean			
Linguistique Vs Sans f.	-580	65.88	-708.65 -450.39
Professionnelle Vs Sans f.	-422	45.25	-510.60 -333.24
Académique Vs Sans f.	-304	49.22	-400.25 -207.32
Professionnelle Vs Lingu.	158	68.59	23.18 292.03
Académique Vs Lingu.	276	71.29	136.01 415.46
Académique Vs Profes.	118	52.68	14.89 221.38
q25			
Linguistique Vs Sans f.	-591	83.78	-755.21 -426.79
Professionnelle Vs Sans f.	-394	44.76	-481.72 -306.28
Académique Vs Sans f.	-243	51.98	-344.89 -141.11
Professionnelle Vs Lingu.	197	84.75	30.90 363.10
Académique Vs Lingu.	348	87.16	177.17 518.83
Académique Vs Profes.	151	47.56	57.79 244.21
q50			
Linguistique Vs Sans f.	-690	101.86	-889.64 -490.36
Professionnelle Vs Sans f.	-477	88.03	-649.52 -304.47
Académique Vs Sans f.	-381	73.99	-526.03 -235.97
Professionnelle Vs Lingu.	213	119.14	-20.51 446.51
Académique Vs Lingu.	309	108.24	96.86 521.14
Académique Vs Profes.	96	95.23	-90.65 282.65
q75			
Linguistique Vs Sans f.	-629	188.37	-998.20 -259.80
Professionnelle Vs Sans f.	-491	74.97	-637.94 -344.06
Académique Vs Sans f.	-495	76.22	-644.38 -345.62
Professionnelle Vs Lingu.	138	204.99	-263.77 539.77
Académique Vs Lingu.	134	200.93	-259.82 527.82
Académique Vs Profes.	-4	101.09	-202.14 194.14

des distributions. Il convient de rappeler que les résultats sont interprétés sous l'hypothèse de l'invariance des rangs, c'est-à-dire que, l'effet observé est la différence entre l'effet quantile estimé quand tous les individus d'un quantile participeraient à une formation formelle et l'effet quantile estimé quand tous les individus dudit quantile ne participeraient à aucune formation. De plus, les écarts-types des estimateurs quantiles ont été obtenus par 1 000 tirages de Monte-Carlo. Les résultats de comparaison entre les formations formelles et le groupe de contrôle sont tous significatifs et statistiquement différents de 0. Les effets des formations linguistiques, professionnelles et académiques sur la médiane des distributions de la durée du chômage sont respectivement de - 690, - 477 et - 243 jours de chômage. Lorsque les formations professionnelles et académiques sont prises en compte, les effets du 25^e quantile sont respectivement de - 394 et - 294 jours de chômage. Ensuite, ceux de la médiane sont respectivement de - 477 et - 381 jours de chômage, puis, ceux du 75^e quantile sont de - 491 et - 495 jours de chômage. Il est donc plus avantageux aux immigrants du 75^e quantile des distributions de participer aux formations professionnelles et académiques comparativement aux immigrants de la médiane des distributions et du 25^e quantile des distributions. Les immigrants du 75^e quantile sont ceux qui ont plus duré au chômage avant d'obtenir un premier emploi. Concernant les effets croisés par quantile, les résultats obtenus pour le 25^e quantile sont significatifs alors que la

FIGURE 4.1 – Intervalles de confiance des comparaisons inter-niveaux pour la durée du chômage



Légende : 0 = Sans formation ; 1 = Linguistique ; 2 = Professionnelle ; 3 = Académique .

plupart des résultats obtenus pour le 75^e quantile et la médiane ne le sont pas. Toutefois, l'effet quantile observé pour le passage des formations linguistiques aux formations académiques des immigrants de la médiane des distributions est significatif et s'établit à 309 jours.

Tableau 4.2 – Résultats du test d'égalité des effets pour la durée du chômage

Variable dépendante	Test
Durée du chômage	$[\text{mean}]^1 \text{ Sans f.} - [\text{mean}] \text{ Lingu.} = 0$
	$[\text{mean}] \text{ Sans f.} - [\text{mean}] \text{ Profes.} = 0$
	$[\text{mean}] \text{ Sans f.} - [\text{mean}] \text{ Acad.} = 0$
	$[\text{q25}]^2 \text{ Sans f.} - [\text{q25}] \text{ Lingu.} = 0$
	$[\text{q25}] \text{ Sans f.} - [\text{q25}] \text{ Profes.} = 0$
	$[\text{q25}] \text{ Sans f.} - [\text{q25}] \text{ Acad.} = 0$
	$[\text{q50}]^3 \text{ Sans f.} - [\text{q50}] \text{ Lingu.} = 0$
	$[\text{q50}] \text{ Sans f.} - [\text{q50}] \text{ Profes.} = 0$
	$[\text{q50}] \text{ Sans f.} - [\text{q50}] \text{ Acad.} = 0$
	$[\text{q75}]^4 \text{ Sans f.} - [\text{q75}] \text{ Lingu.} = 0$
	$[\text{q75}] \text{ Sans f.} - [\text{q75}] \text{ Profes.} = 0$
	$[\text{q75}] \text{ Sans f.} - [\text{q75}] \text{ Acad.} = 0$
	$\text{chi2}(12) = 220.39$
$\text{Prob} > \text{chi2} = 0.0000$	

1. désigne l'estimateur moyen pour un niveau de traitement donné.
2. désigne l'estimateur du 25^e quantile pour un niveau de traitement donné.
3. désigne l'estimateur de la médiane pour un niveau de traitement donné.
4. désigne l'estimateur du 75^e quantile pour un niveau de traitement donné.

Le test observé dans le tableau 4.2 a été effectué pour vérifier l’hypothèse nulle d’égalité entre les effets moyens et quantiles des différents niveaux de la variable « Formations formelles ». Les résultats du test rejettent l’hypothèse nulle postulée avec la P-value du test étant égal à 0. L’impact différencié, selon les différents niveaux de la variable « Formations formelles », est donc réel.

4.1.2 Le revenu

Le tableau 4.3 présente les résultats des effets moyens et quantiles du traitement multiple sur les salaires horaires du premier emploi des participants. En moyenne, les non-participants gagnent le salaire horaire le plus élevé, soit 16 \$ ($e^{2.8}$) et les participants des formations professionnelles gagnent le salaire horaire le plus faible, soit 14 \$ ($e^{2.65}$). En comparaison au groupe de contrôle, les formations linguistiques, professionnelles et académiques baissent respectivement les salaires horaires des participants de 4 %, de 14 % et de 9,5 % en moyenne. Néanmoins, seuls les résultats des formations professionnelles et académiques sont significatifs.

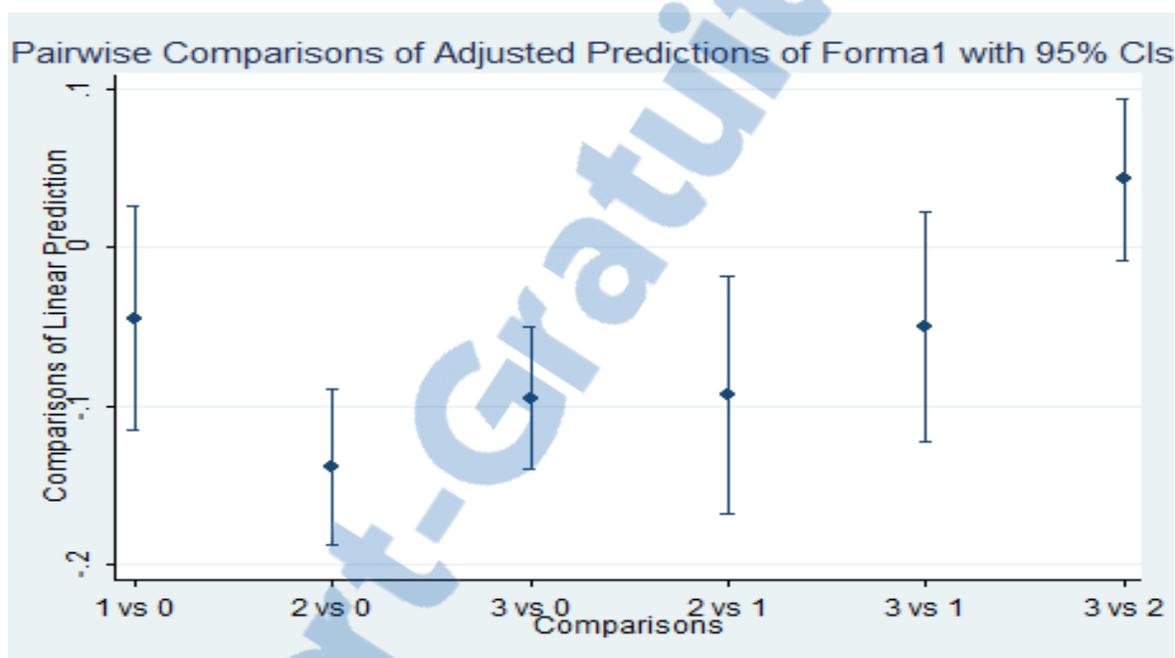
Tableau 4.3 – Effets estimés des comparaisons inter-niveaux des formations formelles pour le revenu

Log(Sal. horaire du 1 ^{er} emploi)		Delta-method	Unadjusted
	Contrast	Std. Err.	[95% Conf. Interval]
mean			
Linguistique Vs Sans f.	-.045	.036	-.116 .026
Professionnelle Vs Sans f.	-.138	.025	-.187 -.089
Académique Vs Sans f.	-.095	.023	-.140 -.050
Professionnelle Vs Lingu.	-.093	.038	-.168 -.019
Académique Vs Lingu.	-.050	.037	-.123 .022
Académique Vs Profes.	.043	.026	-.008 .093
q25			
Linguistique Vs Sans f.	-.043	.030	-.101 .016
Professionnelle Vs Sans f.	-.110	.026	-.161 -.059
Académique Vs Sans f.	-.087	.022	-.130 -.044
Professionnelle Vs Lingu.	-.067	.035	-.136 .001
Académique Vs Lingu.	-.044	.032	-.108 .019
Académique Vs Profes.	.023	.029	-.033 .079
q50			
Linguistique Vs Sans f.	-.093	.070	-.231 .045
Professionnelle Vs Sans f.	-.165	.044	-.252 -.078
Académique Vs Sans f.	-.129	.046	-.219 -.038
Professionnelle Vs Lingu.	-.071	.070	-.210 .067
Académique Vs Lingu.	-.035	.070	-.172 .102
Académique Vs Profes.	.036	.044	-.049 .122
q75			
Linguistique Vs Sans f.	-.082	.082	-.243 .078
Professionnelle Vs Sans f.	-.244	.043	-.328 -.159
Académique Vs Sans f.	-.134	.047	-.225 -.042
Professionnelle Vs Lingu.	-.161	.080	-.318 -.004
Académique Vs Lingu.	-.051	.083	-.214 .112
Académique Vs Profes.	.110	.046	.020 .200

En partant des comparaisons croisées entre les formations formelles, il appert que le passage

des formations linguistiques aux formations professionnelles diminue de 9,5 % en moyenne le salaire horaire des participants des formations linguistiques. Il convient de noter que les résultats des comparaisons pour les autres formations ne sont pas significatifs. Les intervalles de confiance des résultats de comparaison sont illustrés par le graphique 4.2. D'une part, ce graphique dévoile que les comparaisons d'effets entre les formations académiques et linguistiques ne sont pas significatives, ainsi que les comparaisons entre les formations académiques et professionnelles. D'autre part, il montre que les effets moyens des formations linguistiques ne sont pas significatifs.

FIGURE 4.2 – Intervalles de confiance des comparaisons inter-niveaux pour le revenu



Légende : 0 = Sans formation ; 1 = Linguistique ; 2 = Professionnelle ; 3 = Académique .

Quant à la régression par quantile, les résultats obtenus sont étroitement liés à ceux de l'appariement multiple. Les interprétations des effets quantiles sont faites sur la base de l'hypothèse d'invariance des rangs. Peu importe les quantiles, les formations formelles ont un impact négatif sur les quantiles des participants. Ainsi, les formations linguistiques, professionnelles et académiques baissent respectivement de 4 %, 11 % et 9 % les salaires horaires du 25^e quantile des distributions du revenu des participants par rapport au 25^e quantile du groupe de contrôle. Ces baisses sont de 9 %, 16 % et 13 % pour la médiane et de 8 %, 24 % et 13 % pour le 75^e quantile. Par contre, les résultats des formations linguistiques ne sont pas significatifs. Bien que certains résultats des comparaisons croisées entre les formations formelles ne soient pas significatifs, une baisse de salaire horaire significative de 11 % est observé sur le 75^e quantile du revenu des immigrants lorsque ces derniers participent aux formations académiques au lieu de participer aux formations professionnelles. En somme, les résultats révèlent que les formations formelles n'améliorent pas les revenus des participants. Toutefois, les formations formelles di-

Tableau 4.4 – Résultats du test d'égalité des effets pour le revenu

Variable dépendante	Test
Log(Sal.horaire.emp_1)	[mean] ⁵ Sans f. - [mean] Lingu. = 0
	[mean] Sans f. - [mean] Profes. = 0
	[mean] Sans f. - [mean] Acad. = 0
	[q25] ⁶ Sans f. - [q25] Lingu. = 0
	[q25] Sans f. - [q25] Profes. = 0
	[q25] Sans f. - [q25] Acad. = 0
	[q50] ⁷ Sans f. - [q50] Lingu. = 0
	[q50] Sans f. - [q50] Profes. = 0
	[q50] Sans f. - [q50] Acad. = 0
	[q75] ⁸ Sans f. - [q75] Lingu. = 0
	[q75] Sans f. - [q75] Profes. = 0
	[q75] Sans f. - [q75] Acad. = 0
	chi2(12) = 53.47
	Prob > chi2 = 0.0000

minuent moins les salaires horaires du 25^e quantile (les immigrants ayant les revenus les plus faibles) dans l'échantillon d'étude.

Le tableau 4.4 présente les résultats du test d'égalité entre les effets moyens et quantiles des différents niveaux de la variable « Formations formelles » pour le revenu. L'hypothèse nulle du test, qui stipule que les effets quantiles de ces niveaux sont égaux, est rejetée avec la P-value du test inférieure au seuil de significativité de 1 %. Par conséquent, il existe un impact quantile différencié entre les différents niveaux de la variable « Formations formelles »

4.2 Discussion des résultats

Les résultats de cette étude montrent que les formations linguistiques ont un impact négatif sur la durée du chômage des participants, soit un impact positif sur l'emploi. Les travaux de Dustmann et Fabbri (2003) [9] aboutissent à un impact positif de la connaissance de l'anglais sur l'emploi. En effet, ils trouvent que l'anglais augmente environ de 15 % les probabilités d'être en emploi des minorités ethniques sélectionnés à partir des enquêtes nationales en Angleterre de 1993 à 1995.

Ensuite, les résultats révèlent que les formations professionnelles baissent la durée de chômage des participants de 118 jours en moyenne, comparativement à la durée de chômage des participants des formations académiques. Autrement dit, l'impact des formations professionnelles est positif sur l'insertion en emploi. De même, Hanushek et al. (2011) [15] trouvent que, les forma-

5. désigne l'estimateur moyen pour un niveau de traitement donné.

6. désigne l'estimateur du 25^e quantile pour un niveau de traitement donné.

7. désigne l'estimateur de la médiane pour un niveau de traitement donné.

8. désigne l'estimateur du 75^e quantile pour un niveau de traitement donné.

tions professionnelles augmentent de 7 % la probabilité de s'insérer en emploi des participants comparativement à celui des participants des formations académiques.

Quant aux résultats obtenus pour le revenu, l'analyse descriptive des participants des formations formelles et de leurs contrefactuels révèle que ces immigrants étaient en majorité des diplômés universitaires et postsecondaires avant l'immigration. Il convient de rappeler que cette étude porte sur des immigrants de la classe des travailleurs qualifiés du Québec. La différence salariale observée entre les participants des formations formelles et leurs contrefactuels serait due à la surqualification des participants. Ainsi, les résultats de l'étude de Kilolo-Malambwe (2013) [18] pourraient justifier cette affirmation. En fait, il montre qu'un travailleur surqualifié occupe un emploi qui nécessite des compétences inférieures aux siennes alors que la rémunération est souvent en fonction des compétences requises pour l'emploi et non en fonction des compétences du travailleur. Les résultats de son étude montrent que les travailleurs surqualifiés gagnent 13 % de moins que les travailleurs correctement qualifiés. Dans la présente étude, l'analyse descriptive de la base de données dévoile qu'environ 60 % des participants des formations formelles étaient surqualifiés pour leur premier emploi alors que la moitié des non-participants (50,48 %) étaient correctement qualifiés pour leur premier emploi¹. Ce qui justifierait pourquoi les non-participants des formations formelles gagnent plus que les participants une fois en emploi.

Lorsque les formations académiques sont dispensées à des travailleurs qui possèdent déjà les qualifications requises, Krueger et Rouse (1994) [19] montrent que l'impact de ces formations sur le revenu de ces travailleurs est faible. Dans la présente étude, les participants aux formations académiques étaient très scolarisés avant leur arrivée au Québec, ce qui expliquerait le faible impact des formations académiques québécoises sur leurs salaires horaires.

Gonzalez (2000) [14] montre que l'âge des immigrants est un déterminant non-négligeable qui affecte l'impact des cours de langue sur le revenu des participants. En effet, les résultats de son étude révèlent que l'impact des cours de langue était négatif sur les revenus des immigrants qui étaient arrivés aux États-Unis au delà de 30 ans en 1992, soit une baisse environ de 10 % comparativement à ceux qui ont immigré aux États-Unis à l'âge de 20 ans. De même, les statistiques descriptives de la présente étude révèlent que l'âge moyen des participants aux formations linguistiques était de 34 ans environ. De plus, les non-participants avaient en moyenne un score de 14/16 au test de français alors que les participants des formations linguistiques avaient le score le plus faible au test de français, soit 11/16. En somme, le niveau de connaissance du français des non-participants était plus élevé que celui des participants qui seraient moins susceptibles de maîtriser la langue française une fois au Québec, étant donné leur âge moyen. Ce qui expliquerait aussi le faible impact des formations linguistiques sur le revenu des participants.

1. Voir le tableau ??

Conclusion

Cette étude cherche à évaluer l'adéquation entre les formations formelles et la situation économique des immigrants de la classe des travailleurs qualifiés à leur arrivée au Québec. Pour ce faire, l'impact des formations formelles est évalué sur le salaire horaire et la durée du chômage avant l'obtention du premier emploi, aux moyens de l'appariement multiple et de la régression par quantile. Selon Gertler et al. (2011) [13] la première méthode est recommandée lorsqu'il n'existe aucune règle d'assignation claire aux formations formelles. La régression par quantile a été utilisée pour désagréger, au niveau des différents quantiles de l'échantillon d'étude, les résultats obtenus par l'appariement multiple.

Maintes études se sont intéressées à l'évaluation d'impact des formations linguistiques, professionnelles puis académiques sur le revenu et la durée du chômage des participants. Leurs résultats révèlent que l'impact de chacune des formations formelles varie en fonction du niveau de scolarité, du lieu de formation, de l'ethnicité, de l'âge, de l'expérience, du domaine de la formation et de l'emploi.

La base de données employée dans cette étude a été fournie par le Ministère de l'Emploi, de la Solidarité Sociale et de la Famille du Québec en collaboration avec le Ministère de l'Immigration et des Communautés Culturelles (MICC). Le nombre total d'observations de la base est de 3 009 immigrants dont 1 477 ont pris part aux formations formelles. Comparativement au 1 004 contrefactuels réunis dans le groupe de contrôle, les résultats de l'appariement multiple révèlent que les formations linguistiques, professionnelles et académiques diminuent respectivement le nombre de jours de chômage (avant l'obtention du premier emploi) des participants de 580, 422 et 304 jours en moyenne. De plus, les formations linguistiques baissent plus la durée de chômage de la médiane des distributions alors que les formations professionnelles et académiques baissent plus la durée du chômage du 75^e quantile des distributions, soient des baisses respectives de 690, 491 et 495 jours.

Cependant, l'impact moyen des formations formelles sur les salaires horaires du premier emploi des participants est négatif, avec une faible baisse des salaires horaires des participants du 25^e quantile. Ainsi, seuls les résultats des formations professionnelles et académiques sont significatifs, soit des baisses moyennes de 1 \$ environ. Les études de Dustmann et Fabbri (2003)[9], Hanushek et al. (2011)[15] pourraient justifier les résultats obtenus pour la durée de chômage

alors que celles de Gonzalez (2000) [14] et de Kilolo-Malambwe (2013) [18] pourraient expliquer les résultats obtenus pour le revenu des immigrants. En effet, les formations formelles n'augmentent pas le revenu des participants insérés en emploi. Ce résultat serait dû à la surqualification des participants aux formations formelles. De plus, l'âge avancé des participants des formations linguistiques serait l'une des cause du faible impact de ladite formation sur les revenus des participants.

De tout ce qui précède, le gouvernement du Québec doit bien clarifier ses objectifs d'immigration afin de savoir comment orienter ses politiques. Si l'objectif principal du gouvernement était d'améliorer le niveau de revenu des immigrants, les structures d'orientation des immigrants ne devraient pas diriger ces derniers vers les formations formelles. Mais, si le gouvernement souhaitait augmenter l'employabilité des immigrants, il aurait pu les orienter vers les formations formelles en tenant compte des profils des immigrants. En effet, le gouvernement devrait rediriger les immigrants appartenant au 75^e quantile des distributions de la durée du chômage vers les formations professionnelles et académiques. Puis, ceux appartenant à la médiane vers les formations linguistiques. Dans ce cas-ci, une formation des conseillers d'orientation des immigrants pourrait se faire afin de faciliter la bonne marche des politiques qui seront mises en place. Cette formation permettrait au personnel d'orientation de pouvoir identifier les immigrants pour lesquels les formations formelles auraient un impact significatif.

Néanmoins, pour une optimisation des ressources, le gouvernement devrait sélectionner les immigrants qui ont des scores élevés aux tests linguistiques et orienter ceux qui ont les scores moyens à ces tests vers les formations linguistiques parce qu'elles diminuent plus la durée du chômage. Cette politique permettrait de réduire les ressources alloués aux formations formelles. Pour ce qui est du rattrapage salarial, le gouvernement devrait chercher la meilleure politique qui permettrait de réduire l'écart salarial entre les participants et les non-participants des formations formelles.

Annexe A

Équations

Dans les équations qui vont suivre, $\mathbf{D}(t) = \{0 \text{ si } t = 0 \text{ et } 1 \text{ si } t \neq 0\}$ est un vecteur de dimension 4×1 qui permet de distinguer le groupe de contrôle et le groupe des participants, Y_i est la variable d'intérêt observée pour chaque individu de la base de données, puis $Y(t)$ est la somme des Y de tous les individus ayant participé au traitement t . De plus, $r(t, \mathbf{X})$ est un scalaire, c'est-à-dire le score de propension généralisé attribué au traitement t en fonction des caractéristiques observables (\mathbf{X}) des immigrants qui ont participé audit traitement. \mathbf{X} est une matrice de dimension $2\,481 \times 14$ qui regroupe l'ensemble des 14 caractéristiques observables retenues pour les 2 481 individus présents dans la variable « Formations formelles ».

A.1 Les conditions du premier moment

La vérification des conditions du premier moment est très importante pour la validation des estimateurs moyens et quantiles. Pour chaque traitement t , ces conditions stipulent qu'en moyenne, les fonctions influentes, partielle ($\phi_{(ipw,t)}$) et complète ($\phi_{(eif,t)}$) doivent sommer à 0 lorsque tous les individus ayant participé audit traitement sont pris en compte. D'une part, les fonctions influentes sont des procédures d'estimation proposées par Cattaneo (2010) [5] pour construire les estimateurs des effets moyens et quantiles du traitement multiple. D'autre part, ces fonctions influentes permettent d'effectuer des inférences conjointes sur les effets moyens et quantiles estimés. Dans les sous-sections suivantes, une présentation détaillée des conditions du premier moment étayera les paragraphes.

A.1.1 Les conditions du premier moment des estimateurs moyens

En partant des deux procédures d'estimation, les conditions du premier moment des estimateurs sont données par :

$$E[\phi_{(ipw,t)}] = 0 \text{ avec } \phi_{(ipw,t)} = \frac{\mathbf{D}_i(t)(Y_i - \mu_t)}{r(t, \mathbf{X})} \text{ et } \mu_t = E[Y(t)]; \quad (\text{A.1})$$

$$E[\phi_{(eif,t)}] = 0 \text{ avec } \phi_{(eif,t)} = \frac{\mathbf{D}_i(t)(Y_i - \mu_t) - e_t(X_i, \mu_t)[\mathbf{D}_i(t) - r(t, \mathbf{X})]}{r(t, \mathbf{X})}; \mu_t = E[Y(t)] \quad (\text{A.2})$$

et $e_t(X_i, \mu_t) = E[Y_i(t) - \mu_t | X_i, T = t]$.

La différence majeure entre les deux équations est la fonction non paramétrique $e_t(X_i, \mu_t)$ qui se retrouve dans la fonction influente complète $\phi_{(eif,t)}$. En effet, cette fonction non paramétrique est la différence moyenne entre la somme et la moyenne des Y des participants au traitement t , étant données leurs caractéristiques observables et ledit traitement. Pour chaque traitement t , la fonction influente partielle $\phi_{(ipw,t)}$ est obtenue en pondérant, la différence entre les Y_i et la moyenne (μ_t) des Y des participants, par le score généralisé du traitement. La fonction influente complète $\phi_{(eif,t)}$ quant à elle, est obtenue en rajoutant la fonction non paramétrique $\phi_{(eif,t)}$ à la fonction influente partielle.

A.1.2 Les conditions du premier moment des estimateurs quantiles

Les estimateurs quantiles obtenus par les deux procédures d'estimation semi-paramétriques doivent également respecter les conditions du premier moment à savoir :

$$E[\phi_{(ipw,t)}] = 0 \text{ avec } \phi_{(ipw,t)} = \frac{\mathbf{D}_i(t) \left(1_I(Y_i \leq Q_t(\tau)) - \tau \right)}{r(t, \mathbf{X})}; \quad (\text{A.3})$$

$$E[\phi_{(eif,t)}] = 0 \text{ avec } \phi_{(eif,t)} = \frac{\mathbf{D}_i(t) \left(1_I(Y_i \leq Q_t(\tau)) - \tau \right) - e_t(X_i, Q_t(\tau))[\mathbf{D}_i(t) - r(t, \mathbf{X})]}{r(t, \mathbf{X})} \quad (\text{A.4})$$

et $e_t(X_i, Q_t(\tau)) = E \left(1_I[Y_i(t) \leq Q_t(\tau)] - \tau | X_i, T = t \right)$.

Dans les équations (A.3) et (A.4), tau (τ) est la proportion dans laquelle sont scindés les 100 % de probabilité totale de l'échantillon d'étude. $Y_i(t)$ est le Y d'un participant du traitement t . $1_I(Y_i \leq Q_t(\tau))$ désigne les participants au traitement t qui sont inclus dans le τ^{e} quantile.

Pour chaque τ^{e} quantile et chaque traitement t , l'équation (A.3) révèle que la fonction influente partielle $\phi_{(ipw,t)}$ est obtenue en pondérant, la différence entre $1_I(Y_i \leq Q_t(\tau))$ et la proportion

τ , par le score généralisé du traitement. En revanche, la fonction influente complète est obtenue en rajoutant la fonction non paramétrique $e_t(X_i, Q_t(\tau))$ à la fonction influente partielle. Étant données les caractéristiques observables des participants au traitement t , cette fonction non paramétrique n'est rien d'autre que la différence moyenne entre $1_I(Y_i \leq Q_t(\tau))$ et la proportion (τ).

A.2 Détermination des quantiles

La détermination des quantiles passe par la minimisation selon le modèle général pondéré et l'identification des quantiles aux moyens des méthodes semi-paramétriques proposées par Firpo (2007) [11]. Les sous-sections suivantes étayeront plus en détail ces procédés de détermination des quantiles.

A.2.1 Minimisation selon le modèle général pondéré

$$\text{Min} \sum_{i=1}^N \rho \tau \hat{\omega}_{(t,i)} |Y_i - X_i \beta| \text{ avec } \hat{\omega}_{(t \neq 0, i)} = \frac{t}{N \hat{r}(t, \mathbf{X})} \text{ et } \hat{\omega}_{(t=0, i)} = \frac{1-t}{N[1 - \hat{r}(t, \mathbf{X})]}. \quad (\text{A.5})$$

Dans l'équation (A.5), ρ est la fonction de pondération « check function » et $\hat{\omega}_{(t,i)}$ est le poids affecté aux individus selon leur participation ou pas à un traitement t . Ensuite, $\hat{r}(t, \mathbf{X})$ est le score généralisé estimé du traitement t et N la taille de l'échantillon.

Le modèle est une pondération de la somme des valeurs absolues des résidus (issus de la régression des Y des individus sur leurs caractéristiques observables X) du traitement t . Ainsi, la minimisation du produit de ce modèle avec la fonction de pondération ρ permet d'estimer les quantiles conditionnels.

A.2.2 Identification des quantiles

$$Q_{Y_1}(\tau|T=t) \text{ identifié par } \tau = E\left[\frac{(t1_I[Y \leq Q_{Y_1}(\tau)])}{r(t, \mathbf{X})}\right] \quad (\text{A.6})$$

$$Q_{Y_0}(\tau|T=t) \text{ identifié par } \tau = E\left[\frac{r(t, \mathbf{X})}{1-r(t, \mathbf{X})} \frac{t_0 1_I[Y \leq Q_{Y_0}(\tau)]}{p}\right] \quad (\text{A.7})$$

p étant la proportion des participants au traitement t et 1_I étant la fonction indicatrice.

$t1_I[Y \leq Q_{Y_1}(\tau)]$ est le nombre de participants au traitement t inclu dans le τ^{e} quantile.

$t_0 1_I[Y \leq Q_{Y_0}(\tau)]$ est le nombre de contrefactuels des participants du traitement t inclu dans le τ^{e} quantile.

τ est une proportion moyenne des participants au traitement t dans l'équation (A.6).

τ est une proportion moyenne des contrefactuels des participants au traitement t dans l'équation (A.7).

A.2.3 Les estimateurs quantiles semi-paramétriques

$$\hat{Q}_{Y_1}(\tau|T=t) = \hat{Q}_{TETESP}(\tau|D_i(t)=1, T=t) \quad (\text{A.8})$$

$$\hat{Q}_{Y_0}(\tau|T=t) = \hat{Q}_{TETESP}(\tau|D_i(t)=0, T=t). \quad (\text{A.9})$$

L'équation (A.8) présente l'estimateur du τ^e quantile des participants du traitement t alors que l'équation (A.9) présente l'estimateur du τ^e quantile des contrefactuels du même traitement. C'est la différence de ces deux estimateurs qui donne l'effet du τ^e quantile du traitement t .

Bibliographie

- [1] Battu Harminder, Belfield Clive, and Sloane Peter. Overeducation among graduates : a cohort view. *Education economics*, 7(1) :21–38, 1999.
- [2] Bellakha Rihab and Mahjoub Mohammed-Badrane. The impact of tunisian vocational training programs on employment and wage, 2013.
- [3] Berman Eli, Lang Kevin, and Siniver Erez. Language-skill complementarity : returns to immigrant language acquisition. *Labour Economics*, 10(3) :265–290, 2003.
- [4] Cameron Lisa, Dowling Malcolm, and Worswick Christopher. Education and labor market participation of women in asia : Evidence from five countries*. *Economic Development and Cultural Change*, 49(3) :459–477, 2001.
- [5] Cattaneo Matias. Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics*, 155(2) :138–154, 2010.
- [6] Cattaneo Matias, Drukker David, Holland Ashley, et al. Estimation of multivalued treatment effects under conditional independence. *Stata Journal*, 13(3) :407–450, 2013.
- [7] Chiswick Barry and Miller Paul. The endogeneity between language and earnings : International analyses. *Journal of labor economics*, pages 246–288, 1995.
- [8] Dehejia Rajeev and Wahba Sadek. Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, 84(1) :151–161, 2002.
- [9] Dustmann Christian and Fabbri Francesca. Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal*, 113(489) :695–717, 2003.
- [10] Ferrer Ana and Riddell Craig. Sheepskin effects and the returns to education. *Unpublished Paper. Retrieved April, 20 :2005*, 2001.
- [11] Firpo Sergio. Efficient semiparametric estimation of quantile treatment effects. *Econometrica*, 75(1) :259–276, 2007.
- [12] Flores-Lagunes Alfonso and Light Audrey. Identifying sheepskin effects in the returns to education. *University of Arizona, memo*, 2004.

- [13] Gertler Paul, Martinez Sebastian, Premand Patrick, Rawlings Laura, and Vermeersch Christel. *L'évaluation d'impact en pratique*, 2011.
- [14] Gonzalez Arturo. The acquisition and labor market value of four english skills : new evidence from nals. *Contemporary Economic Policy*, 18(3) :259–269, 2000.
- [15] Hanushek Eric, Woessmann Ludger, and Zhang Lei. General education, vocational education, and labor-market outcomes over the life-cycle. Technical report, National Bureau of Economic Research, 2011.
- [16] Jaeger David and Page Marianne. Degrees matter : New evidence on sheepskin effects in the returns to education. *The review of economics and statistics*, pages 733–740, 1996.
- [17] Kang Suk and Bishop John. Vocational and academic education in high school : complements or substitutes? *Economics of education review*, 8(2) :133–148, 1989.
- [18] Kilolo-Malambwe Jean-Marc. La surqualification au sein des grands groupes professionnels au québec : état des lieux en 2012. page 21, 2013.
- [19] Krueger Alan and Rouse Cecilia. New evidence on workplace education. Technical report, National Bureau of Economic Research, 1994.
- [20] Leslie Derek and Lindley Joanne. The impact of language ability on employment and earnings of britain's ethnic communities. *Economica*, 68(272) :587–606, 2001.
- [21] Neuman Shoshana and Ziderman Adrian. Vocational schooling, occupational matching, and labor market earnings in israel. *Journal of Human Resources*, pages 256–281, 1991.
- [22] Rosenbaum Paul and Rubin Donald. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 212–218, 1983.