Three Essays on the Return on Investment in Human Capital of Skilled Immigrants in Quebec and Internal Labor Migration in Developing Countries

Thèse

Marie Albertine Djuikom Tamtchouong

Doctorat en économique
Philosophiæ doctor (Ph. D.)

Québec, Canada

Three Essays on the Return on Investment in Human Capital of Skilled Immigrants in Quebec and Internal Labor Migration in Developing Countries

Thèse

Marie Albertine Djuikom

Sous la direction de:

Guy Lacroix, Directeur de recherche
Bernard Fortin, Co-directeur de recherche

Résumé

Cette thèse de doctorat s'intéresse à la migration interne et internationale. Dans un premier temps, je m'intéresse à l'intégration professionnelle des immigrants de la catégorie des travailleurs qualifiés au Québec. Le Québec comme la plupart des autres provinces du Canada, sélectionnent leurs immigrants sur la base de caractéristiques particulières telles que le niveau d'éducation, l'expérience professionnelle, les compétences en français et ou en anglais. Ces compétences devraient faciliter l'insertion professionnelle de ces immigrants et il est donc surprenant de voir que près de la moitié d'entre eux retournent aux études une fois arrivés au Québec afin d'obtenir un diplôme universitaire ou collégial. De ce fait, les deux premiers chapitres de cette thèse s'attèlent à comprendre pourquoi ces immigrants, malgré une telle dotation en capital humain à l'entrée du marché du travail Québécois, décident de retourner aux études et quels sont les effets de cet investissement en éducation tout d'abord sur les fréquences d'emplois et les durées en emploi et ensuite, sur le profil de revenus. Dans un deuxième temps, cette thèse s'intéresse à la participation à la migration interne en Ouganda et l'effet de cette participation sur la productivité agricole des ménages vivant en milieu rural.

Le premier chapitre s'intéresse à l'effet dynamique de la formation post-migratoire sur l'offre de travail des immigrants. A cet effet, je fais la distinction entre un emploi qualifié et un emploi non qualifié. Ici, un emploi qualifié est celui-là qui correspond au plus haut diplôme obtenu par l'immigrant à l'entrée. J'utilise un modèle de durée à plusieurs états et à plusieurs épisodes qui permet de tenir compte de l'hétérogénéité observable et inobservable entre les individus. Le principal résultat révèle que les immigrants originaires de pays riches n'ont pas besoin d'investir davantage dans l'éducation Québécoise. En revanche, les immigrants originaires de pays pauvres quant à eux, bien que hautement qualifiés, bénéficient largement d'une telle formation à long terme car cela facilite leur transition vers des emplois qualifiés et non qualifiés et hors du chômage. Mes résultats indiquent également que la sélection dans l'éducation doit être prise en compte afin d'éviter des problèmes de sélection significatifs.

À la différence du premier où on suppose que l'effet causal de la formation est le même pour chaque individu, le deuxième chapitre quant à lui s'intéresse à l'hétérogénéité de l'effet causal de la formation sur les revenus. Autrement dit, pour chaque individu il est possible d'estimer un effet moyen en comparant son revenu dans le cas où il a obtenu un diplôme au

Québec avec la situation où il n'aurait pas eu un diplôme au Québec, et vice-versa. Ceci est possible grâce à l'introduction de l'approche bayésienne dans l'analyse d'évaluation d'impact mettant en exergue l'estimation du contre-factuel de la variable d'intérêt. Les principaux résultats révèlent que les gains de l'éducation acquise au Québec par rapport à ceux de l'éducation acquise à l'étranger diffèrent d'un immigrant à l'autre. En outre, il y a un gain négatif à entreprendre des études au Québec pour tous les immigrants. Particulièrement, plus la probabilité d'entreprendre des études au Québec est élevée plus le retour sur investissement est faible. Il semblerait que les employeurs rémunèrent les immigrants non pas seulement par rapport à leur diplôme or sa provenance mais aussi par rapport à la qualité de leur précédent emploi. Ainsi, on s'attendrait à ce que les immigrants, toute suite après leur formation, acceptent un emploi relativement moins rénuméré que celui qu'il aurait eu étant donné son éducation. Par ailleurs, bien que l'approche bayésienne suggère que, comparativement aux immigrants qui ont obtenu un diplôme collégial au Québec, ceux qui obtiennent un diplôme universitaire sont les plus négativement affectés par un tel investissement en éducation, l'approche Fréquentiste suggère que ces derniers obtiennent le meilleur rendement des études acquises au Québec. Cela soulève à nouveau la question du biais de sélection qui peut subvenir lorsque l'hétérogénéité de l'effet n'est pas prise en compte.

Le troisième chapitre a pour objectif d'estimer la distribution de l'effet dynamique de la participation des ménages à la migration interne de la main d'œuvre sur la productivité agricole. Les résultats révèlent que même si en moyenne la migration interne affecte positivement la productivité agricole, il y a des ménages pour lesquels l'effet est négatif. De plus, les ménages pour qui l'effet est négatif sont pour la plupart de petits agriculteurs et sont par conséquent plus susceptibles d'être pauvres et donc plus susceptibles d'être sujet à la volatilité des prix au niveau local. Par ailleurs, l'effet moyen de la migration tend à augmenter avec la probabilité de participer à la migration interne signifiant que les individus décident de participer à la migration parce qu'il anticipent des gains futurs plus élevés. Parallèlement, j'examine dans quelle mesure les taux de migration antérieurs, largement utilisés dans la littérature en tant qu'instrument de la décision de participer à la migration, sont exogènes à la productivité agricole. Les résultats suggèrent que ces variables ne sont pas exogènes car elles sont intimement corrélées avec la productivité agricole.

Abstract

This doctoral thesis is interested in international and internal migration. First, it focuses on the professional integration of immigrants in the category of skilled workers in Quebec. Quebec is one of the ten provinces of Canada that, like most other provinces, implemented a program back in 1996 that explicitly selected highly qualified workers based on particular characteristics such as the level of education (Bachelors', Masters' or PhD's), work experience, French and/or English proficiency. Despite these skills that should facilitate their professional integration, 48% of immigrants return to school once they arrive in Quebec in order to obtain a university or college diploma. The first two chapters of this thesis investigates why these immigrants decide to go back to school with such an endowment of human capital and what the effects of this investment in education are on the job frequencies and job durations and, on the earnings profile. This thesis then focuses on the households participation in internal labor migration and the dynamic effect of this participation on the agricultural productivity of households living in rural area of Uganda.

The first chapter investigates the extent to which the return to foreign-acquired human capital is different from the education acquired in Quebec. Specifically, it seeks to estimate the benefits of post-migration education over foreign-education on the transitions between qualified and unqualified jobs and unemployment by means of a multiple-spells and multiple-states model. Here, a qualified job is one that corresponds to the highest degree obtained by the immigrant before they come in Quebec. The main results suggest that immigrants originating from well-off countries have no need to further invest in domestic education. Meanwhile, immigrants from poor countries, despite being highly qualified, benefit greatly from such training in the long run as it eases their transitions into qualified and unqualified jobs and out of unemployment. Our results also indicate that selection in education must be taken into account in order to avoid significant selection problems.

Unlike the first chapter in which only the average effect of schooling is estimated, the goal of the second chapter is to estimate the distribution of the causal effect of Quebec-acquired education on migrants' earnings. In other words, it is possible to estimate an average effect for each individual by comparing his income in the case he has obtained a Quebec diploma to the situation where he has not obtained a diploma from Quebec, and vice versa. This is possible

thanks to the introduction of the Bayesian approach in the treatment analysis allowing to account for the heterogeneity of the effect. The main results reveal that on average and for each immigrant, there is a negative gain to study in Quebec. However, the magnitude of the effect differs from one immigrant to another. Particularly, the gains tend to decrease with the likelihood of enrolling in school and with the level of *ability*. Thus, our results suggest that employers pay migrants not only based on their level of education or its origin but more importantly based on the quality of prior jobs held. Furthermore, one would expect immigrants to accept, right after their training, a relatively less paid job than the one he would have had given his education. While the Bayesian approach suggests that immigrants who have enrolled to obtain a university degree are the most negatively affected, the Frequentist approach suggests that those immigrants obtain the highest positive return from Quebec-acquired education. This raises again the issue of mis-evaluation when the *essential heterogeneity* is not taking into account.

The goal of the third chapter is to estimate the distribution of the dynamic effect of household participation in internal labor migration on agricultural productivity in Uganda. Since household can have both observed and unobserved factors that can affect both the decision to participate or not in migration and the return from it, this study account for the heterogeneity of the effect. Results reveal that although, on average, internal labor migration positively affects agricultural productivity, there are households for which the effect is negative. In addition, households for which the effect is negative are mostly small farmers, therefore more likely to be poor and more likely to be subject to local price volatility. It seems that return to migration helps poor household to meet other needs. Moreover, the average effect of migration tends to increase with the probability of participating in internal migration, meaning that households decide to participate in migration because they anticipate higher future returns. At the same time, we also examine the extent to which past migration rates, widely used in the literature as an instrument for the decision to participate in migration, are exogenous to agricultural productivity. Results show that these variables are not exogenous because they are highly correlated with agricultural productivity.

Table des matières

Liste des tableaux

Liste des figures

To the memory of my mother
who left us so early. We miss you.

Migration is an expression of the human aspiration for dignity, safety and a better future. It is part of the social fabric, part of our very make-up as a human family.

Ban ki-Moon

ACKNOWLEDGMENTS

The contribution of many people made this thesis possible. I owe special thanks to Mr. Guy Lacroix, who has accepted to supervise this work during this last five years. He is the best supervisor that I could have ever asked for ; he always had an answer for all questions and has always been able to cheer me up, whatever the circumstances. His willingness to give his time so generously has been very much appreciated. I would like to express my very great appreciation to Mr. Bernard Fortin for his valuable and constructive comments and suggestions, and for all the interesting discussions that we had, which helped me to improve this thesis. My grateful thanks are also extended to Mr. Sylvain Dessy, Mrs. Dominique Van De Walle and Mr. Carlos Gradín for their contribution.

I would also like to extend my thanks to all the teaching and administrative staff of the department of economics of université Laval for their advice and help in providing me with the resources that I need during my time there.

I would like to express my deep gratitude to UNU-WIDER team that provided me with an excellent working environment when I started working on the third chapter of my thesis through their Ph.D internship program. I also thank all participants of the Africa Meeting of the Econometric Society conference in Benin, the UNU-WIDER Development Conference on “migration and mobility” conference in Ghana, the Canadian Economics Association conference in Ottawa, the Canadian Economic Science Society conference in Quebec City, the "Association Francophone pour le Savoir" conference in Montreal and the 9th QICSS New Researchers Conference in Montreal.

Special thanks to my best friend Kodjo, for his useful comments, his unconditional support and his help in revising several versions of this thesis. To my friends, Kodjo, Sylvie and Adeline, I extend absolute and sincere thanks for all your support and the smile you put on my lips during difficult times. I wish to thank various other persons for their contribution to this thesis ; Alain and Leïla, for their help in revising this document ; my family specially, my father, *maman* Esther, *papa* Olivier, *tonton* Dubois, *maman* Julie, *maman* Carole, *maman* Oriane, *maman* Josephine, Hermann, Aubin, Michel, Vincent and Sonia.

This journey was particularly full of obstacles and there were times I might have given up but Lord gave me strength to keep going. I thank Lord for giving me health, ability and making a way to put me in the position I am in today.

Thanks to all of my colleague for all debates, recreation and exchanges that we had and that helped me to get out of daily work.

The achievement of this work could not have been possible without the financial support of the Faculty of Social Sciences and the department of economics of the Université Laval, Quebec Inter-University Center for Social Statictics (QICSS), Industrial Alliance Research Chair on the economics of demographic change and *"Centre de recherche sur les risques, les enjeux économiques, et les politiques publiques (CRREP)"*.

Thanks to Immigration & Citizenship Canada, the Ministry of Immigration and Cultural Communities and the Ministry of Employment and Social Solidarity of the Quebec Government for giving me access to the data used in the first two chapters of this thesis. Thanks also to the World Bank and the Uganda Bureau of Statistics for providing the Uganda National Panel data used in the third chapter of my thesis.

Introduction

It is now easier for individuals to move around the world due in part to globalization and the expansion of communication technologies. No country is spared by the phenomenon, as both developed countries and developing countries are affected. Migration takes place internally (within a country) or internationally (between countries) for many reasons, ranging from conflicts or natural disasters in the case of involuntary migration, to the search of better paying jobs or more generally for economic reasons in the case of voluntary migration. Driven by those reasons and the hope for a better life for themselves and for their relatives, people leave every day, temporarily or permanently, their countries, regions and families. There are many push and pull factors in the decision to migrate. In the case of international migration, the monetary cost is the main impediment to leave. Therefore, only individuals from wealthy families participate in the international migration.

From the host countries perspective, policies in favour of migration reflect their willingness to overcome the problem of an aging population and to cover their needs in labor for some specific sector of activities. In these countries, migration policies are well defined and enable them to select workers with particular educational and professional background that are supposed to ease migrants' professional integration. Canada is one of the few countries with migration policies specifically designed to attract qualified workers. The Center for Global Development, with its "Commitment to Development Index" that ranks rich countries' migration policies, places Canada among the four countries with the "*best migration policies*", the three others being New Zealand, Norway and Australia[1].

Once Landed in the host country, migrants encounter professional and socio-cultural assimilation problems. The required time to assimilate depends significantly on the migrants' status. For instance, empirical evidence shows that the likelihood of finding a job in the host country is far lower for refugees than for people who have migrated for economic reasons (Grant, 1999). Moreover, the assimilation process is not facilitated on the one hand by the quality of jobs available to migrants and the knowledge of local language and, on the other hand, by the fact that the returns to foreign-acquired education depend on the level of development of the country that issued the diploma (Coulombe *et al.* , 2014b,a; Chiswick &

1. https://www.cgdev.org/blog/which-countries-have-best-migration-policies

Miller, 2007). As a result, migrants encounter problems finding a job that matches their level of education. In order to get access to better jobs, some immigrants decide to invest in the host country education and obtain a college or university degree. So far, studies that attempt to estimate the gains of host country acquired-education over foreign-acquired education are inconclusive (Bratsberg & Ragan, 2002; Hou & Lu, 2017). Some find small gains and other insignificant gains of host country acquired-education in terms of earning profile.

In Canada, existing literature reveals that the return to foreign-acquired education in terms of earnings is highly correlated with the level of development of the migrant' origin country. So far, less is known about the causal effect of education acquired in Canada on the probability of finding a job. On the one hand, the first chapter of this thesis contributes to the literature by estimating the causal effect of post-migration schooling on labor market transitions. On the other hand, the second chapter contributes to the ongoing discussion about the gains of host-country acquired-education over the foreign-acquired one in terms of earnings profile. Unlike other studies, this study allows the causal effect of host-country acquired-education to be individual-specific and takes into account the fact that each migrant has observed and unobserved characteristics such as social network or *ability* that can affect both his decision to invest in education and his earning profile.

The first two chapters focus on the professional integration of skilled migrants living in the province of Quebec, one of the provinces of Canada. This province has its own migration policy with the goal to select people who are highly educated, speak French and (or) English and who have previously worked in their country of residence. The migrants selected through this process are skilled workers who enter the province with the permanent resident status. Additionally, the government of Canada gives the opportunity to international students to become permanent residents after completing their degree. Despite these policies, skilled workers face some barriers to integrate Quebec's labor market.

Between January 2002 and December 2009, approximately 90,000 immigrants were admitted in through the Quebec Skilled Workers program. Our dataset is a random sample drawn from the so-called "landing file", a large administrative dataset managed by Immigration & Citizenship Canada. [2] The file contains all relevant pre-migration information such as highest degree schooling, results from language tests in French or/and English, the exact arrival date in Canada, *etc.* The administrative information was complemented by a retrospective survey since arrival until May 2011 through which detailed information on all relevant periods of employment, unemployment, schooling, earnings, *etc.* are reported. Our sample contains 3,009 skilled immigrant workers with 42% coming from Europe, 34% form Africa, 7% from Asia and 16% from America. Those migrants are on average more educated than the natives. Indeed, 19% of migrants have a master or doctorate degree and 49% have a Bachelor degree. It is thus surprising that with such endowment in human capital, 48% of those migrants go back

2. This period corresponds to a significant increase in the proportion of immigrants in the population(Citizenship & Immigration Canada, 2014).

to school in order to get Quebec based diplomas.

This thesis investigates why migrants decide to invest in Quebec education and seeks to estimate the returns to this investment in terms of labor mobility and earning profile. In the first chapter, the effect of having a Quebec diploma on the labor market transitions is studied by means of a multiple states and multiple spells model. In this study migrants transitions between four different states -qualified job, unqualified job, unemployment and training- are estimated. A qualified job is a job that matches the highest diploma. Data reveal that when immigrant starts with an unqualified job, it is difficult to find a qualified job after. Since migrants can have both observed and unobserved characteristics that make them enrol or not in school, the identification strategy takes into account the potential correlation between this investment in education and the probability of having a job. The results reveal that immigrants originating from well-off countries have no need to further invest in domestic education. On the other hand, immigrants from poor countries, despite being highly qualified, benefit greatly from such training in the long run as it eases their transitions into qualified and unqualified jobs and out of unemployment. Results also indicate that selection into domestic education needs to be accounted for to avoid significant selection problems.

In the second chapter, the gains of Quebec-acquired education over the foreign-acquired one in terms of weekly earnings are estimated. Existing empirical evidence reveals that host-country education provides, on average, a small benefit over the foreign-acquired one, but very little return to foreign-acquired experience (Skuterud & Su (2012)). This study thus contributes to the ongoing discussion about the gains of host-country education over the foreign-acquired one. To achieve this goal, we estimate the distribution of the gains using a Bayesian treatment analysis that allows to account for : (1) the heterogeneity of the gains between individuals since people can be affected differently even after controlling for their personal characteristics and, (2) the unobserved heterogeneity as *ability* or social network that affects both the decision to invest in education and the earnings profile. In fact, migrants are more likely to make decisions based on information they get from their social network. Moreover, with this approach, it is possible to construct for each migrant the counterfactual outcome in a potential analysis framework. Furthermore, Bayesian treatment analysis do not require to find instruments for endogenous variable. On the other hand, In other to evaluate how the causal effect is biased when the unobserved heterogeneity is not taking into account, we also estimate the average causal effect of host-country acquired education using the sequential causal model based on the propensity score.

Our results reveal that the gains of Quebec-acquired education over foreign-acquired one are different from one migrant to another. Therefore, computing a single average effect, as is customarily done, does not allow to capture this heterogeneity. In fact, even if on average and for almost all migrants, there is negative gain of Quebec-acquired education over the foreign-acquired one, the magnitude of the effect is different from one migrant to another. Particularly, the gains tend to decrease with the likelihood to enroll in school and with the level of *ability*.

Thus, our results suggest that employers pay migrants not only based on their level of education or their origin but more importantly based on the quality of prior jobs held. Indeed, the fact that migrants previously held a qualified job affects positively the earnings and the effect is more important for migrants who have enrolled in university degree in Quebec. It seems that the previous job is a better indicator of migrant' productivity than the level of education. While the Bayesian approach suggests that immigrants who have enrolled in university degree in Quebec are the most negatively affected, the sequential causal model suggests that those immigrants obtain the highest positive return from QC-acquired education. This raises again the issue of mis-specification when the *essential heterogeneity* is not taken into account

In the third chapter, we are interested in the dynamic effects of internal labor migration on agricultural productivity of households living in the rural areas of Uganda. Uganda is one of the poorest country in the world where only 54% of its population has completed the primary school. With such low educational attainment, people participate mainly in internal migration (95%) and most of time, the migration is temporary. Households choose to participate in this kind of migration because the monetary cost is relatively low, although the lost of available labor can be a big concern in the context where markets do not work well. Moreover, even if there is empirical evidence showing that remittances obtained from internal labor migration is relatively low compared to the one from international migration, internal migration may be a way for many poor families to get ahead. In fact, the returns to migration can help households left behind to invest in education, start a business, buy new land or buildings. Furthermore, previous studies reveal that migration decreases the inequality at the level of the community.

This chapter focuses on agricultural productivity because the agricultural sector contributes to 25% of the GDP growth of the country. Further, no study has previously investigated the causal effect of internal labor migration on agricultural productivity in Uganda. The *New Economics labor Migration (NELM)* theory developed by Stark & Bloom (1985) states that labor migration should have positive effect on the agricultural sector since money sent back to households left behind can help them invest in this sector. In fact, according to this theory, labor migration decision is a mutual beneficial agreement between the migrant and other members of his(her) household. Empirical evidence on other African countries such as Ghana reveals that internal labor migration tends to decrease agricultural productivity. However, those studies do not account for the fact that migration can affect households differently due to observed characteristics and unobserved *essential heterogeneity* that might affect both the decision to participate in migration and the investment in agricultural sector.

Using Bayesian treatment analysis that explicitly allows to estimate the counterfactual outcomes, the distribution of the causal effects is estimated. The results show that even if on average internal labor migration positively affects agricultural productivity, there are some households for which the effect is negative. Those households are mostly small farmers, and are therefore more likely to be poor and thus more sensitive to local price volatility. Moreover, the average effect of labor migration tends to increase with the likelihood of participating in the internal labor migration. In parallel, this chapter also examines the extent to which previous migration rates, widely used in the literature as instrument for the migration decision, are exogenous to agricultural productivity. It turns out that previous households' decisions to participate in migration are highly correlated with their current agricultural productivity.

- Chapitre 1 -

Dynamic Causal Effects of Post-Migration Schooling on Labour Market Transitions

Abstract [a]

Immigrants often experience difficulties integrating the local labor market. In Canada, the government of Quebec implemented a program back in 1996 that explicitly selected highly qualified workers (Bachelors', Masters' or PhD's). This paper investigates the extent to which the return to foreign-acquired human capital is different from the education acquired in Quebec. Specifically, we seek to estimate the benefits of post-migration education over foreign-education on the transitions between qualified and unqualified jobs and unemployment by means of a multiple-spells and multiple-states model. Our results indicate that immigrants originating from well-off countries have no need to invest in domestic education. On the other hand, immigrants from poorer countries, despite being highly qualified, benefit greatly from such training in the long run as it eases their transitions into qualified and unqualified jobs and out of unemployment. Our results also indicate that selection into domestic education needs to be accounted for to avoid significant selection problems.

Keyword : Post-migration schooling, foreign education, labour market histories, multiple-spells multiple-states models.

JEL-codes : C31, C41, J15, J24, J64, J61

a. This chapter is co-written with Guy Lacroix

1.1. Introduction

Migrants often experience difficulties integrating the labor market in most host countries. In particular, they encounter difficulties finding a qualified job, *i.e.* one that corresponds to their level of schooling prior to migrating. In Canada, previous studies have found that the professional integration of migrants has worsened over time and that the wage gap with the natives has grown over the last decade. [1] These results are surprising given that recent cohorts are highly educated, have French or English as their native language, and have had valuable work experience prior to migrating.

In Canada, migrants may be admitted under four separate categories skills-assessed principal applicants entering under the Federal Skilled Worker Program (independent economic immigrants), other economic immigrants, family class immigrants and refugees. The Skilled Worker Program is designed to identify applicants who are likely to become economically established upon migrating to Canada by weighing more heavily foreign acquired skills, work experience and knowledge of French or English, among other characteristics. The Province of Quebec has a separate set of criteria to select immigrants entering under the Skilled Worker Program. Quebec Skilled Workers applicants are thus not assessed based upon the selection factors of the Federal Skilled Worker Class. Not surprisingly, skills-assessed independent economic immigrants, both males and females, consistently have the highest median annual earnings among the four admission categories in the long run (Abbott & Beach, 2011; Sweetman & Warman, 2013), although other immigration classes fare just as well if not better in terms of employment. Yet, immigrants admitted under the Quebec Skilled Worker Program (QSWP) suffer from poorer recognition of foreign acquired skills than elsewhere in Canada. Many end-up working in occupations for which they are largely over-qualified (Lacroix, 2013). [2]

Analysing the links between foreign-acquired skills and domestic labor market integration raises difficult empirical issues. First, most surveys do not distinguish between country of origin and country of study. Second, if foreign-acquired skills are indeed ill-recognized on the domestic market, then many may be induced to enhance their skills domestically, thus raising issues of endogeneity or self-selection into domestic education. In this study we avoid these issues in two ways. First, our unique dataset focuses on skilled immigrant workers and contains detailed information on their schooling and past work experience. We can thus separate out the "country of origin" from the "country of study" effects. Second, the data are rich enough to allow the estimation of the causal effect of post-migration education on the transitions between qualified and unqualified jobs and unemployment. Specifically, we estimate a multiple-states and multiple-spells model which accounts for the endogeneity of the decision to invest in domestic education. Our identification strategy exploits the fact that the immigrants in our

1. Borjas (2000), Boudarbat & Boulet (2007), Grant (1999),Hansen (2000), Coulombe *et al.* (2014b,a).

2. According to Coulombe *et al.* (2014a), the return to foreign education is intimately related to the level of development of one's birth country or country of study. Countries of origin of immigrants to Quebec and Canada differ substantially.

sample were selected under two sets of criteria. The majority of the sample was selected under a set that was in place between 1996 and 2006. A smaller subsample was selected under the new criteria that were implemented in 2006. These new criteria put more emphasis on language proficiency and work readiness, among other things. Our goal is to estimate the extent to which the new criteria expedited the transitions into employment, and the jobs constituted better matches to the workers' skills. If the latter holds, then we should observe fewer immigrants participating in domestic skill enhancing programs as get back to school.

Our approach extends the one proposed by Chesher & Lancaster (1983) and is loosely related to those considered in Uhlendorf & Zimmermann (2006) and Hansen (2000). In both cases, they investigate the determinants of the differentials in the duration of unemployment and employment spells between immigrants and natives in Germany and Sweden. Unlike these papers, our work focuses on the immigrants alone, not on a comparison between natives and immigrants. While it is widely acknowledged that the returns to education and work experience vary considerably between immigrants and natives, it has also been found that they vary greatly amongst immigrants (Coulombe *et al.* , 2014a). A novel feature of our approach is to explicitly account for schooling quality of foreign-acquired skills. Following Coulombe *et al.* (2014b) and Chiswick & Miller (2007), we do this by proxying the quality by a function of the per capita GDP in the country of study.

The paper is organized as follows. Section 1.2 provides detailed informations on the data. Section 1.3.1 present a multiple-states and multiple-spells model which account for the endogeneity of investment in education. The effect of post-migration schooling is measured through a dummy variable which is set to one if the individual has attended school in Quebec and zero otherwise. In addition, the schooling status is treated as a separate state in one version of the model. Our focus is on the education that leads to graduation from post-secondary education or more. We also estimate the model with schooling treated as exogenous and compare the parameter estimates with its endogenous variant. In Section 1.3.3, we discuss the quality of foreign schooling and how we proxy it. In addition, we present an empirical strategy to estimate the instantaneous probability of holding a qualified or unqualified job over time for each individual. This allows us to compare the employment trajectories for those who have attended school in Quebec relative to those who have not. This outcome is interesting from a policy perspective since it allows comparisons between individuals who differ in terms of characteristics such as country of study. Finally, Section 1.4 presents and discuss main results. We conclude the paper in Section 1.5.

1.2. Data and Preliminary Analysis

Between January 2002 and December 2009, approximately 90,000 immigrants were admitted in through the Quebec Skilled Workers program. Our dataset is a random sample drawn from the so-called "landing file", a large administrative dataset managed by Immigration & Citizenship Canada.[3] The file contains all relevant pre-migration information such as highest degree schooling, results from language tests in French or/and English, the exact date of the entry in Canada, *etc.* The administrative information was complemented by a retrospective survey since arrival until May 2011 through which detailed information on all relevant periods of employment, unemployment, schooling, earnings, *etc.* are reported. Our sample contains 3,009 skilled immigrant workers. The retrospective survey was conducted conjointly by the Ministry of Immigration and Cultural Communities and the Ministry of Employment and Social Solidarity of the Quebec Government. According to Benzakour *et al.* (2013), the sample is representative of the population. Immigrants from Asia, and particularly those from the Middle-East, are however slightly under-represented in the sample. All other demographic characteristics (gender, marital status, educational background, *etc.*) are representative of the population.[4]

Most immigrants in our sample were selected according to the grid that was implemented in 1996 (2,607/3,009) and remained constants until 2006. The remaining observations (402/3,009) were selected on the basis of the October 2006 selection grid. The new grid weighs more heavily language proficiency (French and English), work experience and educational background. The new grid also introduced a Pass/Fail financial self-sufficiency contract. The "Contract respecting financial self-sufficiency" is a legal agreement by which the applicant undertakes to provide for his basic needs and, where applicable, those of his spouse and dependent children, for at least three months. The basic needs covered by this contract include food, housing, clothing and all other personal necessities. Failure to sign the contract leads to the automatic rejection of the application. This requirement implies that the applicants are not entitled to any public support (*e.g.*, welfare) for at least three months upon landing.

Table 1.1 provides interesting insights into the characteristics of the sample. The table is divided by household type and selection grid. Thus, the sample consists of 1,552 couples and 1,457 singles. The first panel of the table shows that the applicants from the 2006 grid have a significantly higher level of education than those selected with the former grid. This is true for both couples and single people. Thus, the proportion of individuals with a master's or doctoral degree increased from about 17% to over 30%.The proportion of individuals with secondary or lower education is the same for the two selection grids,

The second section of the table ("Grid Points") reports the scores obtained in each of the grids. Only the criteria common to the two grids are presented. The human capital criteria are relatively similar across the two grids, with the exception of English proficiency and age. In the latter case, the higher 2006 scores are mainly due to the fact that the average age of the applicants decreased in 2006, from 34 years to only 31.6. The selection of younger candidates has as a corollary a decrease in the score associated with experience. Likewise, the increase in points associated with English proficiency is partly due to an increased valuation in the 2006

3. This period corresponds to a significant increase in the proportion of immigrants in the population(Citizenship & Immigration Canada, 2014).

4. See Benzakour *et al.* (2013) for a detailed discussion.

Table 1.1 – *Sample Characteristics, by household type and Selection Grid*

	Couples		Singles	
Selection Grid	1996	2006	1996	2006
Education				
Post-Secondart (%)	24.446	12.183	33.386	14.285
Bachelor (%)	53.397	48.223	42.113	45.812
Master-PhD (%)	15.879	30.064	18.735	30.049
	Grid Points			
Human Capital				
Experience	8.417	7.061	6.199	4.941
French proficiency	12.774	12.624	14.421	14.517
English Proficiency	3.725	4.305	3.884	4.527
Age	7.689	16.173	9.319	17.626
Other Criteria				
Prev. stay in Quebec	0.640	1.741	1.797	2.857
Guaranteed Employment (%)	1.182	2.538	0.960	0.487
Spouse				
Education	2.696	2.147		
Experience	1.436	0.741		
Age	1.546	2.386		
Children	1.778	2.914		
Number of observations	1 355	197	1 252	205

grid (0-6 points) relative to the 1996 grid (0-3 points).

The "Other Criteria' section focuses on dimensions other than individual characteristics. For instance, the points allocated for previous stays in Québec were increased in the 2006 schedule. Points allocated for guaranteed employment were also increased in the 2006 grid. However, fewer than 1% of applicants had secured job prior to their arrival. [5]

The last section of the table relates to the spouses of the principal applicants. The points awarded to the spouse's schooling and experience were decreased in the 2006 grid. On the other hand, (younger) age and presence of children were awarded greater scores. This explains the main differences between the mean values of the two grids.

From Table 1.1 it is readily apparent that individuals selected from the 2006 grid have more schooling and are younger. To the extent that these characteristics are more or less favourable to employment, one should expect to observe differences in the time required to find a job. Table 1.2 reports the elapse time between landing and finding a first job (column (1)) or a first which corresponds to qualification. Individuals who did not find employment at the time of the survey are excluded from the calculations in column (1). Those who had not yet found a suitable job are excluded from the calculations in column (2). [6] The number of individuals Found in square brackets in each cell and that number is converted to a percentage

5. Only 34 individuals had guaranteed employment on arrival. Of these, 28 had been selected with the 1996 grid and 6 had been selected with the 2006 grid.

6. In other words, censored episodes are not included in the calculations.

and is shown in parentheses.

The time required to find a job is significantly lower for applicants selected using the 2006 grid. This is true for both single and couples, and regardless of whether the jobs corresponds to his qualification or not. In most cases, the time required has decreased between 50 % and 66 %.[7] It is also found that singles are proportionally more likely to find a job (qualified or not) than couples. However, the proportion of employed individuals decreased slightly with the adoption of the 2006 grid. This is hardly surprising given that individuals selected using the 2006 grid arrived much later and have had consequently much less time to find a job. Like unattached individuals, proportionally fewer couples are employed in the 2006 grid sample. The gap between the employment rate of unattached individuals and couples remains at about 5 percentage points in both samples. With respect to the first skilled job, the differences observed between household types and samples are relatively similar to those observed for the first job.

TABLE 1.2 – *Average Duration Between Landing and First Job, First Qualified Job, by Household Type and Selection Grid*

	First Job		First Qualified Job	
Household type	Grid		Grid	
	1996	2006	1996	2006
Singles	236.29	121.89	518.23	239.52
	[1 124]	[180]	[611]	[96]
	(89.8%)	(87.7%)	(48.8%)	(46.8%)
Couples	331.65	118.32	622.39	182.78
	[1 164]	[161]	[569]	[78]
	(85.9%)	(81.7%)	(42.0%)	(39.6%)
Total	284.81	120.21	568.46	214.09
	[2 288]	[341]	[1 180]	[174]
	(87.8%)	(84.8%)	(45.3%)	(43.3%)
Number of observations	1 355	197	1 252	205

— [#] of individuals who found a job between brackets.
— (%) of individuals who found a job.
— Censored observations not included.

The majority of the applicants succeed in finding a job. The time required to do so varies greatly between the different groups. Of particular concern is that most of them find it difficult to find a job whose requirements correspond to their academic skills. In fact, the table shows that 48.8% of single persons admitted under the 1996 grid have, after a long period of time, been able to find a job that corresponds to their skills. In addition, only 39.6% of couples admitted under the 2006 grid were able to do the same. It may be instructive to analyse the number of jobs that are filled before finding a first qualified one. Table 1.3 reports the sample frequencies of prior "unqualified jobs" by household type and selection grid. The first line of the table, "0", refers to the number of individuals whose first job was qualified. Thus, nearly 58% of the 1,233 individuals admitted under the 1996 grid were originally employed in a qualified

7. The average durations do not take censored observations into account. However, the calculation of the expected durations based on the survival curves, which takes censored episodes into account, nevertheless show that the expected duration for the sample of the single persons in the 1996 grid is 414.5 days whereas that of the sample of the 2006 grid is only 288.9 days. For couples, the results are 650.9 and 337.7 days, respectively.

job. As many as 75% of those admitted under the 2006 grid did the same. Better educated and younger workers under the 2006 grid, favourable economic conditions, *etc.* may all be partly responsible for the differences between the two samples. Yet, over half of immigrants admitted under the Skilled Workers program never managed to find a job that corresponded to their qualifications. The large majority (90%) of our sample was employed at one time between

Table 1.3 – *Number of jobs prior to finding a qualified one*

Number of	Grid		Household	
Unqualified Jobs	1996	2006	Singles	Couples
0	713	132	439	406
	(57.83)	(74.58)	(60.05)	(59.10)
1	266	27	145	148
	(21.57)	(16.00)	(19.84)	(21.80)
2	147	10	89	68
	(11.92)	(5.65)	(12.18)	(10.01)
3	71	8	38	41
	(5.76)	(4.52)	(5.20)	(6.04)
4	36	0	20	16
	(2.92)	(0.00)	(2.74)	(2.36)
Total	1 233	177	731	679
No Qualified Job	1 374	225	726	873
	(52.70)	(55.97)	(49.83)	(56.25)

% between parentheses.

2002–2011. As stressed above, half of them never managed to find a qualified job, others did so after transiting through a number of under-qualified jobs, and still others have gone through successive spells of employment and unemployment without ever finding a suitable job. The information in our dataset entitles us to identify four different statuses on the labour market : (1) Unqualified job ; (2) Qualified job ; (3) Unemployment ; (4) Training. [8] Table 1.4 reports the total number of transitions between the four states over the period ranging from 2002 to 2011. The entire work history of 2,946 immigrants could be gathered from the data. According to the table, the 2,946 individuals in our sample experienced 11,788 spells over eight years, or approximately 4.0 per individual. Give that the yearly cohorts in our sample are about 375, it is likely that those who landed early witnessed yet many more transitions. Using the start and end dates of every employment spells, we can calculate the daily employment rates for the immigrants of each selection grid and distinguish between qualified and unqualified jobs. Figure 1.1 focuses on the first three years upon landing. [9]

What the figure shows is that the employment rate of the latter is clearly higher than that of the applicants of the 1996 grid : their participation rate is higher both in skilled and unskilled jobs. It can also be seen that as of the second year upon landing, their participation rates in unskilled jobs falls rapidly and increases proportionately in skilled jobs. Nothing of the sort is observed for those who were selected using the 1996 grid. As stressed earlier, the 2006

8. In order to keep the model tractable, spells of concurrent employment and training are considered as training spells. Treating them as a separate states would considerably increase the number of transitions and consequently the number of parameters to estimate.

9. Those who were selected under the 2006 grid have at most a three-year history on the labour market.

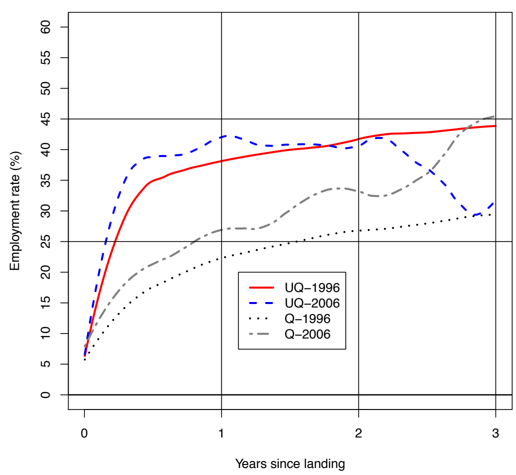
TABLE 1.4 – *Transitions between labour market statuses*

Origin \ Destination	Unqualified job	Unemployment	Qualified job	Training
Unqualified Job	**955** (32%)	1,702 (58%)	82 (3%)	204 (7%)
Unemployment	2,118 (41%)	**921** (17%)	1,253 (24%)	927 (18%)
Qualified Job	52 (2%)	833 (41%)	**1,070** (52%)	93 (5%)
Training	700 (45%)	383 (24%)	495 (31%)	**0** **(0%)**

Number of Individuals :2,946 ; Number of spells : 11,788

group was both significantly more educated and younger than those from the 1996 group. [10] It is not clear to what extent the observed differences in the employment dynamics of the two groups result from differences in human capital. It could be argued that younger and better skilled immigrants may find it yet even more difficult to find an adequate job. It may also be that some choose to invest in domestic education to facilitate their transition into employment. In order to net out the impact of Quebec-education from the foreign-human capital component, we must turn to a formal econometric model that accounts for potential selectivity in education.

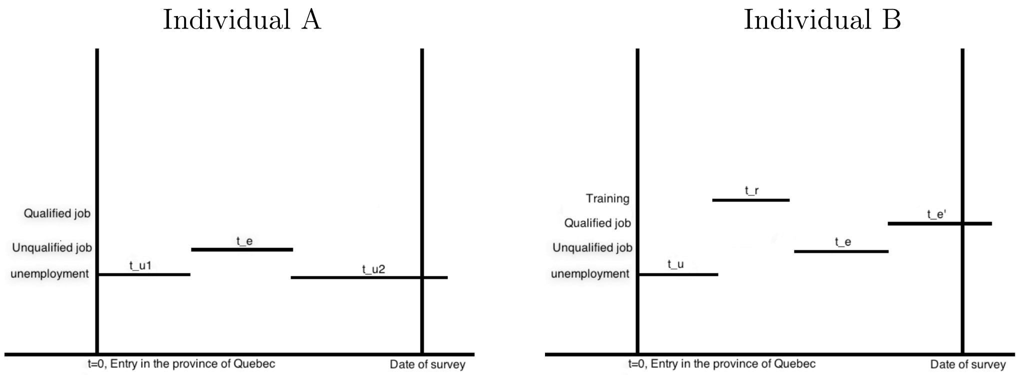
FIGURE 1.1 – *Daily Participation rates, by Selection Grid, First Three Years*



Note : UQ stands for Unqualified Job and Q stands for Qualified job.

10. It should also be stressed that the 2006 group included proportionately fewer immigrants from the Maghreb, a region whose nationals experience more difficulties finding a job (Oreopoulos, 2011; Brière *et al.* , 2017).

Figure 1.2 – *Employment History of Two Hypothetical Individuals*



Note : t_u, t_e, t'_e and t_r are the periods of time spend in unemployment, unqualified job, qualified job and training respectively.

1.3. Empirical Strategy

The typical employment histories of two immigrants are depicted in Figure 1.2. Upon landing, they likely experience both a spell of unemployment. After a while, one may choose to accept an unqualified job (A) whereas the other may elect to invest in domestic education (B). [11] Following his employment in an unqualified job, individual A may experience a new unemployment spell that lasts past the date of the survey (right censoring). Individual B, on the other hand, may become unemployed following his graduation, after which he may find a qualified job that lasts past the survey date.

The previous section has stressed that the characteristics and the employment histories vary significantly across selection grids. Table 1.5 shows, on the other hand, that the individual characteristics of the Quebec-educated migrants and the foreign-educated migrants are relatively similar. Indeed, they more or less have the same age and the same level of proficiency in French and English. The proportion of females, of married couples and foreign work experience are almost identical across groups. They are also more likely to hold a bachelor's degree and less likely to hold a masters' or a PhD. Previous research has shown that those who hold a bachelor's degree don't perform as well on the labour market (Renaud, 2005; Lacroix, 2013). This is perhaps why they are more likely to back to school. Yet there are a few noteworthy differences. The first concerns the indicator of development. Recall that this indicator is given by the logarithm of the ratio between the home country GDP to the Canadian GDP (PPP). Thus Quebec-educated migrants originate on average from poorer countries than foreign-educated migrants. In addition, fewer Quebec-educated have had previous stays in Quebec prior to being granted the permanent resident status.

11. In the data, we do observed multiple job holding. We have chosen to give priority to the qualified job whenever a qualified and an unqualified jobs at held simultaneously. In such cases, an unqualified job starts once the qualified job has ended and ends whenever a (simultaneous) qualified job begins. When two jobs of the same type are held concurrently, we concatenate the durations into a single one.

The bottom panel of the table reports the duration of the spells on the labour market. Note first that the Quebec-educated experience an increase in the duration of each state upon completing their education. The post-schooling average durations are almost identical to those of the foreign-educated migrants. Prior to their schooling, their employment duration is roughly half that of the foreign-educated, save for the unemployment spells which have more or less the same average duration.

Table 1.5 – *Summary statistics of immigrants depending on whether or not they undergone training*

	Trainees(1426)			Non-Trainees(1533)		
Variable	**Mean**	**Std. Dev.**	**# Spells**	**Mean**	**Std. Dev.**	**# Spells**
Age admission	33.61	5.52	-	33.88	6.57	-
Indicator of Devel.($ln\left(\frac{GDP^o}{GDP^c}\right)$)	-1.80	1.23	-	-1.21	1.22	-
French	13.78	3.14	-	13.57	3.75	-
English	3.38	1.85	-	3.84	1.84	-
Female (%)	33.33	-	-	34.0	-	-
Married (%)	50.42	-	-	52.44	-	-
Previous stay in QC (%)	25.63	-	-	44.01	-	-
Work experience (%)	82.7	-	-	84.65	-	
Education						
Master/PhD (%)	15.06	-	-	22.79	-	-
Bachelor (%)	52.87	-	-	44.66	-	
Post sec/less (%)	32.07	-	-	32.54	-	-
Duration (in months)						
Before training						
Unqualified Job	9.18	11.9	617	19.31	21.60	1384
Unemployment	6.88	9.85	1653	8.98	16.13	2583
Qualified Job	13.90	15.48	228	29.33	26.74	1123
After training						
Unqualified Job	17.49	19.12	942			
Unemployment	12.08	16.90	983			
Qualified Job	21.82	20.13	698			

Note : GDP^o is the GDP of country who issue migrant's highest diploma and GDP^c is the GDP of Canada.

1.3.1 Econometric Model : Education as an Exogeneous Variable

In order to model labour market history of our sample, we estimate a multiple spells-multiple states model with unobserved heterogeneity. Individuals are modelled as entering four distinct states : qualified job, unqualified job, unemployment and training. [12]. A qualified job is one which corresponds to an individual's educational degree at the time he received his

12. We do not model the inactivity state because only six percent of individuals in our sample reported not seeking work for reasons of pregnancy, family problems, *etc.*

permanent resident status.[13]

Let $k \in \{u, e, e'\}$ index the states, where u is unemployment, e stands for employment in an unqualified job, and e' for employment in a qualified job. Denote the set of transitions by K, where $K = \{ue, ue', eu, e'u\}$; the first letter denotes the origin state and the second stands for the destination state. Let T_{kj} be a continuous random variable representing the duration of a spell in state k that ends in state j. We thus have four random duration variables for each possible transitions, T_{ue},$T_{ue'}$, T_{eu} and $T_{e'u}$.

The density function of T_{jk} for a completed spell is defined as follows

$$f_{jk}(t) = \lambda_{jk}(t) \times S_{jk}(t) \times S_{jh}(t), j \neq k \quad and \quad j \neq h \tag{1.1}$$

and, $j, k, h = u, e, e'$, where λ_{jk} and S_{jk} are respectively the hazard function and survival function of the exit from state j to state k. We can write $f_{jk}(\cdot)$ conditional on individual-specific characteristics at time t, that is,

$$\begin{aligned} f_{jk}(t|X_1, X_2(t)) = \lambda_{jk}(t|X_1, X_2(t)) \times \exp(-\int_0^t \lambda_{jk}(\tau|X_1, X_2(t))d\tau) \\ \times \exp(-\int_0^t \lambda_{jh}(\tau|X_1, X_2(t))d\tau), \end{aligned} \tag{1.2}$$

$j \neq k \quad and \quad j \neq h$. Here, $\lambda_{jk}(t|X_1, X_2(t))$ is the probability of moving from state j to state k after t, conditional on not having left for either states k and h prior time t and conditional on the fixed (X_1) and time varying ($X_2(t)$) individual characteristics.

The above specification must be modified to account for right censoring. In our data, every last spell is right censored at the time of the survey. The density function thus becomes :

$$f_{jk}(t|X_1, X_2(t)) = \exp(-\int_0^t \lambda_{jk}(\tau|X_1, X_2(t))d\tau) \times \exp(-\int_0^t \lambda_{jh}(\tau|X_1, X_2(t))d\tau).$$

The likelihood function of a complete work history composed of K_i spells is given by,

$$\begin{aligned} L_i(t_s; s = 1, \ldots, K_i|X_1, X_2(t_s)) &= \prod_{s=1}^{K_i} f_{j_s k_s}(t_s|X_1, X_2(t) & (1.3) \\ &= \prod_{s=1}^{K_i} [\lambda_{j_s k_s}(t_s|X_1, X_2(t_s))]^{1-c_{i_s}} & (1.4) \\ &\times \exp(-\int_0^{t_s} \lambda_{j_s k_s}(\tau|X_1, X_2(t_s))d\tau) \\ &\times \exp(-\int_0^{t_s} \lambda_{j_s h_s}(\tau|X_1, X_2(t_s))d\tau), \end{aligned}$$

where, $c_{i_s} = \begin{cases} 1 & \text{if spell } s \text{ is right censored} \\ 0 & \text{Otherwise.} \end{cases}$

13. The required level of education for each job is based on Statistics Canada's National Occupational Classification (NOC).

To illustrate, we can write the likelihood function of individual A in Figure 1.2. He experiences three spells, starting with unemployment, and transiting successively into an unqualified job and then into another unemployment spell. The joint probability density for this hypothetical individual is given by,

$$
\begin{aligned}
L(t|X(t)) = {} & \lambda_{ue}(t_u|X_1, X_2(t_u)) \times \exp\left(-\int_0^{t_u} \lambda_{ue}(\tau|X_1, X_2(t_u))d\tau\right) \times \\
& \exp\left(-\int_0^{t_u} \lambda_{ue'}(\tau|X_1, X_2(t_u))d\tau\right) \times \lambda_{eu}(t_e|X_1, X_2(t_e))\times \\
& \exp\left(-\int_0^{t_e} \lambda_{eu}(\tau|X_1, X_2(t_e))d\tau\right) \times \exp\left(-\int_0^{t_u} \lambda_{ue}(\tau|X_1, X_2(t_u))d\tau\right) \times \\
& \exp\left(-\int_0^{t_u} \lambda_{ue'}(\tau|X_1, X_2(t_u))d\tau\right).
\end{aligned}
$$

Each line represents the density of the duration in each state ; as the last state is right-censored, the duration density is equal to the probability that this hypothetical individual remains in state u.

The likelihood function for the entire sample is as follows :

$$L(t_n; n = 1, \ldots, K|X_1, X_2(t_n)) = \prod_{i=1}^{N} L_i(t_s; s = 1, \ldots, K_i|X_1, X_2(t_s)), \tag{1.5}$$

with $K = \sum_{i=1}^{N} K_i$ and where N corresponds to the number of individuals and K_i, the total number of spells for individual i.

1.3.2 Education as an Endogenous Choice

To the extent enrolling in some form of training is endogenous to the labour market status, it is necessary to modify the econometric specification accordingly. The endogeneity may arise from unobserved variables which are linked to the duration of unemployment and employment spells as well as the probability of being in a qualified or unqualified job.

In this framework, we must distinguish between four distinct states. The set of potential transitions includes the following nine possibilities, $K = \{eu, er, ue, ue', ur, e'u, ru, re, re'\}$, where r stands for training. The conditional density of moving from state j to state k at t is given by,

$$f_{jk}(t|X_1, X_2(t)) = \lambda_{jk}(t|X_1, X_2(t)) \times \prod_{k \neq j} \exp\left(-\int_0^{t} \lambda_{jk}(\tau|X_1, X_2(t))d\tau\right), \tag{1.6}$$

with $j, k = u, e, e', r$. The interpretation of the hazard function is the same as in the previous section. The likelihood for each individual is obtained by replacing the expression of density in equation (1.6) into equation (1.3).

1.3.3 Functional Form Specification

In the literature, several probability density functions are customarily used to write the hazard function. Based on non-parametric analyses, both the Weibull and the log-logistic functional forms appear appropriate to our needs. However, in order to test the robustness of estimated parameters with respect to the functional form, we also estimate the model using a log-logistic density function.

The hazard function of transiting from state j to state k can be written as follows :[14]

$$\begin{aligned} &Weibull\ Distribution: && \lambda_{jk}(t|X_1, X_2(t)) = \alpha_j t^{(\alpha_j - 1)} \exp\left[X_1\beta_{jk} + X_2(t)\gamma_{jk}\right] \\ &Log\ Logistic\ Distribution: && \lambda_{jk}(t|X_1, X_2(t)) = \frac{\alpha_j t^{(\alpha_j-1)} \exp\left[X_1\beta_{jk} + X_2(t)\gamma_{jk}\right]}{1 + t^{\alpha_j} \exp\left[X_1\beta_{jk} + X_2(t)\gamma_{jk}\right]} \end{aligned}$$

where β_{jk} is the parameter of time invariant variables, which is allowed to vary across transitions. For instance, we allow the effect of the training status (dummy variable) to be different when an individual moves from unemployment to a qualified job and when he moves from unemployment to employment in an unqualified job. The vector γ_{jk} captures the effect of the time-varying covariates. Furthermore, we assume that the shape parameter, α_j is origin-specific only.[15]

On the other hand, for the probability of employment or training after experiencing an unemployment spell, we assume that the shape parameter is different regarding the destination state. In the job search literature, the probability of employment depends on the probability of receiving an offer and the probability of accepting an offer, which in turn depends on the reservation wage. As shown in Section 1.2, few individuals get a qualified job after experiencing an unqualified job and *vice versa*. Thus we suspect that the parameter α will differ according on whether an individual moves from unemployment to a qualified job or to an unqualified one. This is also in line with the job search theory which stipulates that employers rely on the employment record as a signal of potential productivity (Lynch, 1986).

1.3.4 Accounting for Human Capital Quality

Previous studies have shown that human capital quality explains a significant part of wage gap and employment history among immigrants. Following this literature, we proxy the human capital quality by the relative Gross Domestic Product (GDP) of the home country to

14. For simplicity, the subscript for spells are omitted since only origin and destination states matter here. On the other hand, the Weibull distribution is the first choice to map the hazard rate function because the probability density function of the duration in each state looks like the Weibull distribution (see figure A.1 that represents the duration density function in unemployment spells. We have the same pattern for other states.).

15. A value of $\alpha_j < 1$ indicates that the failure rate decreases over time. This is often referred to as "duration dependence" in the literature. A value of $\alpha_j = 1$ indicates that the failure rate is constant over time. The Weibull distribution thus reduces to an exponential distribution. Finally, a value of $\alpha_j > 1$ indicates that the failure rate increases with time.

that of the Canadian GDP.[16] Let HCQ_i stand for Human capital quality of immigrant i, and GDP_i be the per capita GDP of the country where he obtained his highest diploma before getting his permanent residence status in Canada. Likewise, let GDP_{CAN} represent the GDP per capita of Canada. Our indicator is giving by[17] :

$$HCQ_i = \ln\left(\frac{GDP_i}{GDP_{CAN}}\right). \tag{1.7}$$

HCQ_i is negative whenever the highest diploma was obtained in a country with lower level of development than Canada and positive otherwise. It is equal to zero for immigrants who have got their highest degree in Quebec and, therefore were already in Quebec before getting permanent resident status. In our sample, roughly 6% of immigrants obtained their highest degree in Quebec.

We introduce this variable as time-invariant since is obtained by averaging the GDP per capita over the period covers by the survey[18]. Moreover, we measure its effect in two ways. First, we measure its direct and indirect effect on labour mobility. The indirect effect runs through the level of education. Secondly, we investigate to what extent the effect of foreign-education could change regarding the level of HCQ. Formally, the total effect of HCQ, denoted $\Lambda^{i,HCQ}_{HCQ}$, is given by[19] :

$$\Lambda^{i,HCQ}_{jk} = \beta^{hcq,d}_{jk} HCQ_i + \beta^{hcq,id}_{jk} HCQ_i \times S = HCQ_i(\beta^{d}_{jk} + \beta^{hcq,id}_{jk} \times S),$$

with S standing for the level of education and the total effect of Education, denoted $\Upsilon^{s,i}_{jk}$ is given by :

$$\Upsilon^{s,i}_{jk} = S_i(\beta^{s,d}_{jk} + \beta^{s,id}_{jk} * HCQ_i).$$

$\beta^{s,id}_{jk}$ measures the effect of education in terms of quality whereas $\beta^{s,d}_{jk}$ captures the effect in terms of the level of education when an individual moves from state j to state k. If $\beta^{s,id}_{jk}$ and $\beta^{s,d}_{jk}$ have the same sign, it implies that the HCQ reduces the effect of the level of education since the log(HCQ) is negative for 90% of immigrants.

The GDP per capita is not assuredly the best proxy for the human capital quality since it does not account for the level of wealth inequality in each country, which can affect migrant's education quality. An alternative would have been to have psycho-technical or vocational tests done for all immigrants upon their arrival in Canada.

16. In the literature, it has been found that the immigrants' human capital quality is closely related to the level of development of the country of origin.
17. GDP per capita for each country are drawn from World Bank data base
18. However, the GDP per capita does not change that much over this period.
19. The subscripts d and id stand for direct and indirect effects, respectively.

1.3.5 Allowing for Unobserved Heterogeneity

As mentioned previously, it is important to take into account the unobserved heterogeneity among immigrants. Heckman & Singer (1984) stresses that a wrongly specified functional form for the unobserved heterogeneity could seriously bias the parameter estimates. To overcome this issue, the unobserved heterogeneity is modelled as a weighted mixture of *iid* standard normal random variates.[20] Hence, let ω_j represents the unobserved heterogeneity specific to state j. In this setting, we assume that ω_j may be written as

$$\omega_j = \phi_j \xi_1 + \psi_j \xi_2 \tag{1.8}$$

where ξ_1 and ξ_2 are drawn from a standard normal distribution. To insure the model is identified, we impose a minimal set of restrictions on the loading factors, namely that $\psi_e = 1$ and $\phi_j = 1$, $\forall$ j $\in \{u, e', r\}$.

The hazard rate and the likelihood function for an individual i are then given by :

$$\lambda_{jk}(t|X_1, X_2(t), \omega_j) = \alpha_j t^{\alpha_j - 1} \exp\left[X_1 \beta_{jk} + X_2(t)\gamma_{jk} + \omega_j\right]$$

$$\widehat{L_i} = \frac{1}{M} \sum_{m=1}^{M} L_i(t|X_1, X_2(t), \omega_{jm})$$

where $L_i(\cdot)$ is the contribution of individual i to the total likelihood and is defined as previously. Finally, the total log-likelihood that we will maximize is defined as follow :

$$\widehat{\log L} = \sum_{i=1}^{N} \log \left[\frac{1}{M} \sum_{m=1}^{M} L_i(t|X_1, X_2(t), \omega_{jm})\right] \tag{1.9}$$

M is the size of vector ξ_1 and ξ_2. The maximization of the simulated likelihood function yields consistent and efficient parameter estimates whenever $\frac{\sqrt{N}}{M} \longrightarrow 0$ when $N \longrightarrow \infty$ and $H \longrightarrow \infty$ (see Gouriéroux & Monfort (1996, 1991); Brouillette & Lacroix (2011)). Although convergence is achieved when $M \geq 20$, we have chosen to set $M = 100$. In this setting the correlation between states j and k is given by :[21]

$$Corr(\omega_j, \omega_k) = \frac{\phi_j \phi_k + \psi_j \psi_k}{\sqrt{\left(\phi_j^2 + \psi_j^2\right)\left(\phi_k^2 + \psi_k^2\right)}} \tag{1.10}$$

1.3.6 Expected Duration and Instantaneous Probability

The sign of the parameter estimates indicates how the hazard function varies relative to a marginal change in a given variable. The interpretation of the size of the parameter is much

20. Identification issues may arise when there are too many mass points, as show by Baker & Melino (2000).

21. The results using a fully non-parametric distribution function are available upon request. The main differences from our preferred specification are twofold : (1) Age do not affect the hazard rate while it does in the simulated likelihood and (2) the shape parameters, α, all tend to be lower.

more difficult to assess. These are best understood if we translate them into their marginal effect on the expected duration in a given state. By computing the expected duration in different states, we can then estimate the expected proportion of time an individual spends, conditional on his characteristics, in each state over a given period of time.

Let ED_j be the expected duration in state j conditional on individual characteristics :

$$ED_j = \frac{1}{M} \sum_{m=1}^{M} \times \int_0^{\infty} \tau f_j \left(\tau | X_1, X_2(t); \omega_{jm}\right) d\tau, \tag{1.11}$$

where $f_j(\cdot)$ is the density of the random duration variable in state j, conditional on individual characteristics, $X_1, X_2(t)$, when the destination state is unknown. Up to now, we have relied on the density function $f_{jk}(\cdot)$ rather than $f_j(\cdot)$. We must derive the density function $f_j(\cdot)$ from the densities $f_{jk}(\cdot)$, $j, k = u, e, e, r$ and $j \neq k$. This can be achieved as follows :

$$f_j(\iota_j | X_1, X_2(t), \omega_{jm}) = \frac{1}{M} \sum_{m=1}^{M} \sum_{\substack{h \in \{u, e, e', r\} \\ h, l, k \neq j}} f_{jh}(\iota_j | X_1, X_2(t), \omega_{jm}),$$

with $\iota_j = \min_h(T_{kh})$, $h \in \{u, e, e', r\}, h \neq k$.

As mentioned above, the expected duration is a relevant statistics since we can compute it for each state despite the fact that a given individual may not as yet spent time in a specific one.

Furthermore, it could be interesting to compute the instantaneous probability that an individual be observed in a particular state since landing (as in Figure 1.1). Let $D_j(t)$ be a dummy variable equal to 1 if the individual is in state j at time t and zero otherwise. The probability that an individual occupies a state j at time t is equal to the probability that at time $t - \delta t$ he is already in state j and does not move to another state in δt periods of time, plus the probability that he is in state h, $\quad h \neq j$ at time $t - \delta t$ and moves from h to j during the period of time δt. Formally, $P(D_j(t) = 1)$ is given by [22]

$$\begin{aligned} P(D_j(t) = 1) = P(D_j(t - \delta t) = 1) \times & \sum_{h \in \{u, e, e', r\}, h \neq j} (1 - \lambda_{jh}(t) \partial t) + \\ & \sum_{h \in \{u, e, e', r\}, h \neq j} P(D_h(t - \delta t) = 1) \times \lambda_{hj}(t) \partial t. \end{aligned} \tag{1.12}$$

To compute the instantaneous probability of each state we need to solve this first-order system of equations for t. [23]

22. For simplicity, the conditioning on observed individual characteristics is omitted.
23. See A.2.2 for the details of the derivation.

1.4. Estimation Results and Discussion

As mentioned above, we estimate two versions of the econometric model. The first one treats education as exogenous while the second treats it as endogenous. The model is estimated using both a Weibull and a Log-logistic specifications. For the sake of brevity only the results based on the Weibull specification are reported in the paper. [24]

1.4.1 Education as an Exogenous variable

Table 1.6 reports the parameter estimates when education is treated as exogenous. We define *QC-Education* as a dummy variable that equals one if at any time since landing an individual attended school on a full-time basis. A number of interesting results emerge from the table. According to the parameter estimates, being educated in Quebec appears to have positive effects on labour market outcomes as it negatively impacts the transition rates into unemployment and out of employment. The new selection grid of 2006 has no impact on most transitions, save for the fact that it appears to have hastened the transition from unqualified employment into unemployment. Female immigrants are less likely to transit into employment (qualified or not) and thus spend more time unemployed than males.

Other results worthy of mention include all human capital variables. Having a foreign Master's or a PhD diploma, relative to a secondary degree, increases the duration of unemployment spells prior to obtaining a qualified job. On the other hand, the duration of qualified jobs, conditional on holding one, is longer for these highly qualified workers. Holding a bachelor degree has mixed effects on the labour market transitions. Although unqualified jobs tend to last longer, access to a qualified job is difficult for these workers. It is as if employers deem these workers as overqualified and are reluctant to offer them employment opportunities in line with their abilities. Having a high fluency score in French positively affects the transitions into employment. On the other hand, a high fluency score in English favours the transitions from unemployment into qualified jobs. Married immigrants experience longer employment spells, both qualified and unqualified.

24. The results of the Log-logistics specification are available upon request.

TABLE 1.6 – *Parameter Estimates : Exogenous Education*

Hazard functions	λ_{eu}	λ_{ue}	$\lambda_{ue'}$	$\lambda_{e'u}$
QC-education	-0.60***	0.25***	0.62 ***	0.14†
2006-Grid Reform	0.39**	0.10	0.17	0.23
Female	-0.04	-0.27***	-0.20**	-0.04
Foreign Education				
Master or PhD	-0.12	-0.05	-0.52***	-0.29*
Bachelor	-0.31***	0.07	-0.78*	-0.07
Fluency Score - French	0.01	0.03**	0.04***	0.02†
Fluency Score - English	-0.03	-0.002	0.09***	0.02
Married	-0.16*	-0.09	0.11	-0.16†
HCQ	0.04	0.12***	0.26***	-0.05
HCQ × Master or PhD	-0.11	-0.06	-0.19*	-0.08
HCQ × Bachelor	-0.07	-0.11*	-0.05	0.07
Age at admission	-0.15***	-0.11***	-0.14***	-0.22***
(Age at admission)^2	0.002***	0.001***	0.002***	0.003***
Previous stay	-0.24**	-0.41***	0.28***	-0.53***
Experience	-0.17	0.06	0.31***	-0.15
Landing Year(2002)				
2003		0.24*	0.06	
2004		0.33**	0.13	
2005		0.17†	-0.0	
2006		0.13	-0.17	
2007		0.19 *	-0.01	
2008		0.21†	-0.08	
2009		0.16	-0.39*	
	Ancillary Parameters			
α (Shape parameter)	0.86***	0.67***	0.67***	0.86***
ϕ_e		-0.23**		
ψ_u		0.15**		
ψ'_e		-0.42***		
	Correlation Matrix			
		ω_e	ω_u	ω'_e
	ω_e	–	-0.075	-0.585
	ω_u	-0.075	–	0.853
	ω'_e	-0.585	0.853	–

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Conditioning on the quality of foreign education (HCQ) yields interesting results. Indeed, immigrants originating from high-income countries (relative to Canada) have shorter unemployment spells and faster transition into qualified and unqualified jobs. On the other hand, the relative level of development has no impact on the duration of employment spells. Having experienced a previous stay in the province is perhaps the main factor that explains the dynamics of the labour market transitions. Indeed, this variable alone is responsible longer employment spells, both qualified and unqualified, but also longer and shorter unemployment spells that end in unqualified and qualified jobs, respectively.

Recall that the shape parameter, α_j, indicates how the hazard rate in state j varies with the duration. The parameter estimates reported in Table 1.6 are all below one which implies negative duration dependence. In other words, the longer one remains in a given state the less likely he/she will exit the state. More interestingly, the parameters that govern the unobserved heterogeneity are all statistically significant (see equation (1.8)). These can be expressed in terms of correlation between labour market states (equation (1.10)). According to the table, individuals who experiment short unqualified jobs tend to have long qualified job spells. Likewise, those who experiment short unemployment spells have shorter than average qualified job spells, but slightly longer than average unqualified job spells (although the correlation is relatively weak in the latter case.) These results suggest that those who transit across the two employment states may have different unobserved characteristics.

1.4.2 Endogenous Education

Table 1.7 reports the parameter estimates of a model in which education is a separate state. Its endogeneity is captured through the correlations with the alternative states. The table is set-up in a similar fashion to Table 1.6 except for addition of two variables :

1. *QC-Education*$_1$ is a dummy variable equal to one if a domestic diploma was granted just prior to the current spell.
2. *QC-Education*$_2$ is similarly defined as *QC-Education*$_1$ but is equal to one if a domestic diploma was granted at least two spells prior to the current one.

A comparison between Tables 1.6 and 1.7 underlines the importance of treating domestic education as a choice variable. To begin with, obtaining a diploma in the near past (*QC-Education*$_1$) is now associated with much longer employment and unemployment spells. This is not surprising given that those who invest in domestic education do so soon upon landing. They thus have little work experience and must face the usual school-to-work transition hurdles. Those who elect not to invest in domestic education have gone through many more transitions on the labour market and have thus accumulated more experience. Fortunately, the beneficial effect of domestic education on employment duration does not taper off (*QC-Education*$_2$) whereas the negative impact on unemployment duration either vanishes. Distant education also tend to expedite the duration of unemployment spells prior to landing a qualified job. The impact of education is thus relatively complex and can only be ascertained by turning to numerical simulations, as in the next section. The remaining estimates of the slope parameters are qualitatively similar to those of Table 1.6, save for a few who gain or lose statistical significance. Note however that, contrary to the results of the previous table, a higher HCQ now translates into longer employment spells. Not surprisingly, it is also found that a higher HCQ will lead to fewer transitions into training.

Yet small changes in these parameters may have important consequences on the labour market dynamics. This is investigated below. On the other hand, all ancillary parameters witness important changes. This is to be expected from treating education as an endogenous variable. As in the exogenous specification, their impact on the dynamics will be investigated through numerical simulations.

Most hazard functions exhibit negative duration dependence as most the shape parameters reported in the table are below one. Interestingly, the transition rates between unemployment and schooling is constant ($\alpha = 0.92$). This implies that the probability of entering formal education, conditional on being unemployed, is independent of the duration of the spell. As in Table 1.6, the bottom panel presents the implicit correlation coefficients between the various states considered in the model. The most noteworthy feature of the table concerns the correlations with the schooling duration (ω_r). Individuals with unobserved characteristics that are favourable to long unemployment spells will also have longer than average schooling durations. Likewise, those with long qualified jobs will tend to have longer than average schooling durations. This is consistent with the idea that some have frequent short spells, while other have fewer but longer spells in each state.

Table 1.7 – *Parameter Estimates : Endogenous Education*

Hazard functions	λ_{eu}	λ_{er}	λ_{ue}	$\lambda_{ue'}$	λ_{ur}	$\lambda_{e'u}$
QC-education$_1$	-0.69***		-1.12***	-1.36***		-0.67***
QC-education$_2$	-0.71***		-0.11	0.36***		-0.53***
1996-Grid Reform	0.30*	0.37	0.22	0.37*	-0.17	0.23
Female	-0.005	0.041	-0.23**	-0.14†	-0.25*	0.09
Foreign Education						
Mast-PhD	-0.17	0.09	0.005	-0.53***	-0.19	-0.29†
Bachelor	-0.27†	0.29	0.15	-0.78***	-0.10	-0.15
Fluency Score - French	0.015	0.034	0.034**	0.052***	0.03†	0.029†
Fluency Score - English	-0.016	-0.007	-0.013	0.086***	-0.036	0.002
Married	-0.13	-0.14	-0.14*	0.15†	-0.034	-0.17
HCQ	0.02	-0.22†	0.14**	0.33***	-0.02	-0.11†
HCQ $\times$Master-PhD	-0.010	0.11	-0.05	-0.27**	-0.09	-0.04
HCQ $\times$ Bachelor	-0.03	0.08	-0.08	-0.064	-0.075	0.02
Age at admission	-0.16***	-0.31***	-0.12***	-0.15***	-0.17***	-0.22***
(Age at admission)^2	0.002***	0.004***	0.001***	0.0017***	0.0015***	0.0031***
Previous stay	-0.29**	-0.04	-0.40***	0.32***	-0.28*	-0.50***
Experience	-0.06	-0.24	0.06	0.35***	0.098	-0.09
Landing Year (2002)						
2003			0.23*	-0.09		
2004			0.18	-0.02		
2005			0.26*	-0.27*		
2006			0.15	-0.34**		
2007			0.11	-0.23†		
2008			-0.06	-0.44**		
2009			-0.034	-0.57**		
	Ancillary Parameters					
Shape parameter (α)	0.85***	0.85***	0.66***	0.63***	0.92***	0.88***
ϕ_e			-0.05			
ψ_u			-0.18**			
ψ'_e			-0.59***			
ψ_r			-0.0053			

Correlation Matrix

	ω_e	ω_u	ω'_e	ω_r
ω_e	–	0.130	-0.549	-0.057
ω_u	0.130	–	0.757	0.983
ω'_e	-0.549	0.757	–	0.866
ω_r	-0.057	0.983	0.866	–

$^{\dagger}p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes :

1. QC-education$_1$ is an indicator variable equal to one if a diploma was obtained in the previous spell. QC-education$_2$ is equal to one if a diploma was granted at least two spells prior to the current one.
2. The transitions from education to employment or unemployment are also estimated in this specification. For the sake of brevity, they are not reported in the table but are available upon request.

1.4.3 Expected Proportion of Time Spent in Each State

The parameter estimates of Tables 1.6 and 1.7 capture fairly complex dynamics on the labour market. While their signs are easily interpreted, their marginal effects are not and are best understood by computing the expected proportion of time spent in each state, conditional on the values of the exogenous variables (see equation 1.11). Table 1.8 reports such proportions using the parameter estimates of Table 1.7. According to the model, the immigrants in our sample will spend roughly 68% of their time employed in qualified employment (women slightly less so). Recall that these individuals are highly qualified and most hold at least a bachelor's degree. Holding a domestic degree is associated with slightly more time spent unemployed, although the difference is relatively small. Having spent some time in Quebec in the past significantly reduced unemployment and increases time spent in a qualified job. From Table 1.7, holding a Master's or a PhD degree had a positive impact on the duration of both unemployment and employment spells. It turn out that holding such a degree is highly beneficial in terms of time spent employed in a qualified job and (little) time spent unemployed. Immigrants originating from Africa fare worst on the market and those from Europe fare best.[25] The last lines of the table focus on the interactions between age and education. Interestingly, irrespective of the degree one holds, immigrating at a young age is highly beneficial. Thus youths with a post-secondary or less spend almost as much time employed in qualified jobs and out of unemployment as those with a bachelor's degree or above but who were admitted when above 30 years of age.

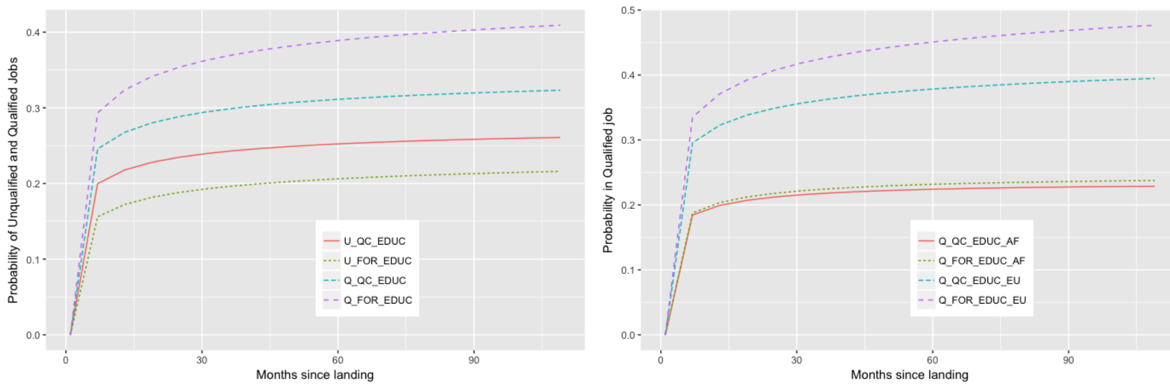
25. The country-of-origin effect is mediated through the HCQ variables.

Table 1.8 – *Conditional Expected Proportion of Time in Each State*

Characteristics		P_e	P_u	$P_{e'}$
Overall		0.21	0.11	0.68
Male		0.21	0.1	0.69
Female		0.22	0.12	0.66
QC-education		0.22	0.12	0.66
Foreign-education		0.21	0.09	0.7
Previous stay in QC		0.19	0.07	0.74
No Previous stay in QC		0.23	0.13	0.65
Master's or PhD degree		0.17	0.08	0.75
Bachelor's degree		0.23	0.11	0.66
Post-secondary less		0.22	0.12	0.67
Africa		0.24	0.15	0.62
Asia		0.21	0.11	0.68
America		0.2	0.11	0.69
Europe		0.2	0.08	0.72
Age at admission ≤ 29	Master's or PhD degree	0.19	0.06	0.76
	Bachelor's	0.24	0.07	0.69
	Post-secondary or less	0.23	0.07	0.7
Age at admission > 29	Master's or PhD	0.17	0.09	0.74
	Bachelor's	0.22	0.12	0.66
	Post-secondary or less	0.21	0.13	0.65

An alternative way to highlight the main features of our model is to focus on the conditional probabilities of being observed in various states as time unfolds (see equation 1.12). These probabilities are reported in Figure 1.3. The left-hand side figure illustrates the probability of occupying either a qualified or an unqualified job according to whether the degree was obtained in Quebec or in the country of origin. As is readily apparent, the probability is higher for those who obtained their degree in their home country. This is somewhat surprising but reflects the dynamics that are implicit in the parameter estimates of Table 1.7. Conversely, holding an unqualified job is more likely for those who are trained in Quebec as opposed to those trained in their home country. The profiles are relatively smooth and are a good fit to the data reported in Figure 1.1. The figure on the right-hand side exhibits the probability of holding a qualified job according to the continent of origin (Europe or Africa) and the country where the degree was granted. The figure shows that immigrants from Europe are much more likely to hold a qualified job than those from Africa. Further, Europeans holding a degree from Quebec are less likely to hold such a job. Conversely, Africans are slightly more likely to hold a qualified job if they received their degree in Quebec. These figures indicate that the labour market dynamics are complex and intimately linked to observed (and unobserved) characteristics. Simple models of labour force participation are likely to miss much of the complex interactions between various states and choices made by immigrants as time unfolds.

FIGURE 1.3 – *Conditional probability to be in a particular state*



Note : Q_QC-EDUC_AF and **Q_QC-EDUC_EU** stand for "Qualified Job, Quebec diploma", from Africa and Europe, respectively. **Q_QC_EDUC**, **U_QC_EDUC** , **Q_FOR_EDUC** and **U_FOR_EDUC** stand for qualified (Q) or unqualified (U) job with a Quebec (QC_EDUC) or Foreign (FOR_EDUC) degree.

1.5. Conclusion

This paper investigates the labour market histories of highly qualified immigrants who landed in the Province of Quebec between 2002 and 2009. In particular, we focus on their transitions between qualified and unqualified employment, unemployment and investment in education. Our goal is to estimate the causal effects of schooling on the likelihood of finding a job that is suited to their abilities. Despite being highly qualified (Bachelor degrees and above), many immigrants experiment long spells of unemployment upon landing. Some are thus induced to undertake formal education to overcome the employment hurdle. Assuredly, a sizable fraction of this population manages to find adequate employment without resorting to domestic schooling. They are likely different from those who do, both from an observable and an unobservable point of view. In other words, domestic education is likely a choice variable whose endogeneity must be accounted for. Our results suggests that immigrants who graduated in Africa have a greater incentive to invest in domestic schooling since it increases their likelihood of holding both qualified and unqualified jobs over time.

Our results suggest that domestic schooling has a positive effect on the medium and long terms employment rates, but a negative one in the short term. In addition, immigrants who have spent some time in the province prior to receiving their permanent residence status also perform better on the labour market. This is perhaps why the provincial government has increased the weight it allocates to previous stays in its selection grids over time.

In recent years, the government of Quebec has selected highly educated immigrants in the hope that they would be more work ready. According to our results, those who hold either a Master's or a PhD degree need to wait longer to find a suitable job, but are also less likely to move back into unemployment once they are hired. The time spent in unemployment is virtually identical irrespective of the educational level. As pointed out by Renaud (2005) and Lacroix (2013), immigrants with a foreign bachelor's degree have a harder time finding a job that matches their skills. Since the majority of immigrants hold a bachelor's degree, 48% versus 19% for a Master's or PhD, it may be advisable from a policy perspective to favour immigrants holding the latter.

Finally, we find that domestic schooling and labour market outcomes are intimately related through unobservable characteristics. This suggests that despite the richness of our data, unobserved factors play an important role in explaining labour market transitions. Assuredly, this makes designing appropriate policies quite challenging for policy-makers.

- Chapitre 2 -

Is it Beneficial to have host-country acquired-education ? evidence of migrants' earnings profile in Quebec.

Abstract

This study investigates the gains of Quebec-acquired education over the foreign-acquired education in terms of the logarithm of weekly earnings of the migrants in the category of skilled workers in Quebec. Unlike previous studies, we observe the level of education of immigrants acquired abroad and the periods of investment in education in Quebec. We then model the migrants self-selection into the schooling investment and, by using the Bayesian treatment analysis we estimate the distribution of the causal effect. We find that the return to Quebec-acquired education is negative and the magnitude is different from one migrant to another. Therefore, computing a single average effect does not allow to capture this essential heterogeneity. Particularly, the gains tend to decrease with the likelihood to enrol in school and with the level of *ability*or the effect of social network. Thus, our results suggest that individual employer pay migrants not regarding their level of education or its origin but more importantly due to the quality of the prior jobs held. While the Bayesian approach suggests that immigrants who have enrolled in university degree in Quebec are the most negatively affected, the frequentist approach suggests that those immigrants obtain the highest positive return from QC-acquired education. This raises again the issue of mis-evaluation when the essential heterogeneity is not taking into account.

Keyword Foreign-acquired education, QC-acquired education, earnings profile, Bayesian treatment analysis, unobserved heterogeneity, Frequentist approach

JEL-codes : C31, C41, J15, J24, J64, J61

2.1. Introduction

Many developed countries have elaborated sophisticated immigration policies to regulate the flow and composition of the immigrant population. In Quebec (Canada), such a program known as the "Skilled Worker Program" selects immigrants whose characteristics are likely to ease integration onto the labour market. The latter are thus generally more highly educated and younger than the average immigrants and natives. Despite the fact that nearly 75% of the selected immigrants have at least university degree prior to landing and 80% of them have considerable work experience in their last country of residence, we observe that as many as 50% of them return to school upon landing in Quebec. It is rather surprising that with such endowments at entry, some need to further their skills in order to successfully integrate the labour market. This paper investigates the extent to which this investment enables them a better access to higher paid jobs than their foreign-acquired skills.

In Canada, Skuterud & Su (2012) find that host-country education provides a small benefit over the foreign-acquired one, but very little return to foreign-acquired experience. Moreover, Canadian-educated immigrants have high-paying job compared to their foreign-educated counterparts, but their earnings are lower than those born in Canada (Hou & Lu, 2017). Other studies conducted in the United states, Israel, Netherlands, Sweden, similarly conclude that the returns to host-country schooling are much larger (Akresh, 2007; Bratsberg & Ragan, 2002; Cohen-Golden & Eckstein, 2008; Kanas & van Tubergen, 2009; Nordin, 2007). On the other hand, it also emerged that the foreign-acquired human capital is not equally evaluated according to its country of origin. Indeed, the foreign-human capital from more developed countries generate higher returns (Coulombe *et al.* , 2014b; Kanas & van Tubergen, 2009).

Another factor that might explain domestic investment is the quality and the quantity of job offers received due to the economic environment or to due market failures. With regard to the first point, the unemployment rates among natives of the same age group and gender as migrants don't vary significantly across cohorts. Therefore, it is probably safe to say that the economic environment matters little in inducing immigrants to return to school even though their unemployment rates are much higher. However, the data show that the duration of the first unemployment spell is higher for immigrants who have decided to obtain a diploma from Quebec. One can conjecture that the lack of appropriate job offers may have led them to go back to school. In addition, the data show that they have lower earnings in their first job once they have graduated than those who have decided not to return to school. [1] This can be explained by the fact that right after graduation, migrants are more willing to accept lower paid jobs.

1. The data also show that earnings at the first job is higher for the recent cohort of immigrants (see table 2.1), which suggests that employers in Quebec may have acquired enough of information over time to well evaluate foreign credential.

In this study, we simultaneously investigate the decision to invest in domestic education and the distribution of the gains of the latter over foreign-acquired education. We use rather unique data from a retrospective survey on the immigrants from Quebec Skilled Workers Program combined with a large administrative data set, the so-called "*landing file*". The latter provides detailed information on foreign human capital. The majority of the previous studies proxy the foreign-education by assuming that school accumulates continuously after the age of five, and that one year after schooling in Canada corresponds to one year of work experience. Those two assumptions are restrictive since most immigrants have worked in their country of origin before migrating, and the second assumption ignores the potential periods of unemployment after completing the studies in Canada. We can then avoid all biases due to measurement errors. To the best of our knowledge, this is the first study that attempts to conduct such an analysis. This is possible because we have detailed information on all periods of enrollment in school, the nature of the training (university or post-secondary degree) as well as their duration. In addition, we observe all periods of employment and unemployment upon landing.

The main contribution of this paper is to estimate the dynamic average return to the post-migration schooling in terms of weekly earnings gain of QC-acquired education over foreign-acquired education. The originality of the paper comes from the fact that we are using the Bayesian treatment analysis to estimate individual average "*gain*" of QC-education over the foreign-education, conditional on investing in domestic education as well as the "*loss*" for those who have not. In this way, we contribute to the ongoing discussion about the advantages of host-country education over the foreign education since our identification strategy allows to self-match each individual. As pointed out in Card (2001), the assumption that the return to education is identical for all individuals could bias the average return to education. It is important to provide accurate distributional results since migration policies are based on such analyses. Indeed, Skuterud & Su (2012) pointed out that because studies find virtually no return from work experience acquired abroad, the Canadian government requires at least one year of work experience prior to requiring the permanent resident status [2]. In Quebec, the immigration policy gives priority to international students.

The Bayesian approach has recently been extended to the treatment effect analysis (see Heckman *et al.* (2012) for a review), thus allowing to tackle in an alternative manner the problem of selection into training as it might be related to individual unobserved heterogeneity. Indeed, the investment in education might be intrinsic to unobserved individual tastes, abilities, social or career prospects. Indeed, the determinants of the decision to return to school are entirely on the demand side as there are no barriers to entry education. This is because in Quebec all immigrants can apply for government loans and grants. In addition the tuition fees are the same as those of natives, which is not the case for international students. Thus, it

2. This requirement targets international students who want to stay in Canada upon graduating and temporary workers.

might be that the same unobserved factors as social network or *ability* that prompt a migrant to return (or not) to school can also affect his earnings profile. Accordingly, it is more likely that the return to post-migration schooling is different from one migrant to another given the migrant's social network or *ability*. To account for that heterogeneity, this paper estimates the "*mean effect*" of post-migration schooling in terms of individual earnings gain for the population of immigrants who were admitted under the SWP in Quebec between 2002 and 2009 with the permanent residency status. [3]

We are not the first to use the Bayesian treatment analysis to estimate the return to schooling. In fact, Carneiro *et al.* (2003) implement this approach to estimate the distribution of the returns to schooling on youth's earnings in the United States. Their results show that the return to education is different across individuals in that the effect is positive for some and negative for others. Unlike Carneiro *et al.* (2003), we don't have access to measurement variable to extract unobserved factors as test scores in mathematics or in science. To solve this problem, we use a method propose by Lindley & Smith (1972) and Chib & Hamilton (2002) to simulate a distribution of the migrants fixed effects and the associated loading factors. Loading factors measure to what extent unobserved variables that affect the earnings profile are correlated with the decision to invest in education in Quebec.

This study differs somewhat from those who have implemented the Bayesian treatment analysis in a dynamic setting. In fact, most of them simultaneously estimate the decision to enroll in schooling and the outcome equation at each period. In our case, since migrants have to complete their training before finding a jobs, we can not estimate the effect and the decision concurrently. Therefore, we first assume that migrants have the possibility to invest in education during the first four years following their arrival, which captures about 80% of people who have actually enrolled in Quebec-education. Afterwards, we model the earnings profile for the following years. This strategy has the advantage of accounting for the full dynamic sequence of schooling investment. [4]

This paper also underlines how omitting the unobserved confounders and assuming identical individual returns may lead to inappropriate inference about the average treatment effect, let alone forbid any inference about the distribution of the effect across individuals. As is now customary, we compute the dynamic average return to education by means of the Sequential Inverse Probability Weighting.

This paper is organized as follows. The next section describes the data (for more details see Djuikom & Lacroix (2018)). Section 3 presents the econometric model and identification strategy. Section 4 presents the results and discussion. Section 5 concludes.

3. We use the term "*mean effect*" to refer to the mean of the posterior distribution of the return to education for each person (see explanation in empirical strategy section 2.4).

4. Lechner (2009) and Sianesi (2004) show, in the context of a potential outcomes analysis, the importance of accounting for the full sequence of treatments to avoid to estimating biased average effect.

2.2. Data Description and Schooling Assignment Process

2.2.1 Data Description

Between January 2002 and December 2009, approximately 90,000 migrants were admitted in Quebec through the Skilled Workers Program. Our data are a random sample drawn from the so-called "landing file", a large administrative data managed by Immigration & Citizenship Canada. The file contains all relevant pre-migration information needed to evaluation eligibility to the program. These include the highest degree of schooling (henceforth foreign-acquired education), results from language tests (French and/or English), the exact date of the entry in Canada, *etc.* The administrative information is complemented by a retrospective survey through which detailed information on all periods of employment, unemployment and periods of schooling are reported. Earnings, occupations (NOC), hours of work, *etc.* are also recorded in the survey. The latter covers the period since arrival up until May 2011. Our sample contains 3,009 individuals, 9.7% of whom have not found (or chosen) a job since landing. The retrospective survey was conducted conjointly by the Ministry of Immigration and Cultural Communities and the Ministry of Employment and Social Solidarity of the Quebec Government. According to Benzakour *et al.* (2013), the sample is representative of the migrant population that arrived in Quebec during the corresponding period. Immigrants from Asia, and particularly those from the Middle-East, are however slightly under-represented in the sample. All other demographic characteristics (gender, marital status, educational background, *etc.*) are representative of the migrant population. [5] The survey design is such that we may follow individuals for a maximum of nine years and a minimum of three. The selection is based on a *"Grid points"* system that consists of allocating points on targeted individual credentials to be evaluated by the Quebec government. Moreover, we can distinguish two cohorts of migrants selected based on different *"Grid points"* system : the first grid was established in 1996 (henceforth "Grid-96") and the second one in (henceforth "Grid-06"). The latter grid puts more weight on human capital and past Quebec experience (past stays for all purposes prior to landing). It is important to stress that many migrants who arrived after 2006 were selected under the Grid-96, and that most migrants selected under the Grid-06 arrived in either 2008 or 2009. [6]

In Table 2.1 we have reported the weekly earnings of migrants in the first and third jobs respectively in nominal and real values adjusted with the consumer price index (CPI). Since migrants who have been selected under the Grid-06 have not yet spent too much time in Quebec at the time of the survey, we focus on the earnings of their first job for comparison. Overall, not only do earnings increase over time, but the earnings at entry in the labour market is higher for newcomers, namely for those who landed in 2008. In fact, in real value

5. See Benzakour *et al.* (2013) for more discussion on the representativeness of the sample.

6. For more details about the data and migrant profile, see Djuikom & Lacroix (2018), we will focus here only on information relevant to this study.

Table 2.1 – *Earnings per week and cohort*

EARNINGS OF	Nominal earnings			Real earnings		
	1996 Grid [2,322]		2006 grid [185]	1996 Grid [2,322]		2006 grid [185]
	1^{st} job	3^{rd} job	1^{st} job	1^{st} job	3^{rd} job	1^{st} job
FOREIGN-EDUCATION(FE)						
Post-sec or less	544.9 [388]	670.7 [138]	888.8 [24]	512.0 [388]	633.0 [138]	829.1 [24]
BAC	601.0 [1,019]	789.2 [384]	575.4 [53]	567.2 [1,019]	744.5 [384]	528.5 [53]
Master or Doctorate	757.4 [649]	882.3 [223]	859.2 [82]	716.5 [649]	836.7 [223]	819.5 [82]
ORIGIN of FE						
Africa	519.1 [628]	712.6 [217]	472.3 [12]	487.9 [628]	612.1 [217]	423.8 [12]
America	670.2 [434]	824.8 [157]	691.9 [39]	632.9 [434]	786.5 [157]	645.4 [39]
Asia	667.0 [106]	662.3 [20]	653.6 [10]	631.1 [106]	640.7 [20]	609.0 [10]
Europe	703.4 [885]	836.7 [351]	841.2 [96]	664.9 [885]	787.8 [351]	798.9 [96]
If HAVE QC-EDUCATION						
Yes	565.9 [1,084]	747.7 [436]	609.9 [49]	531.3 [1,084]	703.3 [436]	569.9 [49]
No	721.1 [977]	859.9 [311]	839.9.4 [110]	683.9 [977]	817.6 [311]	792.5 [110]
GENDER						
Female	568.3 [672]	806.9 [229]	635.6 [61]	537.1 [672]	757.8 [229]	596.5 [61]
Male	673.9 [1,389]	788.9 [518]	852.1 [98]	635.9 [1,389]	747.8 [518]	803.3 [98]
MATRIMONIAL STATUS						
Married	671.1 [1,016]	820.2 [346]	907.8 [76]	634.7 [1,016]	775.7 [346]	861.3 [76]
Singles	608.7 [1,045]	772.1 [401]	642.0 [83]	573.5 [1,045]	729.4 [401]	598.1 [83]
AGE AT ADMISSION						
Age less than 30	613.1 [695]	845.0 [284]	624.2 [78]	577.5 [695]	800.9 [284]	585.2 [78]
Age between 31 and 37	604.7 [910]	750.1 [319]	833.6 [64]	570.0 [910]	710.4 [319]	787.6 [64]
Age greater than 37	749.0 [456]	792.8 [144]	1190.3 [17]	710.7 [456]	741.6 [144]	1123.0 [17]
Total	639.5 [2061]	794.4 [747]	769.0 [159]	603.6 [2061]	750.9 [747]	723.9 [159]

⋆ Frequency of migrants are reported between blackets.
⋆ Earnings are in dollars values.

migrants in the first group (Grid-96) earn about \$603.6 CAD per week in their first job while in the second group they earn about \$723.9 CAD per week, which corresponds to about \$3 CAD per hour more. As expected, earnings are higher the higher the level of foreign-acquired education. Interestingly, the relation in reversed for those selected under the Grid-06. Indeed, those holding a post-secondary degree have the highest earnings while those with a Bachelor's degree have the lowest average earnings.

Moreover, migrants who hold a domestic degree have far lower earnings in their first and third jobs than those who were educated abroad. It thus seems that the acquisition of a domestic education does not allow migrants to close the income gap. We also find that migrants who are married earn more than singles. Surprisingly, while women have lower earnings than men in the first job, the opposite is true in their third job. In addition, first job earnings increase with age, but in the third job, those who have migrated at a younger age earn more than those who were older at landing.

2.2.2 Schooling Assignment Process

The data reveal that individual decide to enrol in school at different times upon arriving. Some have enrolled as early as a few months after landing. The duration of training differ across migrants. To account for these differences, we define a new calender to have comparable post-migration characteristics for all migrants regardless of the date of entry and focusing only on the elapsed time since landing. Time "zero" thus corresponds to the arrival date for each migrant. [7] On the basis of this new calendar, we construct longitudinal data by considering that one period last two years from the date of arrival. We also assume that all activities of investment in education and employment happen within the first four periods, that is eight years after the initial period, given that only those who were admitted in 2002 and 2003 are observed for all periods. In order to have significant sample size to run estimations in the third and forth period, we combine those two periods to make one, which now spans four years.

Table 2.2 – *Sample size by periods and by the date of entry in Quebec*

years of entry / Periods	2002-03	2004-05	2006-07	2008
First	274	463	507	238
Second	166	260	369	234
Third	250	438	272	

The entire education investment process may be lengthy and/or occur more than once. To insure enough observations are available in the post-education phase, we consider that the investment occurs within the first two periods. In fact, approximately 86% of individuals had completed their schooling at time of the survey. Those who were still at school after the second period have been dropped from the sample. For instance, hypothetical individual **D** in Figure 2.1 who is still in training in the third period has been dropped from the sample. In addition, we had to remove those who had not been employed and those who enter the province of Quebec in 2009. [8] In addition, about 21% of those who reported holding a job omitted to report their earnings. For those cases, we follow Carneiro *et al.* (2003) by first estimating a Mincer earnings equation separately for the non training and training groups in order to predict missing earnings. [9] One might think that those who decide to not give their earnings may belong to a selected group with particular characteristics. Table shows the characteristics

7. We nevertheless control for cohort fixed effect by introducing in the model the full set of dummies specifying the arrival year for each migrant.

8. These individuals appear in the sample only during the first period.

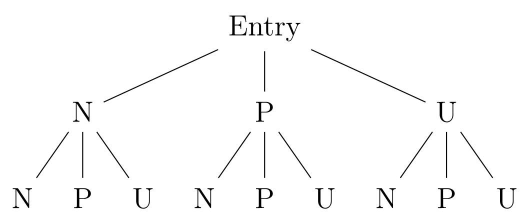
9. The Mincer's earnings equation is expressed as follow :

$$\log Earnings_{it} = \alpha_1 Age_i + \alpha_2 Age_i^2 + \beta_1 work\ experience_{it-1} + \beta_2 work\ experience_{it-1}^2 + \beta X_i,$$

where X_i contains variables such as cohort fixed effect.

of this particular group comparing to others. It appears both groups have virtually the same distribution of characteristics. On the other hand, the model we will been using control for the presence of measurement error on the outcome variable. The final sample contains 1,568 immigrants whose 893 belong to the non training group while 675 are migrants who have QC diploma.

The assignment process is as follow : initially, all individuals are in the same status, namely entry in the province of Quebec and unemployed at time zero. In the next two periods, individual can enrol in a post-secondary program (or less), P, or in a university program, U, or enter the labour market, N. The diagram below illustrates the assignment process into the different statuses upon landing.

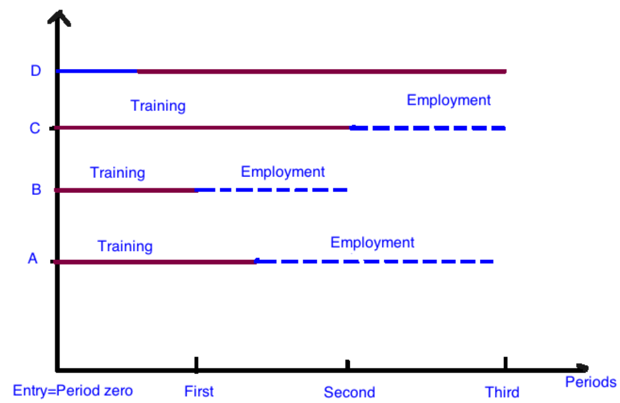


N : No QC-E ; U : university degree ; P : Post secondary or less degree

Since each period is long enough for an individual to complete his training and start working in the same period, the effect of the post-migration schooling is estimated in the second period for migrants who have actually finished their training in the first period (example of hypothetical individual **B**) or in a certain point within the second period (hypothetical individual **A**). Since the third period lasts four years and that it is more likely that someone has more than one job, we average the earnings within this period.

To sum up, our study covers three periods where the first two periods extends over two years and the third period extends over four years. Moreover, training takes place in the first two periods and we measure the causal effect of the training on earnings in the second and third periods.

FIGURE 2.1 – *Hypothetical individuals*



We also assume that an individual can not transit from a post-secondary degree to university degree or *vice versa*.[10] Thus the only feasible sequences of schooling investment are therefore :

- ⋆ UN, UU, NU, for migrants who have graduated from university degree in Quebec in the first period, in the first two periods and in the second period,respectively ;
- ⋆ PN, PP, NP for migrants who have graduated from the post secondary degree or less in the first period, in the first two periods and in the second period,respectively ; and
- ⋆ the sequence of migrants who have never enrolled in school in QC over the survey period is noted by NN. This group will be also refers as the non-training group.

Each letter of the sequence stands for training status in each period for which the investment in education is available.

2.2.3 Individual Characteristics by Schooling Status

Table 2.3 depicts the portrait of individuals in different training status sequences. The statistics, mean and percentage, are computed for each group reported in the columns. The composition of the migrants by the last continent of residence has changed from the initial sample where migrants originating from America are over represented. In fact, while in the initial sample the latter represented 7% of the sample, in our final sample they represent 18%. Likewise, the share of migrants originating from Asia has dropped from 16% to 5%. With regard to migrants from Europe (44% vs 50%) and Africa (34.5% vs 27%), they still represent the largest proportions in the sample. The continent where migrants received their highest degree is distributed in the same way as the continent of last residence.

In the economics of migration, empirical evidence suggests that migrants originating from less developed countries receive lower returns from their education. The fact that the propor-

10. There is only eight migrants in this situation.

tion of African migrants in the non-training group is significantly lower than the proportion of African migrants in other groups, coupled with the fact the converse holds for the European sub-sample, is consistent with this findings. Therefore, going back to school could be a strategy to offset the low recognition of the education acquired abroad. Nonetheless, the problem of endogeneity based on unobserved characteristics does not vanish as a sizeable proportion of Africans migrants do not invest in domestic education. According to Table 2.3, more individuals invest in post-secondary school than in university degree, and these are on average older and less educated at landing.

Furthermore, recent immigrants are less likely to invest in domestic education. This is perhaps because they manage to find an appropriate job with their foreign-acquired education or by the fact that Quebec employers have become more appreciative of the value of the foreign-acquired education. It must be noted that one may ask for an Educational Credential Assessment (ECA) to verify that the foreign degree, diploma, or certificate is valid and equal to a Canadian one. From the table, a higher proportion of migrants who have participated in training have asked for an ECA. The institution issuing the ECA may requires an individual to complete certain courses in order to obtain the credential.

Table 2.3 also reports the average unemployment rate at the time of landing. There do not appear to be any significant differences by school investment statuses. Moreover, for all reported sequences, the probability to get a qualified job increases across period and for each group.[11] Furthermore, the weekly earnings also increase across periods for almost all groups. Interestingly, the non-training group (NN-group) has the highest weekly earnings in all periods. As expected, migrants who have graduated with a domestic university degree have higher earnings than those with a post-secondary degree. We expect to have higher return from post-migration schooling for this group. To investigate whether individuals self-select into training in order to find a better job match, we have defined a variable equal one if training is in the same field as the pre-migration degree. Thus, in most cases, migrants enrolled in a university program are more likely than those involved in a post-secondary program to do so in the same field of education as their foreign-acquired degree. As expected, the work experience is higher for migrants who have not invested in domestic education. Finally, the table shows that many may be induced to invest in domestic education as a result of spending a long spell being unemployed.

11. A "Qualified job" is one whose skill requirements correspond to the level of education, as defined by the National Occupational Classification.

Table 2.3 – *Pre/Post-Immigration characteristics regarding training participation*

	NN	UN	UU	NU	PN	PP	NP
Percentage in each group	57.0	8.0	7.8	4.5	11.1	6.8	5.1
Sample size	893.0	126.0	123.0	71.0	174.0	106.0	80.0
Foreign characteristics							
BAC degree	44.0	62.7	61.0	50.7	45.4	50.9	48.8
Mast. and Doct. Degree	22.4	21.4	21.1	28.2	11.5	8.5	11.2
Female	31.2	34.9	33.3	29.6	25.3	35.8	31.2
If Married	54.2	42.1	41.5	39.4	53.4	57.5	56.2
Age	34.1	32.3	31.8	31.4	34.5	35.6	34.4
French Eval	13.5	13.6	14.1	13.4	13.8	13.7	13.4
English Eval	3.8	3.8	3.6	4.1	3.0	2.7	3.2
Ind. HCQ	0.5	0.4	0.5	0.4	0.3	0.2	0.3
If Foreign Exp	89.5	83.3	82.1	88.7	89.7	80.2	76.2
Last continent of residence							
Africa	20.4	29.6	23.6	26.8	40.8	54.3	43.8
Asia	5.4	5.6	3.3	7.0	4.6	1.9	7.5
Europe	59.0	32.8	43.1	45.1	37.9	28.6	36.2
Entry year							
2002-2003	30.1	24.6	24.4	23.9	27.0	21.7	27.5
2004-2005	28.0	27.8	35.0	33.8	33.3	39.6	37.5
2007-2008	33.7	34.9	34.1	32.4	32.8	34.0	25.0
If Prev. Stay	46.1	42.9	33.3	36.6	21.8	6.6	21.2
Post-migration characteristics							
=1 If have a qualify Job							
during the first two yrs	42.8	40.5	42.3	42.3	21.3	28.3	33.8
during the first four yrs	57.3	63.5	65.0	53.5	35.6	40.6	46.2
=1 if same study field	.	54.0	59.3	43.7	33.9	50.0	42.5
If Equiv-diploma	55.1	59.5	65.9	64.8	83.3	94.3	86.2
If job is under prof. order	42.6	50.0	52.8	46.5	48.3	50.0	48.8
if Loans and Grants from Gvt.	8.7	51.6	61.0	49.3	58.0	78.3	63.8
Work experience in each period(in years)							
Second period	0.2	0.1	0.1	0.3	0.1	0.2	0.2
Third period	3.7	2.2	1.8	2.1	2.8	2.1	2.4
Duration in 1^{st} Unempl. spell	0.0	0.4	0.4	0.6	0.4	0.7	0.8
Weekly earnings in 2^{nd} period	984.9	909.1	675.0	761.5	640.1	613.3	722.5
Weekly earnings in 3^{rd} period	1005.0	882.2	932.7	893.2	798.4	716.3	698.4

In this paper, our main goal is to measure the dynamic return to post-migration schooling in terms of earnings profile by making use of the flexibility allowed by the Bayesian treatment approach. We begin by presenting the Sequential Causal Model (SCM) along with the Sequential Inverse Probability Weighting (SIPW) approaches, which is our non-parametric identification strategy for this framework. The results obtained from this approach will follow. Next, the Bayesian identification strategy is discussed following by some results.

2.3. Causal effects : SCM approach

2.3.1 Sequential Causal Model

As stressed above, we consider that an individual can invest in education only during the first two periods, that is, during the first four years following their entry in Canada. Let S_{it} stand for the individual schooling decision status in period t. It is defined over the set $S = \{N, P, U\}$ for each period $t = 1, 2$ and is always equal to N in the third period. As the result, a sequence of schooling decision in the first period is an element of $S = \{N, P, U\}$ and, in the second period a sequence of schooling decisions is an element of $S = \{NN; UN; UU; NU; PN; PP; NP\}$. [12] The sequence of schooling decisions until period

12. Recall that in our data, an individual cannot drop-out of school without a diploma or re-enroll in school after a temporary spell out of training.

τ is denoted by $\underline{s}_\tau$, $\tau \in \{0,1,2\}$ in keeping Lechner & Miquel (2001)'s notation. To illustrate, $\underline{s}_2 = (S_1, S_2) = NN$.[13]

In the Sequential Causal Model (SCM), the identification strategy relies on the potential outcomes framework. Let $Y_{iu}^{\underline{s}_\tau}$ be the potential earnings of individual i in period u with the sequence of schooling decision $\underline{s}_\tau$ ($\tau \leqslant u$). For each individual, we observe only one realization of the earnings associated with the actual sequence of decisions, denoted Y. The relationship between Y and the potential earnings for each period can be written as follow[14] :

$$\begin{aligned} Y = \;& \mathbb{1}_{(S_1=N)}\mathbb{1}_{(S_2=N)} \times Y^{NN} + \mathbb{1}_{(S_1=U)}\mathbb{1}_{(S_2=N)} \times Y^{UN} + \mathbb{1}_{(S_1=U)}\mathbb{1}_{(S_2=U)} \times Y^{UU} + \mathbb{1}_{(S_1=N)}\mathbb{1}_{(S_2=U)} * Y^{NU} \\ & + \mathbb{1}_{(S_1=P)}\mathbb{1}_{(S_2=N)} \times Y^{PN} + \mathbb{1}_{(S_1=P)}\mathbb{1}_{(S_2=P)} \times Y^{PP} + \mathbb{1}_{(S_1=N)}\mathbb{1}_{(S_2=P)} * Y^{NP} \end{aligned} \tag{2.1}$$

The underlying assumption that sustains this equation is the Stable Unit Value Treatment Assumption (SUVTA), which states that the potential earnings of an individual is independent of the education decision of another persons. The SUTVA assumption is quite close to the *consistency assumption*, which states that the observed earnings is the potential earnings corresponding to the actual sequence of school decisions assigned to an individual (Abbring & Van Berg, 2003)[15].

Dynamic average gain of education investment

Let $\theta_u^{\underline{s}_\tau^k;\underline{s}_\tau^l}$ be the dynamic average gain (DAG) of taking a sequence of school decision $\underline{s}_\tau^k$ over the sequence $\underline{s}_\tau^l$. As in the static setting, we can define the dynamic average gain of QC-education conditional on the probability of belonging to the group of individuals with the sequence of school decision $\underline{s}_\tau^k$ noted by $\theta_u^{\underline{s}_\tau^k;\underline{s}_\tau^l}(\underline{s}_\tau^k)$. This is commonly called the average effect on the treated in the treatment analysis since the targeted or reference group (specified in the bracket) is identical to the group of people for whom we want to measure the effect. More generally $\theta_u^{\underline{s}_\tau^k;\underline{s}_\tau^l}(\underline{s}_a^j)$ is the average gain of the sequence of school decision $\underline{s}_\tau^k$ over the sequence $\underline{s}_\tau^l$ conditional on the probability of belonging to the group of people having the sequence $\underline{s}_a^j$. The expression of the DAG can be written as follows,

Let $\theta_u^{\underline{s}_\tau^k;\underline{s}_\tau^l}$ be the dynamic average gain (DAG) of taking a sequence of schooling decisions $\underline{s}_\tau^k$ over the sequence $\underline{s}_\tau^l$. As in the static setting, we can define the dynamic average gain of QC-education conditional on the probability of belonging to the group of individuals with the sequence of schooling decision $\underline{s}_\tau^k$ noted by $\theta_u^{\underline{s}_\tau^k;\underline{s}_\tau^l}(\underline{s}_\tau^k)$. This is commonly referred to as the "average treatment effect on the treated" in the treatment literature since the targeted or reference group (specified in the bracket) is identical to the group of people for whom we want to measure the effect. More generally, $\theta_u^{\underline{s}_\tau^k;\underline{s}_\tau^l}(\underline{s}_\tau^j)$, is the average gain of the sequence of school

13. Note that the initial period is omitted for the sake of simplicity and because all individuals have the same status. Also, in what follows we will omit the individual subscript, i, except when necessary.

14. $\mathbb{1}(.)$ is the indicator function that takes one if the condition inside the bracket holds

15. This assumption can't be tested with non-randomized data but holds in the experimental data. Then we maintain this assumption as in most previous study using survey data.

decision $\underline{s}_\tau^k$ over the sequence $\underline{s}_\tau^l$ conditional on the probability of belonging to the group of people having the sequence $\underline{s}_a^j$. The expression of the DAG can be written as follows,

$$\theta_u^{\underline{s}_\tau^k;\underline{s}_\tau^l} = E(Y_{iu}^{\underline{s}_\tau^k} - Y_{iu}^{\underline{s}_\tau^l}) = E(Y_{iu}^{\underline{s}_\tau^k}) - E(Y_{iu}^{\underline{s}_\tau^l}), \quad u = 2, 3 \tag{2.2}$$

In our non-randomized framework, we observe only one sequence of education decisions for each individual i. Additional restrictive assumptions must be invoked to identify the DAG as defined in equation (2.2) and its variants because for an given individual with sequence $\underline{s}_\tau^k$, we do not observe $E(Y_{iu}^{\underline{s}_\tau^l})$ since $\underline{s}_\tau^l \neq \underline{s}_\tau^k$.

Identification Strategy

Let X_0 stand for individual characteristics at landing, and X_t be a vector of time-varying variables containing all labour market outcomes such as the type of job, work experience, *etc.*; $\underline{X}_\tau$ ($\tau \geqslant 1$) is a vector of time-varying variables until the period τ. An individual decides to enrol in school at the beginning of each period. As in Lechner (2009), the main assumption to identify pairwise sequences of average gains of the form $\theta_u^{\underline{s}_\tau^k;\underline{s}_\tau^l}$ and $\theta_u^{\underline{s}_\tau^k;\underline{s}_\tau^l}\left(S_1^j\right)$ may be written as follow :

Assumption 1. *Weak Dynamic Conditional Independence Assumption (W-DCIA) :Lechner (2009)*

(*a*) $Y_u^{\underline{S}_2^k} \perp\!\!\!\perp S_1 \mid X_0 = x_0$, *for* $t = 1$, $\forall u \in \{2, 3\}$ *and* $\forall \underline{S}_2^k$

(*b*) $Y_u^{\underline{S}_2^k} \perp\!\!\!\perp S_2 \mid \underline{X}_1 = \underline{x}_1$, $S_1 = s_1$, *for* $t = 2$, $\forall u \in \{2, 3\}$ *and* $\forall \underline{S}_2^k$

(*c*) $0 < P(S_1 = 1 \mid X_0 = x_0) < 1$, *in the first period.*

(*d*) $0 < P(S_2 = 1 \mid S_1 = s_1, \underline{X}_1 = \underline{x}_1) < 1$, *in the second period. Where the notation* $A \perp\!\!\!\perp B \mid C$ *means that* A *is conditionally independent of* B *given* C.

Assumptions 1.(*a*) and 1.(*b*) state that conditional on exogenous co-variates in each period, schooling status is independent of the potential outcome. To paraphrase Lechner (2009), *The question that we have to ask it is if all variables that could caused changes in outcome of interest and training status are all observed.* In this specification, we assume that we actually observed all those potential confounders. Nevertheless, it might exist some unobserved factors such as social network developed by migrants during his training which can influence type of job offer available to them. [16] Another way to test the validity of these assumptions, it is to test for the existence of a *"dip"* in the earnings trajectory prior to school enrolment. On the other hand, since we estimate the gain of training after all combinations of school participation, we are confident about the validity of the independence assumption based on the observed co-variates but less based on unobserved heterogeneity.

16. We will discuss more about that in the next section.

Furthermore, A1.(c) and A1.(d) ensure that the probabilities associated with all potential sequences are distributed on the same support in each period. In addition, all individuals must have a positive probability to enrol or not in school. With these two assumptions, we can compare individuals who are *"equally"* likely to choose a particular sequence of schooling decision, conditional on their person-specific characteristics.

We choose the SIPW approach to identify the dynamic average gain of education because of its theoretical intuition to resolve the selection problem, and because the causal effect is identified in the non-parametric way. For example, if we want to measure the average gains of taking the sequence of schooling decision $\underline{s}_2^k$ over the sequence $\underline{s}_2^j$, where $\underline{s}_1^k \neq \underline{s}_1^j$, and by taking $\underline{s}_1^k$ as the *"reference group"*, the conditional DAG using the SIPW approach is identified as follow :

$$\begin{aligned} \theta^{\underline{s}_2^k;\underline{s}_2^j}\left(\underline{s}_1^k\right) &= E\left(Y^{\underline{s}_2^k} - Y^{\underline{s}_2^j} \mid \underline{S}_1^k = \underline{s}_1^k\right) \\ &= E\left(Y^{\underline{s}_2^k} \mid \underline{S}_1 = \underline{s}_1^k\right) - E\left(Y^{\underline{s}_2^j} \mid \underline{S}_1 = \underline{s}_1^k\right) \end{aligned}$$

The identification of the first term on the right hand side of the last equality is straightforward while the second term can be estimated as follows : [17]

$$\begin{aligned} E(Y^{\underline{s}_2^j} \mid S_1 = s_1^k) &= E_{\underline{X}_0 \mid S_1 = s_1^k}\left\{E\left(Y^{\underline{s}_2^j} \mid S_1 = s_1^k, \underline{X}_0 = \underline{x}_0\right)\right\} \\ &= E_{X_0}\left\{E\left[Y^{\underline{s}_2^j} \mid S_1 = s_1^k, \underline{X}_0 = \underline{x}_0\right]\frac{P(S_1 = s_1^k \mid \underline{X}_0 = x_0)}{P(\underline{S}_1 = \underline{s}_1^k)}\right\} \\ &\overset{A1.b}{=} E_{X_0}\left\{E\left[Y^{\underline{s}_2^j} \mid \underline{S}_1 = \underline{s}_1^j, \underline{X}_0 = \underline{x}_0\right]\frac{P(S_1 = s_1^k \mid \underline{X}_0 = x_0)}{P(\underline{S}_1 = \underline{s}_1^k)}\right\} \\ &\,\vdots\, = E\left(Y^{\underline{s}_2^j}\mathbb{1}_{\{S_2 = \underline{s}_2^j\}}\mathbb{1}_{\{S_1 = s_1^j\}} \times \frac{P(S_1 = s_1^k \mid X_0 = x_0)}{P(S_2^j \mid s_1^j, \underline{X}_1 = \underline{x}_1)P(S_1 = s_1^j \mid X_0 = x_0)}\right) \end{aligned}$$

An empirical estimation of $\theta^{\underline{s}_2^k;\underline{s}_2^j}\left(\underline{s}_1^k\right)$ can be given by the weighted average effect :

$$\tilde{\theta}^{\underline{s}_2^k;\underline{s}_2^j}\left(\underline{s}_1^k\right) = \frac{1}{\sum_{i \in \underline{S}_2^k} W^{\underline{s}_2^k;\underline{s}_1^k}}\sum_{i \in \underline{S}_2^k} W^{\underline{s}_2^k;\underline{s}_1^k} \times Y_i - \frac{1}{\sum_{i \in \underline{S}_2^j} W^{\underline{s}_2^j;\underline{s}_1^k}}\sum_{i \in \underline{S}_2^j} W^{\underline{s}_2^j;\underline{s}_1^k} \times Y_i,$$

With, $W^{\underline{s}_2^k;\underline{s}_1^k} = \frac{1}{P(S_2 = s_2^k \mid S_1 = s_1^k, \underline{X}_1)}$ and $W^{\underline{s}_2^j;\underline{s}_1^k} = \frac{P(S_1 = s_1^k \mid \underline{X}_0)}{P(S_2^j \mid \underline{s}_1^j, \underline{X}_1)P(S_1 = s_1^j \mid X_0)}$. [18] Table 2.4 reports different weights involved in the computation of various average gains of training :

TABLE 2.4 – *Average treatment effects*

Average Treatment Effects (1)	Different Weights (2)	Possible comparisons (3)
$\tilde{\theta}^{\underline{s}_2^k;\underline{s}_2^j}$	$W^{\underline{s}_2^k} = \frac{1}{P(S_2^k \mid s_1^k, \underline{X}_1)P(S_1^k \mid X_0)}$	$\tilde{\theta}^{PP;NN}$, $\tilde{\theta}^{PN;NN}$, $\tilde{\theta}^{NP;NN}$
	$W^{\underline{s}_2^j} = \frac{1}{P(S_2^j \mid s_1^j, \underline{X}_1)P(S_1^j \mid X_0)}$	$\tilde{\theta}^{UU;NN}$, $\tilde{\theta}^{UN;NN}$, $\tilde{\theta}^{NU;NN}$
$\tilde{\theta}^{\underline{s}_2^k;\underline{s}_2^j}(s_1^l)$	$W^{\underline{s}_2^k;s_1^l} = \frac{P(S_1^l \mid X_0)}{P(S_2^k \mid \underline{s}_1^k, \underline{X}_1)P(S_1^k \mid X_0)}$	$\tilde{\theta}^{\underline{s}_2^k;NN}(s_1)\ \forall s_1 = \{N, P, U\}$
	$W^{\underline{s}_2^j;s_l^1} = \frac{P(S_1^l \mid X_0)}{P(S_2^j \mid \underline{s}_1^j, \underline{X}_1)P(S_1^j \mid X_0)}$	and $\forall \underline{s}_k^2 = \{PN, NP, PP, NU, UN, UU\}$

17. See Lechner (2009) for the details.

18. The second sequence of treatments represents the *"reference group"*.

DAGs can also be identified conditional on belonging to a specific reference group characterized by a one-period of decision. This is the average gain of the form $\tilde{\theta}^{\underline{s}_2^k;\underline{s}_2^j}(s_1^l)$. In the two-period decision case, additional restrictions must be met to identify the conditional DAG. Indeed, the decision to enrol in school in the second period is linked to the decision took in the first period, which in turn affects some individual characteristics such as work experience. Therefore, schooling participation is not randomly assigned in the second period. The identification of such an average effect requires further restrictive assumptions known as the *strong dynamic conditional independence assumption(S-DCIA)* (see Lechner & Miquel (2001)). Based on the data available for this study, it is more likely that the *W-SCIA* holds but not the *S-SCIA*. We will therefore focus on the average effects as defined in column (1) in Table 2.4.

FIGURE 2.2 – *Probability of taking or not taking a training*

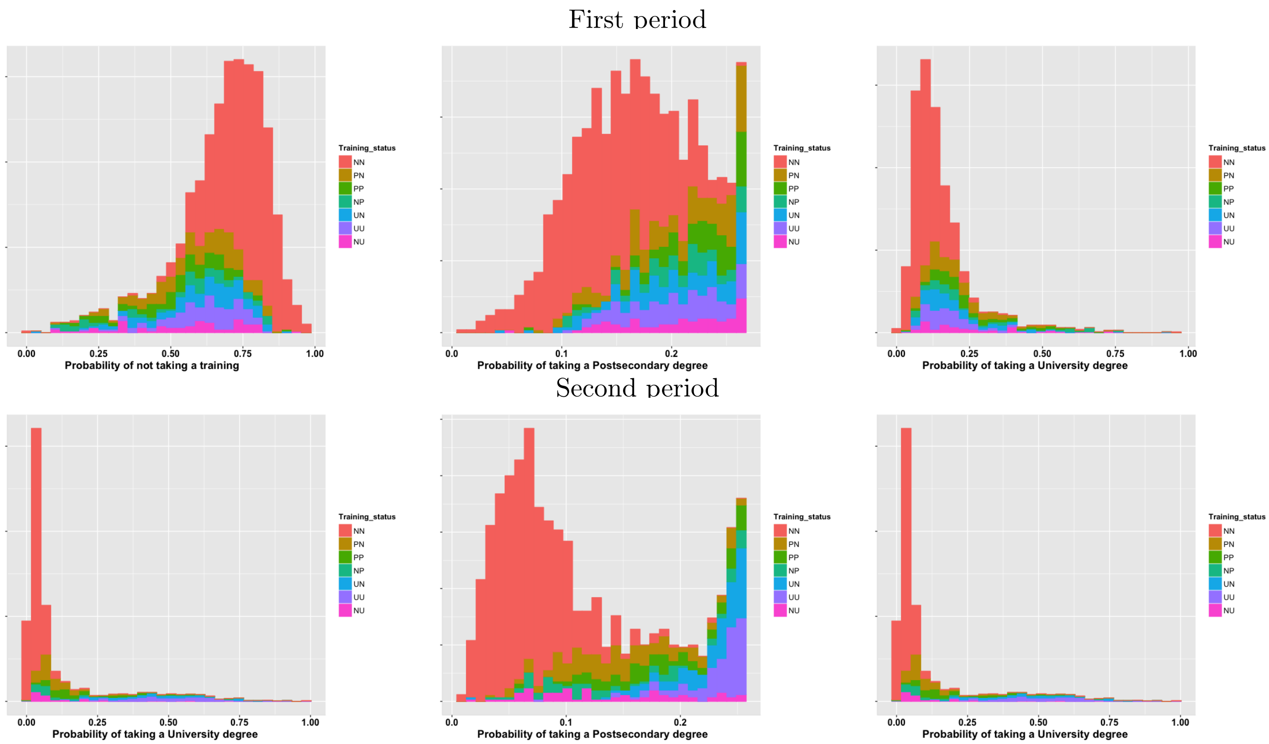


Figure 2.2 shows that whatever the training status considered, there is a common support for the probability of all potential sequences of domestic education in the first period (Variables used to compute propensity score in each period are reported in table B.2). This is not so in the second period. Therefore, we will only compute the average effect by comparing each individual in the training group to the counterfactual situation when he belongs to the group-NN(see third column of Table 2.4).

2.3.2 Results : SCM

In the SIPW framework, the literature is relatively silent about the issue of matching in the common support. The discussions focus entirely on tail probabilities. To deal with this issue, most empirical studies suggest to trim the distribution of the weights used to compute the average effect. We have chosen to remove successively 0.1% , 1% and 5% of the largest weights in each group involved in the computation of the schooling effects. We limit our study to a comparison of migrants with a domestic education to those with a foreign-education only.[19] According to Figure 2.2 above, the existence of a common support would not be an issue for our purposes. The figure also shows that the probability of investing in a domestic education is relatively small for all individuals, regardless of the actual group of training they belong to.

To estimate the different weights, we have estimated in each period a binary probability using Probit model in the corresponding subsample (See table B.1 in appendix for all those subsamples). Table 2.5 presents the set of dynamic average gains of the domestic education over a foreign education, divided into four blocs according to the percentage of largest weights removed from the computation. The standard deviations of the estimators are obtained by bootstrap. The two lines associated to each comparison group stand respectively for the effect in the second and third periods.

The issue arises when it comes to trim the distribution of weights it is to what extent the targeted population has been changed after removing a part of the sample. In table B.3, we have reported migrants' characteristics depending on the percentage of trimming. We can see that the composition of each group doesn't change when the percentage of trimming goes from 0.1% to 1%, however, it changes slightly when it goes from 1% to 5%. This comfort us about the fact that our targeted samples are not significantly affected when we trim the distribution of weights.

The results reveal that the average effect don't change significantly in most cases when the percentage of trimming increases. However, standard error tends to decrease when the percentage of largest weights removed increases. This comforts us about the robustness of our results. Crump *et al.* (2009) find empirically that the optimal percentage of trimming is 0.1%. For this study, based on standard error and how the composition of targeted samples change, the optimal percentage of trimming seems to be 1%. We are then going to focus our analysis basedon results when 1% of the largest weights are removed for each group.

19. We have purposely omitted to compare those with a domestic education between themselves.

Table 2.5 – *Average gain of QC-education over Foreign-education using SIPW*

	$\tilde{\theta}$	se	$\tilde{\theta}(N)$	se	$\tilde{\theta}(P\,or\,U)$	se	$\tilde{\theta}$	se	$\tilde{\theta}(N)$	se	$\tilde{\theta}(P\,or\,U)$	se
	No observation removed						0.1% of largest weights removed					
$PN-NN$	-0.25	0.15	-0.24	0.10	-0.21	0.10	-0.25	0.15	-0.24	0.11	-0.21	0.10
	-0.06	0.06	-0.05	0.08	-0.05	0.04	-0.06	0.06	-0.05	0.08	-0.05	0.04
$NP-NN$	0.25	0.19	0.23	0.01	0.22	0.38	0.25	0.19	0.23	0.01	0.22	0.38
	-0.02	0.10	-0.02	0.01	-0.03	0.19	-0.02	0.10	-0.02	0.01	-0.03	0.19
$PP-NN$	-0.12	0.10	-0.12	0.09	-0.08	0.14	-0.12	0.10	-0.12	0.09	-0.08	0.14
	-0.05	0.06	-0.05	0.06	-0.02	0.08	-0.05	0.06	-0.05	0.06	-0.02	0.09
$UN-NN$	0.05	0.12	0.06	0.11	0.11	0.09	0.05	0.12	0.06	0.11	0.11	0.09
	0.06	0.10	0.07	0.09	0.11	0.09	0.06	0.10	0.07	0.10	0.11	0.09
$NU-NN$	0.05	0.14	0.06	0.02	0.05	0.25	0.05	0.14	0.06	0.01	0.05	0.25
	-0.16	0.08	-0.17	0.01	-0.12	0.16	-0.16	0.08	-0.17	0.01	-0.12	0.16
$UU-NN$	0.20	0.12	0.20	0.08	0.20	0.10	0.20	0.12	0.20	0.08	0.20	0.10
	0.22	0.08	0.22	0.06	0.23	0.08	0.22	0.08	0.22	0.06	0.23	0.08
	1% of the largest weights is removed						5% of the largest weights is removed					
$PN-NN$	-0.25	0.09	-0.24	0.09	-0.23	0.03	-0.23	0.06	-0.23	0.07	-0.22	0.04
	-0.05	0.05	-0.05	0.06	-0.05	0.03	-0.04	0.04	-0.05	0.05	-0.03	0.04
$NP-NN$	0.24	0.16	0.23	0.02	0.22	0.30	0.24	0.22	0.30	0.11	0.19	0.32
	-0.02	0.08	-0.02	0.01	-0.03	0.14	-0.00	0.07	0.03	0.05	0.00	0.13
$PP-NN$	-0.12	0.06	-0.12	0.09	-0.10	0.04	-0.11	0.06	-0.10	0.09	-0.13	0.05
	-0.05	0.04	-0.05	0.05	-0.03	0.03	-0.03	0.04	-0.02	0.06	-0.01	0.05
$UN-NN$	0.05	0.09	0.05	0.11	0.11	0.04	0.03	0.09	0.02	0.11	0.05	0.08
	0.07	0.08	0.07	0.09	0.11	0.04	0.05	0.07	0.03	0.09	-0.01	0.07
$NU-NN$	0.05	0.11	0.05	0.04	0.04	0.18	-0.01	0.11	-0.13	0.08	0.04	0.16
	-0.15	0.06	-0.17	0.01	-0.12	0.11	-0.06	0.06	-0.05	0.03	-0.05	0.10
$UU-NN$	0.20	0.07	0.20	0.07	0.20	0.02	0.23	0.06	0.24	0.07	0.22	0.03
	0.23	0.06	0.23	0.06	0.23	0.03	0.26	0.05	0.27	0.06	0.27	0.03

⋆ $\tilde{\theta}(N)$ and $\tilde{\theta}(P)$ is respectively the average effect conditional on not being treated in the first period and conditional to take a PSS in the first period.
⋆ se is the standard error.

Even though not always statistically significant, the average effect of a domestic PSS degree decreases over periods and is positive or negative depending on the timing and the duration of the training. On the other hand, the average effect is positive in the first period and negative in the second period for those who took-up a university degree only for one period, while it is positive in all periods for those who have taken-up a university degree for the two first periods. Moreover, conditional of being treated in the first period, the average gain of a domestic education over a foreign one ($\theta(P)$ or $\theta(U)$), is greater than the simple average effect θ and the one conditional on not being treated in the first period ($\theta(N)$).

The main result to keep in mind for this identification strategy are those pertaining to migrants who have been enrolled in a university degree for the first four years upon landing for whom the gain of QC-education over the foreign education is positive.

2.4. Causal effects : Bayesian Approach

The treatment effect analysis through a Bayesian approach is still in its infancy. Indeed, few empirical studies have used this approach in a dynamic framework, and on migration issues in particular. One of the interesting contributions of this approach is that it allows to take into account the correlation between the investment in education and the earnings profile due to the unobserved variables and importantly, regardless of the complexity of the likelihood function. For instance, one might think that high-ability individuals with a weak social network could have strong incentives to invest in domestic education and also to have a greater return than low-ability individuals. Evidently, individual ability is not observable. However, the Bayesian approach offers novel workarounds to this problem. For instance, we can simulate a distribution of the unobserved factors given the data, conditional on certain assumptions about their prior distributions (and their hyper-parameters). This section discusses the specification of the likelihood function and derives the posterior distribution of all the model parameters.

2.4.1 Empirical strategy

Likelihood function

To model the schooling decision in each period, we specify an ordered binary model related to the latent utility of the domestic education. We assume that the latent utility associated with a university degree is higher than that of a post-secondary degree, which is itself greater than that of not enrolling in domestic education. We can interpret the latent variable as the additional value individuals place on domestic education above their foreign-acquired one.[20] Considering the foregoing, the likelihood function is defined as follows :

$$\begin{cases} & \textit{School Decision Equations} \\ U_{it}^{*} = & \alpha \underline{x}_{it-1} + \psi_0 \theta_i + \epsilon_i, \forall t \in \{1,2\}, \\ S_{it} = N = 0 & if \quad d_{-1} < U_{it}^{*} \leqslant d_0 \\ S_{it} = P = 1 & if \quad d_0 < U_{it}^{*} \leqslant d_1 \\ S_{it} = U = 2 & if \quad d_1 < U_{it}^{*} \leqslant d_2 \\ & \quad \textit{Outcome Equations} \\ Ln(Y_{iu}^{\bar{s}}) = & \beta_{\bar{s}} \underline{x}_{iu-1} + \psi_{\bar{s}u} \theta_i + \lambda_{iu}^{-1/2} \xi_{\bar{s}}, \quad \forall u \in \{2,3\}. \end{cases} \tag{2.3}$$

U_{it}^{*} is the latent utility of choosing to enroll in school in period t for the individual i. $\underline{x}_{t-1}$ and $\underline{x}_{u-1}$ are respectively the vector of variables up to the period $t-1$ and $u-1$. For instance, $\underline{x}_1 = (x_0, x_1)$ where x_0 is mainly pre-migration characteristics as cohort fixed effects and, x_1 is a vector of covariates in the first period, that is post-migration variables such as the type of job, work experience, work experience (squared), *etc.* $\bar{s} = \left\{\bar{N}, \bar{P}, \bar{U}\right\}$ stands for the

20. Recall that approximately 80% of individuals have at least a bachelor's degree at landing.

three groups of migrants characterized by the level of domestic education, regardless of the timing when the enrolment occurs and the duration of the latter. Indeed, according to the sequences of training, each group is distributed as follows : $\bar{N} = \{NN\}$, $\bar{P} = \{PP, PN, NP\}$ and $\bar{U} = \{UU, UN, NU\}$. In the specification, we assume that the covariates affect earnings in a similar fashion in each group. We also assume that, conditional on observed variables and unobserved factors (θ_i), the latent utility for education is independent of the earnings profile. Conversely, the correlation between the schooling investment equation and the earnings equation is driven entirely by the distribution of the unobserved factors and the magnitude of the loading factors $(\psi_0, \psi_{\bar{N}}, \psi_{\bar{P}}, \psi_{\bar{U}})$. It follows that $Cov(U_t^*, Y_u^{\bar{s}}) = \psi_0 Var(\theta)\psi_{\bar{s}}$ for all periods and $\xi_{u\bar{s}} \perp\!\!\!\perp \xi_{u\bar{s}'}$, $\bar{s} \neq \bar{s}'$. Notice that d_{-1}, d_0, d_1, d_2 represent the so-called "cutoff" points and $(\lambda_{iu})_{u\in\{2,3\}}$ represent individual scale factors for the variance. In the earnings equation, λ_{iu} enables to correct for potential heteroscedasticity due to measurement error or the fact each individual has a probability to get a job (or to be included in the sample) in each period. We can rewrite equation (2.3) as :

$$Z_{i,u} = \begin{pmatrix} U_{i1}^* \\ U_{i2}^* \\ Ln(Y_{iu}^{\bar{N}}) \\ Ln(Y_{iu}^{\bar{P}}) \\ Ln(Y_{iu}^{\bar{U}}) \end{pmatrix} \Bigg| \begin{matrix} x_{i0}, x_{i1} \\ x_{iu-1}, \theta_i \end{matrix} \sim \mathcal{N}\left(\begin{bmatrix} a + \alpha x_{i0} + \psi_0\theta_i \\ a + \alpha x_{i0} + \alpha_1 x_{i1} + \psi_0\theta_i \\ a_{\bar{N}} + \beta_{\bar{N}}\underline{x}_{iu-1} + \psi_{\bar{N}}\theta_i \\ a_{\bar{P}} + \beta_{\bar{P}}\underline{x}_{iu-1} + \psi_{\bar{P}}\theta_i \\ a_{\bar{U}} + \beta_{\bar{U}}\underline{x}_{iu-1} + \psi_{\bar{U}}\theta_i \end{bmatrix}, \overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{iu}^{-1}\sigma_{\bar{N}} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{iu}^{-1}\sigma_{\bar{P}} & 0 \\ 0 & 0 & 0 & 0 & \lambda_{iu}^{-1}\sigma_{\bar{U}} \end{bmatrix}}^{\Sigma_{iu}} \right)$$

or more compactly as,

$$Z_{i,u}|\theta_i, X_i \sim \mathcal{N}\left(X_i\beta + \psi\theta_i, \Sigma_{iu}\right). \tag{2.4}$$

$X_i = diag(x_{i0}, x_{i1}, x_{iu-1})$ is the matrix of all covariates in equation (2.3) and $\beta = (a, a_{\bar{N}}, a_{\bar{P}}, a_{\bar{U}}, \alpha, \alpha_1, \beta_{\bar{N}}, \beta_{\bar{P}}, \beta_{\bar{U}})$ is a vector of parameters excluding the vector of cut-off points, $\psi = (\psi_0, \psi_{\bar{N}}, \psi_{\bar{P}}, \psi_{\bar{U}})$ is the vector of loading factors and $\sigma = (\sigma_{\bar{N}}, \sigma_{\bar{P}}, \sigma_{\bar{U}})$ the vector of the variances of the error terms. Given our specification, the independence assumption can be written as follows :

Assumption 2.

$$Y_u^{\underline{s}_2^k} \perp\!\!\!\perp S_1 \mid x_0, \theta, \forall\, \underline{s}_2^k \tag{2.5}$$

This assumption is more realistic than the one specified in the SCM because it assumes that the unobserved characteristics may affect both the individual choices and the earnings profile. The probability to enrol in a post-secondary degree in the first period is $P(S_{i1} = 1) = P(d_{-1} < U_{i1}^* \leqslant d_0) = \Phi(d_0 - a - \alpha x_{i0} - \psi_0\theta_i) - \Phi(d_{-1} - a - \alpha x_{i0} - \psi_0\theta_i)$
More generally, the probability that the school status in period $t = 1, 2$ takes a value $j \in$

$\{0, 1, 2\}$ is given by [21],

$$\begin{aligned} P(S_{it} = j) &= P(d_{j-1} < U^*_{it} \leqslant d_j) \\ &= \Phi(d_j - a - \alpha \underline{x}_{it-1} - \psi_0 \theta_i) - \Phi(d_{j-1} - a - \alpha \underline{x}_{it-1} - \psi_0 \theta_i). \end{aligned}$$

Without loss of generality, we set $d_{-1} = -\infty$, $d_0 = 0$ and $d_2 = \infty$ so that the latent utility is negative for those who do not attend school in Quebec, akin to a *disutility* for school. At the end, there is only one unknown cutoff point, d_1. The likelihood function for each individual is specified as below :

$$\log\, L_i(Y_i, U^*_{i1}, U^*_{i2} | \beta, \psi, d_1, \sigma, \theta_i) = \sum_{t=1}^{2} \sum_{j=0}^{2} m^t_{ij} log \left[\Phi(d_j - \Upsilon_{it}) - \Phi(d_{j-1} - \Upsilon_{it})\right] + \sum_{u=2}^{3} log\, f\left(log(Y_{iu}) | \beta, \psi, \sigma, \theta_i\right) \tag{2.6}$$

and the total likelihood function as,

$$\log\, L(Y, U^*_1, U^*_2 | \beta, \psi, d_1, \sigma, \theta) = \sum_{i=1}^{N} log\, L_i(Y_i, U^*_{i1}, U^*_{i2} | \beta, \psi, d_1, \sigma, \theta_i), \tag{2.7}$$

where, $m^t_{ij} = 1$ if $S_{it} = j$, $\Upsilon^k_{it} = a + \alpha \underline{x}_{it-1} + \psi_t \theta_i + \alpha_t x_{it-1}$ and the function $f(\cdot)$ is the density function for a normal distribution. To identify the loading factors, we set $\psi_{\bar{N}} = 1$.

Prior and posterior analyses

As mentioned above, our goal is to estimate the joint distribution of the parameters through a Markov Chain Monte Carlo (MCMC) approach. To implement the MCMC simulation, we have chosen as prior distribution for the set of parameters a multivariate normal distribution and for the variance of the error term σ, a multivariate normal distribution restricted to the support $]0, +\infty]$. For the cutoff point, we have chosen a uniform distribution.

We note the prior distribution by $\pi(\cdot)$ and :
$\pi\left(b\right) \sim \mathcal{N}_\kappa\left(B_0, \Gamma_0\right), b = (\beta, \psi)$, $\pi\left(\theta\right) \sim \mathcal{N}\left(0, \vartheta_0\right)$, $\pi(d_1) = a_0, a_0 > 0$, $\pi\left(\lambda_{iu} | \nu_0\right) \sim \mathcal{G}\left(\nu_0/2, \nu_0/2\right)$ and $\pi(\sigma) \sim \mathcal{N}_\iota\left(\nu_0, \Omega_0\right) \mathbb{1}(\sigma \in \Lambda_\iota)$, $\forall u = 2, 3$, where $\mathcal{G}$ and $\mathcal{N}$ are respectively a gamma and a multivariate normal distributions and, κ and ι stand respectively for the dimension of the vectors of parameters b and σ and, Λ_ι stands for the positive orthant of dimension ι.

Finally, the posterior density function is given by :

$$\begin{aligned} \pi'\left(b, d_1, \sigma, (\theta)^N_{i=1}, (\lambda_{i2})^N_{i=1}, (\lambda_{i3})^N_{i=1}\right) \propto\ & \pi\left(b | B_0, \Gamma_0\right) \pi(\sigma | \nu_0, \Omega_0) \pi(d_1 | a_0) \pi\left(\theta | \vartheta_0\right) \\ & \times \pi\left(\lambda_2 | \nu_0\right) \pi\left(\lambda_3 | \nu_0\right) \\ & \times \log\, L(Y_2, Y_3, U^*_1, U^*_2 | b, d_1, \sigma, \theta, \lambda_1, \lambda_2, X), \end{aligned} \tag{2.8}$$

We follow Chib & Hamilton (2000), Albert & Chib (1993), Chib & Hamilton (2002) and sample from the posterior distribution of b, θ and λ from an augmented sample that

21. $\Phi(\cdot)$ stands for the normal cumulative function since we assume that ϵ is drawn from a normal distribution.

includes the distribution of the counterfactuals earnings.[22] With regard to σ we take the advice of Chib & Hamilton (2002) by using a Metropolis-Hastings algorithm to sample first the posterior distribution of the vector σ conditional on the data and prior knowledge about the parameters.[23] The detail of the algorithm of the simulation process is described in the appendix. This process significantly reduces the auto-correlation across the iterations and allows to achieve convergence faster[24].

Algorithm

1. Initialize b, d_1,σ, θ, $(\lambda_{i2})_{i=1}^{N}$,$\lambda_{i3})_{i=1}^{N}$
2. Sample $\sigma|b, d_1, (\theta)_{i=1}^{N}, (\lambda_{i2})_{i=1}^{N}, (\lambda_{i3})_{i=1}^{N}$. The posterior distribution is given by :
$h\left(\sigma|\nu_0, \Omega_0, b, (\theta)_{i=1}^{N}, (\lambda_{i2})_{i=1}^{N}, (\lambda_{i2})_{i=1}^{N}\right) \propto f(\sigma|\nu_0, \Omega_0) \times \log\ L(Y_2, Y_3, U_1^*, U_2^*|b, d_1, \sigma, \theta, \lambda_1, \lambda_2, X)$.
To sample σ from this distribution, the proposal density function is a multivariate-t student $q(\mu, V)$, where μ and V are respectively the mode and the inverse of the negative of the Hessian matrix evaluated at the mode of the function $h(.|.)$.
3. Sample (U_{i1}^*, U_{i2}^*) and the unobserved component of the vectors $Z_{iu}^* = \left(Y_{iu}^{*\bar{N}}, Y_{iu}^{*\bar{P}}, Y_{iu}^{*\bar{U}}\right)$ conditional on $b, d_1, (\theta)_{i=1}^{N}$,$(\lambda_{i2})_{i=1}^{N}$,$(\lambda_{i3})_{i=1}^{N}$, σ and conditional on the data.
   - if $\underline{s}_{2i}^{k} = NP$ then sample first $U_{i1}^*|b, d_1, (\theta)_{i=1}^{N}$,$(\lambda_{i2})_{i=1}^{N}$,$(\lambda_{i3})_{i=1}^{N}$ from a normal distribution truncated to the interval $]-\infty, 0]$ and then sample $U_{i2}^*|b, d_1, (\theta)_{i=1}^{N}$,$(\lambda_{i2})_{i=1}^{N}$,$(\lambda_{i3})_{i=1}^{N}$ from a normal distribution truncated to the interval $]0, d_1]$ and, we do the same for other sequence of school decision.
   - $\forall\, u \in \{2, 3\}$, $i = 1, \cdots, n$, sample either $Y_{iu}^{*\bar{N}}, Y_{iu}^{*\bar{P}}, Y_{iu}^{*\bar{U}}$, independently from i and from u, from a normal distribution depending on whether $\underline{s}_{2i}^{k} \in \{\bar{N}, \bar{P}, \bar{U}\}$.
4. $d_1 \sim\sim \mathcal{U}\left\{max\left[\left(U_{i1}^* + U_{i2}^*\right)/2|i \in \bar{P}\right], min\left[\left(U_{i1}^* + U_{i2}^*\right)/2|i \in \bar{U}\right]\right\}$.
5. Sample $\beta \sim \mathcal{N}(B, \Gamma)$, $\quad with \quad B = \Gamma\left(B_0\Gamma_0^{-1} + \sum_{i=1}^{n} X_i\Sigma_{iu}^{-1}Z_i - G\theta_i\right)$ $and\, \Gamma = (\Gamma_0^{-1} + \sum_{i=1}^{n} X_i\Sigma_{iu}^{-1}X_i)^{-1}$, with $Z_i = \left(U_{i1}, U_{i2}, Y_{i2}^{\bar{N}}, Y_{i2}^{\bar{P}}, Y_{i2}^{\bar{U}}, Y_{i3}^{\bar{N}}, Y_{i3}^{\bar{P}}, Y_{i3}^{\bar{U}}\right)$ $and\, G = [0, 0, 1, 0, 0, 1, 0, 0]$
6. Sample $\theta_i|b, d_1, \sigma, Z_i, X_i, (\lambda_{2i})_{i=1}^{N}, (\lambda_{3i})_{i=1}^{N}$ from a normal distribution.
7. Sample $\lambda_{iu}|Z_{iu}^*, B, X_i, \sigma, (\theta_i)_{i=1}^{N}$, $\forall u = 2, 3$ from a gamma distribution.
8. Repeat step 2 to step 7 to get a full posterior distribution.

Computing the Impact of Education

Assume the MCMC chain has converged. In addition, from Step (3) of the simulation algorithm, we know the posterior distribution of the individual earnings counterfactuals. Furthermore, we assume that the independence assumption 1 holds, which, as argued above, is

22. An alternative strategy would sample from the full posterior distribution as defined in (2.8) using a Metropolis-Hasting method.

23. In fact, Chib & Hamilton (2002) find that sampling σ from the augmented sample that includes the distribution of the counterfactuals delays its convergence.

24. See details of simulation process in the appendix B.2

more robust and realistic than the one posited in the SCM approach. We have nevertheless made a number of restrictive assumptions about the distributions of the observed component of the earnings and their counterfactuals in order to reduce the number of parameters and to increase precision. Yet it is not be possible to estimate the counterfactual of the earnings associated to the particular sequences NP or PP when the individual chooses PN since all three sequences share the same earnings equation. We may nevertheless compare individuals in different, more aggregated, subgroups $\{NN, \bar{U}, \bar{P}\}$. This restriction comes from the fact that we have assumed that $Y_{iu}^{*\bar{P}} = Y_{iu}^{*PP} = Y_{iu}^{*PN} = Y_{iu}^{*NP}$ and that $Y_{iu}^{*\bar{U}} = Y_{iu}^{*UU} = Y_{iu}^{*UN} = Y_{iu}^{*NU}$. It is as if we simulate the average counterfactual earnings for each type of schooling regardless of the period in which it occurs and its duration.

Upon convergence of the posterior distribution after Q iterations (net of the burn-ins), we have a distribution of size Q for b, ψ, θ_i and $(\lambda_{i2}, \lambda_{i3})$. We can self-match individuals to compute their mean returns to education investment. Let Δ_i be this quantity.

Under assumption 2, the mean return to domestic education for individual i in periods $u = 2, 3$ obtained from the posterior distribution is given by :

$$\Delta_{iu} = \sum_{q=1}^{Q} \begin{cases} \left(Y_{iuq}^{NN} - Y_{iuq}^{*\bar{U}}\right) & if\ NN = 1 \\ \left(Y_{iuq}^{NN} - Y_{iuq}^{*\bar{P}}\right) & if\ NN = 1 \\ \left(Y_{iuq}^{UU} - Y_{iuq}^{*NN}\right) & if\ UU = 1 \\ \left(Y_{iuq}^{UU} - Y_{iuq}^{*\bar{P}}\right) & if\ UU = 1 \\ \left(Y_{iuq}^{UN} - Y_{iuq}^{*NN}\right) & if\ UN = 1 \\ \left(Y_{iuq}^{UN} - Y_{iuq}^{*\bar{P}}\right) & if\ UN = 1 \\ \left(Y_{iuq}^{NU} - Y_{iuq}^{*NN}\right) & if\ NU = 1 \\ \left(Y_{iuq}^{NU} - Y_{iuq}^{*\bar{P}}\right) & if\ NU = 1 \\ \left(Y_{iuq}^{PP} - Y_{iuq}^{*NN}\right) & if\ PP = 1 \\ \left(Y_{iuq}^{PP} - Y_{iuq}^{*\bar{U}}\right) & if\ PP = 1 \\ \left(Y_{iuq}^{PN} - Y_{iuq}^{*NN}\right) & if\ PN = 1 \\ \left(Y_{iuq}^{PN} - Y_{iuq}^{*\bar{U}}\right) & if\ PN = 1 \\ \left(Y_{iuq}^{NP} - Y_{iuq}^{*NN}\right) & if\ NP = 1 \\ \left(Y_{iuq}^{NP} - Y_{iuq}^{*\bar{U}}\right) & if\ NP = 1 \end{cases}$$

For each individual, we have two counterfactual earnings : $\Delta_u = \sum_{i=1}^{N} \Delta_{iu}$ and $\Delta = \sum_{u=2}^{3} \sum_{i=1}^{N} \Delta_{iu}$, which can be defined respectively as the Bayesian average mean gain of education ($BAMG$) in period u and over the sample period.

Let $D_j = \left\{P_{ij} | \frac{j-1}{10} < P_{ij} = 1 - P\left(S_{i1} = N | \pi'(\cdot)\right) \times P\left(S_{i2} = N | \pi'(\cdot)\right) < \frac{j}{10}\right\}$ be the group of migrants belonging to the j^{th} percentile of the joint probability to be enrolled in school over the period where school investment is feasible. $\pi'(\cdot$ is the posterior distribution of the set of all parameters. We can then compute an average effect within each percentile group, $\Delta_u^j = \sum_{i \in D_j} \Delta_{iu}$ and $\Delta^j = \sum_{u=2}^{3} \sum_{i \in D_j} \Delta_{iu}$. Unlike the SCM approach, in which matching individuals in the tails of the probabilities is problematic, the Bayesian analysis is accurate because it proceeds by self-matching individuals.

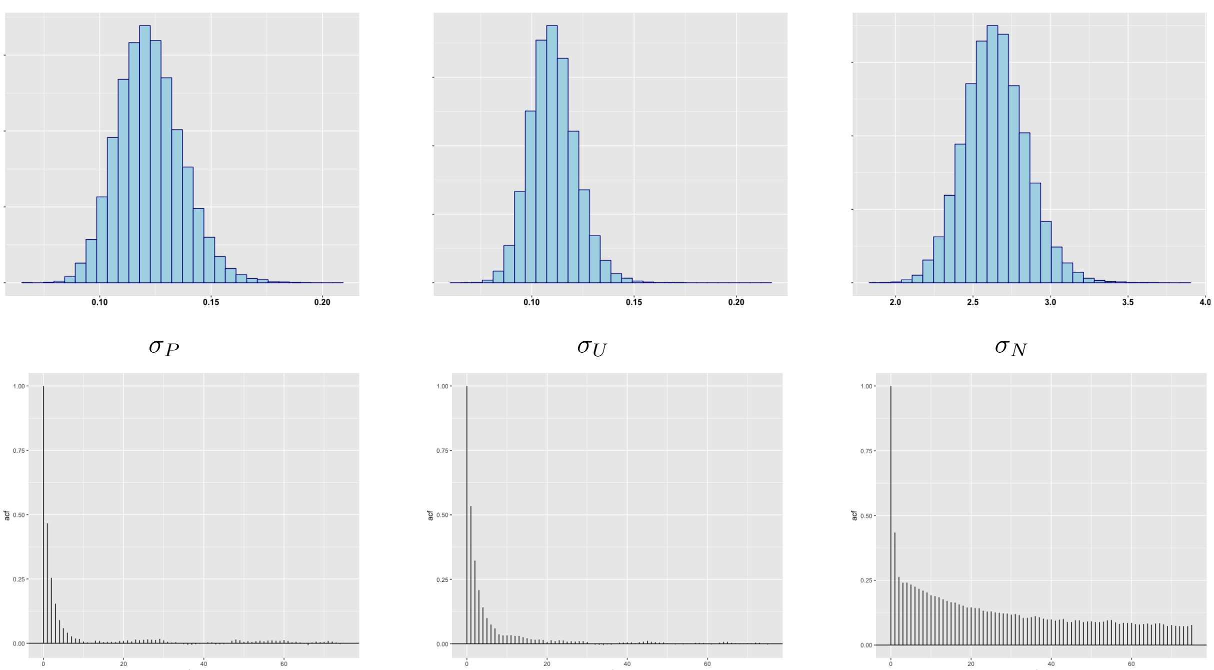
2.4.2 Results

Posterior Distribution : Implementation

We follow the MCMC algorithm discussed above (and detailed in the appendix). In addition to the three components of the variance-covariance matrix that are sampled from a Metropolis-Hasting algorithm, as many as ninety-three parameters must be estimated. The convergence is achieved by running 66,000 iterations with 6000 burn-ins. Figures 2.3 and 2.4 present the distribution and the auto-correlation function of the variance-covariance components and those of the three loading factors, respectively. The figures reveal that the posterior distribution mixed well and that the auto-correlation decreases rapidly for all loading factors and for all components of the variance-covariance matrix.

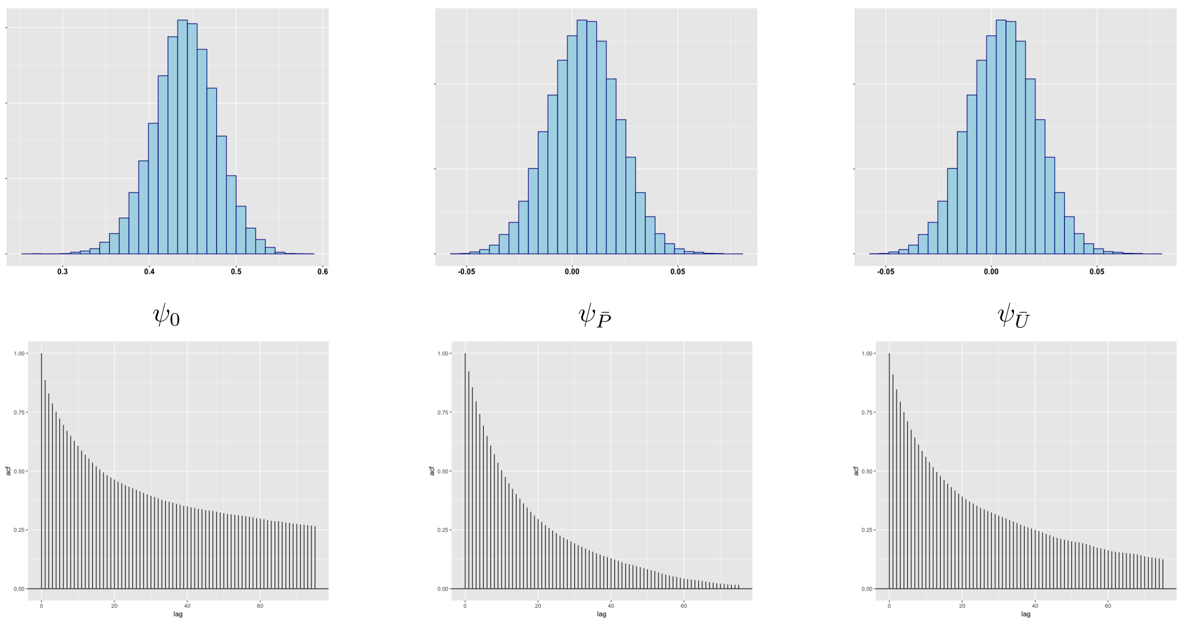
To investigate further the convergence issue, we have computed Heidelberger and Welch(1981) 's convergence test (R-package). This convergence test uses the Cramer-von-Mises statistics and account for the presence of auto-correlation across iterations. In addition, we have computed Geweke (1992)'s mean test convergence which states that a Markov Chain Monte Carlo (MCMC) converges to the proper posterior distribution if the mean of each parameter computed using the sample at the top of the chain is equal to the mean computed over the tail. Geweke (1992) proposes to use 10% of the iterations on the top and 50% of those at the tail.

FIGURE 2.3 – *Distribution and Auto-Correlation of the Variance Components*



Both convergence tests suggest that each component of the variance matrix converged to the posterior distribution. Furthermore, the distribution of the loading factors is spread to the right of zero in the equation of the *latent* utility for education while in the earnings equation the loading factors is centered at zero. Consequently, the way the unobserved variables affect the earnings is the same regardless of the training status and are all significantly equal to one.[25] To investigate the extent to which the model predicts the observed distribution of the log earnings, Figure B.2 presents the observed and the predicted distributions of the weekly earnings. As depicted, the predicted earnings distributions in both periods is quite similar to the observed ones.

FIGURE 2.4 – *Distribution and auto-correlation of loading factors posterior distribution*



Posterior Distribution : Parameters

Table 2.6 provides the mean of the posterior distribution and the credibility interval of each parameter. It also indicates whether the mean is significant at the 5% or 10% levels. There are three earnings equations (columns 3–8) for those who, respectively, received a domestic PSS degree or less ($\bar{P}$), those who received a domestic university degree ($\bar{U}$) and those with a foreign degree (NN). The posterior distribution of the variance reveal that the average variance in the earnings equation for individuals belonging to NN is larger than the other variances, reflecting once again the heterogeneity among individuals of this group.

Attending Domestic Schooling

From the two first columns, it is readily seen that the probability to invest in a domestic education decreases with age at entry and with the level of development of the foreign-degree granting country. Married individuals as well as those with extensive foreign work experience a re less likely to enrol in domestic education. Furthermore, compared to migrants who arrived

25. Recall that the sign of the loading factors provides information about the sign of the correlation between earnings and schooling decision.

in 2002, those who arrived between 2003 and 2004 have a much lower probability to do so, while those who landed between 2005 and 2006 have higher probability. It also appears as though longer unemployment spells upon landing are likely to induce individuals to enrol in domestic education.

Education and Earnings

The table shows that women have lower earnings among those who have a post-secondary or a university degree. On the other hand, their earnings are slightly greater among those who did not attend formal schooling (NN). Not surprisingly, married individuals earn more, and more so among the $\bar{P}$ and $\bar{U}$ samples. For all groups, earnings increase with age and are negatively correlated with the squared of age, which is consistent with the human capital theory.

The origin of the education acquired abroad significantly affects earnings. Hence, within each group those who were trained in more developed countries have higher earnings than those whose education was acquired in less developed countries. This is consistent with Coulombe *et al.* (2014b); Kanas & van Tubergen (2009), in that the origin of foreign-acquired education is correlated with the migrants' earnings in the host country. Furthermore, from columns (3), (5) and (7) it appears that the level of education acquired abroad positively affects earnings even for those who have a domestic degree. In fact, for those belonging to the $\bar{P}$ and $\bar{U}$ groups, having a bachelor's degree acquired abroad increases the earnings respectively by 14% and 15% comparing to those with post-secondary degree or less, which in line with the underlying assumption that migrants go back to school to enhance their foreign education. [26]

As expected, when significantly different from zero, the experience acquired abroad is positively correlated with the earnings and negatively related to the probability to invest in domestic education. However, experience acquired in Quebec does not seem to affect earnings. This result is consistent with that found by Hou & Lu (2017). Finally, and not surprisingly, weekly earnings decrease with time spent in the first unemployment spell for those the $\bar{P}$ group.

The loading factors in all earning equations are not significantly different from each other and are not significantly different from one. There is thus little evidence of unobserved heterogeneity in the earnings. In the latent utility, however, the loading factor is equal to $\psi_0 = 0.44$. Thus, the magnitude of the correlation between the latent utility for education and the earnings profile due to unobserved factors depends only on the loading factor in the latent utility equation and on the variance of θ_i.

26. the marginal effect is given by $100 \times (\exp(0.13) - 1)$ and $100 \times (\exp(0.14) - 1)$.

TABLE 2.6 – *Mean Posterior Parameters*

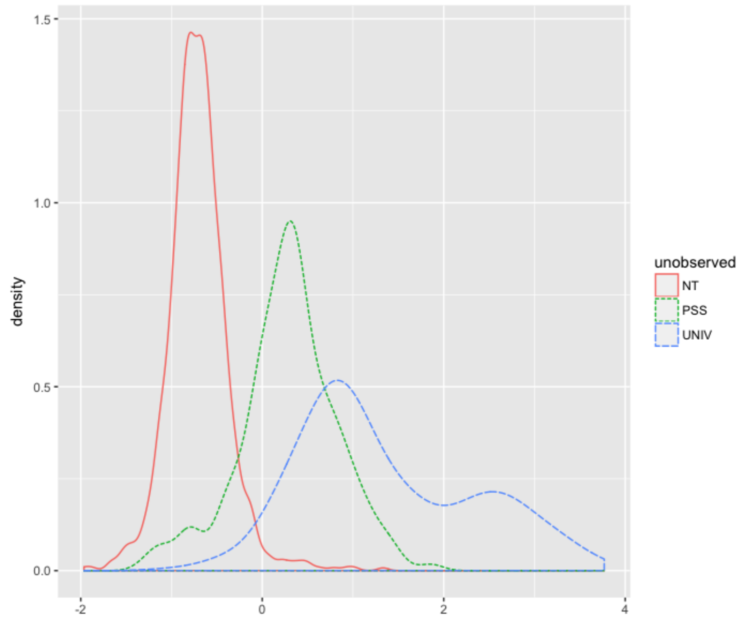
	Latent utility		**Income for $\bar{P}$**		**Income for $\bar{U}$**		**Income for NN**	
	Mean	*se*	Mean	*se*	Mean	*se*	Mean	*se*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age at admission	-0.03**	0.01	0.32**	0.01	0.35**	0.01	0.39**	0.02
Age at admission^2	0.00**	0.00	-0.004**	0.00	-0.01**	0.00	-0.01**	0.00
Female	0.00	0.07	-0.08*	0.04	-0.09**	0.04	0.01**	0.00
Married	-0.12*	0.06	0.06	0.04	0.10**	0.05	0.01**	0.00
Development Index	-0.07**	0.03	0.05**	0.02	0.04**	0.02	0.09*	0.05
Schooling :								
Foreign Bachelor's	0.09**	0.04	0.13**	0.04	0.14**	0.05	0.01	0.01
Foreign M.A/PhD	-0.02	0.03	0.14**	0.06	0.11*	0.06	0.01**	0.01
Foreign Experience	-0.07**	0.01	0.12**	0.05	0.04	0.05	0.02**	0.00
Previous Stay in QC	-0.12**	0.03	0.10*	0.05	0.16**	0.05	0.03**	0.01
French Score	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.02
English Score	-0.01	0.02	0.07**	0.01	0.07**	0.01	0.06**	0.03
ECA	0.15**	0.05	0.10*	0.05	0.01	0.05	-0.01	0.01
2003-2004	-0.14**	0.04	-0.03	0.05	0.06	0.06	0.00	0.01
2005-2006	0.08**	0.04	0.06	0.04	0.04	0.05	0.01**	0.00
2007-2008	0.04	0.04	0.03	0.05	0.03	0.06	0.01**	0.00
Time until 1^{st} job/School	0.64**	0.05	-0.11**	0.03	-0.05	0.04	-0.00	0.00
Job is Professional Order			0.08**	0.04	0.04	0.04	0.01**	0.00
Language Course	0.05**	0.01	-0.00	0.05	0.02	0.05	0.01**	0.00
Change Field Study			0.10**	0.04	-0.01	0.04		
Gov. Loan or Grants		0.03	0.04	-0.07	0.04			
Prior Qualified Job			0.08*	0.04	0.24**	0.04	0.02**	0.01
Experience	-0.06**	0.01	0.05	0.03	0.04	0.03	0.01	0.05
Experience^2			0.00	0.01	0.00	0.01	0.00	0.01
Intercept	-0.00	0.00	0.02**	0.00	0.03**	0.00	0.02**	0.00
Loading Factors	0.44**	0.04	0.01	0.02	-0.01	0.01		
σ			0.12**	0.01	0.11**	0.01	2.64**	0.20
Cut-Off Points	0.70**	0.04						

Notes :

- ⋆ ** $p < 5\%$; * $p < 10\%$.
- ⋆ For cohort fixed effects represented by dummy years, the year 2002 is left out.
- ⋆ The first column of each equation corresponds to the mean of the posterior distribution, while the second reports the Credibility Interval.
- ⋆ Development Index : DI $= ln\left(\frac{GDP^o}{GDP^c}\right)$ is used to proxy the quality of foreign education.
- ⋆ ECA : Educational Credential Assessment

Nevertheless, even though not significantly different from zero, if we account for the sign of the loading factors in the earnings equation of the$\bar{u}$ group, it may be said that the correlation between the selection in training and the earnings equation due to unobserved confounders is negative. The unobserved heterogeneity is positive for almost all individuals in the who have invested in a domestic university degree and the distribution is the most spread on the right(see Figure 2.5).

FIGURE 2.5 – *Distribution of the unobserved variables.*



Conversely, for those who have chosen not to invest in a domestic education, the unobserved confounders are negative for almost everyone. Thus, those with a domestic university degree are more likely to have a low return on their investment. Moreover, unobserved heterogeneity that positively affects the likelihood to invest in education is less correlated to earnings. Therefore, we can be confident that the unobserved factors estimated here are actually the latent distribution for *ability*.

Bayesian return to post-migration education

The first four columns of the table 2.7 show that the return to post-migration education are heterogeneous both with respect to the duration and the timing of the investment. Overall, the gain of post-migration education over the foreign-acquired one is negative but tends to increase as time unfolds, irrespective of education degree, the timing and the duration of the investment. Thus, if the upward trend holds over time, domestic education may turn out to be worthwhile in the long run. Moreover, by comparing the actual earnings of migrants who only have a foreign education to their counterfactuals earnings in the case they also have a QC-education, we find that the gains of the foreign-education over domestic education decrease over periods. This finding reinforces the idea that a domestic education may be beneficial in the long run. Note that the return is expressed in terms of difference in log earnings and could be interpreted as a $(\exp(\delta) - 1)$ percentage changes.

We also find that the timing of schooling matters in terms of return and thereby, any postponement may reduce its net gains over the foreign-acquired education. This is so presumably because education acquired in Quebec takes times and yields its potential benefit long after completion. In fact, those who spent two periods in domestic education are more adversely affected, and so are those attending a university degree.

One of the main contribution of this paper is to provide evidence that limiting oneself to the average treatment effect may hide heterogeneous returns to education across individuals. Figure 2.6 reports individuals returns to domestic education by education status and only for the last period. On the top left of the Figure, we compare the weekly (log) earnings of individuals in the group-NN to their counterfactual earnings if they had earned a domestic post-secondary degree.

Table 2.7 – *Posterior distribution*

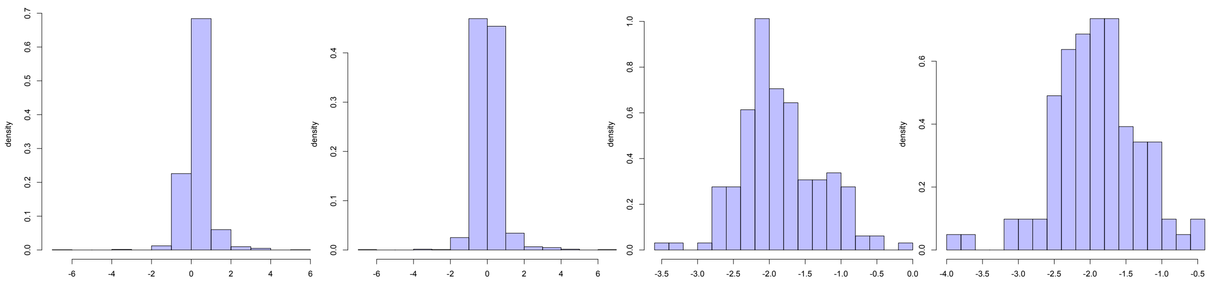
	Whole sample				sample of migrants with $\theta_i > \bar{\theta}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimator	Δ_1	se	Δ_2	se	Δ_1	se	Δ_2	se
$NN\, vs\, \bar{P}$	0.34	0.02	0.15	0.02	0.70	0.02	0.46	0.02
$NN\, vs\, \bar{U}$	0.16	0.02	-0.02	0.02	0.48	0.03	0.25	0.03
$PN\, vs\, NN$	-1.33	0.05	-1.15	0.04	-1.61	0.05	-1.48	0.05
$PN\, vs\, \bar{U}$	-0.11	0.03	-0.06	0.03	-0.09	0.04	-0.10	0.04
$NP\, vs\, NN$	-1.08	0.06	-1.10	0.07	-1.52	0.06	-1.53	0.06
$NP\, vs\, \bar{U}$	-0.06	0.04	-0.18	0.06	-0.02	0.05	-0.11	0.06
$PP\, vs\, NN$	-1.88	0.05	-1.77	0.05	-2.03	0.04	-1.94	0.04
$PP\, vs\, \bar{U}$	-0.04	0.03	-0.03	0.02	-0.03	0.03	-0.03	0.03
$UN\, vs\, NN$	-1.68	0.07	-1.55	0.05	-2.51	0.09	-2.35	0.10
$UN\, vs\, \bar{P}$	0.15	0.05	0.15	0.04	0.06	0.04	0.11	0.02
$NU\, vs\, NN$	-1.65	0.09	-1.42	0.07	-2.47	0.10	-2.33	0.15
$NU\, vs\, \bar{P}$	0.08	0.07	0.19	0.04	0.07	0.12	0.00	0.09
$UU\, vs\, NN$	-3.38	0.07	-3.12	0.08	-3.36	0.07	-3.20	0.07
$UU\, vs\, \bar{P}$	0.02	0.04	0.17	0.03	0.09	0.03	0.15	0.03

Note : For each average treatment effect, the first sequence of treatment represents the actual treatment and the second represents the contrefactual.

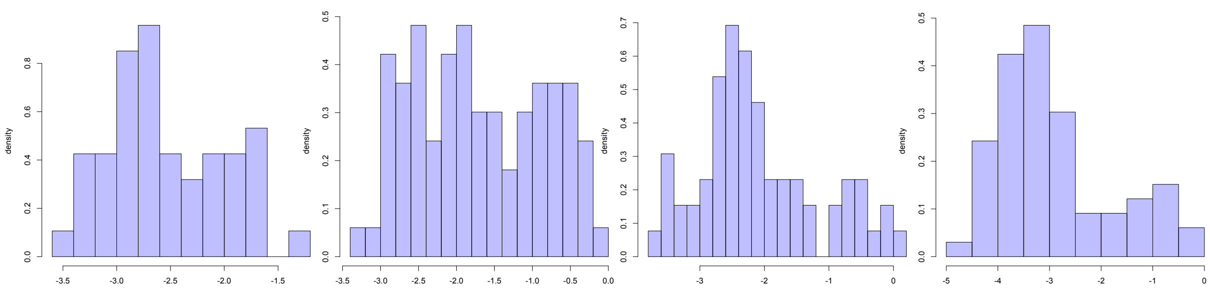
For other diagrams of this Figure, the first sequence of school investment represents the actual education decision and the second sequence represents the counterfactual outcome. It emerges that, regardless of the degree, the return to post-migration education varies greatly across individuals. Interestingly, the gain of a domestic education over a foreign-acquired one is negative for all individuals, irrespective of the degree, and that fewer than 3% of migrants with post-secondary degree acquired in Quebec benefit from a net positive return. Our findings support the view that migrants face a discrimination on Quebec labor market going beyond the country origin of their education.

Although the parameters of the likelihood function suggested that migrants whose education was acquired in less developed countries had lower earnings, Table B.4 shows that those originating from Africa have higher returns to post-migration education than migrants from Europe. This means that migrants from Africa have certainly other attributes that allow them to reap large returns or, alternatively, that migrants whose foreign education have acquired in Europe and do invest in a domestic education are a very particular and selected group. The last suggestion is in line with the results of Djuikom & Lacroix (2018).

Figure 2.6 – *Distribution of the effect in last period for each individual given the training status. Pr indicates the proportion of migrants with positive effect.*



$NN\,vs\,\bar{P}$; $Pr = 32.7\%$ $NN\,vs\,\bar{U}$; $Pr = 53.1\%$ $PN\,vs\,NN$; $Pr = 97\%$ $NP\,vs\,NN$; $Pr = 97\%$

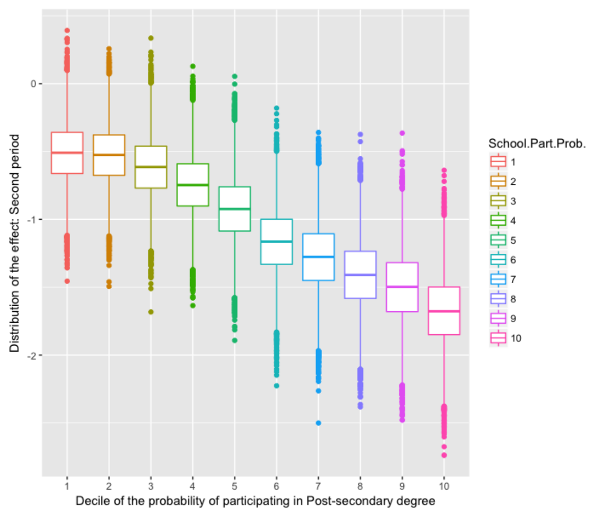


$PP\,vs\,NN$; $Pr = 100\%$ $UN\,vs\,NN$; $Pr = 100\%$ $NU\,vs\,NN$; $Pr = 100\%$ $UU\,vs\,NN$; $Pr = 100\%$

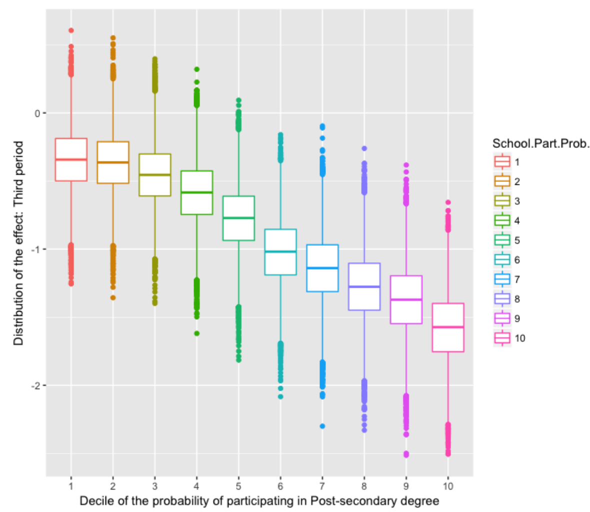
In columns (5) to (8) of table 2.7, we report the Bayesian-average mean gains of post-migration education of migrants for whom the latent *"ability"* is greater than the average in the corresponding group. It is found that they also have lower gain from a domestic education than those with low level of ability. Moreover, Figure B.3 shows that the gain of Quebec education over the foreign-acquired education decreases when moving from the lowest decile to the highest decile of the distribution of unobserved variables. Even if the ability and level of education can give an idea about the productivity of migrants, our results reveal that it does not really matter. Actually, it seems that individual employers pay migrants by considering particularly their previous jobs, if it is a qualify job or not. This is even more important for migrants who have invested in University degree because having a qualified job in the past increase the earnings by 27% versus an increase of 8% for migrants who have Quebec Post-secondary degree and 2% for those who only have a foreign-education(see columns (3), (5) and (7) of Table 2.6).

FIGURE 2.7 – *Distribution of the effect by decile of likelihood of participating in education*

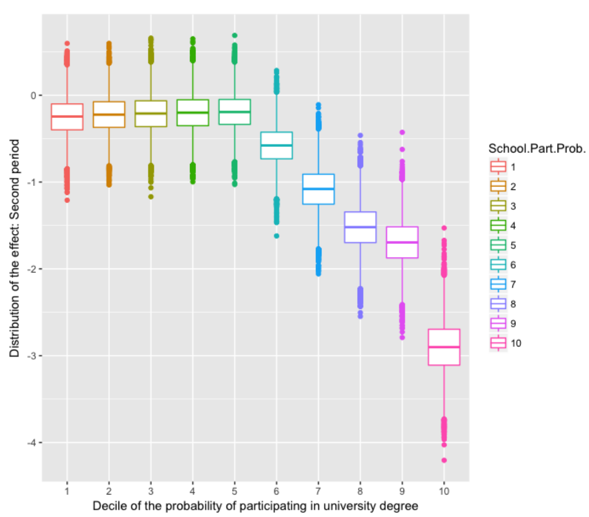
(a) *Period 2*



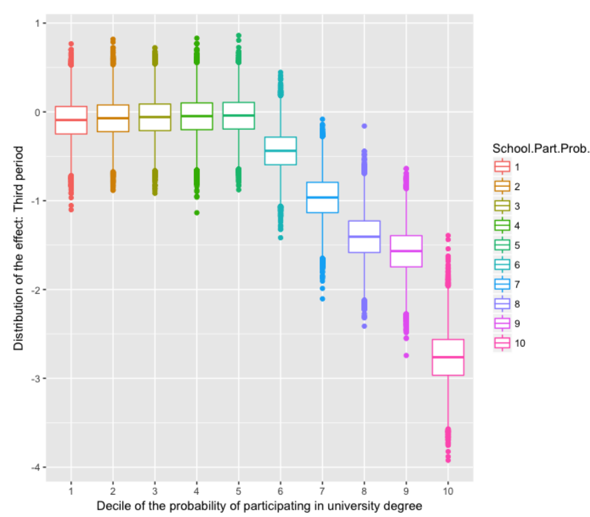
(b) *Period 3*



(c) *Period 2*



(d) *Period 3*



Additionally, as shown in Figure B.4, the returns to post-migration education decrease with the likelihood of participating and therefore individuals with higher probabilities also have lower returns. In other words, migrants who are on high risk to invest in QC-education also have lower returns from that investment. This suggests again that the school achievement is not the main factors that determine the earnings profile of migrants but, the most important is the quality of previous jobs held. Those findings are consistent to those of Hou & Lu (2017) and according to which the earnings of migrants who have a Canadian diploma depend on whether they held a well paid job in Canada in the past. However, we should stress that one would expect immigrants to accept, right after their training, a relatively less paid job than the one he would have had given his education, which could explain this negative returns in the short-term.

2.5. Conclusion

This study investigates the returns to domestic *versus* foreign-acquired education in terms of weekly earnings for immigrants admitted in Quebec under the Skilled Workers Program. Unlike previous studies, we observe foreign-acquired education and the periods of investment in domestic education. Moreover, we account for the dynamic assignment into schooling by assuming that the latter is feasible only during the first four years upon landing. Immigrants have thus the possibility to enrol in university degree or a post-secondary degree. Furthermore, we measure the return to education four and six years after migration. Our estimation strategy relies on the Bayesian analysis that enables to account for the selection into education based on observed and unobserved heterogeneity. As a result, we can self-match each immigrant and then compare his actual earnings to two counterfactuals earnings. For example, for those who earned domestic university degree, we can compare their earnings to their counterfactuals earnings which consists in their foreign-acquired education or a domestic post-secondary degree.

Our results show that focusing on the average effect may lead to ill-conceived policies, despite the fact that the return is negative for most, as it varies greatly across individuals. Furthermore, we find that the gains are increasing over time so that it is conceivable that a domestic education may yield positive net returns in the long run.

In this paper, we interpret and model the unobserved heterogeneity as latent individual ability. Our results indicate that those who earn a university degree have higher average abilities. It is also found that the average gains from a domestic education tend to increase when the correlation between earnings and participation in a domestic degree decreases. Consequently, immigrants with lower *abilities* benefit the most from a domestic education, which indicates substantial negative selection into the latter.

According to the Bayesian approach, immigrants who have enrolled in a university program in Quebec have the highest negative return to education. This result can be explained by the that right after their training, immigrants are more likely to accept a relatively less paid job than the one he would have had given his education. Conversely, the more conventional Sequential Inverse Probability Weighting estimation strategy yields opposite results : Those who earn a domestic university degree reap the highest benefits. Both estimation strategies rely upon a different set of identification assumptions. The former relies upon parametric distributions to account for the unobserved heterogeneity, while the latter assumes them away through the SUTVA (Stable Unit Value Treatment Assumption (SUVTA)) assumption. Further investigation is certainly warranted to help understand why the two approaches yield such different results. In all fairness, our study might suffer from small its sample size as many observation had to be removed to make both approaches amenable. Therefore, our results need to be interpreted with caution.

- Chapitre 3 -

Incentives to labor migration and agricultural productivity : The Bayesian perspective

Abstract

Understanding how internal labor migration affects the agricultural sector is important for all developing countries whose markets do not work well or are non-existent. The availability and the quality of labor and the cost of hiring people to work on farms is an example of a problem that farmers may face in the presence of a critical level of labor migration. Since households select themselves into migration this raises the endogeneity problem. In order to account for this endogeneity and the fact that the effect might be different from one household to another, we model the households' decisions to participate in migration along with their investment in agricultural production using the Bayesian treatment analysis. This approach allows us to self-match each household and to estimate a distribution for the counterfactual outcome. The results show that even if on average internal labor migration positively affects the agricultural productivity, there are some households for which the effect is negative. Those households for which the effect is negative are mostly small farmers, and are therefore more likely to be involved in subsistence farming and to be poor. Moreover, the average effect of the labor migration tends to increase with the likelihood of participating in the internal labor migration. In parallel, we also find that previous migration rates, widely used in the literature as instrument for the migration decision, are intimately correlated to the agricultural productivity.

Keywords : labor migration, agriculture, Bayesian treatment analysis, instrumental variables, rural, Uganda.

JEL : O15, O13, O18, J61, Q12.

3.1. Introduction

Understanding how internal labor migration affects the agricultural sector is important for all developing countries whose markets do not work well or are non-existent. In fact, even if the movement out of the agricultural sector can be viewed as a process to reach development for many African countries, this could lead to a negative effect on the farmers' activities and thus on the rural economy. The availability and the quality of labor and the cost of hiring people to work on farms are examples of problems that farmers may face in the presence of a critical level of labor migration. In addition, farmers are less likely or unable to invest in technologies that allow them to rely less on labor. In Uganda, the agricultural sector is highly labor-intensive and it employs about 80% of the active population. Surprisingly, so far there are few empirical studies devoted to the issue (De Brauw, 2010; Mwesigye & Matsumoto, 2016; Mendola, 2008; De Haan, 1999).

From a theoretical point of view, labor migration can have a positive or negative effect on agricultural production (De Brauw, 2010). The effect could be positive if the migrant-sending households can hire individuals on the local labor market to substitute for the migrant or use remittances to invest in inputs such as buying new land, invest in new crops or buying fertilizer to increase the production. If, on the other hand, the households cannot find a replacement because of the scarcity or the cost of the labor force [1], labor migration could negatively affect the agricultural production. For these reasons, the effect of the labour migration is likely to be different across households depending on factors such as the migrant's productivity and the ability of households to adjust their decisions to the local market constraints. It is therefore important to account for this heterogeneity in order to figure out which households are positively or negatively affected by the migration and then to be able to provide relevant policy recommendations for the most vulnerable population.

Unlike the existing literature that assumes that the effect of internal labor migration is homogeneous across households, this study estimates the distribution of the effect of the internal labor migration on the agricultural productivity of households living in the rural areas of Uganda. We use the four survey rounds of the unique nationally representative panel surveys, the Uganda National Panel Survey (UNPS) that started in 2005. To achieve our goal, we estimate simultaneously and over time the households' decisions to invest in agricultural production and to participate in internal labor migration using a Bayesian treatment approach. This approach has been used to allow the heterogeneity of the effect between individuals and to account for the endogeneity of their decision. In fact Carneiro *et al.* (2003) use this approach to estimate the distribution of the return to school in terms of earnings for youths in the United States.

Furthermore, we consider that the labour migration decision is a common agreement among household members to send one or more member(s) outside the village (in another district) to find additional sources of income so as to increase household resilience to negative shocks. Indeed, for the households living in the rural areas, the labour migration provides a kind of insurance in times of bad harvest and a source of financial support to smooth household consumption, expand household's business(es), launch a new business, or invest in education. This insurance is even more important for the poorest households in the rural areas of most developing countries such as Uganda, because the access to credit is virtually non-existent without any collateral. Therefore, labour migration is not a unilateral decision taken by one person. Even if a household member can decide to migrate on his own

1. It can be difficult to find a substitute if the migrant is highly productive, especially so if we consider villages with high migration rates.

initiative, it is less likely that it will be done without the consent of other household members. Indeed, the loss of available labor to the household directly affects the labor supply in and out of the farm sector of all members left behind, and particularly in farming since the agricultural sector is highly labor-intensive. For instance, Mu & Van de Walle (2011) find that labor migration increases the time that women left behind spend doing domestic and farm work.

In this way, this study is fully integrated into the *New Economics labor Migration (NELM)* theory developed by Stark & Bloom (1985) for which the labor migration decision is a common agreement by household members. Using the Bayesian approach, we are able to account for the endogeneity of the migration decision and to estimate an average effect for each household. Moreover, we allow for the selection into migration to depend on households' time-invariant unobserved heterogeneity such as households' willingness to take risks since it is not sure that households participating in internal labor migration will have positive returns from it. These confounders affect both the labor migration decision and the investment in agricultural production. Most studies do not account for the self-selection of households into migration based on their unobserved heterogeneities, which can lead to bias estimations of the impact of labor migration. As a result, this paper thus contributes to the literature by evaluating to what extent this type of selection is related to the effect we attempt to identify and investigates whether the internal labor migration decreases (increases) the agricultural productivity of all households in the rural areas of Uganda. The Bayesian approach has recently been introduced in the treatment analysis and provides, from an implementation point of view, an easier way to account for households' time-invariant fixed effects and to estimate a distribution of the counterfactual outcome for each household.

To sum up, the contribution of this paper is twofold. First, to the best of our knowledge this is the first study that attempts to estimate the causal effect of internal labor migration on households' agricultural productivity in Uganda. Moreover, we go beyond estimating an average effect by estimating the distribution of the effect on households. Second, as shown by Conley *et al.* (2012) the methodology used here allows to verify the exclusion assumption on the variables used as instruments to correct for the endogeneity of the participation in labor migration. In fact, it is generally admitted in the literature that the past history of households' participation in internal labor migration only affect the current and future migration participation but not the pattern of the outcome of interest.

We find evidence that the average impact of the internal labor migration on agricultural productivity is positive. In fact, labor migration tends to increase the agricultural productivity of households participating in labor migration by 44%. This is in line with the NELM theory which argues that the labor migration enables households to invest in the agricultural sector. However, there are some households for whom the effect is negative, about 30% of households participating in migration. These households are mostly small farmers, and are therefore more likely to be poor. Furthermore, households with a higher likelihood of participating in migration also have a higher average effect, meaning that households select themselves into migration in order to increase their agricultural productivity. Moreover, our results suggest that previous participation in the internal labor migration violate the Instrumental variable exclusion restriction assumption and thus have to be included in the agricultural productivity equation.

The remainder of this paper is organized as follows. The next section provides a brief review of the literature on internal labor migration in Africa and an overview of the socio-economic environment in Uganda. In section 3.3, we describe the data and provide a preliminary analysis. Section 3.4 presents the empirical model and discusses the identification strategy. Section 3.5 presents the main results and section 3.6 is the conclusion.

3.2. Context and brief literature review

The incentives for internal labor migration and its (average) effect on the outcomes of the household members left behind as well as on the migrants have not attracted much attention (Garip, 2008; Mendola, 2008; Mu & Van de Walle, 2011; Garlick *et al.* , 2016; De Haan, 1999; Mwesigye & Matsumoto, 2016; Muto, 2009; Kuhn, 2015; Stark & Taylor, 1991; De Brauw, 2010). Moreover, through the available data it is hard to identify who migrates internally and further to evaluate how important the internal labor migration is in developing countries.

Existing studies reveal that the push factors of the internal labor migration range from the economic costs (household wealth), the household social capital network and relative deprivation, the variation in rainfall, to the increase in the investment in human capital. Furthermore, labor migration enables households to invest in the non-farm activity and in education, acquire new land, improve the health of household members, smooth household consumption and more generally improve the well-being of the household left behind. However, the internal labor migration does not lead to a huge transformation in the agricultural sector but helps households to meet their basic needs, and lightly increases agricultural productivity (De Haan, 1999). Based on seven papers investigating the effect of migration on households agricultural production in the rural areas of six developing countries, Davis *et al.* (2010) find that migration facilitates a transition away from agriculture or leads to less labor-intensive agriculture. Empirical evidence shows that the *NELM* failed in most cases since the labor migration does not necessarily leads to an investment in the agricultural sector.

Given the expansion of internal labor migration in the developing countries, if migration negatively affects the largest producers, it might be urgent for the government to invest more in their agricultural sector in order to stabilize or even increase the domestic agricultural production. When the domestic production is more volatile, households living in the rural areas whose daily food consumption is intimately related to the local products can see their purchasing power decrease significantly. Meanwhile, many recommendations in terms of development policies have long focused on investment in education and in non-farm sectors as a natural way out of poverty. As a result, from the moment that households are able to have access to capital through migration, they are more likely to invest in non-farm businesses that lead to a decrease in total agricultural production.

The case of Uganda is interesting since its economy is heavily dependent on the agricultural sector. It is an East African country emerging from decades of conflict and security challenges in the Northern part of the country. In 2006, all stakeholders signed a cessation of hostilities agreement. These conflicts have caused a massive displacement of people from the affected areas and delayed their development.. However, many people have already returned homes and to their day-to-day life. It is then likely that the average agricultural productivity will be lower in this part of the country. Nevertheless, northern Uganda is not the region with the highest rate of migration (labor migration or other). In fact, similarly to central Uganda, only 50% of people who were born in the north continue to live in the north. Besides, this study covers the period 2009-2011 which corresponds to at least three years after the end of the conflicts.

Additionally, less is known about the incentives of internal labor migration and its effects on the well-being of households left behind in Uganda. In fact, most studies devoted to labor migration in Uganda are primarily descriptive. Rutaremwa (2011) and International Organization for Migration

& Uganda Rapid Migration Profile technical working group (2015) give the profile of migrants and describe the uses of remittances from both international migration and internal migration in Uganda. Their results show that the remittances are mainly used for school investment, savings and to invest in buildings. Besides, Rutaremwa (2011) points out that compared to international migrants, internal migrants come from poorer households and that remittances are lower in value. This implies that the return to internal labor migration in Uganda may not balance the cost in terms of labor loss. Nevertheless, Jagger *et al.* (2012) find a positive effect of circular migration in the logging business on the households living in the community of origin of migrants. Particularly, their results reveal that migration reduces inequality in the community of origin, yet, the study focuses only on households living in the southwestern part of Uganda. Our paper thus fills a gap in the literature by providing evidence of the effect of the internal labor migration using nationally representative data. However, in the rural areas of Uganda, Muto (2009)) finds that internal labour migration increases with the household's social network proxied by being a member of the larger ethnic group present in the capital (Kampala).

Moreover, Uganda is one of the poorest countries in Africa with 80% of the population living in rural areas and where the agricultural sector has contributed up to 28.3% of the GDP in 2011 and 25.5% in 2015 [2]. In addition, Uganda is the country with the highest proportion of its population being aged less than 30 years old ; they represent around 78% of the entire population and about half of the population is under 16 years old. Meanwhile, school dropout is a big concern, Ssewamala *et al.* (2011) report that, in 2007, only one third of children enrolled in the first grade of primary school were likely to participate in the seventh year (last grade of primary school). In fact, as we can see in Figure C.1.(a), even though the education level of people aged between 15 and 35 years has increased over time [3], only 50% of these individuals have actually completed their primary education, which corresponds to seven years of education in the Figure. Along with that, the internal labor migration rate at household level has increased over time. Indeed, the percentage of households involved in migration was 10% in 2005 versus 24% in 2011 and, the share of migrating household members is increasing with the average education level of members living in the same house (Figure C.2). As a result, many more Ugandans are involved in labor migration now with the recent boom in educational attainment.

The dataset contains information on the duration of the labor migration and whether the migrants are still considered a household member. In the sample, temporary labor migration accounts for around 95% of the internal labor migration, i.e. when a household member lives away from home for at least three months and is still considered a member of the household and is not a permanent migrant. The average migration duration within each household ranges from three to ten months. Consequently, our focus in this paper is on the temporary internal labor migration.

Furthermore, there is evidence that migrants are more highly educated than non-migrants implying that households participating in labor migration have members with higher ability than non-migrant households. Figure C.3 shows that the agricultural production increases with the household head's years of education and the average years of education among the members of the household. Thus, the existence of unobserved confounders such as ability that is positively correlated with the level of education may affect both the agricultural production decision and incentives to involve in

2. Trading economics website :https://tradingeconomics.com/uganda/agriculture-value-added-percent-of-gdp-wb-data.html

3. This result might be due to the "*Universal Secondary Education (USE)*" implemented by the government in 2007 to make tuition fees free for ordinary secondary schools. Asankha & Takashi (2011) find a positive effect of USE on girls in secondary school enrollment.

migration. Therefore, the existence of such a confounder can introduce one source of endogeneity of household assignment into labor migration.

Another main source of endogeneity is the social capital network. In developing countries it is very common for people belonging to the same ethnic group, or belonging to the same family (called the *"strong ties"*) or living in the same village (called the *"weak ties"*), to help one another. In the case of migration, a household's social network increases the propensity to participate in migration by reducing the uncertainty surrounding the expected gain from migration and by reducing the monetary costs. From its social network, a household can get information about job opportunities in the potential destination, a place to live or even a ride to the desired place. The endogeneity comes from selection into social network and the fact that we do not observe the household's social network while it is likely that it affects both the migration decision and the household production.[4] Furthermore, when the household decides to participate in labor migration, it simultaneously decides the level of agricultural production given the land owned, the labor available (from household members and hired labor) and other inputs. Consequently, the selection into the migration is not exogenous of the agricultural production decision.

3.3. Data and preliminary analysis

This study uses the Uganda National Panel Survey (UNPS) conducted by the Uganda Bureau of Statistics (UBOS) with the technical support of the Government of the Netherlands and the World Bank Group. The units of interest for these surveys are individuals, households, and community/facilities. The surveys cover multiple topics such as health, education, consumption, labor force, etc. with a special module on agricultural activities. This is the first national representative panel data that covers a large number of socio-economic and demographic indicators. Launched in 2009, the surveys are implemented over a 12-month period divided in two visits (over a period of six months each) in order to rule out or at least minimize measurement errors on agricultural inputs and outputs. The first visit is about, among others, inputs and outputs of the last cropping season which could be the short or long season cropping ; besides, in the same period, the consumption module is implemented on half of the sample. Four waves implemented in 2009-2010, 2010-2011, 2011-2012 and 2012-2013 are publicly available[5]. Henceforth, each panel wave will be identified by the last year of the corresponding survey.

Initially, 3,220 households have been selected from 7,426 households interviewed in the 2005-2006 Uganda National Household Survey (UNHS) and tracked and re-interviewed up to the third wave. Thereby, it is possible to connect data from the UNHS to the panel data. At the fourth round, some adjustments have been made to the initial sample. In fact, the households extracted from the Uganda census survey implemented in 2012 have replaced a part of the panel sample. Therefore, the sample has changed significantly, and the sample weights have been corrected. The goal of the sample rotation was to correct for the attrition and random answers that might occur when the households are already used to being interviewed. The average attrition rate at household level is around 17% across waves. This attrition mainly comes from the fact that the new location of some households who have moved is unknown. Since in this study we focus on migration where a household sends a member outside but not the migration for which a household moves, it is less likely that this attrition raises issues of

4. It is well documented in the literature how social networks impact the way that people behave.

5. Data is available on the World Bank website :http ://microdata.worldbank.org/index.php/catalog/lsms

selection bias. Besides, the highest rate of attrition is observed among households living in the urban area and mainly in Kampala while the focus of this paper is on rural households.

The main goal of the UNPS is to provide reliable national representative data for the experimentation and assessment of the national policies and programs (Uganda Bureau of Statistics (2013)). The sample is clustered at the community level and covers all of the four regions and 323 communities of Uganda. In this study, we focus on agricultural households living in rural areas where about 80% of households live. Agricultural households are defined as households with at least one member operating a holding (farming household) or for which the head, reference person or the main earner is economically active in the agricultural sector (see the Glossary of Statistical Terms (2007) [6]). In data, about 75 per cent of households are actually involved in agricultural activities across waves of survey. Since there are movements in and out of the agricultural sector at the household level and that the panel data is not balanced, the analysis will focus on the balanced sub-sample for which we have information on agricultural productivity for all periods ; we will call this sample the *sample A* throughout this study. The sample of households for which the information on agricultural production is not missing at least for the first round of panel survey is noted *Sample B*.

3.3.1 Labor migration prevalence and agricultural productivity

UNPS provides detailed information that helps define migrants and their profiles. In fact, it is possible to know why a household member was absent during a certain period in the past twelve months, the duration of his absence, the district of destination and so forth. In this paper, we are interested in migration decisions at the household level, that is, we estimate the propensity of a household to get involved in labor migration. Thereby, we identify the migrants-sending household (HH) as a household with at least one member who has spent at least three months outside the household dwelling place in the past twelve months preceding the survey [7] and who is still considered a household member. [8] Moreover, it is common for a migrant to be considered as a household member. This is consistent with the way that we treat the migration decision, that is, migration is a household decision rather than an individual one.

Labor migration prevalence

With the available information, it is possible to distinguish between labor migration and migration for other economic reasons. we have grouped these two types of migration together as they lead at the end to the same goal, which is to find a job upon arrival at the destination place. For simplicity, we will identify this group of households as labor migrants-sending HHs or simply migrants-sending HHs. The data reveal that the migration participation rate among households increased over time from 11.2% in 2006 to 24.6% in 2012. However, it is not possible to know if it is rural-urban or rural-rural or urban-urban migration because only the district of destination is known and unfortunately many districts have both urban and rural areas in Uganda. Most often, migration is from one district to a different district and practically never in the same district.

Migration rates vary across regions with the highest rates in the center and in the north. Given the long history of instability in the north, the high migration rate can be driven by the less advantaged

6. https ://definedterm.com/agricultural_ household.
7. A period of three months is the most common use in the literature
8. That is, to still keep the perspective of temporary migration.

Table 3.1 – *Migration Rate and agricultural production*

	Whole sample			Agricultural Households in rural areas					
Year	All	Rural	Urban	*Sample A*	*Sample B*	Central	East	Northern	Western
Migration rate									
2006	11.2	10.2	14.5	10.0	10.5	10.0	9.2	7.3	15.3
2010	17.8	17.1	19.9	17.7	17.7	25.3	12.8	11.5	20.5
2011	23.9	23.3	27.2	25.8	24.7	27.6	20.7	19.9	31.0
2012	24.6	24.2	26.4	26.1	26.4	36.8	21.3	26.2	25.8
Total	22.3	21.8	24.1	23.3	23.0	29.0	18.5	19.3	26.2
Agricultural production per hectare									
2010	4,815.6	5,847.7	1,811.9	7,244.4	7245.2	5,966.7	4,336.9	6,289.8	11,291.8
2011	5,224.3	5,971.8	1,442.2	6,916.5	6277.9	5,146.9	7,226.6	4,511.7	7,370.6
2012	3,709.8	3,901.5	2,873.4	6,519.8	5899.6	6,477.7	3,707.0	5,559.0	7,676.1
Total	4,582.0	5,218.4	2,041.8	6,916.3	6477.6	5,792.6	5,409.5	5,325.9	8,689.3
Nb. of Obs	2,617	2,201	768	1,452	1,749	393	493	439	444

Notes :

1. Percentage is given in each cell.
2. The statistics given by region in the last four columns is computed on the *sample B.*
3. In the sub-table for agricultural production, we have reported the average production per household in kilograms per hectare.

economic environment and the lack of job opportunities. Moreover, while the migration rate increases over time in the other regions, there is a mitigated pattern in the western part of Uganda. In fact, the migration rate increased from 19% in 2010 to 31% in 2011 and decreased to 25 per cent in 2012. In the meantime, the poverty rate also increased between 2010 and 2011 and decreased between 2011 and 2012. On the contrary agricultural production has followed the opposite pattern in the western part by decreasing between 2010 and 2011 and increasing between 2011 and 2012. Therefore, it seems that agricultural production in the west of Uganda is more sensitive to the movement of people. Furthermore, the average duration that migrants spend outside their district is around two and a half months and this duration increases with the years of education of the household head and the average years of education of household members. It is also the case for the share of household members involved in migration. In other words, households where the members are more educated are also more likely to participate in labor migration for a longer period.

In our data, we also have households whose head has migrated permanently for economic or education reasons within the last ten years. This information is sometimes used to capture the migration rate among individuals. To avoid any selection bias due to the household head past migration, in Table C.1 we have reported the household labor migration status, as defined in this study, regarding the head's past permanent migration. It emerges that there is no significant difference in the proportion of the head's past permanent migration by household migration status. Therefore, it is less likely that there is selection into the actual migration participation due to the past permanent migration decision of the head. Nevertheless, we introduce this variable as an independent co-variable in migration participation and agricultural production equations.

Agricultural production

The agricultural production reported in Table 3.1 is the average production of the main crops planted, that is, production of maize, beans, coffee, peanuts, bananas and potatoes. As we can see, the agricultural production tends to decrease across periods with higher average production in the rural areas. Regarding the distribution by region, the higher level of agricultural production is found in Western Uganda and except for the East, the production decreased between 2010 and 2011 and increased between 2011 and 2012. Despite this upward trend over the last two periods, the level of production in 2012 is still lower than the production in 2010 meaning that the agricultural production has decreased over the period covered by the study. In Eastern Uganda, although we end up with the lowest productivity in 2012 (3707.0 kg/ha), there was an increase of the production between 2010 and 2011 (4,336.9 vs 7,226.6 kg/hectares). In addition, the average total area planted per household has also decreased across periods with a large drop between 2011 and 2012, which can explain the lower production in the last period.

It sometimes emerges in the literature that the migration increases land-related conflicts because, due to cultural diversity induced by migration, it may be difficult to resolve conflicts based on customary laws in the absence of formal legislative law. Data reveal that about 12.4% of households have reported having a conflict about at least one piece of land they possess ; however, the share of households involved in land-related conflicts declined across periods with only 10.5% of households reporting land-related conflicts in 2012. On whether the presence of land-related conflict can affect household investment in the agricultural sector, we find that households that faced a land-related conflict have lower agricultural productivity in the first two periods, yet in the last period there is no difference in agricultural production. Moreover, the average total area planted by households facing land-related conflicts is even higher.

We also look if the household's social network driven by ethnicity group can help increase the agricultural productivity by providing advice and tips. Figure C.1 shows the agricultural productivity given the share of households in the same ethnic group in each district. As we can see, there is a non-linear relationship between the agricultural production and ethnic concentration ; the production increases with the size of the ethnic group's social network but after crossing a certain level of ethnic concentration, agricultural productivity tends to decrease.

3.3.2 Household profile and migration status

Table 3.2 presents some households' characteristics by their labor migration status. It appears that among households participating in labor migration (Migrants-sending HHs), the migration propensity increases between the first and the second quartiles of the wealth distribution of households [9] However, from the third quartile of the wealth distribution, the migration propensity starts to decrease. Moreover, by comparing the share of migrants-sending HHs and the share of non-migrants-sending HHs in each quartile of the wealth distribution it is only for the households in the first and in the last quartiles that we have found a significant difference. In fact, for the households belonging to the first quartile of wealth distribution, the share of migrants-sending HHs is significantly lower than the share of non-migrants-sending HHs and, the share of migrants-sending HHs is higher for the households belonging to the fourth quartile. It seems that the monetary cost of labor migration could be

9. Household wealth is measured here by the total expenditures made by household for consumption, education, health, etc.

a disincentive to participate in migration for the poorest households and, for the households in the middle of the distribution, other important factors affect their decision to participate or not in labor migration. Moreover, the migration rate increases across survey rounds for households belonging to the first quartile of the wealth distribution while it decreases for households in the other quartiles. This suggests that households readjust their migration decision across periods by comparing the cost and the gain obtained from the migration and we expect that if there is any positive net gain, they will be more likely to participate in migration. We can see this evidence for the poorest households since their migration rate increased across periods. Beyond the monetary costs, there are other factors that determine households' labor migration decisions. In fact, migration decisions are intimately related to the household head's characteristics and the composition of the household.

Households headed by a widow or by a polygamous person are more likely to participate in migration. It is also the case with a female or a highly educated household head. In addition, the households' heads of migrants-sending HHs are slightly older than the heads of non-migrants-sending HHs. To investigate to what extent the missing values for a head's education can introduce selection bias, we have computed the percentage of households for which the head's education is missing. It turns out that few if not any migrants-sending HHs have heads with missing value for education, yet, there is a very small share (3%) of household heads for which the head has missing years of education among the non-migrants-sending HHs. Generally, the heads' characteristics for migrants-sending HHs and non-migrants-sending HHs do not vary across survey rounds.

Regarding the composition of the households, migrants-sending HHs have larger size households and lower infant-age dependency ratios (share of children aged less than five years old). Besides, the gap of infant-age dependency between migrants-sending HHs and non-migrants-sending HHs increases across periods, meaning that the gain due to migration may not offset the labor cost when the household has many children aged less than five years old and then needs to readjust their migration decision. We also reported in this table the average age and the education of individuals who migrate or not within each household. Results reveal that the years of education and age of persons left behind are almost identical within migrants and non-migrants-sending HHs. However, individuals who migrate are in general more educated than those left behind and the average years of education has increased from one period to another, from around four years in 2010 to seventeen years in 2012.

In order to see whether ethnic concentration or ethnic diversity constitutes a push or pull factor for migration, we have computed the average share of individuals in the same ethnic group of each household and it comes out that migrants-sending HHs have a slightly smaller ethnic network than non-migrant-sending HHs. We also look at the agricultural productivity and find that migrants-sending HHs have a higher average agricultural productivity than non-migrants-sending HHs. This supports the *NELM* theory that the labor migration might actually help invest more in the agricultural sector. In addition, migrants-sending HHs planted in larger areas and harvested many more crops than non-migrants- sending HHs.

3.4. Setting and Empirical strategy

3.4.1 Preliminaries

This study brings new insights to understand how the internal labor migration decision takes place in a dynamic setting and how it affects the well-being of households left behind in the context

Table 3.2 – *Households Profile depending on their migration status*

	2010		2011		2012	
	Migrants-sending HHs	Non-Migrants -sending HHs	Migrants-sending HHs	Non-Migrants -sending HHs	Migrants-sending HHs	Non-Migrants sending HHs
Head characteristics						
Married monogamous	0.53	0.57	0.54	0.57	0.53	0.57
Married polygamous	0.20	0.16	0.22	0.18	0.18	0.19
Widow	0.17	0.15	0.16	0.14	0.18	0.13
Divorcee or sep	0.09	0.09	0.07	0.09	0.10	0.09
=1 if woman headed-HH	0.32	0.26	0.32	0.27	0.37	0.26
Head years of educ	5.59	4.57	5.74	5.08	5.44	4.72
head educ. missing	0.00	0.02	0.00	0.03	0.00	0.03
Head age	48.45	45.46	48.14	45.74	48.36	46.42
HH characteristics						
1^{st} Quartile	0.16	0.32	0.25	0.34	0.26	0.32
2^{nd} Quartile	0.29	0.29	0.27	0.30	0.24	0.32
3^{rd} Quartile	0.28	0.24	0.26	0.23	0.23	0.24
4^{th} Quartile	0.27	0.15	0.22	0.12	0.27	0.13
HH size	7.25	5.79	8.14	6.21	8.67	6.49
Share of memb aged $\leqslant$5	0.17	0.20	0.16	0.19	0.15	0.18
Share of memb aged>65	0.04	0.03	0.03	0.03	0.05	0.04
Average age in HH	20.16	22.35	18.94	22.05	19.55	22.23
Average educ.of migrants	3.86	.	11.29	.	17.52	.
Average educ.of Non-migrants	2.39	2.60	6.91	6.91	11.13	11.13
Average age of migrants	22.37	.	20.58	.	23.59	.
Average age of Non-migrants	22.74	22.85	20.39	20.39	20.95	20.95
Average educ. in HH	5.29	5.04	4.84	4.41	5.43	4.91
=1 if conflict land-related	0.14	0.17	0.10	0.10	0.10	0.11
Ethnicity concentration in Uganda	0.09	0.07	0.06	0.06	0.05	0.06
Ethnicity concentration at district level	0.60	0.64	0.69	0.71	0.68	0.70
Agricultural production						
log Production(kg/ha)	8.20	8.00	8.08	7.95	8.15	7.94
Total area planted	4.71	3.97	4.98	4.18	2.90	2.50
Number of crops	15.49	13.38	11.63	10.60	11.18	10.21
Average days HH hired people to work on farm	7.69	5.47	5.04	6.27	12.07	7.22
Household Adult Equivalence Scale	4.77	4.10	4.10	3.91	4.04	3.89

of a developing country. We want to measure the effect of the migration decision on agricultural productivity. Since the labor migration reduces the labor force available to households, agricultural production can be negatively affected by migration mainly because in developing countries, the agricultural sector is highly labor-intensive. Therefore, the way that migration affects the agricultural productivity depends on whether or not the net gain from migration allows the household to invest enough in the agricultural sector to make up for the cost of forgone labor. To compute the agricultural productivity, we have only accounted for the main crops cultivated in all regions of Uganda to avoid selection bias related to geographical advantages of one crop to another. Therefore, we have added up the productions per hectare of maize, beans, coffee, peanuts, cassavas, bananas and potatoes.

Let us take equation 3.1 as a baseline model where α_t is the coefficient of interest that measures the average effect of labor migration on agricultural productivity in period t (Y_{it}). LM_{it} LMit takes value one if household i participates in labour migration in period t. X_{it} and Z_i are respectively time-variant and time-invariant characteristics of households.

$$Y_{it} = LM_{it}\alpha + X_{it}\gamma + Z_i\theta + \mu_i + \epsilon_{it} \quad (3.1)$$

The estimator of α_t using the Ordinary Least Squared (OLS) might be biased due to the endogeneity of the labor migration decision. Given the nature of the outcome per se, the main sources of endogeneity will be the existence of omitted variables and the simultaneity between migration decision and investment in the agricultural sector. In fact, when a household decides which member(s) to send outside, it simultaneously decides the level of agricultural production, given the (anticipated) net gain of migration, the land owned, and the available labor force (from remaining members and/or from labor they can hire in the village). In equation 3.1, μ_i captures the time-invariant factors that might be correlated with LM_{it} but is not observed by the econometrician. For instance, households involved in labor migration might have more members with higher abilities and a larger social network and may be more willing to take risks since there is no guarantee that the investment in migration will produce a positive gain.

To solve the simultaneity issue, we can add to equation 3.1 a second equation that estimates the labor migration decision. Yet, this does not solve the issue of omitted and unobserved factors. Another source of endogeneity might be the measurement error on outcomes of interest. Sometimes, it is difficult to measure,precisely, the households' agricultural productivity. Nevertheless, to reduce potential measurement error, the data on the agricultural production are collected in 6-month intervals, corresponding to the short and long cropping seasons. Even if measurement error is minimized in this way, we might still face a heteroscedasticity problem.

To correct for the potential endogeneity problem and the heteroscedasticity issue, we present in the next two sections an identification strategy that minimizes if not totally eliminates the bias.

3.4.2 Identification strategy

Due to the selection into the labor migration, the OLS provides biased estimates of the effect. The selection problems arise from many sources discussed in section 3.4.1. In the literature, different approaches have been developed to correct for the endogeneity problem due to the selection bias introduced by observed and unobserved variables. Different methods such as propensity score matching and its variants enable correcting for the bias based on observed variables in the static and dynamic analyses [10]. However, it is more difficult to be convinced that we have corrected for the selection due to the unobserved heterogeneity. Nevertheless, the Instrumental variables (IVs) approach attempts to correct for both types of selection by finding an exogenous shock that affects the variable source of endogeneity and that is not directly correlated with the outcome of interest. The problem is that finding such a variable is not an easy task. In the literature, some authors use the variations of rainfall as exogenous shock that motivates people in the rural area to move to find jobs elsewhere (Konseiga, 2007; Lucas, 1987). In our data, we did not find significant variation in rainfall for each district and it seems not to affect migration incentives. Besides, the variation in rainfall may be directly correlated with the agricultural productivity.

The IVs approach has been largely implemented in the literature to estimate the return to education using an exogenous variation in the supply side of education such as the reduction of the tuition fees, which is not correlated with the wages of individuals (see Card (2001) for a review).

10. See, e.g. Lechner (2009) for the evidence of the sequential matching approach.

In the economics of migration, empirical studies stress that the costs related to the migration are the main disincentive for households to involve in migration (Mendola, 2008; Mckenzie & Rapoport, 2007; Garip, 2008). Therefore, the instruments for labor migration are related to the household's social capital network, from which the household can have many resources (transportation to the destination place, a place to live and useful information about job opportunities). In this way, the monetary cost of migration and the uncertainty surrounding the potential returns from migration are significantly reduced. For this reason, the larger the household's social network is, the higher the propensity of migrating is. In this setting, authors distinguish the *"strong"* social network from the *"weak"* social network. The former is the network made up of household members and relatives who have experienced the labour migration and the latter is the network formed by people from the same village who have also experienced internal labour migration in the past. The weak social network is relevant in developing countries since it is common that people in the same village help one another. Therefore, the social network is a strong push factor to labor migration when the household faces a monetary constraint. We then expect to have a higher and stronger effect from the *"strong"* social network.

Furthermore, other data sources may contain specific information about the people with whom the household members interact. Since in the data there are no specific questions about the household's social network, we follow the current literature by taking as a proxy for the household's *"strong"* social network the number of household members who have experienced internal labor migration in the past and for the *weak* social network, the labor migration rate at the district level. In this study, the household's *"strong"* social network is captured by the number of household members who experienced labor migration in 2006, that is at least four years before the beginning of the period covered by the analysis. we take the much earlier prevalence of migration as an instrument to manage the potential correlation that might exist between the migration that occurred just before the period covered by the data and the agricultural productivity of the households. Nevertheless, the factors that led households to participate in migration five years ago may have led them to participate in migration today and then could be correlated with the current agricultural productivity.

In addition, it is not excluded that there are still unobserved time-invariant and/or time-varying variables that affect the migration propensity and production function over time, that is $E\left[LM_{it} \times \mu_i\right] \neq 0$ and $E\left[LM_{it} \times \epsilon_{it}\right] \neq 0$. To provide an illustration for the unobserved time-varying variables, based on the theory of learning by doing, it is expected that the more a household participates in migration, the higher the probability of succeeding in the destination place is. That could then increase the return to labor migration. Thereby, the migration prevalence in 2006 may not verify the restriction assumption as instrument. This raises the issue of the trade-off between the power of the instrument and the restriction assumption. Conley *et al.* (2012) offer a way to achieve efficiency in the case that the restriction assumption fails. The idea is to introduce the instruments in the outcome equation (the second stage of the two-stage least squared estimations) in order to test if their parameters are significantly different from zero and to have efficient confidence intervals for the causal effect. At this point, we do not know if the instruments for the household's social network (as presented in the literature) actually verify the exclusion restriction assumption because the authors usually just assume it is verified. we propose to test this assumption in this paper. Moreover, we will also test to see if the social network through ethnicity can be a push factor for the internal labor migration in Uganda.

To estimate the effect of migration over time on agricultural productivity, selection bias due to observed and unobserved variables [11]. We specify the model as follows :

11. See Heckman *et al.* (2012) for a review of using Bayesian analysis in treatment analysis and Chib &

$$\begin{cases} MU_{it}^{*} & = Z_i\beta + W_{it}\alpha_m + \theta_i\gamma + \lambda_{it}^{-1/2}\varepsilon_i \\ Prod_{1it} & = X_{it}\alpha_1 + \theta_i\gamma_1 + \lambda_{it}^{-1/2}\epsilon_{1i} \\ Prod_{0it} & = X_{it}\alpha_0 + \theta_i\gamma_0 + \lambda_{it}^{-1/2}\epsilon_{0i} \end{cases} \tag{3.2}$$

Where the system of equations that tests for the exclusion restriction assumption for the set of instruments is given by :

$$\begin{cases} MU_{it}^{*} & = Z_i\beta + W_{it}\alpha_m + \theta_i\gamma + \lambda_{it}^{-1/2}\varepsilon_i \\ Prod_{1it} & = Z_i\beta^1 + X_{it}\alpha_1 + \theta_i\gamma_1 + \lambda_{it}^{-1/2}\epsilon_{1i} \\ Prod_{0it} & = Z_i\beta^0 + X_{it}\alpha_0 + \theta_i\gamma_0 + \lambda_{it}^{-1/2}\epsilon_{0i} \end{cases} \tag{3.3}$$

$\forall t \in \{1, 2, 3\}$ stands for the time period ; the first period corresponds to the survey implemented in 2009. MU_{it}^{*} is the latent variable representing the migration utility function which is related to the migration status, LM_{it}, by the preference relation :

$$LM_{it} = \begin{cases} 1 & if \;\; MU_{it}^{*} > 0 \\ 0 & if \;\; MU_{it}^{*} \leqslant 0 \end{cases}$$

$Prod_{1it}$ stands for the agricultural production per hectare of households involved in migration in period t and $Prod_{0it}$ the counterfactual agricultural productivity if households are not involved in labor migration in period t. In the model specification, the distribution of λ_{it} allows to account for the heteroskedasticity induced by the measurement error in agricultural productivity or by the fact that the way some variables are related with dependent variable can be different among individuals. In fact, the household social capital can affect households' migration decisions differently.

Z_i is the vector of instruments for the labor migration. In addition to the prevalence of migration at household level and at district level, we have added as instrument for the labor migration the household's relative income deprivation computed by using the monthly household total expenditures and the average total expenditure of the reference group for each household in the district. [12] In the relative deprivation model of migration, Stark & Taylor (1991) argue that once we control for the absolute income gain from migration, relative income deprivation can be an incentive for households to participate in migration if both the HHs and its migrants feel less deprived. However, in the case where the migrants-sending HHs substitute its reference group with the group of households in the district of destination so that the income gain does not compensate the higher relative income deprivation given the new reference group, neither the household's absolute income, nor the relative income is going to be significant in the household's propensity to participate in temporary internal labor migration. Additionally, equation 3.3 allows to test for the exclusion restriction assumption of the instruments. In this case, if the set of parameters β^1 and β^0 are significantly different from zero, it means that the exclusion restriction assumption failed.

X_{it} is the vector of covariates that we have categorized into four sub-groups as detailed in the Table 3.3. These are households' head attributes, households' characteristics, inputs for agricultural production and risk management. In the absence of formal insurance on agricultural production, some households plant many crops to manage the risk related to negative shock, therefore, we expect to

Hamilton (2002) in the setting of Bayes treatment in longitudinal data

12. We have reported in appendix the details about the way that we compute the Index of relative deprivation for each household.

TABLE 3.3 – *Set of variables*

Inputs for agricultural production	- Hired labor(in terms of the number of days) - Adult equivalent - Total area planted - Percentage of households involved in agricultural sector within a radius of 5 km
Risk management	- Number of crops managed
Head attributes	- Marital status - Education - Age
Household characteristics	- Household wealth measured by the household total expenditures - Size - Share of members aged less than 5 years - Share of members aged more than 65 years - Share of members aged between 6 and 14 years - Share of female - Geographical deprivation
Instruments for migration decision	- Number of members involved in migration in 2006 - Migration rate at district level in 2006 - Wealth Deprivation

have a positive correlation between the number of crops and the agricultural productivity. Larson *et al.* (2015) report some irregularities in the information about the number of household members working on the farm. Therefore we take as proxy for the household labor force the household adult equivalence scale. we also include some spatial variables such as the percentage of households involved in agricultural activities within a radius of 5 km and the relative geographical deprivation [13]. The first variable can measure the extent to which households can learn new agricultural techniques from others near them. The set of variables W_{it} in the migration likelihood equation contains the same covariates as X_{it} except those related to the inputs of the agricultural production.

By assuming that $\epsilon_{si} \sim \mathcal{N}\left(0, \sigma_s^2\right)$, for $s \in \{0, 1\}$, and $\varepsilon_{it} \sim \mathcal{N}\left(0, \varsigma\right)$, the matrix format of the system 3.2 can be written as follows :

$$H_{it} = \left(\begin{array}{c} MU_{it}^* \\ Prod_{1it} \\ Prod_{0it} \end{array} \right) \Bigg|_{\substack{\alpha_m, \alpha_0, \alpha_1, \gamma \\ \gamma_0, \gamma_1, \lambda_{it}, \theta_i}} \sim \mathcal{N}\left(\left[\begin{array}{c} Z_i\beta + W_{it}\alpha_m + \theta_i\gamma \\ Z_i\beta^1 + X_{it}\alpha_1 + \theta_i\gamma_1 \\ Z_i\beta^0 + X_{it}\alpha_0 + \theta_i\gamma_0 \end{array} \right], \lambda_{it}^{-1} \left(\begin{array}{ccc} \varsigma & 0 & 0 \\ 0 & \sigma_1^2 & 0 \\ 0 & 0 & \sigma_0^2 \end{array} \right) \right) \tag{3.4}$$

Since M_{it}^* is unobservable, we normalize ς to one. In addition, we assume that there are unobserved time-invariant variables θ_i, that differently affect the migration decision and the agricultural production.

One implication of the model is that in each period, conditionally to the observed and unobserved variables, the vector of loading factors $(\gamma, \gamma_1, \gamma_0)$ and the variance of the distribution of $(\theta_i)_{i=1}^N$ drive

13. we compute the geographical deprivation using information on the geolocation of the household dwelling relative to the main road, main market, border post, administrative services, etc. in the district.

the correlation between the migration decision and the production. In fact, given the set of parameters, $\forall\, t$,

$$cov(MU_t^*, Prod_{1it}) = \gamma \times \gamma_1 \times Var(\theta_i) \text{ and } cov(MU_t^*, Prod_{0it}) = \gamma \times \gamma_0 \times Var(\theta_i).$$

In addition, across periods,

$$cov(MU_{it}^*, MU_{it+1}^*) = \varsigma \times cov(\lambda_{it}, \lambda_{it+1}) \text{ and } cov(Prod_{it}, Prod_{it+1}) = \sigma_1 \times cov(\lambda_{it}, \lambda_{it+1}).$$

On the other hand, in order to identify all the parameters, one loading factor for each outcome has to be set to one, we choose $\gamma_0 = 1$. The likelihood function is defined as follows :

$$\begin{aligned}
L(Prod_t, LM_t|B, \sigma, \lambda, \theta) &= \prod_{i=1}^{N}\prod_{t=1}^{3} f(Prod_{1it}, LM_{it} = 1) \times f(Prod_{0it}, LM_{it} = 0) \\
&= \prod_{i|LM_{it}=1}\prod_{t=1}^{3} f(Prod_{1it}|Z_i\beta^1 + X_{it}\alpha_1 + \theta_i\gamma_1, \lambda_{it}^{-1}\sigma_1)P(LM_{it} = 1|Z_i\beta + X_{it}\alpha_m + \theta_i\gamma, \lambda_{it}^{-1}) \\
&\times \prod_{i|LM_{it}=0}\prod_{t=1}^{3} f(Prod_{0it}|Z_i\beta^0 + X_{it}\alpha_0 + \theta_i\gamma_0, \lambda_{it}^{-1}\sigma_0)P(LM_{it} = 0|Z_i\beta + X_{it}\alpha_m + \theta_i\gamma, \lambda_{it}^{-1}) \\
&= \prod_{i|LM_{it}=1}\prod_{t=1}^{3} f(Prod_{1it}|Z_i\beta^1 + X_{it}\alpha_1 + \theta_i\gamma_1, \lambda_{it}^{-1}\sigma_1)\Phi(Z_i\beta + X_{it}\alpha_m + \theta_i\gamma, \lambda_{it}^{-1}) \\
&\times \prod_{i|LM_{it}=0}\prod_{t=1}^{3} f(Prod_{0it}|Z_i\beta^0 + X_{it}\alpha_0 + \theta_i\gamma_0, \lambda_{it}^{-1}\sigma_0)\Phi(-Z_i\beta - X_{it}\alpha_m - \theta_i\gamma, \lambda_{it}^{-1})
\end{aligned}$$

$\Phi(.)$ is the standard normal cumulative function and $f(.)$ is a density function for a normal distribution. In the likelihood function $L(.|.)$, $\lambda = \big((\lambda_{1i})_{i=1}^N, (\lambda_{2i})_{i=1}^N, (\lambda_{3i})_{i=1}^N\big)$ and $B = (\beta, \alpha_m, \alpha_1, \alpha_0, \gamma, \gamma_1, \gamma_0)$, $\sigma = (\sigma_1, \sigma_0)$ are the set of parameters.

Considering the complexity of the likelihood function and because we want to estimate a mean effect of the internal labor migration for each household, we use the Bayesian approach that assumes that each parameter of the model has a distribution with non-zero mean and variance. In the setting of treatment analysis, this approach provides a simple way, from the computational point of view, to account for the selection due to unobserved variables. Also, it enriches the analysis by enabling the effect of internal labor migration to be heterogeneous between all households, which is important in terms of public policy implications. In fact, public policies should be more efficient if the specific population who suffer from the labor migration is better targeted.

Recently introduced in the treatment analysis, there are few empirical studies that have attempted to use this approach (See Heckman *et al.* (2012) and Chib & Hamilton (2002) for a review) and, particularly in the literature of internal labor migration, there is no study so far that we are aware of. In the next section, the procedure used to implement the distribution for each component of $B = (\beta, \alpha_m, \alpha_1, \alpha_0, \gamma, \gamma_1)$, the distribution of σ, θ_i and $\lambda_i = (\lambda_{1i}, \lambda_{2i}, \lambda_{3i})$ is described.

3.4.3 Simulation procedure : posterior distribution

First, we define a prior distribution for each parameter :

Parameters	**Prior distribution**
$\sigma = (\sigma_1, \sigma_1, \sigma_0)$	$\mathcal{N}_3\left(\ell_0, L_0\right)$
$B = (\beta, \alpha_m, \alpha_1, \alpha_0, \gamma, \gamma_1)$ B is $k \times 1$ vector	$\mathcal{N}_k\left(b_0, B_0\right)$
θ_i	$\mathcal{N}\left(\mu_0, \nu_0\right)$
λ_{it}, $t = 1, 2, 3$	$\mathcal{G}\left(\frac{\varphi_{0t}}{2}, \frac{\varrho_{0t}}{2}\right)$, $\mathcal{G}(.,.)$ is a gamma density function

Therefore the posterior distribution is given by :

$$\pi(B, \sigma, \theta, \lambda) = \pi(\sigma|\ell_0, L_0)\pi(B|b_0, B_0)\pi(\theta|\nu_0)\pi(\lambda|\frac{\varphi_{0t}}{2}, \frac{\varrho_{0t}}{2})L(Prod_t, LM_t|B, \sigma, \lambda, \theta). \tag{3.5}$$

To sample a distribution for each parameter, unobserved heterogeneity and time-varying scale, we follow the strategy proposed by Chib & Greenberg (1998), Chib & Hamilton (2002) and Lindley & Smith (1972) which can be resumed by the following steps [14]

1. Initialize $B, \sigma, \theta_i, \lambda_{1i}, \lambda_{2i}, \lambda_{3i}$
2. Sample σ from a Metropolis Hastings algorithm. The posterior distribution is $h\left(\sigma|\ell_0, L_0, B, \theta, (\lambda_{1i})_{i=1}^N, (\lambda_{2i})_{i=1}^N, (\lambda_{3i})_{i=1}^N\right) = f(\sigma|\ell_0, L_0) \times L\left(Prod_t, LM_t|B, \sigma, \theta, \lambda\right)$. To sample σ from this distribution, the proposal density function in a multivariate-t student $q(\mu, V)$, where μ and V are respectively the mode and the inverse of the negative of the Hessian matrix of the function $h(.)$ evaluated at the mode.
3. Sample the unobserved component of the vector $H_{it}^* = (MU_{it}^*, Prod_{1it}^*, Prod_{0it}^*)$, $\forall t = 1, 2, 3$.
   - if $LM_{it} = 1$ then sample first $MU_{it}^*|B, \sigma, \theta_i, \lambda_{1i}, \lambda_{2i}, \lambda_{3i}$ from a normal distribution truncated to the interval $]0, +\infty[$. Instead, if $LM_{it} = 0$ then sample $MU_{it}^*|B, \sigma, \theta_i, \lambda_{1i}, \lambda_{2i}, \lambda_{3i}$ from a normal distribution truncated to the interval $] - \infty, 0]$.
   - $\forall\, t \in \{1, 2, 3\}$, $i = 1, \cdots, n$, sample either $Prod_{1it}^*$ or $Prod_{0it}^*$, independently from i and t, from a normal distribution depending on whether LM_{it} is equal to 0 or 1.
4. Sample the set of parameters $B|H_{it}, b_0, B_0, \sigma, (\theta_i)_{i=1}^N, (\lambda_{1i})_{i=1}^N, (\lambda_{2i})_{i=1}^N, (\lambda_{3i})_{i=1}^N$, from a normal distribution.
5. Sample $\theta_i|H_{it}, B, \sigma, (\lambda_{1i})_{i=1}^N, (\lambda_{2i})_{i=1}^N, (\lambda_{3i})_{i=1}^N$. Note that the posterior mean and the posterior variance of θ_i are different from one household to another.
6. Sample $\lambda_{it}|H_{it}, B, \sigma, (\theta_i)_{i=1}^N$, $\forall t = 1, 2, 3$. The posterior parameters are also intrinsic to each household.
7. Repeat steps 2 to 6 to get a full distribution of the posterior distribution.

With the posterior distribution in hand, it is possible to compute various estimators of the effect of labor migration on household's production.

14. See appendix for the details of the sample algorithm

3.4.4 Bayesian treatment effect of internal labor migration

At this stage, we assume that we have the posterior distribution for all parameters, the time-invariant unobserved variables and the time scale variation. From the third step of the simulation algorithm, we have the posterior distribution of the outcome and its counterfactual for each household, that is $(Prod^*_{1it}, Prod^*_{0it})$ for which two of them are observed and the other two are simulated depending on the household's migration status. In our specification, the independence assumption is similar to the one posited in the standard matching analysis and can be expressed as follows :

Assumption 3. $Prod^*_{1it},\ Prod^*_{0it} \perp\!\!\!\perp LM_{it} \mid X_{it},\ Z_{it},\ W_{it},\ \theta_i, \lambda_{it}\ ;\ \forall\, t = 1, 2, 3$

This assumption states that in each period, conditional on the data and on the posterior distributions of θ_i and λ_{it}, the agricultural distribution is independent of the migration decision. Moreover, we assume here that there are no time-varying confounding factors that affect both the migration decision and the production decision. In fact, the time scale variation λ_{it} captures the deviation from the mean variance to correct for the measurement error and heterogeneous effect of some variables. Nevertheless, it goes beyond the assumption made in the potential outcome analysis because the independence assumption accounts for the presence of time-invariant confounders.

Under assumption 3, the effect of migration on household i in period t obtained from the posterior distribution is given by :

$$\rho_{it} = \begin{cases} Prod_{1it} - Prod^*_{0it} & if\ LM_{it} = 1 \\ Prod^*_{1it} - Prod_{0it} & if\ LM_{it} = 0 \end{cases}$$

where $Prod^*_{0it}$ and $Prod^*_{1it}$ are the simulated components of the agricultural productivity and $Prod_{0it}$ and $Prod_{1it}$, the observed and actual agricultural productivity.

By assuming that we have achieved the convergence to the posterior distribution after Q iterations from the simulation process, $\bar{\rho}_{it} = \frac{1}{Q}\sum_{q=1}^{Q} \rho^q_{it}$ is the mean effect for the household i in period t. The denotation of the mean effect for $\bar{\rho}_{it}$ is because we averaged on the posterior distribution. We can also aggregate the effect across time periods by $\bar{\rho}_i = \frac{1}{3Q}\sum_{t=1}^{3}\sum_{q=1}^{Q} \rho^q_{it}$ that is the mean effect of migration on the entire period covered by the analysis.

The Bayesian-average Mean Treatment Effect in period t ($BAMTE_t$) and over the three periods ($BAMTE$) can be expressed as follow :

$$BAMTE_t = \frac{1}{N}\sum_{i=1}^{N} \bar{\rho}_{it} \quad and \quad BAMTE = \frac{1}{N}\sum_{i=1}^{N} \bar{\rho}_i$$

At the same time, we can compute the BAMTE on the treated ($BAMTET_t$) or on the non-treated ($BAMTENT_t$) in period t. They can be obtained from the expressions below, with N_1 and N_0 representing respectively the sample size of migrants-sending HHs and non migrants-sending HHs.

$$\begin{array}{llll} BAMTET_t & = \frac{1}{N_1}\sum_{i|LM_{it}=1}^{N} \bar{\rho}_{it} & BAMTET & = \frac{1}{3N_1}\sum_{t=1}^{N}\sum_{i|LM_{it}=1}^{N} \bar{\rho}_{it} \\ BAMTENT_t & = \frac{1}{N_0}\sum_{i|LM_{it}=0}^{N} \bar{\rho}_{it} & BAMTENT & = \frac{1}{3N_0}\sum_{t=1}^{N}\sum_{i|LM_{it}=0}^{N} \bar{\rho}_{it} \end{array}$$

we also follow Chib & Hamilton (2002) by grouping households depending on their probability of experiencing labor migration in the period t conditional to covariate and unobserved heterogeneity, that is $P_{it} = \Phi(Z_i\beta + X_{it}\alpha_m + \theta_i\gamma, \lambda^{-1}_{it})$. At the q^{th} iteration, $P^q_{it} = \Phi\left(Z_i\beta^q + X_{it}\alpha^q_m + \theta^q_i\gamma^q, (\lambda^q_{it})^{-1}\right)$. Therefore, by discretizing the distribution of probability at each period and at each iteration per decile, we can match households given that random probability inside each decile. Let $D^q_{ht} = \left\{i | P^q_{it} \in \left(\frac{h-1}{10}, \frac{h}{10}\right)\right\}$ be the different groups for $h = 0, 1, \cdots, 10$ and $t = 1, 2, 3$. As pointed out by Chib & Hamilton (2002), the matching of individuals based on P^q_{it} is well defined even at the bottom tail of the distribution because households are self-matched since we are able to compute counterfactual for each household. In the frequentist analysis, the individuals with extreme values of propensity score are generally dropped from the estimation. The average effect in each decile group is given by :

$$\delta_{ht} = \frac{1}{Q}\sum_{q=1}^{Q}\frac{1}{M^q_h}\sum_{i\in D^q_{ht}}\rho^q_{it} \quad and \quad \delta_h = \frac{1}{3Q}\sum_{t}^{3}\sum_{q=1}^{Q}\frac{1}{M^q_h}\sum_{i\in D^q_{ht}}\rho^q_{it}$$

where $M^q_h = |D^q_{it}|$, $|.|$ is the cardinality function.

Moreover, we can also estimate the average effect for a group of households categorized by households' characteristics and head's characteristics. For example, one can be interested in the effect of migration on poorer households or female-headed households. To be more general, the average effect for a group Ω is :

$$\delta_{\Omega t} = \frac{1}{Q\times\Omega}\sum_{i\in\Omega}\sum_{q=1}^{Q}\rho^q_{it} \quad and \quad \delta_{\Omega} = \frac{1}{3Q\times\Omega}\sum_{t}^{3}\sum_{i\in D_t}\sum_{q=1}^{Q}\rho^q_{it}.$$

3.5. Results and discussion

In this section, we discuss the issues of convergence of the posterior distribution. we also comment the parameters that intervene in the likelihood function and the distribution of the effect of the internal labor migration on household agricultural production.

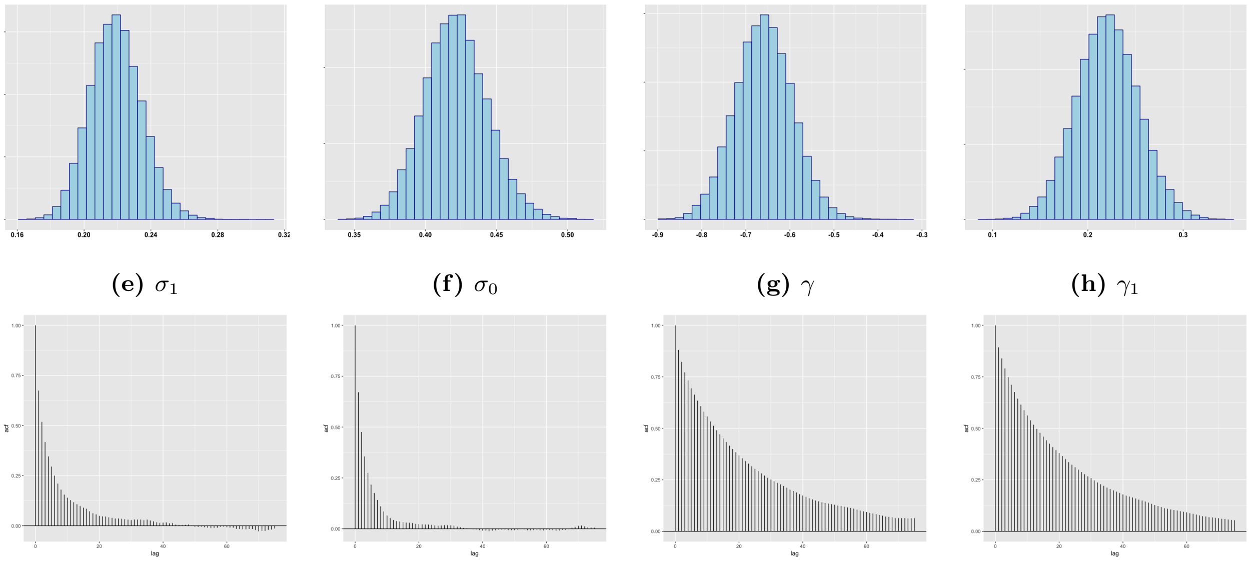
3.5.1 Convergence check and strength of instruments

Convergence check

We follow the algorithm reported in section 3.4.3 to reach the full set of posterior distribution for each parameter. To start with reliable parameters for each prior distribution and to thereby achieve the convergence more rapidly, we first run 4000 iterations with 400 burn-ins from the simulation algorithm on a random training sample made up of a quarter of the initial sample (365 households) as suggested by Chib & Hamilton (2002). The algorithm is then run over 55,000 iterations with 5,000 burn-ins on the entire sample using the parameters obtained from the posterior distribution of the training sample. Figure 3.1 shows the distribution and the autocorrelation functions for each component of the variance-covariance matrix and the two loading factors.

As we can see, the autocorrelation functions reveal that our algorithm mixed well and the convergence is reached rapidly. Indeed, the autocorrelation over iterations is dropped very fast for all variances and a little less faster for the loading factors. Moreover, the posterior distribution of the loading factors in the migration decision equation is negative, meaning that households' time-invariant unobservables tend to decrease the probability of participating in migration. On the other hand, the positive sign of the loading factor in the production function suggests that households' unobservables tend to increase the agricultural production. Therefore, there is a negative correlation between the participation in migration and the investment in agricultural production due to time-invariant unobservables.

FIGURE 3.1 – *Distribution and auto-correlation of the Variance posterior distribution*



(e) σ_1 **(f)** σ_0 **(g)** γ **(h)** γ_1

To go further with the convergence check, we perform both the Geweke means convergence test and the stationary test proposed by Heidelberger & Welch (1983) on all parameters. The latter test states that a Markov chain Monte Carlo (MCMC) converges to the right posterior distribution if the mean of each parameter computed on a proportion of the sample on the top of the iteration is equal to the mean computing on the tail of the distribution. Geweke (1992) proposes using 10% on the top and 50% on the tail of the distributions. The convergence tests reveal that each component of the variance-covariance matrix reaches the convergence.

Strength of instruments : testing the exclusion restriction

We have reported in Tables 3.4 and C.2 the mean and the standard deviation of the posterior distribution of all parameters of the equations 3.2 and 3.3 respectively. For both specifications, the posterior distribution of the two components of the variance-covariance matrix are all significant at the 5% level, meaning that the 95% credibility interval doesn't include zero. Moreover, the variances are higher in the first specification (equation 3.2) than in the specification that allows the instruments of migration to be correlated with the agricultural productivity. This might suggest that some instruments are correlated with the error term. Indeed, in Table 3.2 we can see that the migration rate experienced in 2006 at household and district levels tends to significantly decrease the household productivity. Alternatively, when we do not allow the instruments to intervene in the productivity equation, the migration rate at the district level is strongly positively correlated with the migration decision (see Table 3.3) ; however, when it is included in the productivity equation, it is no longer significant in the migration decision.

3.5.2 Posterior distribution

Here, we comment on the way that different variables affect the households' likelihood to invest in labor migration and also on how households' attributes impact the agricultural productivity. Upper and lower represent respectively the upper and lower bands of the interval of credibility for a p-value of 5%. All results reported here are obtained from estimations on the balanced sample (*sample A*). To test if restraining our analysis to this sample could bias our results and in what way, we also run the model on the *sample B*. We find that the average effect of the labor migration on agricultural productivity is lower than the one obtained from *sample A*. This result is due to the fact that the average agricultural productivity in the *sample B* is lower than the agricultural productivity of its counterpart in *sample A*. It seems that households who are absent from some survey rounds tend to have lower agricultural productivity [15].

Migration likelihood

Unlike other studies that find a concave relationship between the probability to participate in migration and the level of households' wealth, we find that in the rural areas of Uganda this relationship is convex [16]. In fact, the propensity to invest in labor migration decreases with the household's wealth and the squared of the log of the wealth positively affects the migration decision. This means that below a certain level of wealth, households are less likely to participate in migration, perhaps because their expected gain does not offset other costs of migration that they cannot absorb. An increase of 1% of wealth (measured in log) decreases the probability to participate in labor migration by 3.2%. Moreover, households headed by women are more likely to participate in migration ; yet, the marital status of the head does not seem to differently affect the migration participation, only the households headed by singles have lower probability compared with households headed by a monogamously married head.

The composition of the household in terms of members and their educational attainment seem to be the strong determinants of the decision to participate in migration. Households with a higher proportion of children aged 5 or less have a significantly lower incentive to participate in labor migration. On the other hand, households with a higher proportion of adults aged more than 65 years have a higher incentive to get involved in migration. Furthermore, in the literature, some authors argue that once we control for the head's education, the average education among households' members might be a strong instrument for the labor migration. Our results in this paper also confirm this finding as we find that the average education level in households significantly increases the participation in migration and has no effect on the agricultural productivity. Yet, the magnitude of the effect is too small, it only increases the probability by .006. In the same vein, the household's relative deprivation in terms of total hours worked in domestic tasks by adult members seems to be a disincentive to get involved in migration and it has no significant effect on household's agricultural productivity at a 95% credibility interval. Again, this variable appears to be a strong instrument favoring the labor migration.

The propensity to migrate also increases with the head's education and the size of the household, and decreases with the head's age ; particularly, an increase in 1% of household size increases the migration probability by 0.98%. The geographical position of the household's dwelling place has a negative effect on migration participation as expected. We also test the hypothesis that the household's relative wealth deprivation can be a strong push factor for migration once controlling for the level of the wealth. It turns out that the relative wealth deprivation has no effect on the migration propensity. However, all other instruments appear to be intimately related with household migration decisions. In fact, having one more member who has experienced labor migration inside the household at least five

15. Results are available on request
16. A two stage Heckman selection model leads to the same conclusion.

years earlier increases the likelihood of migrating by .045 of percentage points and, an increase of the migration rate at the district level by 1% increases household migration by .033 of percentage points.

Table 3.4 – *Posterior distribution of parameters : Sample A*

Variables	Migration likelihood			Agr. Productivity for MIG-sending HHs			Agr. Productivity for Non MIG-sending HHs		
	mean	**lower**	**upper**	**mean**	**lower**	**upper**	**mean**	**lower**	**upper**
Log wealth	-0.75	-1.07	-0.44	0.05	-0.22	0.33	0.36	0.08	0.64
Log wealth sq. 10	0.06	0.03	0.09	0.01	-0.02	0.04	-0.02	-0.05	0.01
If head is									
Married polygamous	0.04	-0.11	0.19	-0.08	-0.19	0.04	0.07	-0.06	0.19
Divorced or sep	0.06	-0.18	0.30	-0.13	-0.30	0.04	0.01	-0.20	0.21
Widow	-0.02	-0.23	0.20	-0.08	-0.24	0.06	-0.01	-0.20	0.17
Single	-0.67	-1.18	-0.17	-1.30	-1.92	-0.73	0.02	-0.40	0.45
=1 if head is female	0.57	0.40	0.75	0.13	0.01	0.25	-0.27	-0.41	-0.12
log Head age	0.05	-0.15	0.24	0.44	0.29	0.60	0.12	-0.05	0.29
Head Educ	0.03	0.01	0.05	0.02	0.01	0.03	0.00	-0.02	0.02
Children less than 5	-1.15	-1.52	-0.78	-0.04	-0.35	0.27	-0.40	-0.70	-0.10
Individuals aged more than 65	1.10	0.62	1.57	-1.01	-1.46	-0.57	-0.43	-0.87	0.01
log HH size	1.44	1.28	1.60	0.33	0.18	0.48	0.26	0.12	0.41
Average educ in hh	0.04	0.01	0.08	0.00	-0.02	0.03	0.00	-0.02	0.03
Adult domes. lab DP.	-0.10	-0.17	-0.02	0.01	-0.05	0.08	-0.01	-0.06	0.03
Children domes.lab. DP.	0.05	-0.06	0.15	-0.04	-0.12	0.05	0.03	-0.04	0.09
GEO DP.	-0.07	-0.13	-0.02	-0.06	-0.10	-0.03	0.00	-0.03	0.03
Head-MIG	0.01	-0.25	0.27	-0.24	-0.48	-0.01	-0.05	-0.23	0.14
Ethnicity (at country)	-1.46	-2.04	-0.90	1.66	1.12	2.20	0.38	-0.19	0.96
Ethnicity (at district)	-0.05	-0.26	0.15	-0.24	-0.40	-0.09	0.04	-0.13	0.20
Center	0.02	-0.15	0.20	-0.17	-0.29	-0.04	-0.10	-0.25	0.06
East	-0.59	-0.76	-0.43	-0.22	-0.35	-0.10	-0.25	-0.39	-0.10
North	-0.41	-0.58	-0.24	-0.09	-0.24	0.06	-0.05	-0.21	0.11
Hired lab	-	-	-	-0.02	-0.06	0.01	0.03	0.00	0.06
Proxy of HH labor	-	-	-	-0.10	-0.22	0.03	-0.16	-0.30	-0.02
Area	-	-	-	-0.57	-0.64	-0.51	-0.49	-0.55	-0.44
Num crops	-	-	-	0.84	0.73	0.94	0.75	0.67	0.84
labor Agriculture in 1 km radius	-	-	-	0.00	-0.22	0.23	0.08	-0.19	0.34
Nb. of migrants in hh(2005)	0.30	0.20	0.41	-0.08	-0.14	-0.01	0.01	-0.10	0.11
Migration rate in the District	0.22	-0.39	0.83	-0.74	-1.34	-0.14	-2.15	-2.80	-1.51
Log Wealth DP	-0.00	-0.02	0.02	-0.01	-0.03	0.01	-0.00	-0.03	0.03
Intercept	-1.45	-2.03	-0.88	4.60	4.06	5.14	4.93	4.35	5.51
Loading fact.	-0.66	-0.79	-0.53	0.22	0.16	0.29	-	-	-
σ	-	-	-	0.22	0.19	0.25	0.42	0.38	0.47

Note :

- ⋆ Head-MIG equal 1 if the current place of residence of household is different from the place of household's head place of birth, that is when the household's head has permanently migrated in the past (less than 11 years and more than two years).
- ⋆ In the column of the likelihood migration, we have reported estimated parameters and they can't not be interpreted as a marginal effect. Only the sign is significant. However, in the text we will sometimes refer to the marginal effect computed for some variables.

Moreover, the distribution of the unobservables shows that about 75% of migrants-sending HHs have negative value for θ while it is the contrary for non migrants-sending HHs. Besides, Figure C.5 plots the inverse demand of migration as a function of unobservables. It emerges that the inverse demand of migration decreases with the time-invariant unobservables, meaning that the distribution of θ_i captures the unobserved distribution of the cost of migration that doesn't vary over time (see Figure C.5), these are factors that discourage households year after year to not participate in labor migration and which tend to increase the agricultural productivity.

Agricultural productivity

As mentioned above, the total production aggregates the production of maize, beans, coffee, peanuts, bananas and potatoes per hectare. Furthermore, these crops constitute the most cultivated crops within the country and the agricultural productivity is expressed in kilogram per hectare. We can see from Table 3.2 that the investment process in the agricultural productivity differs according to the household migration status. Nevertheless, regardless of the migration status, the agricultural productivity does not change given the marital status and household's relative deprivation in terms of wealth, the domestic hours worked by adult household members and the geo-spatial position of the household dwelling place. On the other hand, both the migrants-sending HHs and non-migrants-sending HHs production is positively affected by the household size and by the number of crops planted while their productivity is negatively affected by the total area plotted. The number of crops managed is usually used as a proxy of how households manage the risk associated with potential events such as the contamination of a particular crop by insects. The results suggest the importance of this factor for households involved in migration. In fact, cultivating one additional crop increases the agricultural productivity by 84% for migrants-sending HHs and 76% for non-migrants-sending HHs. Moreover, an increase of one hectare of planted area decreases the agricultural productivity by 53% and 56% respectively for migrants-sending HHs and non-migrants-sending HHs.

The results suggest that the agricultural productivity of female-headed households is higher in the migrants-sending HHs group than in the non-migrants-sending HHs group. Since female-headed households are more likely to participate in migration, it seems that returns to migration allow them to invest more in the agricultural sector than households headed by males. Moreover, while belonging to the larger ethnic group at the district and country levels is respectively negatively and positively correlated with the migrants-sending HHs production, belonging to an ethnic group has no effect on non-migrants-sending HHs. Mwesigye & Matsumoto (2016) find in the case of Uganda that ethnic diversity tends to lower agricultural production ; instead, the results suggest the contrary since the increase of the share of individuals belonging to the same ethnic group lowers the investment in agricultural productivity.

Regarding the labor hired to work on the farms, we find that the number of days that the households hired people to work on their farms does not affect migrants-sending HHs' productivity ; yet, it increases the non-migrants-sending HHs' productivity. Moreover, while the agricultural productivity increases by 3% with each additional year of education of the migrants-sending HHs head, the head's education has no effect within households not participating in migration. In parallel, the share of household members aged less than 5 years old tends to lower the agricultural productivity among the non-migrants-sending HHs group while it has no effect on the migrants-sending HHs' productivity. This means that non-migrants-sending HHs' agricultural productivity strongly depends on its household composition as compared to the migrants-sending HHs, which is in line with their decision to not participate in migration.

The distribution of the individual's variance scale, $(\lambda_{i1}, \lambda_{i2}, \lambda_{i3})$ reveals that the conditional distributions of the agricultural productivity of the households belonging to the non-migrants-sending households are more heteroscedastic than the counterpart for the migrants-sending households. Regarding the source of heteroscedasticity, this result means that the measurement error in agricultural productivity and the omission of some time-variant unobservables in the productivity equation are more likely to occur for the households that are not involved in labor migration.

3.5.3 Distribution of the effect of labor migration on agricultural production

In this section, we compare each household productivity to the counterfactual productivity obtained by simulation. Prior to this, we first evaluate to what extent the model accurately predicts the actual agricultural productivity distributions. Figure C.4 shows that the predicted distributions are very close to the actual distributions, however, the variance is smaller than the one of the observed distribution. Besides, the average gap between the actual and the predicted value does not attain a production of two kilograms per hectare.

Table 3.5 – *Average effect of internal labor migration on agricultural productivity*

		MIG-sending HHs($BAMET$)					Non-MIG-sending HHs ($BAMENT$)				
Average effect	All	All	Female headed	Male headed	Production>med	Production≤med	All	Female headed	Male headed	Production>med	Production≤med
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Period 1	0.26	0.27	0.39	0.21	0.89	-0.44	0.26	0.41	0.21	-0.50	1.00
	(0.03)	(0.07)	(0.14)	(0.08)	(0.07)	(0.10)	(0.03)	(0.07)	(0.04)	(0.03)	(0.04)
Period 2	0.17	0.54	0.69	0.47	0.99	0.01	0.02	0.23	-0.05	-0.55	0.56
	(0.03)	(0.05)	(0.09)	(0.05)	(0.04)	(0.07)	(0.03)	(0.06)	(0.04)	(0.04)	(0.04)
Period 3	0.16	0.30	0.37	0.26	0.78	-0.24	0.10	0.25	0.05	-0.51	0.68
	(0.02)	(0.04)	(0.08)	(0.05)	(0.04)	(0.06)	(0.03)	(0.06)	(0.03)	(0.03)	(0.03)
Total	0.20	0.37	0.48	0.31	0.89	-0.22	0.13	0.30	0.07	-0.52	0.75
	(0.03)	(0.05)	(0.10)	(0.06)	(0.05)	(0.08)	(0.03)	(0.06)	(0.04)	(0.03)	(0.04)

Note :

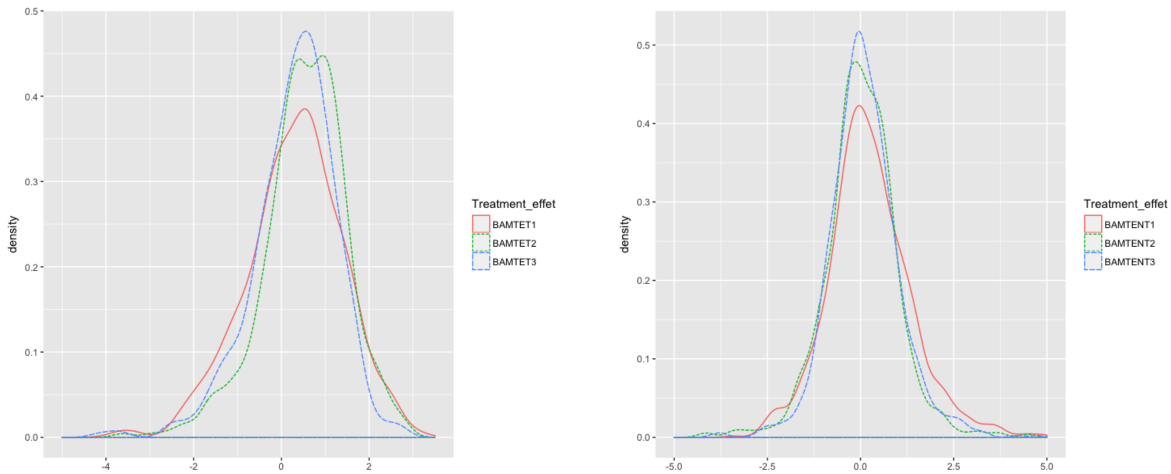
⋆ " *med* " stands for the median of the corresponding distribution.

Although the Bayesian average mean effect ($BAME$), the Bayesian average effect on the treated ($BAMET$) and the Bayesian average effect on the non treated ($BAMENT$) of the internal labor migration are all positives[17] over the entire sample (see columns (1), (2) and (3) of Table 3.5), the distribution of the effect represented in Figure 3.2 and the average effect on the specific subgroup tell us a different story. Indeed, there are households for whom the effect is negative while for other households the effect is positive. This result then suggests that the internal labor migration affects the households' agricultural productivity differently, thus, aggregating the effect over the entire population could hide other facets of the actual impact of the labor migration.

For all periods, the $BAMET$ is higher than the average mean effect on the non treated ($BAMENT$). Furthermore, while the $BAMET$ increases between the first and the second periods and decreases between the second and the third periods, the $BAMENT$ decreases over time. The households headed by women always have a higher return to migration than households headed by men. Therefore, internal labor migration might be a way for female-headed households to increase their agricultural productivity, which is an interesting result since the literature on poverty usually depicts worse livelihood conditions for this group of households. Nevertheless, since female headship is highly correlated with being a migrant household, one might think that this is mostly due to the fact that the husband has migrated. Thus, it is more likely that transfers of the husband to the household left behind are higher. However, only about 3% of female-headed households participating in migration have a husband who has migrated and those have in fact the highest effect on migration.

17. The average effect does not exceed the gain of two kilograms per hectare

Figure 3.2 – *Distribution of the effect of internal migration on agricultural production*



Percentage of HHs with BAMET>0 : 63.0%,76.8%,67.7%

Percentage of HHs with BAMTENT>0 : 56.6%,51.3%,52.7%

Among migrants-sending HHs, there are respectively 63%, 76.8% and 67.7% of the households for which the effect is positive from the first to the third period. Results in columns (5) and (6) reveal that, among households who decide to participate in migration, the larger farmers are those who benefit the most from the internal labor migration compared to the smaller farmers. This implies that the larger farmers are more likely to invest the return to migration in the agricultural sector. Among households belonging to the non-migrants group, we obtain a similar result that larger farmers would have had positive return by investing in the labor migration. Instead, for the smaller farmers, the results suggest that the internal labor migration tends to decrease the agricultural productivity. Everything happens as if smaller farmers bear all of the labor cost induced by the migration and that the major part of the positive return to migration, if there is any, is devoted to non-agricultural productivity.

Regarding the distribution of the effect by region, we have a higher share of migrants-sending HHs who benefit from the labor migration (about 75%) in the western part of Uganda and it is in this region that the agricultural production is also the highest. This result and the fact that the larger farmers are positively affected by the labor migration lead to think that there is no reason to believe that the labor migration will be a great threat to food security and the stability of the food prices in Uganda. Another thing we could look at is if risk diversification (by planting many crops) allows households to increase the benefit of migration. Our results lead to mitigated effects since the households who have planted many crops have higher returns from migration for a period and sometimes they have the smallest return in another period.

To go further, one might be interested in knowing how the effect changes when the likelihood of participating in migration increases. Figure C.6 shows that the effect of migration on agricultural productivity increases with the likelihood of participating in migration. Moreover, the dispersion of the effect within each decile increases when the households are more likely to participate in migration. Furthermore, the average effect of migration is higher for those with negative time-invariant unobservables compared to those with positive unobservables. Therefore, the selection into the participation in migration due to unobservables is also correlated with the average effect and the correlation is positive.

3.6. Conclusion and discussion

To our knowledge few studies have investigated the effect of internal labor migration on agricultural productivity. This might be due to the difficulty of successfully identifying the causal impact of migration since households select themselves into labor migration. Therefore, the migration participation is endogenous to the agricultural productivity. This paper fills a gap in the literature by investigating the distribution of the effects of the temporary internal labor migration on households living in the rural areas of Uganda. The outcome of interest is the agricultural productivity in kilograms per hectare of six crops (maize, beans, coffee, peanuts, bananas and potatoes) planted in all regions of Uganda. We find that the average effect of the internal labor migration on the agricultural productivity is positive ; the average effect on the households participating in migration is around .37 in terms of difference of the logarithm of agricultural productivity, corresponding to a 44% increase in agricultural productivity.

However, as we allow the effect to be heterogeneous between households, it emerges that even if the average effect is positive, there are some households for which the labor migration decreases their agricultural productivity. These households are mostly small farmers, and are therefore more likely to be poor. Moreover, about half of the households that do not participate in the labor migration across rounds would have had higher levels of agricultural productivity if they had participated in the labor migration. This study thus brings new insights into how internal labor migration affects each household' agricultural productivity. This kind of analysis is possible through the introduction of the Bayesian approach in the treatment analysis by allowing to self-match each household. For the decision maker, this kind of analysis makes it possible to target a particular population that absorbs the negative shock of the internal labor migration.

Moreover, the Bayesian framework enables to test the exclusion restriction assumption for the instruments and to account for its violation. Indeed, in the migration literature, some papers use previous participation in migration at the household and community levels as instruments for the current migration participation. The problem is that there is no evidence that these variables are exogenous, meaning that they are not correlated with the households' livelihoods. In this paper we test this hypothesis when the outcome of interest is the agricultural productivity and it emerges that the migration decision taken five years earlier is highly correlated with the current agricultural productivity. Moreover, when we introduce the instruments in the agricultural productivity, the effect of the time-invariant unobserved factors on the likelihood of participating in migration increases significantly while its effect on the agricultural productivity does not change. This may suggest that factors that encouraged households to participate in migration in the past still have an impact on their agricultural production decision today.

We also estimate the average effect within each percentile of the probability of participating in the labor migration. It turns out that the effect of migration on agricultural productivity increases with the likelihood of participating in migration which is good news in terms of achieving optimality.

Economic theory argues that internal labor migration can affect positively the agricultural productivity through the remittances and through useful information that migrant can have from his new social network. Information can be the use of fertilizer, the adoption of new technology or new crop. In fact, remittances send back to household left behind can allow them to buy new land, invest in new crops, hire people to work on the farm. For this first point, our results suggest that internal labor migration helps households to buy new lands, diversify their crops and hire people to work on the farm. As we can see in table 3.2 Migrants-sending households plant on larger land and their crops are more diversify and they hire people to work on farm for many days than Non-sending-Migrants households. This is in line with the *NELM* theory which argues that migration enables households to invest in the agricultural sector.

Although we attempt to limit the bias on the estimation results, they might suffer from many weaknesses. First, we don't allow the parameters in each equation to vary across the survey rounds which could bias the posterior distribution of the counterfactual outcomes since the changes across periods are only due to the households' attributes and not to the way that these attributes affect the agricultural productivity. Second, we don't know the when exactly migrant leaves his household to find job elsewhere. This can also bias our estimation results.

Conclusion

This thesis is interested in internal and international migration. Around the world people leave their countries, regions and families to find better life for themselves and for their relatives. In Canada, migration policies are designed to attract highly skilled workers. However, once arrived in Canada, those skilled workers encountered professional assimilation problem. In fact, their foreign-acquired academic and professional credentials are not fully recognized by local employers. Empirical evidence reveal that the return to foreign-acquired education is highly correlated with the level of development of migrants' origin country. As a result, migrants have difficulties finding a job that matches their highest diploma. In order to get access to better jobs, some migrants decide to invest in host country education and obtain a university or college degree. So far, existing studies are inconclusive about the earnings gains of the host-country acquired education over the foreign-acquired one. The first two chapters of this thesis contribute to this literature by estimating the causal effect of education acquired in Canada on the frequencies of jobs and their durations and, on the weekly earnings profile. Particularly, our focus in on the skilled workers who enter the province of Quebec with permanent residence status.

By means of multiple states and multiple spells model, the transitions between four different states : qualified job, unqualified job, unemployment and training are estimated. Our results suggest that immigrants from rich countries do not need to invest more in Quebec education since the return is relatively low. On the other hand, immigrants from poor countries, although highly skilled, benefit greatly from such training in the long-term as this facilitates their transition to qualified and unqualified jobs and out of unemployment. Moreover, we use a Bayesian treatment analysis to estimate the distribution of the gains of Quebec-acquired education over the foreign one. We find that there is negative relative gain of Quebec-acquired education for almost all migrants. Besides, the magnitude of the gains are different from one migrant to another. Furthermore, the gains tend to decrease with the likelihood of investing in Quebec-acquired education and with the level of *ability*. More importantly, the return to Quebec-acquired education is higher for migrants with previous high quality job. This suggests that local employers evaluate the migrants' productivity based on the quality of previously held jobs.

Both studies reveal the importance of taking into account the unobserved *essential heterogeneity* in order to minimize if not eliminate the selection bias. The importance of unobserved factor such as *ability* or social network in job search issue is well known in the literature. For migrants, their social network can be an important source of information about job search process and job opportunities.

Our results then send a clear message that investing in host country education can not alone help migrant to find high quality job. In this case, policymarkers have to invest more in migrants job search program. We suggest to design program that assists migrants through their job search process not only few months after their arrival but through their entire professional assimilation process. Our results show that only a quarter and a half of individuals respectively from Africa and from Europe are able to find a job that matches their education after nine years in Quebec. Moreover, government should educate employers about the potential productivity of migrants based on their foreign-acquired education and work experience. It has been found that since migrants are self-selected group, they may be more able and more motivated than the natives (Borjas, 1987). This should be an asset for employers. Furthermore, since high quality jobs are rare for both natives and migrants, migration policies should target immigrants willing to work in sector of activities for which external labor supply is needed. For instance, policymakers should facilitate and promulgate program allowing some industries, where the need of foreign workers is urgent, to hire temporarily or permanently foreign workers.

Developing countries are more concerned about the internal labor migration since it can have positive or negative effect on both the destination and origin communities of migrants. In the third chapter of this thesis, we investigate the causal effect of household participation in internal labor migration on agricultural productivity. The originality of this study is that we estimate the distribution of the effect by accounting for the unobserved *essential heterogeneity*. We use to that aim a Bayesian treatment analysis and find that although, on average, internal labor migration positively affects agricultural productivity, there are households for which the effect is negative. Those for which the effect is negative are mostly small farmers and are therefore more likely to be poor and more likely to be subject to local price volatility. The negative effect on the small farmers can be explained by the fact that hiring people to work in the farm become costly and that these farmers are more involved in subsistence farming. In addition, our results lead to the conclusion that previous migration rates, widely used in the literature as an instrument for the decision to participate in migration, are not exogenous to agricultural productivity but are intimately correlated with that.

This chapter suggests that internal labor migration is optimal from an economic point of view since it increases the agricultural productivity of farmers more likely to contribute significantly to the growth of the sector while it decreases the production of households more involved in subsistence farming. Therefore, policymakers should put in place programs to make gains from migration more profitable. We propose to the government to create a fund in agriculture just like the Green Climate Fund (GCF) using immigrants (internal and international) as the main sources of funding. The existence of such a fund should be promoted for farmers and any person with innovative ideas in the agricultural sector. For small farmers, policymarker should invest in their children education and make cash transfer in order to meet their basic needs.

Bibliographie

Abbott, Michael G., & Beach, Charles M. 2011 (August). *Immigrant Earnings Differences Across Admission Categories and Landing Cohorts in Canada.* CLSRN, Working paper No. 81. 7

Abbring, Jaap H. 2003. Dynamic Econometrics Program Evaluation. *IZA Discussion paper*, **804**.

Abbring, Jaap H., & Van Berg, Gerard J. 2003. The Nonparametric Identification of Treatment Effects in Duration Models. *Econometrica*, **71**(05), 1491–1517. 42

Adaku, Abigail A. 2013. The effect of rural migration on agricultural production in the Northern region of Ghana. *Journal of Agricultural Science and Applications*, **2**(4), 193–201.

Akresh, Ilana R. 2007. U.S. Immigrants' labor market adjustment : Additional Human capital investment and Earnings Growth. *Empirical Economics*, **44**(4), 865–881. 32

Albert, James H., & Chib, Siddhartha. 1993. Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, **88(422)**. 50

Arkes, Jeremy. 2010. Using Unemployment Rates as Instruments to Estimate Returns to School. *Southern Economic Journal*, **76(3)**, 711–722.

Asankha, Pallegedara, & Takashi, Yamano. 2011. Impacts of Universal Secondary Education Policy on Secondary School Enrollments in Uganda. *Journal of Accounting, Finance and Economics*, **1**, 16–30. 66

Baker, Michael, & Melino, Angelo. 2000. Duration Dependence and Nonparametric Heterogeneity : A Monte Carlo Study. *Journal of Econometrics*, **96**(2), 357–393. 20

Bandiera, Oriana, & Rasul, Imran. 2002. Social Network and Technology Adoption in Northern Mozambique. *DEDPS-London school of Economics and Political Science*, **35**.

Baudassé, Thierry, & Bazillier, Rémi. 2014. Gender Inequality and Emigration : Push factor or selection process. *International Economics*, **139**, 19–47.

Benzakour, Chakib, Lechaume, Aline, Castonguay, Marie-Hélène, & Santos, Paula. 2013. Enquête auprès des Immigrants de la Catégorie des Travailleurs Qualifiés . Portrait des répondants au moment de l'enquête. *Québec, Ministère de l'Immigration et des communautés culturelles et mininstère de l'Emploi et de la Solidarité sociale, Gouvernement du Québec,*, **54p**. 9, 35

Blasco, Sylvie, Crépon, Bruno, & Kamionka, Thierry. 2010. Evaluation of the Impact of Training on Individual Labor Market Transitions. *Mimeo, Crest, Paris*, **vol 44, No 4**.

Borjas, George J. 1987. Self-Selection and the Earnings of Immigrants. *The American Economic Review*. 90

Borjas, George J. 2000. Introduction to " Issues in the Economics of Immigration". *NBER*, 1–14. 7

Boudarbat, Brahim, & Boulet, Maude. 2007. Détérioration des salaires des nouveaux immigrants au Québec par rapport à l'Ontario et à la Comlombie-Britannique entre 1991 et 2007. *IRRP Choix*, **13 n-7. Montréal : Institut de recherche en Politiques Publiques**. 7

Bratsberg, Bernt, & Ragan, James F. 2002. The Impact of Host-Country Schooling on Earnings : A Study of Male Immigrants in the United States. *The Journal of Human Resources*, **37**(1), 63–105. 2, 32

Brière, Simon, Fortin, Bernard, & Lacroix, Guy. 2017. *Discrimination à l'embauche des candidates d'origine maghrébine dans la région de la Capitale-Nationale*. Forthcoming, L'Actualité Économique. 13

Brooks, Stephen P., & Gelman, Andrew. 1998. General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, **7(4)**, 434–455.

Brouillette, Dany, & Lacroix, Guy. 2011. Assessing the Impact of a Wage Subsidy for single parents on Social Assistance. *Canadian Journal of Economics*, **vol 44, No 4**. 20

Bryan, Gharad, Chowdhury, Shyamal, & Mobarak, Ahmed M. 2014. Underinvestment in a Profitable Technology : the Case of Seasonal Migration in Bangladesh. *Econometrica*, **82**(5), 1671–1748.

Card, David. 2001. Estimating the return to Schooling : Progress on Some Persistent Econometric Problems. *Econometrica*, **69(5)**, 1127–1160. 33, 73

Carneiro, Pedro, Hansen, Karsten T., & Heckman, James J. 2003. Estimating Distributions of Treatment Effetcs with an Apllication to Return to schooling and measurenment of the effects of uncertainty on college choice. *International Econometric Reviews*, **44**, 361–422. 34, 37, 63

Carrasco, Raquel, & García-Pérez, J. Ignacio. 2008. Unemployment Duration among Immigrants and Natives : Unobserved heterogenéity in a multi-spell duration Model. *Working Paper. Economic*, **8 No 33**.

Chesher, Andrew, & Lancaster, Tony. 1983. The Estimation of Models of Labour Market Behaviour. *The Review of Economics Studies*, **50**(4), 609–624. 8

Chib, Siddhartha, & Greenberg, Edward. 1998. Analysis of multivariate probit models. *Biometrika*, **85**(2), 347–361. 78, 106, 113

Chib, Siddhartha, & Hamilton, Barton H. 2000. Bayesian Analysis of Cross-section and clustered data treatment models. *Journal of Econometrics*, 25–50. 50

Chib, Siddhartha, & Hamilton, Barton H. 2002. Semiparametric Bayes analysis of longitudinal data treatment models. *Econometric Reviews*. 34, 50, 51, 74, 77, 78, 79, 80, 107

Chiswick, Barry R., & Miller, Paul W. 2007. The International Transferability of Immigrants' Human Capital Skills. *IZA Discussion Paper*, **2670**. 1, 8

Citizenship, & Immigration Canada. 2014. Canada – Résidents permanents selon le sexe et la catégorie, 1989-2013. *Statistiques Canada*. 2, 9

Cloutier, Luc, Benard, Paul, & Diane-Gabrielle, Tremblay. 2009. Job Quality and Gender Inequality : Key Changes in Québec over the Last Decade. *Studies in Social Justice*, **3, Issues 2**, 191–212.

Cohen-Golden, Sarit, & Eckstein, Zvi. 2008. Labor mobility of Immigrants : Training, Experience, Language, and Opportunities. *International Economic Review*. 32

Conley, Timothy G., Hansen, Christian B., & Rossi, Peter E. 2012. Plausibly Exogenous. *The Review of Economics and Statistics*, **94**(1), 260–272. 64, 74

Coulombe, Serge, Grenier, Gilles, & Nadeau, Serge. 2014a. Human Capital Quality and the Immigrant wage gap. *IZA Journal of Migration*, **3**. 1, 7, 8

Coulombe, Serge, Grenier, Gilles, & Nadeau, Serge. 2014b. Quality of work experience and economic development-estimates using Canadian Immigrant data. *Journal of Human Capital*, **8**. 1, 7, 8, 32, 55

Cox, D. R. 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, **34 No. 2**, 187–220.

Crump, Richard K., Hotz, V. Joseph, Imbens, Guido W., & Mitnik, Oscar A. 2009. Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, **96**(1), 187–199. 46

Davis, Benjamin, Carletto, Gero, & Winters, Paul C. 2010. Migration, Transfers and Economic Decision Making among Agricultural Households : an Introduction. *The Journal of Development Studies*, **46**, 1–13. 65

De Brauw, Alan. 2010. Seasonal Migration and Agricultural Production in Vietnam. *The Journal of Development Studies*, **46**, 114–139. 63, 65

De Haan, Arjan. 1999. Livelihoods and poverty The role of migration a critical review of the migration literature. *The Journal of Development Studies*, **36(2)**, 1–47. 63, 65

Deininger, Klaus. 2003. Does cost of schooling affect enrollment by the poor ? Universal primary education in Uganda. *Economics of Education Review*, **22**, 291–305.

Dinkelman, Taryn, & Mariotti, Martine. 2016. The long Run Effects of Labor Migration on Human Capital Formation in Communities of Origin. *Cambridge, MA : National Bureau of Economic Research (NBER)*.

Djuikom, Marie A., & Lacroix, Guy. 2018. Dynamic Causal Effects of Post-Migration Schooling on Labour Market Transitions. *Working paper(CEDIA)*. 34, 35, 59

Esses, Victoria M., Dovidio, John F., Jackson, Lynne M., & Armstrong, Tamara L. 2001. The Immigration Dilemma- The Role of Perceived Group Competition, Ethnic Prejudice, and National Identity. *Journal of Social Issues*, **57 No 3**, 389–412.

Garip, Filiz. 2008. Social Capital and migration : how do similar Resources Lead to Divergent Outcomes. *Demography*, **45**(3), 591–617. 65, 74

Garlick, Julia, Leibbrandj, Murray, & Levinsohm, James. 2016. Individual Migration and Household Incomes. *National Bureau of Economic Research*. 65

Geweke, John. 1992. Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments. *In Bayesian Statistics 4 (ed JM Bernardo, JO Berger, AP Dawid, and AFM Smith). Clarendon Press, Oxford, UK*. 53, 81

Glick, Peter, & Sahn, David E. 2006. The Demand for Primary Schooling in Madagascar : Price, Quality, and the choice between Public and Private Providers. *Journal of Development Economics*, **79**(2), 118–145.

Gonzales, Rolando, Aranda, Patricia, & Mendizábal, Aranda J. 2016. Is Microfinance Truly Useless for Poverty Reduction and Women Empowerment ? A Bayesian Spatial-Propensity Score Matching Evaluation in Bolivia. *Partnership for Economic Policy (PEP)*, **10**.

Gouriéroux, Christian, & Monfort, Alain. 1991. Simulation based econometrics in models with heterogeneity. *Annales d'économie et de statistique*. 20

Gouriéroux, Christian, & Monfort, Alain. 1996. *Simulation-based Econometric Methods*. Oxford University Press. 20

Grant, Mary L. 1999. Evidence of new Immigrant Assimilation in Canada. *Canadian Journal of Economies*, **32 No 4**, 930–955. 1, 7

Green, David A., & Worswick, Christopher. 2010. Entry earnings of immigrant men in Canada : The roles of labour market entry effects and returns to foreign experience. *Canadian immigration : Economic evidence for a dynamic policy environment*, 77–110.

Gritz, R. Mark. 1993. The impact of training on the frequency and duration of employment. *Journal of Econometrics*, **57**, 21–51.

Hansen, Jörgen. 2000. The duration of Immigrants Unemployment Spells :Evidence from Sweden. *IZA Discussion Paper*, **155**. 7, 8

Heckman, James J., & Singer, Burton. 1984. A method for Minimizing the Distribution Asumptions in Econometrics Models for Duration Data. *Econometrics*, 271–320. 20

Heckman, James J., Ichimura, Hidehiko, Smith, Jeffrey, & Todd, Petra. 1998a. Characterizing Selection Bias Using Experimental Data. *Econometrica*, **66**(05), 1017–1098.

Heckman, James J., Ichimura, Hidehiko, Smith, Jeffrey, & Todd, Petra. 1998b. Characterizing Selection Bias Using Experimental Data. *Econometrica*, **66**(05), 1017–1098.

Heckman, James J., Urzua, Sergio, & Vytlacil, Edward. 2006. Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, **LXXXVIII**(3).

Heckman, James J., Lopes, Hedibert F., & Piatek, Rémi. 2012. Treatment Effects : A Bayesian Perspective. *Econometric Reviews*. 33, 74, 77

Heidelberger, Philip, & Welch, Peter D. 1983. Simulation run length control in the presence of an initial transient. *Opns Res.*, **31**, 1109–1144. 81

Hirano, Keisuke, Imbens, Guido W., & Ridder, Geert. 2003. Efficient Estimation of Average Treatment Effets Using the Estimated Propensity Score. *Econometrica*, **71**(04), 1161–1189.

Hou, Feng, & Lu, Yuqian. 2017. International students, immigration and earnings growth : the effect of a pre-immigration host-country university education. *IZA Journal of Development and Migration*, **7 :5**. 2, 32, 55, 60

Huylebroeck, Lisa, & Kristof, Titeca. 2015. Universal Secondary Education(USE) in Uganda : blessing or curse ? The impact of USE on educational attainment and performance. *In : Reyntjens,F., Vandeginste, S. and M. Verpoorten(eds) L'Afrique des Grands Lacs : Annuaire 2014-2015. Antwerp :University Press Antwerp,*, 349–372.

Imbens, Guido W. 2004. Nonparametric Estimation of Average Treatment Effects Under Exogeneity : A Review. *The Review of Economics and Statistics*, **86(1)**, 4–29.

Imbens, Guido W., & Angrist, Joshua D. 1994. Identification and Estimation of Local Average Treatment effects. *Econometrica*, **62**(2), 467–475.

Imbens, Guido W., & Wooldridge, Jeffrey M. 2009. Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, **47**(1), 5–86.

International Organization for Migration, & Uganda Rapid Migration Profile technical working group. 2015. Migration in Uganda : A rapid country profile 2013. *International Organization for Migration(IOM)*. 65

Jagger, Pamela, Shively, Gerald, & Arinaitwe, Arthur. 2012. Circular migration, small-scale logging, and household livelihoods in Uganda. *Popul Environ*, **34**, 235–256. 66

Kanas, Agnieszka, & van Tubergen, Frank. 2009. The Impact of Origin and Host Country Schooling on the Economic Performance of Immigrants. *Social Forces*, **88**(2), 893–915. 32, 55

Konseiga, Adama. 2007. Household Migration Decisions as Survival Strategy : The Case of Burkina Faso. *Journal of African Economies*, **16**, 198–233. 73

Kuhn, Randall. 2015. Internal Migration : Developing Countries. *International Encyclopedia of the Social and Behavioral Sciences*, **12**(2nd edition), 433–442. 65

Lacroix, Guy. 2013. *L'intégration en emploi des immigrants de la catégorie des travailleurs qualifiés au Québec, 2002–2009.* https://www.mess.gouv.qc.ca/publications/pdf/ADMIN_ETQ_4.pdf. 7, 14, 30

Larson, Donald F., Savastano, Sara, Murray, Siobhan, & Palacios-López, Amparo. 2015. Are Women Less Productive Farmers ? How Markets and Risk Affect Fertilizer Use, Productivity, and Measured Gender Effects in Uganda. *Policy Research Working Paper. Washington, DC : World Bank Group*, **7241**. 76

Lechner, Michael. 2009. Sequential Causals Model for the Evaluation of Labor Market Programs. *Journal of Business and Economic statistics*, **27**(1), 71–83. 34, 43, 44, 73

Lechner, Michael, & Miquel, Ruth. 2001. *Potential Outcomes Approach to Dynamic Programme Evaluation-Part I : Identification* . IZA Discussion Papers. Institute for the Study of Labor (IZA). 42, 45

Lindley, Dennis V., & Smith, Adrian F. M. 1972. Bayes Estimates for the Linear Model. *Journal of the Royal Statictics Society. Series B (Methodological)*, **34(1)**, 1–41. 34, 78, 107

Lucas, Robert E. B. 1987. Emigration to south Africa's Mines. *America Economic Review*, **77**(3), 313–330. 73

Lynch, Lisa M. 1986. The Youth labor market in the 80s : Determinants of Re-employment probabilities for Young Men and Women. *National, Bureau of Economic Research*, **No 2021**. 18

Mckenzie, David, & Rapoport, Hillel. 2007. Network Effects and the Dynamics of Migration and Inequality : Theory and evidence from Mexico. *Journal of Development Economics*, **84**, 1–24. 74

Mckenzie, David, Stillman, Steven, & Gibson, John. 2010. How Important is Selection : Experimental vs Non-experimental Measures of the Income Gains from Migration. *Journal of the European Economic Association*, **8**(4), 913–945.

Mendola, Mariapia. 2008. Migration and Technological Change in Rural Household : Complements or Substitutes ? *Journal of Development Economics*, **85**, 150–175. 63, 65, 74

Mu, Ren, & Van de Walle, Dominique. 2011. Left behind to farm ? Women's labor re-allocation in rural china. *Labour Economics*, **18**, 83–97. 64, 65

Muto, Megumi. 2009. The impacts of mobile phone coverage expansion and personal networks on migration : evidence from Uganda. *Prepared for presentation at the international association of Agricultural Economists conference, Beijing, China, August 16-22*. 65, 66

Mwesigye, Francis, & Matsumoto, Tomoya. 2016. The Effect of Population Pressure and Internal Migration on Lands Conflicts : Implications for Agricultural Productivity in Uganda. *World Development*, **79**(01), 25–39. 63, 65, 84

Ndiaye, Ameth S., Niang, Oumoul K., Ndione, Ya C., & Dedehouanou, Sessinou E. 2016. Migration, Remittances, Labor Market and Human capital in Senegal. *Partnership for Economic Policy (PEP)*, **10**.

Nordin, Martin. 2007. Immigrants' return to schooling in Sweden. *Institute for Labour Market Policy Evaluation(IFAU)*, **12**. 32

Orazem, Peter F., & King, Elizabeth M. 2008. *Schooling in Developing Countries : The Roles of Supply, Demand and Government Policy.* Book chapter. Handbook of Development Ecconomics.

Oreopoulos, Philip. 2011. Why Do Skilled Immigrants Struggle in the Labor Market ? A Field Experiment with Thirteen Thousand Résumés. *American Economic Review : Public Policy*, **3**, 148–171. 13

Penny, Alan, Ward, Michael, Read, Tony, & Bines, Hazel. 2008. Education sector reform : The Ugandan experience. *International Journal of Educational Development*, **28**, 268–285.

Renaud, Jean. 2005. Limites de l' Accès à l' Intégration des Immigrants au Québec : Quelques exemples à partir d'enquêtes. *Santé, Société at Solidarité*, **No 1**. 14, 30

Robins, James M., Greenland, Sander, & Hu, Fu-Chang. 1999. Estimation of causal effect of a time-varying exposure on the marginal Mean of a repeated Binary outcome. *Journal of the American Statistical Association*, **94**, 687–700.

Rutaremwa, Gideon. 2011. Internal and International Labor Migration in Uganda : The contribution of Remittances to Household Livelihoods. *Makerere University, presented in African conference on Population*. 65, 66

Schaafsma, Joseph, & Sweetman, Arthur. 2001. Immigrant Earnings : Age at Immigration Matters. *The Canadian Journal of Economics*, **34 No 4**.

Shannon, Claude E. 1948. A Mathematical Theory of Communication. *The Bell System Technical Journal*, **27**, 379–423,623–656.

Sianesi, Barbara. 2004. An Evaluation of the Swedish System of Active Labor market Programs in the 1990s. *The Review of Economics and Statistics*, **86**(01), 133–155. 34

Skuterud, Mikal, & Su, Mingcui. 2012. The influence of measurement error and unobserved heterogeneity in estimating immigrant returns to foreign and host-country sources of human capital. *Empirical Economics*, **43**(3), 1109–1141. 3, 32, 33

Ssewamala, Fred M., Wang, Julia S.-H., Karimli, Leyla, & Nabunya, Proscovia. 2011. Strengthening Universal Primary Education in Uganda : The potential role of an asset-based development policy. *International Journal of Education Development*, **31**, 472–477. 66

Stark, Oded, & Bloom, David E. 1985. The New Economics of Labor Migration. *American Economics Review*, **75**(2), 173–178. 4, 64

Stark, Oded, & Taylor, J. Edward. 1989. Relative Deprivation and International Migration. *Demography*, **26**(1), 1–14.

Stark, Oded, & Taylor, J. Edward. 1991. Migration Incentives, Migration types : Role of relative deprivation. *The Economic Journal*, **101**, 1163–1178. 65, 75

Sweetman, Arthur. 2004. Immigrant Source Country Educational Quality and Canadian Labour Market Outcomes Catalogue. *Statistics Canada,Ottawa*, **234**.

Sweetman, Arthur, & Warman, Casey. 2013. Canada's Immigration Selection System and Labour Market Outcomes. *Canadian Public Policy*, **39**(May), 141–160. 7

Uganda Bureau of Statistics, UBOS. 2013. *Uganda Panel National Survey-Wave III report*. Tech. rept. 68

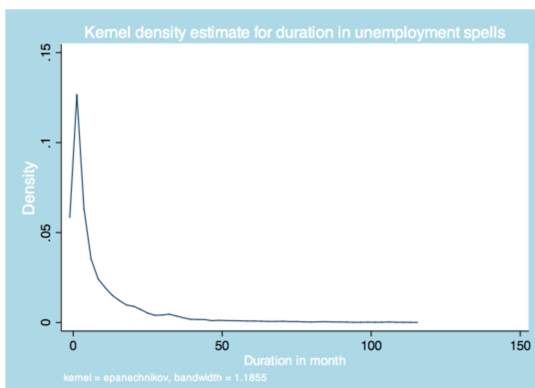
Uhlendorf, Arne, & Zimmermann, Klaus F. 2006. Unemployment dynamics among migrants and natives. *IZA,Institute for the Study of Labor*, **2299**. 8

Vasilaky, Kathryn N., & Leonard, Kenneth L. 2016. As Good as the Network they Keep ? : Improving the Outcomes Through the Weak Ties in Rural Uganda. *Forthcoming, Economic Development and Cultural Change*.

- Annexe A -

A.1. Density function of duration in unemployment spells

Figure A.1 – *Density function of duration in unemployment spells*



A.2. Proofs

A.2.1 Correlation between unobserved heterogeneities

Taking the characterization of unobserved heterogeneity given in section 1.3.5, we have,

$$\omega_j = \mu_j + \nu_j \theta^*,$$

where

$$\theta^* = \begin{cases} \theta_1 & with \quad probability \quad p_1 \\ \theta_2 & with \quad probability \quad p_2 \\ \theta_3 & with \quad probability \quad (1 - p_1 - p_2), \end{cases}$$

and $cov(\omega_j, \omega_k) = \nu_j \nu_k var(\theta^*)$, $var(\omega_j) = \nu_j^2 var(\theta^*)$. Therefore,

$$\rho_{jk} \quad = \tfrac{cov(\omega_j, \omega_k)}{\sigma_j \sigma_k} = \tfrac{\nu_j \nu_k var(\theta^*)}{\sqrt{\nu_j^2 var(\theta^*)}\sqrt{\nu_k^2 var(\theta^*)}} = 1$$

A.2.2 Instantaneous probability

In the case where we have three states, that is $k \in e, u, e'$, the instantaneous probability to be in state k is given by,

$$\begin{aligned} P(D_{e'}(t) = 1) = P\left(D_{e'}(t - \delta t) = 1\right) \times [1 - \lambda_{e'u}(t)\delta t] + P(D_u(t - \delta t) = 1) \times \lambda_{ue'}(t)\delta t \\ + P(D_r(t - \delta t) = 1) \times \lambda_{re'}(t)\delta t \end{aligned}$$

Arranging gives,

$$\begin{aligned} \frac{P(D_{e'}(t) = 1) - P(D_{e'}(t - \delta t) = 1)}{\delta t} = -P\left(D_{e'}(t - \delta t) = 1\right) \times \lambda_{e'u}(t) \\ + P(D_u(t - \delta t) = 1) \times \lambda_{ue'}(t) + P(D_u(t - \delta t) = 1) \times \lambda_{re'}(t) \end{aligned}$$

Letting δt pass to zero, we obtain,

$$\frac{dP(D_{e'}(t) = 1)}{dt} = -P\left(D_{e'}(t) = 1\right) \times \lambda_{e'u}(t) + P(D_u(t) = 1) \times \lambda_{ue'}(t) + P(D_r(t) = 1) \times \lambda_{re'}(t)$$

We get similar expression for $P(D_e(t) = 1)$ in the same way, given by,

$$\begin{aligned} P(D_e(t) = 1) = P\left(D_e(t - \delta t) = 1\right) \times [1 - \lambda_{eu}(t)\delta t]\,[1 - \lambda_{er}(t)\delta t] + P(D_u(t - \delta t) = 1) \times \lambda_{ue}(t)\delta t \\ + P(D_r(t - \delta t) = 1) \times \lambda_{re}(t)\delta t \end{aligned}$$

$$\begin{aligned} = P\left(D_e(t - \delta t) = 1\right) \times \left[1 - \lambda_{eu}(t)\delta t - \lambda_{er}(t)\delta t + \lambda_{er}(t)\lambda_{eu}(t)\delta t^2\right] + P(D_u(t - \delta t) = 1) \times \lambda_{ue}(t)\delta t \\ + P(D_r(t - \delta t) = 1) \times \lambda_{re}(t)\delta t \end{aligned}$$

As before, we get :

$$\begin{aligned} \frac{dP(D_e(t) = 1)}{dt} = -P\left(D_e(t) = 1\right) \times [\lambda_{er}(t) + \lambda_{eu}(t)] + P(D_u(t) = 1) \times \lambda_{ue}(t) + \\ -P(D_r(t) = 1) \times \lambda_{re}(t) \end{aligned}$$

Once again, we get similar expression for $P(D_u(t) = 1)$ in the same way, given by,

$$\begin{aligned} P(D_u(t) = 1) = P\left(D_u(t - \delta t) = 1\right) \times [1 - \lambda_{ue}(t)\delta t]\,[1 - \lambda_{ue'}(t)\delta t]\,[1 - \lambda_{ur}(t)\delta t] + \\ P(D_e(t - \delta t) = 1) \times \lambda_{eu}(t)\delta t + P(D_e'(t - \delta t) = 1) \times \lambda_{e'u}(t)\delta t + P(D_r(t - \delta t) = 1) \times \lambda_{ru}(t)\delta t \end{aligned}$$

As before, by arranging and letting δt tend to zero, we get :

$$\begin{aligned}\frac{dP(D_u(t)=1)}{dt} &= -P\left(D_u(t)=1\right)\times[\lambda_{ue}(t)+\lambda_{ue'}(t)+\lambda_{er}(t)]+P(D_e(t)=1)\times\lambda_{eu}(t)\\ &+P(D'_e(t)=1)\times\lambda_{e'u}(t)+P(D_r(t)=1)\times\lambda_{ru}(t)\end{aligned}$$

and ,

$$\begin{aligned}\frac{dP(D_r(t)=1)}{dt} &= -P\left(D_r(t)=1\right)\times[\lambda_{re}(t)+\lambda_{re'}(t)+\lambda_{ru}(t)]+P(D_e(t)=1)\times\lambda_{er}(t)\\ &+P(D'_e(t)=1)\times\lambda_{e'r}(t)+P(D_u(t)=1)\times\lambda_{ur}(t)\end{aligned}$$

As the results we have obtained the equation system :

$$\begin{pmatrix}\frac{d}{dt}P(D_{e'}(t)=1)\\ \frac{d}{dt}P(D_e(t)=1)\\ \frac{d}{dt}P(D_u(t)=1)\\ \frac{d}{dt}P(D_r(t)=1)\end{pmatrix}=\begin{pmatrix}A_{e'}(t) & B_{e'}(t) & C_{e'}(t) & D_{e'}(t)\\ B_e(t) & A_e(t) & C_e(t) & D_e(t)\\ B_u(t) & C_u(t) & A_u(t) & D_u(t)\\ B_r(t) & C_r(t) & D_r(t) & A_r(t)\end{pmatrix}\times\begin{pmatrix}P(D_{e'}(t)=1)\\ P(D_e(t)=1)\\ P(D_u(t)=1)\\ P(D_r(t)=1)\end{pmatrix}\tag{A.1}$$

With,

$$\begin{cases}A_{e'}(t)=-\lambda_{e'u}(t)\\ B_{e'}(t)=0\\ C_{e'}(t)=\lambda_{ue'}(t)\\ D_{e'}(t)=\lambda_{re'}(t)\\ A_e(t)=-(\lambda_{eu}(t)+\lambda_{er}(t)\\ B_e(t)=0\\ C_e(t)=\lambda_{ue}(t)\\ D_e(t)=\lambda_{re}(t)\\ A_u(t)=-(\lambda_{ue}(t)+\lambda_{ue'}(t)+\lambda_{er}(t))\\ B_u(t)=-\lambda_{e'u}(t)\\ C_u(t)=\lambda_{eu}(t)\\ D_u(t)=\lambda_{ru}(t)\\ B_r(t)=\lambda_{e'r}(t)\\ C_r(t)=\lambda_{er}(t)\\ D_r(t)=\lambda_{ur}(t)\\ A_r(t)=-(\lambda_{re}(t)+\lambda_{re'}(t)+\lambda_{ru}(t))\end{cases}$$

- Annexe B -

B.1. Tables and Figures

Figure B.1 – *Probability of taking or not taking a training*

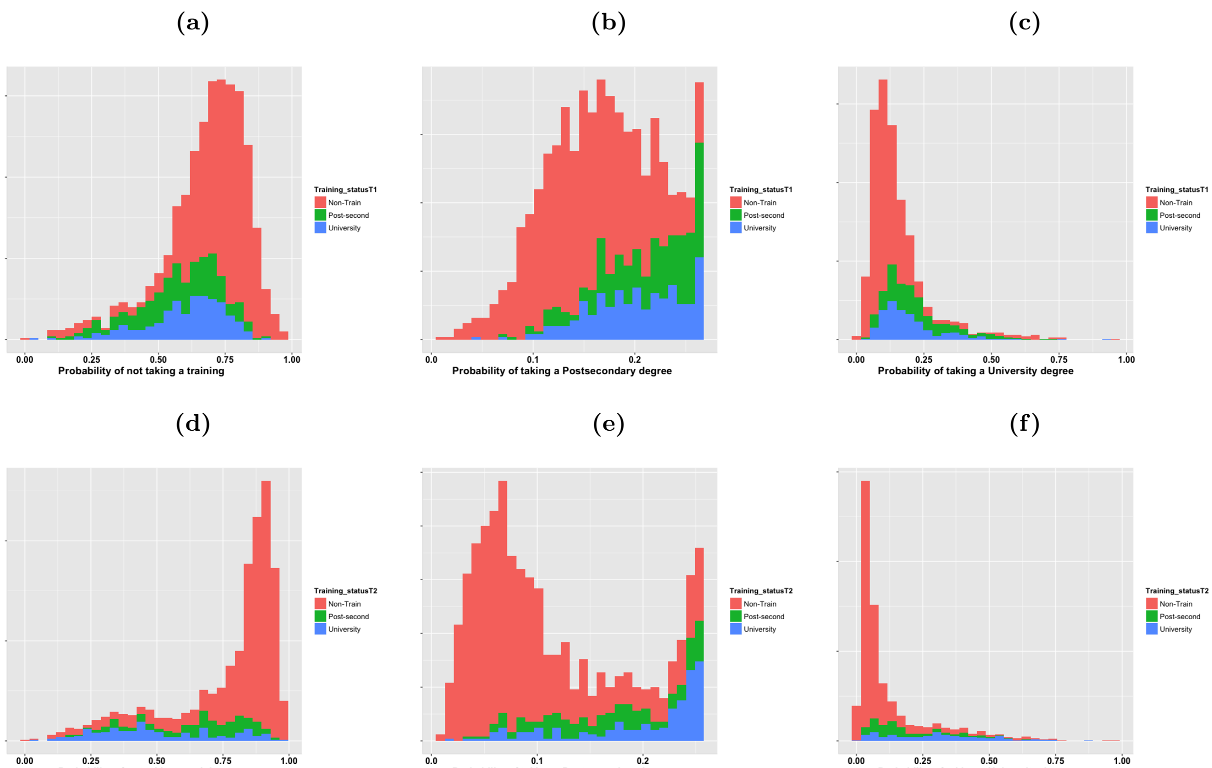


TABLE B.1 – *Subsamples used to estimate weights*

Weight	Period	Subsample	Dummy	Weight	Subsample	Dummy
W^{PN}	2	All	P	W^{UN}	All	U
	3	P	P		U	U
W^{NP}	2	All	N	W^{NU}	All	N
	3	N	P		N	N
W^{PP}	2	All	P	W^{UU}	All	U
	3	P	P		U	U
W^{NN}	2	All	N			
	3	N	N			

TABLE B.2 – *List of variables used to compute the propensity scores*

In the first period	- Age, Age squared, language test scores, foreign-education,foreign-experience - Indicator for human capital quality,gender, dummies for continent of origin, - If individual has a qualified job before in the past, dummies for entry date - If domain of QC-education is different from the foreign-education
In the second period	- Same variables as above plus the variable standing for the work experience in the first period.

Table B.3 – *Migrants characteristics depending on the percentage of trimming on propensity score*

	0.1% of trimming						1% of trimming					
	UN	NU	UU	PN	NP	PP	UN	NU	UU	PN	NP	PP
BAC degree	0.62	0.48	0.61	0.44	0.45	0.50	0.62	0.48	0.61	0.44	0.45	0.50
Mast. or Doc. degree	0.21	0.30	0.21	0.13	0.13	0.09	0.21	0.30	0.21	0.13	0.13	0.09
Female	0.36	0.28	0.33	0.28	0.33	0.35	0.36	0.28	0.33	0.28	0.33	0.35
If Married	0.45	0.39	0.43	0.51	0.54	0.56	0.45	0.39	0.43	0.51	0.54	0.56
Age	31.04	31.34	32.08	33.98	34.30	35.64	31.04	31.34	32.04	34.00	34.30	35.63
French Eval	12.52	13.24	14.02	13.91	13.50	13.72	12.52	13.24	14.03	13.90	13.50	13.73
English Eval	3.78	4.12	3.60	3.19	3.29	2.72	3.78	4.12	3.61	3.17	3.29	2.71
Ind. HCQ	0.52	0.44	0.44	0.32	0.32	0.22	0.52	0.44	0.45	0.32	0.32	0.22
If Foreign Exp	0.77	0.91	0.83	0.87	0.77	0.79	0.77	0.91	0.83	0.87	0.77	0.79
Africa	0.17	0.27	0.23	0.37	0.44	0.54	0.17	0.27	0.23	0.38	0.44	0.55
Asia	0.08	0.08	0.03	0.04	0.09	0.02	0.08	0.08	0.03	0.04	0.09	0.02
Europe	0.34	0.44	0.45	0.42	0.36	0.28	0.34	0.44	0.45	0.42	0.36	0.28
2002-2003	0.24	0.24	0.22	0.23	0.23	0.20	0.24	0.24	0.22	0.23	0.23	0.20
2004-2005	0.41	0.34	0.35	0.33	0.39	0.40	0.41	0.34	0.35	0.33	0.39	0.40
2007-2008	0.29	0.34	0.36	0.38	0.28	0.35	0.29	0.34	0.37	0.38	0.28	0.35
If ask for diploma equiv.	0.51	0.39	0.32	0.27	0.24	0.07	0.51	0.39	0.32	0.26	0.24	0.07
If diploma is under prof order	0.49	0.67	0.67	0.80	0.84	0.96	0.49	0.67	0.67	0.81	0.84	0.96
If ask for loans	0.47	0.47	0.52	0.46	0.48	0.49	0.47	0.47	0.52	0.46	0.48	0.49
Duration before	0.22	0.58	0.36	0.32	0.71	0.67	0.22	0.58	0.36	0.33	0.71	0.67

	5% of trimming					
	UN	NU	UU	PN	NP	PP
BAC degree	0.62	0.48	0.63	0.45	0.45	0.49
Mast. or Doc. degree	0.22	0.30	0.21	0.11	0.13	0.08
Female	0.36	0.28	0.32	0.27	0.33	0.35
If Married	0.45	0.39	0.43	0.52	0.54	0.56
Age	31.08	31.29	31.91	34.10	34.30	35.63
French Eval	12.50	13.24	14.05	13.87	13.48	13.78
English Eval	3.77	4.15	3.62	3.11	3.30	2.66
Ind. HCQ	0.52	0.45	0.45	0.30	0.32	0.21
If Foreign Exp	0.78	0.91	0.82	0.87	0.77	0.79
Africa	0.17	0.27	0.22	0.39	0.43	0.56
Asia	0.07	0.08	0.02	0.04	0.09	0.01
Europe	0.35	0.44	0.45	0.40	0.36	0.27
2002-2003	0.24	0.24	0.21	0.23	0.23	0.19
2004-2005	0.42	0.34	0.35	0.33	0.39	0.40
2007-2008	0.28	0.34	0.37	0.38	0.28	0.36
If ask for diploma equiv.	0.50	0.40	0.31	0.24	0.24	0.06
If diploma is under prof order	0.50	0.67	0.67	0.84	0.84	0.96
If ask for loans	0.47	0.46	0.52	0.46	0.48	0.49
Duration before	0.23	0.56	0.36	0.34	0.69	0.69

Figure B.2 – *Predicted and actual distribution of the log Earnings*

(a) *Period 2* **(b)** *Period 3*

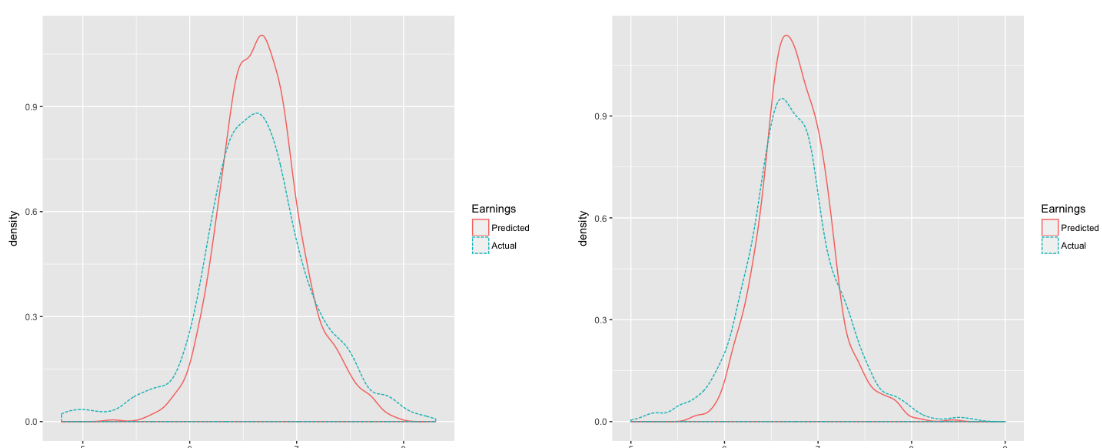
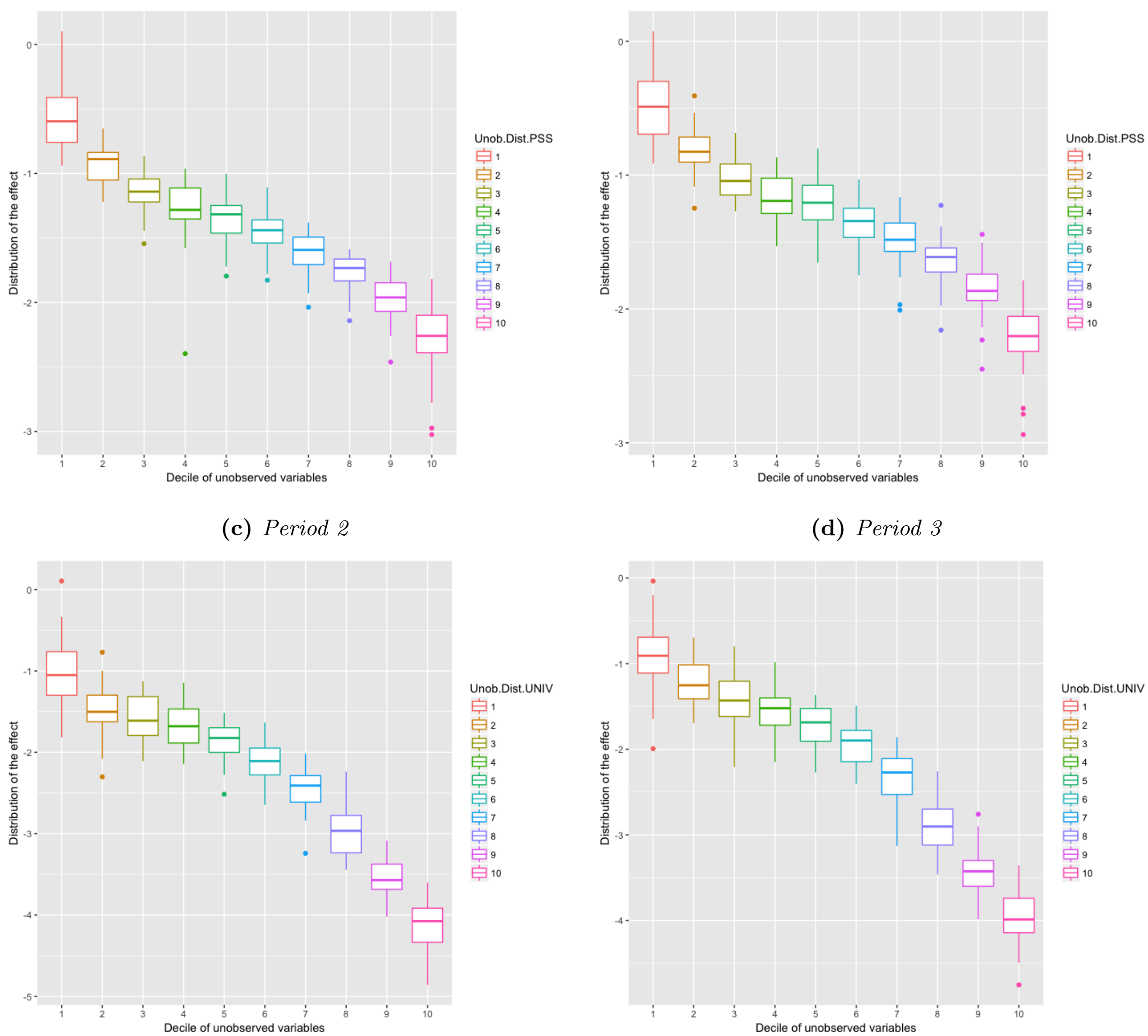


TABLE B.4 – *Posterior distribution*

	Second period		Third period	
	Mean	*Sd*	*Mean*	*Sd*
$P\ vs\ NN$:Africa	-0.47	0.02	-0.44	0.02
$P\ vs\ NN$:Europe	-0.59	0.03	-0.52	0.02
$U\ vs\ NN$:Africa	-0.82	0.05	-0.74	0.05
$U\ vs\ NN$:Europe	-0.88	0.04	-0.83	0.04

FIGURE B.3 – *Distribution of the effect by decile of unobserved heterogeneity*

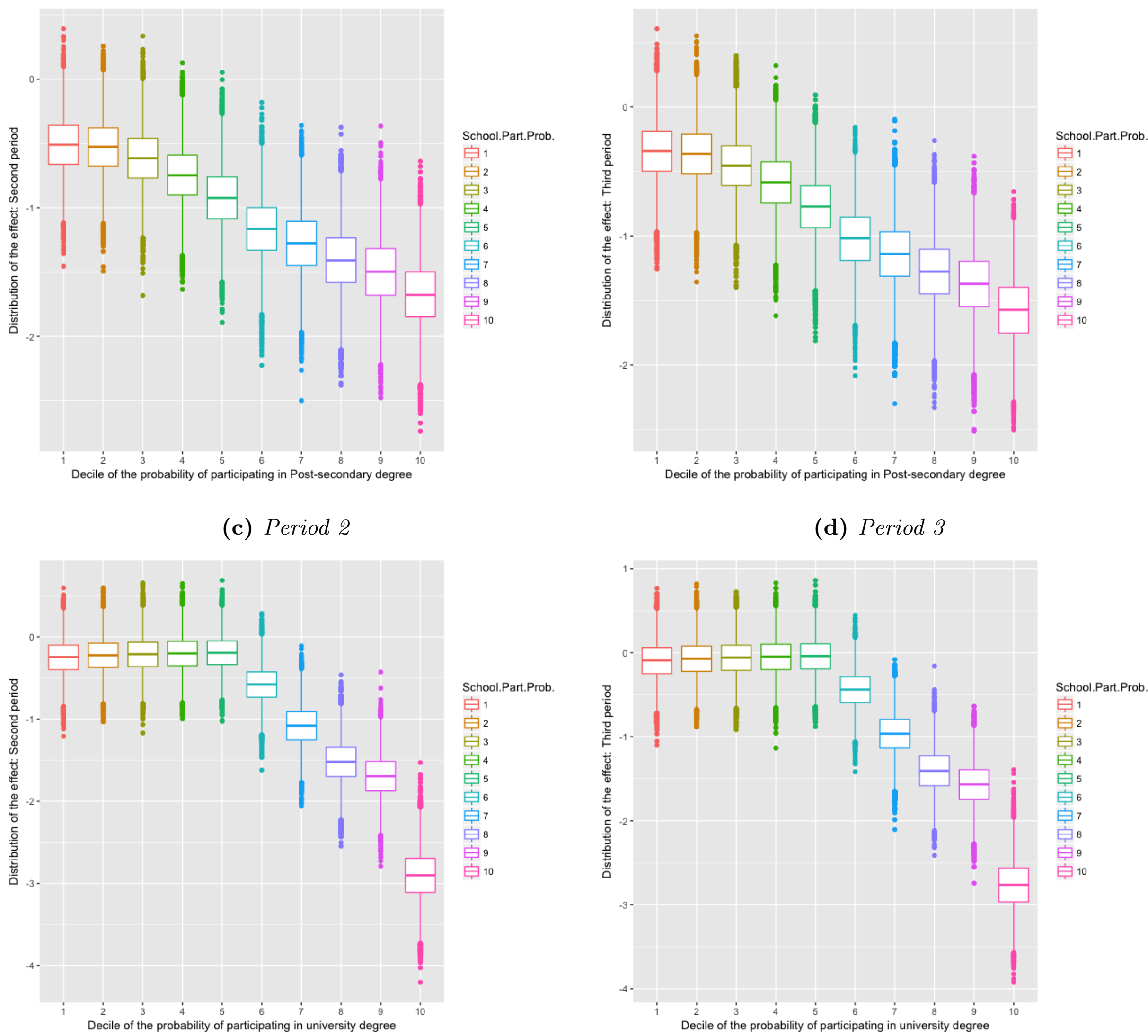


(a) *Period 2* **(b)** *Period 3*

(c) *Period 2* **(d)** *Period 3*

Figure B.4 – *Distribution of the effect by decile of likelihood of participating in education*

(a) *Period 2* **(b)** *Period 3* **(c)** *Period 2* **(d)** *Period 3*



B.2. Algorithm for posterior sampling

1. Initialize b, d_1,σ, θ, $(\lambda_{i2})_{i=1}^N$,$\lambda_{i3})_{i=1}^N$

2. Sample $\sigma|b,d_1,(\theta)_{i=1}^N,(\lambda_{i2})_{i=1}^N,(\lambda_{i3})_{i=1}^N$. The posterior distribution is given by :
$h\left(\sigma|\nu_0,\Omega_0,b,(\theta)_{i=1}^N,(\lambda_{i2})_{i=1}^N,(\lambda_{i2})_{i=1}^N\right) = f(\sigma|\nu_0,\Omega_0)\times\log L(Y_2,Y_3,U_1^*,U_2^*|b,d_1,\sigma,\theta,\lambda_1,\lambda_2,X)$.
To sample σ from this distribution, Chib & Greenberg (1998) propose to sample from a multivariate-t student $q(\mu,V)$, where μ and V are respectively the mode and the inverse of the negative of the Hessian matrix evaluated at the mode of the function $h(.|.)$.Therefore, we move from σ to σ' with probability :

$$p = \min\left\{\frac{h\left(\sigma'|\eta_0,\Omega_0,b,d_1,\theta,(\lambda_{1i},\lambda_{2i},\lambda_{3i},\lambda_{4i})_{i=1}^N\right)q(\sigma|\nu,V)}{h\left(\sigma|\eta_0,\Omega_0,b,\theta,(\lambda_{1i},\lambda_{2i},\lambda_{3i},\lambda_{4i})_{i=1}^N\right)q(\sigma'|\nu,V)},1\right\}$$

3. Sample (U_{i1}^*, U_{i2}^*) and the unobserved component of the vectors $Z_{iu}^* = \left(Y_{iu}^{*\bar{N}}, Y_{iu}^{*\bar{P}}, Y_{iu}^{*\bar{U}}\right)$ conditional on $b, d_1, (\theta)_{i=1}^N$,$(\lambda_{i2})_{i=1}^N$,$(\lambda_{i3})_{i=1}^N$, σ and conditional on the data.
   - if $\underline{s}_{2i}^k = NP$ then sample first $U_{i1}^*|b, d_1, (\theta)_{i=1}^N$,$(\lambda_{i2})_{i=1}^N$,$(\lambda_{i3})_{i=1}^N$ from a normal distribution truncated to the interval $]-\infty, 0]$ and then sample $U_{i2}^*|b, d_1, (\theta)_{i=1}^N$,$(\lambda_{i2})_{i=1}^N$,$(\lambda_{i3})_{i=1}^N$ from a normal distribution truncated to the interval $]0, d_1]$ and, we do the same for other sequence of school decision.
   - $\forall\, u \in \{2, 3\}$, $i = 1, \cdots, n$, sample either $Y_{iu}^{*\bar{N}}, Y_{iu}^{*\bar{P}}, Y_{iu}^{*\bar{U}}$, independently from i and from u, from a normal distribution depending on whether $\underline{s}_{2i}^k \in \{\bar{N}, \bar{P}, \bar{U}\}$.
4. $d_1 \sim \mathcal{U}\left\{max\left[\left(U_{i1}^* + U_{i2}^*\right)/2|i \in \bar{P}\right], min\left[\left(U_{i1}^* + U_{i2}^*\right)/2|i \in \bar{U}\right]\right\}$.
5. Sample $\beta \sim \mathcal{N}(B, \Gamma)$, $\quad with \quad B = \Gamma\left(B_0\Gamma_0^{-1} + \sum_{i=1}^n X_i\Sigma_{iu}^{-1}Z_i\right)\ and\, \Gamma = (\Gamma_0^{-1} + \sum_{i=1}^n X_i\Sigma_{iu}^{-1}X_i)^{-1}$, with $Z_i = \left(U_{i1}, U_{i2}, Y_{i2}^{\bar{N}}, Y_{i2}^{\bar{P}}, Y_{i2}^{\bar{U}}, Y_{i3}^{\bar{N}}, Y_{i3}^{\bar{P}}, Y_{i3}^{\bar{U}}\right)$

$$X_i = \begin{bmatrix} 1 & x_{i0} & 0 & \theta_i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & x_{i0} & x_{i1} & \theta_i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \underline{x}_{i1} & \theta_i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \underline{x}_{i1} & \theta_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \underline{x}_{i1} & \theta_i & \\ 0 & 0 & 0 & 0 & 1 & \underline{x}_{i2} & \theta_i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \underline{x}_{i2} & \theta & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \underline{x}_{i2} & \theta_i \end{bmatrix}$$

   of dimension $(8 \times k)$, k is the length of B and $G = [0, 0, 1, 0, 0, 1, 0, 0]$
6. Sample $\theta_i|b, d_1, \sigma, Z_i, X_i, (\lambda_{2i})_{i=1}^N, (\lambda_{3i})_{i=1}^N$ from the normal distribution $\mathcal{N}(\mu_i, \vartheta_i)$ [1]
   with $\mu_i = \vartheta_i\left(\psi'\Sigma^{-1}\tilde{\lambda}_i(Z_i - M_i)\right) \quad and \quad \vartheta_i = \left(\psi'\tilde{\lambda}_i\Sigma^{-1}\psi + 1/\vartheta_0\right)^{-1}$, where $\psi = (\psi_0, 1, \psi_{\bar{P}}, \psi_{\bar{U}}$ is the vector of loading factors, $\lambda_i = (\lambda_{i2}, \lambda_{i3})$, $\tilde{\lambda_i} = (1, \lambda_{i2} + \lambda_{i3}, \lambda_{i2} + \lambda_{i3}, \lambda_{i2} + \lambda_{i3})$ and finally,

$$M_i = \Big(\ a + \alpha x_{i0}, a + \alpha x_{i0} + \alpha_1 x_{i1}, a_{\bar{N}} + \beta_{\bar{N}}\underline{x}_{i1}, a_{\bar{P}} + \beta_{\bar{P}}\underline{x}_{i1}, a_{\bar{U}} + \beta_{\bar{U}}\underline{x}_{i1}, a_{\bar{N}} + \beta_{\bar{N}}\underline{x}_{i2}, a_{\bar{P}} + \beta_{\bar{P}}\underline{x}_{i2}, a_{\bar{U}} + \beta_{\bar{U}}\underline{x}_{i2}\ \Big)$$

7. Sample $\lambda_{iu}|Z_{iu}^*, B, X_i, \sigma, (\theta_i)_{i=1}^N, \forall u = 2, 3$ from the gamma distribution $\mathcal{G}\left(\frac{\nu_0+3}{2}, \frac{\nu_0+\underline{z}'_{iu}\Sigma^{-1}\underline{z}_{iu}}{2}\right)$, with

$$\underline{z}_{iu} = Z_{iu} - \begin{pmatrix} a + \alpha x_{i0} + \psi_0\theta_i \\ a + \alpha x_{i0} + \alpha_1 x_{i1} + \psi_0\theta_i \\ a_{\bar{N}} + \beta_{\bar{N}}\underline{x}_{iu} + \theta_i \\ a_{\bar{P}} + \beta_{\bar{P}}\underline{x}_{iu} + \psi_{\bar{P}}\theta_i \\ a_{\bar{U}} + \beta_{\bar{U}}\underline{x}_{iu} + \psi_{\bar{U}}\theta_i \end{pmatrix}$$

   .
8. Repeat step 2 to step 7 to get a full posterior distribution.

1. This is an adaption of the version of parameters propose in Chib & Hamilton (2002) and Lindley & Smith (1972)

B.3. Some proofs

We demonstrate here how the parameters of the posterior distribution of θ_i in the simulation process have been obtained. Given b, d_1, σ, Z_i, X_i, $(\lambda_{2i})_{i=1}^N$, $(\lambda_{3i})_{i=1}^N$, we have from equation 2.4 that

$$\begin{aligned} \bar{U}_i = &\quad \sum_{t=1}^{2}\left(U_{it}-\Delta_{it}^{1}\tilde{\alpha}\right)\sim\mathcal{N}\left(\psi_0\theta_i,1\right) \\ \bar{Z}_i = &\quad \sum_{u=2}^{3}\lambda_{it}\left(H_{iu}-\Delta_{iu}^{2}\tilde{\beta}\right)\sim\mathcal{N}\left(\sum_{u=2}^{3}\lambda_{iu}\underline{\psi}\theta_i,\Sigma'\right),\Sigma'=\begin{pmatrix}\sigma_{\bar{N}} & 0 & 0\\ 0 & \sigma_{\bar{P}} & 0\\ 0 & 0 & \sigma_{\bar{U}}\end{pmatrix}\end{aligned} \tag{B.1}$$

with, $\Delta_{i1}^1=(1,x_{i0})$, $\Delta_{i2}^1=(1,x_{i0},x_{i1})$, $\underline{\psi}=(1,\psi_{\bar{P}},\psi_{\bar{U}})$, and $\Delta_{iu}^2=\begin{pmatrix}1 & \underline{x}_{iu} 0 & 0 & 0 & 0 & \\ 0 & 0 & 1 & \underline{x}_{iu} & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & \underline{x}_{iu}\end{pmatrix}$.

We can rewrite the system as follow, $V_i=\begin{pmatrix}\bar{U}_i\\ \bar{Z}_i\end{pmatrix}|_{\theta_i}\sim\mathcal{N}\left(\tilde{\lambda}_i\tilde{\psi}\theta_i,\Sigma\right)$ With $\Sigma=\begin{pmatrix}1 & 0\\ 0 & \Sigma'\end{pmatrix}$, $\tilde{\psi}=(\psi_0,1,\psi_{\bar{P}},\psi_{\bar{U}})$ and $\tilde{\lambda}_i=\left(1,\sum_{u=2}^{3}\lambda_{iu}\right)$.

The posterior distribution of θ_i following the Bayes rules is given by : $P(\theta_i|V_i,\cdot)\propto P(V_i|\theta_i)P(\theta_i)$. In the right hand side of the proportionate sign, both distributions are normal distributions meaning that they are proportionate to $\exp(-\frac{1}{2}Q)$. With $\theta_i\sim\mathcal{N}(0,\nu_0)$, [2] we have

$$\begin{aligned} Q &= \left[V_i-\tilde{\lambda}_i\tilde{\psi}\theta_i\right]'\left(\tilde{\lambda}_i\right)^{-1}\Sigma^{-1}\left[V_i-\tilde{\lambda}_i\tilde{\psi}\theta_i\right]+\theta_i'\nu_0^{-1}\theta_i \\ &= V_i'\tilde{\lambda}_i\Sigma^{-1}V_i-V_i'\left(\tilde{\lambda}_i\right)^{-1}\Sigma^{-1}\tilde{\lambda}_i\tilde{\psi}\theta_i-\theta_i\tilde{\psi}'\Sigma^{-1}V_i+\theta_i\tilde{\psi}'\Sigma^{-1}\tilde{\lambda}_i\tilde{\psi}\theta_i+\theta_i\nu_0^{-1}\theta_i \\ &= \theta_i\left[\tilde{\lambda}_i\tilde{\psi}'\Sigma^{-1}\tilde{\psi}+\nu_0^{-1}\right]\theta_i-V_i'\Sigma^{-1}\tilde{\psi}\theta_i-\theta_i\tilde{\psi}'\Sigma^{-1}V_i+V_i'\tilde{\lambda}_i\Sigma^{-1}V_i \\ &= \left[\theta_i-AK\right]'A^{-1}\left[\theta_i-AK\right]+V_i'\tilde{\lambda}_i\Sigma^{-1}V_i;A^{-1}=\tilde{\lambda}_i\tilde{\psi}'\Sigma^{-1}\tilde{\psi}+\nu_0^{-1};\quad K=\tilde{\psi}'\Sigma^{-1}V_i\end{aligned}$$

The last term of the equality is independent from θ_i therefore, the posterior distribution $\theta_i\sim\mathcal{N}(AK,A)$ which ends the demonstration with $A=\Sigma_\theta^i=\tilde{\lambda}_i\tilde{\psi}'\Sigma^{-1}\tilde{\psi}+\nu_0^{-1}$ and $\mu_\theta^i=AK=\Sigma_\theta^i\tilde{\psi}'\Sigma^{-1}V_i$.

2. Since θ_i is a constant, it is equal to its transpose.

- Annexe C -

C.1. Distribution of education, ethnicity and agricultural production in Uganda

Figure C.1 – *Distribution and auto-correlation of the Variance posterior distribution*

(a) *Education distribution in Uganda*

(b) *Ethnicity and Agricultural productivity*

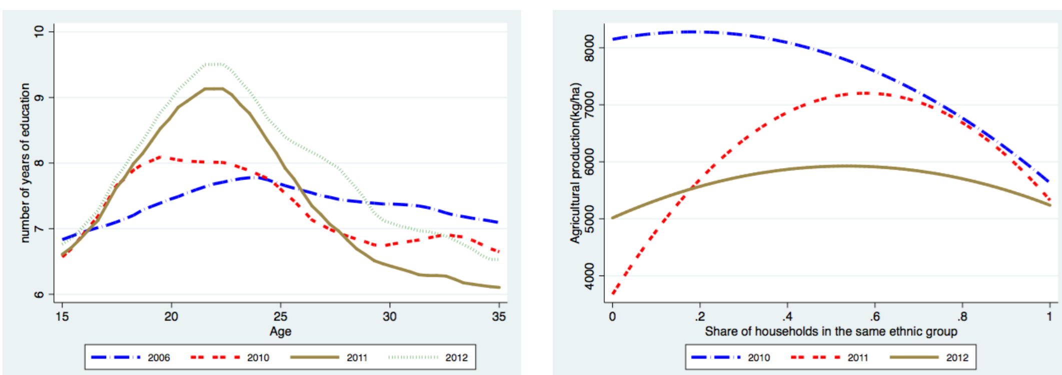


Figure C.2 – *Duration of migration and share of household members involved in migration*

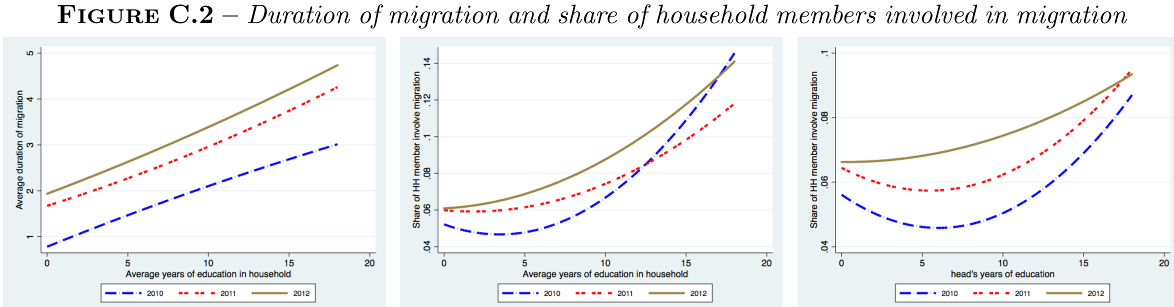


Figure C.3 – *Agricultural productivity of grain by head's years of education and by average years of education in household*

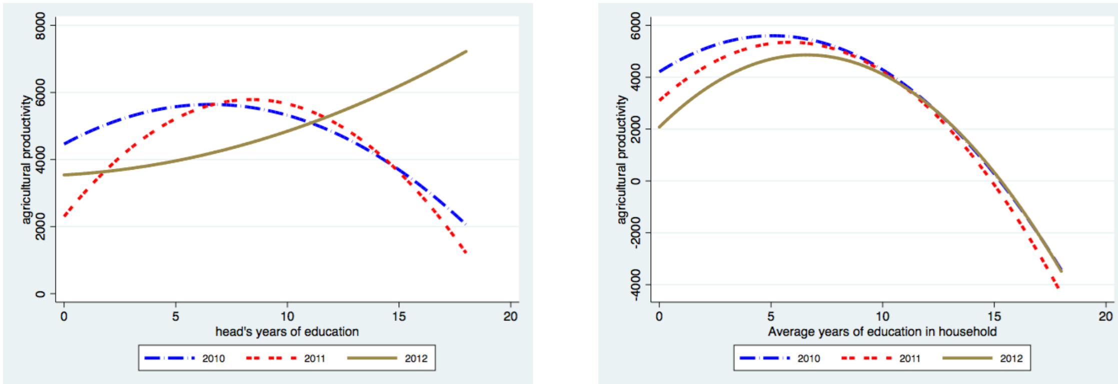


Table C.1 – *Household migration status and previous permanent migration status of household head*

	Migrants-sending HHs			Non Migrants-sending HHs		
	2009	2010	2011	2009	2010	2011
$Head - Migrants$	13.1	12.4	13.4	86.9	87.6	86.6
$Head - NonMigrants$	12.3	11.6	12.0	87.7	88.4	88.0

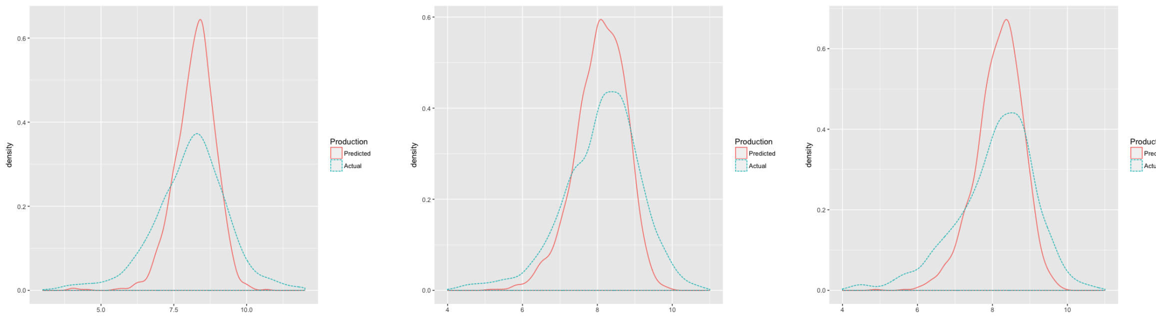
Table C.2 – *Posterior distribution of parameters assuming that instruments verify the restriction assumption*

Variables	Migration likelihood		Agr. Productivity for MIG-sending HHs		Agr. Productivity for Non-MIG-sending HHs	
	mean	SD	mean	SD	mean	SD
Log wealth	-1.04	0.17	0.56	0.14	0.12	0.14
Log wealth sq	0.09	0.02	-0.03	0.01	0.00	0.01
Married poly	-0.01	0.07	-0.06	0.06	0.08	0.07
Div. sep	-0.00	0.12	-0.09	0.09	0.14	0.11
Widow	-0.01	0.11	-0.18	0.08	0.03	0.10
Single	-0.64	0.25	0.04	0.23	-0.08	0.21
Head gender	0.50	0.09	0.09	0.06	-0.26	0.08
Log Head age	0.05	0.10	0.64	0.08	0.07	0.08
Head Educ	0.03	0.01	0.02	0.01	0.00	0.01
Head-MIG	0.06	0.13	-0.11	0.12	-0.01	0.10
Children less than 5	-1.11	0.19	0.15	0.16	-0.43	0.15
Individuals aged more than 65	0.96	0.24	-1.02	0.22	-0.58	0.22
Log HH size	1.36	0.08	0.35	0.08	0.19	0.07
AVE educ in hh	0.04	0.02	-0.01	0.01	0.01	0.01
Adult domes. lab DP.	-0.06	0.04	-0.01	0.03	-0.00	0.02
Children domes.lab. DP.	0.05	0.05	-0.03	0.04	0.02	0.03
Geo-spatial DP.	-0.02	0.03	-0.05	0.02	-0.00	0.02
Hired lab	0.00	0.00	-0.02	0.02	0.01	0.02
Proxy of HH labor	-	-	-0.14	0.07	-0.07	0.07
Area planted	-	-	-0.61	0.04	-0.51	0.03
Nb. crops	-	-	0.85	0.05	0.75	0.04
Agriculture in 1 km radius	-	-	-0.27	0.11	-0.02	0.14
Ethnicity concentration (at country level)	-1.47	0.29	2.61	0.27	0.06	0.30
Ethnicity concentration(at district level)	-0.05	0.10	-0.20	0.08	0.12	0.08
Center	0.04	0.08	-0.21	0.07	-0.06	0.08
East	-0.53	0.08	-0.28	0.07	-0.23	0.08
North	-0.39	0.08	-0.03	0.08	-0.04	0.08
Nb. of migrants in hh(2005)	0.24	0.05	-	-	-	-
Migration rate in the District	1.30	0.31	-	-	-	-
Log Wealth DP.	-0.00	0.01	-	-	-	-
Intercept	-0.76	0.29	2.27	0.28	5.65	0.30
Loading fact.	-0.55	0.06	0.23	0.03	1	-
σ	1	-	0.24	0.02	0.49	0.03

⋆ Head-MIG equals 1 if the current place of living of household is different from the place of household's head place of birth, that is when household's head has migrated permanently in the past (less than 11 years and more than two years).

⋆ In the column of the likelihood migration, we have reported the parameters and they cannot be interpreted as a marginal effect. Only the sign is significant.

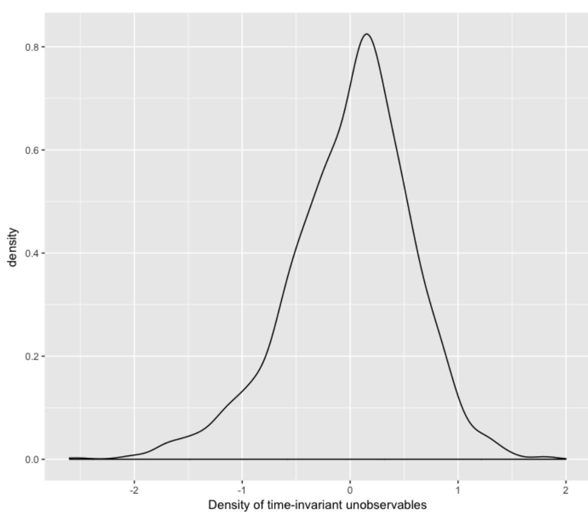
FIGURE C.4 – *Actual and predicted distribution of agricultural production*



Period 1 *Period 2* *Period 3*

FIGURE C.5 – *Distribution of the unobserved factors*

(a) *Density funct. of unobservables,* θ_i



(b) *Inverse demand of migration as a funct. of unobservables* θ_i

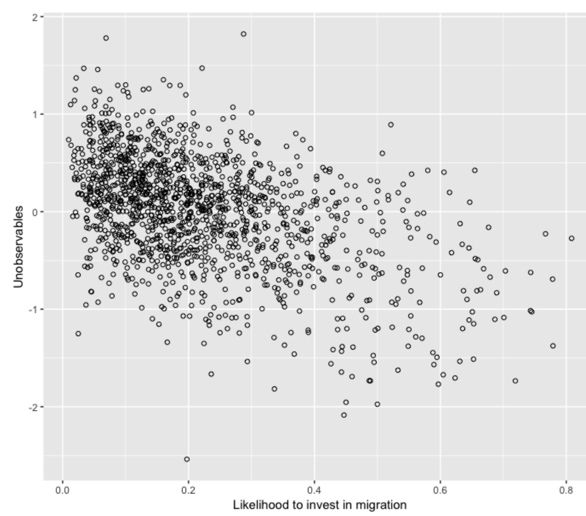


FIGURE C.6 – *Heterogeneity of the effect by prob. of participating in MIG*

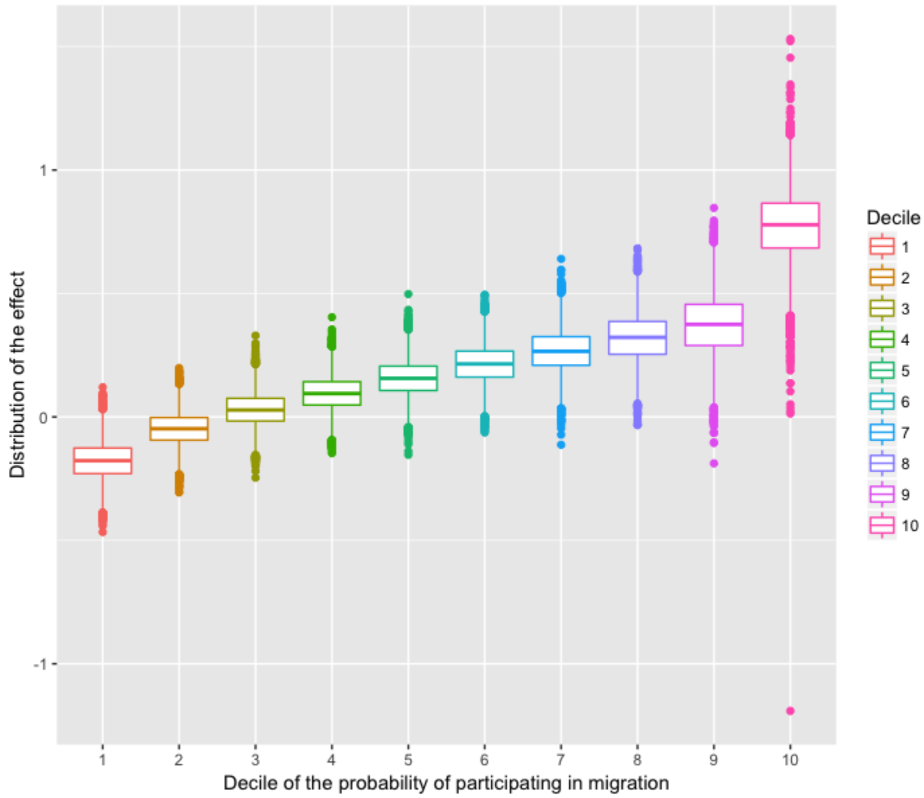


TABLE C.3 – *Percentage of households with positive return to migration*

	Migrant-HHs					
	Central	Eastern	Northern	Western	Num crops>13	Num crops<=13
Period 1	61.46	53.33	60.47	72.94	63.69	62.20
Period 2	75.00	80.00	75.56	76.42	80.35	72.73
Period 3	75.74	58.26	57.66	77.98	72.73	61.93
	Non Migrant-HHs					
Period 1	63.51	62.40	51.95	47.91	56.50	56.67
Period 2	61.21	45.23	37.93	66.22	54.81	48.26
Period 3	40.66	61.25	48.75	54.39	51.00	54.14

C.2. Sampling Algorithm

1. Sample σ from a Metropolis Hastings strategy. The posterior distribution is $h\left(\sigma|\ell_0, L_0, B, \theta, (\lambda_{1i})_{i=1}^N, (\lambda_{2i})_{i=1}^N, (\lambda_{3i})_{i=1}^N\right) = f(\sigma|\ell_0, L_0) \times L\left(Prod_t, LM_t|B, \sigma, \lambda, \theta\right)$ and f(.|.) is a multivariate normal distribution of order two.

   To sample σ from function $h(.|.)$, Chib & Greenberg (1998) propose to sample σ from a multivariate-t distribution $q(\nu, V)$ where ν and V are respectively the mode and the inverse of the negative of the hessian of $log(h)$. Therefore, we move from σ to σ' with probability :

   $p = \min\left\{\frac{h\left(\sigma'|\ell_0,L_0,B,\theta,(\lambda_{1i})_{i=1}^N,(\lambda_{2i})_{i=1}^N,(\lambda_{3i})_{i=1}^N\right)q(\sigma|\nu,V)}{h\left(\sigma|\ell_0,L_0,B,\theta,(\lambda_{1i})_{i=1}^N,(\lambda_{2i})_{i=1}^N,(\lambda_{3i})_{i=1}^N\right)q(\sigma'|\nu,V)}, 1\right\}$

This strategy enables to reach the convergence of σ more rapidly.

2. Sample the unobserved component of the vector $H^*_{it} = (MU^*_{it}, Prod^*_{1it}, Prod^*_{0it})$, $\forall t = 1,2,3$.
   - if $LM_{it} = 1$ then sampled first $MU^*_{it}|B,\sigma,\theta_i,\lambda_{1i},\lambda_{2i},\lambda_{3i}$ a normal distribution truncated to the interval $]0,+\infty[$. Instead, if $LM_{it} = 0$ then sampled $MU^*_{it}|B,\sigma,\theta_i,\lambda_{1i},\lambda_{2i},\lambda_{3i}$ a normal distribution truncated to the interval $]-\infty,0]$.
   - $\forall\, t \in \{1,2,3\}$, $i = 1,\cdots,n$, sample either $Prod^*_{1it}$ or $Prod^*_{0it}$, independently from i and t, from a normal distribution depending on whether LM_{it} is equal to zero or one.

3. Sample the set of parameters $B|H^*_{it}, b_0, B_0, \sigma, (\theta_i)_{i=1}^N, (\lambda_{1i})_{i=1}^N, (\lambda_{2i})_{i=1}^N, (\lambda_{3i})_{i=1}^N$, from the normal distribution $\mathcal{N}(g,G)$, with
$g = G\left(b_0B_0^{-1} + \sum_{i=1}^N R_i'\Omega_i^{-1}(H_i - \Lambda\theta_i)\right)$; $G = \left(B_0^{-1} + \sum_{i=1}^N R_i'\Omega_i^{-1}R_i\right)^{-1}$ where
$\Omega_i = diag\,(\lambda_{i1},\lambda_{i2},\lambda_{i3}) \bigotimes diag(1,\sigma_1,\sigma_0)$ [1], $H_i = (H_{i1},H_{i2},H_{i3})$
$R_i = \begin{bmatrix} \Delta_{i1} \\ \Delta_{i2} \\ \Delta_{i3} \end{bmatrix}$ is a matrix, with $\Delta_{it} = \begin{bmatrix} Z_i & W_{it} & 0 & 0 \\ Z_i & 0 & X_{1it} & 0 \\ Z_i & 0 & 0 & X_{0it} \end{bmatrix}$ of dimension $(3\times T, k)$;$T = 3$
and k is the length of B and $\Lambda_i = [0,0,1,0,0,1,0,0,1]$.

4. Sample $\theta_i|H_{it},B,\sigma,(\lambda_{1i})_{i=1}^N,(\lambda_{2i})_{i=1}^N,(\lambda_{3i})_{i=1}^N$ from the normal distribution with mean $\mu^i_\theta = \Sigma^i_\theta D'C^{-1}\bar{H}_i$ and $\Sigma^i_\theta = \left(1/\nu_0 + D'C^{-1}\lambda^*_iD\right)^{-1}$. $D = (\gamma,\gamma_1,1)$ is the vector of loading factors, $\lambda^*_i = \sum_{i=t}^N \lambda_{it}$ and
$$\bar{H}_i = \sum_{t=1}^3 \lambda_{it}\left(H_{it} - [Z_i\beta + X_{it}\alpha_m, X_{1it}\alpha_1, X_{0it}\alpha_0]\right)$$

5. Sample $\lambda_{it}|H_{it},B,\sigma,(\theta_i)_{i=1}^N$, $\forall t = 1,2,3$ from a gamma distribution $\mathcal{G}\left(\frac{\lambda_0+3}{2}, \frac{\lambda_0+\underline{H'}_{it}C^{-1}\underline{H}_{it}}{2}\right)$, with
$$\underline{H}_i = H_{it} - \begin{pmatrix} Z_i\beta + X_{it}\alpha_m + \theta_i\gamma \\ Z_i\beta^1 + X_{1it}\alpha_1 + \theta_i\gamma_1 \\ Z_i\beta^0 + X_{0it}\alpha_0 + \theta_i \end{pmatrix}$$

6. Complete the sampling procedure by repeating step 1 to step 5.

C.3. Some proofs

We demonstrate now how the parameters of the posterior distribution of θ_i in the simulation process have been obtained. Given H_{it},B,σ, $(\lambda_{1i})_{i=1}^N,(\lambda_{2i})_{i=1}^N,(\lambda_{3i})_{i=1}^N$, we have from equation 3.4 that $H_{it} \sim \mathcal{N}\left(\Delta_{it}A + D\theta_i, \lambda_{it}^{-1}\Sigma\right)$, with A be the vector of all parameters of the model except the loading factors.
Thereby, $\lambda_{it}\left(H_{it} - \Delta_{it}A\right) \sim \mathcal{N}\left(\lambda_{it}D\theta_i, \lambda_{it}\Sigma\right) \iff \sum_{t=1}^3\left[\lambda_{it}\left(H_{it} - \Delta_{it}A\right)\right] \sim \mathcal{N}\left(\lambda^*_iD\theta_i, \lambda^*_i\Sigma\right)$, $\lambda^*_i = \sum_{t=1}^3 \lambda_{it}$

1. $diag(C)$ represents the diagonal matrix with the elements of vector C on the diagonal and $A \bigotimes B$ stands for the kronecker product of A and B.

If we set $M_i = \sum_{t=1}^{3} \lambda_{it} \left(H_{it} - \Delta_{it} A\right)$, the posterior distribution of θ_i following the Bayes rules is given by : $P(\theta_i | M_i, \cdot) \propto P(M_i|\theta_i)P(\theta_i)$. In the right hand side of the proportionate sign, both distributions are normal distributions meaning that they are proportionate to $\exp(-\frac{1}{2}Q)$, with $\theta_i \sim \mathcal{N}\left(0, \nu_0\right)$, [2] and

$$\begin{aligned}
Q \quad &= \left[M_i - \lambda_i^* D\theta_i\right]' \left(\lambda_i^*\right)^{-1} \Sigma^{-1} \left[M_i - \lambda_i^* D\theta_i\right] + \theta_i' \nu_0^{-1} \theta_i \\
&= M_i' \lambda_i^* \Sigma^{-1} M_i - M_i' \left(\lambda_i^*\right)^{-1} \Sigma^{-1} \lambda_i^* D\theta_i - \theta_i D' \Sigma^{-1} M_i + \theta_i D' \Sigma^{-1} \lambda_i^* D\theta_i + \theta_i \nu_0^{-1} \theta_i \\
&= \theta_i \left[\lambda_i^* D' \Sigma^{-1} D + \nu_0^{-1}\right] \theta_i - M_i' \Sigma^{-1} D\theta_i - \theta_i D' \Sigma^{-1} M_i + M_i' \lambda_i^* \Sigma^{-1} M_i \\
&= \left[\theta_i - AK\right]' A^{-1} \left[\theta_i - AK\right] + M_i' \lambda_i^* \Sigma^{-1} M_i ; A^{-1} = \lambda_i^* D' \Sigma^{-1} D + \nu_0^{-1}; \quad K = D' \Sigma^{-1} M_i
\end{aligned}$$

The last term of the equality is independent from θ_i therefore, the posterior distribution $\theta_i \sim \mathcal{N}\left(AK, A\right)$ which ends the demonstration with $A = \Sigma_\theta^i = \lambda_i^* D' \Sigma^{-1} D + \nu_0^{-1}$ and $\mu_\theta^i = AK = \Sigma_\theta^i D' \Sigma^{-1} M_i$.

2. Since θ_i is a constant, it is equal to its transpose.