ESSAYS IN IMMIGRATION AND MIGRATION

STEIN MONTEIRO

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN DEPARTMENT OF ECONOMICS
YORK UNIVERSITY
TORONTO, ONTARIO
JULY 2018

Abstract

This research explores the socio-structural features of the migration and assimilation decision. The socio-structural features explored are the impact of extended family members on the migration decision of individuals within a household, and productivity differences on the assimilation rate of new immigrants.

Extended families are a common feature of developing country households. I generalize the Mincer (1978) model of husband-wife migration by including decision makers from the extended family. The model with extended families predicts that migration decisions may become freer than in the husband-wife model because spouses are not more likely to be tied to their partners than members of the extended family. That is, marital status is a smaller deterrent to migration in extended family settings relative to nuclear families. I provide justification for the implications of the model using data from Nepal.

Immigrants from poorer source countries have lower assimilation rates compared to immigrants from richer countries. Theory suggests that new immigrants from poor countries are exposed to co-ethnics more often than comparable immigrants from richer countries, which lead to lower assimilation rates. However, many new immigrants come with pre-immigration experience with the local culture which decreases learning costs. I insert investment into the matching model of Konya (2007). All immigrants face a cost to assimilating by investing in a process of cultural assimilation, but some new immigrants with large pre-immigration experience have significantly lower costs to investing. I provide evidence from the Longitudinal Survey of Immigrants in Canada: Waves 1-3. Source country richness has a significant positive effect on assimilation rates. But conditional on pre-immigration experience with the local culture, the exposure channel through which source country richness affects assimilation rates becomes insignificant.

However, exposure to co-ethnics is not random, new immigrants face location choices among neighbourhoods in the host country. These location choices determine the level of exposure to other immigrants and the costs of learning the local native-born culture. I expand the model to include neighbourhood choice. Among neighbourhoods with fewer co-ethnics, immigrants from richer source country groups will sort into assimilating neighbourhoods. And neighbourhoods with a relatively large number of co-ethnics will receive some non-assimilating types. Using data from the Longitudinal Survey of Immigrants in Canada: Waves 1-3, I show that sorting is an important component of the exposure channel through which productivity differences affect assimilation rates. However, controlling for sorting, source country richness still has a significant positive effect on assimilation rates. There appears to be an alternate channel through which productivity differences affect assimilation rates.

For my parents.

Acknowledgements

I would like to express my sincere gratitude to my advisor Dr. Sam Bucovetsky for the continuous supervision of my Ph.D study and research, for his patience, enthusiasm, and immense knowledge. His guidance helped me in all aspects of research and writing of this thesis.

I would also like to thank my thesis committee members, Dr. Uros Petronijevic and Dr. Benjamin Sand, for their encouragement, insightful comments and recommendations to improving the research presented in the thesis. I would also like to thank Dr. Mahmudul Anam for sharing his time and being available to serve on the committee.

My sincere thanks to Dr. Tasso Adamopolous for offering me the research assistantship opportunities on diverse and exciting projects, and for his comments and difficult questions in the presentation of my research at the seminars. I would like to thank Dr. Nippe Lagerlöf for his insightful and thoughtful contributions to the various drafts of the job market paper and providing insight into the job market.

A big thank you to all the seminar participants, discussants and friends: Akram Panahi Dargahloo, Andrew Hencic, Andrew Dickens, Chiara Marazzi, Imran Aziz, Michelle Tong, Rachel Harris, Sara Tumpane, Shafkat Ali, and Vajiheh Mahdiyan Amirabadi, for the discussions and helping each other succeed.

Last but not the least, I would like to thank my family: my parents Romel and Lynette Monteiro for their unquestioning support, my siblings Sternin and Stacy Monteiro for their cheerfulness, and my wife Dewlyn D'Souza for her patient and caring nature.

Table of Contents

List of Tables

List of Figures

1 Introduction

The factors driving potential migrants to leave their home country and learn the culture of the host country may be separated by socio-structural elements and personal motivators. Economic and sociological research on the influence of personal motivations on the migration and integration decisions of migrants are numerous. Ability, talent, drive to succeed, demographic characteristics, pre-immigration wealth, social status, among other observables and unobservables are important drivers of the decision to migrate and assimilate. However, very little research has studied the various aspects of the structural elements in the home and host country that also influence the migration and assimilation decisions.

To name a few important aspects of the socio-structural dimension that might influence potential migrants' decisions: institutions in the home country, institutional differences between the home and host country, family structure, historical experiences of the home country, the extent of cultural frictions between the home and host country, and exposure to co-ethnics.

In this thesis I study the effect of family structure, cultural frictions and co-ethnic exposure on the migration and assimilation decision. Family structure is shown to have an effect on how migration decisions are formed. Similarly, the productivity differences between the source and host country are shown to have effects on assimilation rates in the host country through exposure to co-ethnics.

The first chapter of this thesis analyzes the impact of the presence of extended family members in the household on the husband-wife migration decision. Extended families are a common feature of developing country households. These households are very relevant in the decision to migrate. This paper generalizes the Mincer (1978) model of husband-wife migration by including decision makers from the extended family. The model with extended families predicts that migration decisions may become freer than in the husband-wife model because spouses are not more likely to be tied to their partners than members of the extended family. That is, marital status is a smaller deterrent to migration in extended family settings relative to nuclear families. I provide justification for the implications of the model using data from the Demographic and Health Surveys: Nepal 2011.

Once the decision to migrate has been formed, potential migrants face cultural frictions in the host country. Cultural frictions prevent new immigrants from accessing the benefits of immigration, for instance, higher wages, better job mobility, and shorter job searches. This research is interested in communication frictions. Communication frictions are alleviated by learning the local culture. Cultural assimilation of immigrants is an integral part of a well functioning multicultural society that leads to better communication and frictionless interactions between immigrants and native-born. Cultural assimilation is a form of learning that occurs through engagement with the local culture overtime or investment into cultural capital accumulation.

Productivity differences between the source and host country incentivize immigration from poorer countries. Potential migrants from poorer source countries have larger incentives to migrate than comparables from richer source countries. Hence immigrants from poorer countries form a larger portion of the immigrant stock in the host country and are more often exposed to co-ethnics. A larger exposure to co-ethnics is associated with lower assimilation rates. On the other hand, many new immigrants come with pre-immigration experience with the local culture which influences their decision to invest in further learning by decreasing the costs of learning.

The model developed in chapter 3 inserts investment into the matching model of Konya

(2007) with communication frictions represented by lack of knowledge of the local culture and exposure to co-ethnics. All immigrants face a cost to assimilating by investing in a process of cultural assimilation. But some new immigrants with large pre-immigration experience with the local culture become selected into the assimilating group without investing. Additionally I consider formal and informal learning. I provide evidence from the Longitudinal Survey of Immigrants in Canada: Waves 1-3.

Cultural assimilation is measured by English speaking proficiency, source country richness is measured by RGDP per capita in immigrant's place of birth, formal learning is proxied by learning of English through courses or schooling, informal learning is proxied by learning of English through media, family/friends or self-study, and formal/informal pre-immigration experience with the local culture is measured in the same way as formal and informal learning but it is undertaken prior to landing. Source country richness has a significant positive effect on assimilation rates through exposure to co-ethnic and allowing for investment. However, conditioning on pre-immigration experience with the local culture and further learning, source country richness has no significant effect on assimilation rates. This implies that pre-immigration experience with the local culture and further learning are important components of the channel through which productivity differences affect assimilation rates. Finally, formal learning is shown to be an important contributor to cultural assimilation, but not informal learning.

That being said, exposure to co-ethnics is not random, new immigrants face location choices among neighbourhoods in the host country. These location choices determine the level of exposure to other immigrants and the costs of learning the local native-born culture. I expand the model of chapter 3 to include neighbourhood choice. Among neighbourhoods with fewer co-ethnics, immigrants from richer source country groups will sort into assimilating neighbourhoods. And neighbourhoods with a relatively large number of co-ethnics will receive some non-assimilating types. On the other hand, immigrants from poor source countries will have some non-assimilating immigration in all neighbourhoods.

Using data from the Longitudinal Survey of Immigrants in Canada: Waves 1-3, sorting is proxied using an ethnic enclave measure, where living and working in an ethnic enclave is determined jointly by an over-representation of co-ethnics in CMA/CA of first arrival and non-official language spoken at work. Including the sorting variable in the regression overadjusts for the exposure channel through which productivity differences affect assimilation rates. Since sorting is probably endogenous I instrument with average housing costs in CMA/CA of arrival. Sorting has a significant negative effect on assimilating immigration and is certainly an important component of the exposure channel. However, overadjusting the exposure channel, source country richness still has a significant positive effect on assimilation rates. This implies the presence of an alternate channel. I find that source country richness appears to be significantly positively related to sorting, conditional on the size of the co-ethnic group. New immigrants from richer source countries are more likely to cluster compared to similar co-ethnics from poorer countries. The hypothesized alternate channel is probably related to quality, as opposed to size, of the co-ethnic group.

2 Extended Family Migration Decisions: Evidence from Nepal

2.1 Introduction

Family plays a decisive role in the decision to migrate. The relevance of husband-wife (nuclear) families in forming migration decisions and how these decisions differ from the motivations of single persons is a recurring topic of concern in the economic and sociological literature. Married couples are understood to have a smaller degree of freedom in the decision to migrate as they aim to maximize the welfare of the family as a whole rather than their own individual net benefit. Long (1974) showed empirically that married men in the US are less likely to migrate within/between counties and across states than unmarried men. This evidence was formalized in a neoclassical model of family migration decision in Mincer (1978) and further formalized within a model of altruism in Becker (1981). Massey (1990) reiterates the evidence through a sociological perspective with further evidence provided in Maxwell (1988). Miller (1976), Speare Jr. & Goldscheider (1987) and Maxwell (1988) further stress the importance of the family life cycle in the decision to migrate. More recent evidence by McKinnish (2008) and Tenn (2010) show tied-ness of the wife to the husband's migration decision through smaller wage returns and opportunities for employment after migration. Moreover, Miller (1976) is the first paper to take up extended families and how the presence of extended family members will impact the family migration decision. Root & De Jong (1991) introduce extended family members within the context of developing countries where such families are a common feature. They find that family migration is a function of the extended families' involvement in the migration decision process, such that, the extended family has the power to pressure the decision to migrate for or against the nuclear families' own welfare.

The husband-wife model is an important contribution to understanding migration in developed countries where households are essentially 'nuclear'. However, the extended family household is the dominant family structure in developing countries (Bongaarts & Zimmer, 2002). The complexity of the extended family differs by region but is nonetheless prolific.

Extended family structures are complex as they include the nuclear family, plus grandparents, uncles, aunts, cousins, etc. Migrants from developing countries leave behind some portion of the family structure and the migrating part of the household is only a subset of the complete family structure. It is highly unlikely that migration leaves an 'empty household'.

The relevance of the household in individual migration decisions cannot be understated (Mincer, 1978 and Massey, 1990). The household's characteristics play an important role in determining whether migration is valuable, or not, to the point that it determines the extent of the push towards it. Kley (2011) draws attention to the migration decision process. Potential migrants have 'place utility' and at the pre-decision stage form expectations about the attainment of valued goals in the host country (taking into consideration the pull factors). However, within the pre-decision stage potential migrants will weigh the effects of their actions on peers and household members. Certain household characteristics will lead to a lower likelihood of migration because of high economic, social and emotional costs. The constraints imposed by marital status on the migration decision within nuclear families do not carry over without complications when considering the presence of extended family members. Extended families may further constrain the nuclear family migration decision or encourage nuclear family migration (Root & De Jong, 1991). However, lesser understood is the fact that the presence of extended family members may dissolve nuclear family migration and present a new situation where one spouse migrates while the other remains. The migration observed in these situations may be predominately temporary but is certainly an important dimension to consider. The tied-ness of the spouse is given a new dimension as he/she becomes no more tied to members of the extended family than his/her own spouse. The goal of this research is to uncover the conditions under which such situations appear in the Mincer (1978) model and present empirical evidence for these cases.

Very little research has attended to the intersections between individual decisions and household characteristics within the context of migration (Mincer, 1978; Becker, 1981; Borjas & Bronars, 1991). Mincer's (1978) model with husband-wife decision makers works well in understanding the type of migration that occurs in western industrialized countries where households are commonly nuclear. As such, marital status is a deterrent to individual migration decisions and married men are significantly less likely to migrate than unmarried men. Mincer's (1978) model when applied to the developing world should predict a lot less married men migration than what is currently observed. Instead, a significant amount of married men migration (without their spouses) does take place. This is shown to be the case using data on Nepali households ("Descriptive Analysis: The Extended Family" in section 3 of this paper).

Furthermore, Mincer's (1978) model predicts significant tied-ness of the spouse to the husband's decision in the US where households are typically nuclear. In "Regression Analysis: Migration from the Extended Family" in section 4 of this paper, I show that married men (without their spouse) migration are more likely than couple migration in extended families than nuclear ones. That is, the presence of extended family members frees the spouse's decision to migrate.

The estimation strategy of this paper is to provide justification for the implications of the theoretical model using data from Nepal. Unlike Mincer (1978) and Becker (1981) where individual migrants within the household are expected to be 'tied-stayers' or 'tied-movers' to the migration decisions of the spouse; in the model of this paper there are further interdependencies with members of the extended family that may free the migration decision. The deterrent effect of marital status on migration decisions as predicted in Mincer's (1978) model are only special conditions of a more general picture of migration decisions in an extended family. The presence of extended family members will dampen the effect of marital status on migration decisions and the prevalence of tied-movers.

The set up of the paper is as follows: a theoretical model is provided, followed by empirical justification for the implications of the theoretical model. And finally the conclusion.

2.2 Theoretical model

I will develop a model of the individual's migration decision in an extended family structure. Similar to Mincer (1978), I assume that all migration decisions exclude children. However, children are quite relevant for household decisions, so their presence is treated for appropriately in the empirical section. The model extends migration in a husband-wife household and considers the role of extended family members. There are only two countries: the sending country and the receiving country. Let G_i be the net benefit to individual i such that if $G_i > 0$ then the individual's private calculus indicates a gain from migrating. Likewise $G_i \leq 0$ indicates a preference to staying. The net benefit of migrating to an individual i is decomposed as

$$G_i = R_i - C_i + v(s_i),$$

where $R_i - C_i$ is the standard economic returns minus cost and $v(s_i)$ is the value to individual i from living as a unit with s other family members. It is assumed that $v'(s_i) > 0$, $v''(s_i) < 0$ and $v(0) = 0$. The individual net benefit G_i for each individual i in the family is a measure of the net economic benefit of migrating had this individual made an independent decision. The total net benefit for a household with n members, if all migrate, will be[1]

$$G_f = G_1 + G_2 + ... + G_n.$$

In a husband-wife household ($n = 2$ and $m = 2$) any differences in the net benefits of migration between spouses are associated with tied migration. The subscripts denote $1 =$ husband and

[1] The model presented in Mincer (1978) and extended in this paper assume away sociological considerations concerning gender-roles ideology or some other within-household roles held by members of the extended family. Under this assumption, the net benefit is simply a sum of un-weighted individual net benefits. Certainly, gender-roles ideology play a significant role in family migration decisions as it brings into question the sufficiency of net economic benefit as an explanation for tied-ness. For instance, if the husband assumes a role of provider in the family, the mutually recognized right to exercise power in the family migration decision will discount the wife's net economic gain (or loss) from a prospective geographic move (Bielby & Bielby, 1992).

2 = wife. The net benefit of migration to the husband and wife from migrating together are

$$G_1 \mid (s_1 = 1) = R_1 - C_1 + v(1)$$

$$G_2 \mid (s_2 = 1) = R_2 - C_2 + v(1),$$

where conditional notation is adopted to indicate i is living with s_i other family members. On the other hand, the net benefit to the husband from migrating alone is

$$G_1 \mid (s_1 = 0) = R_1 - C_1$$

$$G_2 \mid (s_2 = 0) = 0.$$

The wife is a tied mover if she faces a negative return to migration, $R_2 - C_2 < 0$, and the family gains from migrating as a unit, $G_1 \mid (s_1 = 1) + G_2 \mid (s_2 = 1) > G_1 \mid (s_1 = 0) + G_2 \mid (s_2 = 0)$ or $R_2 - C_2 + 2v(1) > 0$. This is the conclusion produced in Mincer (1978). In fact, in the $n = 2$ case, the spouse with lower returns to migration becomes tied to the other's decision.

This model differs in the case when $n > 2$, the presence of an extended family member. Suppose the household consists of a 1 = husband, 2 = wife and 3 = extended family member. If the husband and wife migrate together their net benefits are identical to the $n = 2$ case. But if the husband migrates alone leaving the wife behind with the extended family member the wife receives a larger payoff

$$G_1 \mid (s_1 = 0) = R_1 - C_1$$

$$G_2 \mid (s_2 = 1) = v(1).$$

Now the wife is a tied mover if she faces a negative return to migration, $R_2 - C_2 < 0$, and the nuclear family gains from migrating as a unit, $G_1 \mid (s_1 = 1) + G_2 \mid (s_2 = 1) > G_1 \mid (s_1 = 0) + G_2 \mid (s_2 = 1)$ or $R_2 - C_2 + v(1) > 0$. Since $R_2 - C_2 + v(1) > 0$ is strictly less than $R_2 - C_2 + 2v(1) > 0$ from the $n = 2$ case, the wife is less tied to migrate in the presence of the extended family member.

The extended family contains some husband-wife couples which I denote by the marital pairing function $\phi(i,j)$ and let m be the number of married persons. A non-marital pairing function $\eta(i,j)$ denotes a pair of any two persons in the household. In reality the family decision problem is more complex and the assumption of independence between net benefits of family members must be relaxed. A multivariate distribution of $G_1, ..., G_n$ has means $\mu_1, ..., \mu_n$, standard deviation $\sigma_1, ..., \sigma_n$, and correlations coefficients ρ_{ij} $\forall i, j$. The probability of migrating for any person in the household is $P(G_i > 0)$, the area in the distribution of G_i to the right of $G_i = 0$. Converting this area to a standardized value[2] and using it in $\mathcal{Z}_i = (0 - \mu_i)/\sigma_i = -1/CV_i$ at zero will yield the coefficient of variation measure CV_i. The probability $P(G_i > 0)$ is assumed to be the same for all potential migrants so the coefficient of variation does not differ among individuals of the household, $CV_i = CV_j$ $\forall i \neq j$.

The means and standard deviations among individuals of the household are related as $\mu_i = k_{ij}\mu_j$ and $\sigma_i = k_{ij}\sigma_j$. The parameter $k_{ij} \in [0,1]$ is a measure of the relative dominance in net benefit of individual j to i so that a smaller k_{ij} signifies greater dependence of j on the net earnings of i. Similarly, $\rho_{ij} \in [-1,1]$ measures the correlation in net benefits between i and j so that a smaller ρ_{ij} signifies greater frictions in the migration decision between family members i and j. Based on this information the CV in the extended family context can be constructed as

$$CV_f = \frac{\sigma\left(\sum_{i=1}^{n} G_i\right)}{\sum_{i=1}^{n}\mu_i} = \frac{\sigma\left(\sum_{i=1}^{m} G_{\phi(i,i')} + \sum_{i=m+1}^{n} G_i\right)}{\sum_{i=1}^{n}\mu_i} = \frac{\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i\sigma_j\rho_{ji}\right)^{1/2}}{\sum_{i=1}^{n}\mu_i}$$

where i' denotes the spouse of i in the marital pairing function $\phi(i,i')$. Additionally, $\sigma_i = k_{ij}\sigma_j$ and $\mu_i = k_{ij}\mu_j$. Adjusting the parameters k_{ij} and ρ_{ij} gives rise to multifarious situations and interdependence among individuals within the extended family structure.

The existence of one particular situation is of relevance in this study: a tied-mover of the husband-wife model becoming untied in the extended family framework. There are two

[2] Using the standard normal table (z-value).

arguments to support the intuition in regards to the existence of such a situation: (1) there is no reason to assume *a priori* that a spouse is more (or less) tied to their partner than they are to members of the extended family, and (2) a common feature of migration in developing countries are household risk distribution through remittances (Stark & Levhari, 1982; Stark & Bloom, 1985; Katz & Stark, 1986; Gubhaju & De Jong, 2009; and Mendola, 2012). Within certain parameter configurations for k_{ij} and ρ_{ij}, the independence of an individual's migration decision in an extended family structure could lead to outcomes no different from a model with only individual decision makers. The frictions in the migration decision of individuals are relaxed within an extended family setting as opposed to the predictions in Mincer (1978). These situations and related others are worth exploring using an example.

Example. Consider a three person household with one couple, $n = 3$ and $m = 2$. To be specific, let 1 = husband, 2 = wife and 3 = wife's mother. Using $\mu_2 = k_{21}\mu_1$, $\sigma_2 = k_{21}\sigma_1$, $\mu_3 = k_{31}\mu_1$ and $\sigma_3 = k_{31}\sigma_1$, the CV for the entire extended family is

$$CV_f = CV_1 \left(\frac{1 + 2k_{21}\rho_{21} + 2k_{31}\rho_{31} + 2k_{31}k_{21}\rho_{32} + k_{21}^2 + k_{31}^2}{1 + 2k_{21} + 2k_{31} + 2k_{31}k_{21} + k_{21}^2 + k_{31}^2} \right)^{1/2},$$

and the following sub-pairing assumes $k_{31} \to 0$. This gives the CV from the husband-wife model

$$CV_{\phi(1,2)} = CV_1 \left(\frac{1 + 2k_{21}\rho_{21} + k_{21}^2}{1 + 2k_{21} + k_{21}^2} \right)^{1/2}.$$

The extended family model approaches the husband-wife model under the special case when the extended family member is not dependent on the migrating married couple. But this only proves that the extended family model is a more general model than the husband-wife model. This result does not provide a testable hypothesis; instead proposition 1 below is empirically testable. The next sub-pairing is the coefficient of variation in the joint probability of migrating between the wife and her mother

$$CV_{\eta(2,3)} = CV_1 \left(\frac{k_{21}^2 + 2k_{31}k_{21}\rho_{32} + k_{31}^2}{k_{21}^2 + 2k_{31}k_{21} + k_{31}^2} \right)^{1/2}.$$

The following proposition characterizes the parameter space in which husband migration (without the wife) occurs.

Proposition. *$P(G_{\phi(1,2)} > 0 \mid s_1 = 0, s_2 = 1) > P(G_{\phi(1,2)} > 0 \mid s_1 = s_2 = 1)$ where $k_{21} \in (0,1]$ and $\rho_{21} \in [-1,1]$. The husband's probability of migrating (without the wife) is greater than the probability of migrating as a couple.*

Proof. Assume $P(G_{\phi(1,2)} > 0 \mid s_1 = 0, s_2 = 1) \leq P(G_{\phi(1,2)} > 0 \mid s_1 = s_2 = 1)$. Since $G_{\phi(1,2)} \mid (s_1 = 0, s_2 = 1) \equiv G_1 \mid (s_1 = 1, s_2 = 1)$. Then, $P(G_1 > 0) \leq P(G_{\phi(1,2)} > 0)$. Let $P(G_1 > 0) = 16\%$, which translates as $\mathcal{Z}_1 = 1$ and $CV_1 = -1$. This implies that $P(G_{\phi(1,2)} > 0) \geq 16\%$, which in turn would translate as $\mathcal{Z}_{\phi(1,2)} \leq 1$ and $CV_{\phi(1,2)} \leq -1$. Setting $CV_1 = -1$ and $CV_{\phi(1,2)} \leq -1$

$$CV_{\phi(1,2)} = (-1)\left(\frac{1 + 2k_{21}\rho_{21} + k_{21}^2}{1 + 2k_{21} + k_{21}^2}\right)^{1/2} \leq -1.$$

Some algebraic manipulation reduces the inequality to

$$k_{21}(\rho_{21} - 1) > 0.$$

Since $k_{21} \in (0,1]$ and $\rho_{21} \in [-1,1]$, the case of $k_{21} \leq 0$ and $\rho_{21} - 1 \leq 0$ can be ruled out. This only leaves $k_{21} > 0$, then $\rho_{21} - 1 > 0$ and $\rho_{21} > 1$ which contradicts $\rho_{21} \in [-1,1]$. Thus, $P(G_1 > 0) > P(G_{\phi(1,2)} > 0)$. □

The proposition states that husband migration (without the wife) is the outcome of a family migration decision process under very reasonable parameter conditions of $\rho_{21} \in [-1,1]$ and $k_{21} \in (0,1]$. The above example is a useful exposition, and more cases may be derived, to present the case that migration decisions in an extended family setting will reduce the potential of tied-movers. Migrant husbands (or wives) are likely to appear as the extended family structure becomes more complex. The interactions between the strength of correlations between family members and the level of dominance of some members of the household create a variety of instances in which migrant husbands (or wives) emerge.

The relationship between martial status and migration decisions when viewed from the perspective of extended families, assuming all other factors unchanged (such as occupational choice and location decisions), the nuclear family's net gains in migration are not a sufficient condition for migration of the marital couple. The net gains to the couple must be accompanied by normal conditions in relation to the other members of the extended family. For the marital couple to migrate, either spouse should have sufficient correlations in the net gains (or losses) with each other and much less correlation with other members of the family. Also, the similarity in returns from migration between family members will affect how tied they are to others in the household.

2.3 Empirical findings

To test the implications of the model I use households recorded in the Demographic and Health Survey (DHS) for Nepal 2011. Nepali individuals and/or families have migrated to many different countries around the world but for the purposes of this study we are concerned with households that have supplied migrants to India or within Nepal. The DHS records demographic, health and economic information of 10,826 households in Nepal, out of which 5,876 (54.28 percent) are migrant households. This data set is a more reliable indicator of actual migration than the census because of the extent of undocumented migration that occurs between India and Nepal[3]. The 2011 Nepal census recorded 25.42 percent of all households with an absent migrant member, small in comparison to the 54.28 percent recorded in the DHS. Furthermore, the DHS is better for the purpose of studying the structure of the household. Since the unit of analysis is the household in the sending country, there is likely to be some attenuation bias; that is, entire households have migrated leaving no extended family behind[4]. Attenuation bias is not a major issue using the DHS because the number of migrants from these households (11,215; 18 percent of all individuals) closely resembles the Nepal Living Standards

[3]More so due to the trafficking of women and children (Datta, 2005). The difficulty in accurately estimating migration numbers is presented in Sharma & Thapa (2013).

[4]This was mentioned previously as leaving an 'empty household'.

Survey 2010-2011 figure of 20 percent.[5]

Nepal-India migration is a widely studied area of economic and sociological research. India and Nepal maintain open borders and have close historical, cultural and economic ties (Dutt, 1981). Nepali migration to India is concentrated in certain occupations and geographic regions of India. Nepali's are involved in the tea gardens of West Bengal (Datta, 2005), as sex workers in Calcutta (Datta, 2005), as watchmen in factories (Dutt, 1981), and as *gurkhas* of the Indian army (Dutt, 1981). The economic factors that push Nepali migration are caused by "increasing fragmentation of landholding, indebtedness, ecological crisis through intense cultivation and deforestation, rising population without further land to cultivate and chronic deficits in food production" (Dutt, 1981). The pull factors from remittances (Seddon, Adhikari & Gurung, 2002) are important as well. By excluding Nepali's that have migrated to other countries the sample is reduced by 38 percent. Restricting to this smaller sample of migrants I avoid a bulk of the issues relating to unobservable (and unmeasurable) political and cultural differences between Nepal and the receiving country. Some of the receiving countries that are excluded from this sample are Saudi Arabia, Qatar, Malaysia and Dubai. The Gulf countries do not offer citizenship to Nepali's and have very strict visa requirements for migrant workers. The visa requirement for Gulf countries entail that wives must be sponsored by the husband and all husbands must be holding a full-time job. Thus all migration to the Gulf countries must involve a 'tied-spouse' or a 'trailing spouse' to the husband's migration decision. Tied-ness in this sense would not be due to the controlled net benefits of migration but from exogenous visa requirements. Including this sample would not serve as an appropriate testing ground for the implications of the theoretical model as it would severely bias the results.

The empirical hypothesis of this study is: spouses are less tied in their migration decisions within extended family settings relative to nuclear families. Firstly using statistics from the

[5] I did not use the Nepal Living Standards Survey 2010-2011 because it is not easily available to researchers from non-Nepali institutions.

DHS, I present the relevance of the extended family in Nepal within the migration context. Secondly, I show that migrant husbands (without their wives) are significantly more likely in extended families than nuclear families.

Descriptive analysis: the extended family

In this section, I present the relevance of extended families and their characteristics for migration in Nepal by the following three implications: (1) how relevant is the extended family structure among Nepali households? (2) among the households' migrant members, is the dominant motive to migrate based on factors related to private calculus? and (3) what is the extent of migrant husbands among Nepali households?

The justification of the relevance of the extended family structure in Nepal is easily established. Based on the relationship to the head of the household, 7,834 (72 percent of all households) Nepali households are male headed and 2,992 (28 percent) are female headed. Among the male headed households presented in table A1 located in the appendix: 1,786 (23 percent) households have a son/daughter/parent in-law living within the household, 1,706 (22 percent) have grandchildren living within the household, 464 (6 percent) have a brother/sister living within the household, 874 (11 percent) have some other relative (uncle/aunt/cousin/etc.) living within the household, and 247 (3 percent) have an unrelated person (renter or friend) living within the household. Similarly, within the female headed households: 648 (22 percent) households have a son/daughter/parent in-law living within the household, 538 (18 percent) have grandchildren living within the household, 147 (5 percent) have a brother/sister living within the household, 372 (12 percent) have some other relative (uncle/aunt/cousin/etc.) living within the household, and 56 (2 percent) have an unrelated person (renter or friend) living within the household. The extended family structure does not differ between male and female headed households; thus, there is no advantage to separating male and female headed households for the purpose of analyzing the extended family structure.

Whether there is a difference in the extended family structure between migrant and non-migrant households must also be tested. Based on the results of table A2 located in the appendix there are 4,950 (46 percent of all households) Nepali non-migrant households and 5,876 (54 percent) migrant households. Among non-migrant households: 676 (14 percent) have a son/daughter/parent in-law living within the household, 610 (12 percent) have a grandchild living within the household, 304 (6 percent) have a brother/sister living within the household, 560 (11 percent) have some other relative (uncle/aunt/cousin/etc.) living within the household, and 173 (3 percent) have an unrelated person (renter or friend) living within the household. On the other hand, among migrant households: 1,758 (30 percent) have a son/daughter/parent in-law living within the household, 1,634 (28 percent) have a grandchild living within the household, 307 (5 percent) have a brother/sister living within the household, 686 (12 percent) have some other relative (uncle/aunt/cousin/etc.) living within the household, and 130 (2 percent) have an unrelated person (renter or friend) living within the household. There is a noticeable difference in the extended family structure between migrant and non-migrant households. Many more migrant households have in-laws and grandchildren present[6]. Thus it seems that family structure has some relation to migration.

To show that the dominant motive for migration is based on private calculus alone, that is, the justification of implication (2), I examine the motive for migration (these are illustrated in table A3 of the appendix). There are many motives for migration: work, study, marriage, other family, or security. It would be unreasonable to suppose that an individual's incentive to migrate involves economic motives but removes from the non-economic ones. That being

[6]Note that migrant households are defined as a household that has atleast one member living outside of Nepal or within another region in Nepal. As such, the statistics presented of migrant households does not include the migrant members that have moved away. This is because the data only captures the characteristics of the household after the member has moved away. Moreover, I cannot know with certainty the member's relationship (son/daughter/cousin/etc.) to the head of the household. Migrant households are larger than the picture that is presented here.

said, I am interested in a particular incentive with an ultimate view in mind, that is, incentives aimed at understanding the tied-movers (if any) involved. This particular aim removes study and marriage as a relevant incentive because it should not involve any tied-movers. Work, other family and security related motives are relevant for private calculus and may involve some tied-movers. Only 7 (0.07 percent) households reported members moving for security reasons, and 890 (8.22 percent) households reported members that moved due to other family. I will rule out security as a relevant motive for migration due to its small number. Other family, although small in proportion, does seem to matter for migration. Other family is an important motive for understanding tied-movers because there is good reason to suppose that pioneering migrant members from a household had economic motives for migrating but subsequent members migrated for other family reasons, although these might actually be economic reasons. Additionally, the 'trailing spouse' would fall under this category.

Migration might occur in groups or as a chain. To determine the intensity of 'chain migration' (migrant members that leave the household and move to the same location as previous migrant members) as a motive behind Nepali migration it must be compared to 'group migration' (a group of individuals that move out of the household and are destined to the same location and in the same time period). Frequencies for group and chain migration are displayed in table 3. It seems that neither chain migration nor group migration is common for Nepali migration. Nepali households supply migrants to various locations with some randomness in the years. This is indicative of a risk diversification motivation in the supply of migrants as proposed by Radu (2008). The evidence for risk diversification is indicated by the frequency of households that experience chain migration to those that experience group migration, 798 (14 percent of all migrant households) and 426 (7 percent) respectively. A much larger number, 4,100 (70 percent) households exhibit neither group nor chain migration[7].

[7] And 552 households (9 percent) exhibit group and chain migration.

Table 3: Number of households experiencing group and/or chain migration, or neither

		Group Migration			
		No		Yes	
Chain Migration	No	4100	(69.78)	426	(7.250)
	Yes	798	(13.58)	552	(9.394)

Percent in parentheses

The final implication (3) is difficult to establish as the DHS does not record the migrant's relationship to the head of the household. This relationship information is essential to determine the extent of tied-movers as would be predicted by the husband-wife model. Tied-movers in group migration is commonly understood as a nuclear portion of the extended family migrating while tied-movers in chain migration is associated with the 'trailing spouse', and none of these were of great importance among migrant households. However, this evidence does not completely remove the possibility of tied-movers in the decision to migrate from extended families because I am not dealing solely with husband-wife couples. The inability to identify migrant members' relationship to non-moving members of the household within the DHS data, I reason in the following way: children under the age of 5 are unlikely to migrate without their mother/father[8], and thus the children among non-movers/movers must have their mother/father present as well. Table 4 depicts the incidence of migrant households with children living elsewhere or within the household, presumably with their father and/or mother. A very small number of migrant households have children under the age of 5 living elsewhere, presumably with their mother/father, 439 (7 percent of migrant households); and many more migrant households have children under the age of 5 living with non-movers, presumably among them is their mother/father, 2,351 (40 percent)[9].

Implication (3) requires that children and mothers (fathers) that remain with the non-movers must also have husbands (wives) that migrated. As previously mentioned, I am not

[8]Only two children were reported to be without a parent or caretaker present.

[9]The remainder 53 percent of migrant households have no children under the age of 5 within the household.

Table 4: The number of migrant households with children that live at home and abroad, by number of migrant children

Number of children	Migrant children		Non-migrant children	
0	5437	(92.53)	3525	(59.99)
1	329	(5.599)	1520	(25.87)
2	95	(1.617)	680	(11.57)
3	11	(0.187)	125	(2.127)
4	3	(0.0511)	22	(0.374)
5			4	(0.0681)
6	1	(0.0170)		
Total	5876	(100.00)	5876	(100.00)

Percent in parentheses

able to determine the relationship of the persons that migrated to the persons that remained. Instead I do have parent alive (or dead) and living within households (or not) information of children under the age of 5 among those that remained in Nepal. I identify children under 5 years as the starting point; then determine whether the child's father and/or mother is alive; if the father/mother is alive, whether he/she lives in the household and is currently identified as married; if he/she does not live in the household then he/she must be migrant; I pair the alive mother living in the household with the alive father living elsewhere and call them the 'migrant husband' pair. Similarly, I am able to identify the 'migrant wife' pair. This reasoning is presented as figure 1 in the appendix and the results are presented in table 5. No household reported a migrant wife, but a larger number, 1,418 households (24 percent), reported a migrant husband.

The final step is to determine the frequency of 'migrant husband-wife' pairs; that is, the incidence of husbands and wives moving together. I deconstruct group migration into husband-wife pairs. Due to the inability of directly extracting relationship information among migrant members, I use an alternate route: two migrant members must be migrating to the same location but be of different sex[10]. Although a very inaccurate and generalized proxy for migrant

[10]They may migrate at different times to account for the trailing spouse.

Table 5: The number of households with migrant husband and migrant wife pairs, by number of migrant members

Number of migrant members	Migrant wives		Migrant husbands	
0	5876	(100.00)	4458	(75.87)
1			954	(16.24)
2			404	(6.875)
3			48	(0.817)
4			11	(0.187)
7			1	(0.0170)
Total	5876	(100.00)	5876	(100.00)

Percent in parentheses

couples[11], the frequency estimated is an upper bound; a more accurate measure of husband-wife couples would give a smaller frequency. The results presented in table 6 indicate that 1,219 (21 percent of migrant households) households supplied couples. Compared to the 1,418 households that supplied migrant husbands, a more accurate measure of couples would only justify implication (3) further. Thus, husband-wife migration is not more likely than migrant husband pairs and removes the prominence of tied-movers when considering extended families.

I have shown that extended families are an important feature of Nepal and that husband-wife migrant pairs are not more likely than migrant husbands. It is clear from the descriptive exposition so far that migration within an extended family is possible, that migration is based on private calculus, and tied-movers, i.e. migrant husband-wife pairs, are not as common as migrant husbands within the extended family framework of Nepal.

Regression analysis: migration from the extended family

Household wealth distribution, gender distribution within the household, age and gender of the household head, education/literacy level of the household, location of the household, stage of family life cycle and presence of extended family members, need to be considered as factors

[11]The measure of couples used here would include a brother/sister pair, unmarried couples, mother/son, father/daughter, etc.

Table 6: The number of households with migrant husband-wife pairs, by number of migrant pairs

Number of migrant pairs	Husband-wife pairs	
0	4657	(79.25)
1	731	(12.44)
2	289	(4.918)
3	119	(2.025)
4	51	(0.868)
5	15	(0.255)
6	8	(0.136)
7	4	(0.0681)
8	1	(0.0170)
12	1	(0.0170)
Total	5876	(100.00)

Percent in parentheses

influencing migration of household members. Further, the location of migration in India or within Nepal and age at migration of the migrant are included as controls. The thematic discussion of this article has been to infer the relevance of the extended family dynamic in the context of the migration decision. Within the regression framework I show that migrant husbands (relative to migrant husband-wife and non-migrant households) are significantly more likely to be selected from extended families than non-extended ones. Migrant husbands are about 2.5 times more likely to appear from a household with an extended family member present. This finding satisfies the implication that wives become untied to their spouses' decision to migrate in the presence of an extended family member. The empirical model to estimate is a logistic regression of the form

$$Y = 1(X'\gamma + E\alpha + \epsilon > 0)$$

where E is a dummy variable indicating the presence of the extended family member, Y is a dummy variable indicating the husband in the household is migrant, X is the set of observable controls, and ϵ is the model error. Extended families are defined as households with the presence of in-laws, grandparents, grandchildren, brother/sisters, other relatives and/or unrelated persons. The nuclear family (or non-extended families) are households that do not include these persons. I also show that migrant husbands are more likely be from Nepali households that have an extended family member present, are poor, are female headed, and the head is younger.

I present odds ratios of the logistic regression framework to study the implications of household characteristics. To be completely robust in the analysis I present the logistic regression results in sub-samples. Considering the fact that migrant households may differ by the years since migration, I run a separate regression in sub-samples 'less than one year', 'less than two years', 'less than five years' and 'less than ten years'. The results of the regressions are provided in table 7.

There is potential endogeneity in the relationship between wealth and likelihood of migration. This occurs through remittances that are not explicitly captured in the DHS. Remittances have been proven to be a great source of income and development for households and villages in Nepal and these are captured by the wealth index. Remittances are the main pull factor leading the supply of migrant members to India and other countries such as United Arab Emirates, Saudi Arabia, Qatar, and Malaysia. Seddon et. al. (2002) find that the value of remittances from outside South Asia, although a small share of total remittances, are worth far more than those from India and within Nepal. However, there has been decreasing relevance of India as a destination for remittances: 32.9 percent in 1995/96, 23.2 percent in 2003/04, and 11.3 percent in 2010/11; Nepalis are opting for better paying destinations outside of India and Nepal. That being said, although the share of remittances from India had decreased, the volume of remittances in real terms had increased: NPR 1355 million in 1991/92 to NPR 12,100 million in 2005/06 (Sharma & Thapa, 2013). The wealth index indirectly captures total

Table 7: Nepali households that have supplied at most one migrant to India or within Nepal.

Dependent Variable: Migrant husband	Less than 1 year	Less than 2 years	Less than 5 years	Less than 10 years
Extended family	2.048***	2.253***	2.418***	2.614***
	(0.374)	(0.364)	(0.328)	(0.326)
Poor	1.837**	1.664*	1.648**	1.633**
	(0.409)	(0.332)	(0.291)	(0.259)
Urban	0.683	0.722	0.704	0.635*
	(0.188)	(0.178)	(0.150)	(0.125)
Poor * Urban	0.413	0.584	0.607	0.668
	(0.255)	(0.361)	(0.330)	(0.338)
Eastern	0.565*	0.682	0.612*	0.630*
	(0.162)	(0.175)	(0.138)	(0.133)
Central	0.636	0.750	0.722	0.723
	(0.172)	(0.180)	(0.152)	(0.140)
Western	0.659	0.650	0.652	0.662
	(0.196)	(0.173)	(0.151)	(0.140)
Mid-western	0.803	0.855	0.830	0.910
	(0.204)	(0.195)	(0.163)	(0.163)
Cluster altitude	1.000	1.000	1.000	1.000
	(0.000142)	(0.000126)	(0.000108)	(0.000100)
Age of oldest child	1.005	1.005	1.010	1.017*
	(0.0138)	(0.0122)	(0.00951)	(0.00883)
Secondary education	0.289	0.357	0.503	0.466
	(0.327)	(0.365)	(0.446)	(0.380)
Higher education	0.183	0.323	0.442	0.432
	(0.207)	(0.281)	(0.334)	(0.313)
Literacy program	0.546	0.658	0.655	0.631
	(0.332)	(0.367)	(0.343)	(0.296)
Male members	0.359*	0.408*	0.369**	0.341***
	(0.159)	(0.163)	(0.130)	(0.109)
Male headed	0.244***	0.250***	0.282***	0.295***
	(0.0540)	(0.0486)	(0.0445)	(0.0417)
Age of head	0.970**	0.973**	0.973***	0.964***
	(0.0106)	(0.00957)	(0.00758)	(0.00681)
Age at migration	1.035***	1.038***	1.038***	1.035***
	(0.00913)	(0.00830)	(0.00707)	(0.00665)
Destination city: Kathmandu	0.654	0.515	0.654	0.795
	(0.342)	(0.236)	(0.253)	(0.302)
Destination city: Other city in Nepal	0.412	0.418*	0.515	0.691
	(0.204)	(0.182)	(0.190)	(0.251)
Destination city: Rural area in Nepal	0.273**	0.324**	0.342**	0.445*
	(0.133)	(0.139)	(0.123)	(0.157)
Destination city: Mumbai	0.597	0.676	0.652	1.000
	(0.379)	(0.384)	(0.311)	(0.451)
Destination city: Delhi	0.847	0.825	0.930	1.151
	(0.472)	(0.408)	(0.392)	(0.479)
Destination city: Punjab	0.420	0.399	0.452	0.598
	(0.270)	(0.231)	(0.238)	(0.308)
Destination city: Other cities in India	0.664	0.692	0.749	0.908
	(0.319)	(0.296)	(0.272)	(0.325)
Observations	1,023	1,266	1,698	2,077
Pseudo R^2	0.277	0.255	0.228	0.228

Exponentiated coefficients; Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

wealth such as income (formal and informal), inheritances, savings, as well as remittances[12]. Historical research emphasized differentials in employment opportunities between Nepal and India as a relevant impetus for hill emigration, as was the granting of land entitlements in the 1960s to keep Nepali's from moving to the Assam region of India. Wage differentials were not seen as important for driving out-migration (Subedi, 1991). The complication arises because the wealth index is capturing opportunities for members in the household from staying, and the fact that neither remittances nor opportunities are observed directly. A large wealth index may be due to accumulated remittances from supplying migrants, but an increase in the supply of migrants may be due to fewer opportunities in the poor household and a risk diversification strategy. To work around the possibility of such omitted variables bias, I estimate by years since migration. The set of estimates for recent migrants, i.e. less than one year since migration, provide conditions under which households send their first member not too long ago to avail of the benefits from remittances; thus, the lack of opportunities is the main driver for migration in this set of regressions. Interestingly, increasing the number of years since migration from one to two, two to five, and from five to ten years does not change the regression estimates greatly. Thus, omitted variables bias due to a time-dependent process is not a major concern.

The construction of the wealth index differs for rural and urban households. Some goods in the household's consumption basket differ between urban and rural households and the weighting used to construct the wealth index reflects this. Thus, the wealth index is interacted with an urban/rural indicator to take account of these differences. Surprisingly, household wealth does not seem to be a great determinant of husband migration in any of the results presented. A very different result is presented in Kotorri (2010) where the probability of

[12]The wealth index is a composite measure of a household's cumulative living standard. It does not directly capture the sources of wealth but indirectly infers them from the quantity and quality of consumption goods in the household. For instance, households that own a color TV are richer than those with a black/white TV. Similarly, those that use tap water are richer than those that use well water. The wealth index falls in line with the 'relative deprivation hypothesis' of Stark & Taylor (1989). Moreover, the wealth index captures non-market activities; a very relevant source of income for households in the developing world (Shields & Shields, 1989).

emigration among Kosovar-Albanian households is negatively related to household income[13]. Similar findings are found among Micronesian-US and Micronesian-Hawaii migrants (Akee, 2010). Hatton & Williamson (2002) provide an alternate theory that poverty is a constraint to migration since it is generally expensive to migrate. This view is not carried in the case of Nepal-India migration due to open borders and close geographic distance. On the other hand, Chiquiar & Hanson (2005) and Lacuesta (2010) find that Mexican-US migrants are selected from the middle of the wage distribution. The lack of significant selection in Nepal-India migration is a new and interesting finding that deserves its own study.

A second endogeneity issue appears in the relationship between extended families and migrant husbands. Namely, the decision to migrate by husbands may induce extended families to form. For instance, a husband from a nuclear household plans to migrate alone but invites his mother or mother-in-law to live with the family in his absence to assist in household duties. The presence of an unobserved confounding variable is the source of bias in the model. In the example described, the omitted variable is household duties. The identification strategy I use is two-stage residual inclusion (Cai, Small & Have, 2000). This approach is ideal because it always for the controlling of omitted variables more completely in the second stage of the estimation.

I identify households that have 'usable agricultural land' and use this variable as exogenous variation in the model. The reasoning is as follows: nuclear households that possess agricultural land will have a greater need for someone to work it. Thus, the migration decision of husbands would require the wife to receive some support which would come from extended family members. Usable agricultural land appears to be a strong instrument[14]. The results are displayed in table 8. Including usable agricultural land as an instrumental variable for the

[13]The empirical methodology pursued in Kotorri (2010) is similar to what I have conducted in this paper by studying the characteristics of the household with migrant members.

[14]The Wald-χ^2 estimate in the first stage regression is large and significant.

presence of extended family members increases the likelihood of migrant husbands appearing from extended families. Moreover, the change in the coefficient estimates are large. Thus, the coefficient estimates are suffering from severe omitted variables bias, but appear to increase the magnitude of the coefficient estimates rather than make them insignificant.

Table 8: Nepali households that have supplied at most one migrant to India or within Nepal, IV = usable agricultural land.

Dependent Variable: Migrant husband	Less than 1 year	Less than 2 years	Less than 5 years	Less than 10 years
Extended family	13.84**	20.11***	26.53***	13.36***
	(11.82)	(15.03)	(13.07)	(6.278)
First stage residual	0.0657**	0.0449***	0.0365***	0.0758***
	(0.0563)	(0.0340)	(0.0190)	(0.0376)
Observations	1,023	1,266	1,698	2,077
AIC	898.2	1071.5	1387.1	1628.5

Exponentiated coefficients; Standard errors in parentheses

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

Linkages to the migration system have been proved to be of importance (Root & De Jong, 1991). That is, the presence of family/kin at the destination (as opposed to the origin) might be driving the results of this model. The presence of ties at the destination implies a smaller need for familial support so that the husband would consider migrating alone rather than have his wife along with him. These linkages are not captured in the estimation model due to the lack of post-migration characteristics. Instead the regression subsample used is households that have supplied only one migrant to India or within Nepal over the last ten years.

The lack of post-migration information constraints the regression results of this paper because I cannot capture individual specific benefits from migration. The effect of individual specific characteristics are important for explaining the motivation for migration of unmarried men and women (Gubhaju & De Jong, 2009). Hence the small pseudo-R^2 estimates.

Overall, there is clear evidence that the presence of extended family members does relieve the wife's tied-ness to the husband's migration decision. The migration constraints imposed by the nuclear family dynamic are weaker in an extended family.

Selection on observables and un-observables

The DHS records a large amount of household characteristics, which were used as controls in the empirical model. However, only a few characteristics of the migrant were included in the survey, so that, the robustness of the effect of extended family members on the likelihood of husband migration might be put into question. There will be selection bias due to unobserved migrant characteristics that are interacting with the extended family variable which would in turn have no significant independent effect on the likelihood of husband migration. For instance, high skill migrants might be more likely to come from extended families and are also more likely to migrate.

This section of the paper performs a robustness check on the selection bias that are bothering the results of the regression analysis. Selection on observables and un-observables is a strategy developed in Altonji, Elder & Taber (2002b) to be able to measure the degree of selection on un-observables by using the degree of selection on observables as a benchmark. This strategy is useful in identifying the extent of omitted variables bias if indeed there is large amounts of hidden migrant information. It serves as a useful quantification tool to judge whether the regression results relating the presence of extended family members and the likelihood of husband migration is robust to inclusion of the unobserved part. I estimate jointly, the bivariate Probit model

$$E = 1(X'\beta + u > 0)$$

$$Y = 1(X'\gamma + E\alpha + \epsilon > 0)$$

where E is a dummy variable indicating the presence of the extended family member, Y is a dummy variable indicating the husband in the household is migrant, X is the set of observable controls, u and ϵ are unobserved selection bias. The second equation is the estimated model of the regression analysis and the first equation is the projection of the variable of interest E

onto the model observables $X'\gamma$ and un-observables ϵ. The selection model is formalized as

$$Proj(E|X'\gamma, \epsilon) = \phi_0 + \phi_{X'\gamma} X'\gamma + \phi_\epsilon \epsilon$$

where $\phi_{X'\gamma}$ and ϕ_ϵ are the observed and unobserved selection terms. The following condition formalizes the idea of "selection on observables is the same as selection on un-observables"

$$\phi_{X'\gamma} = \phi_\epsilon$$

The derivation of this condition and assumptions are stated in Altonji, Elder & Taber (2002b). An informal characterization of condition 1: the ratio of unobserved selection required to explain the extended family effect is

$$\frac{\widehat{\alpha}}{\left(\frac{\sigma_E^2}{\sigma_{\hat{u}}^2}\right)\left(\mathbb{E}(\hat{\epsilon}|E=1) - \mathbb{E}(\hat{\epsilon}|E=0)\right)} \equiv M.$$

If the estimate of this ratio is too high then it is highly unlikely that the un-observables, if they were observed, would explain the entire E effect. The estimate of this ratio is provided in table 2.7 below

Table 2.7: Relative amount of selection required on un-observables to explain the E effect

Parameter	Value
$\widehat{\alpha}$	1.72
σ_E^2	0.50
$\sigma_{\hat{u}}^2$	0.64
$\mathbb{E}(\hat{\epsilon}|E=1)$	0.59
$\mathbb{E}(\hat{\epsilon}|E=0)$	1.72
M	-2.50

The estimate of the ratio M indicates that a shift in the un-observables would have to be two and half times as large as the shift in the observables to be able to explain away the E effect; but this is too large for practical purposes. So it is unlikely that including more variables to this model would explain away the effect of extended family households on the likelihood of husband migration. Therefore, the estimated relationship between the presence of extended family members and the likelihood of husband migration is robust to the inclusion of unobserved factors.

2.4 Conclusion

The husband-wife model of migration is an appropriate model under certain instances e.g. to characterize migration between/within developed countries. However, it is unrealistic to suppose all migration involves a husband-wife decision. Migration from developing countries is more likely to be from extended families and this type of family structure is proven to be of importance among several developing countries around the world (Bongaarts & Zimmer, 2002). As was presented in the theoretical model of section 2, tied-movers become less likely within the extended family structure and frees up the migration decision within certain situations. The husband-wife model continues to be of relevance from the perspective of the receiving country if these pairs of migrants do not appear to have extended families in the sending country. Studies that do not account for the family dynamic in the sending country and assume tied-movers are involved in the migration decision of the husband-wife pair may only be observing special cases of the sending country's family dynamic, and the extent of tied-movers will depend on the relevance of the extended family structure in the sending country.

I present conclusive evidence to justify the implications of the model using data from Nepal. I show that extended family structures are of importance in Nepal, the dominant motive for migration is based on factors closely related to private calculus, and those normally considered to be tied-movers in the husband-wife model are not as common as migrant husbands within the extended family framework of Nepal. Within a regression framework I show that migrant husbands are significantly more likely to be supplied to India (or within Nepal) from extended families than nuclear ones. This result remains robust to alternative identification strategies and the inclusion of unobserved factors.

3 Optimal Immigration and Investment into Cultural Assimilation

3.1 Introduction

Immigration applicants are a self-selected group of individuals in the sending country (Borjas, 1987). Immigrants will select themselves for migration given their skills, ability, wealth status and their drive to succeed. Immigrants choose to migrate to other countries to improve their economic and social condition relative to their home country outcomes. But immigration is a costly process. A large portion of the costs that immigrants must incur is the cost of integration and assimilation. Researchers have studied assimilation in various aspects. A large body of work has emerged in the study of economic assimilation with emphasis on wage assimilation (Borjas, 1987, 1994, 1999; Chiswick, 1978) and job search assimilation (Bowlus, Miyairi & Robinson, 2016). The wage assimilation literature aims to identify how quickly new immigrants, facing an initial wage discount, will achieve wage parity with the native-born. Analogously, researchers in job search assimilation aim to identify the length of time required for new immigrants to mimic the job search parameters of comparable native-born. A smaller literature exists to identify degenerative health parameters among new immigrants in the health assimilation context (Antecol & Bedard, 2005; Biddle, McDonald & Kennedy, 2007). Finally, a small literature exists on civic assimilation (Vigdor, 2008). This paper is concerned with the processes in cultural assimilation (Konya, 2007; Lazear, 1999).

The seminal article in cultural assimilation is by Lazear (1999). The model developed in Lazear's (1999) article is based on an understanding that gains from diversity is the reason people of uncommon culture interact. Incentives arise for people of uncommon culture to interact with each other and exchange useful knowledge. The conditions required for such an exchange of useful information is (1) there should be very little overlap between the two cultures' sphere of knowledge, (2) the knowledge gained must be useful, and (3) learning should be low cost. Learning costs were shown to depend inversely on the size of the own cultural group, so that, if a person's own cultural group is large then there are higher costs from learning the other culture. Konya (2007) contributes to the diversity model by specifying the coordinating cultural groups as the dominant native-born and the minority immigrant. Productivity differences between the source and host country are introduced, so that, there are larger gains to migrating from poorer source countries. Without any barriers to movement, the number of immigrants from poor source countries will be larger. The new immigrants in the host country incur migration costs and may choose to incur additional learning costs to learn about the local native-born culture. Learning the local culture allows new immigrants to communicate more effectively with native-born. However, their incentives to learn are inversely related to the size of (exposure to) the immigrant group (van Tubergen & Kalmijn, 2009). Source country richness affects cultural assimilation through exposure. In this paper I introduce the role of pre- and post-immigration experience in making formal or informal learning decisions. Where pre-immigration learning alleviates communication frictions through an endowment or selection effect, while further learning post-immigration through a form of investment. Pre-immigration learning controls for selection from within the source country.

Konya (2007) and Lazear (1999) show that years spent in the host country leads to higher levels of integration. They show this using cross-sectional census data from the U.S. This same result is reiterated in Chiswick, Lee & Miller (2004) using longitudinal data from Australia. They find unanimously that years since immigration is a significant determinant of integration, but the number of years suggested by these papers is much longer than what I find in the data. I find that the majority of new immigrants will assimilate within a few years. Using data from the Canadian 2001 Census of Population I find that 81.91% of new immigrants that come from non-English speaking households whose mother tongue is not English, are 18-64 years old, and do not reside is Quebec will have learnt to speak English within their first year arrival. Similarly, 88.36% within 2 years, 89.84% within five years and 91.47% within 10 years. The preliminary evidence supports the suggestion that assimilation through investments are more likely.

New immigrants face cultural barriers that make communication with native-born difficult. Given that communication with others is an integral part of a productive society, authorities in major immigrant receiving country are increasingly interested in issues surrounding immigrant assimilation because faster assimilation and/or close integration with native-born are assumed to be beneficial for all. Alberto, Johann & Rapoport (2013) show increasing birthplace diversity among *all* persons living in major immigrant receiving countries such as USA, UK, Austria, Norway, Germany, Belgium and Canada between 1990 and 2000 (see table A1). On the other hand, the birthplace among *migrants* of these countries did not became much more diverse in the ten years[15]. In fact, the USA grew in total immigration and became less diverse. This indicates that immigrants are a growing proportion of the population but are being selected from fewer source countries. In 2011, Canada had a foreign-born population of 6,775,800 people (20.6 % of the total population). The highest proportion among the G8 countries. Asia (including the Middle East) was Canada's largest source of immigrants over the five years 2006-2011. Although the share of immigration from Africa, Caribbean, Central and South America increased slightly. Canada appears to be becoming more diverse in it's source country selection (Statistics Canada, 2013). Ultimately, governments are not interested in removing *all* immigration; *some* immigration is worthwhile and preferred. In the sphere of immigration policy design, the *type* of immigration is of greater concern.

Unlike discussions pertaining to economic assimilation where immigrants are competing with native-born for jobs and higher wages, cultural assimilation looks closely at the complementarities that exist between the two groups. The economic literature studying immigrants has tended to focus on the low/high skilled dichotomy. An additional high-skilled immigrant encourages growth in wages of comparable native-born. On the other hand, evidence exists

[15]Except for New Zealand. New Zealand took in a lot of new immigrants and from a variety of source countries

to the effect that low-skilled immigrants reduce wages of comparable native-born (Peri, 2012). The policies that are designed then reflect the preferred type of immigration. Bianchi (2013) discusses the effective policy tools available to immigration authorities for control of the skill composition among migrants. The policy tool is the cost of migration. Since skill level is directly determined by individual wealth in the sending country and the cost of migration impacts differently by altering the returns to migration for different wealth levels, it is clear that increasing the cost of migration will have the effect of increasing the amount of high skilled migration and reducing the amount of low skilled migration.

Besides economic factors; safety and security, loss of national identity, and cultural conflicts are also equally relevant factors in designing immigration policy. Thus, it does not seem likely that the low/high skilled dichotomy of Bianchi (2013) to be of increasing relevance in studying immigration policy when immigration authorities and public opinion should be just as concerned about the social and political impact of the type of immigration. A large reason for Brexit was the immigration and refugee crisis in Europe; the British voted for sovereignty over immigration policy, which is regarded as a national issue. In this paper I focus attention on the wider scope of assimilating/non-assimilating immigration dichotomy.

Why is the assimilating/non-assimilating immigration dichotomy more relevant than the high/low skilled immigration dichotomy? as previously mentioned, it doesn't seem likely that immigration authorities will screen potential migrants solely on their technical skill level. In 2013 the Canadian immigration point system allocated 28% of maximum allowable points to language proficiency, accounting for the largest portion of total skill allocation. This understanding stems from the fact that skill level may not be directly relevant to success. One could easily imagine a high-skill immigrant working a low-skill position because they do not have the networking and other tertiary skills to market their potential. Assimilation encompasses a wider notion of success for immigrants than the narrower conception of skill. Similarly, one could imagine a low skilled immigrant performing extremely well in his/her skill class because she can network and does not face the insurmountable cultural barriers that some other immigrants would face. In particular, assimilating immigration is conducive to complimentarities with native-born. Cultural assimilation focuses less on the depressing effect of low-skilled immigrants on comparable native-born wages (when native-born and immigrants are easily substitutable) and more on the complimentarities that are created through social relationships, larger business networks, and a multicultural society that caters to heterogeneous tastes.

Cultural barriers lead to frictions in communication and lower economic productivity. The alleviation of these frictions is the assimilation process through learning about the local culture. New immigrants will learn through time spent in the host country as well as through formal/informal investments into cultural accumulation. The sum of all an immigrant learns forms the cultural capital that may be leveraged to reduce the barriers in communication. Most studies view assimilation solely as a time dependent process. In this paper I take the view that assimilation is also a form of learning that accumulates through investment. Cultural capital is a form of social capital that facilitates integration of new immigrants into the local culture. van Tubergen & Kalmijn (2009) generalize three mechanisms that influence language accumulation: exposure, incentive and efficiency. Social capital negatively interacts with language accumulation through exposure. In this paper I identify the exposure channel through which source country richness affects assimilation rates by conditioning on pre- and post-immigration experience with the local culture.

The empirical strategy of this paper is as follows. Firstly, I test the predictions of the theoretical model with data from the Longitudinal Survey of Immigrants in Canada: Waves 1 to 3 where cultural assimilation is measured by proficiency in English, and investment into native-born culture is measured by formal/informal learning of English. Since the data is longitudinal, it is an improvement over cross-sections used in prior studies because I am able to observe new immigrants integrating over the course of the first 5 years in Canada and any investments into the local culture they might have made along the way. Secondly, I provide causal estimates of learning by taking advantage of the data's panel structure. Finally, I infer the presence of unobserved heterogeneities that are functioning in the assimilation process because some immigrants have a greater ability to adapt to new environments and learn.

In the next section I formulate a model that treats investment dichotomously, discuss the externalities, solve for equilibrium and the planner's optimum. In section 3 I discuss policy implications and the receiving countries' best response to equilibrium levels and composition of immigration. Section 4 of the paper tests the model using Canadian data and section 5 looks at extensions. And finally the conclusion.

3.2 Model

A random matching model is used to describe the equilibrium levels of migration and type of migration. There are two sets of agents in the model: potential migrants and native-born. I suppose there are only two countries in the world: North and South where the South is less developed relative to the North. Thus migration flows from South to North. The native-born agents are located in the North and potential migrants are located in the South. The potential migrants in the South make a decision about whether to stay in the South or migrate to the North. Additionally, potential migrants must also decide whether to assimilate or not. The decision to assimilate in this model depends on the potential migrant's investment into learning Northern culture.

Whence the migrate and/or assimilate problem has been solved, in the North there will be two subsets of immigrants: assimilating and non-assimilating. Besides these subsets there is the set of native-born. The native-born have a very small role in the model because they face no specific decision problem. The native-born have preferences over the outcomes of the model, the level of immigration and the level of assimilating immigration, but have no impact on the outcome. Potential migrants on the other hand face the same set of outcomes and are able to choose a strategy based on a set of individual-specific characteristics.

Individuals in the North are drawn together randomly so that matches are created. Efficient matching is the primary mode of production. The model introduces communication frictions that inhibit ease of production by randomly matching agents. A match between persons of similar culture are able to generate a surplus but matches between persons of different culture create no surplus. Matches between non-assimilating and native-born create no surplus while matches between non-assimilating and assimilating immigrants, or assimilating immigrants and native-born, do generate a surplus. Four scenarios can be imagined in such a model: (i) there is no migration; in this scenario there are only productive matches between native-born in the North. However, non-migrants in the South prefer to migrate North because a match between an assimilating immigrant and a native-born in the North leads to a higher surplus than in the South (although this is only implied by the assumption of a less developed South and there being atleast some Southerners willing to make an investment in Northern culture), (ii) there is only assimilating immigration; in this scenario there is a great benefit to native-born because many more productive matches can be facilitated than having no migration at all, (iii) there is assimilating and non-assimilating immigration; in this scenario there may be too much non-assimilating immigration or too little depending on the size of the immigrating group, and (iv) finally, there is only non-assimilating immigration; although the least likeliest scenario there is the possibility of a large initial non-assimilating immigrant influx that can benefit from matches with its own kind, but at the expense of native-born losing out on the matches that could have happened had atleast some of the immigrants been assimilating instead. Clearly, the first scenario is dominated by the second while the second scenario is the most efficient relative to the other two. The third scenario has contestable results as was discussed in Konya (2007) because there may be too much non-assimilating immigration. And the fourth scenario is a possibility when one thinks of a new immigrant refugee group making its way into a country.

Unlike other two-sided matching problems there is no competition among the set of agents as in the worker-employer (Roth, 1985; Crawford & Knoer, 1981; Crawford & Kelso, 1982; Albrecht & Vroman, 2002) or marriage market models (Gale & Shapley, 1962)[16]. There is no competition among potential migrants for better outcomes in the North. In fact, there may be considerable amount of cooperation because a large group of non-assimilating immigrants can be of benefit to each other when they all migrate together since it increases the chances of a productive match (as in scenario (iv) above). For instance, the ability of Miami's labor market in the 1980's to absorb the Mariel boatlift (where a large influx of unskilled Cuban immigrants made their way to the shores of Miami) was attributed to large waves of previous immigrants in the last two decades before the Mariel boatlift (Card, 1990; and Portes & Shafer, 2006). The cooperation may only be implicit through forced coordination by asylum seekers and refugees facing political strife or social issues in the sending country. Similarly, coordination can also be had through efforts in the receiving country through worker import for special purposes; for example, Chinese workers were brought from the Guandong province to build the Canadian Pacific Railway, many of these Chinese workers stayed as immigrants. Similarly, temporary foreign worker programs are common in many developed countries experiencing labor shortages in some industries. The type of coordinated effort is relevant. In the case of asylum seekers and refugees there is a negative impact on native-born. In the case of worker import, a labor market model with complimentarities between native-born and low-skilled immigrants may be better suited. That being said, I do not go into the details of the impact of coordination and cooperation and how these will arise in the current model, it is sufficient to have it as a possibility even if it is a rare one.

[16] A complete demonstration of matching and its application is provided in Sonmez & Unver (2009)

The marginal migrant creates a positive and/or negative externality. The marginal migrant has a positive externality on other immigrants because matches with positive surplus will occur. The marginal migrant may impose a negative or positive externality on native-born. The negative externality to native-born occurs if the marginal migrant is non-assimilating. The negative externality creates conflict between the two sides of the market. Non-assimilating immigrants and native-born are in conflict because of the surplus foregone had the marginal migrant been assimilating instead. The positive externality to native-born occurs if the marginal migrant is assimilating.

Positive and negative externalities

Assuming the population of the North is given as one and the population in the South is L, let m be the level of immigration and a the level of assimilating immigration. The two variables are aggregate variables but have a specific impact on native-born and immigrants. Let $\mathcal{I} = \{i|i = 1, ..., L\}$ be the set of potential migrants, $\mathcal{N} = \{n|n = 1, ..., N\}$ be the set of native-born, $\mathcal{M} = \{p|p = 1, ..., m\} \subseteq \mathcal{I}$ be the set of all immigrants, and $\mathcal{A} = \{j|j = 1, ..., a\} \subseteq \mathcal{M}$ is the set of all assimilating immigrants. Moreover, let (d_i, a_i) be the strategy pair for the potential migrant i so that d_i takes a value of one if the potential migrant decides to migrate and a value of zero otherwise. Similarly, a_i takes a value of one if the potential migrant decides to assimilate, and zero otherwise. This is the basic strategy (d_i, a_i) and outcome (m, a) profile of Konya (2007). Once the strategy pair is given for the set of all potential migrants $i \in \mathcal{I}$ then the aggregate outcomes are constructed as

$$m = \sum_{i=1}^{L} d_i \quad \text{and} \quad a = \sum_{i=1}^{L} a_i. \tag{3.1}$$

Based on the dis-aggregated outcomes, the marginal migrant imposes an externality to all other immigrants that is positive, an externality to native-born that is negative if the immigrant is non-assimilating but positive if assimilating.

Proposition 1. *(a) The externality imposed on immigrants by the marginal migrant is*

$$\frac{1+\sum_{i=1}^{L} d_i}{2+\sum_{i=1}^{L} d_i} - \frac{\sum_{i=1}^{L} d_i}{1+\sum_{i=1}^{L} d_i} = \frac{1+m}{2+m} - \frac{m}{1+m} > 0 \tag{3.2}$$

(b) On the other hand, the externality imposed on native-born by the marginal migrant is

$$\frac{2+\sum_{i=1}^{L} a_i}{2+\sum_{i=1}^{L} d_i} - \frac{1+\sum_{i=1}^{L} a_i}{1+\sum_{i=1}^{L} d_i} = \frac{2+a}{2+m} - \frac{1+a}{1+m} > 0 \tag{3.3}$$

if the marginal migrant is assimilating. If the marginal migrant is non-assimilating, the externality on native-born is

$$\frac{1+\sum_{i=1}^{L} a_i}{2+\sum_{i=1}^{L} d_i} - \frac{1+\sum_{i=1}^{L} a_i}{1+\sum_{i=1}^{L} d_i} = \frac{1+a}{2+m} - \frac{1+a}{1+m} < 0 \tag{3.4}$$

Proof. The proof of proposition 1 is in the appendix □

There isn't much more to take away from equations (2)-(4) at this time but it formalizes the externalities and will be a useful tool for analysis at a later point in the paper. The following corollary will also be useful.

Corollary 1. *The positive externality on native-born from an assimilating marginal migrant is larger than the positive externality on other immigrants from any type of marginal migrant.*

$$\frac{2+a}{2+m} - \frac{1+a}{1+m} > \frac{1+m}{2+m} - \frac{m}{1+m}$$

Proof. Equation (3) is larger than equation (2) for any m, a and $m \geq a$ which must also be true as the set of assimilating immigrants are a subset of all immigrants. □

Size and composition effect

As was discussed in proposition 1, the marginal migrant creates a positive externality for other immigrants and native-born if they are assimilating, and a negative externality on native-born if they are non-assimilating. The trade-off between size and composition works through the positive and negative externalities. The externality of equation (2) is measured simply as the gain in the expected benefit to immigrants from an additional immigrant. This additional immigrant increases the level of immigration and may change the composition. However, the definition of the externality on immigrants (exclusively a measure of benefit gain/loss for immigrants) does not capture the composition effect, only the size effect. This is because equation (2) does not require the marginal migrant to be specified as assimilating or not; purely a size effect. On the other hand, the externality on native-born of equations (3) and (4) are measured as the expected gain/loss to native-born from an additional immigrant. The externality on native-born captures the composition effect. If the marginal migrant is assimilating then there is a positive externality as the numerators of equation (3) increased at the same rate as their respective denominators. If the marginal migrant is non-assimilating then a negative externality appears. Thus the size effect is neutral to the type of immigration and is measured by the externality to all other immigrants, while the composition effect is sensitive to changes in the type of immigration relative to changes in the level of immigration and is measured by the externality to native-born. The Northern immigration authorities are interested in increasing immigration levels upto the point where the negative and positive externalities cancel each other out.

Bianchi (2013) discusses size and composition effects in the context of the equilibrium skills ratio (high-skilled relative to low-skilled immigrants). The high-skill/low-skill dichotomy is substituted for the assimilating/non-assimilating dichotomy in this model. Changes in the policy will have an impact on the size and composition of immigration by internalizing the externalities to all other immigrants and native-born. The following proposition discusses this trade-off in the size and composition of immigration.

Proposition 2. *For any $m, a > 0$ the size and composition effects trade-off according to*

$$m \begin{cases} > a(2a+1)/2(1+a), & \textit{the size effect dominates the composition effect} \\ = a(2a+1)/2(1+a), & \textit{externalities cancel out} \\ < a(2a+1)/2(1+a), & \textit{the composition effect dominates the size effect} \end{cases}$$

Proof. Using the following condition to equate the total positive and negative externalities

$$m\left(\frac{1+m}{2+m} - \frac{m}{1+m}\right) + a\left(\frac{2+a}{2+m} - \frac{1+a}{1+m}\right) = (m-a)\left(\frac{1+a}{2+m} - \frac{1+a}{1+m}\right)$$

and some algebra will solve for the level of m as $a(2a+1)/2(1+a)$ which separates the size and composition effects. □

Proposition 2 is a description of the level of immigration that satisfies Northern interests. The size of immigration is positively related to the composition, $dm/da > 0$, and is increasing at an increasing rate, $d^2m/da^2 > 0$. This is a different planner outcome from maximizing the total welfare of the incumbent native-born population. Optimizing immigration authorities in the North would solve the following maximization problem

$$\underset{a,m}{\text{maximize}} \quad \frac{1+a}{1+m} \qquad \text{Subject to} \quad (m-a) \geq 0$$

But such a policy dictates that $m = a$ is the only solution. The optimizing Northern planner is minimizing negative externalities and maximizing atleast some of the positive externalities. This policy is highly restrictive and more involved than the one described by proposition 2.

This result on size and composition effects deserves greater attention because of its direct policy relevance. The immigration policies administered by governments in major immigrant receiving countries are set by the pre-existing "base" population, which are either predominately native-born or immigrants. If the receiving country's base population is mostly native-born then there is likely to be averseness to higher levels of immigration, especially if it is the non-assimilating type. On the other hand, if the receiving country's base population is mostly immigrants then it is more open to higher levels of immigration.

Thresholds

From this point onwards I will drop the subscript i indicator for the potential migrant and discuss further the potential migrant's decision problem, how heterogeneity is introduced into the model, derive the equilibrium thresholds for migrating and assimilating, and discuss the equilibria that come out of the random matching immigration problem. The migration cost is

$$\text{Migration cost} = \mu c$$

Where c is an individual-specific migration cost that is distributed over all potential migrants in the South, $c \sim F(c)$ and $c \in [0, 1]$. And $\mu < 1$ is an index of physical distance between the North and the South. Migration costs are included in the model because the migration models developed in Chiswick (1999) and Borjas (1987) attribute these costs to selection bias. The specification of migration costs are identical to Konya (2007) but the assimilation cost is different. I introduce the decision to invest as a part of assimilation costs. That is, potential migrants are allowed to invest in themselves incurring costs of

$$\begin{aligned} \text{Assimilation cost} &= \theta\tau \\ &= \theta\nu\iota \end{aligned}$$

Where $\theta < 1$ is an index of cultural differences between the North and the South and τ is an individual-specific distribution of assimilation costs in the South as in Konya (2007). In this paper τ is replaced with investment into cultural assimilation ι and a unit cost of investing ν. Assimilation costs are determined by the level of investment that is undertaken. It encompasses all aspects of potential migrants' investments into making integration as simple as possible. This is accomplished through several avenues; through learning the official language, understanding cultural differences when living in a multicultural society, or building effective communication skills. Investment is a part of the process of cultural capital accumulation that is either amassed over time in a passive learning process. Or, formal and informal investments may be made towards learning the culture as a decisive step towards integration. This formulation of assimilation costs better accomodates differences in immigrants' potential investment into cultural capital accumulation.

Learning is introduced through the capital accumulation process. Learning is distinguished at their pre- and post-immigration levels where pre-immigration experience with the local culture is an individual-specific distribution in the source country, while post-immigration learning comes as a form of investment into learning the local culture after the migration decision has been made. Further learning in the host country are individual-specific decisions that are a function of the investment returns and costs.

Cultural capital has two components, $x = \phi_1 \iota + \phi_2 \omega$ where $x \in [0, 1]$ is the total number of native-born that the migrant can communicate with given the level of investment ι and pre-immigration experience ω. Cultural capital is dichotomous; immigrants can either talk to all native-born ($x = 1$) or some ($x = \omega$). Pre-immigration experience ω is accumulated capital or endowment. Accumulated capital ω is distributed $W(\omega)$, $\omega \in [0, 1]$. It represents accumulated learning that includes the potential migrant's experience with host country culture prior to migrating. Investment is a dichotomous variable representing further learning, $\iota \in \{0, 1\}$. New immigrants with large pre-immigration experience (ω is large) face the lowest cost of investment ($\iota = 1 - \omega$).

The parameters $\phi_1, \phi_2 \geq 0$ are constant conversion factors of pre- and post-immigration experience into the units of x. Alternatively, ϕ_1 and ϕ_2 may be interpreted as the returns to further learning and pre-immigration experience. For simplicity I suppose that $\phi_1 = \phi_2 = 1$.

Potential migrants that are assimilating can generate a surplus of one when they are matched with other immigrants or native-born. Potential migrants that are non-assimilating can generate surplus with other immigrants and the chance of meeting another immigrant is $m/(1 + m)$. Non-assimilating create no surplus when they are matched with native-born.

Those potential migrants that stay in the South will generate a surplus of $h < 1$ (this condition ensures that migration only moves from South to North; that is, matches are more efficient in the North than in the South). The potential migrant's decisions are summarized by the following set of value functions

$$V_a = 1 - \theta\nu(1-\omega) - \mu c \tag{3.5}$$

$$V_n = \frac{m}{1+m} - \mu c \tag{3.6}$$

$$V_s = h \tag{3.7}$$

The first value function is the utility from assimilating immigration. The assimilating immigrant can trade with anyone and earn a surplus of one, but must incur the cost of migration and assimilation. The second value function is the utility from non-assimilating immigration. The non-assimilating immigrant only trades with other immigrants but faces no assimilation cost. The third value function is the utility from not migrating. All meetings with Southerners will produce a surplus of h with no migration and assimilation costs incurred.

Based on equations (5) to (7), the threshold levels for migrating (or not) and assimilating (or not) can be derived in (ω, c) space. Equating (5) and (6), and solving for ω derives the threshold level for assimilating immigration such that $\omega \geq \omega_a$ there is assimilating immigration, and non-assimilating immigration otherwise. Equating (5) and (7), and solving for c determines the threshold level for migration such that $c \leq c_a(\omega)$ there is migration (assimilating or non-assimilating), and no migration otherwise (the potential migrant remains in the South). The threshold levels are

$$\omega_a = 1 - \frac{1}{\theta\nu(1+m)} \tag{3.8}$$

$$c_a(\omega) = \frac{1 - h - \theta\nu(1-\omega)}{\mu} \tag{3.9}$$

The different regions and threshold levels are depicted in figure 1 below. Note that $c_a'(\omega) > 0$, $c_a''(\omega) = 0$, and when $c_a(\omega) = 0$, $\omega = 1 - (1-h)/\theta\nu \equiv \underline{\omega}$. This determines the shape of the

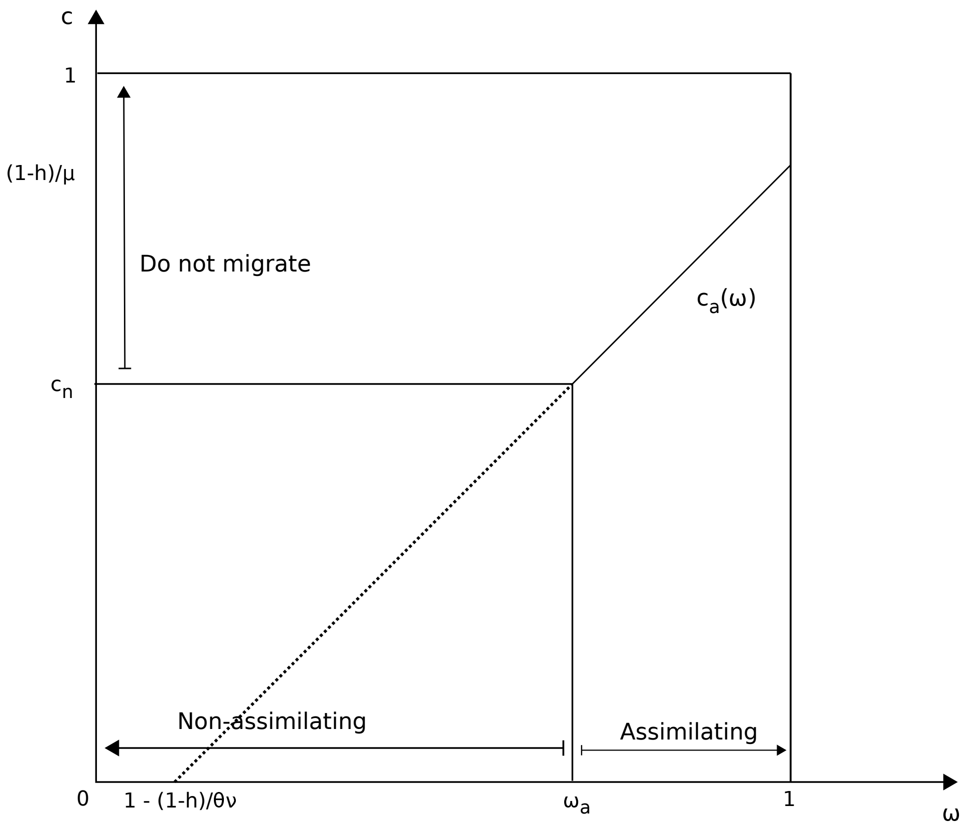

Figure 3.1: The interior equilibrium

threshold function $c_a(\omega)$. Moreover, let $c_n \equiv c_a(\omega_a)$ so that

$$c_n = \frac{1 - h - \frac{1}{1+m}}{\mu} \tag{3.10}$$

This equation characterizes the threshold for non-assimilating immigration. The level of c_n depends explicitly on m; equilibrium is not unique because non-assimilating immigration is more lucrative as the level of immigration increases.

The cultural capital term x represents formal/informal investments into cultural accumulation through learning in a class, self-study or purposefully engaging with native-born people, institutions and media. To distinguish the two components of x further, the following example will be helpful: two potential migrants A and B in a less developed and non-English speaking country are currently in limbo about migrating or not to a more developed English speaking country. Moreover, both A and B feel that assimilation is the way to go so they both must atleast learn the local language. However, relative to the predefined endowment threshold ω_a, that distinguishes between an assimilating and non-assimilating immigrant, A finds that his current language speaking characteristic does not meet the requirement for assimilating immigration, his accumulated capital in English (the culmination of his lifetime experiences with the English language) is small. For A to come on par with the threshold ω_a, he must make a formal/informal investment of at least $\iota = 1 - \omega$ into English speaking skills to be considered assimilating. The potential migrant B on the other hand does have the necessary English speaking skills that are already a part of her pre-immigration experience because she went to an English speaking school and worked for an American company for most of her life in the home country. She finds her accumulated capital ω to be more than sufficient to be part of the assimilating group, $\omega > \omega_a$. B does not require to make any more of an investment because she is already assimilating based on the distribution of the component ω, thus she sets ι to zero.

Similar to the experiences of A and B I can identify several cases where the potential migrant may face a decision on how much to invest, conditional on their endowment ω and whatever the equilibrium assimilation threshold ω_a happens to be. Accumulated capital or endowment is given before migration and those with a little accumulated capital must make a decision on whether to assimilate and invest. Four distinct cases are possible and these are depicted in Figure 3.2 below.

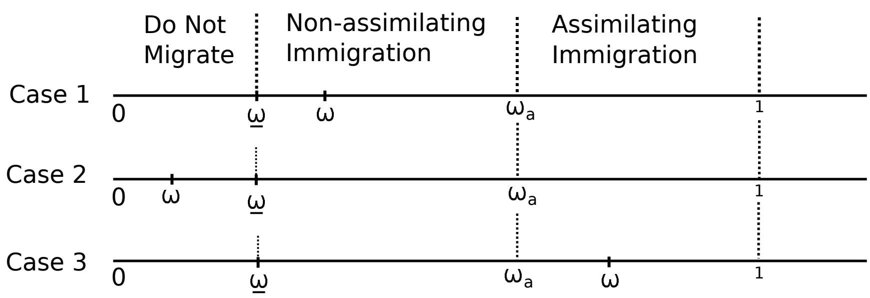

Figure 3.2: Cases for the assimilation-specific investment

The three cases in the diagram will occur when assimilation is beneficial ($V_a > V_n$). In the first case $\underline{\omega} \leq \omega < \omega_a$: the investment will be the difference $\iota = 1 - \omega$. In the second case $\omega < \underline{\omega}$: do not migrate. In the third case $\omega > \omega_a$: the investment will be $\iota = 0$. The fourth case is trivial: when assimilation is not beneficial ($V_a < V_n$), the investment will be zero. The investment is dichotomous and takes two values: zero when there is non-assimilating immigration and $\max\{1 - \omega, 0\}$ when there is assimilating immigration.

Equilibrium

The first two conditions describe a mixed equilibrium where both types of immigration is present.

$$m - a = LF(c_n) \tag{3.11}$$

$$a = L \int_{\omega_a}^{1} F(c_a(\omega))dW(\omega) \tag{3.12}$$

When the assimilation threshold is large enough $\omega_a > \underline{\omega}$, the equilibrium is said to be *interior* with assimilating and non-assimilating types. From $\omega_a > \underline{\omega}$ I can infer that $h < m/(1 + m)$, the source country must be sufficiently poor for mixed immigration. The left-hand side of (3.11) is the actual number of non-assimilating immigrants and the right-hand side is the expected number of non-assimilating immigrants. Similarly, the left-hand side of (3.12) is the actual number of assimilating immigrants equated to the expected number of assimilating immigrants.

In a *corner equilibrium* all immigration is assimilating. The corner equilibrium occurs when the assimilation threshold is small enough $\omega_a = \underline{\omega}$. That is,

$$m = L \int_{\underline{\omega}}^{1} F(c_a(\omega))dW(\omega) \tag{3.13}$$

and $a = m$. Given that $\omega_a = \underline{\omega}$, I can infer that the source country must be rich enough for purely assimilating immigration $h = m/(1 + m)$. Combining equations (3.11) - (3.13), a single equation characterizes the three possibilities

$$m = LF(c_n) + L \int_{\max\{\underline{\omega}, \omega_a\}}^{1} F(c_a(\omega)) dW(\omega). \tag{3.14}$$

The first term of equation (3.14) is the number of non-assimilating immigrants, which becomes zero when the corner occurs (i.e. $c_n = 0$ and $F(0) = 0$). The second term is the number of assimilating immigrants due to a large endowment effect, which is always positive but varies in size depending depending on whether the corner (i.e. $\omega_a = \underline{\omega}$) or interior (i.e. $\omega_a > \underline{\omega}$) case occurs. The following table summarizes the outcomes given ω_a

Table 1: Interior or corner equilibrium outcomes given ω_a.

k	ω_a	h	Outcome
≤ 1	$= 1 - (1-h)/\theta\nu$	Rich	Corner
< 1	$> 1 - (1-h)/\theta\nu$	Poor	Interior

It is clear that if the sending country is rich enough then all immigration is assimilating. When the sending country is poor then immigration is mixed. The following proposition states this result formally.

Proposition 3. *Let (a^*, m^*) be a unique solution to the immigration problem. There exists an $\overline{h}(\theta, \mu, \nu, L)$ satisfying*

$$\frac{\overline{h}}{1-\overline{h}} = L \int_{\underline{\omega}}^{1} F\left(\frac{1-\overline{h}-\theta\nu(1-\omega)}{\mu}\right) dW(\omega) \tag{3.15}$$

such that if $h > \overline{h}$, then a corner equilibrium occurs. And if $h < \overline{h}$, then the equilibrium is interior.

Proof. In the interior/investor equilibrium $\omega_a > \underline{\omega}$ or analogously $c_n > 0$. Define $\overline{m}$ by the equation $c_n > 0$ so that $\overline{m} > h/(1-h)$. Since c_n is increasing in m, the interior/investor equilibrium occurs for any $m^* > \overline{m}$, and corner otherwise. Thus, $\overline{m}$ is the cut-off between the two equilibrium, and at the cut-off both equilibrium occur. So substituting $\overline{m}$ into (3.15) gives the left-hand side of (3.15). Finally replace everywhere $h = \overline{h}$ so (3.15) solves for the seperating level of h.

The left-hand side of (3.15) is increasing and the right-hand side is decreasing in $\overline{h}$. Additionally, the left-hand side is less than the right-hand side at $\overline{h} = 0$, and the left-hand side is greater than the right-hand side at $\overline{h} = 1$. Thus, there is a unique $\overline{h} \in (0, 1)$ that satisfies (3.15). Finally, since $dm/dh < 0$[17], if $h < \overline{h}$ then $m^* > \overline{m}$, and if $h > \overline{h}$ then $m^* < \overline{m}$. □

Rich countries would send only assimilating immigrants and poor countries send some non-assimilating immigrants aswell. The model results are identical to those produced by Konya (2007) but re-stated in terms of pre-immigration experience rather than assimilation costs.

The intuition behind proposition 3 is that efficiency of matches in the South has an impact on the level of immigration. As the South gets poorer there is more migration from the South because matches in the North become relatively more productive. That is, incurring the migration (and possibly assimilation) costs, and foregoing the matches that could have been made in the South (the opportunity cost), the potential migrant finds these costs are small relative to the gains from matches in the North. This is the *productivity effect* described in Konya (2007). Alternatively, non-assimilating immigrants in the North experience a debilitating *communication effect* because they cannot interact efficiently with native-born. In Konya (2007), if the South is sufficiently poor then the productivity effect dominates the communication effect so the benefit of migrating without assimilating gets larger. In the current setup of the model the communication effect is fully determined by a *selection* effect. The likelihood of investing in further learning given the endowment has the potential to strengthen the communication effect post-immigration. The selection effect functions through the endowment. A large endowment implies a larger pre-immigration exposure to the local culture.

To be able to clearly see the relationship between h and m, a convert the variables to

[17] See comparative statics in the appendix.

common units $\alpha = m/L$ and $\beta = a/L$. Using the fact that

$$\frac{\partial \beta}{\partial h} = -\frac{1}{\mu} \int_{\omega_a}^{1} F'(c_a(\omega))\, dW(\omega) < 0,$$
$$\frac{\partial^2 \beta}{\partial h^2} = \frac{1}{\mu^2} \int_{\omega_a}^{1} F''(c_a(\omega))\, dW(\omega) > 0,$$
$$\frac{\partial \alpha}{\partial h} = \frac{\partial(\alpha - \beta)}{\partial h} + \frac{\partial \beta}{\partial h}$$
$$= -\frac{1}{\mu} F'(c_n) + \frac{\partial \beta}{\partial h} < 0$$
$$\frac{\partial^2 \alpha}{\partial h^2} = \frac{1}{\mu^2} F''(c_n) + \frac{\partial^2 \beta}{\partial h^2} > 0$$

and $(\alpha - \beta)$ is simply the vertical distance between α and β. Figure 3.3 depicts the cases when the source country is poor ($h < \overline{h}_1$) or rich ($h \geq \overline{h}_1$).

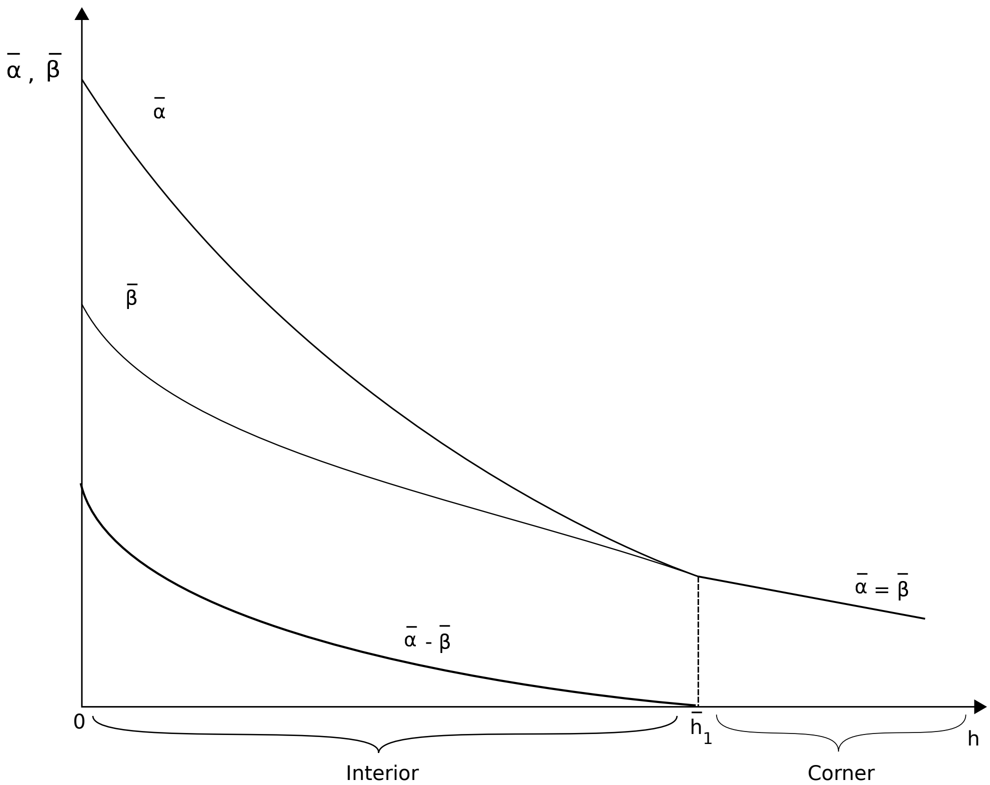

Figure 3.3: The equilibrium outcomes

Social planner optimum

This section discusses outcomes from the perspective of a hypothetical social planner. The social planner discussed in this paper is a world planner. A discussion of socially optimal outcomes is necessary to be able to judge the welfare properties of the equilibrium outcomes in the previous section. The world planner is not simply interested in maximizing welfare in the North, but also the South. The world planner solves the following problem

$$\underset{\alpha,\beta}{\text{maximize}} \quad (1-\alpha)hL + \beta L + (\alpha-\beta)\frac{\alpha L}{1+\alpha L}L + \frac{1+\beta L}{1+\alpha L} - C(.)L$$

$$\text{Subject to} \qquad \alpha - \beta \geq 0$$

The first term is the benefit to remaining in the South, the second term is the benefit to assimilating immigration, the third term is the benefit to non-assimilating immigration and the fourth term is the benefit to native-born. The inequality constraint accounts for the possibility of the corner outcome in which $\alpha = \beta$. The last term in the planner's problem is the cost function. The cost function is given by

$$C(\alpha, \beta; \omega, \nu, \theta, \mu, c) = \beta(\theta\nu(1-\omega) + \mu c) + (\alpha - \beta)(\mu c) \tag{3.16}$$

The first term in the cost function is the cost of assimilating and migrating. The second term is the cost of non-assimilating immigration. The first-order conditions from the maximization problem are

$$\phi_\alpha \equiv -hL + \frac{\alpha L}{1+\alpha L}L + (\alpha-\beta)\frac{L}{(1+\alpha L)^2}L - \frac{1+\beta L}{(1+\alpha L)^2}L - L(\mu c) + \lambda = 0 \tag{3.17}$$

$$\phi_\beta \equiv L - \frac{\alpha L}{1+\alpha L}L + \frac{L}{1+\alpha L} - L(\theta\nu(1-\omega) + \mu c) + L(\mu c) - \lambda = 0. \tag{3.18}$$

The second-order condition in terms of h yields the following results for the planner optimal $\alpha = \widetilde{\alpha}$ and $\beta = \widetilde{\beta}$

$$\phi_{\alpha h} = -L < 0$$

$$\phi_{\beta h} = 0.$$

Plotting the relationship between h and α, β shows the threshold level of $h = \widetilde{h}_1$ between the interior and corner outcomes. The two cases are depicted in Figure 3.4 below: the interior case which occurs for $h < \widetilde{h}_1$, and the corner case which occurs for $h \geq \widetilde{h}_1$.

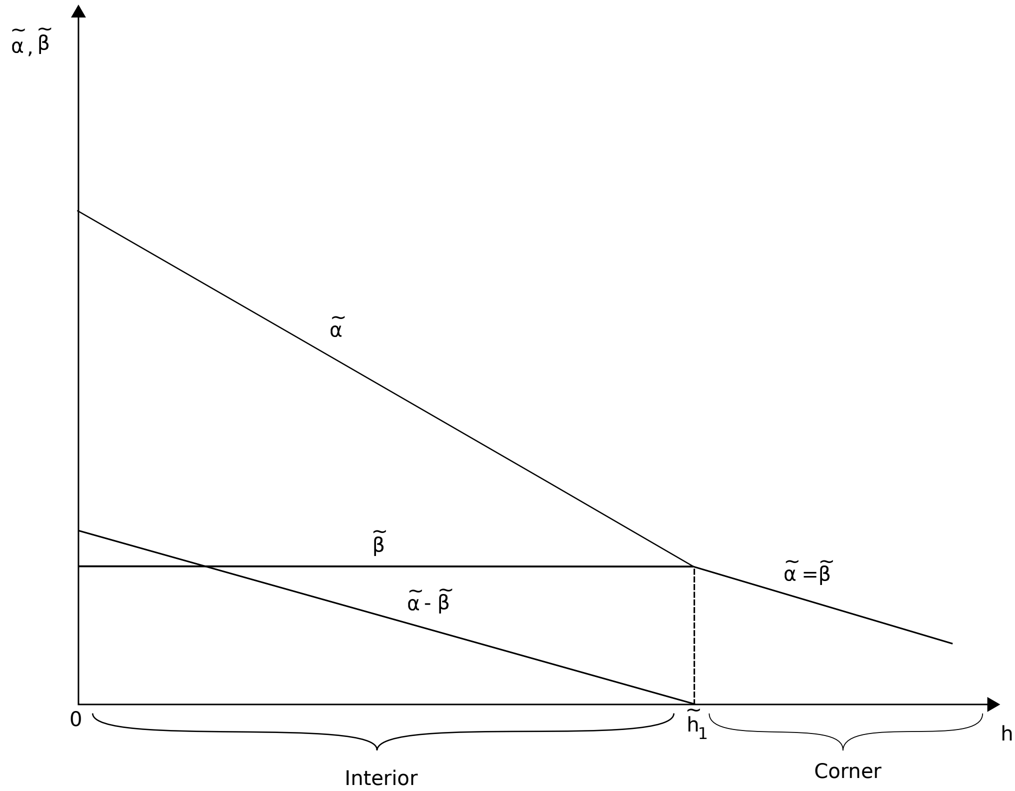

Figure 3.4: The social planner optimum

Using the first order condition (3.17) and applying the conditions for the corner case $\alpha = \beta$ and $c_n = 0$, as well as the interior solution $\lambda = 0$ the system becomes

$$\phi_\alpha \equiv -\widetilde{h}L + \frac{\alpha L}{1+\alpha L}L - \frac{L}{1+\alpha L}L = 0 \tag{3.19}$$

$$\phi_\beta \equiv L - \frac{\alpha L}{1+\alpha L}L + \frac{L}{1+\alpha L} - L(\theta\nu(1-\omega) + \mu c) = 0 \tag{3.20}$$

Solving (3.19) for αL in terms of $\widetilde{h}$ and since $\alpha L = m$,

$$m = \frac{1+\widetilde{h}}{1-\widetilde{h}}. \tag{3.21}$$

The $\widetilde{h}$ required to solve (3.19) is smaller than the $\overline{h}$ required to solve for the left-hand side of (3.15), therefore $\widetilde{h} < \overline{h}$.

There are three regions to consider. The trivial first case is for very rich sending countries, $h > \overline{h}$, the corner optimum and equilibrium coincide. This is in fact a global optimum (already proved in footnote (8) of Konya (2007)). Among the poorest sending countries, $h < \widetilde{h}$, the interior optimum and equilibrium coincide. In this case, it is optimal to have some non-assimilating immigrants. For intermediate countries, $\widetilde{h} < h < \overline{h}$, the interior equilibrium coincides with the corner optimum. There is too much non-assimilating immigrants when there should be none. Finally, for poor countries, $h < \widetilde{h}$, is not optimal in the composition and size of immigration.

The social planner cut-offs for assimilating immigration, non-assimilating immigration and migration are obtained by solving the first-order conditions (3.17) and (3.18). These are as follows:

$$\widetilde{\omega}_a = 1 - \frac{2}{\theta\nu(1+m)}, \tag{3.22}$$

$$\widetilde{c}_a(\omega) = \frac{1 - h - \theta\nu(1-\omega)}{\mu} + \frac{2(m-a)}{\mu(1+m)^2}, \tag{3.23}$$

$$\widetilde{c}_n = \frac{1 - h - \theta\nu(1-\omega_a)}{\mu} + \frac{2(m-a)}{\mu(1+m)^2}. \tag{3.24}$$

A relationship between m and a may be constructed that shows how changes in the size of immigration are associated with changes in the composition of immigration. In a corner optimum, this relationship is trivially $m = a$. In the interior, the relationship is given by

$$m = L \int_{\widetilde{\omega}_a}^{1} F(\widetilde{c}_a(\omega)) dW(\omega) + a \tag{3.25}$$

where $\widetilde{c}_n$, $\widetilde{c}_a(\omega)$ and $\widetilde{\omega}_a$ denote the social planner cut-offs of (3.22)-(3.24). Since $\widetilde{c}_n > c_n$, $\widetilde{c}_a(\omega) > c_a(\omega)$, $\widetilde{\omega}_a < \omega_a$, and $\underline{\widetilde{\omega}} < \underline{\omega}$, it can be shown that the planner's optimal level of immigration must be less than or equal to the equilibrium level. The following section depicts these results graphically.

3.3 Policy

In this section I will discuss the policy options available for immigration authorities in the North to achieve first-best results in the size and composition of immigration. Assuming that

Northern interests coincide with the joint maximization of immigrant and native-born interests (excluding those that remain in the South), it is shown that Northern interests will diverge from the social planner optimum and a unique (m, a) combination exists where Northern and social planner interests coincide.

Figure 3.5 describes the choices available to the Northern planner when choosing the size and composition of immigration with m and a. The equation $m = a(2a+1)/2(1+a)$ identifies the set of optimal combinations $m, a \geq 0$ such that the total positive and negative externalities associated with the size and composition of immigration cancel each other out. This was derived as proposition 2 and depicted as the "externalities" line in figure 3.5. The equation $m = a$ describes the corner optimum case when negative externalities associated with non-assimilating immigration are set to zero and atleast some of the positive externalities associated with assimilating immigration are available; the "corner" solution in figure 3.5. The third optimality proposition is the "interior optimum" given by the function (3.25). And equation $m = LF(c_n) + a$ is the "interior equilibrium" line derived from the single equation equilibrium (3.14) when the interior occurs and the integral denotes a. Moreover, $LF(c_n) \geq 0$.

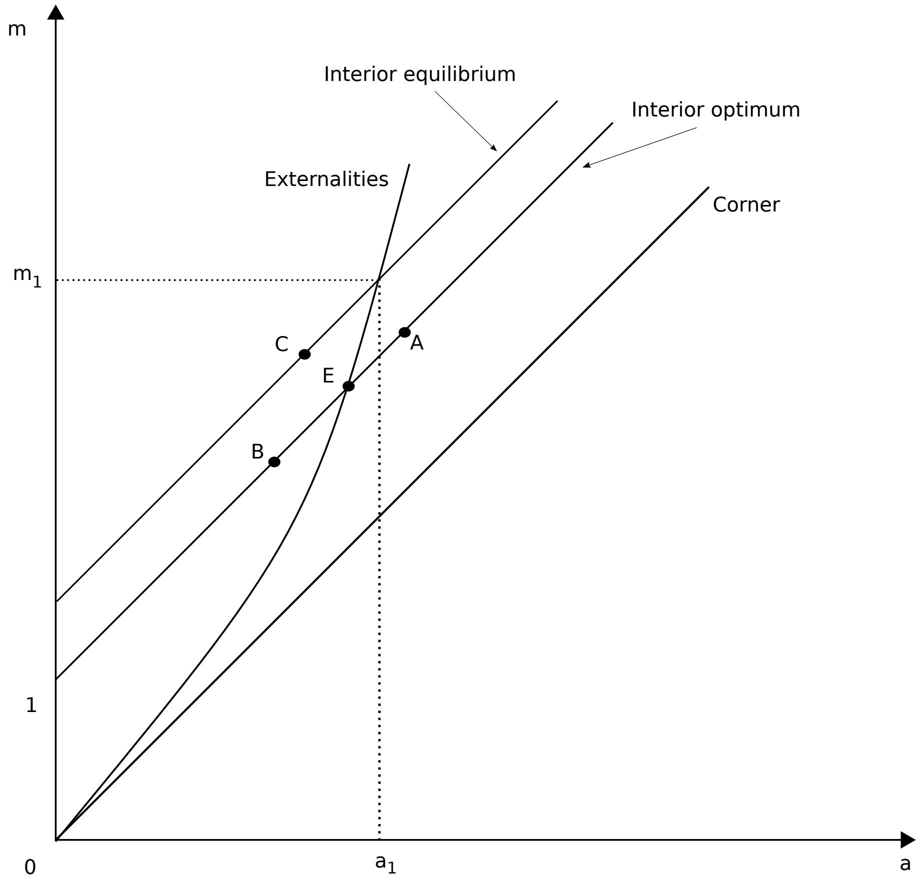

Figure 3.5: Policy outcomes in Northern interests

The line denoting the interior equilibrium outcome approaches an interior optimum in the special case when $F(c_n) = F(\widetilde{c_n})$. Or the corner optimum when either $c_n = 0$. However, the second best outcome (a_1, m_1) is still available. In this next best case, the total positive and negative externalities cancel each other out. Total immigration of m_1 will be accepted with a_1 levels of assimilating and $m_1 - a_1$ levels of non-assimilating immigration.[18]

Point system

Language is a crucial component to admitting potential migrants. In Canada, the point system for admitting immigration applicants was introduced in 1967. Language accounted for 21% of total points required to be granted permanant residency status in 1996; education also recieved the same number of points (Green & Green, 1999). In 2013, language accounted for 28% of total

[18]Visually, the level of non-assimilating immigration is the vertical distance between m_1 and the $m = a$ line.

points; 3 percentage points higher than education (Tani, 2014). In this section of the paper I discuss the role of pre-immigration experience in serving as a signal to immigration authorities of the receiving country in selecting potential migrants. I also discuss the implications of a stricter selection policy.

The total cost to assimilating immigration $\theta\nu(1-\omega)+\mu c$ is decreasing in ω. Since assimilation costs are decreasing in pre-immigration experience, those potential migrants with the largest ω will be the first to migrate. Since ω is fully observable to the planner. Let $\overline{\omega}$ be the minimum language skill requirement for admittance to the North so that all potential migrants with $\omega \geq \overline{\omega}$ are admitted, and denied entry otherwise. Moreover, suppose that $\overline{\omega} < \omega_a$. Under this system, the size of immigration will decrease by the number of non-assimilating immigrants. As follows

$$m - a = L\left(W(\omega_a) - W(\overline{\omega})\right)F(c_n)$$
$$a = L\int_{\omega_a}^{1} F(c_a(\omega))dW(\omega).$$

There is a negative effect on the size of immigration and a positive effect on the composition. In figure 3.5, a restrictive policy on the selection of immigration will move the interior equilibrium line rightwards, closer towards the interior optimum policy line.

On the other hand, when $\overline{\omega}$ is set too high so that $\overline{\omega} > \omega_a$, then all immigration would be assimilating:

$$m = L\int_{\overline{\omega}}^{1} F(c_a(\omega))dW(\omega).$$

This is associated with an extreme rightward shift of the interior equilibrium line in figure 3.5 till it coincides with the corner optimum policy line. However, there is potential loss of surplus from the minimum skill requirement being set too high. The lost surplus is the assimilating immigration given by

$$a_0 = L\int_{\omega_a}^{\overline{\omega}} F(c_a(\omega))dW(\omega).$$

Which directly affects native-born through smaller positive externalities. Since the Northern planner could achieve the same result with a smaller surplus loss by a more lenient selection policy, so that no assimilating immigration is given up, the best policy for the Northern planner is to set $\overline{\omega} = \omega_a$.

3.4 Empirical Results

This section will test the implications of the theoretical model. There are five subparts to this section: a description of the data, a statement of the model's theoretical predictions, the empirical model identification strategy, and estimation results.

Data

The data used in this study is from the Longitudinal Survey of Immigrants in Canada (LSIC). The survey is a three wave study conducted on new immigrants and refugees to Canada, atleast 15 years of age, in the period October 2000 and September 2001 (approximately 65,000 new immigrants). The survey excludes applications for immigration or asylum made within Canada. The cohort of immigrants captured in the survey were subject to the non-discriminatory character of the 1976 Immigration Act (Bodvarsson & Berg, 2013) and were admitted prior to the 2002 Immigration and Refugee Protection Act. The first wave is collected between April 2001 and March 2002, six months after arrival (12,040 immigrants of the 65,000 were recorded). A second wave of data is collected on this same group of individuals, six months later, between December 2002 and November 2003 (9,500 immigrants were re-recorded). A final wave is conducted, one year later, between November 2004 and October 2005 (7,715 immigrants were re-re-recorded) (Haan, 2012). The attrition rate is 21.1% and 18.8% in waves 2 and 3. Attrition is especially important in a study such as this because those immigrants that were lost from the sampling between waves 1-2 and waves 2-3 have important information about their level of assimilation or non-assimilation.

The focus of this study is new immigrants with little pre-immigration experience and low levels of exposure to Canadian culture. The subsample used in this study are immigrants whose mother tongue is non-English and do not reside in Quebec. This subsample is used throughout unless mentioned otherwise.

Immigrants may become missing from subsequent waves for various reasons: change in address and no follow-up contact information was provided, become deceased, or return to the home country. The reason for becoming missing from subsequent waves is not recorded. This becomes an issue because immigrants that returned to their home country due to difficulties in the assimilation process will bias the final results of the model. Only the remaining successful assimilates will be recorded and immigrant assimilation is overestimated. Characteristics of the returnees may be nonrandom and the bias will be exacerbated.

The advantages of using LSIC data for this research over the Canadian Census of Population are: (1) the data is longitudinal so that causal effects of learning on assimilating immigration are obtainable, (2) all landed immigrants were admitted under the same policy, and (3) there is a large amount of pre-immigration information.

Model predictions

An empirical test of the implications in the model requires a direct test of equilibrium proposition 3

$$\rho = \begin{cases} 1 & \text{if } k = 1 \text{ or } h \geq \overline{h} \\ \rho(m(L, h, \mu, \theta), h, \theta, \mu) & \text{if } k < 1 \text{ and } h < \overline{h} \end{cases}$$

Where $\rho = a/m$ is the likelihood of assimilating immigration. The model makes predictions on the effect of parameters L, h, μ and θ on the likelihood of assimilating. The predictions are summarized in table 2 below.

The threshold effect of changes in the parameters on $\rho(.)$ is the effect on assimilation levels through changes in equation (3.15). On the other hand, direct effects on $\rho(.)$ are established through the interior equilibrium conditions (3.11) - (3.12). And indirect effects on $\rho(.)$ are determined through transmissions from increases in the size of the immigrant group.[19]

Table 2: The effect of parameters on likelihood of assimilation

Parameter	$\rho(m(L,h,\mu,\theta),h,\theta,\mu)$		
	Threshold	Direct	Indirect
L	$-$	0	$-$
h	$+$	?	$+$
μ	$+$	?	$+$
θ	$+$	$-$	?

Empirical model

The estimation equation is a regression model to determine the effect of exogenous time-varying individual and contextual characteristics (X_{ist}), exogenous time-invariant source country and personal characteristics (X_{is}), investment into language accumulation (ι_{ist}), and source country richness (h_{is}). The X_{ist} factors are exogenous controls; they include demographic information on immigrants that vary overtime, for instance, marital status and months since arrival. ι_{ist} is formal and informal investment into language accumulation; it enters endogenously into the regression model. The X_{is} factors include the source country's linguistic distance, population, geographic distance, and pre-immigration experience with the local culture; pre-immigration and contextual factors are exogenous. I also introduce the share of co-ethnics in CMA/CA of arrival within X_{is} to control for sorting at the CMA/CA level. The assimilation variable is proxied by English speaking proficiency ρ_{ist}, where immigrant i is from source country s and measured in period t.

[19]Further information on deriving the direct, indirect and threshold effects are provided in the appendix.

The regression model treats the three waves of the survey as a panel and provides pooled probit estimates to identify the exposure channel through source country richness affects assimilation rates.

$$\rho_{ist} = \gamma_0 + \gamma_1 h_{is} + \gamma_2 \iota_{ist} + \gamma_3 X_{is} + \gamma_4 X_{ist} + \eta_i + \epsilon_{ist} \tag{3.26}$$

Where η_i is the random error component of the structural model. I expect η_i to capture factors that inhibit or ease the selection into formal language programs not currently captured in (4.30), such as distance to nearest ESL course, time cost, program costs, and ability to learn new languages. And η_i is correlated with ι_{ist} giving rise to the simultaneity. A first differences estimator consistently estimates γ_2 and coefficients on other time-varying covariates.

To test for unobserved effects I estimate a bivariate probit model to identify unobserved correlation in English speaking proficiency across waves of the survey,

$$\begin{aligned} P(\rho_{ist} = j, \rho_{ist+1} = j') = \Phi(&c(j) - \alpha_{1t} h_{is} - \alpha_{2t} \iota_{ist} - \alpha_{3t} X_{is} + \alpha_{4t} X_{ist}, \\ &c(j+1) - \alpha_{1t+1} h_{is} - \alpha_{2t+1} \iota_{ist+1} - \alpha_{3t+1} X_{is} + \alpha_{4t+1} X_{ist+1}, r) \end{aligned} \tag{3.27}$$

where $j, j' = \{1, 0\}$ indicate high proficiency (1) and low proficiency (0) in speaking English. The disturbances are distributed bivariate standard normal Φ with correlation r. The amount of unobserved correlation across waves is measured by $r = \text{cov}(\epsilon_{ist-1}, \epsilon_{ist})$. Previous studies have estimated an unfavorably large $|r|$. Given the set of covariates, the estimated $|r|$ in this paper is relatively small and decreasing over time. This phenomenon is indicative of decreasing influence of unobservables on English speaking proficiency.

A standard set of controls are used throughout. The controls that enter the model are based on previous studies that have shown the specific variable to be an important determinant of English speaking proficiency. The variables used in this study are summarized in table A1. Given that the language proficiency variable is subjectively determined there is potential response error arising from lack of a stable benchmark. That is, respondents in the LSIC may report decreasing language proficiency across waves. Whether the decreases in language proficiency are actually due to worsening language skills or due to a lack of a benchmark is difficult to determine. This is problematic because, in the estimation strategy described above, I may wrongly categorize a respondent to having worsening language proficiency when the problem was simply a response error. This error will underestimate assimilation levels. The extent of the measurement error is described in table 3 below

Table 3: Measurement error in English speaking proficiency across waves.

	Wave 1 - 2		Wave 2 - 3	
	Frequency	Percentage	Frequency	Percentage
Worse	14,798	15%	19,286	20%
Better	28,379	29%	21,812	22%
Same	53,870	56%	55,949	58%
Total	97,047	100%	97,047	100%

Note: The summary statistics produced in this table uses the raw categorical five-level English speaking proficiency variable to be as accurate as possible.

Atmost 15% of new immigrants that provided responses in waves 1 and 2, and 20% in waves 2 and 3, could be attributed to the measurement error. Since this error would underestimate assimilation levels, the results of this study may be interpreted as a lower bound. For the remainder of the paper, the five-level categorical English speaking variable will be collapsed into a dummy variable which would minimize the influence of the measurement error. Those who speak English "very well, native" or "well" are considered assimilating, others are non-assimilating. But if the error is non-random then the bias will persist. It is expected that the measurement error is more likely to occur among refugees, those with low English speaking proficiency, those who are less educated, and have less experience with the English language. Table 4 shows there is no systematic response in worsening English speaking proficiency between refugees and non-refugees, and between those with less than a high school education and those with atleast a high school education. Moreover, those with pre-immigration experience with the English language are more likely to respond with worsening English speaking skills than those without; either because they are understating their knowledge of the English language or they discover, after immigration, that the quality of their pre-immigration experience is low.

On the other hand, those with lower English speaking proficiency in wave t are more likely to report worsening English speaking proficiency in wave $t+1$ compared to those with higher proficiency in wave t. This evidence indicates there will be some bias associated with measurement error. The measurement error in the dependent variable will bias coefficient estimates of the time-invariant variables in the model.

Table 4: Worsening of English speaking proficiency across waves t and $t+1$ by refugee status, English speaking proficiency, education, and formal pre-immigration experience with the English language.

	Refugee		Low proficiency		Less than high school		Pre-immigration experience	
	Yes	No	Yes	No	Yes	No	Yes	No
	Waves 1 - 2							
Worse	14.76	15.28	29.39	9.62	13.86	15.49	17.74	7.92
Better	46.61	28.20	21.80	32.21	35.32	28.16	29.45	28.67
Same	38.63	56.53	48.81	58.17	50.82	56.34	52.81	63.41
	Waves 2 - 3							
Worse	22.96	19.69	39.09	12.57	20.75	19.72	21.42	15.41
Better	34.33	21.76	15.30	25.20	22.26	22.51	23.67	19.01
Same	42.71	58.55	45.61	62.23	56.99	57.77	54.91	65.58

Note 1: The summary statistics produced in this table uses the raw categorical five-level English speaking proficiency variable.
Note 2: Low proficiency is measured by speaking English fairly well, poorly and not at all in wave $t+1$.
Note 3: Less than high school is measured by highest level of formal education attained outside Canada in wave $t+1$.

Results

Equation (4.30) treats the data as a panel. The coefficient γ_2 is consistently estimated by a first differences estimator. Table 5 provides the set of regression results. The first column is a pooled probit estimator of the structural model (4.30) by excluding pre- and post-immigration investment. In this regression source country richness is allowed to affect assimilation rates through any type of learning and exposure to co-ethnics. As expected, the effect is estimated to be significant and positive.

The second column includes pre-immigration experience with the local culture. The estimated coefficient of γ_1 becomes insignificant. This implies that a large part of the observed effect of source country richness on assimilating immigration through the exposure channel is due to pre-immigration learning of the local culture. Pre-immigration formal and informal learning of English proxies for ω in the theoretical model. This variable is a major advantage to this study as many other studies do not sufficiently treat the learning that occurs prior to immigration. New immigrants with larger pre-immigration formal and informal learning of English have lower learning costs. Pre-immigration learning of English enters significantly and positive.

Table 5: Probit estimation results of English speaking proficiency: mother tongue is not English and do not reside in Quebec.

	Pooled	Pooled	Pooled
RGDP per capita (h)	0.156**	0.113	0.119
	(0.0555)	(0.0584)	(0.0614)
Informal pre-immigration experience (ω)		0.347***	0.360***
		(0.0490)	(0.0501)
Formal pre-immigration experience (ω)		0.610***	0.618***
		(0.0497)	(0.0510)
Informal investment (ι)			-0.272***
			(0.0434)
Formal investment (ι)			-0.241***
			(0.0482)
Share of co-ethnic group in Canada (m)	1.709	0.268	0.0104
	(1.038)	(1.029)	(1.037)
Population level (L)	-0.0257***	-0.0154*	-0.0168**
	(0.00600)	(0.00601)	(0.00612)
Geographic distance (μ)	0.00000163	-0.0000154	-0.0000242*
	(0.0000118)	(0.0000120)	(0.0000123)
Linguistic distance (θ)	-0.0526	-0.0249	-0.00734
	(0.219)	(0.220)	(0.222)
Months since arrival	-0.0399*	-0.0371	-0.0363
	(0.0196)	(0.0200)	(0.0205)
Time-varying controls (X_{ist})	✓	✓	✓
Time-invariant controls (X_{is})	✓	✓	✓
Survey wave controls (t)	✓	✓	✓
N	216,672	216,672	204,122

$^{*}p < 0.05^{**}p < 0.01^{***}p < 0.001$

Note 1: All models are weighted due to Statistics Canada's RDC disclosure process.

Note 2: Cluster robust standard errors in parentheses.

Column 3 of table 5 includes further learning of the local culture through formal and informal means into the model. Formal and informal investment, as expected, enter endogenously and the coefficient estimates are signed negative when it was expected to be positive. The inclusion of further learning has no discernible impact on the coefficient of RGDP per capita.

The least squares estimates for the models are provided in table A4 of the appendix. The first differences estimator in the fourth column. The estimated coefficients for formal and informal learning are significant. The coefficient estimates are unbiased and standard errors are clustered. The estimated coefficients are similar to those obtained by Orlov (2017). Orlov (2017) estimates are larger 0.2910^{***} using an IV-GMM approach[20].

Months since immigration becomes insignificant after including pre- and post-immigration learning into the model. As expected, learning through investment is more relevant than learning as a time dependent process. However, this result is inconclusive because the survey used in this study records a single entry cohort (October 2000 - September 2001 landed immigrants) so there are insignificant differences in immigrants' time of arrival to account for changes in language proficiency.

Country of birth variables such as population levels (L) relative to Canada, geographic distance (μ) of the capital city in country of birth to Ottawa and linguistic distance (θ) of the official language to English are included in the empirical model. The effect of source country population levels on English speaking proficiency is significant negative. Densely populated source countries are less likely to send assimilating immigrants.

Geographic distance is significant and negative which indicates that immigrants from more distant source countries are less likely to send assimilating immigrants. Linguistic distance has a negative effect on English speaking proficiency, but insignificant. This is an expected result

[20]George Orlov (2017) uses the LSIC to estimate the language transfer equation where the dependent variable is a continuous speaking score calculated by Principal Component Analysis on five questions of English speaking and comprehension competence. The independent variable of interest is time spent in an English as a Second Language (ESL) course, which is a continuous measure of formal language training. The instrument used is distance to the nearest ESL program. The coefficient estimate obtained by Orlov (2017) was 0.2910^{***}.

and coincides with the prediction of the theoretical model.

Potential migrants are selected by the immigration policy of the host country. A group share variable is calculated from the LSIC: wave 1. Share of co-ethnics admitted into Canada at the time of arrival proxies for m in the theoretical model and exogenously controls for the immigration policy. This variable is insignificant.

Refugees in the LSIC represent 15% of the sample. Although this is considerable, excluding this group from the estimation model makes no difference to the estimated coefficient on language investment. Also, the survey weighting accounts for the oversampling.

The joint probability model of equation (4.29) allows for the extraction of correlation coefficients across waves. The correlation across waves captures the effect of unobservables in the model. If the effect of the unobservables is purely random overtime, then there will be zero correlation between the error terms. A non-zero correlation between the error terms would be expected if the effect of unobservables are relevant. For instance, if there were heterogeneity in innate skills for language acquisition then there is strong correlation. If the correlation is lower for non-adjacent (more distant) waves than for adjacent waves, there is a "regression to the mean" phenomenon. That is, there is a decreasing effect of unobservables on English speaking skills. Compared to Chiswick, Lee & Miller (2004) the estimated correlation coefficient is much smaller but the regression to the mean phenomenon is present. The correlation coefficients are presented in table 6 below.

Table 6: Estimates of correlation coefficients from bivariate probit models of English speaking proficiency: mother tongue is not English and do not reside in Quebec.

	Wave 1 - 2	Wave 2 - 3	Wave 1 - 3
Correlation coefficient	0.560^{***}	0.599^{***}	0.537^{***}
Standard error	(0.007)	(0.007)	(0.008)
N	58,121	47,104	47,346

Note 1: Correlation coefficient estimates are estimated from weighted models due to Statistics Canada RDC's disclosure process on summary statistics.

Note 2: Cluster robust standard errors in parentheses.

3.5 Extension: spatial sorting

Immigrants have strong location preferences in the host country. New immigrants tend to locate where there are existing concentrations of immigrant groups, called ethnic enclaves. The location of immigrants determines their observed patterns of cultural and economic assimilation (Massey, 1985). Enclaves serve as a means to minimize costs in communication with co-ethnics and transportation from consumption of ethnic goods. Enclaves provide benefits to new immigrants through social networks creating greater levels of economic mobility (Edin, Fredriksson & Aslund, 2003), higher returns to foreign experience and education, and are places of entrepreneurialism where individuals start ethnic businesses (Portes & Shafer, 2006). Living in an ethnic enclave is essential to new immigrants beginning the integration process into host country culture. However, new immigrants living among co-ethnics also have lower levels of English language skill acquisition (Allen & Turner, 1992) and earnings (Allen & Turner, 1992; Xie & Gough, 2011). In this paper I study communication frictions between new immigrants and native-born and the means of cultural capital accumulation among new immigrants that lead to decreasing frictions overtime. The accumulation of language skills introduced in this model contributes to the literature of Bailey & Waldinger (1991) and Waldinger (1993), where enclaves are viewed as a "training system" for new immigrants before entering the mainstream economy. The model may be extended by including a spatial dimension in which new immigrants facing communication frictions will sort themselves into distinct ethnic enclaves to minimize costs asscociated with these frictions.

The model developed in this paper supposes potential migrants in the South are *randomly* matched with other immigrants and native-born in the North. Given the evidence that immigrants choose to be around other immigrants, non-assimilating immigrants avoid frictions associated with being matched with native-born by sorting into communities of co-ethnics.

All non-assimilating immigrants will choose to live in neighbourhoods with large numbers of existing immigrants. On the other hand, assimilating immigrants are indifferent between living among native-born and other immigrants. Given that the only source of friction in this model is finding a partner to communicate with and all surplus created is through communicating effectively, spatial segregation leads to increasing surplus within immigrant neighbourhoods. That being said, English speaking profciency within immigrant neighbourhoods would decrease and further segregate the enclave from the native-born. Moreover, as spatial segregation intensifies and immigration increases, neighbourhoods converge in productivity because communication becomes frictionless within neighbourhood.

There are two necessary conditions for such a model to explain actual new immigrant integration. Firstly, the potential migrant must have access to information on the share of immigrants living in neighbourhoods around the host country before forming the location decision. The presence of strong migration networks allows for almost perfect information exchange between sending and receiving country. Secondly, native-born flight to non-immigrant neighbourhoods must also be incorporated into the model. Given these elements frictionless communication is achieved by location decisions. In fact, the economic outcomes would be similar to investment into cultural capital of the current model.

Since sorting and learning are competing explanations of the same phenomenon it is necessary to disentangle the two effects. To test whether the estimated results are driven by learning or sorting we follow the method proposed by Lazear (1999) and Ortega & Verdugo (2015). An interaction term between formal/informal investment and ethnic share in CMA/CA of arrival (R_{is}) is introduced into the regression equation of (4.30). The equation to estimate is

$$\rho_{ist} = \gamma_0 + \gamma_1 h_{is} + \gamma_2 \iota_{ist} + \gamma_3 X_{is} + \gamma_4 X_{ist} + \gamma_5 \left(\iota_{ist} \times R_{is}\right) + \eta_i + \epsilon_{ist} \qquad (3.28)$$

If the learning effect dominates, then γ_5 will be significant and large. However, the results of the robustness test depend crucially on the unobservables, such as, ability or willingness to learn the local language, being insignificant and does not confound the sorting effect. High ability individuals might choose to locate in areas with fewer co-ethnics because they are confident in their learning ability. The presence of unobservables would confound the effect of sorting. Unobservables do have a large impact on language acquisition based on the correlation coefficient estimate (table 6) of the joint probability model. Although, the influence of unobservables appears to be decreasing overtime.

The estimated γ_5 is significant for all of the interaction terms. The results of these estimates are provided in table A5 of the appendix. I can conclude that sorting is influencing the estimated coefficients in the model, and learning is not an isolated decision in driving cultural assimilation for this cohort of new immigrants in Canada.

3.6 Conclusion

Immigrants face frictions in communication in their life within the host country due to cultural gaps. Cultural assimilation plays a primary role in determining the success of immigrants by overcoming those cultural barriers. Immigrants assimilate by a process of learning. Learning is achieved through engagement with the local culture over long periods of time living in the host country. However, this paper shows that learning through investments into language accumulation is an important part of attaining frictionless communication with native-born. Immigrants learn formally by attending language courses or informally through self-study, media and interactions with others.

This paper has shown that formal and informal language accumulation is a strong contributor to cultural accumulation, especially during the early stages of living in the host country. While learning over time can be an important source of cultural accumulation, it may take years before any real growth is experienced. Investment allows new immigrants to assimilate within a short period of time.

New immigrants from poorer source countries have greater incentives to migrate, increasing the share of co-ethnics in the host country and subsequently increase exposure to co-ethnics. This exposure creates a positive externality to new immigrants because they face smaller cultural frictions and do not need to invest in learning the local culture. However, new immigrants come with pre-immigration experience with the local culture which reduces their learning costs. Those with very large pre-immigration experience enter the assimilating group, others incur an investment cost.

The empirical model estimated in this paper identifies the exposure channel through which source country richness affects assimilation rates by controlling for alternate channels. The effect of source country richness on the assimilation rate becomes insignificant when I control for pre-immigration experience with the local culture. This is because pre-immigration experience with the local culture decreases learning costs.

Using first differences, I provide causal effects of further learning on assimilating immigration. Formally learning the local culture has a significant positive effect on assimilation rates, but informal learning has a negative effect.

Finally, it is shown that unobserved characteristics of new immigrants, including exposure to co-ethnics, adaptability to a new culture or ability in learning new languages is a major determinant of cultural assimilation. But the influence of unobservables are decreasing overtime.

4 Cultural Assimilation: Learning and Sorting

4.1 Introduction

Immigrants exhibit strong location preferences in the host country. New immigrants tend to locate where there are existing concentrations of immigrant groups, called ethnic enclaves. The location of immigrants determines their patterns of cultural and economic assimilation (Massey, 1985). Enclaves serve as a means to minimize costs in communication with co-ethnics and transportation from consumption of ethnic goods. Enclaves provide benefits to new immigrants through social networks creating greater levels of economic mobility (Edin, Fredriksson & Aslund, 2003), higher returns to foreign experience and education, and are places of entrepreneurialism where individuals start ethnic businesses (Portes & Shafer, 2006). Living in an ethnic enclave is essential for new immigrants beginning the integration process into host country culture. However, new immigrants living among co-ethnics also have lower levels of English language skill acquisition (Allen & Turner, 1992) and earnings (Allen & Turner, 1992; Warman, 2006; Xie & Gough, 2011). Further empirical evidence is provided by Bauer, Epstein & Gang (2005) to substantiate the claim that migrants choose smaller networks as English language proficiency improves, and the quality of those networks matter greatly for the economic and social outcomes of new immigrants. In this paper I study communication frictions between new immigrants and native-born and how sorting and/or learning among new immigrants alleviates those frictions. The size of the co-ethnic group combined with the extent of clustering into co-ethnic communities are determinants of the level of exposure to co-ethnics that new immigrants will experience in the host country. The assimilating decision is compromised by choosing to live in a neighbourhood that is predominately co-ethnic but reduces communication costs for the new immigrant. The theoretical model of Konya (2007) is extended by including a spatial dimension.

This paper views ethnic enclaves as a "training system" for new immigrants before entering the mainstream economy. The training system is a composite of labour market information, recruitment practices and skill accumulation. The enclave is an institution of formal and informal traditions differing from the secondary and primary sectors of the economy. New immigrants will choose to live in enclaves as a means to minimize communication costs while learning the local culture (Bailey & Waldinger, 1991; and Waldinger, 1993).

The literature studying the relationship between language proficiency and enclaves, given data limitations, cannot simultaneously include the role of learning. The likelihood of further learning is complementary or substitutable to sorting into co-ethnic communities. Learning and sorting are naturally opposing forces when viewed in a static environment. Davila & Mora (2000) generalize this view where the association between language accumulation and location choices are either temporary or permanent. It is temporary if immigrant labour is highly mobile and will move to where returns to current English speaking skills are highest, thus dissipating the enclave's association with language skills. On the other hand, it is permanent if new immigrants are caught in a mobility trap leading to lower levels of English speaking proficiency, remaining in the enclave, and thus strengthening the implied association. Their empirical study of Mexican-US border workers, a predominately minority language region, indicates that the association is temporary, and regional wage gaps encourage migration of Mexican immigrants away from the enclave. I do not address the issue of permanency because of the short panel available and also because this is not borne out by the theoretical model. Instead I am interested in whether exposure in an enclave to co-ethnics has any effect on language proficiency in the short-run.

Preliminary evidence suggests that exposure to co-ethnics is associated with lower assimilation rates. Using the 2001 Canadian Census of Population and focusing on the subsample of immigrants with no English speaking background, that is, immigrants from non-English speaking households whose mother tongue is non-English, 18-64 years old, and do not reside in Quebec. New immigrants that have lived in Canada for under a year, some are living within enclaves consisting of predominately co-ethnics. The preliminary evidence suggests that 81.68% of those living in an ethnic enclave[21] spoke English and 83.37% of those living outside the enclave did speak English. A difference of -1.69%** (0.0064). Among immigrants that have lived in Canada for 2 years the difference increases to -2.71%*** (0.0031). Among immigrants that have lived in Canada for 5 years the difference decreases drastically to -0.84%** (0.0029), then -0.50%* (0.0028) after living in Canada for 10 years. In the short-run there are growing differences in the composition of immigration inside ethnic enclaves compared to outside.

The model developed in this paper supposes potential migrants from the sending country are *randomly* matched with other immigrants and native-born in the host country. Due to productivity differences between the source and host country, immigration from poorer countries is expected to be larger than from richer ones. Even conditional on individual location preferences, such as relative size of family/friend networks, immigrants from poorer countries are more likely to encounter co-ethnics than comparable immigrants from richer source countries. This implies greater exposure to co-ethnics among immigrants from poorer source countries. Given that the only source of friction in this model is finding a partner to communicate with and all surplus created is through communicating effectively, the distribution of immigrants among neighbourhoods in the source country matters greatly to new immigrants beginning the integration process into host country culture. Non-random location choices among immigrant groups imply greater levels of clustering, increased exposure to co-ethnics and decreasing assimilation rates.

Using data from the Longitudinal Survey of Immigrants in Canada: Waves 1-3 where cultural assimilation is measured by proficiency in English, investment into native-born culture is measured by formal/informal learning of English, and location in an ethnic enclave is determined jointly by use of an unofficial language at work and clustering of co-ethnics in CMA/CA

[21] Construction of the ethnic enclave variable is discussed in the appendix.

of arrival. I provide estimates of sorting on cultural assimilation, and test the implications of the theoretical model. I find that sorting is an important component of the channel through which source country richness affects assimilation rates, but conditional on learning and sorting, source country richness still has a significant positive effect on assimilating immigration. There is an alternate channel unaccounted for by the current model through which sountry richness affects assimilation rates. I show that this alternate channel functions prior to sorting and it is probably related to quality of the immigrant group, rather than size of immigration.

In the next section I formulate a simplified model that incorporates location choice among neighbourhoods in the host country and show that multiple equilibria are possible. In section 3 of the paper I empirically test these findings.

4.2 Model

The mechanism functioning to transmit personal and contextual characteristics into language proficiency are generalized as exposure, incentives and efficiency (van Tubergen & Kalmijn, 2009). Exposure and incentives function interactively to determine the level of cultural integration. However, the mechanism relating to exposure is determined by location preferences of new immigrants upon arrival. Selection into neighbourhood must be introduced into the model to better identify the true relationship between exposure and language accumulation. Exposure is measured by the size of the immigrant group. An increase in the size of the immigrant group decreases incentives to learn the local culture. Additionally, exposure is determined by the size of immigration in the host country. This would have a feedback effect through changes in the immigration levels as more immigrants choose to migrate; further increasing exposure to other immigrants and decreasing incentives for new immigrants to invest in learning the local culture. These exposure effects are stronger within ethnic enclaves, leading to drastic effects on the composition of immigration within the enclave and outside of it. I will present a theoretical model accounting for these features and derive the equilibria that arise from this environment.

Communication frictions between native-born and immigrants are modelled into the host country environment through a random matching framework. I assume there are only two countries in the world: North and South where the South is less developed relative to the North. Thus migration flows from South to North. The North is further segmented into neighbourhoods, but for simplicity I assume that there are only two neighbourhoods. There are two sets of agents: native-born and immigrants. The native-born agents are located in the North and potential migrants are located in the South. The potential migrants in the South make a decision about whether to stay in the South or migrate to the North. Additionally, potential migrants must also decide whether to assimilate or not, and simultaneously choose which neighbourhood in the North to locate.

Whence the migrate problem has been solved the new immigrant must choose a neighbourhood to live/work in the North. This decision depends on the relative size of their networks among the neighbourhoods of the North. Potential migrants choose a strategy that gives the best outcome based on a set of individual-specific and contextual characteristics.

Individuals in the North are drawn together randomly so that matches are created. Efficient matching is the primary mode of production. The model introduces communication frictions that inhibit ease of production by randomly matching agents. A match between persons of similar culture are able to generate a surplus but matches between persons of different culture create no surplus. Matches between non-assimilating and native-born create no surplus while matches between non-assimilating and assimilating immigrants, or assimilating immigrants and native-born, do generate a surplus. The randomness in meeting people of common cultures are effectively reduced in the presence of location choice. Potential migrants that fail to assimilate choose less communication frictions by living in neighbourhoods with a large number of immigrants. As new immigrants enter the host country their location decisions are determined by the immigrant distribution and in turn influence their learning decision. The location decision of other new immigrants is influenced by the location and learning decisions of prior immigrant cohorts. Equilibrium is achieved when no individual has an incentive to relocate from their current neighbourhood and their learning decisions are optimized.

Thresholds

In this section I will discuss further the potential migrant's decision problem, how heterogeneity is introduced into the model, derive the equilibrium thresholds for migrating and assimilating, and discuss the equilibria that come out of the random matching immigration problem. The migration cost is

$$\text{Migration cost} = \mu c$$

where c is an individual-specific migration cost that is distributed over all potential migrants in the South, $c \sim F(c)$ and $c \in [0, 1]$. And $\mu < 1$ is an index of physical distance between the North and the South. Migration costs do not differ across neigbourhoods in the North. Potential migrants may face assimilation costs of

$$\text{Assimilation cost} = \theta \nu \iota$$

where $\theta < 1$ is an index of cultural differences between the North and the South and ι is investment into cultural assimilation by unit cost of investing ν. Assimilation costs are determined by the level of investment that is undertaken. Investment is a part of the process of cultural capital accumulation that is either amassed over time in a passive learning process where immigrants learn about the host country's culture through time spent in the host country. Or, formal and informal investments may be made towards learning the culture as a decisive step towards integration.

Learning is introduced through the capital accumulation process. Learning is distinguished at their pre- and post-immigration levels where pre-immigration experience with the local culture is an individual-specific distribution in the source country, while post-immigration learning comes as a form of investment into learning the local culture after the migration decision has been made. Further learning in the host country are individual-specific decisions that are a function of the investment returns and costs.

Cultural capital has two components, $x = \phi_1 \iota + \phi_2 \omega$ where $x \in [0, 1]$ is the total number of native-born that the migrant can communicate with given the level of investment ι and pre-immigration experience ω. Cultural capital is dichotomous; immigrants can either talk to all native-born ($x = 1$) or some ($x = \omega$). Pre-immigration experience ω is accumulated capital or endowment. Accumulated capital ω is distributed $W(\omega)$, $\omega \in [0, 1]$. It represents accumulated learning that includes the potential migrant's experience with host country culture prior to migrating. Investment is a dichotomous variable representing further learning, $\iota \in \{0, 1\}$. New immigrants with large pre-immigration experience (ω is large) face the lowest cost of investment ($\iota = 1 - \omega$).

Assimilating immigrants can generate surplus with other immigrants or native-born. Non-assimilating immigrants can generate surplus with other immigrants but the chance of meeting another immigrant depends on the neighbourhood they will choose to live and the size of their network. Suppose there are two neighbourhoods $i = A, B$, the potential migrant's neighbourhood choice is given by a general preference parameter, $\beta \sim B(\beta)$ and $\beta \in [0, 1]$, which is an individual-specific variable measuring the size of the immigrant's network in neighbourhood i relative to neighbourhood j.[22] New immigrants with a large network in $i = A, B$ prefer this neighbourhood over any other $j \neq i$. The potential migrant's decisions are summarized by the

[22]This formulation is similar to cultural segregation in a variant of the model presented in Bonn (2012a).

following set of value functions

$$V_a^i = \beta\left(1 - \theta\nu(1-\omega) - \mu c\right) \tag{4.1}$$

$$V_a^j = (1-\beta)\left(1 - \theta\nu(1-\omega) - \mu c\right) \tag{4.2}$$

$$V_n^i = \beta\left(\frac{m^i}{1+m^i} - \mu c\right) \tag{4.3}$$

$$V_n^j = (1-\beta)\left(\frac{m^j}{1+m^j} - \mu c\right) \tag{4.4}$$

$$V_s = h \tag{4.5}$$

where $i, j = A, B$ and $j \neq i$, the total number of native-born in neighbourhood i is normalized to one, and m^i is the total number of immigrants in neighbourhood $i = A, B$. Equations (4.1) and (4.2) are the utility from assimilating immigration within neighbourhood $i, j = A, B$ and $j \neq i$. The assimilating immigrant can trade with anyone within their own neighbourhood and earn a surplus of one, but must incur the cost of migration and assimilation. Equations (4.3) and (4.4) are the utility from non-assimilating immigration within neighbourhood $i, j = A, B$ and $j \neq i$. The non-assimilating immigrant only trades with other immigrants within the neighbourhood but face no assimilation cost. New immigrants with the largest ω have the lowest assimilating cost and the largest incentive to learn the local culture. Assimilating immigrants are not limited to trading with other immigrants, they can also trade with native-born. However, even assimilating immigrants are limited to trading within the neighbourhood. Equation (4.5) is the utility from not migrating. All meetings with Southerners will produce a surplus of h with no migration and assimilation costs incurred. Among those that stay in the South will generate a surplus of $h < 1$ (this condition ensures that migration only moves from South to North; that is, matches are more efficient in the North than in the South).

Equating (4.1) and (4.3), and solving for ω derives the threshold level for assimilating in neighbourhood $i = A, B$, such that $\omega \geq \omega_a$ there is assimilating immigration, and non-assimilating immigration otherwise. There are two seperate threshold conditions for each

$i = A, B$.

$$\omega_a^i = 1 - \frac{1}{\theta\nu(1+m^i)}, \tag{4.6}$$

$$\omega_a^j = 1 - \frac{1}{\theta\nu(1+m^j)}, \tag{4.7}$$

where $i, j = A, B$ and $j \neq i$. Among those in the assimilating group, the neighbourhood choice is given by the threshold β_a. The location choice for the assimilating group is derived by equating (4.1) and 4.2), and solving for β, as such

$$\beta_a = 1/2 \tag{4.8}$$

Those assimilating immigrants with $\beta \geq \beta_a$ there is immigration to neighbourhood i, otherwise neighbourhood j. Similarly among non-assimilating immigrants, the neighbourhood choice is derived by equating (4.3) and (4.4), and solving for β (or for c). This is the threshold $\beta_n(c)$ (or $c(\beta_n)$)

$$\beta_n(c) = \frac{m^j/(1+m^j) - 2\mu c}{m^i/(1+m^i) + m^j/(1+m^j) - 2\mu c}, \tag{4.9}$$

$$\text{or } c(\beta_n) = \frac{m^j/(1+m^j) - (m^i/(1+m^i))(\beta_n/(1-\beta_n))}{2\mu} \tag{4.10}$$

Non-assimilating immigrants with $\beta > \beta_n(c)$ will choose to live in neighbourhood i, others will choose neighbourhood j. In addition to threshold conditions (4.6) - (4.9) which describe the individual's assimilation and location incentives, the following four conditions determine the individual's migration choices which must also be satisfied. Equating (4.5) to either (4.1) or (4.2) provides the threshold for assimilating immigration to neighbourhood i or j, respectively. Similarly, equating (4.5) to either (4.3) or (4.4) is the threshold for non-assimilating immigration to neighbourhood i or j, respectively. The set of immigration thresholds are

$$c_a^i(\omega,\beta) = \frac{1-\theta\nu(1-\omega)-h/\beta}{\mu} \tag{4.11}$$

$$c_a^j(\omega,\beta) = \frac{1-\theta\nu(1-\omega)-h/(1-\beta)}{\mu} \tag{4.12}$$

$$c_n^i(\beta) = \frac{m^i/(1+m^i)-h/\beta}{\mu} \tag{4.13}$$

$$c_n^j(\beta) = \frac{m^j/(1+m^j)-h/(1-\beta)}{\mu} \tag{4.14}$$

Only those Southerners with small migration costs will migrate. Non-assimilating immigration to neighbourhood $i = A, B$ must satisfy $c \leq c_n^i(\beta)$, $\omega < \omega_a^i$ and $\beta > \beta_n(c)$. Non-assimilating immigration to neighbourhood j must satisfy $c \leq c_n^j(\beta)$, $\omega < \omega_a^j$ and $\beta < \beta_n(c)$. These conditions for the non-assimilating group may be plotted in (c, β) space (see figure 4.1). Similarly,

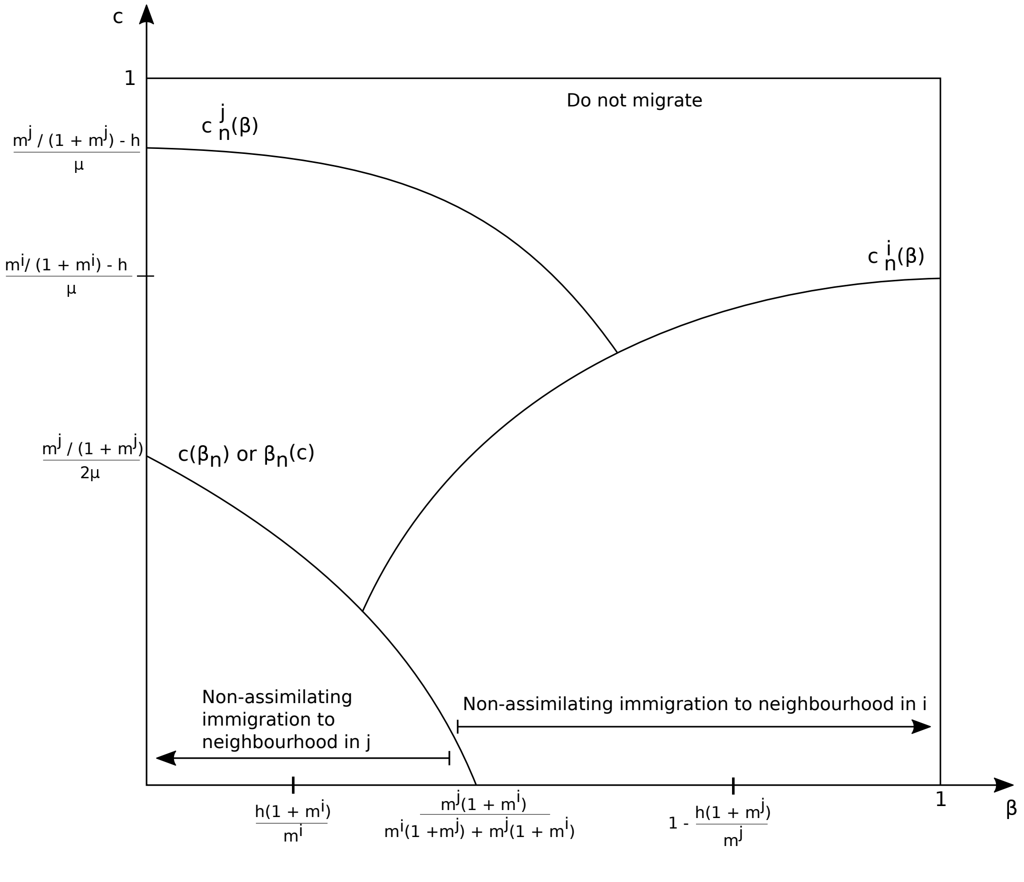

Figure 4.1: Non-assimilating immigration; $m^j \geq m^i$

assimilating immigration to neighbourhood $i = A, B$ must satisfy $c \leq c_a^i(\omega, \beta)$, $\omega > \omega_a^i$, and $\beta > \beta_a = 1/2$. Assimilating immigration to neighbourhood j must satisfy $c \leq c_a^j(\omega, \beta)$, $\omega > \omega_a^j$, and $\beta < \beta_a = 1/2$. Figure 4.2 shows these conditions for the assimilating group in (c, β) space assuming $\omega > \omega_n^i$.

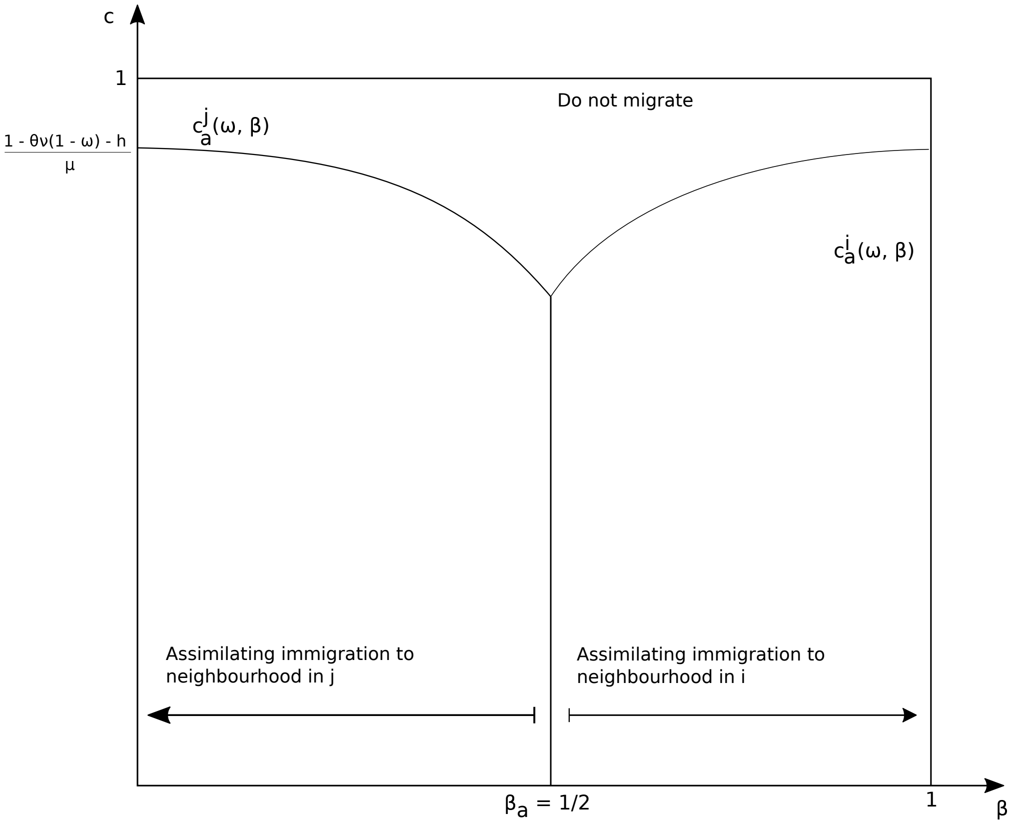

Figure 4.2: Assimilating immigration; $m^j \geq m^i$ and $\omega > \omega_n^i$

It is more useful to view the decision space for both groups in a single diagram. However, this requires one of the two variables, ω or β, to be held constant while the other is graphed as a function of c. I have chosen to hold β constant and depict the various regions in (c, ω) space. Note that the assimilating immigration thresholds $c_a^i(\omega, \beta)$ and $c_a^j(\omega, \beta)$ are a function of β and ω. For given values of β, the two functions are related as follows

$$c_a^i(\omega, \beta) \begin{cases} > c_a^j(\omega, \beta) & \text{if } \beta > 1/2 \\ = c_a^j(\omega, \beta) & \text{if } \beta = 1/2 \\ < c_a^j(\omega, \beta) & \text{if } \beta < 1/2 \end{cases}$$

Figure 4.3 depicts the non-assimilating and assimilating immigration groups for different levels of ω and c in the case when $\beta > 1/2$. A similar graph exists for $\beta < 1/2$ which would depict assimilating immigration to neighbourhood j. Non-assimilating immigration to neighbourhood

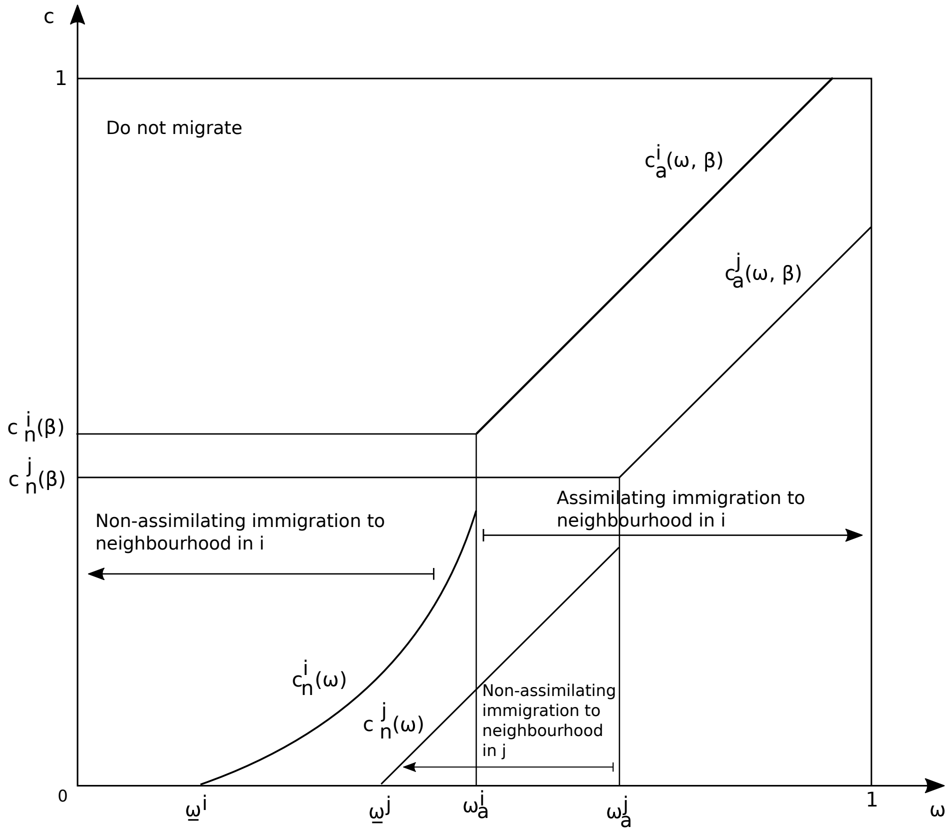

Figure 4.3: Mixed equilibrium; $\beta > 1/2$, $m^j \geq m^i$ and $h < m^i m^j/(m^i + m^j + 2m^i m^j)$

j must satisfy the condition that $c \leq c_n^j(\beta)$, $\omega < \omega_a^j$ and $\beta < \beta_n(c)$. The first two conditions are identifiable in figure 4.3, but the last condition is not a function of ω. To be able to plot this condition in (c, ω) space insert $\beta_n(c)$ into $c_a^j(\omega, \beta)$ and solve for c. The result is

$$c_a^j(\omega, \beta_n(c)) \equiv c_n^j(\omega) = \frac{m^i/(1+m^i)(1-\theta\nu(1-\omega)) - h(m^i/(1+m^i) + m^j/(1+m^j))}{\mu(m^i/(1+m^i) - 2h)} \quad (4.15)$$

Similarly, non-assimilating immigration to neighbourhood i must satisfy the condition that $c \leq c_n^i(\beta)$, $\omega < \omega_a^i$ and $\beta > \beta_n(c)$. Substitute $\beta_n(c)$ into $c_a^i(\omega, \beta)$ and implicitly define $c_n^i(\omega)$ as

$$c_n^i(\omega) + \frac{h}{\mu}\left(\frac{m^i/(1+m^i)}{m^j/(1+m^j) - 2\mu c_n^i(\omega)}\right) = \frac{1-\theta\nu(1-\omega)-h}{\mu} \quad (4.16)$$

Equations (4.15) and (4.16) are depicted in figure 4.3. This information is sufficient to identify the non-assimilating immigrant groups in both neighbourhoods.

Equilibria

Two equilibrium states are possible: a sorting and mixed equilibrium. One of the two states will emerge given levels of h. Each equilibrium has certain qualitative features. For simplicity of exposition let the the level of ω at which $c_n^i(\omega) = 0$ and $c_n^j(\omega) = 0$ be defined respectively as

$$\begin{aligned}\underline{\omega}^i &\equiv 1 - \frac{1 - h(1 + m^i(1 + m^j)/m^j(1 + m^i))}{\theta\nu} \\ \underline{\omega}^j &\equiv 1 - \frac{1 - h(1 + m^j(1 + m^i)/m^i(1 + m^j))}{\theta\nu}\end{aligned}$$

In a mixed equilibrium, immigration is mixed within neighbourhoods. That is, there is positive levels of immigration in both neighbourhoods and there is assimilating and non-assimilating immigration present. The mixed equilibrium outcome in neighbourhood $i = A, B$ is given by

$$a^i = L \int_{\omega_a^i}^1 \int_{1/2}^1 F(c_a^i(\omega, \beta)) dB(\beta) dW(\omega) \tag{4.17}$$

$$m^i - a^i = LW(\omega_a^i) \int_{\frac{h(1+m^i)}{m^i}}^1 \max\{0, F(c_n^i(\beta)) - F(c(\beta_n))\} dB(\beta) \tag{4.18}$$

and $\omega_a^i > \underline{\omega}^i$.

The left-hand side of (4.17) - (4.18) is the actual number of assimilating immigrants (a^i) and non-assimilating immigrants ($m^i - a^i$), respectively. These are equated to their respective expected values. Similarly, the outcomes in a mixed equilibrium for neighbourhood $j = A, B$ where $j \neq i$ is given by

$$a^j = L \int_{\omega_a^j}^1 \int_0^{1/2} F(c_a^j(\omega, \beta)) dB(\beta) dW(\omega) + l^j \tag{4.19}$$

$$m^j - a^j = L \left(W(\omega_a^j) - W(\underline{\omega}^j) \right) \int_0^{\frac{m^j(1+m^i)}{m^i + m^j + 2m^i m^j}} F(c(\beta_n)) dB(\beta) \tag{4.20}$$

and $\omega_a^j > \underline{\omega}^j$.

The left-hand side of (4.19) - (4.20) is the actual number of assimilating immigrants (a^j) and non-assimilating immigrants ($m^j - a^j$), respectively. These are equated to their expected values. Total immigration is $m = m^i + m^j$ and total assimilating immigration is $a = a^i + a^j$.

Table 1: Mixed or sorting equilibrium outcomes given h.

k	ω_a^i	ω_a^j	h	Outcome
< 1	$\geq \underline{\omega}^i$	$= \underline{\omega}^j$	Rich	Sorting
≤ 1	$\geq \underline{\omega}^i$	$> \underline{\omega}^j$	Poor	Mixed

In a sorting equilibrium there is mixed immigration in one neighbourhood and only assimilating immigration in the other. The sorting equilibrium is given by equations (4.17) - (4.20) and

$$k(1-\omega) < 1,\ \omega_a^i \geq \underline{\omega}^i \text{ and } \omega_a^j = \underline{\omega}^j$$

for all $i, j = A, B$ and $j \neq i$. Immigration to neighbourhood i is mixed but only assimilating in neighbourhood j. Table 1 further summarizes the outcomes given h.

Although the equilibria are not unique, they are stable under very reasonable conditions. Stability in the two neighbourhood case must satisfy the following condition

$$\|J(m^i, m^j)\| < 1 \tag{4.21}$$

on $D = \{(m^i, m^j)\,|\, m^i, m^j > 0\}$ where $J(m^i, m^j)$ is the Jacobian matrix of first-order derivatives of the equation system describing the mixed equilibrium outcomes in neighbourhoods i and j.

$$m^i = L\int_{\omega_a^i}^{1}\int_{1/2}^{1} F(c_a^i(\omega,\beta))dB(\beta)dW(\omega) \tag{4.22}$$
$$+ W(\omega_a^i)^2 \int_{\frac{h(1+m^i)}{m^i}}^{1} \max\{0, F(c_n^i(\beta)) - F(c(\beta_n))\}dB(\beta) \equiv \Gamma^i(m^i, m^j)$$
$$m^j = L\int_{\omega_a^j}^{1}\int_{0}^{1/2} F(c_a^j(\omega,\beta))dB(\beta)dW(\omega) \tag{4.23}$$
$$+ L\left(W(\omega_a^j) - W(\underline{\omega}^j)\right)^2 \int_{0}^{\frac{m^j(1+m^i)}{m^i+m^j+2m^i m^j}} F(c(\beta_n))dB(\beta) \equiv \Gamma^j(m^i, m^j)$$
$$\omega_a^i \geq \underline{\omega}^i \text{ and } \omega_a^j > \underline{\omega}^j.$$

The first-order derivatives are provided in the comparative statics section of the appendix. Stability of the mixed equilibrium system requires that the eigenvalues of $J(m^i, m^j)$ are signed as follows

$$\frac{\partial \Gamma^i}{\partial m^i} < 0 \text{ and } \frac{\partial \Gamma^j}{\partial m^j} < 0. \tag{4.24}$$

The stability conditions also imply that $\partial \Gamma^j / \partial m^i > 0$. The stability of a sorting equilibrium must also satisfy condition (4.24), although $\partial \Gamma^j / \partial m^j < 0$ and $\partial \Gamma^j / \partial m^i > 0$ hold without assumption. The first-order derivatives in the sorting case are also provided in the comparative statics section of the appendix.

Finally, the emergence of either equilibrium type is dependent on the level of h as proposed in column 4 of table 1. Proposition 4 below proves the existence of a level of h that switches the equilibrium outcome from a mixed to a sorting.

Proposition 4. *If $m^j \geq m^i$, where $i, j = A, B$ and $j \neq i$, and $k(1-\omega) < 1$. There exists an $\overline{h}$, such that, for $h < \overline{h}$ the equilibrium is sorting, and mixed for $h > \overline{h}$.*

Proof. Let $\overline{m}^j$ be the solution to $\omega_a^j > \underline{\omega}^j$ in a mixed equilibrium outcome. Similarly, the sorting outcome is given by substituting $\omega_a^j = \underline{\omega}^j$ into the right-hand side of equation (4.23). Equating the former $\overline{m}^j$ with the latter gives the following implicit function defining $\overline{h}$ by

$$\frac{\overline{h}}{1 - 2\overline{h} - \overline{h}/m^i(L, h, \mu, \theta)} = L \int_{\underline{\omega}^j}^{1} \int_{0}^{1/2} F(c_a^j(\omega, \beta)) dB(\beta) dW(\omega). \tag{4.25}$$

Note that both sides of the equation include $m^i(L, h, \mu, \theta)$, as defined by equation 4.22, which is also a function of h. Let $\phi(m^i(L, \overline{h}, \mu, \theta), m^j(L, \overline{h}, \mu, \theta), L, \overline{h}, \mu, \theta) = 0$ be the difference between the left-hand and right-hand side of equation (4.25). The comparative statics in the appendix show that $\partial m^i / \partial h < 0$, as such, the left-hand side of equation (4.25) is monotonically increasing in $h \in [0, 1)$, but is discontinuous at $h = 1$, at which point it becomes zero. The right-hand side is continuous and monotonically decreasing in $h \in [0, 1]$, and zero at $h = 1$. At $h = 0$ the right-hand side is larger than the left-hand side. And, at $h = 1$ the left-hand side and right-hand side are both zero and equal. Figure 4.4 shows clearly the existence of $\overline{h}$. The left-hand hand side of equation (4.25) is clearly shown by the upward sloping linear function with

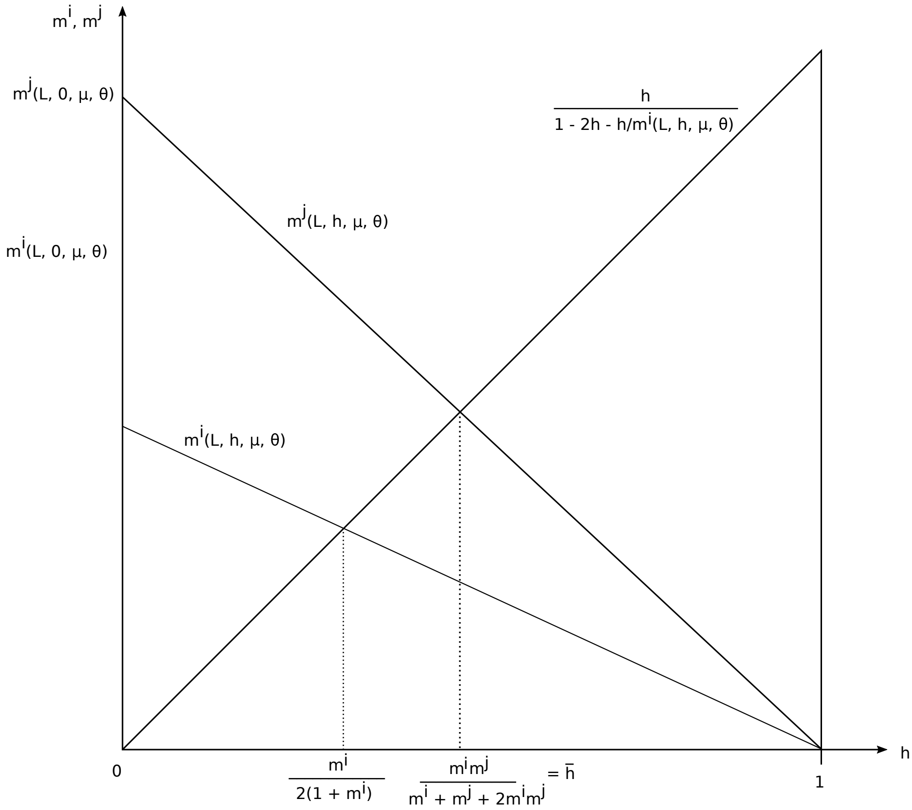

Figure 4.4: Existence of $\overline{h}$; $m^j \geq m^i$

a discontinuity at $h = 1$. Similarly, the right-hand side of equation (4.25) and $m^i(L, h, \mu, \theta)$ is also depicted as the downward sloping functions, where $m^j \geq m^i$. The three equations depicted in figure 4.4 are not linear in h, instead they are concave/convex at different levels of $h \in (0, 1)$; linearity is imposed for the sake of simplicity in exposition. Since the functions are monotonically increasing/decreasing in $h \in (0, 1)$, linearity suffices to depict the existence of $\overline{h}$. The threshold level of h at which point the equilibrium changes from the mixed to the sorting case is indicated by

$$\overline{h} = \frac{m^i m^j}{m^i + m^j + 2m^i m^j}. \tag{4.26}$$

Consistent with the graphical depiction is the result that a higher h is associated with a lower level of immigration in both neighbourhoods. Moreover, changes in the parameter set $\{L, h, \mu, \theta\}$ will have implications for the level of $\overline{h}$ and, in turn, on the equilibrium outcome. □

Under proposition 4 a richer source country does not necessarily imply that all immigration will be assimilating, as was the case in Konya (2007). Instead I find that one of the two neighbourhoods in the North continues to receive the non-assimilating type even if the source country is rich.

The intuition behind this result is best understood as a *productivity*, *network* and *communication effect*, where the communication effect is composed of a *selection effect*. Efficiency of matches in the South has an impact on the level of immigration. As the South gets poorer there is more migration from the South because matches in the North become relatively more productive. Incurring the migration costs and foregoing the matches that could have been made in the South (the opportunity cost), the potential migrant finds these costs are small relative to the gains from matches in the North. This represents a productivity effect increasing the total number of immigrants from poorer source countries. Non-assimilating immigrants experience a debilitating communication effect if their pre-immigration experience with the local culture is not large. Some of them will have large enough pre-immigration experience with the local culture to enter the assimilating group; a selection effect. The selection effect determines the composition of immigration within neighbourhoods. Only if the communication effect is larger than the productivity effect will all immigration become assimilating.

Given the productivity and communication effects, networks function to induce new immigrants to choose neighbourhoods where they have more family/friends. If the productivity effect is large enough then total immigration is increasing. But based on relative size of pre-existing networks new immigrants will have a general preference for some neighbourhoods over others.

4.3 Empirical Results

This section will test the implications of the theoretical model. There are four subparts to this section: a description of the data, a statement of the model's theoretical predictions, the empirical model identification strategy, estimation results, and further insights.

Data

The data used in this study is from the Longitudinal Survey of Immigrants in Canada (LSIC). The survey is a three wave study conducted on new immigrants and refugees to Canada, atleast 15 years of age, in the period October 2000 and September 2001 (approximately 65,000 new immigrants). The survey excludes applications for immigration or asylum made within Canada. The cohort of immigrants captured in the survey were subject to the non-discriminatory character of the 1976 Immigration Act (Bodvarsson & Berg, 2013) and were admitted prior to the 2002 Immigration and Refugee Protection Act. The first wave is collected between April 2001 and March 2002, six months after arrival (12,040 immigrants of the 65,000 were recorded). A second wave of data is collected on this same group of individuals, six months later, between December 2002 and November 2003 (9,500 immigrants were re-recorded). A final wave is conducted, one year later, between November 2004 and October 2005 (7,715 immigrants were re-re-recorded) (Haan, 2012). The attrition rate is 21.1% and 18.8% in waves 2 and 3. Attrition is especially important in a study such as this because those immigrants that were lost from the sampling between waves 1-2 and waves 2-3 have important information about their level of assimilation or non-assimilation.

The focus of this study is new immigrants with little pre-immigration experience and low levels of exposure to Canadian culture. The subsample used in this study are immigrants whose mother tongue is non-English and do not reside in Quebec. This subsample is used throughout unless mentioned otherwise.

Immigrants may become missing from subsequent waves for various reasons: change in address and no follow-up contact information was provided, become deceased, or return to the home country. The reason for becoming missing from subsequent waves is not recorded. This becomes an issue because immigrants that returned to their home country due to difficulties in the assimilation process will bias the final results of the model. Only the remaining successful assimilates will be recorded and immigrant assimilation is overestimated. Characteristics of the returnees may be nonrandom and the bias will be exacerbated.

The advantage of using the LSIC over cross-sections in Konya (2007) and Lazear (1999) is that all the respondents were admitted under the same immigration policy and empirical issues associated with differences in the quality of the immigration cohort do not enter.

Model predictions

An empirical test of the relationships implied by the model requires the following definition for assimilating immigration $\rho^j = a^j/m^j$ in neighbourhood $j = A, B$. The measure of assimilating immigration ρ measures within neighbourhood composition of immigration in the North. Since the j^{th} neighbourhood receives no non-assimilating immigration in the sorting equilibrium under the assumption of $m^j \geq m^i$, the predicted outcomes of the model are summarized simply by

$$\rho^j = \begin{cases} 1 & \text{if } h \geq \overline{h} \text{ or } k = 1 \\ \rho^j(m^j(h,\theta,\mu,L), m^i(h,\theta,\mu,L), h, \theta, \mu) & \text{if } h < \overline{h} \text{ and } k \leq 1 \end{cases} \tag{4.27}$$

where $\overline{h}$ is defined by $\phi(m^i(\overline{h},\theta,\mu,L), m^j(\overline{h},\theta,\mu,L), \overline{h}, \theta, \mu, L)$ from equation (4.25), $i, j = A, B$ and $i \neq j$. The parameters (L, h, μ, θ) impact the type of equilibrium through a threshold, direct or indirect effect. The threshold effect determines the type of equilbrium in neighbourhood j through changes in $\phi(.)$. Changes in the parameters have a direct impact on the composition of immigration within the mixed equilibrium through changes in $\rho^j(.)$. The indirect effects of the model are associated with changes in the composition through the size of immigration. Since there are two neighbourhoods to consider, the parameters effect the composition of immigration through $m^i(.)$ and $m^j(.)$. Table 2 summarizes the various effects.

Table 2: The effect of parameters on likelihood of assimilation ρ^j

Parameter	$\rho^j(m^j(h,\theta,\mu,L), m^i(h,\theta,\mu,L), h,\theta,\mu)$			
	Threshold	Direct	Indirect $m^j(h,\theta,\mu,L)$	Indirect $m^i(h,\theta,\mu,L)$
L	$-$	0	?	?
h	$+$	?	?	?
μ	$+$	$-$	?	?
θ	?	?	?	?

The full effect of the parameters on the composition of immigration are ambiguous, the empirics cannot be used as a test of the model. On the other hand, although most immigrant receiving countries will restrict immigration through a quota or points system, the total size of immigration is typically exogeneously fixed at $m = m^i + m^j$. However, the model predictions function through the relative size of immigration across neighbourhoods, which is determined by individual preferences, some source country characteristics, and not the immigration authorities. Let $\zeta \equiv m^i/(m^i + m^j)$ measure the distribution of immigrants across neighbourhoods in the North. The effects of the parameters on the distribution of immigration are also estimated and discussed in the further insights section.[23]

Empirical model

Exposure to co-ethnics is the effective channel through which source country richness impacts assimilating immigration in the model. Since sorting amplifies the effects of exposure I expect source country richness' effect on assimilating immigration to function through the sorting variable.

Identifying the exposure channel through which source country richness affects assimilating immigration requires a regression framework that controls for alternate channels, such as,

[23] See the appendix for further discussion on the direct, indirect and threshold effects.

selection from within the source country, exposure within the household and immigration policy in the host country. Similarly, confounding factors, such as, pre-immigration experience with the local culture, individual learning costs and other unobservable characteristics of the immigrant will probably affect the sorting and learning decisions simultaneously. Each of these are discussed and treated appropriately in the following sections.

Sorting increases exposure to co-ethnics, through which, source country richness affects assimilation rates. A proxy for the sorting variable is included in the model which overadjusts for the exposure channel. The proxy variable for sorting is expected to interact with source country richness. This addition to the model may be used to determine if sorting is indeed a relevant component of the path through which source country richness affects assimilation rates.

The estimation equation is a regression model to determine the effect of exogenous time-varying characteristics (X_{ist}), exogenous time-invariant variables (X_{is}), investment into language accumulation (ι_{ist}), living/working in an ethnic enclave in CMA/CA of arrival (ζ_{is}), and source country richness (h_{is}) on the immigrant's decision to assimilate overtime. The X_{ist} and X_{is} factors are treated as exogenous; they include demographic and economic information on immigrants as well as pre-immigration experience with the local culture and friend/family networks. The time-invariant variable X_{is} includes contextual variables, such as, source country's linguistic distance, population, geographic distance, and a measure of the share of co-ethnics in Canada at the time of arrival m_{is}. The variable ι_{ist} includes formal and informal investment into language accumulation. Living/working in an ethnic enclave ζ_{is} is measured by choosing to initially live in a CMA/CA where co-ethnics are clustered and working in an organization that is predominately co-ethnic; this variable is endogenous. The assimilation variable ρ_{ist} is proxied by language proficiency, where immigrant i is from source country s and measured in period t.

The three waves of the survey are treated as a panel. The benchmark regression model to estimate is simply the effect of source country richness (h_{is}) on assimilating immigration (ρ_{ist}).

$$\rho_{ist} = \gamma_0 + \gamma_1 h_{is} + \gamma_3 \iota_{ist} + \gamma_4 X_{is} + \gamma_5 X_{ist} + \eta_i + \epsilon_{ist}. \tag{4.28}$$

Equation (4.28) is the equation of interest. The estimated effect of h_{is} on ρ_{ist} through the hypothesized exposure channel is $\widehat{\gamma}_1$. Moreover, exposure to co-ethnics occurs within or across neighbourhoods, and allows for non-random location choices. The time-varying and -invariant variables in the model control for alternate channels and confounding relationships. If the observed relationship between source country richness and assimilating immigration is through the exposure channel then the distribution measure is overadjusting the exposure effect by including it in the model. The following regression model includes the measure for distribution of immigrants across neighbourhoods (ζ_{is}),

$$\rho_{ist} = \gamma_0 + \gamma_1 h_{is} + \gamma_2 \zeta_{is} + \gamma_3 \iota_{ist} + \gamma_4 X_{is} + \gamma_5 X_{ist} + \eta_i + \epsilon_{ist}. \tag{4.29}$$

The coefficient estimate of language investment is biased because of simultaneity; non-assimilating immigrants are more likely to invest into learning the local culture. There is correlation between ι_{ist} and the individual-level error component η_i, implying that the coefficient on language investment will not be signed precisely without a fixed effects estimator. That being said, causal effects of investment are not the purpose of this paper.

I expect η_i to capture factors that inhibit or ease the selection into formal language programs not currently captured in (4.28), such as distance to nearest ESL course, time cost, program costs, and ability to learn new languages. Finally, the time-invariant variable ζ_{is} is endogenous because new immigrants with low English speaking skills will choose smaller communication frictions by locating in ethnic enclaves where access to immigration services are concentrated and learning the local culture can happen in a costless environment.

Immigrants that live and work in an ethnic enclave are identified using data from the Longitudinal Survey of Immigrants in Canada: wave 1 and the 2001 Canadian Census of

Population. The LSIC was used to provide information on whether the immigrant worked in an organization that predominately spoke a language other than English or French, in wave 1. The Census was used to identify whether a given ethnic group is over-represented in a census tract (CT) relative to the CMA/CA in which it is located. If more than 50% of CTs are over-representative of the ethnic group then the CMA/CA is considered to be clustering co-ethnics. Immigrants in the LSIC are identified as living/working in an ethnic enclave if they work in an organization predominately speaking a langauge other than English or French in wave 1, and if they belong to an ethnic group that lived in a CMA/CA of first arrival that clustered co-ethnics as determined by the Census.

A standard set of controls are used throughout. The controls that enter the model are based on previous studies that have shown the specific variable to be an important determinant of English speaking proficiency. The variables used in this study are described in table A1 of the appendix. Given that the language proficiency variable is subjectively determined there is potential response error arising from lack of a stable benchmark. That is, respondents in the LSIC may report decreasing language proficiency across waves. Whether the decreases in language proficiency are actually due to worsening language skills or due to a lack of a benchmark is difficult to determine. This is problematic because, in the estimation strategy described above, I may wrongly categorize a respondent to having worsening language proficiency when the problem was simply a response error. This error will underestimate assimilation levels. The extent of the measurement error is described in table 3 below

Atmost 15% of new immigrants that provided responses in waves 1 and 2, and 20% in waves 2 and 3, could be attributed to the measurement error. Since this error would underestimate assimilation levels, the results of this study may be interpreted as a lower bound. For the remainder of the paper, the five-level categorical English speaking variable will be collapsed into a dummy variable which would minimize the influence of the measurement error. But if

Table 3: Measurement error in English speaking proficiency across waves.

	Wave 1 - 2		Wave 2 - 3	
	Frequency	Percentage	Frequency	Percentage
Worse	14,798	15%	19,286	20%
Better	28,379	29%	21,812	22%
Same	53,870	56%	55,949	58%
Total	97,047	100%	97,047	100%

Note: The summary statistics produced in this table uses the raw categorical five-level English speaking proficiency variable to be as accurate as possible.

the error is non-random then the bias will persist. It is expected that the measurement error is more likely to occur among refugees, those with low English speaking proficiency, those who are less educated, and have less experience with the English language. Table 4 shows there is no systematic response in worsening English speaking proficiency between refugees and non-refugees, and between those with less than a high school education and those with atleast a high school education. Moreover, those with pre-immigration experience with the English language are more likely to respond with worsening English speaking skills than those without; either because they are understating their knowledge of the English language or they discover, after immigration, that the quality of their pre-immigration experience is low.

On the other hand, those with lower English speaking proficiency in wave t are more likely to report worsening English speaking proficiency in wave $t+1$ compared to those with higher proficiency in wave t. This evidence indicates there will be some bias associated with measurement error. The measurement error in the dependent variable will bias coefficient estimates of the time-invariant variables in the model.

Results

Firstly, I estimate the benchmark model of equation (4.28). The probit estimates are presented below in the first column of table 5. The effect of source country richness on assimilating immigration is measured by the coefficient on RGDP per capita. Source country richness has

Table 4: Worsening of English speaking proficiency across waves t and $t+1$ by refugee status, English speaking proficiency, education, and formal pre-immigration experience with the English language.

	Refugee		Low proficiency		Less than high school		Pre-immigration experience	
	Yes	No	Yes	No	Yes	No	Yes	No
	Waves 1 - 2							
Worse	14.76	15.28	29.39	9.62	13.86	15.49	17.74	7.92
Better	46.61	28.20	21.80	32.21	35.32	28.16	29.45	28.67
Same	38.63	56.53	48.81	58.17	50.82	56.34	52.81	63.41
	Waves 2 - 3							
Worse	22.96	19.69	39.09	12.57	20.75	19.72	21.42	15.41
Better	34.33	21.76	15.30	25.20	22.26	22.51	23.67	19.01
Same	42.71	58.55	45.61	62.23	56.99	57.77	54.91	65.58

Note 1: The summary statistics produced in this table uses the raw categorical five-level English speaking proficiency variable.
Note 2: Low proficiency is measured by speaking English fairly well, poorly and not at all in wave $t+1$.
Note 3: Less than high school is measured by highest level of formal education attained outside Canada in wave $t+1$.

a positive effect on assimilating immigration. But this effect is insignificant because learning is included in the right-hand side.

The second column includes the network variable (β) into the model. This variable controls for individual preference in neighbourhood choice. Networks are proxied by whether the immigrant had family/friends in their current area (or nearby area) of residence and chose the area because family/friends live there. The effect of this variable on assimilating immigration is negative but insignificant. The source country richness coefficient also remains insignificant.

The third column includes the ethnic enclave variable, which has the effect of changing the coefficient estimate of RGDP per capita and also making it significant. Ethnic enclave, or non-random distribution of immigrants among neighbourhoods, is interacting with RGDP per capita. Either, source country richness affects assimilation through sorting and exposure to co-ethnics, or there is an alternate channel through which source country richness affects assimilation rates.

The ethnic enclave variable ζ is endogenous because immigrants with low English speaking skills are more likely to choose to live and work within the enclave. Additionally, the ethnic enclave variable is interacting with source country richness. I instrument the ethnic enclave variable with housing costs in CMA/CA of arrival from the 2001 Canadian Census of Population. Housing costs are significantly and positively correlated with living and working in an ethnic enclave because new immigrants will choose to live in the low cost ethnic enclave if the CMA/CA has high average cost of housing. Table 6 provides a sorting regression where the coefficient estimate of housing costs in CMA/CA of arrival is significant and positive. Also, housing costs in CMA/CA of arrival has no observed association with assimilation rates. The fourth column instruments the ethnic enclave variable with housing costs in CMA/CA of arrival. The estimated effect of living and working in an ethnic enclave becomes larger. The coefficient effect of source country richness is still significant, implying the presence of an alternate channel. The further insights section discusses a possible alternate pathway.

Investment ι is measured by further learning of English in Canada. Further learning is differentiated by formal and informal means. Formal and informal investment have a significant effect on assimilating immigration. The sign of formal and informal learning is negative because of simultaneity.

A dummy variable for pre-immigration formal and informal learning of English proxies for ω in the theoretical model. Pre-immigration learning of English enters significantly and positive. This is expected since individuals with higher levels of pre-immigration experience with the local culture are more likely to enter the assimilating group if their experience is large, or have a low investment cost to learning the local culture.

The restrictions implied by the immigration policy in place at the time of arrival are proxied by ethnic share in Canada at the start of the survey (m). The coefficient estimate of this variable on assimilating immigration is positive but insignificant.

Finally, population levels (L) relative to Canada, linguistic distance (θ) and geographic distance (μ) of the capital city in country of birth to Ottawa are estimated. The effect of source country population on English speaking proficiency is negative. Geographic distance and cultural differences have no significant effect on assimilating immigration.

In table A3 of the appendix I estimate the benchmark model of equation 4.28 by OLS. The last column of table A3 subsets the data to those source countries that scored "low" or "very low" on the English proficiency index (EF EPI, 2017)[24]. There is much less variation in RGDP per capita and the countries were predominately low income. As expected the coefficient estimate on RGDP per capita is insignificant but still positive.

Refugees in the LSIC represent 15% of the sample. Although this is considerable, excluding this group from the estimation model makes no difference to the estimated coefficients on language investment and living/working in an ethnic enclave. Also, the survey weighting accounts for the oversampling.

Further insights

Source country richness impacts assimilating immigration through exposure, of which, in the model developed in this paper, is determined by sorting into co-ethnic neighbourhoods. The distribution of co-ethnics, or sorting, measure is defined as $\zeta \equiv m^j/(m^i + m^j)$. The effect of source country richness on the distribution of co-ethnics is given by the estimating equation

$$\zeta_{is} = \alpha_0 + \alpha_1 h_{is} + \alpha_2 X_{is} + \alpha_3 X_{ist} + \alpha_4 I_{is} + \epsilon_{ist}, \tag{4.30}$$

where I_{is} is housing costs in CMA/CA of arrival. The regression model (4.30) is estimated by probit, the results of the estimates are provided in table 6 below. This result supports the hypothesis of the enclave as a "training system" for new immigrants (Bailey & Waldinger, 1991; and Waldinger, 1993). Those without pre-immigration experience with the local culture

[24] Many African and Eastern European countries were not surveyed for the index. These were also omitted in the sub-sample.

Table 5: Probit estimation results of English speaking proficiency: households whose mother tongue is not English and do not reside in Quebec.

Dependent variable: English speaking proficiency	Pooled	Pooled	Pooled	Pooled IV
RGDP per capita (h)	0.119	0.116	0.129*	0.150**
	(0.0614)	(0.0614)	(0.0630)	(0.0569)
Network (β)		-0.0827	-0.0716	0.0363
		(0.0572)	(0.0573)	(0.0574)
Ethnic enclave (ζ)			-0.368***	-2.588***
			(0.0602)	(0.384)
Informal investment (ι)	-0.272***	-0.271***	-0.273***	-0.144*
	(0.0434)	(0.0435)	(0.0431)	(0.0641)
Formal investment (ι)	-0.241***	-0.242***	-0.248***	-0.203***
	(0.0482)	(0.0482)	(0.0483)	(0.0542)
Formal pre-immigration experience (ω)	0.618***	0.617***	0.616***	0.367**
	(0.0510)	(0.0510)	(0.0512)	(0.119)
Informal pre-immigration experience (ω)	0.360***	0.358***	0.348***	0.174*
	(0.0501)	(0.0501)	(0.0504)	(0.0837)
Share of co-ethnic group in Canada (m)	0.0104	0.0522	0.553	3.200**
	(1.037)	(1.036)	(1.042)	(1.074)
Population level (L)	-0.0168**	-0.0169**	-0.0183**	-0.0197***
	(0.00612)	(0.00611)	(0.00620)	(0.00586)
Geographic distance (μ)	-0.0000242*	-0.0000237	-0.0000231	-0.00000563
	(0.0000123)	(0.0000123)	(0.0000123)	(0.0000116)
Linguistic distance (θ)	-0.00734	-0.0102	0.00866	0.0378
	(0.222)	(0.222)	(0.222)	(0.176)
Months since arrival	-0.0363	-0.0366	-0.0373	-0.0290
	(0.0205)	(0.0205)	(0.0206)	(0.0180)
Time-varying controls (X_{ist})	✓	✓	✓	✓
Time-invariant controls (X_{is})	✓	✓	✓	✓
Survey wave controls (t)	✓	✓	✓	✓
N	204,122	204,122	203,858	203,858

$^{*}p < 0.05^{**}p < 0.01^{***}p < 0.001$

Note 1: All models are weighted due to Statistics Canada's RDC disclosure process.

Note 2: Cluster robust standard errors in parentheses.

are more likely to live and work in the ethnic enclave, although the effects reported are not significant. The ethnic enclave provides new immigrants with an environment that minimizes communication frictions with native-born until the local culture is learned.

Table 6: Probit estimation results of living/working in an ethnic enclave: households whose mother tongue is not English and do not reside in Quebec.

Dependent variable: Living/working in an ethnic enclave	Pooled
RGDP per capita (h)	0.211**
	(0.0771)
Housing costs in CMA/CA of arrival (I)	0.104***
	(0.0309)
Informal pre-immigration experience (ω)	-0.0920
	(0.0736)
Formal pre-immigration experience (ω)	-0.0892
	(0.0782)
Share of co-ethnic group in Canada (m)	6.923***
	(1.648)
Network (β)	0.162
	(0.0961)
Population level (L)	-0.0268**
	(0.00885)
Geographic distance (μ)	0.0000296
	(0.0000207)
Linguistic distance (θ)	0.0862
	(0.356)
Months since arrival	-0.0150
	(0.0236)
Time-varying controls (X_{ist})	✓
Time-invariant controls (X_{is})	✓
Survey wave controls (t)	✓
N	216,414

$^{*}p < 0.05^{**}p < 0.01^{***}p < 0.001$
Note 1: All models are weighted due to Statistics Canada's RDC disclosure process.
Note 2: Cluster robust standard errors in parentheses.

Consistent with the results presented in the previous section is a network interpretation of the model. The positive effect of source country richness on assimilating immigration can be explained by greater exposure to co-ethnics. But greater exposure to co-ethnics cannot be solely attributed to size of immigration and clustering of immigrants. There are other factors that determine clustering of co-ethnics; for instance, the qualitative features of the immigrant network must also matter.

Relative RGDP per capita may be interpreted as the immigrant group's average quality.

New immigrants that belong to a higher quality group are incentivised to locate among co-ethnics because the group, as a whole, assimilates more easily to the local culture. In this sense, quality of the network, rather than size, also plays a role. This is clear from table 6. Source country richness has a significant positive impact on clustering into co-ethnic communities, conditional on size of immigration. Similarly, connectedness of the co-ethnic community in addition to group quality and size might also be relevant.

4.4 Conclusion

Immigrants face frictions in communication in their life within the host country due to cultural gaps. Cultural assimilation plays a primary role in determining the success of immigrants by overcoming those cultural barriers. Immigrants assimilate by a process of learning. However, location choices are expected to crowd incentives to further learning so that exposure to other immigrants becomes the primary determinant on cultural accumulation.

Immigrants from poorer source countries have a lower assimilation rate than comparable immigrants from richer countries. This paper shows that exposure to co-ethnics is the primary determinant of lower assimilation rates because immigrants from poor source countries experience the largest productivity gains from migrating, leading to an increase in their presence within the host country. Location preferences among new immigrants in the host country further increases exposure to co-ethnics, further clustering into enclaves, lower communication frictions and fewer learning the local culture.

There are externalities (feedback) associated with the size of immigration within neighbourhoods. Some neighbourhoods will recieve more immigrants relative to others. Those neighbourhoods with a large number of immigrants will attract more of the non-assimilating type. Since immigrants from poor source countries are a bigger group, their exposure to co-ethnics will be larger and co-ethnic neighbourhoods will have lower assimilation rates. Among rich source countries total immigration to the host country is smaller. There are fewer externalities and the composition of immigration within neighbourhoods is favorable towards the assimilating type. Immigrants will sort, so that, neighbourhoods with fewer immigrants will receive all assimilating types.

Empirically, the implications of the model are tested using longitudinal Canadian data. As predicted immigrating from a richer source country has a positive effect on assimilating immigration. Exposure to co-ethnics through sorting is a major determinant of assimilation rates in the host country. Conditional on sorting, source country richness still has a significant positive effect on assimilation rates. There appears to be an alternate channel present. Presumably, this alternate channel is the quality, rather than size, of immigration.

Bibliography

Chapter 2

[1] Akee, Randall (2010). Who Leaves? Deciphering Immigrant Self-Selection from a Developing Country. *Economic Development and Cultural Change*, 58(2), pp. 323-344.

[2] Altonji, Joseph G.; Elder, Todd E. & Taber, Christopher R. (2002a). An Evaluation of Instrumental Variables Strategies for Estimating the Effects of Catholic Schools. *NBER Working Paper Series*, Working Paper No. 9358.

[3] Altonji, Joseph G.; Elder, Todd E. & Taber, Christopher R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1), pp. 151-184.

[4] Becker, Gary S. (1981). Altruism in the Family and Selfishness in the Market Place. *Economica*, 48(189), pp. 1-15.

[5] Bielby, William T. & Bielby, Denise D. (1992). I Will Follow Him: Family Ties, Gender-Role Beliefs, and Reluctance to Relocate for a Better Job. *American Journal of Sociology*, 97(5), pp. 1241-1267.

[6] Borjas, George J. (1987). Self-Selection and the Earnings of Immigrants. *The American Economic Review*, 77(4), pp. 531-553.

[7] Borjas, George J. & Bratsberg, Brent (1996). Who Leaves? The Out-Migration of the Foreign-Born. *The Review of Economics and Statistics*, 78(1), pp. 165-176.

[8] Borjas, George J. & Bronars, Stephen G. (1991). Immigration and the Family. *Journal of Labor Economics*, 9(2), pp. 123-148.

[9] Cai, Bing; Small, Dylan S. & Have, Thomas R. Ten (2000). Two-stage instrumental variables methods for estimating the causal odds ratio: analysis of bias. *Statistics in Medicine*, 30(15), pp. 1809-24.

[10] Chiswick, Barry R. (1999). Are Immigrants Favorably Selected? *The American Economic Review*, 89(2), pp. 181-185.

[11] Chiquiar, Daniel & Hanson, Gordon H. (2005). International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States. *Journal of Political Economy*, 113(2), pp. 239-281.

[12] Chen, Yiu Por (2005). Skill-Sorting, Self-Selectivity, and Immigration Policy Regime Change: Two Surveys of Chinese Graduate Students' Intention to Study Abroad. *The American Economic Review*, 95(2), pp. 66-70.

[13] Cobb-Clark, Deborah A. (1993). Immigrant Selectivity ad Wages: The Evidence for Women. *The American Economic Review*, Vol. 83, No. 4 (September 1993), pp. 986-993.

[14] Datta, Pranati (2005). Nepali Female Migration and Trafficking. *Journal of Sociological Science*, 11(1), pp. 49-56.

[15] Dutt, Srikant (1981). Migration and Development: The Nepalese in Northeast. *Economic and Political Weekly*, 16(24), pp. 1053-1055.

[16] Gubhaju, Bina & De Jong, Gordon F. (2009). Individual versus Household Migration Decision Rules: Gender and Marital Status Differences in Intentions to Migrate in South Africa. *International Migration*, 47(1), pp. 31-61.

[17] Hatton, Timothy, J. & Williamson, Jeffrey G. (2002). What Fundamentals Drive World Migration?. *NBER Working Paper Series*, Working Paper No. W9159.

[18] Katz, Eliakim & Stark, Oded (1986). Labor Migration and Risk Aversion in Less Developed Countries. *Journal of Labor Economics*, 4(1), pp. 134-149.

[19] Kley, Stefanie (2011). Explaining the Stages of Migration within a Life-course Framework. *European Sociological Review*, Vol. 27, No. 4 (2011), pp. 469-486.

[20] Kotorri, Mrika (2010). The Probability of Economic Emigration in Kosovo. *Eastern European Economics*, 48(6), pp. 17-35.

[21] Lacuesta, Aitor (2010). A Revision of the Self-Selection of Migrants Using Returning Migrant's Earnings. *Annals of Economics and Statistics*, 97/98, pp. 235-259.

[22] Long, Larry H. (1974). Women's Labor Force Participation and the Residential Mobility of Families. *Social Forces*, 52(3), pp. 342-348.

[23] Massey, Douglas S. (1990). Social Structure, Household Strategies, and the Cumulative Causation of Migration. *Population Index*, 56(1), pp. 3-26.

[24] Maxwell, Nan L. (1988). Economic Returns to Migration: Marital Status and Gender Differences. *Social Science Quarterly*, 69(1), pp. 108-121.

[25] McKinnish, Terra (2008). Spousal Mobility and Earnings, *Demography*, 45(4), pp. 829-849.

[26] McKenzie, David & Rapoport, Hillel (2010). Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks. *The Review of Economics and Statistics*, 92(4), pp. 811-821.

[27] Mendola, Mariapia (2012). Rural Out-migration and Economic Development at Origin: A Review of the Evidence.*Journal of International Development*, 24(1), pp. 102-122.

[28] Miller, Sheila J. (1976). Family Life Cycle, Extended Family Orientations, and Economic Aspirations as Factors in the Propensity to Migrate. *The Sociological Quarterly*, 17(3), pp. 323-335.

[29] Mincer, Jacob (1978). Family Migration Decisions. *Journal of Political Economy*, 86(5), pp. 749-773.

[30] Mont, Daniel (1989). Two earner family migration: A search theoretic approach. *Journal of Population Economics*, 2(1), pp. 55-72.

[31] Radu, Dragos (2008). Social Interactions in Economic Models of Migration: A Review and Appraisal. *Journal of Ethnic and Migration Studies*, 34(4), pp. 531-548.

[32] Root, Brenda Davis & De Jong, Gordon (1991). Family Migration in a Developing Country. *Population Studies: A Journal of Demography*, 45(2), pp. 221-233.

[33] Ruhs, Martin (2008). Economic research on labour immigration policy. *Oxford Review of Economic Policy*, 24(3), pp. 403-426.

[34] Seddon, David; Adhikari, Jagannath & Gurung, Ganesh (2002). Foreign Labor Migration and the Remittance Economy of Nepal. *Critical Asian Studies*, 34(1), pp. 19-40.

[35] Sharma, Sanjay & Thapa, Deepak (2013). Taken For Granted: Nepali Migration to India. *Centre for the Study of Labour and Mobility*, Working Paper III.

[36] Shields, Gail M. & Shields, Micheal P. (1989). Family Migration and Nonmarket Activities in Costa Rica. *Economic Development and Cultural Change*, 38(1), pp. 73-88.

[37] Speare Jr., Aldin & Goldscheider, Frances Kobrin (1987). Effects of Marital Status Change on Residential Mobility. *Journal of Marriage and Family*, 49(2), pp. 455-464.

[38] Stark, Oded & Bloom, David E. (1985). The new economics of labor migration. *American Economic Review*, 75(2), pp. 173-178.

[39] Stark, Oded & Levhari, David (1982). On Migration and Risk in LDCs. *Economic Development and Cultural Change*, 31(1), pp. 191-196.

[40] Stark, Oded & Taylor, J. Edward (1989). Relative Deprivation and International Migration. *Demography*, 26(1), pp. 1-14.

[41] Subedi, Bhim Prasad (1991). International Migration in Nepal: Towards an Analytical Framework. *CNAS Journal*, 18(1), pp. 83-102.

[42] Takenaka, Ayumi & Pren, Karen A. (2010). Determinants of Emigration: Comparing migrants' Selectivity from Peru and Mexico. *Annals of the American Academy of Political and Social Sciences* 630(1), pp. 178-193.

[43] Tenn, Steven (2010). The relative importance of husband's and wife's characteristics in family migration, 1960-2000. *Journal of Population Economics*, 23(4), pp. 1319-1337.

[44] USAID (2015). Demographic and Health Surveys: Standard Household File for Nepal 2011. *USAID*. Using The DHS Program. http://dhsprogram.com/Data/ (accessed October 4, 2018). Accompanying documentation: Codebook and User Guide.

Chapter 3

[45] Albrecht, James & Vroman, Susan (2002). A Matching Model with Endogenous Skill Requirements. *International Economic Review*, 43(1), pp. 283-305.

[46] Alesina, Alberto; Harnoss, Johann & Rapoport, Hillel (2015). Birthplace Diversity and Economic Prosperity. *Journal of Economic Growth*, 21(2), pp. 101-138.

[47] Aliprantis, C. D.; Camera, G. & Puzzello, D. (2007). A Random Matching Theory. *Games and Economic Behavior*, 59(1), pp. 1-16.

[48] Allen, James P. & Turner, Eugene (1996). Spatial Pattern of Immigrant Assimilation. *Professional Geographer*, 48(2), pp. 140-155.

[49] Antecol, Heather & Bedard, Kelly (2005). Unhealthy Assimilation: Why Do Immigrants Converge to American Health Status Levels? *IZA Discussion Paper Series*, Discussion Paper No. 1654.

[50] Bianchi, Milo (2013). Immigration Policy ad Self-Selecting Migrants. *Journal of Public Economic Theory*, 15(1), pp. 1-23.

[51] Biddle, Nicholas; Kennedy, Steven & McDonald, James T. (2007). Health Assimilation Patterns Amongst Australian Immigrants. *Economic Record*, 83(260), pp. 16-30.

[52] Bonn, Moritz (2012a). Costs and Benefits of Immigration and Multicultural Interaction. *MAGKS*, Joint Discussion Paper Series in Economics, No. 40-2012.

[53] Bonn, Moritz (2012b). Migrants' Acquisition of Cultural Skills and Selective Immigration Policies. *MAGKS*, Joint Discussion Paper Series in Economics, No. 47-2012.

[54] Borjas, George J. (1987). Self-selection and the earnings of immigrants. *American Economic Review*, 77(4), pp. 531-52.

[55] Borjas, George J. (1991). Immigration and Self-selection. In J. Aboud & R. Freeman (eds.) *Immigration, trade, and the labor market*, pp. 29-76.

[56] Borjas, George J. (1994). The Economics of Immigration. *Journal of Economic Literature*, 32, pp. 1667-1717.

[57] Borjas, George J. (1994). The Economic Analysis of Immigration. In O. Ashenfelter & D. Card (eds.) *Handbook of Labor Economics*, 3A, pp. 1697-1760.

[58] Bowlus, Audra J.; Miyairi, Masashi & Robinson, Chris (2016). Immigrant job search assimilation in Canada. *Canadian Journal of Economics*, 49(1), pp. 5-51.

[59] Card, David (1990). The Impact of the Mariel Boatlift on the Miami Labor Market. *Industrial and Labor Relations Review*, 43(2), pp. 245-257.

[60] Carliner, Geoffrey (2000). The Language Ability of US Immigrants: Assimilation and Cohort Effects. *International Migration Review*, 34(1), pp. 158-182.

[61] Chiswick, Barry R. (1978). The Effect of Americanization on the Earnings of Foreign-born Men. *Journal of Political Economy*, 86(5), pp. 897-921.

[62] Chiswick, Barry R. (1999). Are immigrants favorably selected? *The American Economic Review*, 89(2), pp. 181-185.

[63] Chiswick, Barry R. & Miller, Paul W. (1996). Ethnic networks and language proficiency among immigrants. *Journal of Population Economics*, 9(1), pp. 19-35.

[64] Chiswick, Barry R.; Lee, Yew Liang & Miller, Paul W. (2004). Immigrants' Language Skills: The Australian Experience in a Longitudinal Survey. *International Migration Review*, 38(2), pp. 611-654.

[65] Crawford, Vincent P. & Knoer, Elsie Marie (1981). Job Matching with Heterogeneous Firms and Workers. *Econometrica*, 49(2), pp. 437-450.

[66] Edin, Per-Anders; Frederiksson, Peter & Aslund, Olof (2003). Ethnic Enclaves and the Economic Success of Immigrants - Evidence from a Natural Experiment. *The Quarterly Journal of Economics*, 118(1), pp. 329-357.

[67] Espenshade, Thomas J. & Fu, Haishan (1997). An Analysis of English-Language Proficiency among U.S. Immigrants. *American Sociological Review*, 62(2), pp. 288-305.

[68] Green, Alan G. & Green, David A. (1999). The Economic Goals of Canada's Immigration Policy: Past and Present. *Canadian Public Policy*, 25(4), pp. 425-451.

[69] Haan, Micheal (2012). The Housing Experiences of New Canadians: Insights from the Longitudinal Survey of Immigrants in Canada (LSIC). *Citizenship and Immigration Canada, Research and Evaluation*, Reference No. RR20120301.

[70] Heston, Alan; Summers, Robert & Aten, Bettina (2005). Penn World Tables 6.3. *Centre for International Comparisons*, University of Pennsylvania.

[71] Horn, Daniel & Konya, Istvan (2016). Determinants of the Cultural Integration of Immigrants. In Z. Blask (eds.) *In Focus: International Migration*, pp. 150-158.

[72] Kelso, Alexander S. & Crawford, Vincent P. (1982). Job Matching, Coalition Formation, and Gross Substitutes. *Econometrica*, 50(6), pp. 1483-1504.

[73] Konya, Istvan (2007). Optimal Immigration and Cultural Assimilation. *Journal of Labor Economics*, 25(2), pp. 367-391.

[74] Lazear, Edward P. (1999). Culture and Language. *Journal of Political Economy*, 107(S6), pp. S95-S126.

[75] Massey, Douglas S. (1985). Ethnic residential segregation: A theoretical and empirical review. *Sociology and Social Research*, 69(3), pp. 315-350.

[76] Mayer, Thierry & Zignago, Soledad (2011). Notes on CEPII's distances measure: The GeoDist database. *CEPII Working Paper*, Working Paper No. 2011-25.

[77] Melitz, Jacques & Toubal, Farid (2014). Native language, spoken language, translation and trade. *Journal of International Economics*, 93(2), pp. 351-363.

[78] Orlov, George (2017). The Impact of Language Training on the Transfer of Pre-Immigration Skills and the Wages of Immigrants'. *Job Market Paper*, University of Western Ontario, accessed February 13, 2017.

[79] Ortega, Javier & Verdugo, Geroge (2015). Assimilation in multilingual cities. *Journal of Population Economics*, 28(3), pp. 785-815.

[80] Peri, Giovanni (2012). The Effect Of Immigration On Productivity: Evidence From U.S. States. *The Review of Economics and Statistics*, 94(1), pp. 348-358.

[81] Portes, Alejandro & Shafer, Steven (2006). Revisiting the Enclave Hypothesis: Miami Twenty-Five Years Later. *The Center for Migration and Development*, Working Paper Series, Working Paper No. 06-10.

[82] Rabe-Hesketh, Sophia & Skrondal, Anders (2012). Multilevel and Longitudinal Modeling Using Stata. Volume I: Continuous Responses. *STATA Press*, 3rd Edition.

[83] Rabe-Hesketh, Sophia & Skrondal, Anders (2012). Multilevel and Longitudinal Modeling Using Stata. Volume II: Categorical Responses, Counts and Survival. *STATA Press*, 3rd Edition.

[84] Roth, Alvin E. (1985). Common and Conflicting Interests in Two-sided Matching Markets. *European Economic Review*, 27(1), pp. 75-96.

[85] Statistics Canada (2013). Immigration and Ethnocultural Diversity in Canada. *Statistics Canada*. Statistics Canada Catalogue no. 99-010-X2011001.

[86] Statistics Canada (2007). Longitudinal Survey of Immigrants to Canada (LSIC). *Statistics Canada*. Using York University Research Data Center. http://researchguides.library.yorku.ca/data/microdata (accessed October 4, 2018). Accompanying documentation: Microdata User Guide.

[87] Sonmez, Tayfun & Unver, M. U. (2009). Matching, Allocation, and Exchange of Discrete Resources. In A. Bisin, J. Benhabib & M. Jackson (eds.) *Handbook of Social Economics*, 1A, pp. 781-852.

[88] Tani, Massimiliano (2014). Using a point system for selecting immigrants. *IZA World of Labor*, 24.

[89] van Tubergen, Frank & Kalmijn, Matthijs (2009). A Dynamic Approach to the Determinants of Immigrants' Language Proficiency: The United States, 1980-2000. *The International Migration Review*, 43(3), pp. 519-543.

[90] Vigdor, Jacob L. (2008). Measuring Immigrant Assimilation in the United States. *Manhattan Institute for Policy Research*, Civic Report No. 53.

[91] Xie, Yu & Gough, Margaret (2011). Ethnic Enclaves and the Earnings of Immigrants. *Demography*, 48(4), pp. 1293-1315.

Chapter 4

[92] Allen, James P. & Turner, Eugene (1996). Spatial Pattern of Immigrant Assimilation. *Professional Geographer*, 48(2), pp. 140-155.

[93] Bauer, Thomas; Epstein, Gill S. & Gang, Ira N. (2005). Enclaves, language, and the location choice of migrants. *Journal of Population Economics*, 18(4), pp. 649-662.

[94] Bailey, Thomas & Waldinger, Roger (1991). Primary, Secondary, And Enclave Labor Markets: A Training Systems Approach. *American Sociological Review*, 56(4), pp. 432-445.

[95] Bonn, Moritz (2012a). Costs and Benefits of Immigration and Multicultural Interaction. *MAGKS*, Joint Discussion Paper Series in Economics, No. 40-2012.

[96] Bonn, Moritz (2012b). Migrants' Acquisition of Cultural Skills and Selective Immigration Policies. *MAGKS*, Joint Discussion Paper Series in Economics, No. 47-2012.

[97] Chiswick, Barry R. & Miller, Paul W. (1996). Ethnic networks and language proficiency among immigrants. *Journal of Population Economics*, 9(1), pp. 19-35.

[98] Chiswick, Barry R.; Lee, Yew Liang & Miller, Paul W. (2004). Immigrants' Language Skills: The Australian Experience in a Longitudinal Survey. *International Migration Review*, 38(2), pp. 611-654.

[99] Edin, Per-Anders; Frederiksson, Peter & Aslund, Olof (2003). Ethnic Enclaves and the Economic Success of Immigrants - Evidence from a Natural Experiment. *The Quarterly Journal of Economics*, 118(1), pp. 329-357.

[100] EF EPI (2017). EF English Speaking Proficiency. *Education First*. Seventh Edition.

[101] Haan, Micheal (2012). The Housing Experiences of New Canadians: Insights from the Longitudinal Survey of Immigrants in Canada (LSIC). *Citizenship and Immigration Canada, Research and Evaluation*, Reference No. RR20120301.

[102] Heston, Alan; Summers, Robert & Aten, Bettina (2005). Penn World Tables 6.3. *Centre for International Comparisons*, University of Pennsylvania.

[103] Horn, Daniel & Konya, Istvan (2016). Determinants of the Cultural Integration of Immigrants. In Z. Blask (eds.) *In Focus: International Migration*, pp. 150-158.

[104] Konya, Istvan (2007). Optimal Immigration and Cultural Assimilation. *Journal of Labor Economics*, 25(2), pp. 367-391.

[105] Lazear, Edward P. (1999). Culture and Language. *Journal of Political Economy*, 107(S6), pp. S95-S126.

[106] Lazear, Edward P. (2017). Why are Some Immigrant Groups More Successful than Others? *NBER Working Paper Series*, Working Paper No. 23548.

[107] Massey, Douglas S. (1985). Ethnic residential segregation: A theoretical and empirical review. *Sociology and Social Research*, 69(3), pp. 315-350.

[108] Mayer, Thierry & Zignago, Soledad (2011). Notes on CEPII's distances measure: The GeoDist database. *CEPII Working Paper*, Working Paper No. 2011-25.

[109] Melitz, Jacques & Toubal, Farid (2014). Native language, spoken language, translation and trade. *Journal of International Economics*, 93(2), pp. 351-363.

[110] Portes, Alejandro & Shafer, Steven (2006). Revisiting the Enclave Hypothesis: Miami Twenty-Five Years Later. *The Center for Migration and Development*, Working Paper Series, Working Paper No. 06-10.

[111] Rabe-Hesketh, Sophia & Skrondal, Anders (2012). Multilevel and Longitudinal Modeling Using Stata. Volume I: Continuous Responses. *STATA Press*, 3rd Edition.

[112] Rabe-Hesketh, Sophia & Skrondal, Anders (2012). Multilevel and Longitudinal Modeling Using Stata. Volume II: Categorical Responses, Counts and Survival. *STATA Press*, 3rd Edition.

[113] Statistics Canada (2007). Longitudinal Survey of Immigrants to Canada (LSIC). *Statistics Canada*. Using York University Research Data Center. http://researchguides.library.yorku.ca/data/microdata (accessed October 4, 2018). Accompanying documentation: Microdata User Guide.

[114] van Tubergen, Frank & Kalmijn, Matthijs (2009). A Dynamic Approach to the Determinants of Immigrants' Language Proficiency: The United States, 1980-2000. *The International Migration Review*, 43(3), pp. 519-543.

[115] Waldinger, Roger (1993). The Ethnic Enclave Debate Revisited. *International Journal of Urban and Regional Research*, 17(3), pp. 428-436.

[116] Warman, Casey (2007). Ethnic Enclaves and Immigrant Earnings Growth. *The Canadian Journal of Economics*, 40(2), pp. 401-422.

[117] Xie, Yu & Gough, Margaret (2011). Ethnic Enclaves and the Earnings of Immigrants. *Demography* 48(4), pp. 1293-1315.

Appendix

Table A1: Number of female and male headed households with in-laws, grandchildren, brothers/sisters, other relatives, and unrelated persons living in the household (excludes migrants), by number of persons.

Number of persons in the household	Female headed households									
	In-laws		Grandchildren		Brothers/sisters		Other relatives		Unrelated persons	
0	2344	(78.34)	2454	(82.02)	2845	(95.09)	2620	(87.57)	2936	(98.13)
1	551	(18.42)	234	(7.821)	115	(3.844)	241	(8.055)	44	(1.471)
2	89	(2.975)	163	(5.448)	22	(0.735)	75	(2.507)	5	(0.167)
3	6	(0.201)	88	(2.941)	10	(0.334)	33	(1.103)	4	(0.134)
4	2	(0.0668)	30	(1.003)			10	(0.334)	3	(0.100)
5			16	(0.535)			7	(0.234)		
6			5	(0.167)			2	(0.0668)		
7			1	(0.0334)			2	(0.0668)		
8			1	(0.0334)			1	(0.0334)		
9										
12							1	(0.0334)		
Total	2992	(100.00)	2992	(100.00)	2992	(100.00)	2992	(100.00)	2992	(100.00)
	Male headed households									
0	6048	(77.20)	6128	(78.22)	7370	(94.08)	6960	(88.84)	7587	(96.85)
1	1498	(19.12)	682	(8.706)	357	(4.557)	566	(7.225)	187	(2.387)
2	255	(3.255)	550	(7.021)	80	(1.021)	159	(2.030)	34	(0.434)
3	31	(0.396)	265	(3.383)	23	(0.294)	76	(0.970)	19	(0.243)
4	1	(0.0128)	121	(1.545)	3	(0.0383)	30	(0.383)	3	(0.0383)
5	1	(0.0128)	46	(0.587)			15	(0.191)	2	(0.0255)
6			24	(0.306)			9	(0.115)	1	(0.0128)
7			7	(0.0894)	1	(0.0128)	8	(0.102)		
8			6	(0.0766)			4	(0.0511)		
9			3	(0.0383)			5	(0.0638)		
12			1	(0.0128)						
13			1	(0.0128)						
10							1	(0.0128)		
16							1	(0.0128)		
14									1	(0.0128)
Total	7834	(100)	7834	(100)	7834	(100)	7834	(100)	7834	(100)

Column percent in parentheses

Table A2: Number of migrant and non-migrant households with in-laws, grandchildren, brothers/sisters, other relatives, and unrelated persons living in the household (excludes migrants), by number of persons.

Number of persons in the household	**Non-migrant households**									
	In-laws		Grandchildren		Brothers/sisters		Other relatives		Unrelated persons	
0	4274	(86.34)	4340	(87.68)	4646	(93.86)	4390	(88.69)	4777	(96.51)
1	589	(11.90)	231	(4.667)	233	(4.707)	378	(7.636)	120	(2.424)
2	78	(1.576)	185	(3.737)	54	(1.091)	107	(2.162)	28	(0.566)
3	8	(0.162)	115	(2.323)	16	(0.323)	40	(0.808)	16	(0.323)
4			51	(1.030)	1	(0.0202)	17	(0.343)	6	(0.121)
5	1	(0.0202)	15	(0.303)			7	(0.141)	2	(0.0404)
6			10	(0.202)			1	(0.0202)	1	(0.0202)
7							5	(0.101)		
8			2	(0.0404)			1	(0.0202)		
9			1	(0.0202)			3	(0.0606)		
12										
13										
10							1	(0.0202)		
Total	4950	(100)	4950	(100)	4950	(100)	4950	(100)	4950	(100)
	Migrant households									
0	4118	(70.08)	4242	(72.19)	5569	(94.78)	5190	(88.33)	5746	(97.79)
1	1460	(24.85)	685	(11.66)	239	(4.067)	429	(7.301)	111	(1.889)
2	266	(4.527)	528	(8.986)	48	(0.817)	127	(2.161)	11	(0.187)
3	29	(0.494)	238	(4.050)	17	(0.289)	69	(1.174)	7	(0.119)
4	3	(0.0511)	100	(1.702)	2	(0.0340)	23	(0.391)		
5			47	(0.800)			15	(0.255)		
6			19	(0.323)			10	(0.170)		
7			8	(0.136)	1	(0.0170)	5	(0.0851)		
8			5	(0.0851)			4	(0.0681)		
9			2	(0.0340)			2	(0.0340)		
12			1	(0.0170)			1	(0.0170)		
13			1	(0.0170)						
16							1	(0.0170)		
14									1	(0.0170)
Total	5876	(100.00)	5876	(100.00)	5876	(100.00)	5876	(100.00)	5876	(100.00)

Column percent in parentheses

Table A3: Number of households with members that migrated for work, study, marriage, other family, and security reasons, by number of migrants.

Number of migrant members	Work		Study		Marriage		Other family		Security	
0	6986	(64.53)	9441	(87.21)	8871	(81.94)	9936	(91.78)	10819	(99.94)
1	2930	(27.06)	1013	(9.357)	1488	(13.74)	444	(4.101)	6	(0.0554)
2	719	(6.641)	279	(2.577)	380	(3.510)	237	(2.189)	1	(0.00924)
3	151	(1.395)	72	(0.665)	77	(0.711)	118	(1.090)		
4	29	(0.268)	18	(0.166)	8	(0.0739)	44	(0.406)		
5	9	(0.0831)	2	(0.0185)	2	(0.0185)	26	(0.240)		
6	1	(0.00924)	1	(0.00924)			10	(0.0924)		
7	1	(0.00924)					3	(0.0277)		
8							2	(0.0185)		
9							3	(0.0277)		
10							1	(0.00924)		
11							1	(0.00924)		
12							1	(0.00924)		
Total	10826	(100)	10826	(100)	10826	(100)	10826	(100)	10826	(100)

Percent in parentheses

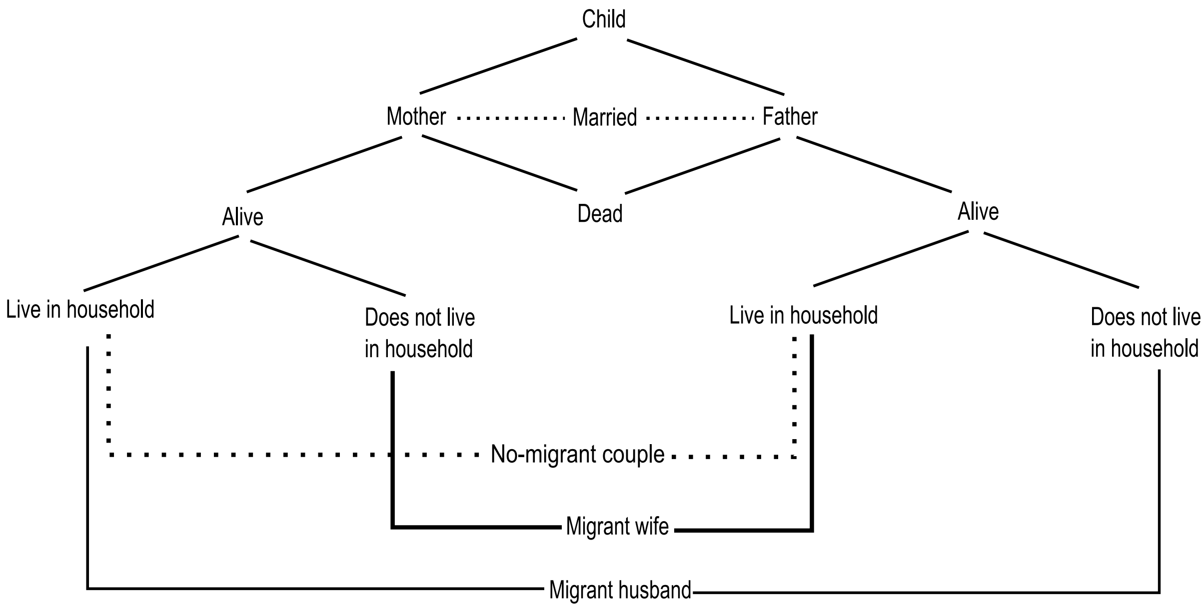

Figure 4.5: Inferring migrant husbands, migrant wives, and non-migrant couples from available information in the DHS, Nepal 2011 (household file).

Table A.S1: Summary Statistics of Nepali households that supply migrants to India.

Variable	Description	Mean	Standard Deviation	Count
Migrant Husband	(0 or 1) Household has a migrant husband (see figure 1)	0.389	0.488	548
Extended Family	(0 or 1) Presence of extended family member in the household	0.429	0.495	548
Poor	(0 or 1) Household is poor based on wealth index	0.582	0.494	548
Urban	(0 or 1) Household is located in an urban area	0.201	0.401	548
Eastern	(0 or 1) Household is located in the Eastern region of Nepal	0.128	0.334	548
Central	(0 or 1) Household is located in the Central region of Nepal	0.137	0.344	548
Western	(0 or 1) Household is located in the Western region of Nepal	0.161	0.367	548
Mid-western	(0 or 1) Household is located in the Mid-western region of Nepal	0.223	0.416	548
Cluster Altitude	(Continuous) Location of household above sea level (in meters)	678.726	651.090	548
Age of Oldest Child	(Continuous) Family life cycle variable	13.929	9.029	548
Secondary Education	(Continuous) Proportion of household members aged 25+ with a secondary education	0.013	0.066	548
Higher Education	(Continuous) Proportion of household members aged 25+ with higher education	0.015	0.083	548
Literacy Program	(Continuous) Proportion of household members aged 25+ that have participated in literacy program	0.072	0.133	548
Male Members	(Continuous) Proportion of male members in the household	0.412	0.200	548
Male Headed	(0 or 1) Household is male headed	0.432	0.496	548
Age of Head	(Continuous) Age of head of the household	40.797	14.209	548
Age at Migration	(Continuous) Age at migration of the migrant	19.607	15.083	759
Mumbai	(0 or 1) Location of migrant: Mumbai	0.088	0.284	759
Delhi	(0 or 1) Location of migrant: Delhi	0.126	0.333	759
Punjab	(0 or 1) Location of migrant: Punjab	0.050	0.218	759
Other city in India	(0 or 1) Location of migrant: Other city in India	0.391	0.488	759

Table A.S2: Summary Statistics of Nepali households that supply migrants within Nepal.

Variable	Description	Mean	Standard Deviation	Count
Migrant Husband	(0 or 1) Household has a migrant husband (see figure 1)	0.156	0.363	1534
Extended Family	(0 or 1) Presence of extended family member in the household	0.471	0.499	1534
Poor	(0 or 1) Household is poor based on wealth index	0.419	0.494	1534
Urban	(0 or 1) Household is located in an urban area	0.264	0.441	1534
Eastern	(0 or 1) Household is located in the Eastern region of Nepal	0.225	0.418	1534
Central	(0 or 1) Household is located in the Central region of Nepal	0.267	0.443	1534
Western	(0 or 1) Household is located in the Western region of Nepal	0.151	0.358	1534
Mid-western	(0 or 1) Household is located in the Mid-western region of Nepal	0.198	0.398	1534
Cluster Altitude	(Continuous) Location of household above sea level (in meters)	805.688	674.130	1534
Age of Oldest Child	(Continuous) Family life cycle variable	16.450	10.348	1534
Secondary Education	(Continuous) Proportion of household members aged 25+ with a secondary education	0.034	0.103	1534
Higher Education	(Continuous) Proportion of household members aged 25+ with higher education	0.040	0.128	1534
Literacy Program	(Continuous) Proportion of household members aged 25+ that have participated in literacy program	0.082	0.134	1534
Male Members	(Continuous) Proportion of male members in the household	0.459	0.200	1534
Male Headed	(0 or 1) Household is male headed	0.716	0.451	1534
Age of Head	(Continuous) Age of head of the household	46.175	13.907	1534
Age at Migration	(Continuous) Age at migration of the migrant	10.928	12.181	2917
Kathmandu	(0 or 1) Location of migrant: Mumbai	0.113	0.316	2917
Other city in Nepal	(0 or 1) Location of migrant: Delhi	0.143	0.350	2917
Rural area in Nepal	(0 or 1) Location of migrant: Punjab	0.268	0.443	2917

Proofs

Proof to Proposition 1(a). Let $i = i'$ be the marginal migrant that comes to the North from the South and changes the level of migration from m to m'. The decision to migrate indicator for i' is $d_{i'} = 1$. The positive externality to immigrants from i' is given by

$$\begin{aligned}\frac{m'}{1+m'} - \frac{m}{1+m} &= \frac{d_{i'} + \sum_{i=1}^{L} d_i}{1 + d_{i'} + \sum_{i=1}^{L} d_i} - \frac{\sum_{i=1}^{L} d_i}{1 + \sum_{i=1}^{L} d_i} \\ &= \frac{1 + \sum_{i=1}^{L} d_i}{2 + \sum_{i=1}^{L} d_i} - \frac{\sum_{i=1}^{L} d_i}{1 + \sum_{i=1}^{L} d_i} \\ &= \frac{1+m}{2+m} - \frac{m}{1+m} > 0\end{aligned}$$

The first equality comes from dis-aggregating m, m', and separating $d_{i'}$ from the summation. The final equality uses the assumption that i' is an immigrant $d_{i'} = 1$. □

Proof to Proposition 1(b). Let $i = i'$ be the marginal migrant that comes to the North from the South and changes the level of migration from m to m' as well as the level of assimilating immigration a to a'. The decision to migrate indicator for i' is $d_{i'} = 1$ and the decision to assimilate indicator is $a_{i'} = 1$. The negative externality to native-born from i' is given by

$$\begin{aligned}\frac{1+a'}{1+m'} - \frac{1+a}{1+m} &= \frac{1 + a_{i'} + \sum_{i=1}^{L} a_i}{1 + d_{i'} + \sum_{i=1}^{L} d_i} - \frac{1 + \sum_{i=1}^{L} a_i}{1 + \sum_{i=1}^{L} d_i} \\ &= \frac{2 + \sum_{i=1}^{L} a_i}{2 + \sum_{i=1}^{L} d_i} - \frac{1 + \sum_{i=1}^{L} a_i}{1 + \sum_{i=1}^{L} d_i} \\ &= \frac{2+a}{2+m} - \frac{1+a}{1+m} > 0\end{aligned}$$

if the marginal migrant i' is assimilating. The first equality comes form dis-aggregating m, m', a, a', and separating $d_{i'}$ and $a_{i'}$ from the summation. The final equality uses the assumption that i' is an immigrant $d_{i'} = 1$ and non-assimilating $a_{i'} = 1$. On the other hand, if the marginal migrant were non-assimilating we have $a_{i'} = 0$ in the second equality and have the following instead

$$\frac{1 + \sum_{i=1}^{L} a_i}{2 + \sum_{i=1}^{L} d_i} - \frac{1 + \sum_{i=1}^{L} a_i}{1 + \sum_{i=1}^{L} d_i} = \frac{1+a}{2+m} - \frac{1+a}{1+m} < 0$$

□

Comparative statics over the thresholds

The following are a set of useful comparative statics used throughout the main body of the text in derivations, proofs and analysis. The effect of θ and m on the threshold for assimilating is given by

$$\omega_a = 1 - \frac{1}{\theta\nu(1+m)}, \qquad \frac{\partial \omega_a}{\partial \theta} = \frac{1}{\theta^2\nu(1+m)} > 0, \qquad \frac{\partial \omega_a}{\partial m} = \frac{1}{\theta\nu(1+m)^2} > 0.$$

The equation $c_a(\omega)$ is the thresholds for assimilating immigration. The following derivatives provide the effect of changes in h, μ and θ.

$$c_a(\omega) = \frac{1 - \theta\nu(1-\omega) - h}{\mu},$$

$$\frac{\partial c_a(\omega)}{\partial h} = -\frac{1}{\mu} < 0, \qquad \frac{\partial c_a(\omega)}{\partial \mu} = -\frac{1 - \theta\nu(1-\omega) - h}{\mu^2} < 0, \qquad \frac{\partial c_a(\omega)}{\partial \theta} = -\frac{\nu(1-\omega)}{\mu} < 0.$$

The equation c_n is the threshold for non-assimilating immigration. The following set of derivatives provide the effect of changes in h, μ and m.

$$c_n^i(\beta) = \frac{m^i/(1+m^i) - h/\beta}{\mu},$$

$$\frac{\partial c_n}{\partial h} = -\frac{1}{\mu} < 0, \qquad \frac{\partial c_n}{\partial \mu} = -\frac{m/(1+m) - h}{\mu^2} \begin{cases} \geq 0 & \text{if } h \geq \overline{h} \\ < 0 & \text{if } h < \overline{h} \end{cases}, \qquad \frac{\partial c_n}{\partial m} = \frac{1}{\mu(1+m)^2} > 0,$$

where $\overline{h} = m/(1+m)$ is the threshold between the interior and corner equilibria. The effect of geographic distance on non-assimilating immigration depends on the prevailing equilibrium. Finally, the level of ω at which the corner equilibrium occurs is $\underline{\omega}$. Changes associated with $\underline{\omega}$ due to h and θ are provided

$$\underline{\omega} = 1 - \frac{1-h}{\theta\nu}, \qquad \frac{\partial \underline{\omega}}{\partial h} = \frac{1}{\theta\nu} > 0, \qquad \frac{\partial \underline{\omega}}{\partial \theta} = \frac{1-h}{\theta^2\nu} > 0$$

Comparative statics over $\Gamma(m)$

Consider an interior equilibrium outcome $c_n > 0$. The set of first-order partial differentials of $\Gamma(m)$ in terms of m and $\mathbb{P} \equiv \{L, h, \mu, \theta\}$ are as follows,

$$\begin{aligned}
m &= LF(c_n)W(\omega_a) + L\int_{\omega_a}^{1} F(c_a(\omega))dW(\omega) \\
\frac{\partial \Gamma}{\partial m} &= LF'(c_n)\frac{\partial c_n}{\partial m}W(\omega_a) + LF(c_n)W'(\omega_a)\frac{\partial \omega_a}{\partial m} - LF(c_a(\omega_a))\frac{\partial \omega_a}{\partial m} \\
\frac{\partial \Gamma}{\partial L} &= F(c_n)W(\omega_a) + \int_{\omega_a}^{1} F(c_a(\omega))dW(\omega) > 0 \\
\frac{\partial \Gamma}{\partial h} &= LF'(c_n)\frac{\partial c_n}{\partial h}W(\omega_a) + L\int_{\omega_a}^{1} F'(c_a(\omega))\frac{\partial c_a(\omega)}{\partial h}dW(\omega) < 0 \\
\frac{\partial \Gamma}{\partial \mu} &= LF'(c_n)\frac{\partial c_n}{\partial \mu}W(\omega_a) + L\int_{\omega_a}^{1} F'(c_a(\omega))\frac{\partial c_a(\omega)}{\partial \mu}dW(\omega) < 0, \\
\frac{\partial \Gamma}{\partial \theta} &= LF(c_n)W'(\omega_a)\frac{\partial \omega_a}{\partial \theta} - LF(c_a(\omega_a))\frac{\partial \omega_a}{\partial \theta} + L\int_{\omega_a}^{1} F'(c_a(\omega))\frac{\partial c_a(\omega)}{\partial \theta}dW(\omega) \lesseqqgtr 0,
\end{aligned}$$

In a corner equilibrium outcome $c_n = 0$. The first-order differentials of $\Gamma(m)$ in terms of m and $\{L, h, \mu, \theta\}$ are as follows,

$$\begin{aligned}
m &= L\int_{\underline{\omega}}^{1} F(c_a(\omega))dW(\omega), \\
\frac{\partial \Gamma}{\partial m} &= 0, \\
\frac{\partial \Gamma}{\partial L} &= \int_{\underline{\omega}}^{1} F(c_a(\omega))dW(\omega) > 0, \\
\frac{\partial \Gamma}{\partial h} &= L\int_{\underline{\omega}}^{1} F'(c_a(\omega))\frac{\partial c_a(\omega)}{\partial h}dW(\omega) - LF(c_a(\omega))\frac{\partial \underline{\omega}}{\partial h}dW(\omega) < 0, \\
\frac{\partial \Gamma}{\partial \mu} &= L\int_{\underline{\omega}}^{1} F'(c_a(\omega))\frac{\partial c_a(\omega)}{\partial \mu}dW(\omega) < 0, \\
\frac{\partial \Gamma}{\partial \theta} &= L\int_{\underline{\omega}}^{1} F'(c_a(\omega))\frac{\partial c_a(\omega)}{\partial \theta}dW(\omega) - LF(c_a(\omega))\frac{\partial \underline{\omega}}{\partial \theta} < 0,
\end{aligned}$$

Comparative statics over $m(L, h, \mu, \theta)$

Since $m = \Gamma(m; \mathbb{P})$ implicitly defines m, where $\mathbb{P} \equiv \{L, h, \mu, \theta\}$. Then totally differentiating with respect to the parameters in $\mathbb{P}$ is given by

$$\frac{dm}{d\mathbb{P}} = \frac{\partial \Gamma}{\partial m}\frac{\partial m}{\partial \mathbb{P}} + \frac{\partial \Gamma}{\partial \mathbb{P}},$$
$$\frac{dm}{d\mathbb{P}} = \frac{\partial \Gamma / \partial \mathbb{P}}{1 - \partial \Gamma / \partial m}$$

The signs of the partial differentials of $\partial \Gamma / \partial \mathbb{P}$ have already been determined in the previous section. Moreover, the stability conditions require that $\partial \Gamma / \partial m < 1$. Using this information the effect of $\mathbb{P}$ on m in the interior and corner cases are given as

$$\frac{dm}{dL} > 0, \qquad \frac{dm}{dh} < 0, \qquad \frac{dm}{d\mu} \begin{cases} \lesseqgtr 0 & \text{if } h \geq \overline{h} \\ < 0 & \text{if } h < \overline{h} \end{cases}, \qquad \frac{dm}{d\theta} \lesseqgtr 0.$$

Direct effects

The direct effects of table 2 are determined by changes in $\mathbb{P} = \{L, h, \mu, \theta\}$ on $\rho = a/m$. To simplify the calculations, let $\eta \equiv \mathbb{A}/\mathbb{B}$, where

$$\mathbb{A} = L \int_{\omega_a}^{1} F(c_a(\omega) dW(\omega) > 0$$
$$\mathbb{B} = LF(c_n) > 0$$

Now ρ can be written in a simpler form, as a function of η. It can then be inferred that ρ is a monotonic increasing function of η.

$$\rho = \frac{\eta}{\eta + W(\omega_a)}, \qquad \frac{\partial \rho}{\partial \eta} > 0$$

The parameters in the model $\mathbb{P} = \{L, h, \mu, \theta\}$ are expected to have a direct effect on $\rho = a/m$. These effects are estimated by the relationship

$$\frac{\partial \rho}{\partial \mathbb{P}} = \frac{\partial \rho}{\partial \eta} \times \frac{\partial \eta}{\partial \mathbb{P}}, \qquad \text{where} \quad \frac{\partial \eta}{\partial \mathbb{P}} = \frac{\partial \mathbb{A}/\partial \mathbb{P} \times \mathbb{B} - \partial \mathbb{B}/\partial \mathbb{P} \times \mathbb{A}}{\mathbb{B}^2}.$$

All that remains to sign the direct effect by differentiating $\mathbb{A}$ and $\mathbb{B}$ in terms of the elements in $\mathbb{P} = \{L, h, \mu, \theta\}$.

$$\begin{aligned}
\frac{\partial \mathbb{A}}{\partial L} &= \int_{\omega_a}^{1} F(c_a(\omega))dW(\omega) > 0,\\
\frac{\partial \mathbb{A}}{\partial h} &= L\int_{\omega_a}^{1} f(c_a(\omega))\frac{\partial c_a(\omega)}{\partial h}dW(\omega) < 0,\\
\frac{\partial \mathbb{A}}{\partial \mu} &= L\int_{\omega_a}^{1} f(c_a(\omega))\frac{\partial c_a(\omega)}{\partial \mu}dW(\omega) < 0,\\
\frac{\partial \mathbb{A}}{\partial \theta} &= -LF(c_a(\omega_a))\frac{\partial \omega_a}{\partial \theta} + L\int_{\omega_a}^{1} f(c_a(\omega))\frac{\partial c_a(\omega)}{\partial \theta}dW(\omega) < 0,\\
\frac{\partial \mathbb{A}}{\partial m} &= -LF(c_a(\omega))\frac{\partial \omega_a}{\partial m} < 0,\\
\frac{\partial \mathbb{B}}{\partial L} &= F(c_n) > 0,\\
\frac{\partial \mathbb{B}}{\partial h} &= Lf(c_n)\frac{\partial c_n}{\partial h} < 0,\\
\frac{\partial \mathbb{B}}{\partial \mu} &= Lf(c_n)\frac{\partial c_n}{\partial \mu} < 0\\
\frac{\partial \mathbb{B}}{\partial \theta} &= 0,\\
\frac{\partial \mathbb{B}}{\partial m} &= Lf(c_n)\frac{\partial c_n}{\partial m}dB(\beta) > 0.
\end{aligned}$$

Indirect effects

The indirect effects of table 2 are determined by changes in $\mathbb{P} = \{L, h, \mu, \theta\}$ on $\rho = a/m$ through changes in $m(L, h, \mu, \theta)$. These effects are estimated by the relationship

$$\frac{\partial \rho}{\partial \mathbb{P}} = \frac{\partial \rho}{\partial \eta} \times \frac{\partial \eta}{\partial m} \times \frac{\partial m}{\partial \mathbb{P}}, \qquad \text{where} \quad \frac{\partial \eta}{\partial m} = \frac{\partial \mathbb{A}/\partial m \times \mathbb{B} - \partial \mathbb{B}/\partial m \times \mathbb{A}}{\mathbb{B}^2}.$$

The signs for $\partial \mathbb{A}/\partial m$ and $\partial \mathbb{B}/\partial m$ were determined in the previous section. The signs for the indirect effects are as follows,

$$\frac{\partial \rho}{\partial L} < 0, \qquad \frac{\partial \rho}{\partial h} > 0, \qquad \frac{\partial \rho}{\partial \mu} > 0, \qquad \frac{\partial \rho}{\partial \theta} \lesseqqgtr 0.$$

Threshold effects

The threshold effects are identified by changes in $\mathbb{P} = \{L, h, \mu, \theta\}$ on equation (3.15), or $\Phi(m(L, \overline{h}, \mu, \theta), L, \overline{h}, \mu, \theta)$. These effects are provided below.

$$\frac{d\Phi}{dL} < 0, \qquad \frac{d\Phi}{dh} > 0, \qquad \frac{d\Phi}{d\mu} > 0, \qquad \frac{d\Phi}{d\theta} > 0.$$

Tables

Table A1: Ethnic, linguistic and changes in birthplace diversity indices across major immigrant receiving countries

Country	Ethnic	Linguistic	% change in birthplace diversity 1990-2000	
			All	Migrants
Australia	0.093	0.335	-4%	3%
Austria	0.107	0.152	119%	3%
Belgium	0.555	0.541	15%	2%
Canada	0.712	0.577	8%	1%
France	0.103	0.122	-1%	1%
Germany	0.168	0.164	33%	1%
Netherlands	0.514	0.335	2%	2%
Norway	0.059	0.067	46%	1%
New Zealand	0.397	0.166	8%	17%
Switzerland	0.531	0.544	0%	3%
USA	0.490	0.251	40%	-3%
UK	0.121	0.053	21%	2%

Source: Alesina, Alberto; Harnoss, Johann & Rapoport, Hillel (2015). Birthplace Diversity and Economic Prosperity. *Journal of Economic Growth*, Vol. 21, issue 2 (June 2016), pp. 101-138.

Table A2: Descriptions of variables used in this study.

Variable Name	**Description**
Speaking proficiency	English speaking proficiency of the respondent (Dummy)
Formal investment	Learned to speak English in language class or school (Dummy)
Informal investment	Learned to speak English through family & friends, work, non-language classes, media, self-study and/or other (Dummy)
CMA of arrival: Toronto	The respondent lives in Toronto at the time of landing in Canada (Dummy)
CMA of arrival: Vancouver	The respondent lived in Vancouver at the time of landing in Canada (Dummy)
Share of co-ethnic group in Canada	Share of the co-ethnic population that the respondent belongs to admitted into Canada in 2000 (Continuous)
Share of co-ethnic group in CMA/CA of arrival	Share of the ethnic population that the respondent belongs to living in the same CMA/CA of first arrival (Continuous; Canadian Census of Population, 2001)
Education outside Canada: Bachelor's or higher	Highest level of formal education attained outside of Canada is Bachelor's or higher (Dummy)
Formal learning of English outside of Canada	Learned most of their English before coming to Canada through language classes, a private tutor or schooling (Dummy)
Informal learning of English outside of Canada	Learned most of their English before coming to Canada from family/friends, self-study, work, media, everyday interactions or other (Dummy)
Age at immigration	Age of the respondent (Continuous)
Male	Gender of the respondent = Male (Dummy)
Married	Marital status of the respondent = Married (Dummy)
Employed	The respondent is employed full-time (Dummy)
Savings brought from outside Canada	Total amount of savings brough from outside Canada (Continuous)
Family class	Immigration category = Family class (Dummy)

Table A2: Descriptions of variables used in this study.

Variable Name	**Description**
Number of immigrating members	Size of the immigrating party at the time of landing (Continuous)
Number of joining members	Number of members that have joined the household since landing (wave 1) or since the last interview (waves 2 and 3) (Continuous)
Number of household members	Size of household (Continuous)
Number of children	Percentage of children living in the household (Continuous)
Months since arrival	Number of months elapsed since landing in Canada (Continuous)
RGDP in place of Birth	Real GDP of place of birth of the respondent relative to Canada (Continuous; Heston, Summers & Aten, 2005)
Population in place of birth	Population of place of birth of the respondent relative to Canada (Continuous; Heston, Summers & Aten, 2005)
Distance from capital city in place of birth	Distance of the capital city in the place of birth of the respondent relative to Canada (Continuous; Mayer & Zignago, 2011)
Linguistic distance	Linguistic distance of the official language in the place of birth of the respondent relative to English (Continuous; Melitz & Toubal, 2014)

Table A3: Summary statistics.

	Full sample			Sub-sample		
Variable Name	Count	Mean	SD	Count	Mean	SD
Speaking proficiency	291,135	0.692	0.462	229,362	0.700	0.458
Formal investment	276,555	0.314	0.464	216,129	0.349	0.477
Informal investment	290,853	0.503	0.500	229,080	0.554	0.497
CMA of arrival: Toronto	291,141	0.484	0.500	229,368	0.554	0.497
CMA of arrival: Vancouver	291,141	0.135	0.341	229,368	0.159	0.366
Ethnic share in Canada	291,141	0.069	0.070	229,368	0.078	0.072
Ethnic share in CMA/CA of arrival	291,141	0.014	0.013586	229,368	0.015	0.014
Atleast a Bachelor's education outside Canada	291,141	0.283	0.451	229,368	0.307	0.461
Formal pre-immigration investment	291,066	0.744	0.436	229,368	0.807	0.395
Informal pre-immigration investment	291,066	0.262	0.440	229,368	0.290	0.454
Age at immigration	291,141	35.307	12.688	229,368	35.395	12.847
Male	291,141	0.494	0.500	229,368	0.491	0.500
Married	291,141	0.760	0.427	229,368	0.765	0.424
Employed	290,294	0.491	0.500	228,715	0.497	0.500
Savings brought from outside Canada	275,625	6.311	4.560	217,260	6.227	4.617
Family class	291,141	0.328	0.470	229,368	0.348	0.476
Number of immigrating members	291,141	2.718	1.541	229,368	2.782	1.549
Number of joining members	291,141	0.626	1.222	229,368	0.671	1.288
Number of household members	291,141	3.833	1.723	229,368	3.981	1.753
Number of children	291,141	0.662	0.902	229,368	0.679	0.908
Months since arrival	291,141	27.554	17.454	229,368	27.558	17.456
RGDP in place of birth	291,141	0.847	0.899	229,368	0.882	0.911
Population in place of birth	291,141	10.113	14.104	229,368	11.459	14.623
Distance from capital city in place of birth	291,141	10080.31	2,751.119	229,368	10645.9	2306.757
Linguistic distance in place of birth	291,141	0.864	0.194	229,368	0.897	0.135

Note 1: Sub-sample is of those whose mother tongue is not English and do not reside in Quebec
Note 2: All variables are weighted using the weights provided with the LSIC by Statistics Canada.

Table A4: OLS estimation results of English speaking proficiency: mother tongue is not English and do not reside in Quebec.

	Pooled	Pooled	Pooled	First differences
RGDP per capita (h)	0.0479***	0.0395**	0.0409**	
	(0.0135)	(0.0140)	(0.0141)	
Informal pre-immigration experience (ω)		0.0875***	0.0871***	
		(0.0116)	(0.0116)	
Formal pre-immigration experience (ω)		0.192***	0.194***	
		(0.0156)	(0.0159)	
Informal investment (ι)			-0.00651	-0.0107
			(0.0137)	(0.0149)
Formal investment (ι)			-0.0940***	0.0513*
			(0.0171)	(0.0227)
Share of co-ethnic group in Canada (m)	0.379	0.0341	0.113	
	(0.257)	(0.250)	(0.251)	
Population level (L)	-0.00719***	-0.00468***	-0.00528***	
	(0.00142)	(0.00140)	(0.00138)	
Geographic distance (μ)	-0.00000244	-0.00000652*	-0.00000793*	
	(0.00000337)	(0.00000330)	(0.00000332)	
Linguistic distance (θ)	0.0297	0.0447	0.0583	
	(0.0561)	(0.0546)	(0.0529)	
Months since arrival	-0.0115*	-0.0101*	-0.00977	
	(0.00509)	(0.00501)	(0.00503)	
Time-varying controls (X_{ist})	✓	✓	✓	✓
Time-invariant controls (X_{is})	✓	✓	✓	✓
Survey wave controls (t)	✓	✓	✓	✓
N	216,672	216,672	204,122	138,783

$^{*}p < 0.05^{**}p < 0.01^{***}p < 0.001$

Note 1: All models are weighted due to Statistics Canada's RDC disclosure process.

Note 2: Cluster robust standard errors in parentheses.

Table A5: Probit estimation results of English speaking proficiency: mother tongue is not English and do not reside in Quebec.

Dependent variable: English speaking proficiency	Pooled
Informal investment (ι) × Share of co-ethnics in CMA/CA of arrival (R)	-8.759***
	(2.245)
Formal investment (ι) × Share of co-ethnics in CMA/CA of arrival (R)	9.226**
	(2.915)
Informal investment (ι)	-0.142**
	(0.0546)
Formal investment (ι)	-0.368***
	(0.0635)
Time-varying controls (X_{ist})	✓
Time-invariant controls (X_{is})	✓
Survey wave controls (t)	✓
N	204,122

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

Note 1: All models are weighted due to Statistics Canada RDC's disclosure process.

Note 2: Cluster robust standard errors in parentheses.

Comparative statics over the thresholds

The following are a set of useful comparative statics used throughout the main body of the text in derivations, proofs and analysis. The effect of θ and m^i on the threshold for assimilating is given by

$$\omega_a^i = 1 - \frac{1}{\theta\nu(1+m^i)}, \quad \frac{\partial \omega_a^i}{\partial \theta} = \frac{1}{\theta^2\nu(1+m^i)} > 0, \quad \frac{\partial \omega_a^i}{\partial m^i} = \frac{1}{\theta\nu(1+m^i)^2} > 0.$$
$$\omega_a^j = 1 - \frac{1}{\theta\nu(1+m^j)}, \quad \frac{\partial \omega_a^j}{\partial \theta} = \frac{1}{\theta^2\nu(1+m^j)} > 0, \quad \frac{\partial \omega_a^j}{\partial m^j} = \frac{1}{\theta\nu(1+m^j)^2} > 0.$$

The function $c(\beta_n)$ is the threshold between choosing to locate in neighbourhood j over i. The following derivatives provide the effect of changes in μ, m^i and m^j.

$$\begin{aligned}
c(\beta_n) &= \frac{m^j/(1+m^j) - (m^i/(1+m^i))(\beta_n/(1-\beta_n))}{2\mu}, \\
\frac{\partial c(\beta_n)}{\partial \mu} &= -\frac{m^j/(1+m^j) - (m^i/(1+m^i))(\beta_n/(1-\beta_n))}{2\mu^2} > 0 \text{ if } \beta > 1/2, \\
\frac{\partial c(\beta_n)}{\partial m^i} &= -\frac{\beta_n}{2\mu(1+m^i)^2(1-\beta_n)} < 0, \\
\frac{\partial c(\beta_n)}{\partial m^j} &= \frac{1}{2\mu(1+m^j)^2} > 0.
\end{aligned}$$

The equations $c_a^i(\omega,\beta)$ and $c_a^j(\omega,\beta)$ are the thresholds for assimilating immigration in neighbourhoods i and j. The following derivatives provide the effect of changes in h, μ and θ.

$$\begin{aligned}
c_a^i(\omega,\beta) &= \frac{1-\theta\nu(1-\omega)-h/\beta}{\mu}, & c_a^j(\omega,\beta) &= \frac{1-\theta\nu(1-\omega)-h/(1-\beta)}{\mu}, \\
\frac{\partial c_a^i(\omega,\beta)}{\partial h} &= -\frac{1}{\mu\beta} < 0, & \frac{\partial c_a^j(\omega,\beta)}{\partial h} &= -\frac{1}{\mu(1-\beta)} < 0, \\
\frac{\partial c_a^i(\omega,\beta)}{\partial \mu} &= -\frac{1-\theta\nu(1-\omega)-h/\beta}{\mu^2} < 0, & \frac{\partial c_a^j(\omega,\beta)}{\partial \mu} &= -\frac{1-\theta\nu(1-\omega)-h/(1-\beta)}{\mu^2} < 0, \\
\frac{\partial c_a^i(\omega,\beta)}{\partial \theta} &= -\frac{\nu(1-\omega)}{\mu} < 0. & \frac{\partial c_a^j(\omega,\beta)}{\partial \theta} &= -\frac{\nu(1-\omega)}{\mu} < 0.
\end{aligned}$$

The equations $c_n^i(\beta)$ and $c_n^j(\beta)$ are the threshold between non-assimilating immigration in neighbourhoods i and j. The following set of derivatives provide the effect of changes in h, μ, m^i and/or m^j.

$$\begin{aligned}
c_n^i(\beta) &= \frac{m^i/(1+m^i)-h/\beta}{\mu}, & c_n^j(\beta) &= \frac{m^j/(1+m^j)-h/(1-\beta)}{\mu} \\
\frac{\partial c_n^i(\beta)}{\partial h} &= -\frac{1}{\mu\beta} < 0, & \frac{\partial c_n^j(\beta)}{\partial h} &= -\frac{1}{\mu(1-\beta)} < 0, \\
\frac{\partial c_n^i(\beta)}{\partial \mu} &= -\frac{m^i/(1+m^i)-h/\beta}{\mu^2} < 0, & \frac{\partial c_n^j(\beta)}{\partial \mu} &= -\frac{m^j/(1+m^j)-h/(1-\beta)}{\mu^2}, \\
\frac{\partial c_n^i(\beta)}{\partial m^i} &= \frac{1}{\mu(1+m^i)^2} > 0, & \frac{\partial c_n^j(\beta)}{\partial m^j} &= \frac{1}{\mu(1+m^j)^2} > 0,
\end{aligned} \tag{4.31}$$

Comparative statics over $\underline{\omega}^i$ and $\underline{\omega}^j$

The equations $\underline{\omega}^i$ and $\underline{\omega}^i$ are defined as the ω at which $c^i_n(\omega)=0$ and $c^j_n(\omega)=0$. The effect of changes in h, θ, m^i and m^j are given as

$$\underline{\omega}^i = 1 - \frac{1-h(1+m^i(1+m^j)/m^j(1+m^i))}{\theta\nu}, \qquad \underline{\omega}^j = 1 - \frac{1-h(1+m^j(1+m^i)/m^i(1+m^j))}{\theta\nu},$$

$$\frac{\partial\underline{\omega}^i}{\partial h} = \frac{1+m^i(1+m^j)/m^j(1+m^i)}{\theta\nu} > 0, \qquad \frac{\partial\underline{\omega}^j}{\partial h} = \frac{1+m^j(1+m^i)/m^i(1+m^j)}{\theta\nu} > 0,$$

$$\frac{\partial\underline{\omega}^i}{\partial m^i} = \frac{h}{\theta\nu}\frac{1+m^j}{m^j(1+m^i)^2} > 0, \qquad \frac{\partial\underline{\omega}^j}{\partial m^i} = -\frac{h}{\theta\nu}\frac{m^j}{(m^i)^2(1+m^j)} < 0,$$

$$\frac{\partial\underline{\omega}^i}{\partial m^j} = -\frac{h}{\theta\nu}\frac{m^i}{(m^j)^2(1+m^i)} < 0, \qquad \frac{\partial\underline{\omega}^j}{\partial m^i} = \frac{h}{\theta\nu}\frac{1+m^i}{m^i(1+m^j)^2} > 0,$$

$$\frac{\partial\underline{\omega}^i}{\partial \theta} = \frac{1-h(1+m^i(1+m^j)/m^j(1+m^i))}{\theta^2\nu} < 0, \qquad \frac{\partial\underline{\omega}^j}{\partial \theta} = \frac{1-h(1+m^j(1+m^i)/m^i(1+m^j))}{\theta^2\nu} < 0,$$

Comparative statics over $\Gamma^i(m^i,m^j)$ and $\Gamma^j(m^i,m^j)$

Consider a mixed equilibrium outcome $k(1-\omega)\leq 1, \omega^i_a \geq \underline{\omega}^i$ and $\omega^j_a > \underline{\omega}^j$. The elements of the Jacobian matrix $J(m^i,m^j)$ are the set of first-order partial differentials of $\Gamma^i(m^i,m^j)$ and $\Gamma^j(m^i,m^j)$ in terms of m^i and m^j. As follows,

$$\begin{aligned}
m^i &= L\int_{\omega^i_a}^1\int_{1/2}^1 F(c^i_a(\omega,\beta))dB(\beta)dW(\omega) \\
&\quad + LW(\omega^i_a)^2\int_{\frac{h(1+m^i)}{m^i}}^1 \max\{0, F(c^i_n(\beta)) - F(c(\beta_n))\}dB(\beta) \equiv \Gamma^i(m^i,m^j) \\
\frac{\partial\Gamma^i(m^i,m^j)}{\partial m^i} &= -L\int_{1/2}^1 F(c^i_a(\omega,\beta))dB(\beta)\frac{\partial\omega^i_a}{\partial m^i} \\
&\quad + 2LW(\omega^i_a)W'(\omega^i_a)\frac{\partial\omega^i_a}{\partial m^i}\int_{\frac{h(1+m^i)}{m^i}}^1 \max\{0, F(c^i_n(\beta)) - F(c(\beta_n))\}dB(\beta) \\
&\quad + LW(\omega^i_a)^2\int_{\frac{h(1+m^i)}{m^i}}^1 \max\{0, f(c^i_n(\beta))\frac{\partial c^i_n(\beta)}{\partial m^i} - f(c(\beta_n))\frac{\partial c(\beta_n)}{\partial m^i}\}dB(\beta) \\
\frac{\partial\Gamma^i(m^i,m^j)}{\partial m^i} &= LW(\omega^i_a)^2\int_{\frac{h(1+m^i)}{m^i}}^1 \max\{0, -f(c(\beta_n))\frac{\partial c(\beta_n)}{\partial m^j}\}dB(\beta) = 0
\end{aligned}$$

$$\begin{aligned}
m^j &= L\int_{\omega^j_a}^1\int_0^{1/2} F(c^j_a(\omega,\beta))dB(\beta)dW(\omega) \\
&\quad + L\left(W(\omega^j_a) - W(\underline{\omega}^j)\right)^2\int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} F(c(\beta_n))dB(\beta) \equiv \Gamma^j(m^i,m^j) \\
\frac{\partial\Gamma^j(m^i,m^j)}{\partial m^i} &= -2L\left(W(\omega^j_a) - W(\underline{\omega}^j)\right)W'(\underline{\omega}^j)\frac{\partial\underline{\omega}^j}{\partial m^i}\int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} F(c(\beta_n))dB(\beta) \\
&\quad + L\left(W(\omega^j_a) - W(\underline{\omega}^j)\right)^2\int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} f(c(\beta_n))\frac{\partial c(\beta_n)}{\partial m^i}dB(\beta) \\
\frac{\partial\Gamma^j(m^i,m^j)}{\partial m^j} &= -L\int_0^{1/2} F(c^j_a(\omega,\beta))dB(\beta)\frac{\partial\omega^j_a}{\partial m^j} \\
&\quad + 2L\left(W(\omega^j_a) - W(\underline{\omega}^j)\right)\left(W'(\omega^j_a)\frac{\partial\omega^j_a}{\partial m^j} - W'(\underline{\omega}^j)\frac{\partial\underline{\omega}^j}{\partial m^j}\right)\int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} F(c(\beta_n))dB(\beta) \\
&\quad + L\left(W(\omega^j_a) - W(\underline{\omega}^j)\right)^2\int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} f(c(\beta_n))\frac{\partial c(\beta_n)}{\partial m^j}dB(\beta)
\end{aligned}$$

Since $\Gamma^i(m^i,m^j)$ and $\Gamma^j(m^i,m^j)$ are also directly affected by the parameters $\{L,h,\mu,\theta\}$. These comparative statics are provided as well in the mixed equilibrium case.

$$\begin{aligned}
\frac{\partial \Gamma^i(m^i,m^j)}{\partial L} &= \int_{\omega_a^i}^1 \int_{1/2}^1 F(c_a^i(\omega,\beta))dB(\beta)dW(\omega) \\
&\quad + W(\omega_a^i)^2 \int_{\frac{h(1+m^i)}{m^i}}^1 \max\{0, F(c_n^i(\beta)) - F(c(\beta_n))\}dB(\beta) > 0 \\
\frac{\partial \Gamma^i(m^i,m^j)}{\partial h} &= L\int_{\omega_a^i}^1 \int_{1/2}^1 f(c_a^i(\omega,\beta))\frac{\partial c_a^i(\omega,\beta)}{\partial h}dB(\beta)dW(\omega) < 0 \\
\frac{\partial \Gamma^i(m^i,m^j)}{\partial \mu} &= L\int_{\omega_a^i}^1 \int_{1/2}^1 f(c_a^i(\omega,\beta))\frac{\partial c_a^i(\omega,\beta)}{\partial \mu}dB(\beta)dW(\omega) < 0 \\
\frac{\partial \Gamma^i(m^i,m^j)}{\partial \theta} &= -L\int_{1/2}^1 F(c_a^i(\omega,\beta))dB(\beta)\frac{\partial \omega_a^i}{\partial \theta} + L\int_{\omega_a^i}^1 \int_{1/2}^1 f(c_a^i(\omega,\beta))\frac{\partial c_a^i(\omega,\beta)}{\partial \theta}dB(\beta)dW(\omega) \\
&\quad + 2LW(\omega_a^i)^2 W'(\omega_a^i)\frac{\partial \omega_a^i}{\partial \theta}\int_{\frac{h(1+m^i)}{m^i}}^1 \max\{0, F(c_n^i(\beta)) - F(c(\beta_n))\}dB(\beta) \lesseqgtr 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \Gamma^j(m^i,m^j)}{\partial L} &= \int_{\omega_a^j}^1 \int_0^{1/2} F(c_a^j(\omega,\beta))dB(\beta)dW(\omega) \\
&\quad + \left(W(\omega_a^j) - W(\underline{\omega}^j)\right)^2 \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} F(c(\beta_n))dB(\beta) > 0 \\
\frac{\partial \Gamma^j(m^i,m^j)}{\partial h} &= L\int_{\omega_a^j}^1 \int_0^{1/2} f(c_a^j(\omega,\beta))\frac{\partial c_a^j(\omega,\beta)}{\partial h}dB(\beta)dW(\omega) \\
&\quad - 2L\left(W(\omega_a^j) - W(\underline{\omega}^j)\right)W'(\underline{\omega}^j)\frac{\partial \underline{\omega}^j}{\partial h}\int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} F(c(\beta_n))dB(\beta) < 0 \\
\frac{\partial \Gamma^j(m^i,m^j)}{\partial \mu} &= L\int_{\omega_a^j}^1 \int_0^{1/2} f(c_a^j(\omega,\beta))\frac{\partial c_a^j(\omega,\beta)}{\partial \mu}dB(\beta)dW(\omega) \\
&\quad + L\left(W(\omega_a^j) - W(\underline{\omega}^j)\right)^2 \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} f(c(\beta_n))\frac{\partial c(\beta_n)}{\partial \mu}dB(\beta) < 0 \\
\frac{\partial \Gamma^j(m^i,m^j)}{\partial \theta} &= -L\int_0^{1/2} F(c_a^j(\omega,\beta))dB(\beta)\frac{\partial \omega_a^j}{\partial \theta} + L\int_{\omega_a^j}^1 \int_0^{1/2} f(c_a^j(\omega,\beta))\frac{\partial c_a^j(\omega,\beta)}{\partial \theta}dB(\beta)dW(\omega) \\
&\quad + 2L\left(W(\omega_a^j) - W(\underline{\omega}^j)\right)\left(W'(\omega_a^j)\frac{\partial \omega_a^j}{\partial \theta} - W'(\underline{\omega}^j)\frac{\partial \underline{\omega}^j}{\partial \theta}\right)\int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} F(c(\beta_n))dB(\beta) \lesseqgtr 0
\end{aligned}$$

In a sorting equilibrium outcome $k(1-\omega) \leq 1, \omega_a^i \geq \underline{\omega}^i$ and $\omega_a^j = \underline{\omega}^j$. The first-order differentials of $\Gamma^i(m^i,m^j)$ in terms of m^i and m^j, and $\{L, h, \mu, \theta\}$ are the same as in the mixed case. But the set of first-order differentials of $\Gamma^j(m^i,m^j)$ are as follows,

$$\begin{aligned}
\frac{\partial \Gamma^j(m^i,m^j)}{\partial m^i} &= -L\int_0^{1/2} F(c_a^j(\omega,\beta))dB(\beta)\frac{\partial \underline{\omega}^j}{\partial m^i} < 0 \\
\frac{\partial \Gamma^j(m^i,m^j)}{\partial m^j} &= -L\int_0^{1/2} F(c_a^j(\omega,\beta))dB(\beta)\frac{\partial \underline{\omega}^j}{\partial m^j} < 0 \\
\frac{\partial \Gamma^j(m^i,m^j)}{\partial L} &= \int_{\underline{\omega}^j}^1 \int_0^{1/2} F(c_a^j(\omega,\beta))dB(\beta)dW(\omega) > 0 \\
\frac{\partial \Gamma^j(m^i,m^j)}{\partial h} &= -L\int_0^{1/2} F(c_a^j(\omega,\beta))dB(\beta)\frac{\partial \underline{\omega}^j}{\partial h} + L\int_{\underline{\omega}^j}^1 \int_0^{1/2} f(c_a^j(\omega,\beta))\frac{\partial c_a^j(\omega,\beta)}{\partial h}dB(\beta)dW(\omega) < 0 \\
\frac{\partial \Gamma^j(m^i,m^j)}{\partial \mu} &= L\int_{\underline{\omega}^j}^1 \int_0^{1/2} f(c_a^j(\omega,\beta))\frac{\partial c_a^j(\omega,\beta)}{\partial \mu}dB(\beta)dW(\omega) < 0 \\
\frac{\partial \Gamma^j(m^i,m^j)}{\partial \theta} &= -L\int_0^{1/2} F(c_a^j(\omega,\beta))dB(\beta)\frac{\partial \underline{\omega}^j}{\partial \theta} + L\int_{\underline{\omega}^j}^1 \int_0^{1/2} f(c_a^j(\omega,\beta))\frac{\partial c_a^j(\omega,\beta)}{\partial \theta}dB(\beta)dW(\omega) \lesseqgtr 0
\end{aligned}$$

Comparative statics over $m^i(L,h,\mu,\theta)$ and $m^j(L,h,\mu,\theta)$

Since $m^i = \Gamma^i(m^i, m^j; \mathbb{P})$ and $m^j = \Gamma^j(m^i, m^j; \mathbb{P})$ implicitly define m^i and m^j, where $\mathbb{P} \equiv \{L, h, \mu, \theta\}$. Then totally differentiating with respect to the parameters in $\mathbb{P}$ is given by

$$\begin{aligned}\frac{dm^i}{d\mathbb{P}} &= \frac{\partial\Gamma^i}{\partial m^i}\frac{\partial m^i}{\partial\mathbb{P}} + \frac{\partial\Gamma^i}{\partial m^j}\frac{\partial m^j}{\partial\mathbb{P}} + \frac{\partial\Gamma^i}{\partial\mathbb{P}},\\ \frac{dm^j}{d\mathbb{P}} &= \frac{\partial\Gamma^j}{\partial m^i}\frac{\partial m^i}{\partial\mathbb{P}} + \frac{\partial\Gamma^j}{\partial m^j}\frac{\partial m^j}{\partial\mathbb{P}} + \frac{\partial\Gamma^j}{\partial\mathbb{P}}.\end{aligned}$$

Solving the two equation systems simultaneously provides the following set of reduced form solutions to the differentials of m^i and m^j in terms of $\mathbb{P}$.

$$\begin{aligned}\frac{dm^i}{d\mathbb{P}} &= \frac{\partial\Gamma^i/\partial m^j \times \partial\Gamma^j/\partial\mathbb{P} + \partial\Gamma^i/\partial\mathbb{P} \times (1 - \partial\Gamma^j/\partial m^j)}{(1-\partial\Gamma^i/\partial m^i)(1-\partial\Gamma^j/\partial m^j) - \partial\Gamma^i/\partial m^j \times \partial\Gamma^j/\partial m^i},\\ \frac{dm^j}{d\mathbb{P}} &= \frac{\partial\Gamma^j/\partial m^i \times dm^i/d\mathbb{P} + \partial\Gamma^j/\partial\mathbb{P}}{1-\partial\Gamma^j/\partial m^j}.\end{aligned}$$

The signs of the partial differentials, $\partial\Gamma^i/\partial\mathbb{P}$ and $\partial\Gamma^j/\partial\mathbb{P}$, have already been determined in the previous section. Additionally, $\partial\Gamma^i/\partial m^j = 0$ which simplifies the problem further. Moreover, the stability conditions require that $\partial\Gamma^i/\partial m^i < 0$ and $\partial\Gamma^j/\partial m^j < 0$. Finally, $\partial\Gamma^j/\partial m^i > 0$ is also implied by the stability conditions. Using this information the effect of $\mathbb{P}$ on m^i and m^j in the mixed and sorting cases are given as

$$\begin{aligned}&\frac{dm^i}{dL} > 0, \quad \frac{dm^i}{dh} < 0, \quad \frac{dm^i}{d\mu} < 0, \quad \frac{dm^i}{d\theta} \lesseqqgtr 0,\\ &\frac{dm^j}{dL} > 0, \quad \frac{dm^j}{dh} < 0, \quad \frac{dm^j}{d\mu} < 0, \quad \frac{dm^j}{d\theta} \lesseqqgtr 0.\end{aligned}$$

Direct effects

The direct effects of table 2 are determined by changes in $\mathbb{P} = \{L, h, \mu, \theta\}$ on $\rho^j = a^j/m^j$. To simplify the calculations, let $\eta \equiv \mathbb{A}/\mathbb{B}$, where

$$\begin{aligned}\mathbb{A} &= L\int_{\omega_a^j}^{1}\int_0^{1/2} F(c_a^j(\omega,\beta))dB(\beta)dW(\omega) > 0\\ \mathbb{B} &= L\left(W(\omega_a^j) - W(\underline{\omega}^j)\right)\int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^im^j}} F(c(\beta_n))dB(\beta) > 0\end{aligned}$$

Now ρ^j can be written in a simpler form, as a function of η. It can then be inferred that ρ^j is a monotonic increasing function of η.

$$\rho^j = \frac{\eta + \int_{\underline{\omega}^j}^{\omega_a^j} k(1-\omega)dW(\omega)}{\eta + \left(W(\omega_a^j) - W(\underline{\omega}^j)\right)}, \qquad \frac{\partial\rho^j}{\partial\eta} > 0 \quad \forall\, k(1-\omega) \le 1$$

The parameters in the model $\mathbb{P} = \{L, h, \mu, \theta\}$ are expected to have a direct effect on $\rho^j = a^j/m^j$. These effects are estimated by the relationship

$$\frac{\partial\rho^j}{\partial\mathbb{P}} = \frac{\partial\rho^j}{\partial\eta} \times \frac{\partial\eta}{\partial\mathbb{P}}, \qquad \text{where} \quad \frac{\partial\eta}{\partial\mathbb{P}} = \frac{\partial\mathbb{A}/\partial\mathbb{P} \times \mathbb{B} - \partial\mathbb{B}/\partial\mathbb{P} \times \mathbb{A}}{\mathbb{B}^2}.$$

All that remains to sign the direct effect is differentiating $\mathbb{A}$ and $\mathbb{B}$ in terms of the elements in $\mathbb{P} = \{L, h, \mu, \theta\}$.

$$\begin{aligned}
\frac{\partial \mathbb{A}}{\partial L} &= \int_{\omega_a^j}^{1} \int_0^{1/2} F(c_a^j(\omega,\beta)) dB(\beta) dW(\omega) > 0, \\
\frac{\partial \mathbb{A}}{\partial h} &= L \int_{\omega_a^j}^{1} \int_0^{1/2} f(c_a^j(\omega,\beta)) \frac{\partial c_a^j(\omega,\beta)}{\partial h} dB(\beta) dW(\omega) < 0, \\
\frac{\partial \mathbb{A}}{\partial \mu} &= L \int_{\omega_a^j}^{1} \int_0^{1/2} f(c_a^j(\omega,\beta)) \frac{\partial c_a^j(\omega,\beta)}{\partial \mu} dB(\beta) dW(\omega) < 0, \\
\frac{\partial \mathbb{A}}{\partial \theta} &= -L \int_0^{1/2} F(c_a^j(\omega,\beta)) dB(\beta) \frac{\partial \omega_a^j}{\partial \theta} dW(\omega) + L \int_{\omega_a^j}^{1} \int_0^{1/2} f(c_a^j(\omega,\beta)) \frac{\partial c_a^j(\omega,\beta)}{\partial \theta} dB(\beta) dW(\omega) < 0, \\
\frac{\partial \mathbb{A}}{\partial m^i} &= 0, \\
\frac{\partial \mathbb{A}}{\partial m^j} &= -L \int_0^{1/2} F(c_a^j(\omega,\beta)) \frac{\partial \omega_a^j}{\partial m^j} dB(\beta) dW(\omega) < 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathbb{B}}{\partial L} &= \left(W(\omega_a^j) - W(\underline{\omega}^j)\right) \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^i m^j}} F(c(\beta_n)) dB(\beta) > 0, \\
\frac{\partial \mathbb{B}}{\partial h} &= -L W'(\underline{\omega}^j) \frac{\partial \underline{\omega}^j}{\partial h} \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^i m^j}} F(c(\beta_n)) dB(\beta) < 0, \\
\frac{\partial \mathbb{B}}{\partial \mu} &= L \left(W(\omega_a^j) - W(\underline{\omega}^j)\right) \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^i m^j}} f(c(\beta_n)) \frac{\partial c(\beta_n)}{\partial \mu} dB(\beta) > 0, \\
\frac{\partial \mathbb{B}}{\partial \theta} &= L \left(W'(\omega_a^j) \frac{\partial \omega_a^j}{\partial \theta} - W'(\underline{\omega}^j) \frac{\partial \underline{\omega}^j}{\partial \theta} \right) \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^i m^j}} F(c(\beta_n)) dB(\beta) \lesseqqgtr 0, \\
\frac{\partial \mathbb{B}}{\partial m^i} &= -L W'(\underline{\omega}^j) \frac{\partial \underline{\omega}^j}{\partial m^i} \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^i m^j}} F(c(\beta_n)) dB(\beta) \\
&\quad + L \left(W(\omega_a^j) - W(\underline{\omega}^j)\right) \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^i m^j}} f(c(\beta_n)) \frac{\partial c(\beta_n)}{\partial m^i} dB(\beta) \lesseqqgtr 0, \\
\frac{\partial \mathbb{B}}{\partial m^j} &= L \left(W'(\omega_a^j) \frac{\partial \omega_a^j}{\partial m^j} - W'(\underline{\omega}^j) \frac{\partial \underline{\omega}^j}{\partial m^j} \right) \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^i m^j}} F(c(\beta_n)) dB(\beta) \\
&\quad + L \left(W(\omega_a^j) - W(\underline{\omega}^j)\right) \int_0^{\frac{m^j(1+m^i)}{m^i+m^j+2m^i m^j}} f(c(\beta_n)) \frac{\partial c(\beta_n)}{\partial m^j} dB(\beta) \lesseqqgtr 0.
\end{aligned}$$

Indirect effects

The indirect effects of table 2 are determined by changes in $\mathbb{P} = \{L, h, \mu, \theta\}$ on $\rho^j = a^j/m^j$ through changes in $m^i(L, h, \mu, \theta)$. These effects are estimated by the relationship

$$\frac{\partial \rho^j}{\partial \mathbb{P}} = \frac{\partial \rho^j}{\partial \eta} \times \frac{\partial \eta}{\partial m^i} \times \frac{\partial m^i}{\partial \mathbb{P}}, \qquad \text{where} \quad \frac{\partial \eta}{\partial m^i} = \frac{\partial \mathbb{A}/\partial m^i \times \mathbb{B} - \partial \mathbb{B}/\partial m^i \times \mathbb{A}}{\mathbb{B}^2}.$$

But since $\partial \mathbb{B}/\partial m^i$ was ambiguously signed, all the indirect effects through $m^i(L, h, \mu, \theta)$ are ambiguous as well. Similarly, I can estimate the indirect effects through $m^j(L, h, \mu, \theta)$ as

$$\frac{\partial \rho^j}{\partial \mathbb{P}} = \frac{\partial \rho^j}{\partial \eta} \times \frac{\partial \eta}{\partial m^j} \times \frac{\partial m^j}{\partial \mathbb{P}}, \qquad \text{where} \quad \frac{\partial \eta}{\partial m^j} = \frac{\partial \mathbb{A}/\partial m^j \times \mathbb{B} - \partial \mathbb{B}/\partial m^j \times \mathbb{A}}{\mathbb{B}^2}.$$

Again, $\partial \mathbb{B}/\partial m^j$ is ambiguously signed, all the indirect effects through $m^j(L, h, \mu, \theta)$ are ambiguous. The results are depicted as such in table 2.

Threshold effects

The threshold effects are identified by changes in $\mathbb{P} = \{L, h, \mu, \theta\}$ on equation (4.25), or $\phi(.)$. These effects are provided below.

$$\frac{d\phi}{dL} < 0, \qquad \frac{d\phi}{dh} > 0, \qquad \frac{d\phi}{d\mu} > 0, \qquad \frac{d\phi}{d\theta} \lesseqgtr 0.$$

Tables

Table A1: Descriptions of variables used in this study.

Variable Name	**Description**
Speaking proficiency	English speaking proficiency of the respondent (5-level categorical)
Ethnic enclave	Arrived in a CMA/CA that is overrepresented by co-ethnics (Census of Population 2001), and worked in an organization that is mostly co-ethnic (Dummy)
Housing costs in CMA/CA of arrival (in $'000's)	Average cost of housing in CMA/CA of arrival. (Continuous; Census of Population 2001)
Formal investment	Learned to speak English in language class or school (Dummy)
Informal investment	Learned to speak English through family & friends, work, non-language classes, media, self-study and/or other (Dummy)
CMA of arrival: Toronto	The respondent lives in Toronto at the time of landing in Canada (Dummy)
CMA of arrival: Vancouver	The respondent lived in Vancouver at the time of landing in Canada (Dummy)
Share of co-ethnic group in Canada	Share of the co-ethnic population that the respondent belongs to admitted into Canada in 2000 (Continuous)
Education outside Canada: Bachelor's or higher	Highest level of formal education attained outside of Canada is Bachelor's or higher (Dummy)
Formal learning of English outside of Canada	Learned most of their English before coming to Canada through language classes, a private tutor or schooling (Dummy)
Informal learning of English outside of Canada	Learned most of their English before coming to Canada from family/friends, self-study, work, media, everyday interactions or other (Dummy)
Friend/family networks	Respondent had friends/family in current area (or nearby area) and chose this area because friends/family live here, in wave 1 (Dummy)
Age at immigration	Age of the respondent (Continuous)
Male	Gender of the respondent = Male (Dummy)
Married	Marital status of the respondent = Married (Dummy)
Employed	The respondent is employed full-time (Dummy)
Savings brought from outside Canada	Total amount of savings brough from outside Canada (Continuous)
Family class	Immigration category = Family class (Dummy)
Number of immigrating members	Size of the immigrating party at the time of landing (Continuous)

Table A1: Descriptions of variables used in this study.

Variable Name	**Description**
Number of joining members	Number of members that have joined the household since landing (wave 1) or since the last interview (waves 2 and 3) (Continuous)
Number of household members	Size of household (Continuous)
Number of children	Percentage of children living in the household (Continuous)
Months since arrival	Number of months elapsed since landing in Canada (Continuous)
RGDP in place of Birth	Real GDP of place of birth of the respondent relative to Canada (Continuous; Heston, Summers & Aten, 2005)
Population in place of birth	Population of place of birth of the respondent relative to Canada (Continuous; Heston, Summers & Aten, 2005)
Distance from capital city in place of birth	Distance of the capital city in the place of birth of the respondent relative to Canada (Continuous; Mayer & Zignago, 2011)
Linguistic distance	Linguistic distance of the official language in the place of birth of the respondent relative to English (Continuous; Melitz & Toubal, 2014)

Table A2: Summary statistics.

	Full sample			Sub-sample		
Variable Name	Count	Mean	SD	Count	Mean	SD
Speaking proficiency	291,135	0.692	0.462	229,362	0.700	0.458
Ethnic enclave	290,877	0.123	0.329	229,104	0.136	0.342
Formal investment	276,555	0.314	0.464	216,129	0.349	0.477
Informal investment	290,853	0.503	0.500	229,080	0.554	0.497
CMA of arrival: Toronto	291,141	0.484	0.500	229,368	0.554	0.497
CMA of arrival: Vancouver	291,141	0.135	0.341	229,368	0.159	0.366
Housing costs in CMA/CA of arrival	291,141	13.111	2.207	229,368	13.630	1.848
Ethnic share in Canada	291,141	0.069	0.070	229,368	0.078	0.072
Atleast a Bachelor's education outside Canada	291,141	0.283	0.451	229,368	0.307	0.461
Formal pre-immigration investment	291,066	0.744	0.436	229,368	0.807	0.395
Informal pre-immigration investment	291,066	0.262	0.440	229,368	0.290	0.454
Family/friend networks	291,141	0.818	0.386	229,368	0.832	0.374
Age at immigration	291,141	35.307	12.688	229,368	35.395	12.847
Male	291,141	0.494	0.500	229,368	0.491	0.500
Married	291,141	0.760	0.427	229,368	0.765	0.424
Employed	290,294	0.491	0.500	228,715	0.497	0.500
Savings brought from outside Canada	275,625	6.311	4.560	217,260	6.227	4.617
Family class	291,141	0.328	0.470	229,368	0.348	0.476

Table A2: Summary statistics.

	Full sample			Sub-sample		
Variable Name	Count	Mean	SD	Count	Mean	SD
Number of immigrating members	291,141	2.718	1.541	229,368	2.782	1.549
Number of joining members	291,141	0.626	1.222	229,368	0.671	1.288
Number of household members	291,141	3.833	1.723	229,368	3.981	1.753
Number of children	291,141	0.662	0.902	229,368	0.679	0.908
Months since arrival	291,141	27.554	17.454	229,368	27.558	17.456
RGDP in place of birth	291,141	0.847	0.899	229,368	0.882	0.911
Population in place of birth	291,141	10.113	14.104	229,368	11.459	14.623
Distance from capital city in place of birth	291,141	10080.31	2,751.119	229,368	10645.9	2306.757
Linguistic distance in place of birth	291,141	0.864	0.194	229,368	0.897	0.135

Note 1: Sub-sample of 18-64 year old from non-English speaking households whose mother tongue is not English, and do not reside in Quebec
Note 2: All variables are weighted using the weights provided with the LSIC by Statistics Canada.

Table A3: OLS estimation results of English speaking proficiency: households whose mother tongue is not English and do not reside in Quebec.

Dependent variable: English speaking proficiency	Pooled	Pooled	Pooled, sub-sample	Pooled	Pooled IV
RGDP per capita (h)	0.0407**	0.0399**	0.0395	0.0424**	0.0678**
	(0.0141)	(0.0141)	(0.0202)	(0.0142)	(0.0231)
Network (β)		-0.0227	-0.00591	-0.0199	0.0124
		(0.0131)	(0.0218)	(0.0130)	(0.0236)
Ethnic enclave (ζ)				-0.0984***	-1.033*
				(0.0161)	(0.406)
Informal investment (ι)	-0.0465***	-0.0459***	-0.0416*	-0.0447***	-0.0321*
	(0.0102)	(0.0102)	(0.0172)	(0.0101)	(0.0151)
Formal investment (ι)	-0.0648***	-0.0652***	-0.0875***	-0.0666***	-0.0834***
	(0.0131)	(0.0130)	(0.0205)	(0.0130)	(0.0191)
Formal pre-immigration	0.195***	0.195***	0.104***	0.194***	0.181***
experience (ω)	(0.0158)	(0.0158)	(0.0241)	(0.0158)	(0.0223)
Informal pre-immigration	0.0881***	0.0876***	0.0645***	0.0851***	0.0652***
experience (ω)	(0.0116)	(0.0116)	(0.0188)	(0.0116)	(0.0196)
Share of co-ethnic group in	-0.0390	-0.0322	-2.782***	0.0901	1.210
Canada (m)	(0.247)	(0.246)	(0.612)	(0.246)	(0.628)
Population level (L)	-0.00495***	-0.00496***	0.00620	-0.00529***	-0.00831**
	(0.00138)	(0.00138)	(0.00638)	(0.00138)	(0.00255)
Geographic distance (μ)	-0.00000820*	-0.00000800*	0.000000995	-0.00000760*	-0.00000377
	(0.00000328)	(0.00000329)	(0.00000434)	(0.00000327)	(0.00000443)
Linguistic distance (θ)	0.0519	0.0505	-0.350***	0.0515	0.0624
	(0.0530)	(0.0530)	(0.104)	(0.0528)	(0.0672)
Months since arrival	-0.0103*	-0.0103*	-0.0143	-0.0103*	-0.0121
	(0.00504)	(0.00504)	(0.00850)	(0.00503)	(0.00673)
Time-varying controls (X_{ist})	✓	✓	✓	✓	✓
Time-invariant controls (X_{is})	✓	✓	✓	✓	✓
Survey wave controls (t)	✓	✓	✓	✓	✓
N	204,122	204,122	78,405	203,858	203,858

$^{*}p < 0.05^{**}p < 0.01^{***}p < 0.001$

Note 1: All models are weighted due to Statistics Canada's RDC disclosure process.

Note 2: Cluster robust standard errors in parentheses.

Note 3: The sub-sample excludes new immigrants from source countries that scored "medium" or "high" on the EF EPI (2017).