The Impact of Immigrant Language Skills on Canadian Wages

by

Seda Gunduz

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for the Doctorate in Philosophy Degree

in the
Department of Economics
Faculty of Graduate and Postdoctoral Studies

Abstract

This thesis consists of three chapters investigating the impact of immigrant language skills on Canadian wages. The first chapter, "Linguistic diversity among Canadian immigrants: 1981-2006", describes the changes in linguistic diversity among Canadian immigrants, as measured by a preferred linguistic distance measure, the Levenshtein Distance (LD) Index, and documents socio-demographic characteristics of recent immigrants as well as their labour market performance based on their language capital at the time of entry. The LD is an approximation of immigrants' language skills in the Canadian official languages and represents the "distance" of an immigrant's reported language to the Canadian official languages. Using the 20% micro-data files of the Canadian Censuses between 1981 and 2006, I assign each immigrant an index number based on two language measures: mother tongue and home language. French and English are defined as the Canadian official languages in Quebec and outside of Quebec, respectively. The main findings suggest that although immigrants' mother tongues became more "distant" to the Canadian official languages in both regions over time, the language skill of an average immigrant based on home language remained almost the same in Quebec, in particular, between 1981 and 1996. In terms of immigrants' socio-demographic characteristics and their labour market performance, general patterns were similar across the two regions, although there were significant differences by language groups. In particular, the change in immigrants' wages by language groups is suggestive of the role of language skills in determining wages.

The second chapter, "Immigrant versus native men? Substitutability and the role of linguistic diversity in Canada", estimates the degree of substitutability between immigrant and native men by incorporating immigrants' language skills into the analysis

and calculates the potential wage effects of immigration on Canadian wages. Using the 20% micro-data files of the Canadian Censuses between 1981 and 2006 and imposing a nested-CES production function on the Canadian economy, I estimate immigrant-native substitutability based on immigrants' language skills in addition to education levels and years of labour market experience. I use the LD Index to represent immigrants' language skills by the distance of the mother tongue and home language of an immigrant to English outside of Quebec and to French in Quebec. I define three language groups for immigrants as the high language-skilled, the medium language-skilled, and the low language-skilled. The key findings are as follows. First, home language-based estimates suggest imperfect substitutability in Canada outside of Quebec in some cases. Second, by language skill groups, the low language-skilled immigrants are more likely to be imperfect substitutes for the Canadian-born. Third, the findings for Quebec are substantially different from those for Canada outside of Quebec. My simulations suggest that the long-run effect of immigration on immigrants' wages was negative between 1981 and 2006 while the long-run effect of immigration on the wages of the Canadian-born was small but positive over the same period.

The third chapter, "Gender, linguistic diversity, and labour market substitutability", uses the same methodology and data sources as in the second chapter to incorporate female workers into the analysis of immigrant-native substitutability. This study estimates the elasticity of substitution between immigrant language groups and natives for female workers and the pooled sample of male and female workers. The findings suggest that the degree of substitutability between female immigrants and female natives is similar to the degree of substitutability between male immigrants and male natives. The main results do not change for the pooled sample. Due to potential differences between language accumulation processes between female and male immigrants, the third chapter also estimates female-male immigrants substitutability based on language skills, education levels, and years of labour market experience. The findings suggest that female and male immigrants are imperfect substitutes outside of Quebec regardless of language measures.

Dedication

This thesis is dedicated to my mother and my father, Muzeyyen and Dogan Gunduz, for their unconditional love and support and their constant encouragement; to my grandmother and my grandfather, Hacer and Nihat Arda, whom I could not say my last goodbyes and thank for the sweetest memories of my childhood; and to my best friend, Oya Uzum, for always reminding me that even distances separate and years pass, the real and true one stays real and true.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to convey my special thanks to my supervisors, Dr. Catherine Deri Armstrong and Dr. Miles Corak, whose stimulating ideas, guidance, and continuous support helped me complete this thesis. I am deeply indebted to their guidance, encouragement and inspiration. I would also like to thank my committee members, Dr. Ana Ferrer, Dr. Christopher Worswick, Dr. Gilles Grenier, Dr. Pierre Brochu, for serving as my committee members. They have helped me to be a better researcher with their invaluable comments. I would also like to thank Sami Alpanda and Sermin Gungor for their great comments and support during my studies.

Introduction to the thesis

Immigration in Canada has seen some major changes over the last decades; source countries have shifted in important ways and immigrant inflows have more than tripled. This thesis consists of three essays that examine the change in linguistic diversity among Canadian immigrants and the impact of increased immigrant inflows on Canadian wages considering diversified language skills.

The first chapter has three objectives. First, it considers the advantages and drawbacks of existing indices for measuring linguistic diversity, all approximations for the challenges in learning a second language, and choses a preferred linguistic distance index, the Levenshtein Distance (LD). Second, it documents the changes in linguistic diversity among Canadian immigrant cohorts between 1981 and 2006 by using the LD. Third, it considers sociodemographic characteristics and the labour market performance of newcomers by the language skill groups that are constructed based on the LD of mother tongue at the time of entry. The LD index represents the "distance" between any two languages, which suggests a level of difficulty in learning the second language. Using the 20% micro-data files of the Canadian Censuses between 1981 and 2006, linguistic diversity is calculated based on two language measures: mother tongue, a language learned in early childhood and still understood at the time of the census, and home language, the language spoken most often at home. Each immigrant is assigned an LD index based on their reported language, representing the "distance" to English in Canada outside of Quebec and to French in Quebec. This chapter accounts for important differences between Quebec and the rest of Canada, including the linguistic characteristics of Quebec, Quebec's extensive responsibility in immigrant selection process since the early 1990s, and Quebec's labour market. The main

contribution of this chapter is to use a fairly new measure to describe linguistic diversity among Canadian immigrants between 1981 and 2006. There are three key findings for this chapter that are summarized as follows. First, the average LD based on mother tongue increased in both regions over time. Nonetheless, the average LD based on home language increased outside of Quebec whereas it remained almost the same in Quebec between 1981 and 1996, suggesting that an average immigrant's language skills in French did not vary although mother tongues shifted towards more "distant" languages. Second, the shifts in birth countries differed in two regions. In Canada outside of Quebec, English-speaking countries lost their ranks among the top five most common birth countries over time while French-speaking countries were still among the top five during the 2000s in Quebec. Third, while sociodemographic characteristics of recent immigrants in general were similar between Quebec and the other provinces, significant differences emerged by language skill groups. In addition, wages vary across language groups, suggesting that language skills play a role in determining wages.

The second chapter analyzes the substitutability between immigrant and native men by incorporating immigrant language skills and considers the impact of immigration on Canadian wages. I build my model on Borjas (2003) and Ottaviano and Peri (2008, 2012) and highlight the Canadian experience using the model. Economic theory predicts that immigration-induced labour supply shifts would reduce wages given that immigrants and natives are perfect substitutes. However, empirical studies in general do not find a strong evidence to support this prediction. Borjas (2003) criticizes the methodology in these empirical studies and emphasizes the significance of market equalizing forces across local labour markets while analyzing the impact of immigration on wages. He suggests that the impact of immigration should be examined by using variations across different skill types at the national level. By imposing a Constant Elasticity of Substitution (CES) function on the aggregate economy and assuming a perfect substitutability between immigrants and natives, his findings indicate that the immigration between 1980 and 2000 reduced overall wages by 3.2% in the United States. Ottaviano and Peri (2008, 2012) challenge the main assumption of this model by considering the factors such as the lack of knowledge

in labour market institutions and practices, networking, and language skills that are likely to contribute to immigrant-native substitutability. They find that immigrants and natives are imperfectly substitutable, which yields a result that immigrant wages are more severely affected by immigration than wages of native-born workers. The main contribution of Chapter 2 is to examine immigrant-native substitutability related to language skills and to discuss potential wage effects of immigration in the Canadian labour market. Two linguistic distance measures based on mother tongue and home language are derived from the 20% micro-data files of the Canadian Censuses conducted every five years between 1981 and 2006. Three different language skill groups are constructed: high, medium, and low. These are based on the linguistic distance of both mother tongue and home language to English outside of Quebec. In Quebec, linguistic distance is defined as the distance to French. My notable findings are summarized as follows. First, most estimates from home language-based categorizations point to perfect immigrant-native substitutability for Canada and Quebec. In Canada outside of Quebec, empirical evidence for imperfect substitutability exists in some cases. Second, examining language groups separately, the findings suggest imperfect substitutability between low language-skilled immigrants and the Canadian-born in Canada outside of Quebec. For example, I estimate the elasticity of substitution to be about 20 between the Canadian-born and the low language-skilled immigrants. Third, the findings for the substitutability between immigrant language groups and natives for Quebec substantially differ from those for the rest of Canada. In Quebec, even low language-skilled immigrants seem perfectly substitutable with the Canadian-born. On the basis of my simulations, immigrant wages are more negatively affected by immigration as compared to Canadian-born. Immigrant wages declined by 6.9% whereas the wages of the Canadian-born declined by 4.2% due to immigration between 1981 and 2006 in the short-run. In particular, the most educated immigrants with the lowest language skills in English are most severely affected.

The third chapter expands on Chapter 2 by focusing on women in particular. In the literature, female workers are generally not included in the analysis due to discontinuities in their labour market participation decisions. However, this chapter includes female workers

in the analysis by using the same data sources as in Chapter 2 and compares the estimates of the elasticity of substitution between immigrant language groups and natives for female, male, and the pooled sample of workers as well. As in Chapter 2, three language skill groups are defined for immigrants: the high, the medium, and the low language-skilled in English for Canada outside of Quebec and in French for Quebec, respectively. Chapter 3 finds that substitutability between female immigrants and female natives is similar to substitutability between male immigrants and male natives. The pooled sample of female and male workers does not change the main conclusions. As an extension, Chapter 3 estimates the degree of substitutability between female and male immigrants due to potential differences in language acquisition between female and male immigrants. The results suggest that the elasticity of substitution between female and male immigrants is around 10 on average. By using the estimates for separate language groups, it is simulated that the relative wages of female-male immigrants have reduced by around 5% between 1981 and 2006. There is no empirical evidence to suggest a degree of substitutability between female and male immigrants in Quebec.

Language skills play an essential role on wages, but language skills per se might not be responsible for imperfect substitutability between immigrants and natives. Other factors such as the lack of knowledge in labour market institutions and practices and networking within ethnic communities might also contribute to immigrant-native substitutability. In case of imperfect substitutability, simulations suggest that the most severely affected group is the most educated with low levels of language skills. Furthermore, female immigrants are more negatively affected as compared to male immigrants. On this basis, understanding the reasons for imperfect substitutability between immigrants and natives might help policy makers design ways in which they can help immigrants transfer their skills into the labour market in a more efficient way and can combat declining economic performances among recent immigrant cohorts.

Chapter 1

Linguistic Diversity among Canadian Immigrants: 1981-2006

1.1 Introduction

The tremendous shift in the birth countries of immigrants to Canada since the 1970s is well-documented. Traditional major sources such as the United States and Europe have been predominantly replaced by the countries in the Middle East and Asia (Beach, Green, and Worswick, 2011). This change has been significantly influenced by immigration policies at both the federal and the provincial levels. The abandonment of the nationality-based immigrant selection in 1962, the introduction of the point-based selection process in 1967, and the inclusion of refugees as a distinct class of immigrants in The Immigration Act of 1976 are key examples.[1] Federal-provincial agreements have also allowed provinces to address their unique linguistic and economic needs.[2]

[1]The idea of the point-based system is to assess potential economic class migrants objectively based on their observable characteristics including education, years of labour market experience, age, and language skills. Economic immigrants are defined as individuals who are selected based on their skills and ability to contribute to Canada (Citizenship and Immigration Canada, 2015). This federal point system has continuously evolved. For example, the Immigration and Refugee Protection Act enacted in 2001 changed the weighting scheme so as to attribute greater weight to applicants' education, language skills, and adaptability.

[2]The first among these agreements was the 1991 Canada-Quebec Accord, which gave full responsibility to Quebec in selecting both economic immigrants and refugees (Challinor, 2005). The federal government remained in charge of finalizing the admission process and setting up immigration categories and immigration levels proportional to the population of Quebec (Nijboer, 2010). During the 2000s, the agreements covering a wide range of matters related to migration were signed between the federal government and the following

While much attention has been given to the shifts in the birth countries of immigrants at the national level, much less attention has been given to the changes in the languages spoken by immigrants. Striking changes can be found both across immigrant cohorts over time and within cohorts across provinces. For example, the share of immigrants speaking a non-official language as mother tongue more than quadrupled between 1981 and 2006, from below 20% in 1981 to almost 80% in 2006 (Beach et al., 2011). Among the most commonly spoken immigrant languages in 2011, Cantonese, Punjabi, and Chinese were spoken at home by 25% of immigrants in Toronto and 40% of immigrants in Vancouver, respectively. In contrast, Arabic and Spanish made up almost one third of the languages spoken at home for immigrants in Montreal (Statistics Canada, 2012).

Language skills are important for many reasons, including the social and political integrations of immigrants and their access to health care and other social benefits (Deri, 2005; Derwing and Waugh, 2012; Remennick, 2004; Rublee and Shaw, 1991). Importantly language skills have been cited as a contributing factor to Canadian immigrants' worsening economic hurdles (Ferrer, Green and Riddell, 2006; Green and Worswick, 2010). A number of studies have documented an increasing wage gap between immigrant cohorts and their Canadian born counterparts.[3] Recent immigrants from non-English/non-French-speaking countries have been found to face substantial difficulties in integrating into the Canadian labour market as compared to previous immigrant cohorts (Aydemir and Skuterud, 2005).

This chapter has three main objectives. The first objective is to choose a preferred linguistic distance index from existing indices for measuring linguistic diversity in the Canadian context, where there are two official languages. The advantages and drawbacks of three indices, all approximations for the challenges in learning a specific language, are considered. The second and third objectives make use of the one-in-five samples of the confidential micro data files of the Canadian Censuses between 1981 and 2006. The second objective is to use the preferred linguistic distance index, the Levenshtein Distance (LD),

provinces and territories: Alberta, British Columbia, Manitoba, Nova Scotia, Ontario, Prince Edward Island, Saskatchewan, and Yukon.

[3] A few studies among many include Aydemir and Skuterud, 2005; Baker and Benjamin, 1994; Beach, Green and Worswick, 2011; Bloom, Grenier and Gunderson, 1995; Boyd and Cao, 2009; Ferrer, Green and Riddell, 2006; Grant, 1999; Green and Worswick, 2010, 2012; Skuterud, 2011.

to document changes in linguistic diversity among Canadian immigrant cohorts between 1981 and 2006. Linguistic diversity is calculated based on two different measures of immigrants' language: mother tongue, a language learned in childhood and still understood, and home language, the language spoken most often at home. Because mother tongue is time invariant, documenting linguistic diversity based on mother tongue across cohorts can be considered an alternative way to document changes in birth countries. Given the important role of language skills for labour market integration, the adoption of official languages as home language becomes particularly important for economic integration.[4]

The third objective is to examine the sociodemographic characteristics and the labour market performance of successive entry cohorts, defined as the groups of immigrants entering Canada within five years of a given census, by language skill group. Language skill groups are constructed based on the LD calculated using mother tongue at the time of arrival. By focusing on different levels of language skills, I aim to examine whether low language skills in the Canadian official languages are compensated with other observable characteristics such as high education levels and younger ages. Further, I aim to document the association between various levels of immigrant language skills and the labour market performance of immigrants.

The main contribution of this chapter is to use a fairly new linguistic distance measure, the LD, to describe linguistic diversity among Canadian immigrants. The LD, as a linguistic distance measure, is an approximation for the challenges in learning a second language (English or French in this chapter). The higher the LD, the more "distant" the official languages. The LD is the preferred index because of wide coverage of world languages, its ability to measure (linguistic) distance to both English and French, and its sensitivity to subtle differences between any two languages.

In this chapter, as well as throughout the thesis, care is taken to account for the important differences between Quebec and the rest of Canada, including the linguistic characteristics of Quebec, Quebec's extensive responsibility in immigrant selection process since the early 1990s, and Quebec's labour market. To this end, analyses are carried out

[4]In some cases alternative measures of language are used: the majority and national language. These measures are defined based on the immigrant's birth country and are defined later in section 1.2.

separately for Canada outside of Quebec and Quebec. Further, while linguistic distance is generally measured based on the "distance" to English, it is measured based on the "distance" to French in Quebec.

This chapter has three notable findings. First, on the one hand, the average LD based on mother tongue increased across Canada over the time period, consistent with the changing birth countries. On the other hand, the average LD based on home language went up outside of Quebec whereas it was stable in Quebec between 1981 and 1996. This suggests that although immigrant mother tongues shifted towards more "distant" languages in both regions, the uptake of the official languages as home language differed across regions and time periods. Second, shifts in birth countries differ in the two regions. Outside of Quebec, English-speaking countries lost their ranks among the top five most common birth countries over time. In Quebec, French-speaking countries were still among the top five during the 2000s. These findings are consistent with the trends in average LD indices in the two regions. Third, while the sociodemographic characteristics of recent immigrants in general are similar between Quebec and the other provinces, interesting differences emerge when immigrants are grouped by language skill group. Recent non-English-speaking immigrants are found to be more educated as compared recent English-speaking immigrants.[5] This implies that language skills and other human capital-related characteristics of immigrants are associated with each other. In addition, on average wages are positively related to language skills, suggesting that language skills are a determining factor for wages.

This chapter is organized as follows. Section 1.2 presents various linguistic distance measures. Section 1.3 presents the data. Section 1.4 describes the key features of linguistic diversity. Section 1.5 documents sociodemographic characteristics of recent immigrant cohorts with diversified language skills in terms of education, age, gender and their labour market performance in terms of employment and wages. The last section concludes.

[5]In two regions, various language skill groups are constructed based on the LD to English outside of Quebec and to French in Quebec. The group of the most skilled immigrants in language, the high language-skilled, includes those reporting English as mother tongue or home language if they reside in Canada outside of Quebec. The high language-skilled include those reporting French if they reside in Quebec. For the high language-skilled, the LD takes the value of zero. The other language groups categorize immigrants' languages into two additional language groups, from the closest to the most distant to English or French.

1.2 The measurement of linguistic distance

The Levenshtein Distance (LD) is the main indicator used in this study.[6] For comparison, two other indices are considered: the Adsera-Pytlikova (A-P) index and Chiswick-Miller (C-M) test scores (Adsera and Pytlikova, 2015; Chiswick and Miller, 1999; 2005).[7] This section describes how each index is constructed and highlights the main advantages and drawbacks of each. The LD measures "the distance" between any two given languages by quantifying phonetic similarities of a common set of words in two languages. I calculate the LD for all immigrants in the Census based separately on both the reported mother tongue and the home language to English for immigrants living outside of Quebec and to French living in Quebec. That the LD index can be constructed for all languages reported in the Census and to both English and French are the important advantages of the LD index as compared to the other two measures. In addition, it is sensitive to subtle differences between certain languages belonging to the same language branch or sub-branch.

The LD is a measure of linguistic distance that can be derived between any two languages using a computer algorithm, called the Automated Similarity Judgement Program, developed by the Max Planck Institute of Evolutionary Anthropology (Isphording and Otten, 2013). The algorithm converts the speech sounds of words with the same meaning in the most stable 40-item Swadesh list into a string of standardized phonetic symbols.[8] Many researchers have examined alternative word lists by using subsets of the original Swadesh list. Alternative subsets of the original 200 item Swadesh list were tested against each other by the Max Planck Institute for Evolutionary Anthropology with the finding that a 40 item subset gives the most stable results. The Automated Similarity Judgement Program

[6]The LD has been used in the immigration literature by Isphording and Otten (2013), Adsera and Pytlikova (2015), and Adsera and Ferrer (2015) along with the Linguistic Proximity index and Chiswick-Miller test scores.

[7]Other studies such as Adsera and Ferrer (2015, 2016) call it the Linguistic Proximity (LP) index. I transform this index to make it comparable with the LD so I call this transformed index the A-P index in this chapter.

[8]The original Swadesh list consists of 200 items, and is often used to determine the time that has passed among speech varieties since the separation from their common original language. The list was developed by emphasizing the importance of everyday language vocabulary rather than cultural language. Everyday language, associated with common things found anywhere in the world and known to each member of a society, is more resistant to borrowing from other languages than cultural language. The higher the percentage of shared words with a common origin in the item list, the more recently common language is divided into two or more speech varieties.

is based on this 40 item subset, which is presented in Table 1.1. These words refer to parts of the body, bodily functions or the sense of self and others.

The set of the phonetic symbols used by the algorithm includes seven vowels and 34 consonants that can easily be created by using any QWERTY keyboard (Brown et al. 2008; Holman et al. 2011). The computer algorithm measures the minimum number of deletions, additions, and substitutions of symbols required to transform one string of phonetic symbols to the other provided that both strings have the same meaning. This minimum number is used in deriving the LD. For example, by using phonetic symbols to transform the word "tooth" into the string of these symbols, the word "tooth" is transcribed into tu8 in English, and tant in Dutch. The first phonetic symbol, the consonant t, is the same in both strings and no change is required. However, the rest of the English phonetic representation of the word requires substitutions and an addition to transform it into its Dutch translation.

Petroni and Serva (2010) derive the so-called Normalized and Divided Levenshtein Distance, the NDLD, as follows. Consider two languages, α and β, with M pairs of words having the same meaning. The Normalised Levenshtein Distance, NLD, between the phonetic transcriptions of word i in both languages is the minimum number of deletions, substitutions, and additions to transform the word in language α to language β, and then divided by the maximum length of the longer phonetic string.

$$NDLD_i = D(\alpha_i, \beta_i) = \frac{D_{LD}(\alpha_i, \beta_i)}{L(\alpha_i, \beta_i)} \tag{1.1}$$

where $D_{LD}(\alpha_i, \beta_i)$ is the minimum number of changes, and $L(\alpha_i, \beta_i)$ is the maximum number of characters in the longer string of phonetic transcriptions.

The average of the NLD_i is calculated over M pair of words with the same meaning.

$$D(\alpha, \beta) = \frac{1}{M} \sum_i NDL_i = \frac{1}{M} \sum_i D(\alpha_i, \beta_i) \tag{1.2}$$

A second normalization is done in order to take into account coincidental phonetic similarities between languages that may exist for words with different meanings. This accidental similarity between words with different meanings is measured as:

$$\Gamma(\alpha,\beta) = \frac{1}{M(M-1)} \sum_{i \neq æ} NLD_{i \neq æ} = \frac{1}{M(M-1)} \sum_{i \neq æ} D(\alpha_i, \beta_j) \tag{1.3}$$

The $NDLD$ is accordingly defined as:

$$NDLD = \frac{D(\alpha,\beta)}{\Gamma(\alpha,\beta)} \tag{1.4}$$

The $NDLD$, or simply referred as the Levenshtein Distance, takes the value of zero if the two languages, α and β, are the same. The higher the Levenshtein Distance is, the more "distant" the languages. However, linguistic distance can be greater than 100 in cases where the accidental phonetic similarity is lower than the phonetic similarity between the two languages.

As an alternative measure, Adsera and Pytlikova (2015) use an index based upon so-called language trees. They assign a weight, between 0 and 1, to each language pair depending on the number of branches they share in the language-family tree. The linguistic distance between two languages is set at 0 if they are not related at any level, and at 1 if they are the same. If the languages are related at the first-linguistic tree level, they are given a weight of 0.1; and at the second, the third and the fourth levels, weights of 0.25, 0.40, and 0.70 are given respectively.

Figure 1.1 illustrates an example to show how the Adsera-Pytlikova index weights are assigned to languages based on their "distance" to English. English belongs to the Indo-European family of languages. According to the linguistic-family tree, Indo-European and Austro-Asiatic languages are not related at any level. Consequently, the index of linguistic distance assigned to Vietnamese, which belongs to Austro-Asian language family, is 0. Indo-Iranian languages (e.g. Hindi) and English are related at the first level of the language family, Indo-European language family, so these languages are assigned a 0.1 weight. If two languages share another sub-branch in the language family such as English

(West Germanic) and Swedish (North Germanic) as Germanic languages, another sub-branch in the Indo-European language family, they are assigned a weight of 0.25. If both languages share another sub-branch like German and English (West Germanic languages), they are assigned a weight of 0.45. If both languages are related at the forth-level such as Scots and English, they are assigned 0.70. English has weight 1. The Ethnologue, which is a language catalogue of almost 7,000 world languages, is used to determine the number of shared branches between languages so that the A-P index can be constructed for reported languages in the Census.[9] The A-P index is used in a limited sense in this study, assessing the robustness of some of the findings based on the LD index. For the sake of this comparison, the A-P weights are transformed by subtracting them from 1 and multiplying by 100: the larger the transformed index, the greater the distance to English or French.

A third measure of linguistic distance is offered by Chiswick and Miller (2005), and used in Chiswick and Miller (1999, 2001, and 2007). This indicator is based upon tests given to English-speaking Americans who are learning other languages. The U.S. Department of State, School of Language Studies, administers a test to assess the speaking efficiency of government employees who have had 16 to 24 weeks of training in a foreign language. It measures the distance to English only; it cannot be adapted to measure the distance to other languages, and particularly, for the purpose of this chapter, to French. The test scores take a value between one and three, with a higher score being associated with higher speaking proficiency. These are test scores of English speakers with an average set of skills, but who have a strong incentive to perform well since their career opportunities depend on second language acquisition. The Chiswick and Miller (2005) linguistic distance measure derived from these tests has the advantage over other indices in that it is actually measuring the challenges in learning another language. There are two important limitations with the C-M test scores. First, the C-M scores are available for only 43 languages. The authors assigned scores to closely-related languages to make the index available to a greater number of languages. These are listed in Table 1.2, along with their associated scores.[10]

[9]The online version of the Ethnologue is used: http://www.ethnologue.com.

[10]Although the number of languages included in Chiswick and Miller (2005) is more than 100 due to closely-matched connections among languages, Table 1.2 only reports languages included in the Canadian Censuses.

For example, it is reported that the distance to English for Farsi is 2, and this same score is then assigned to Pashto, Kurdish, Balochi and Osseta, languages that were not part of the original assessment due to non-availability of test scores for these languages. Second, the C-M test score does not offer a measure of difficulties of learning other languages for native French speakers.

All three of these indices have the limitation of assumed symmetry in learning difficulty: that is, native speakers of other languages are assumed, in other words, to have the same difficulty in learning English as English-speakers have learning their language.[11] This said, my review suggests that the LD is preferred due to its wide coverage of the world's languages, its sensitivity to subtle differences between certain languages, and its ability to measure linguistic distance to both English and French. It should be noted, however, that the LD is not a direct measure of the difficulty in learning a language although it is implicit in its construction. In order to assess how reasonable it is to approximate the difficulty in language acquisition with the LD, I offer a comparison to the C-M test scores for the subset of 65 languages for which the C-M score is available as a measure of difficulty in learning English. Figure 1.2 illustrates the comparison, with a horizontal movement from left to right indicating an increasing ease with which English speakers learn the language, and a vertical movement from low to high, indicating an increasing distance to English according to the LD. The relationship for this set of countries is negative once the C-M score exceeds 1.5.[12] Figure 1.2 suggests that a greater LD can be interpreted as a greater challenge in learning a new language. Further, the illustration confirms the claim that the LD captures other aspects of languages as well as their phonetic differences. As stressed by Serva (2011), its use is motivated by much broader considerations in the structural features of language.[13]

[11]For example, the C-M test scores of English-speakers learning Japanese and Malay are 1 and 2.75, respectively. This means that the distance between English and Japanese is greater than that between English and Malay, and by assumption would imply that a native Japanese-speaker would experience more difficulty learning English than a native Malay-speaker.

[12]The smooth line is a prediction from a quadratic least squares regression. The number of observation is 65 and determined by the languages for which a Chiswick-Miller score is available.

[13]Ginsburg and Weber (2011) address various aspects of linguistic distance between languages in detail, including vocabulary, phonetics, syntax, grammar, and structure. If a linguistic distance measure is based on the structure of a language, this measure quantifies similarities between any two languages based on the number of shared common roots of words in their vocabularies. The roots might be shared coincidentally or common roots might arise from borrowing or descending from an original language. In this regard, it should be noted that the LD combines phonetics with these structural aspects. In other words, instead of shared

For all these reasons, I use the LD to help illustrate changes in linguistic diversity among Canadian immigrant cohorts in the following sections.[14]

1.3 Data and the representation of linguistic distance

I restrict the sample I draw from the long-form Census files to foreign-born individuals and then further exclude Canadian citizens born abroad, immigrants having arrived in the census year, and immigrants for which LD could not be calculated.[15,16] Table 1.3 provides the sample sizes associated with the number of immigrants in each sample cycle. I define recent immigrants as those having entered Canada within five years of the census year. These cohorts include those entering Canada in the period of 1976-1980, 1981-1985, 1986-1990, 1991-1995, 1996-2000, and 2001-2005.

Information on mother-tongue and home-language is provided for each individual in the long-form files of all the censuses included in the analysis. Mother tongue refers to the

roots, shared speech sounds for a common word list are considered to quantify the distance between any two languages.

[14]Dyen et al. (1992) offer yet another lexico-statistical measure to quantify linguistic distance, which uses the 200-item Swadesh list but apply to 84 Indo-European languages or their dialects. The matching of each possible word pair across two languages is categorized as positive, negative or indeterminate. If two words with the same meaning borrowed any cognates (common roots) from each other over time, these cognates are characterized as positive. If no common characteristic is observed between words, it is claimed that cognates are negative. In cases that it would not be possible to determine whether the corresponding words share cognates or not, they are defined as indeterminate. The index is derived by dividing the number of positive cognates by the number of the positive and negative cognates. If two languages are the same, the index takes a value of 1000, and if two languages presumably share no cognates, it takes the value of zero. Dyen et al. (1992) also provide an alternative to test the strength of different measures of linguistic distance. These indexes are not used in this study since they refer to only a small subset of available languages in the sample, the Indo-European languages. This would limit the analysis to quantify changes in the distribution of linguistic differences since immigrants to Canada come from countries in which other language groups dominate.

[15]There are a number of reasons why the LD cannot be calculated for each immigrant. In the 1981 Census, 1,320 and 1,220 observations are recorded under the residual language category "other" for mother tongue and home language, respectively. The categories "Creole" and "non-official languages" are also present in the 1986 and 1991 Censuses. Immigrants that report more than one non-official language as their mother tongue or home language fall into the latter category. As a result, approximately 3,615 and 2,595 individuals cannot be associated with a language distance measure for their mother tongue and home language in these years. Only the categories "other languages" and "Creole" are recorded as residual language categories in the subsequent Census years. The number of immigrants without a language distance measure for mother tongue is approximately 8,030 in 1996, 10,660 in 2001 and 16,140 in 2006. Between 1996 and 2006, almost half of the immigrants that cannot be assigned a language measure for their mother tongue cannot be assigned the LD for their home language: 4,435 in 1996, 5,310 in 2001, and 7,450 in 2006.

[16]Given the fact that immigrants who arrive in Canada during the census year cannot be enumerated fully, the 1976-1981 cohort might be markedly different in 1981 than the 1976-1981 cohort in 1986. To avoid this problem, I exclude immigrants who arrive in Canada in the census year.

language first learned in childhood and still understood at the time of the census. Home language is the language spoken most often at home. There have been three important changes over censuses in how the answers are recorded. First, the 2001 and 2006 censuses provide more detailed information about languages spoken at home as they ask both about the language spoken most often as well which language(s) are regularly spoken at home. In order to ensure the comparability across censuses, only the information on the language spoken most often at home from 2001 and 2006 censuses is used.

Second, in the 1981 Census, an individual was instructed to report only one language for mother tongue and only one language for home language.[17] Starting with the 1986 Census, the instructions were expanded, clearly instructing the respondent to report more than one language if they are spoken equally often. In order to assign the LD to respondents reporting multiple languages from 1986 onward, I apply the following rule: the respondent is assigned the LD of zero if the respondent reports at least one official language corresponding to his province (French in Quebec and English outside of Quebec). If an individual reports more than two unofficial languages but no Canadian official language, he is assigned the LD of the language spoken by most of the immigrants coming from the same source country as him.

Third, there have been changes in the wording of questions and the collection of data. For example, in 1991, the word "at home" was added to mother tongue question and further instructions were provided for multiple responses. Further, the way data was collected changed such that check boxes were available only for the official languages, while a write-in answer space was provided for all non-official languages. Appendix Table A1.1 provides detailed information on home language and mother tongue questions, including questions' numbers, their wording, and the available check boxes for responses. In 2001 and 2006, additional information on language used at work was collected. Both the changes in the questions and the collection of data make the comparison across censuses a challenge, and require attention in describing the patterns in linguistic diversity among immigrants over time.

[17] Although multiple responses are given by some respondents in 1981, available data does not provide any information for more than one reported language in 1981.

All indices are derived for the following four measures of language skills: (1) mother tongue; (2) home language; (3) the language spoken by the majority in the country of origin; (4) the national language, which refers to the language of mass media, public communication, and the education system in the source country. This is done both for the distance to English in Canada outside of Quebec and distance to French in Quebec. The country of birth information is used to obtain the information on the majority language and the national language in the source country.[18,19]

The LD based on mother tongue, home language, the majority language, and the national language offer different proxies for language skills. Mother tongue can be considered as a measure of language-capital available to an immigrant at the time of the entry. The level of language-capital at the time of the entry is a contributing factor to immigrants' integration and is unlikely to change over time. Although it is least likely to be correlated with other observable characteristics of immigrants, it does not sufficiently capture accumulated language skills in the host country. Home language, on the other hand, is a better approximation of language skills. An immigrant's home language is an indicator of his/her extent of integration into the economic and social life of the host country. It is likely to change based on an immigrant's human capital-related characteristics as well as other factors such as having a spouse speaking an official language. Basing linguistic distance on the national-level language may, in fact, complement individual-level indicators such as mother tongue and home language because previous exposure to English or French may facilitate the development of language skills once in Canada even though neither official languages is

[18]The language spoken by the majority of the population in a country is obtained from the online version of the Ethnologue, which provides detailed information on the language family and different dialects of each language, but also geographical and demographic information on the countries where the language is spoken, and the approximate number of speakers (Lewis et al., 2013). Although demographic information is not continuously updated in the Ethnologue, it still provides a good approximation on the population distribution of ethno-linguistic communities in a country. This information is used to calculate the approximate percentage of each ethno-linguistic population in the total population of a country. If one of the languages in the country is spoken by at least the majority of population in the home country of an immigrant, the LD based on that language is assigned to each individual originating from that country. By applying this rule, some African countries and couple of European countries cannot be associated with an LD value for majority language as not one language is spoken by the majority of the population: Trinidad and Tobago, Nevis, French Guiana, Gibraltar, Yugoslavia, Burkina Faso, Liberia, Mali, Niger, Nigeria, Sierra Leone, Togo, Cote d'Ivoire, Guinea-Bissau, Kenya, Mozambique, Tanzania, Uganda, Zambia, Cameron, Chad, Oman, Qatar.

[19]Since the online version of Ethnologue defines the national language as "the language used in education, work, mass media, and government at the national level", I use the national language synonymously with the official language.

spoken at home most often. In addition, even if the language spoken by the majority in the home country is different than immigrant's mother tongue, previous exposure to various languages may improve an individual's ability to learn new languages in general. In particular, if the language spoken by the majority in home country is actually closer to English or French than an individual's mother tongue, it might foster the development of language skills in official languages. To cite one example, the majority in Gambia speak Mandinka, while the majority in Senegal speak Wolof. However, the national language of these two countries is, respectively, English and French. Mandinka and Wolof are significantly different than English and French, but this may overstate the challenges of adapting once in Canada if immigrants received previous exposure to French and English in the media or in school.[20] Many African countries have either English or French as their official language, and linguistic distance of these languages to the official languages of Canada might not be as far as expected by only focusing on an individual's mother tongue or home language.

1.4 Descriptive patterns in linguistic distance

This section uses the LD index to document the changing patterns in linguistic diversity among the Canadian immigrant population in three separate exercises. First, the average LD for the four measures of language skills are illustrated for recent cohorts of immigrants at the time of their entry. Second, the compositional shifts in source countries are described for both recent immigrant cohorts and all Canadian immigrants across censuses. Third, for each recent immigrant cohort, the changes in the average LD for mother tongue and home language are discussed over the period of 1981-2006.

1.4.1 The patterns of the average LD for recent immigrants

Figures 1.3 and 1.5 plot the average LD for the six cohorts of recent immigrants for the four measures of language skills for the two parts of the country.[21] The first major trend that

[20]The LD between Mandinka and English is 101.91, and the LD between Wolof and French is 96.09

[21]The average LD for English is calculated for all languages taking the values between 0 and 104.06. In this case, Vietnamese is the most distant language with the LD of 104.06. The average LD for French is calculated for all languages taking the value of 0 and 103.59. The most distant language to French is Rundi.

Figure 1.3 demonstrates is a clear increase in the average distance to English in Canada outside of Quebec. Second, the average LD jumps in 1991, reflecting a relatively large increase in LD for the cohort arriving between 1986 and 1990 as compared to previous cohorts. The increase appears for all four measures of language but is most pronounced for mother tongue. For example, the average LD increased by 14.39 points (about 34% of a standard deviation) in 1991. Third, the increase in the average LD stabilizes from 2001 onward. For example, the increase in the average LD for mother tongue was 0.27 points (less than 0.7% of a standard deviation) in 2006. The fact the average LD increases over time is consistent with the abandonment of the nationality-based immigration policy. The particular jump in the average LD for the 1986-1990 cohort might be engendered by the focus of the immigration policy on economic class rather than family class and refugees in the second half of the 1980s through the 1990s. In 2001, the Immigration and Refugee Protection Act increases the emphasis on economic immigrants' observable characteristics such as education, language skills and adaptability in the selection process, which might contribute to the stabilization of the average LD over this later period.

Along with these policy shifts, substantial changes occurred for language-related questions in the long-form questionnaires from 1986 to 1991. These changes include the changes in the order and wording of questions, and the addition of language-related question. While the proportion of individuals providing multiple responses for mother tongue and home language declined, the proportion of individuals reporting non-official languages as mother tongue and home language went up from 1986 to 1991, suggesting that once immigrants report their knowledge in official language(s), they might be less likely to report English or French as their mother tongue or home language.[22] Hence, more people reporting non-official languages will increase the average LD. In light of both the changes in the questionnaire and policy shifts over the period, the steep increase in the average LD between 1986 and 1991 should be considered carefully.

[22]Mother tongue- and home language-related questions were asked after the question related to the ability to carry on a conversation in official languages from 1991 onward. According to Statistics Canada (1993), the percentage of people reporting non-official languages as their mother tongue increased from 11.6 percent to 14.8 percent between 1986 and 1991. The percentage of reported non-official languages for home language went up to 7.7 percent in 1991, from 5.9 percent in 1986.

Turning to the average LD calculated using the other measures of language, the average LD is systematically lower when calculated based on home language as compared to mother tongue, suggesting that some immigrants have enough language capital to speak English as their home language, although they have non-English mother tongue. The average LD is lowest for the measure based on the national language of the source country while the majority language-based measure gives the highest average LD, although both trends are similar to those based on mother tongue and home language. This suggests that using immigrants' national languages overestimates language abilities while using the majority language-based measure underestimates these abilities relative to language abilities measured by mother tongue and home language.

As a robustness check, I repeat the same exercise above with the A-P index to examine if substantial differences exist in general trends of immigrants' linguistic diversity due to using alternative linguistic distance measures. Figure 1.4 presents the average A-P index for recent immigrants outside of Quebec. As illustrated, the general patterns in Figures 1.3 and 1.4 are the same. Parallel to the trends in Figure 1.3, the average A-P index reveals that linguistic diversity among Canadian immigrants increases over time. In particular, the steep increase is present for the 1986-1990 recent immigrant cohort. The increase in linguistic diversity stabilizes starting with 2001. Henceforth, this suggests that alternative linguistic distance measures are equally beneficial, while describing the general trends in linguistic diversity among immigrants.

Figure 1.5 plots the average LD from French for the four measures of language for immigrants residing in Quebec. There are three notable findings in these trends. First, the average LD based on home language did not change between 1981 and 1996 whereas the average LD based on mother tongue increased. Second, the average LD calculated from majority language provides the highest average values, similar to the majority language-based measure in Canada outside of Quebec. Third, compared to Figure 1.3, the average LD for French is consistently higher than the average LD for English at the beginning of the period and varies less dramatically over time. For example, the largest increase in the average LD in Quebec that is observed for mother tongue between 1986 and 1991 is less

than 27% of a standard deviation in 1991, which is 7 percentage points lower than that in Canada outside of Quebec. Given that the changes in the LD for majority and national languages are not as large as those for immigrants residing in other Canadian provinces, this is likely related to changes in immigration policy that changed the composition of immigrants admitted to or choosing to settle down in Quebec. Due to the Quebec-Canada Accord in 1991, which gave more autonomy to Quebec in selecting immigrants settling in Quebec, might have been more closely screened for language ability to further integration, at least, in terms of language so that the preservation of the language of the francophone population can be guaranteed.[23]

1.4.2 The shifts in source countries

To further illustrate the changing language composition of Canadian immigrant cohorts based on mother tongue, I next present the top five most common birth countries for both all immigrants and immigrant groups based on the LD for each Census year separately. Since mother tongue is time-invariant, documenting birth countries based on mother tongue is an alternative way to document the changes in linguistic diversity among Canadian immigrants. The analysis above used only recent immigration (immigrant flows) but in the following section I also consider the full sample of immigrants from each census year (immigrant stocks).[24] Tables 1.4 and 1.5 present top five birth countries for immigrant stocks and flows residing outside of Quebec between 1981 and 2006. Tables 1.6 and 1.7 replicate this exercise for the immigrants from Quebec. Columns 2 and 3 present the top five countries and the share of all recent immigrants from these countries in a given year. Columns 4 to 11 present birth countries and the share of recent immigrants in four language groups defined with respect to the distance of mother tongue to English, using the LD. The first category includes high language-skilled immigrants in English, who report English as mother tongue. The LD

[23]In Quebec and the rest of Canada, admission rules require that, as a proof of language skills, an immigrant should provide either the results of a language test or a recognized diploma in English or French. Before the Immigration and Refugee Protection Act in 2002, whether proof of language ability is acceptable depended on the subjective evaluation of interviewer (Adsera and Ferrer, 2016). Following the changes in the immigration policy in 2008, test results have become mandatory (Alboim and Kohl, 2012). The distribution of points for French and English language skills varies across the two regions.

[24]Recent immigrants and immigrant flows are used interchangeably.

index takes the value of zero for these immigrants. The second category mainly includes immigrants speaking Indo-European languages whose LD is less than 95.22, the distance between English and Hungarian.[25] The third is composed of immigrants with mother tongue with the LD greater than or equal to 95.22 and less than or equal to 100.58.[26] The final language category is comprised of Asian languages primarily and represents immigrants with LD greater than 100.58. In Tables 1.6 and 1.7, the distance of mother tongue to French is used for grouping. The first and second cut-off languages are chosen as Hebrew and Pashto.

Table 1.4 illustrates the top five birth countries of all immigrant stocks and the share of immigrants from these countries. Column 3 reports the share of immigrants from the top five birth countries. Columns 5, 7, 9, and 11 report the percentage of immigrants from the top five birth countries, given that immigrants are grouped into language groups based on mother tongue. Both the top five birth countries and the share of immigrants changed from 1981 to 2006, although the composition of birth countries stayed almost the same between 1981 and 1991 for immigrant stocks. For example, the share of immigrants from United Kingdom dropped from 25.3% in 1981 to 10.6% in 2006 but it was still the top birth country of all immigrants outside of Quebec. China first appeared among the top five source countries in 2001 and became the second top source country of immigrant stocks in both 2001 and 2006. By language group, birth countries do not seem to change much over time although the shares of immigrants coming from these countries vary substantially. For example, among English-speaking immigrant stocks, the shares of immigrants from the top five birth countries fall over time, while the composition almost stays the same. Since the second and fourth language groups include languages which are more or less geographically

[25] In general, Indo-European languages are closer to English or French in comparison to non-Indo European languages due to descending from the common original language. In this regard, after ranking all languages by the LD, Hungarian, as the non-Indo-European language with the lowest LD, is chosen as the first cut-off language. Punjabi, as the Indo-European language with the highest LD, is chosen as the second cut-off language.

[26] The second cut-off language is determined as the Indo-European language with the highest LD. Between the first and second cut-off language, the third group includes both Indo-European and non-Indo-European languages.

concentrated, the countries in these categories change very little.[27] The third language group that includes immigrants speaking languages that are dispersed all around the world, i.e. Indo-European and non-Indo-European languages, birth countries are found to change significantly over time in this group. The fact that the nationality-based immigration policy was abandoned contributed to the changes in birth countries and more diversified language skills among immigrants over the last decades. This is clearly observed in the third language group through the shift from European countries to Asian countries.

Table 1.5 reports the top five birth countries and the share of immigrants from these countries only for those entering Canada within five years of the census. The shift in birth countries is more evident when flows, as opposed to stocks, are considered. In 1981 UK and the USA were the first and second most common birth country, but by 2006 UK was not in the list and the US fell to the fifth position for recent immigrants. The share of immigrants from the US declined from 8.4% in 1981 to 3.8% in 2006. Between 1981 and 2006, India ranked in the top five countries but it only appeared in the top five in 2001 and 2006 among immigrants with English as mother tongue. From 1986 onward, China was regularly in the top five birth countries. The share of recent immigrants from China went up from 6.5% to 13.7% between 1986 and 2006. Looking at language groups separately, on the one hand, there is less change over time in the first and fourth language groups as compared to the second and third language groups. Four countries in the first language group and three countries in the fourth language group ranking among the top five in 1981 still ranked among the top five in 2006. On the other hand, none of the top five countries in 1981 appeared as five most common birth countries in 2006 in the second language group. In addition to the changes in the composition of birth countries, the changes in the distribution of shares among the top five countries are evident from 1981 to 2006, in particular in the first and fourth language groups. For example, the share of the top five countries is less equally distributed in the first language group in 1981 than in 2006, while the opposite is observed for the fourth language group. The changes in the composition of birth countries

[27] This is particularly true for the fourth language group. The second language group includes few Indo-European languages spoken by immigrants coming from the countries in the Middle East or South Asia such as Iran and India.

show that geographical shifts, from traditional sources to non-traditional sources, in birth countries became clear even for immigrants with similar language skill groups such as the second language group. Overall, the shift in birth countries suggests that not only the number of immigrants with more distant languages increased over time, but the number of immigrants from non-traditional source countries reporting languages as close to English as European languages increased as well.

As presented in Tables 1.4 and 1.5, the same country can rank among top five birth countries in two different language groups within the same census. For example, China is among top five birth countries in both the third and fourth language groups in 2001 and 2006. Immigrants from the same birth country do not necessarily report the same language as mother tongue. In fact, there are countries throughout which a wide variety of languages is spoken. As an example, according to Ethnologue, there are around 290 languages spoken throughout China. Since the LD captures subtle differences between these languages and English/French, the same country can be among top five countries in two different language groups based on the distance of the reported language to English/French.

Tables 1.6 and 1.7 present the shift in birth countries among immigrant stocks and flows residing in Quebec. Language groups are based on the distance of mother tongue to French. In Table 1.6, only Italy and France, which ranked among the top five birth countries of immigrants stocks in 1981, kept their place among the top five in 2006. By language group, the differences in the compositional change were substantial as well. The countries in the first and second language groups were almost the same in the 1981 and 2006 censuses while the country composition in the third and fourth language groups changed notably. Only two countries on the top five list in 1981 ranked among the top five again in 2006 in the third and fourth language groups (Greece and (West) Germany in the third language group and Vietnam and Lebanon in the fourth language group). Table 1.7 documents significant changes in the immigrant flows by birth country. France was the only country ranked among the top five countries of immigrant flows in both the 1981 and the 2006 censuses. In comparison to other language groups, the first language group which includes immigrant flows in Quebec reporting French as mother tongue has the most stable composition of birth

countries. Four countries ranking among the top five in 1981 were still on the list in 2006. In the other language groups, almost all countries in 1981 lost their position among top five in 2006.

A number of striking differences between the regions become apparent when comparing Tables 1.4 and 1.5 with Tables 1.6 and 1.7. First, France, as the main source country of French-speakers, not only ranked among the top five countries but also increased its share for recent immigrants, from 6.2% in 1981 to 7.7% in 2006 in Quebec. In the other Canadian provinces, UK as one of the main source country as English-speaking immigrants did not rank among the top five countries of immigrant flows after 1986. The US whose share continuously declined over time barely made it to the top five as the fifth country in 2006. Second, by language group, the composition of the countries including French-speakers in general is the most stable group among immigrants in Canada. The differences across regions might be related to policy changes giving Quebec more autonomy in selecting immigrants since 1991 as well as to the preference of French-speaking immigrants to settle down in Quebec. In either case, this exercise highlights the compositional differences in immigrants' birth countries in Quebec and the rest of Canada. It also illustrates the shift in birth countries for immigrant flows and stocks between 1981 and 2006 in both regions.

1.4.3 Evolution of the average LD across cohorts

This section tracks the average LD based on mother tongue and home language for the six arrival cohorts over the six censuses to describe the evolution of language skills with time. Unless there is non-random emigration from Canada or across the two regions which would change the average LD based on mother tongue, it is expected that the average LD for mother tongue should not change for recent immigrant cohorts over time. If the average LD for mother tongue for each immigrant cohort is fairly stable over time, then, a decline in the average LD for home language clearly ascribes to language skill accumulation of immigrants over time. In other words, the decline in the average LD for home language would not stem from return migration but rather the adoption of the official languages of Canada as their home language.

Table 1.8 reports the average LD for mother tongue and home language for the six arrival cohorts in Canada outside of Quebec. As presented in Table 1.8, compared to earlier cohorts, more recent cohorts have larger gaps between the average LD for home language and mother tongue indicating greater distance between their home language and English upon arrival. In addition, there is a continuous decline in the average LD based on home language, while the average LD based on mother tongue is more or less stable over time. This clearly indicates that immigrants accumulate language skills over time and speak English as their home language. Table 1.9 provides the same averages for immigrants residing in Quebec. The same patterns are observed. The average LD for mother tongue for immigrants in Quebec stays almost the same and the average LD for home language gradually declines over time.

In both tables, the changes in the average LD, in particular based on mother tongue, between 1981 and 1991 is greater than those in other years. For example, the average LD for mother tongue of recent immigrants entering Canada between 1976 and 1980 declined from 54.26 in 1981 to 49.32 in 1986, and then increased to 56.74 in 1991. As discussed previously, in addition to changes in data gathering, more information regarding multiple responses was provided to respondent for language-related questions. It is likely that these contributed to the unstable pattern in the average LD of the 1976-1980 immigrant cohort, rather than non-random emigration. Nevertheless, the changes in questionnaires over time should be recalled, while interpreting the changes in the patterns of LD over time.

1.5 Immigrants' sociodemographic characteristics and labour market performance

It is documented that, in addition to changes in source countries and languages, many other characteristics of immigrants have been changing over time. Compared to previous cohorts, more recent immigrants are younger and more educated. Beach et al. (2011) report that the percentage of prime working age immigrants, who are between 25 and

54, went up among the most recent immigrants in 2006. Further, they report that the percentage of immigrants, who hold a bachelor degree or above, continuously increased over time, from 7.5% in 1980 to 34% in 2000. Nonetheless, recent cohorts have fared less well in the labour market. In this section, I consider recent immigrants as a whole to verify existing evidence and then separate recent immigrants into language groups to examine the association between language skills and immigrants' sociodemographic characteristics and labour market performance. As before, I use the same four language groups based on mother tongue. I consider mother tongue as the approximation of the immigrant's language capital at the time of entry and describe other characteristics associated with human capital and labour market performance conditional on that language capital. I devote particular attention to sociodemographic characteristics associated with the aspects of human capital such as age, gender, and education. In terms of labour market performance, I concentrate on employment and wage. I am going to focus on these particular aspects of immigrant characteristics and labour market performance because they relate to the way I am going to frame the analytical part of the thesis, Chapters 2 and 3. In particular, given education, years of labour market experience and language skills, I will examine the impact of immigrant inflows on wages in Chapter 2. Chapter 3 is going to examine whether there are differences in the impact of immigration by gender.

I first describe the distribution of immigrants across age groups. Age groups are defined based on respondents' age at the time of the census. Four age groups are defined for those: (i) at the age of 13 and younger; (ii) older than 13 and younger than 20; (iii) between 21 and 64; (iv) older than 64. Three education groups are defined for those: (i) who hold at most a high school degree; (ii) who hold a trade certificate or any degree below bachelor level; (iii) who hold at least a bachelor degree. Furthermore, I define immigrants to be employed if they report holding a full-time job at least 40 weeks of the census year, and report positive hours of work and wages in the census year. The sample of the employed is also used to calculate the average logarithm of real weekly wages. The annual wages are adjusted to 2005 prices by using the Canadian CPI and divided by the number of weeks worked. Then, the simple average of logarithm of real weekly wages is used. Education

and employment distributions as well as the average logarithm of real weekly wages are presented for recent immigrants that are between 21 and 64 years old.

Table 1.10 presents the age distribution for the six arrival cohorts residing in Canada outside of Quebec. The percentage of child immigrants aged 13 years and younger decreased from 21% in 1981 to 18% in 2006. The share of working-age immigrants aged 21 to 64 increased from 61% to 67% over the last three decades. The biggest change was observed for immigrants older than 64, whose share decreased by 28%, from 4.6% in 1986 to 3.3% in 2006. By language group, the most substantial changes occurred within the immigrant group with the most distant languages. The percentage of immigrants that are younger than 21 declined by 8 percentage points. The share of working-age immigrants with the most distant languages increased from 63% to 72% between 1981 and 2006. The share of the youngest continuously declined, while the share of working population gradually went up for immigrants with the most distant languages between 1981 and 2006.

Table 1.11 shows that gender distribution stayed more or less the same across the six immigrant cohorts. The number of female immigrants was higher than that of male immigrants between 1981 and 2006 and the gap between the share of females and males was around 4 to 5 percentage points except for the immigrant cohort entering Canada in 1991. By language group, the largest change in gender distribution occurred for immigrants with the most distant languages. The share of female immigrants increased from 51.6% in 1981 to 53.4% in 2006. The difference between the shares of females and males increased from 3.2 to 6.8 percentage points for the immigrants with the most distant languages. The fact that women are main caregivers in the family and they are more likely to have intermittent work experience than men, the increase in the share of women might adversely affect the likelihood of the adoption of Canadian official languages within cohorts. Further, while analyzing labour market performance of cohorts, examining men and women separately rather than the pooled sample might give a more accurate description in their labour market performance.

Table 1.12 presents the distribution of education groups for the subsample of immigrants 21-64 years. While the share of immigrants with at most a high school degree

declined from 55.4% to 28%, the percentage of immigrants with at least a bachelor degree went up to 48.3%, from 16.7%, over the last there decades. The distribution of educational attainment is found to vary dramatically by language skill group. The largest increases in the share of the most educated immigrants are observed in the second and fourth language groups, where these shares more than doubled. In the second language group, the share of the most educated increased from 14.2% in 1981 to 50.7% in 2006, while the share of the most educated with the most distant languages went up from 15.2% to 46.9% between 1981 and 2006. During the 2000s, a non-English-speaking immigrant was more likely to have arrived in Canada with a higher level of education relative to an English-speaking immigrant.

Table 1.13 shows the proportion of employed immigrants by gender for the subsample aged 21-64 years. The likelihood of employment for newly arrived immigrants varied over time. The lowest employment rate occurred for the immigrants entering Canada during the first half of the 1990s. Females experienced more hurdles in finding a job in their first five years in Canada relative to males. By language group, immigrants with the most distant languages experienced the least decline in employment rates within the five years after arrival from 1981 to 2006 whereas the immigrants in the second language group that include those reporting a non-English mother tongue that is relatively close to English saw the largest decline in employment rate. The employment rate for this group declined by 15% over the period. This is particularly interesting. Recall that the largest increases in the share of the most educated occurred for the second and fourth language groups. Assuming the same average education level for the two groups, a recent immigrant with a non-English mother tongue but higher language skills is less likely to be employed within the first five years in Canada. This might be because my measure of language skills is imperfect, or immigrants with high education levels but low language skills might be more likely to accept lower skilled jobs compared to those equally educated but with higher language skills.

Table 1.14 presents the average logarithm of real weekly wages for immigrant groups. Entry level wages declined over time for successive immigrant groups for male immigrants and increased for female immigrants. In 2006, the entry earnings of both

English-speaking immigrants and immigrants with the most distant languages were able to catch up with the entry earnings of these groups in 1981. During the period, the largest decline occurred for non-English-speaking immigrants with languages closest to English. The most striking result in Table 1.14 is that the immigrants with the most distant languages had the lowest average wages between 1981 and 2006, suggesting an important role for language skills in the determination of wages.

Table 1.15 presents the changes in the age distribution of recent immigrants in Quebec. There are two striking differences from the age distribution in Canada outside of Quebec. First, the percentage of the oldest age group among recent immigrants went down gradually and continuously over time as the immigration policy was increasingly targeting working age immigrants. This gradual downward trend became more pronounced after 1991 by language group. Second, by language group, both the share of the youngest and the oldest declined for non-French-speaking immigrants in 2006 as compared to those in 1981. In Table 1.16, the share of female is almost the same for the six immigrant cohorts in Quebec over time. This is slightly different than the gender distribution in other provinces, where the share of females exceeds the share of males by 4 to 5 percentage points for most arrival cohorts. With respect to language groups, in comparison to the other provinces in Canada, the differences in gender distributions in both regions are less notable for the immigrants speaking either French or a language relatively close to French than for the immigrants with more distant languages. Table 1.17 shows the educational composition of immigrants. The patterns for education groups are comparable with those for the rest of Canada in general. In Quebec, education composition most dramatically changed for the immigrants with the most distant languages.

Table 1.18 shows that the rate of employment declined by 8 percentage points for recent immigrants in Quebec between 1981 and 2006. Female and male employment rates went down by around 4 and 12 percentage points over the same period. These results are the same as what is found in Table 1.13 for other Canadian provinces. The most striking difference is observed for French-speaking immigrants. The employment rate of French-speaking immigrants increased by 5 percentage points from 1981 to 2006. The lion's share

went to French-speaking female immigrants and their employment rates increased by 9 percentage points. As compared to all other immigrant groups in both regions, this is the only immigrant group that experienced an increase in employment rates from 1981 to 2006. In Table 1.19, the average logarithm of real weekly wages declined from 6.26 log points in 1981 to 6.15 log points in 2006. The decline in the average wages from 1981 to 2006 is similar to the decline in other Canadian provinces. In Quebec, the largest decline occurred for male immigrants. By language group, all wages declined from 1981 to 2006 except for French-speaking immigrants. In general, language skills are positively associated with average wage levels, which go down as language skills are lower. The lowest average wages were observed among the immigrants with the most distant languages in Quebec between 1981 and 2006.

When analyzed together, immigrants in Quebec and immigrants outside of Quebec show similarities in terms of sociodemographic characteristics. In general, an immigrant became more likely to be aged between 21 and 64 and to be more educated in the two regions over time. However, when analyzed separately by language skill group, some interesting differences emerge. For example, although the age distribution has shifted over time, favouring younger rather than older immigrants in general, the working-age population was particularly favoured for the immigrants with languages other than French in Quebec. As another example, the number of females generally exceeded the number of males regardless of language skills in Canada outside of Quebec. In Quebec, the number of females was particularly lower than the number of males for the immigrants with the most distant languages to French. In terms of labour market performance, similarities for all immigrants and differences across language groups emerge as well in both regions. For example, average wage levels declined for all recent immigrants in both regions from 1981 to 2006, while the lowest wage levels were of the immigrants with the most distant languages. By language group, all immigrant language groups experienced the deterioration in their employment rates and wage levels outside of Quebec. However, in Quebec, French-speaking immigrants experienced an increase in their employment rates and wages while these outcomes for all

other non-French-speaking immigrants followed a similar downward movement as those for their English- or non-English-speaking counterparts in other provinces.

Overall, these results emphasize three main points. First, immigrant characteristics vary by language skill group, suggesting other human capital characteristics of immigrants play a role to compensate low levels of language skills. Second, given that wages decline with lower language skills, language skills play a role in wage determination. Third, given that immigrants' labour market performance differs by language skill group across the two regions, it is meaningful to treat these regions separately. Chapters 2 and 3 are going to focus on particular aspects of sociodemographic characteristics of immigrants as described above and their wage performance in both regions.

1.6 Conclusion

The major objective of this chapter is to describe the changes in linguistic diversity among Canadian immigrants by using a fairly new linguistic distance measure, the LD. As an approximation of immigrant language skills, a linguistic distance measure represents the difficulty experienced by immigrants in learning the host country's official language(s) in cases where no direct measures of language skills are available. The second motivation is to examine previously documented sociodemographic characteristics of Canadian immigrants and their labour market trends by language skill groups. Language groups are defined based on the LD from various measures of language to English for immigrants residing outside of Quebec and to French in Quebec. There are four alternative measures in the descriptive analysis, but the focus is mainly on the two measures based on mother tongue and home language.

There are three notable findings in the descriptive analysis of linguistic diversity. First, the patterns in average LD differ in both regions. For example, the average LD for home language continuously increases outside of Quebec between 1981 and 1996, the average LD for home language in Quebec stays stable. Second, two regions differ regarding the shifts in source countries for recent immigrants as well. Outside of Quebec, predominantly English-speaking countries lose their rank among the top five birth countries between 1981

and 2006. In Quebec, French-speaking countries remain among the top five source countries even during the 2000s. Third, the decline in the average LD for home language of each cohort suggests that immigrants adopt official languages in both regions over time.

Immigrant sociodemographic characteristics and labour market performance are described based on the distance from the reported mother tongue to the prominent languages in both regions. The main conclusion from this exercise is that labour market performance and sociodemographic characteristics show very similar patterns for all immigrants across the two regions, but differ when immigrants are analyzed separately by language group. For example, language skills vary with wage levels in both regions, suggesting a role for language skills in determining wages. The performances of French-speaking immigrants in Quebec and English-speaking immigrants outside of Quebec follow opposite patterns. French-speaking immigrants experienced an increase in employment rates and wage levels between 1981 and 2006, while the employment rates and wage levels of English-speaking immigrants declined over the same time period.

There are a number of directions in which this research could be continued. For example, descriptive work could be expanded to construct alternative language groups to examine if the changes in immigrant characteristics and labour market performance are sensitive to various skill definitions. Further work could be undertaken to describe linguistic diversity in Canada by using alternative data sets, which includes a direct measure of language skills. This said, my findings are similar to the findings in the literature, which highlight the role of language skills in the determination of labour market performance, in particular, wages.

Immigration policies have changed and have led to changes in the composition of immigrants. Since immigrants are an important part of the Canadian labour market, the shifts in immigrant composition engender the compositional shifts in labour supply. Then, combined with the fact that immigrant labour market performance varies across language skill groups, these lead to the question of how the shifts in labour supply composition due to immigration contributes to the differences in workers' performances in the Canadian labour market. In Chapter 2, I explore further this link to estimate the impact of immigrant

inflows with different education levels, years of labour market experience and language skills on Canadian wages.

References 1

[1] **Adsera, Alicia and Ana Ferrer**. "The effect of linguistic proximity on the occupational assimilation of immigrant men". IZA Discussion Paper, no. 9499, 2015.

[2] **Adsera, Alicia and Ana Ferrer**. "Occupational skills and labour market progression of married immigrant women in Canada". *Labour Economics*, 39:88–98, 2016.

[3] **Adsera, Alicia and Mariola Pytlikova**. "The role of language in shaping international migration". *The Economic Journal*, 125(586):49–81, 2015.

[4] **Alboim, Naomi and Karen Cohl**. "Shaping the future: Canada's rapidly changing immigration policies". Policy report. Toronto: Maytree Foundation. Available at http://maytree.com/wp-content/uploads/2012/10/shaping-the-future.pdf, 2012.

[5] **Aydemir, Abdurrahman and Mikal Skuterud**. "Explaining the deteriorating entry earnings of Canada's immigrant cohorts: 1966-2000". *Canadian Journal of Economics*, 38(2):641–671, 2005.

[6] **Baker, Michael and Dwayne Benjamin**. "The performance of immigrants in the Canadian labour market". *Journal of Labor Economics*, 13(3):369–405, 1994.

[7] **Beach, Charles, Alen. G. Green, and Christopher Worswick**. *Toward improving Canada's skilled immigration policy: An evaluation approach.* CD Howe Institute, Toronto, 2011.

[8] **Bloom, David E., Gilles Grenier, and Morley Gunderson**. "The changing labour market position of Canadian immigrants". *Canadian Journal of Economics*, 28(4b):987–1005, 1995.

[9] **Boyd, Monica and Xingshan Cao**. "Immigrant language proficiency, earnings, and language policies". *Canadian Studies in Population*, 36(1-2):63–86, 2009.

[10] **Brown, Cecil H., Eric W. Holman, Soren Wichmann, and Viveka Velupillai**. "Automated classification of the world's languages: a description of the method

and preliminary results". *STUF-Language Typology and Universals*, 61(4):285–308, 2008.

[11] **Challinor, Ashley E.** "Canada's immigration policy: A focus on human capital", 2015. Retrieved July, 2016. `http://www.migrationpolicy.org/article.html`.

[12] **Chiswick, Barry R.** "The effect of Americanization on the earnings of foreign-born men". *Journal of Political Economy*, 86(5):897–921, 1999.

[13] **Chiswick, Barry R. and Paul W. Miller**. "English language fluency among immigrants in the United States". In Solomon W. Polacheck, editor, *Research in Labor Economics*, volume 17, pages 151–200. JAI Press, Oxford, 1999.

[14] **Chiswick, Barry R. and Paul W. Miller**. "A model of destination-language acquisition: Application to male immigrants in Canada". *Demography*, 38(3):391–409, 2001.

[15] **Chiswick, Barry R. and Paul W. Miller**. "Linguistic distance: A quantitative measure of the distance between English and other languages". *Journal of Multilingual and Multicultural Development*, 26(1):1–11, 2005.

[16] **Chiswick, Barry R. and Paul W. Miller**. "Computer usage, destination-language proficiency, and the earnings of natives and immigrants". *Review of the Economics of the Household*, 5(2):129–157, 2007.

[17] **Citizenship and Immigration Canada**. "Frederal-provincial/territorial agreements", 2010. Retrieved on June, 2016. `http://www.cic.gc.ca/english/department/laws-policy/agreements.html`.

[18] **Citizenship and Immigration Canada**. "Facts and figures 2015: Immigration overview: Permanent residents", 2015. Retrieved on January, 2017. `http://www.cic.gc.ca/english/resources/statistics/menu-fact.asp`.

[19] **Deri, Catherine**. "Social networks and health service utilization". *Journal of Health Economics*, 24(6):1076–1107, 2005.

[20] **Derwing, Tracey M. and Erin Waugh**. "Language skills and the social integration of Canada's adult immigrants". *IRPP Study*, 31(1):1–36, 2012.

[21] **Dyen, Isidore, Joseph B. Kruskal, and Paul Black**. "An Indoeuropean classification: A lexicostatistics experiement". *Transactions of the American Philosophical Society*, 82(5):iii–iv+1–132, 1992.

[22] **Ferrer, Ana, David A. Green, and William C. Riddell**. "The effect of literacy on immigrant earnings". *Journal of Human Resources*, 41(2):380–410, 2006.

[23] **Ferrer, Ana, Garnett Picot, and William C. Riddell**. "New directions in immigration policy: Evolving approach to immigration selection". Canadian Labour Market and Skills Researcher Network, Working Paper, no. 107, 2012.

[24] **Ginsburg, Victor and Shlomo Weber**. *How many languages do we need? The economics of linguistic diversity.* Princeton University Press, Princeton, NJ, 2010.

[25] **Grant, Mary L.** "Evidence of new immigrant assimilation in Canada". *Canadian Journal of Economics*, 32(4):930–955, 1999.

[26] **Green, Alan G. and David A. Green**. "The economic goals of Canada's immigration policy: Past and present". *Canadian Public Policy*, 25(4):425–451, 1999.

[27] **Green, David A. and Christopher Worswick**. "Entry earnings of immigrant men in Canada: The roles of labour market entry effects and returns to foreign experience". In Ted McDonald, Elizabeth Ruddick, Arthur Sweetman, and Christopher Worswick, editor, *Canadian Immigration: Economic Evidence for A Dynamic Policy Environment*, pages 77–110. McGill-Queen's University Press, Kingston, ON, 2010.

[28] **Green, David A. and Christopher Worswick**. "Immigrant earnings profiles in the presence of human capital investment: Measuring cohort and macro effects". *Labour Economics*, 19(2):241–259, 2012.

[29] **Holman, Eric W., Soren Wichmann, Andre Muller, Viveka Velupillai, Harold Hammarstrom, Sebastian Sauppe, Hagen Jung, Dik Bakker,**

Pamela Brown, Oleg Belyaev, Matthias Urban, Robert Mailhammer, Johann-Mattis List, and Dimitri Egorov. "Automated dating of the world' language families based on lexical similarity". *Cultural Anthropology*, 52(6):841–875, 2011.

[30] **Isphording, Ingo E. and Sebastian Otten**. "The costs of Babylon–Linguistic distance in applied economics". *Review of International Economics*, 21(2):354–369, 2013.

[31] **Lewis, Paul, Gary S. Simons, and Charles D. Fenning (Ed.)**. *Ethnologue: Languages of the World*. Seventeenth edition. International, Dallas, Texas, 2013. Online version: `http://www.ethnologue.com`.

[32] **Nijboer, Harriet**. "Frederal-provincial relations on immigration: Striking the right balance". Retrieved June, 2016. `https://tspace.library.utoronto.ca/bitstream/1807/25877/3/Nijboer_Harriet_201011_LLM_Thesis.pdf`, 2010.

[33] **Petroni, Filippo and Maurizio Serva**. "Measures of lexical distance between languages". *Physica A: Statistical Mechaniscs and its Applications*, 389(11):2280–2283, 2010.

[34] **Remennick, Larissa**. "Language acquisition, ethnicity and social integration among former Soviet immigrants of the 1990s in Israel". *Ethnic and Racial Studies*, 27(3):431–454, 2004.

[35] **Serva, Maurizio**. "Phylogeny and geometry of languages from normalized Levenshtein distance". Cornell University Library, Ithaca, NY `http://arxiv.org/abs/1104.4426v3`, 2011.

[36] **Skuterud, Mikal**. "Language skills in the new economy and the deteriorating labour market performance of Canada's immigrant workers". Report, Human Resources Development and Skills Canada, 2011.

[37] **Statistics Canada**. "Home language and mother tongue - The nation". 1991 Statistics Canada Census Catalogue no:93-317. Ottawa: Statistics Canada., 1993.

[38] **Statistics Canada**. "Mother tongue - 1991 Census technical reports". 1991 Statistics Canada Census Catalogue no:92-335E. Ottawa: Statistics Canada., 1993.

[39] **Statistics Canada**. "Linguistic characteristics of Canadians". 2011 Statistics Canada Census Catalogue no:98-314-X2011001. Ottawa: Statistics Canada., 2012.

[40] **Swadesh, Morris**. "Lexicostatistics dating of prehistoric ethnic contacts". *Proceedings of the American Philosophical Society*, 96(4):452–463, 1952.

[41] **Xue, Li**. "A comprehensive look at the employment experience of recent immigrants during their first four years in Canada". In Ted McDonald, Elizabeth Ruddick, Arthur Sweetman, and Christopher Worswick, editor, *Canadian Immigration: Economic Evidence for A Dynamic Policy Environment*, pages 11–40. McGill-Queen's University Press, Kingston, ON, 2010.

Table 1.1: The 40-item list of common language words from the Swadesh list

I	ear	knee	nose	sun
blood	eye	leaf	one	tongue
bone	fire	liver	path	tooth
breast	fish	louse	person	tree
come	full	mountain	see	two
die	hand	name	skin	water
dog	hear	new	star	we
drink	horn	night	stone	you

Source: Isphording and Otten (2013).
Notes: The 40-item subset of common language words from the main Swadesh list is used in the development of the Levenshtein Distance index.

Table 1.2: Linguistic distance to English in the original set of tested languages

Chiswick-Miller Index								
3.00	2.75	2.50	2.25	2.00	1.75	1.50	1.25	1.00
Afrikaans	Dutch	French	Catalan	Amharic	Bengali	Arabic	Cantonese	Japanese
Norwegian	Frisian	Haitian Creole	Danish	Azerbaijani	Burmese	Fuchow	Fukien	Korean
Romanian	Malay	Italian	German	Bulgarian	Greek	Lao	Hakka	
Swedish	Swahili	Portuguese	Russian	Cambodian	Gujarati	Mandarin		
			Spanish	Czech	Hindi	Min Nan		
			Yiddish	Dari	Marathi	Vietnamese		
			Bylerossian	Farsi	Nepali	Wu		
			Ukranian	Finnish	Panjabi			
				Hebrew	Sinhala			
				Hungarian				
				Indonesian				
				Kurdish				
				Macedonian				
				Mongolian				
				Pashto				
				Polish				
				Serbo-Croat				
				Slovak				
				Slovene				
				Tagalog				
				Thai				
				Tibetian				
				Turkish				
				Uzbek				

Source: Chiswick and Miller (2005).

Table 1.3: Selection of foreign-born landed immigrants to Canada between 1981 and 2006

Census year	Original sample	Individuals born outside of Canada	Excluding Canadian citizens born outside of Canada
(1)	(2)	(3)	(4)
1981	5,036,955	778,890	769,640
1986	5,241,565	781,985	773,245
1991	5,624,565	910,105	846,490
1996	6,025,565	1,039,805	987,260
2001	6,080,915	1,102,565	1,043,145
2006	6,470,470	1,278,500	1,204,100

Source: The Canadian Censuses, 1981 and 2006.

Note: All samples sizes are unweighted and rounded according to the dissemination rules imposed by Statistics Canada. The final column considers only the number of foreign-born individuals regardless of whether a lingusitic distance measure could be assigned. In all of the Census years respondents are asked the question: "where was this person born?". The wording of this question does not change, and this is the question determining the sample highlighted in column (3). Information on citizenship is derived from the question: "of what country is this person a citizen?", and on this basis those who report being Canadian citizens by birth are excluded from the analysis to obtain the samples highlighted in column (4).

Table 1.4: Top five birth countries for immigrant stocks in Canada outside of Quebec: 1981-2006

Years	**All Immigrants**		$LD = 0$		$0 < LD < 95.22$		$95.22 <= LD <= 100.58$		$LD > 100.58$	
	Countries	%	Countries	%	Countries	%	Countries	%	Countries	%
1981	United Kingdom	25.3	United Kingdom	54.4	Italy	22.9	Greece	19.9	Hong Kong	15.8
	Italy	9.0	USA	15.8	West Germany	9.8	Hungary	17.9	Taiwan	15.0
	USA	8.0	Jamaica	4.7	Netherlands	9.8	USSR	10.1	China	14.6
	West Germany	4.3	Guyana	2.4	Poland	9.8	India	9.0	India	14.2
	Netherlands	4.1	Trinidad and Tobago	2.1	Portugal	9.0	Lebanon	4.8	Phillippines	13.9
1986	United Kingdom	22.5	United Kingdom	45.9	Italy	22.4	Greece	18.6	China	27.3
	Italy	8.4	USA	14.1	Poland	10.4	Hungary	17.0	Hong Kong	17.0
	USA	7.3	Jamaica	5.0	West Germany	10.3	India	9.7	Vietnam	14.4
	West Germany	4.3	Guyana	2.9	Netherlands	9.4	South Korea	6.5	India	13.5
	Poland	4.1	India	2.8	Portugal	9.2	USSR	8.6	Phillippines	12.0
1991	United Kingdom	18.5	United Kingdom	46.5	Italy	19.8	Greece	13.6	China	21.8
	Italy	7.3	USA	14.0	Poland	11.9	Hungary	12.3	Hong Kong	21.4
	USA	5.9	Jamaica	6.5	Germany	11.2	India	9.7	India	14.7
	Germany	4.5	Guyana	4.3	Portugal	10.0	South Korea	7.7	Vietnam	13.2
	Poland	4.4	Trinidad and Tobago	3.0	Netherlands	8.7	Lebanon	7.0	Phillippines	13.0
1996	United Kingdom	14.7	United Kingdom	40.9	Italy	17.4	Greece	11.4	Hong Kong	22.2
	Italy	6.0	USA	13.3	Poland	11.9	India	10.7	China	20.8
	Hong Kong	5.4	Jamaica	7.2	Germany	10.3	Hungary	10.1	India	13.5
	India	5.2	Guyana	4.9	Portugal	9.2	South Korea	9.7	Phillippines	12.1
	USA	5.1	Trinidad and Tobago	3.7	Netherlands	7.8	Lebanon	7.3	Vietnam	10.7
2001	United Kingdom	12.4	United Kingdom	38.4	Italy	15.3	Hong Kong	18.2	China	21.9
	China	6.5	USA	13.1	Poland	10.3	China	11.4	India	18.6
	India	6.3	Jamaica	7.6	Germany	9.1	India	8.5	Phillippines	15.9
	Italy	5.2	Guyana	5.3	Portugal	8.2	South Korea	8.3	Vietnam	10.1
	Hong Kong	4.9	Trinidad and Tobago	3.9	Netherlands	6.8	Greece	6.0	Hong Kong	8.2
2006	United Kingdom	10.6	United Kingdom	35.7	Italy	13.1	Hong Kong	14.8	China	25.3
	China	8.0	USA	13.3	Poland	9.0	China	11.9	India	20.5
	India	8.0	Jamaica	7.6	Germany	8.1	India	11.8	Phillippines	17.5
	Phillippines	5.4	Guyana	5.4	Portugal	7.3	South Korea	9.8	Sri Lanka	6.1
	Italy	4.3	Trinidad and Tobago	3.9	Netherlands	5.8	Greece	4.7	Vietnam	8.7

Notes: The Canadian Censuses, 1981 and 2006.
Source: Immigrant stocks include all immigrants at the time of the census. Language groups are based on mother tongue. The cut-off languages are Hungarian and Punjabi. The percentage represents the share of immigrants coming from the corresponding birth country speaking a language with the corresponding LD.

Table 1.5: Top five birth countries for immigrant flows in Canada outside of Quebec: 1981-2006

Years	**All Immigrants**		$LD = 0$		$0 < LD < 95.22$		$95.22 <= LD <= 100.58$		$LD > 100.58$	
	Countries	%	Countries	%	Countries	%	Countries	%	Countries	%
1981	United Kingdom	15.1	United Kingdom	33.7	Portugal	15.8	India	15.6	Vietnam	27.4
	USA	8.4	USA	18.4	Italy	7.5	Lebanon	11.7	Phillippines	14.8
	Vietnam	8.1	Jamaica	9.1	Netherlands	6.7	Laos	8.7	Hong Kong	12.6
	India	6.5	Guyana	5.6	Chile	6.6	Greece	7.2	India	12.3
	Phillippines	5.6	Phillippines	3.6	West Germany	6.4	South Korea	5.8	Taiwan	11.5
1986	United Kingdom	10.0	United Kingdom	24.6	Poland	24.4	India	15.1	China	23.9
	Vietnam	7.7	USA	17.8	Portugal	7.7	South Korea	11.9	Vietnam	23.2
	USA	7.5	Guyana	8.7	West Germany	7.4	Hungary	7.2	Hong Kong	16.5
	China	7.1	Jamaica	5.9	El Salvador	7.0	Lebanon	6.2	India	11.4
	India	6.5	Phillippines	5.2	Iran	5.4	Japan	5.4	Phillippines	10.4
1991	Hong Kong	10.8	United Kingdom	17.7	Poland	25.4	Lebanon	13.2	Hong Kong	29.1
	Poland	7.2	USA	14.0	Portugal	11.7	South Korea	10.2	China	18.1
	China	6.6	Jamaica	11.1	Iran	7.5	India	9.9	Phillippines	12.9
	India	6.3	Guyana	9.2	El Salvador	6.9	Ethiopia	8.1	India	11.6
	Phillippines	6.2	Trinidad and Tobago	6.2	USSR	3.6	Somalia	4.6	Vietnam	10.7
1996	Hong Kong	11.9	USA	12.8	Poland	16.0	South Korea	14.8	Hong Kong	25.5
	China	9.0	United Kingdom	12.6	Yugoslavia	8.8	India	12.1	China	19.7
	India	7.5	Jamaica	11.3	Iran	7.9	Lebanon	8.4	Phillippines	11.7
	Phillippines	7.5	Guyana	8.2	Romania	4.9	Iraq	7	India	11.0
	Sri Lanka	4.5	Phillippines	8.0	Pakistan	4.7	Phillippines	5.5	Sri Lanka	8.1
2001	China	13.7	USA	14.8	Pakistan	12.0	South Korea	16.2	China	34.9
	India	10.5	United Kingdom	11.7	Iran	11.0	Hong Kong	15.4	India	17.3
	Phillippines	6.3	India	9.0	Russian Federation	8.3	India	14.1	Phillippines	13.5
	Pakistan	4.8	Jamaica	7.8	Yugoslavia	6.2	China	9.3	Taiwan	10.5
	Hong Kong	4.5	Phillippines	6.8	Ukraine	5.8	Bangladesh	5.1	Sri Lanka	6.2
2006	China	14.9	USA	17.7	Pakistan	16.1	India	23	China	38.9
	India	13.7	United Kingdom	14.2	Iran	9.3	South Korea	17.5	India	20.2
	Phillippines	8.1	India	10.2	Russian Federation	6.7	China	7.3	Phillippines	18.7
	Pakistan	5.8	Phillippines	8.6	Colombia	6.0	Bangladesh	5.5	Sri Lanka	4.9
	USA	3.8	Jamaica	5.5	Romania	5.6	Iraq	4.4	Taiwan	3.3

Notes: The Canadian Censuses, 1981 and 2006.
Source: Immigrant flows enter Canada within five years in the given census. Language groups are based on mother tongue. The cut-off languages are Hungarian and Punjabi. The percentage represents the share of immigrants coming from the corresponding birth country speaking a language with the corresponding LD.

Table 1.6: Top five birth countries for stocks of immigrants in Quebec: 1981-2006

Years	**All Immigrants**		$LD = 0$		$0 < LD < 93.24$		$93.24 <= LD <= 97.03$		$LD > 97.03$	
	Countries	%	Countries	%	Countries	%	Countries	%	Countries	%
1981	Italy	17.0	France	28.6	Italy	31.6	Greece	34.4	Vietnam	17.0
	United Kingdom	8.1	Haiti	21.4	United Kingdom	15.1	West Germany	12.5	Hungary	14.6
	USA	7.4	USA	15.1	Portugal	7.4	USSR	10.8	Egypt	8.5
	France	7.0	Belgium	5.9	USA	7.3	Poland	8.3	Lebanon	7.5
	Greece	5.4	Morocco	5.9	Poland	4.7	Egypt	3.9	Taiwan	6.5
1986	Italy	16.0	France	22.8	Italy	30.5	Greece	35.6	Vietnam	23.7
	France	6.7	Haiti	20.4	United Kingdom	12.1	West Germany	12.8	China	12.1
	USA	6.6	USA	11.6	Portugal	7.3	USSR	9.1	Hungary	11.3
	United Kingdom	6.1	Morocco	5.7	USA	6.8	Poland	6.8	Cambodia	7.2
	Haiti	5.7	Belgium	4.9	Poland	5.4	Egypt	3.4	Lebanon	6.9
1991	Italy	13.3	France	29.9	Italy	28.2	Greece	32.6	Vietnam	17.4
	France	6.5	Haiti	13.8	United Kingdom	9.2	Germany	12.7	Lebonan	16.5
	Haiti	6.3	USA	9.3	Portugal	8.6	USSR	7.3	China	9.2
	USA	4.7	Morocco	7.3	USA	5.8	Poland	4.8	Cambodia	6.3
	Lebanon	4.4	Belgium	6.0	Poland	5.5	Lebanon	4.6	Hungary	6.0
1996	Italy	11.3	France	29.9	Italy	25.2	Greece	29.1	Vietnam	15.5
	Haiti	6.9	Haiti	14.3	Portugal	7.6	Germany	11.3	Lebonan	14.6
	France	6.7	USA	7.3	United Kingdom	7.0	Sri Lanka	7.1	China	10.6
	Lebanon	4.3	Morocco	6.7	USA	5.5	India	6.3	Phillippines	5.8
	USA	4.1	Belgium	5.2	Poland	5.0	Lebonan	4.7	Cambodia	5.4
2001	Italy	9.8	France	31.8	Italy	22.8	Greece	26.9	China	13.0
	France	7.1	Haiti	12.1	Portugal	7.4	Germany	11	Lebanon	12.9
	Haiti	6.8	Morocco	6.4	United Kingdom	5.8	Sri Lanka	9.5	Vietnam	12.8
	Lebaoan	4.1	USA	6.3	USA	5.0	India	7.4	Phillippines	6.6
	USA	3.6	Belgium	4.5	Poland	4.6	Bangladesh	5.8	Morocco	6.0
2006	Italy	7.7	France	32.2	Italy	18.4	Greece	26.1	China	15.4
	France	7.0	Haiti	11.4	Romania	6.9	Germany	9.8	Lebanon	11.2
	Haiti	6.7	Morocco	6.8	Portugal	5.9	Sri Lanka	9.7	Algeria	10.1
	China	4.6	USA	5.0	USA	4.8	India	8.3	Vietnam	10.0
	Lebanon	4.1	Algeria	4.6	United Kingdom	4.4	Bangladesh	8	Morocco	9.6

Notes: The Canadian Censuses, 1981 and 2006.
Source: Immigrant stocks include all immigrants at the time of the census. Language groups are based on mother tongue. The cut-off languages are Hebrew and Pashto. The percentage represents the share of immigrants coming from the corresponding birth country speaking a language with the corresponding LD.

Table 1.7: Top five birth countries for immigrant flows in Quebec: 1981-2006

Years	**All Immigrants**		$LD = 0$		$0 < LD < 93.24$		$93.24 <= LD <= 97.03$		$LD > 97.03$	
	Countries	%	Countries	%	Countries	%	Countries	%	Countries	%
1981	Haiti	12.9	Haiti	40.1	USA	11.7	Greece	23	Vietnam	28.8
	Vietnam	8.5	France	18.3	United Kingdom	10.7	Lebanon	11.3	Cambodia	12.8
	France	6.2	USA	4.8	Italy	8.4	Switzerland	11.3	Lebonan	11.2
	USA	5.9	Vietnam	4.4	Portugal	8.2	Israel	8.2	Laos	9.4
	Lebanon	4.7	Morocco	3.8	Chile	6.9	West Germany	7.7	Phillippines	5.3
1986	Vietnam	10.3	Haiti	32.1	USA	9.5	Sri Lanka	18	Vietnam	34.5
	Haiti	10.2	France	14.2	El Salvador	8.4	Greece	11.4	Cambodia	13.1
	USA	4.7	Vietnam	6.7	Poland	7.6	Turkey	9.3	China	9.2
	France	4.6	Morocco	4.7	Iran	6.3	West Germany	8.3	Lebanon	6.9
	Cambodia	4.2	Cambodia	3.6	United Kingdom	5.9	Israel	7.5	Hong Kong	4.2
1991	Lebanon	12.6	France	22.6	El Salvador	9.5	Sri Lanka	18.8	Lebanon	26.5
	Haiti	6.7	Haiti	14.1	Poland	7.2	Lebonan	12.8	Vietnam	8.6
	France	4.3	Lebanon	12.9	Portugal	7.0	India	8.8	China	7.7
	El Salvador	3.6	Morocco	6.3	Iran	6.2	Greece	8.5	Hong Kong	6.0
	Vietnam	3.3	USA	3.8	USA	5.7	Bangladesh	6.9	Syria	5.2
1996	Haiti	6.9	France	30.8	Romania	9.5	Sri Lanka	28.8	Lebanon	16.6
	Lebonan	6.7	Haiti	12.8	Salvador	5.5	Bangladesh	18.3	China	12.8
	France	6.6	China	5.4	Iran	4.7	India	15.2	Vietnam	8.3
	China	5.3	Algeria	5.0	Peru	4.7	Lebanon	5.5	Phillippines	7.5
	Romania	3.7	Lebanon	4.1	India	4.3	Pakistan	3.9	Hong Kong	7.4
2001	France	8.4	France	34.4	Romania	9.3	Sri Lanka	21.7	China	21.1
	China	8.3	Algeria	8.8	Russia Federation	7.4	India	17.9	Algeria	14.7
	Algeria	8.0	China	6.7	Iran	5.1	Bangladesh	16.2	Morocco	11.7
	Haiti	4.8	Haiti	6.2	Pakistan	4.7	Algeria	13.7	Phillippines	6.2
	Morocco	4.8	Congo,D. Rep.	5.0	USA	4.6	Pakistan	6.7	Lebanon	5.7
2006	China	9.5	France	34.1	Romania	17.1	Bangladesh	19.2	China	23.4
	Algeria	8.0	Haiti	8.5	Colombia	13.2	India	16.7	Algeria	18.5
	France	7.7	Algeria	7.1	Mexico	6.4	Sri Lanka	16.7	Morocco	16.8
	Morocco	7.1	Morocco	6.1	Pakistan	4.7	Pakistan	9.9	Lebanon	6.4
	Romania	6.6	China	5.6	Bulgaria	4.3	Ukraine	5.2	Tunisia	4.0

Notes: The Canadian Censuses, 1981 and 2006.
Source: Immigrant flows enter Canada within five years in the given census. Language groups are based on mother tongue. The cut-off languages are Hebrew and Pashto. The percentage represents the share of immigrants coming from the corresponding birth country speaking a language with the corresponding LD.

Table 1.8: The evolution of LD in Canada outside of Quebec: 1981-2006

Mother Tongue						
Entry Cohorts	Census Years					
	1981	1986	1991	1996	2001	2006
1976-1980	54.26	49.32	56.74	56.56	57.94	58.72
1981-1985	–	57.88	63.90	64.12	64.70	66.45
1986-1990	–	–	72.06	71.08	71.91	72.90
1991-1995	–	–	–	77.63	76.71	77.53
1996-2000	–	–	–	–	77.21	79.11
2001-2005	–	–	–	–	–	79.35
Home Language						
Entry Cohorts	Census Years					
	1981	1986	1991	1996	2001	2006
1976-1980	47.91	37.55	36.67	34.61	33.88	32.46
1981-1985	–	48.27	46.71	43.81	40.71	39.84
1986-1990	–	–	59.85	53.62	49.79	47.43
1991-1995	–	–	–	64.87	58.75	55.28
1996-2000	–	–	–	–	64.15	58.53
2001-2005	–	–	–	–	–	62.23

Source: The Canadian Censuses, 1981 and 2006.
Notes: Immigrants flows enter Canada within five years in the given census. The LD for home language and mother tongue are based on the distance to English.

Table 1.9: The evolution of LD in Quebec: 1981-2006

Mother Tongue						
Entry Cohorts	Census Years					
	1981	1986	1991	1996	2001	2006
1976-1980	67.90	56.35	67.29	65.32	67.23	68.89
1981-1985	–	62.46	68.77	70.52	70.49	71.19
1986-1990	–	–	74.75	72.16	73.32	75.20
1991-1995	–	–	–	73.39	74.13	74.02
1996-2000	–	–	–	–	70.79	71.39
2001-2005	–	–	–	–	–	72.40
Home Language						
Entry Cohorts	Census Years					
	1981	1986	1991	1996	2001	2006
1976-1980	63.67	51.31	52.14	50.54	50.12	49.75
1981-1985	–	60.65	54.37	54.51	52.82	51.35
1986-1990	–	–	62.41	58.42	56.05	55.02
1991-1995	–	–	–	60.62	57.85	55.63
1996-2000	–	–	–	–	54.62	51.73
2001-2005	–	–	–	–	–	55.00

Source: The Canadian Censuses, 1981 and 2006.
Notes: Immigrants flows enter Canada within five years in the given census. The LD for home language and mother tongue are based on the distance to French.

Table 1.10: Age distribution of immigrants by arrival cohort and language group in Canada outside of Quebec: 1981-2006

	Recent Immigrant Cohorts					
Age Groups	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
Age <= 13	0.217	0.185	0.192	0.169	0.191	0.187
Age > 13 and < 21	0.126	0.117	0.107	0.117	0.118	0.111
Age >= 21 and <= 64	0.612	0.646	0.662	0.669	0.659	0.669
Age > 64	0.046	0.052	0.039	0.045	0.032	0.033
Age By Language Groups						
LD = 0						
Age <= 13	0.274	0.234	0.249	0.250	0.288	0.287
Age > 13 and < 21	0.127	0.121	0.110	0.114	0.122	0.116
Age >= 21 and <= 64	0.565	0.601	0.599	0.603	0.563	0.573
Age > 64	0.035	0.044	0.043	0.033	0.027	0.024
0 < LD < 95.22						
Age <= 13	0.190	0.188	0.210	0.189	0.197	0.189
Age > 13 and < 21	0.105	0.101	0.101	0.116	0.127	0.114
Age >= 21 and <= 64	0.652	0.670	0.661	0.665	0.649	0.666
Age > 64	0.053	0.041	0.028	0.030	0.027	0.031
95.22 <= LD < 100.58						
Age <= 13	0.153	0.136	0.160	0.146	0.175	0.176
Age > 13 and < 21	0.115	0.101	0.108	0.126	0.127	0.121
Age >= 21 and <= 64	0.684	0.709	0.698	0.690	0.663	0.674
Age > 64	0.049	0.055	0.034	0.037	0.036	0.029
LD >= 100.58						
Age <= 13	0.167	0.127	0.148	0.127	0.141	0.139
Age > 13 and < 21	0.143	0.128	0.110	0.117	0.101	0.099
Age >= 21 and <= 64	0.634	0.671	0.697	0.695	0.720	0.721
Age > 64	0.056	0.074	0.046	0.061	0.038	0.042

Source: The Canadian Censuses, 1981 and 2006.

Notes: The language cut-offs are Hungarian and Punjabi. Language groups are based on the distance of mother tongue to English. Cohorts are defined as having entered Canada within five years of the census year.

Table 1.11: Gender distribution of immigrants by arrival cohort and language group in Canada outside of Quebec: 1981-2006

	Recent Immigrant Cohorts					
Gender Groups	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
Female	0.520	0.522	0.507	0.528	0.518	0.525
Male	0.480	0.478	0.493	0.472	0.482	0.475
Gender By Language Groups						
LD = 0						
Female	0.523	0.531	0.529	0.534	0.514	0.515
Male	0.477	0.469	0.471	0.466	0.486	0.485
0 < LD < 95.22						
Female	0.514	0.496	0.490	0.513	0.508	0.524
Male	0.486	0.504	0.511	0.487	0.492	0.477
95.22 <= LD < 100.58						
Female	0.527	0.514	0.475	0.517	0.524	0.522
Male	0.474	0.486	0.526	0.483	0.476	0.478
LD >= 100.58						
Female	0.516	0.533	0.516	0.535	0.527	0.534
Male	0.484	0.467	0.484	0.465	0.474	0.466

Source: The Canadian Censuses, 1981 and 2006.
Notes: The language cut-offs are Hungarian and Punjabi. Language groups are based on the distance of mother tongue to English. Cohorts are defined as having entered Canada within five years of the census year.

Table 1.12: Education of immigrants by arrival cohort and language group in Canada outside of Quebec: 1981-2006

	Recent Immigrant Cohorts					
Education Groups	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
High Sch. or Less Than High Sch.	0.554	0.560	0.538	0.500	0.360	0.280
Voc. Cert. or Below Bachelor Degr.	0.279	0.250	0.258	0.255	0.223	0.237
Bachelor Degr. Or Above	0.167	0.191	0.204	0.245	0.418	0.483
Education By Language Groups						
LD = 0						
High Sch. or Less Than High Sch.	0.462	0.504	0.475	0.440	0.360	0.258
Voc. Cert. or Below Bachelor Degr.	0.354	0.301	0.330	0.336	0.292	0.304
Bachelor Degr. Or Above	0.184	0.196	0.195	0.224	0.348	0.438
0 <= LD < 95.22						
High Sch. or Less Than High Sch.	0.540	0.472	0.522	0.407	0.303	0.253
Voc. Cert. or Below Bachelor Degr.	0.318	0.318	0.294	0.313	0.253	0.240
Bachelor Degr. Or Above	0.142	0.210	0.184	0.280	0.443	0.507
95.22 <= LD < 100.58						
High Sch. or Less Than High Sch.	0.583	0.533	0.536	0.470	0.401	0.276
Voc. Cert. or Below Bachelor Degr.	0.223	0.231	0.227	0.222	0.198	0.216
Bachelor Degr. Or Above	0.195	0.236	0.238	0.308	0.401	0.508
LD >= 100.58						
High Sch. or Less Than High Sch.	0.688	0.717	0.589	0.577	0.383	0.314
Voc. Cert. or Below Bachelor Degr.	0.161	0.130	0.198	0.202	0.183	0.217
Bachelor Degr. Or Above	0.152	0.153	0.214	0.221	0.434	0.469

Source: The Canadian Censuses, 1981 and 2006.

Notes: The language cut-offs are Hungarian and Punjabi. Language groups are based on the distance of mother tongue to English. Cohorts are defined as having entered Canada within five years of the census year.

Table 1.13: Employment of immigrants by arrival cohort, by gender, and by language group in Canada outside of Quebec: 1981-2006

	Recent Immigrant Cohorts					
Employment Groups	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
Employment	0.614	0.573	0.624	0.467	0.540	0.534
Female Employment	0.466	0.455	0.526	0.374	0.429	0.425
Male Employment	0.782	0.708	0.729	0.578	0.666	0.662
Employment By Language Groups						
LD = 0						
Employment	0.663	0.616	0.683	0.566	0.629	0.600
Female Employment	0.504	0.495	0.596	0.475	0.527	0.497
Male Employment	0.844	0.761	0.787	0.679	0.741	0.712
0 < LD < 95.22						
Employment	0.592	0.559	0.610	0.470	0.530	0.503
Female Employment	0.409	0.409	0.475	0.341	0.390	0.377
Male Employment	0.785	0.707	0.738	0.611	0.682	0.648
LD >= 95.22 and < 100.58						
Employment	0.547	0.503	0.558	0.420	0.469	0.486
Female Employment	0.382	0.360	0.423	0.298	0.358	0.365
Male Employment	0.735	0.654	0.677	0.554	0.598	0.628
LD >= 100.58						
Employment	0.587	0.554	0.622	0.437	0.551	0.558
Female Employment	0.485	0.468	0.547	0.366	0.458	0.466
Male Employment	0.708	0.659	0.706	0.526	0.658	0.672

Source: The Canadian Censuses, 1981 and 2006.

Notes: The languages cut-offs are Hungarian and Punjabi. Language groups are based on the distance of mother tongue to English. Cohorts are defined as having entered Canada within five years of the census year. The sample includes immigrants 21-64 years of age who held full time jobs for at least 40 weeks in the census year and reported positive working hours in that year.

Table 1.14: The average logarithm of real weekly wages of immigrants by arrival cohort, by gender, and by language group in Canada outside of Quebec: 1981-2006

	Recent Immigrant Cohorts					
Average Log of Wages	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
All	6.37	6.20	6.26	6.12	6.29	6.27
Females	6.08	5.98	6.09	5.98	6.11	6.13
Males	6.57	6.36	6.38	6.23	6.42	6.38
Average Log of Wages By Language Groups						
LD = 0						
All	6.48	6.29	6.36	6.24	6.42	6.47
Females	6.15	6.06	6.20	6.10	6.23	6.32
Males	6.71	6.47	6.51	6.37	6.57	6.58
0 < LD < 95.22						
All	6.42	6.24	6.27	6.17	6.35	6.30
Females	6.07	5.97	6.02	5.98	6.15	6.14
Males	6.62	6.39	6.42	6.29	6.47	6.41
95.22 >= LD < 100.58						
All	6.28	6.17	6.21	6.11	6.24	6.22
Females	6.00	5.95	6.03	5.92	6.08	6.04
Males	6.45	6.29	6.30	6.22	6.36	6.34
LD >= 100.58						
All	6.18	6.04	6.19	6.03	6.20	6.19
Females	6.01	5.88	6.07	5.93	6.03	6.08
Males	6.32	6.17	6.30	6.12	6.34	6.29

Source: The Canadian Censuses, 1981 and 2006.

Notes: The language cut-offs are Hungarian and Punjabi. Language groups are based on the distance of mother tongue to English. Cohorts are defined as having entered Canada within five years of the census year. The sample includes immigrants 21-64 years of age who held full time jobs for at least 40 weeks in the census year and reported positive working hours in that year.

Table 1.15: Age distribution of immigrants by arrival cohort and language group in Quebec: 1981-2006

	Recent Immigrant Cohorts					
Age Groups	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
Age <= 13	0.225	0.206	0.201	0.194	0.211	0.197
Age > 13 and < 21	0.122	0.115	0.110	0.107	0.094	0.078
Age >= 21 and <= 64	0.611	0.638	0.650	0.668	0.676	0.709
Age > 64	0.042	0.041	0.039	0.031	0.020	0.017
Age By Language Groups						
LD = 0						
Age <= 13	0.262	0.281	0.291	0.321	0.327	0.301
Age > 13 and < 21	0.114	0.123	0.098	0.095	0.082	0.070
Age >= 21 and <= 64	0.583	0.560	0.578	0.565	0.579	0.618
Age > 64	0.041	0.035	0.033	0.020	0.011	0.011
0 < LD < 93.24						
Age <= 13	0.218	0.182	0.195	0.188	0.190	0.196
Age > 13 and < 21	0.122	0.103	0.108	0.111	0.106	0.096
Age >= 21 and <= 64	0.619	0.675	0.663	0.673	0.679	0.691
Age > 64	0.041	0.040	0.034	0.028	0.026	0.017
93.24 <= LD < 97.03						
Age <= 13	0.192	0.135	0.153	0.143	0.142	0.146
Age > 13 and < 21	0.100	0.083	0.099	0.091	0.077	0.080
Age >= 21 and <= 64	0.640	0.714	0.689	0.717	0.750	0.741
Age > 64	0.067	0.069	0.060	0.050	0.032	0.033
LD >= 97.03						
Age <= 13	0.197	0.163	0.179	0.139	0.167	0.146
Age > 13 and < 21	0.144	0.134	0.116	0.109	0.088	0.061
Age >= 21 and <= 64	0.624	0.660	0.668	0.720	0.728	0.778
Age > 64	0.035	0.043	0.037	0.032	0.016	0.015

Source: The Canadian Censuses, 1981 and 2006.

Notes: The language cut-offs are Hebrew and Pashto. Language groups are based on the distance of mother tongue to French. Cohorts are defined as having entered Canada within five years of the census year.

Table 1.16: Gender distribution of immigrants by arrival cohort and language group in Quebec: 1981-2006

	Recent Immigrant Cohorts					
Gender Groups	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
Female	0.505	0.503	0.486	0.509	0.510	0.508
Male	0.495	0.497	0.514	0.491	0.491	0.492
Gender By Language Groups						
LD = 0						
Female	0.518	0.514	0.482	0.521	0.512	0.505
Male	0.482	0.486	0.518	0.479	0.488	0.495
0 < LD < 93.24						
Female	0.522	0.497	0.493	0.513	0.526	0.524
Male	0.479	0.504	0.507	0.487	0.474	0.477
93.24 <= LD < 97.03						
Female	0.486	0.466	0.479	0.488	0.473	0.486
Male	0.514	0.534	0.521	0.512	0.527	0.515
LD >= 97.03						
Female	0.470	0.509	0.474	0.495	0.497	0.494
Male	0.530	0.491	0.526	0.505	0.503	0.506

Source: The Canadian Censuses, 1981 and 2006.
Notes: The language cut-offs are Hebrew and Pashto. Language groups are based on the distance of mother tongue to French. Cohorts are defined as having entered Canada within five years of the census year.

Table 1.17: Education of immigrants by arrival cohort and language group in Quebec: 1981-2006

	Recent Immigrant Cohorts					
Education Groups	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
High Sch. or Less Than High Sch.	0.562	0.581	0.533	0.474	0.335	0.224
Voc. Cert. or Below Bachelor Degr.	0.270	0.226	0.244	0.259	0.272	0.287
Bachelor Degr. or Above	0.169	0.193	0.223	0.267	0.393	0.489
Education By Language Groups						
LD = 0						
High Sch. or Less Than High Sch.	0.555	0.565	0.397	0.306	0.193	0.151
Voc. Cert. or Below Bachelor Degr.	0.324	0.267	0.343	0.383	0.374	0.347
Bachelor Degr. Or Above	0.121	0.168	0.260	0.311	0.433	0.502
0 < LD < 93.24						
High Sch. or Less Than High Sch.	0.509	0.538	0.549	0.466	0.366	0.246
Voc. Cert. or Below Bachelor Degr.	0.277	0.232	0.257	0.265	0.262	0.254
Bachelor Degr. or Above	0.215	0.230	0.193	0.270	0.372	0.500
93.24 <= LD < 97.03						
High Sch. Dropouts or Grad.	0.615	0.606	0.638	0.669	0.551	0.471
Voc. Cert. or Below Bachelor Degr.	0.236	0.238	0.194	0.155	0.165	0.198
Bachelor Degr. or Above	0.150	0.157	0.168	0.176	0.285	0.333
LD >= 97.03						
High Sch. or Less Than High Sch.	0.629	0.663	0.519	0.472	0.288	0.195
Voc. Cert. or Below Bachelor Degr.	0.203	0.162	0.207	0.229	0.255	0.287
Bachelor Degr. or Above	0.168	0.174	0.274	0.299	0.457	0.518

Source: The Canadian Censuses, 1981 and 2006.

Notes: The language cut-offs are Hebrew and Pashto. Language groups are based on the distance of mother tongue to French. Cohorts are defined as having entered Canada within five years of the census year.

Table 1.18: Employment of immigrants by arrival cohort, by gender, and by language group in Quebec: 1981-2006

	Recent Immigrant Cohorts					
Employment Groups	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
Employment	0.555	0.516	0.534	0.427	0.491	0.474
Female Employment	0.417	0.406	0.428	0.345	0.390	0.377
Male Employment	0.696	0.627	0.633	0.511	0.593	0.573
Employment By Language Groups						
LD = 0						
Employment	0.534	0.516	0.558	0.500	0.593	0.587
Female Employment	0.410	0.415	0.444	0.419	0.496	0.502
Male Employment	0.663	0.623	0.659	0.578	0.678	0.665
0 < LD < 93.24						
Employment	0.622	0.522	0.568	0.424	0.493	0.474
Female Employment	0.461	0.392	0.456	0.334	0.392	0.369
Male Employment	0.796	0.647	0.681	0.520	0.609	0.594
93.24 <= LD < 97.03						
Employment	0.514	0.483	0.546	0.418	0.472	0.436
Female Employment	0.353	0.312	0.394	0.298	0.319	0.325
Male Employment	0.674	0.622	0.668	0.529	0.608	0.544
LD >= 97.03						
Employment	0.501	0.520	0.486	0.391	0.428	0.424
Female Employment	0.383	0.447	0.398	0.329	0.342	0.336
Male Employment	0.607	0.600	0.564	0.454	0.515	0.512

Source: The Canadian Censuses, 1981 and 2006.

Notes: The language cut-offs are Hebrew and Pashto. Language groups are based on the distance of mother tongue to French. Cohorts are defined as having entered Canada within five years of the census year. The sample includes immigrants 21-64 years of age who held full time jobs for at least 40 weeks in the census year and reported positive working hours in that year.

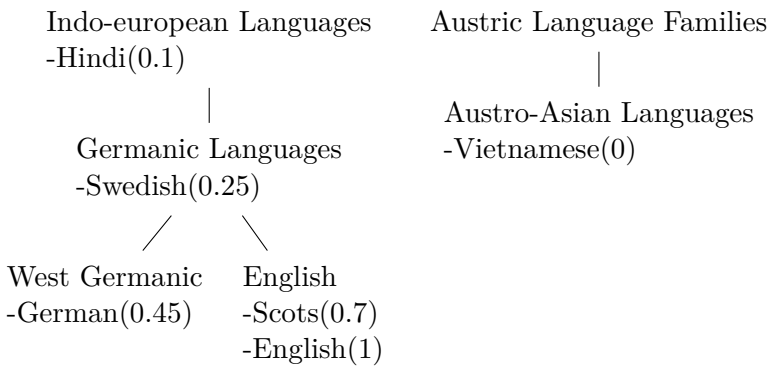
Table 1.19: The average logarithm of real weekly wages of immigrants by arrival cohort, by gender, and by language group in Quebec: 1981-2006

	Recent Immigrant Cohorts					
Average Wages	**1976-80**	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
All	6.26	6.04	6.07	5.97	6.17	6.15
Females	6.03	5.83	5.89	5.84	6.03	6.02
Males	6.41	6.18	6.18	6.07	6.26	6.24
Average Wages By Language Groups						
LD = 0						
All	6.23	6.07	6.20	6.18	6.36	6.29
Females	6.00	5.85	6.10	6.05	6.22	6.15
Males	6.38	6.23	6.25	6.28	6.45	6.38
0 < LD < 93.24						
All	6.34	6.09	6.09	5.97	6.13	6.15
Females	6.07	5.82	5.88	5.81	5.97	6.01
Males	6.51	6.25	6.24	6.08	6.24	6.25
93.24 <= LD < 97.03						
All	6.34	6.02	6.08	5.86	6.10	6.08
Females	6.05	5.86	5.91	5.60	6.04	5.85
Males	6.49	6.08	6.16	6.00	6.12	6.22
LD >= 97.03						
All	6.13	5.93	5.99	5.91	6.12	6.08
Females	5.97	5.81	5.81	5.83	5.96	5.96
Males	6.23	6.03	6.11	5.97	6.22	6.15

Source: The Canadian Censuses, 1981 and 2006.

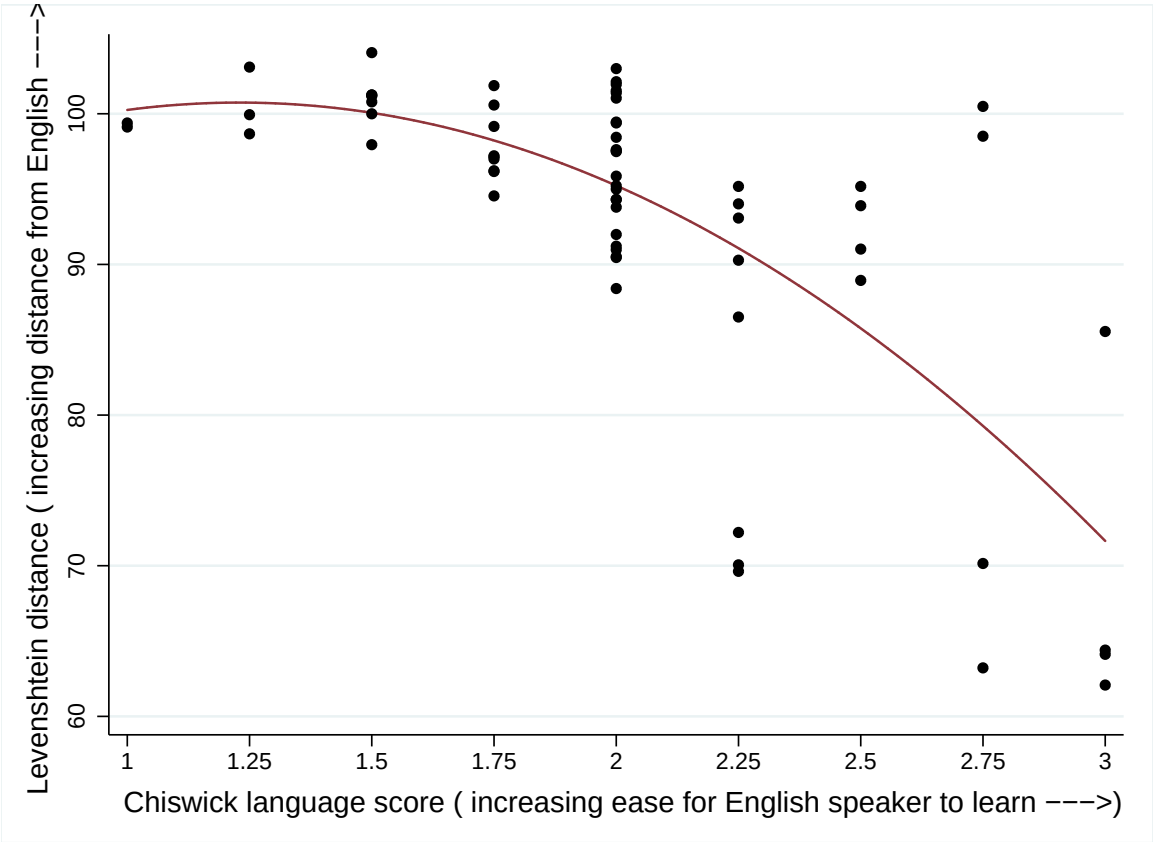
Notes: The language cut-offs are Hebrew and Pashto. Language groups are based on the distance of mother tongue to French. Cohorts are defined as having entered Canada within five years of the census year. The sample includes immigrants 21-64 years of age who held full time jobs for at least 40 weeks in the census year and reported positive working hours in that year.

Figure 1.1: An illustration for the A-P Language tree index



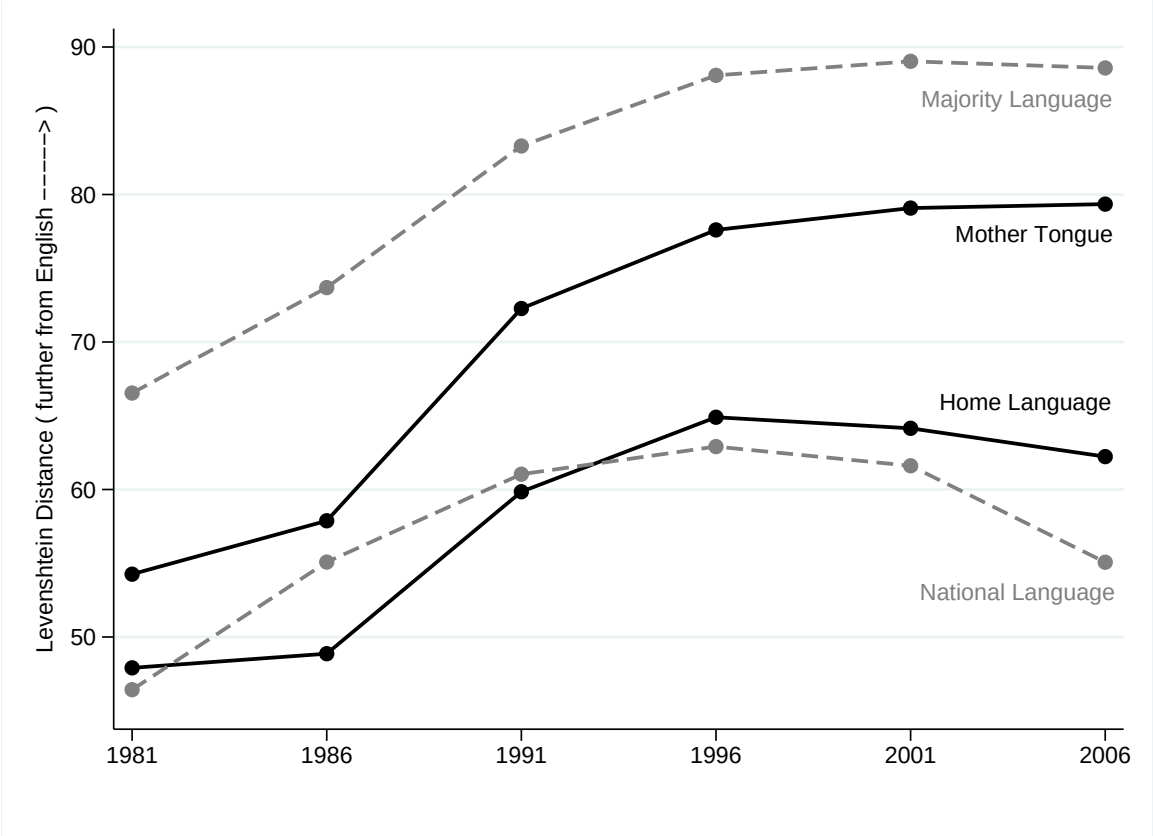
Notes: Family trees are derived by using the Ethnologue and the assigned weights based on the number of shared brances between English and the other language are used to derive the linguistic distance to English.

Figure 1.2: The relation between the Levenshtein Distance to English and the Chiswick-Miller score



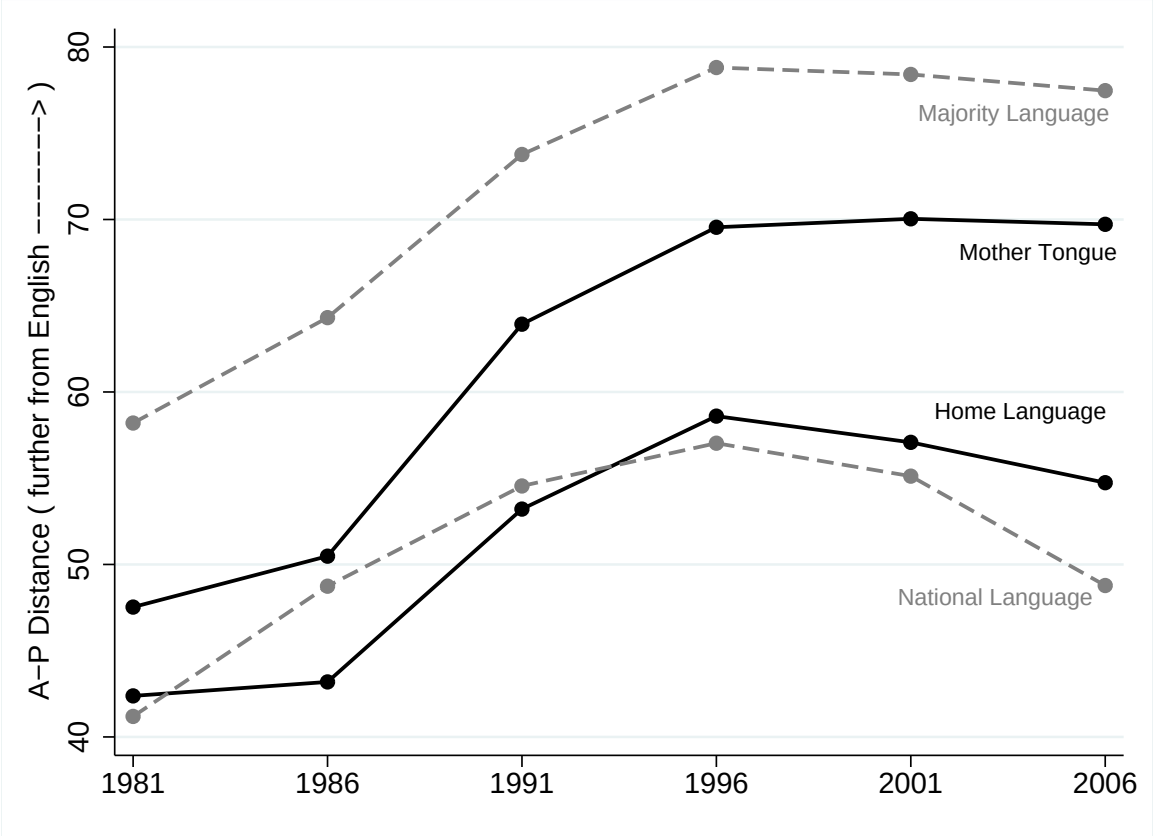
Notes: The number of observation is 65 and determined by the languages for which a Chiswick-Miller score is available.

Figure 1.3: Average Levenshtein Distance to English: Recent immigrants in Canada outside of Quebec



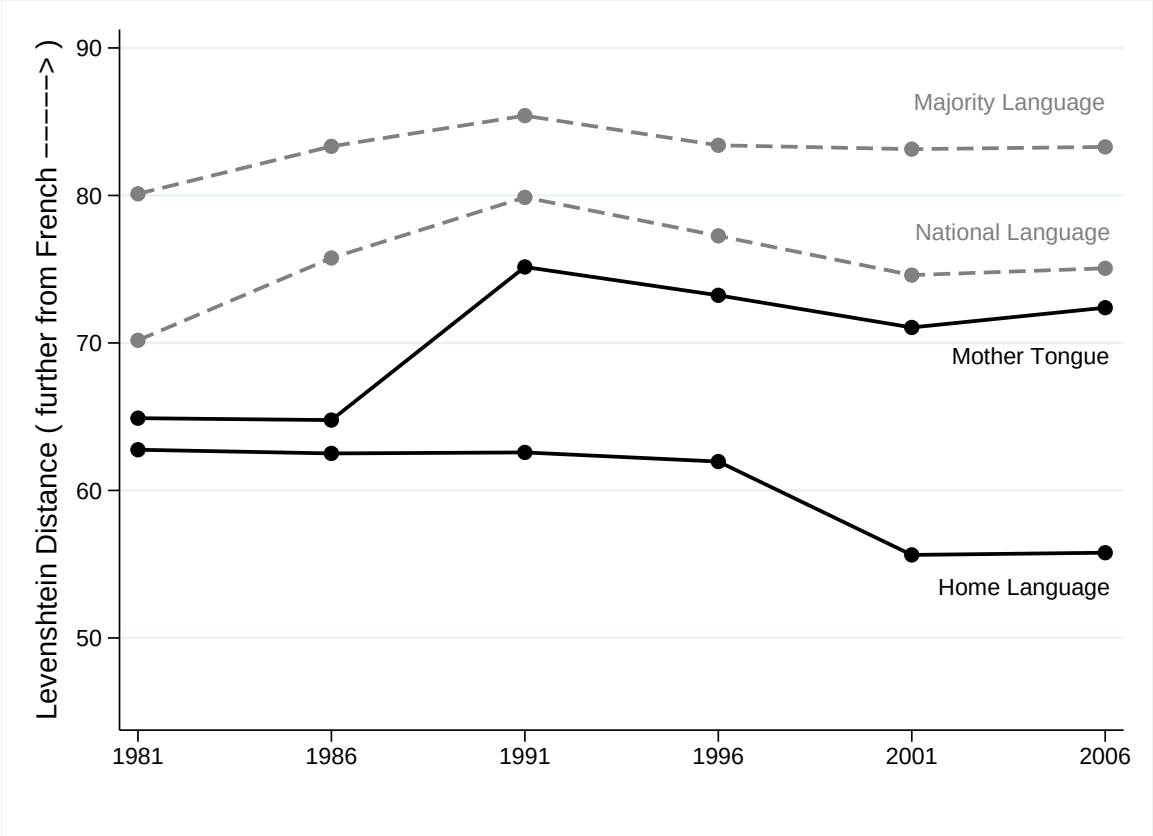
Source: The Canadian Censuses, 1981 and 2006.
Notes: Census years are indicated along the horizontal axis. Recent immigrants refers to the cohort of immigrants who enter Canada within five years of the census year. The LD representing the distance of the reported language to English is assigned to each immigrant based on the reported language. Mother tongue is the language first learned in childhood and still understood at the time of the census. Home language is the language spoken most often at home. The majority language is the the language spoken by the majority in the country of origin. The national language is the language of mass media, public communication, and the education system in the source country.

Figure 1.4: Average A-P Index to English: Recent immigrants in Canada outside of Quebec



Source: The Canadian Censuses, 1981 and 2006.
Notes: Census years are indicated along the horizontal axis. Recent immigrants refers to the cohort of immigrants who enter Canada within five years of the census year. The A-P Index representing the distance of the reported language to English is assigned to each immigrant based on the reported language. Mother tongue is the language first learned in childhood and still understood at the time of the census. Home language is the language spoken most often at home. The majority language is the the language spoken by the majority in the country of origin. The national language is the language of mass media, public communication, and the education system in the source country.

Figure 1.5: Average Levenshtein Distance to French: Recent immigrants in Quebec



Source: The Canadian Censuses, 1981 and 2006.
Notes: Census years are indicated along the horizontal axis. Recent immigrants refers to the cohort of immigrants who enter Canada within five years of the census year. The LD representing the distance of the reported language to French is assigned to each immigrant based on the reported language. Mother tongue is the language first learned in childhood and still understood at the time of the census. Home language is the language spoken most often at home. The majority language is the the language spoken by the majority in the country of origin. The national language is the language of mass media, public communication, and the education system in the source country.

Table A1.1: The "long-form" mother tongue and home language questions

Census Years	MOTHER TONGUE No	Wording	Check Boxes	HOME LANGUAGE No	Wording	Check Boxes
1981	Q6	What is the language you first learned in childhood and still understand?	English, French, German, Italian, Ukranian, Other (Other)	Q28	What language do you yourself speak at home now? (If more than one language, language, which language do you speak most often?	English, French, German, Italian, Ukranian, Other (Other)
1986	Q6	What is the language you first learned in childhood and still understand?	English, French, Italian, German, Ukranian, Other (Specify)	Q18	What language do you yourself speak at home now? (If more than one language, language, which language do you speak most often?	English, French, Italian, Chinese, German, Other (Specify)
1991	Q10	What is the language this person first learned at home in childhood and still understands? (If this person no longer understands the first language learned, indicate the second language learned!	English, French, Other (Specify)	Q9	What language does this person speak most often at home?	English, French, Other (Specify)
1996	Q12	What is the language this person first learned at home in childhood and still understands? (If this person no longer understands the first language learned, indicate the second language learned!	English, French, Other (Specify)	Q11	What language does this person speak most often at home?	English, French, Other (Specify)
2001	Q16	What is the language this person first learned at home in childhood and still understands? (If this person no longer understands the first language learned, indicate the second language learned!	English, French, Other (Specify)	Q15	a) What language does this person speak most often at home? b) Does this person speak any other languages on a regular basis at home.	English, French, Other (Specify) No, Yes-English, Yes-French, Yes-Other (Specify)
2006	Q16	What is the language this person first learned at home in childhood and still understands? (If this person no longer understands the first language learned, indicate the second language learned!	English, French, Other (Specify)	Q15	a) What language does this person speak most often at home? b) Does this person speak any other languages on a regular basis at home.	English, French, Other (Specify) No, Yes-English, Yes-French, Yes-Other (Specify)

Source: Canadian Census Long-Form Questionnaires between 1981 and 2006.
Notes: The additional instructions for mother tongue question from 1991 onward were provided in the 1986 Census Guide.

Chapter 2

Immigrant versus Native Men? Substitutability and the Role of Linguistic Diversity in Canada

2.1 Introduction

Canada is a major immigrant receiving country. According to Statistics Canada (2008), immigrants made up more than one-fifth of the country's employed population in 2006. Further, the annual number of newcomers almost tripled over the last three decades, from 88,276 in 1984 to 186,881 in 2010 (Beach, Green, and Worswick, 2011). Economic theory predicts that such an increase in labour supply should decrease host country wages in the short-run if there exists perfect substitutability between immigrants and natives.

There exists a recent literature examining immigrant-native substitutability (Aydemir and Borjas, 2007; Borjas, Grogger, and Hanson, 2008; Ottaviano and Peri, 2008, 2012). Ottaviano and Peri were the first to tackle this question. Categorizing individuals by skill type, namely by education and years of labour market experience, Ottaviano and Peri (2008, 2012) found the elasticity of substitution between immigrants and natives to be around 20. Their findings suggest that immigration affects the wages of foreign-born workers more negatively than those of native-born workers in the United States. Borjas et

al. (2008) criticized the findings of Ottaviano and Peri and suggested that the findings were mainly driven by the inclusion of a group of young workers that are not fully attached to the labour market, and an unconventional definition of wages. Using the same American data set and addressing these issues, Borjas et al. (2008) find evidence of perfect substitutability between immigrants and natives. Aydemir and Borjas (2007) examine the impact of immigration on wages in the United States and Canada as well as the impact of emigration on wages in Mexico. For all three countries, their results indicate perfect immigrant-native substitutability.

One potentially important factor that has not to date been taken into account in the existing estimates is the changing language skills of immigrants. Since the introduction of the point-based selection system for economic class immigrants in 1967, immigrants to Canada are less likely to come from North America and Western Europe, and more likely to have their origins in Asia and the Middle East. These changes have been accompanied by a significant increase in linguistic diversity, as documented in Chapter 1. Language skills have been shown to play an important role in labour market outcomes. Boyd and Cao (2009) and Chiswick and Miller (1995) document a positive association between language skills and immigrant earnings, and Aydemir and Skuterud (2005) show that a significant fraction of the gap in entry earnings between recent immigrants and the Canadian-born is explained by the changes in place of birth and in language skills. This chapter aims to examine the impact of the increasing immigrant inflows with diversified language skills on Canadian wages. Specifically, I re-examine immigrant-native substitutability in Canada by extending the models in Borjas (2003) and Ottaviano and Peri (2008, 2012) by allowing immigrants' skill types to vary not only by education and labour market experience, but language skills as well.[1]

The analysis is based on the 20% micro-data files of the Canadian Censuses conducted every five years between 1981 and 2006, the so called "long form" data. The Census does not provide direct measures of language skills but does offer information on mother

[1]Although Borjas (2003) does not explicitly estimate immigrant-native substitutability, he emphasizes the importance of taking into consideration market equalizing forces across local labour markets. He suggests that the impact of the immigrant-induced increase in labour on wages should be examined by using variations across skill types at the national level, rather than using variations across local labour markets.

tongue and home language.[2] I use the Levenshtein Distance Index (LD), a measure of linguistic distance based on the phonetic differences between two languages to proxy for language skills. I construct the LD based on both mother tongue and home language to Canada's official languages; distance to English for immigrants living outside of Quebec, and distance to French for immigrants living in Quebec.

Earlier research examining the impact of language skills on immigrants' labour market performance has been based upon mother tongue as well as home language. Since mother tongue does not change, it cannot capture language skill accumulation over time. In contrast, home language better reflects the accumulation of language skills in official languages, which has been shown to be important for labour market outcomes. On this basis, home language is a better proxy for language skills and the preferred specifications of my estimations are based on the distance of home language to English or French. As a robustness check, I also offer estimates using mother tongue as the basis for categorizing immigrant language skills. In this chapter, immigrants are categorized into three language skill groups: high, medium, and low based on the distance of the reported language to English or to French.

There are three notable findings in this chapter. First, using the LD based on home language leads to estimates of the elasticity of substitution between 50 and infinity or perfect immigrant-native substitutability for Quebec, but an estimated elasticity of substitution of around 30, suggesting imperfect substitutability between immigrants and natives outside of Quebec. When language groups are analyzed separately, it is found that the estimates for low language-skilled immigrants refer to imperfect substitutability whereas the estimates for medium and high language-skilled immigrants refer to perfect substitutability, suggesting that imperfect substitutability between immigrants and natives is mainly driven by low language-skilled immigrants outside of Quebec. For example, the degree of substitutability between low language-skilled immigrants in English and natives is estimated to range between 20 and 30 for Canada outside of Quebec. Second, mother tongue-based specifications provide even stronger evidence for imperfect substitutability in Canada outside

[2]Mother tongue is defined as the first language learned in the childhood and still understood at the time of the Census. Home language refers to the language spoken at home most often at the time of the Census.

Quebec, suggesting that language skills per se might not be the only factor contributing to immigrant-native substitutability. Other factors such as the lack of knowledge of local labour markets or the lack of networking opportunities within ethnic communities might also be playing a role. Third, the findings for the substitutability between immigrant language groups and natives for Quebec substantially differ from those for the rest of Canada regardless of language measures. In Quebec, even low language-skilled immigrants seem perfectly substitutable with the Canadian-born.

I then use these results to simulate both short-run and long-run wage effects. Using the estimated elasticity of substitution of 30 for Canada outside of Quebec, the wage rates of the Canadian-born are simulated to have fallen by 4.2% and those of immigrants to have fallen by 6.9% in the short-run. In the long-run, I simulate the wage effects to be positive on natives, around 0.6% and to be negative on immigrants, around 2.1%. I also find that the largest long-run impact to have been for the most educated immigrants with the lowest language skills. Their wages are simulated to have fallen by about 35.2% during the period. In Quebec, I use the value of infinity for the elasticity of substitution. In this case, the short-run immigrant supply induced simulated wage effects are estimated to be around 2.2%.

This chapter is organized as follows. I begin with a literature review of two streams of studies, those addressing the economic assimilation of immigrants to Canada, and those addressing the impact they have on the Canadian born. A detailed discussion of the data follows, describing my use of the Census and how I organized the data into skill categories. The last substantive section presents the empirical methods based on the nested-CES production function, and the estimates I obtain from the model.

2.2 Review of the literature

2.2.1 Language skills and labour market outcomes of immigrants

The economic assimilation of Canadian immigrants has received considerable attention in recent decades. Previous studies suggest that not only entry earnings of Canadian immi-

grants have declined, but the subsequent rate of growth in earnings has slowed as well. Changes in productivity related characteristics, education and labour market experience, have contributed to larger earnings differentials between immigrants and natives (Baker and Benjamin, 1994; Bloom, Grenier, and Gunderson, 1995; Grant, 1999; Frenette and Morissette, 2003). Research has also documented the contribution of other factors such as language skills to immigrant-native earnings differentials.

Most of these studies use proxies based on reported languages in the Canadian Census as the basis for analyzing language skills. As mentioned, the Canadian Census does not collect any direct information on language skills. Rather, it collects information on knowledge of official languages, mother tongue and home languages. Mother tongue refers to the first language learned in childhood that is still understood, while home language is the language spoken most often at home at the time of the Census. While some studies use a single set of information to represent language skill of an immigrant, other studies combine different information sets for the same purpose. For example, Chiswick and Miller (1995) use the information on home language to construct an indicator variable to represent language skill of an immigrant. The indicator takes the simple value of 1 if an immigrant speaks English at home, and zero otherwise. Chiswick and Miller (2001) investigate the acquisition of official languages among Canadian immigrants. They construct a three-level variable using information on both knowledge of an official language and languages spoken at home to measure acquired language-capital. They also introduce a linguistic distance measure into the analysis to control for potential hardships an immigrant might encounter while accumulating language-capital. These hardships might be enhanced if the two languages in question (the language in the host and that in the source country) are linguistically distant. Their linguistic distance measure is based on language test scores in foreign languages by English-speaking Americans at the U.S. Department of State, School of Language Studies. A higher test score represents lower difficulty in verbal communication and smaller distance between English and the second language in question. The results suggest that immigrants with linguistically distant mother tongues are less likely to acquire Canada's official languages.

Aydemir and Skuterud (2005) combine different information together to construct a proxy for immigrant language skills. They use the information on mother tongue and knowledge of official languages, and include the interaction of these to control for linguistic skills across Canadian immigrant cohorts. Their study finds that the changes in birth countries and language skills of recent immigrants explain almost a third of the gap in entry earnings between immigrants and non-immigrants. Boyd and Cao (2009) combine the information regarding home languages, mother tongue, and knowledge of official languages to construct a five-level language ability indicator. In their study, if immigrants report at least one official language as both mother tongue and home language along with knowledge of official languages, then they are considered as being endowed with the highest level of language-capital. An immigrant reporting at least one official language as either mother tongue or home language is considered as being more advantaged in terms of language-capital as compared to an immigrant who does not report any official languages. Their results indicate that lower language skills are associated with lower wages.

More recently, alternative data sets or linguistic measures are used to represent language abilities of Canadian immigrants. For example, Ferrer, Green, and Riddell (2006) use literacy and numeracy test scores from the 1994 International Adult Literacy Survey for non-immigrants and the 1998 Ontario Immigrant Literacy Survey for immigrants. One of the most important advantages of these data sets is that they provide objective measures of language abilities. These test scores indicate that an average immigrant has lower literacy skill levels compared to an average non-immigrant. Lower levels of literacy skills lead to lower returns for foreign education and foreign experience as well. Skuterud (2011) uses the 2003 International Adult Literacy and Life Skill Surveys for Canada to demonstrate that there exists a strong positive association between immigrant earnings and language skills. However, the sorting of immigrants across occupations does not seem to be related to language skills of immigrants. The author suggests that high collinearity among skills required for any occupations prevents immigrants from sorting across occupations. Adsera and Ferrer (2016) examine the integration of recent female immigrants entering Canada between 1986 and 2006, in particular married women, into the Canadian labour market. As a measure of

language skills of immigrants, they use the Linguistic Proximity (LP) Index defined according to the number of shared branches of languages in a language family tree.[3] The LP takes the values between 0 and 1 based on the closeness of the official language in immigrant's birth country to English or French. Higher index number indicates closer languages. The main results suggest that labour market participation of married immigrant women is no longer supplementary to that of their spouses. Female immigrants' participation and their career progression are similar to those of non-immigrant female workers. Further, they find that immigrant women experience lower wages as the LP becomes smaller. As compared to immigrant women with non-university level of education, wage penalties are larger for immigrant women with university level of education.

In addition to linguistic measures mentioned above, the Levenshtein Distance has been used in the economic literature to examine the impacts of language barriers in immigration and international trade. The LD is a measure of the linguistic distance that can be derived between any two languages using a computer algorithm, called the Automated Similarity Judgement Program, developed by the Max Planck Institute of Evolutionary Anthropology (Isphording and Otten, 2013). The higher the Levenshtein Distance, the more "distant" the languages.[4] Isphording and Otten (2013) use the LD to examine the association between language fluency of an immigrant and the distance between his/her language and the language spoken in the host country. They find language fluency is negatively affected by the distance between the two languages. Additionally, the volume of trade between the two countries is negatively affected by the distance between the two languages of trade partners. Adsera and Pytlikova (2015) find that the distance between immigrant languages and the language spoken in the host country affects the choice of a destination country. The likelihood of an immigrant migrating to a particular country decreases if immigrant language is more distant from the official language in the host country. Along with the LP index, Adsera and Ferrer (2015) also incorporate the LD into the analysis

[3] A language family is a group of languages historically originating from a common language. If two languages are not related at all to a common language in a language family tree, the index takes the value of 0. For example, the LP is 0 for Vietnamese and English. The index takes the value of 1 if the languages being compared are the same.

[4] Chapter 1 provides a detailed presentation of the Levenshtein Distance.

of career progression of immigrant men into the Canadian labour market. Parallel to the results obtained with the LP, the high levels of LD are associated with lower wage levels of immigrants.

Indicating the significance of language skills in the economic assimilation of Canadian immigrants, previous studies construct alternative indicators while referring to immigrant language skills. The indicators combine different sets of information such as knowledge of the official languages and reported languages. More recently, alternative measures such as test scores and proximity/distance indices are introduced into the Canadian data analysis. These alternative measures provide more advantages and flexibility over multiple-scale indicators in capturing variations of immigrants' language capital and representing their skills.

2.2.2 Immigrants and natives

Another stream of studies has explored the impact of immigration on the earnings of native-born workers. Labour market theory suggests that wage levels decline with the increases in labour supply due to immigration if immigrants and natives compete in the same labour market.[5] With regard to empirical findings, however, the studies using differences across and within local labour markets as an identification strategy do not provide strong evidence to support the prediction of the model (Friedberg and Hunt, 1995). For example, Card (1990) examines the impact of an increase in the immigrant population by 7% due to Mariel Boatlift in 1980 on wages in the Miami labour market. Using a difference-in-difference methodology, this study shows that the increase in labour supply due to this immigrant inflow does not change wages and the rate of unemployment for the less-skilled workers. Butcher and Card (1991) analyze the impact of immigration in the labour markets of 24 major U.S. cities during the 1980s, and conclude that immigration had a small but positive effect on wages.

[5] There is also a vast literature on the impact of labour supply shocks in labour markets as cohort size changes, Welch (1979) and Berger (1989) being a couple of studies among many others. For example, Morin (2015) examines the impact of Ontario's double cohort of high school graduates on youth earnings as a result of the removal of Grade 13 in Ontario in 2001. These studies are not reviewed here in detail. Instead, I focus on the changes in labour supply due to migration, and hence the impact of immigration on wages.

Borjas (2003) criticized these earlier studies for ignoring the fact that competitive market forces equalize economic conditions across local labour markets. This groundbreaking study uses the variation across labour skill groups at the national level in order to investigate the impact of the immigrant-induced increase in labour supply on native-born workers. The impact of large immigrant influxes are examined in a national labour market conditional on perfect substitutability among immigrants and natives.

In the human capital accumulation theory, not only education but also labour market experience is an important factor in the accumulation of human capital. In this regard, Borjas (2003) defines skill groups with respect to both education levels and labour market experiences of workers. In other words, workers with the same level of education but different years of labour market experience are not considered perfect substitutes. The supply fluctuations across skill groups induced by immigration generate enough variation to identify the impact of inflows of immigrants on labour market outcomes at the national level. This assumes that foreign-born and native-born workers are homogenous.

Borjas (2003) assumes a linear homogenous constant elasticity of substitution function (CES) with capital and labour for the aggregate output in the economy. Then, a nested structure is imposed on the production function to incorporate skill types into the aggregate labour force, which is comprised of heterogeneous workers with respect to education and years of labour market experience. Education levels form the second level, while years of labour market experience form the third level in the production function. In this structure, the identification of the labour demand is provided by the changes in the aggregate supply of labour induced by the shifts in the foreign-born workers' supply. In other words, an immigration-induced labour supply shift becomes a supply-shift factor to identify the impact of changes in the immigration on native-born wage rate.

Using the Public Use Micro Data Samples (PUMS) of the Decennial Census between 1960 and 2000, the analytical sample includes men between 18 and 64 who are labour force participants and have between one and 40 years of labour market experience.[6] Education groups are defined as: high school dropouts, high school graduates, individuals with

[6]To represent the data for 2000, the author pools the 1999, 2000, 2001 Annual Demographic Supplement of the Current Population Surveys (CPS).

some college degree, and college graduates. Years of labour market experience are divided into five-year periods. Both annual and weekly wages are examined for the sample of men who do not live in collective dwellings, report positive hours, weeks and annual earnings, are in the civilian labour force, are not self-employed, and also not enrolled in school.

Borjas's simulation results suggest that immigration between 1980 and 2000 in the U.S. reduced overall wages by 3.2%. However, the impact was not evenly distributed across education groups. It seems that the least educated natives, high school dropouts, experienced the biggest drop. Their wages declined by 8.9%. The least affected group was the group of native-born workers with some college degree whose wages fell by only 0.3%. The decline in college graduates wages was estimated to be 4.9%. Borjas (2003) concludes that considering market equalizing forces through a national analysis is required to observe the actual impact of the increase in the immigrant supply on their native-born counterparts.

Parallel to Borjas (2003), Ottaviano and Peri (2008) employ the aggregate production function with a nested structure to estimate the impact of immigration on the U.S. wages. In contrast with the findings of Borjas, these authors estimate imperfect substitutability between natives and immigrants. In addition to skill groups defined by education and years of labour market experience, Ottaviano and Peri (2008) extend the structure in Borjas (2003) to a fourth-level by asserting that immigrants and natives are not perfect substitutes. They suggest that, due to a variety of reasons such as a lack of information in the labour market institutions, networking, and language skills, immigrants and natives are not homogenous within education-experience skill cells, and another labour-aggregation equation that divides workers into two groups based on immigration status should be added to the model.

They use the U.S. Decennial Census 1960-2000 and the 2006 American Community Survey. Individuals older than 18, not living in group quarters, who worked at least one week in the previous year are included to measure the total hours worked in each cell. To construct the average wage, a subset of employment sample is used and the workers who do not report any wages or who are self-employed are excluded from the sample.

One of their most significant findings is the elasticity of substitution between immigrants and natives estimated to be around 20. By using the estimates of the elasticity of substitution among different skill groups based on education level, years of labour market experience and immigrant status, Ottaviano and Peri (2008) suggest that immigration affects the wages of foreign-born workers more negatively than those of native-born workers. The simulation for the 1990-2006 period reveals that in the long run immigration increases the wage of an average U.S.-born worker by 0.6%, and reduces the wage of an average foreign-born worker by 6.4%.

Using the same nested-CES structure and the U.S. Decennial Census between 1960 and 2000, the elasticity of substitution between immigrants and natives is re-estimated by Borjas, Grogger and Hanson (2008), and it is found that immigrants and natives are perfect substitutes. Borjas et al. (2008) criticize the findings by Ottaviano and Peri regarding the estimates of imperfect substitutability between immigrants and natives. They show that the imperfect substitutability between immigrants and natives arises as a result of the inclusion of the young workers (those between 18 and 21), a group that might not be fully attached to the labour market. Inclusion of high school students as high school dropouts creates skill groups capturing high levels of labour supply with low wages for the youngest native workers. As a result, a spurious negative correlation likely lies behind the imperfect substitutability between natives and immigrants. Borjas et al. (2008) also emphasize the importance of using the correct weights. The corresponding weights within each skill group should be the ones used to calculate the dependent variable in estimable equations, average wage levels. They also underline the differences in results derived using the average of logarithmic wages versus the logarithm of average wages, and suggest using the average of the logarithm of the wages within skill groups, following the convention in the literature. They show that using the logarithm of average wages tilts the results toward finding imperfect substitutability between immigrant and natives.

Ottaviano and Peri (2012) not only reproduce the estimates of the elasticity of substitution in the light of the critique raised by Borjas et al. (2008), but more significantly test four nesting models against each other by using a nested-CES structure for aggregate

labour. The first model starts with four education categories including individuals with no high school degree, a high school degree, some college education and a college degree. Then, different experience groups that are defined in five-year intervals nest into each education group while immigrants and native workers nest into each experience group. This model assumes that any pair of education groups or any pair of experience groups have the same substitutability among each other. For example, the substitutability between workers with no high school degree and a high school degree is the same as the substitutability between the workers with no high school degree and a college degree. The second model has four nests starting with two broad education categories: low educated workers and high educated workers. Then, more detailed education groups nest into these broad education groups. Individuals with less than a high school diploma and a high school degree make up the low-educated group while the individuals with some college education and a college degree make up the high-educated group. This model allows the substitutability to differ between and within broad education categories. The third model starts with the detailed four education categories as does the first model. The difference stems from the detailed experience categories that are grouped into broad experience categories. The broad experience categories include the young and the old. Then, young workers are divided into five-year experience groups: 1-5 years, 6-10 years, 11-15 years and 16-20 years. Older workers are grouped according to their potential experience starting with 25 years and ending with 40 years of potential experience defined in five-year intervals. According to this model, the substitutability differs between and within experience groups. The last model inverts the education-experience order of nesting in the first model.

The results indicate that the second model, allowing separate estimates of the elasticity of substitution within low educated group and within high educated group, should be selected. Moreover, depending on the nesting structure imposed on the CES model, it is shown that once imperfect substitutability is allowed between immigrants and natives, there is a significant negative long-run wage impact on less educated natives over 1990-2006, which changes from -3.1% to +0.6% depending upon the nesting structure. The wage of an average U.S.-born worker increased around 0.6% for the corresponding time period. The

wage of an average foreign-born worker is reduced by around 6% after considering capital adjustment.

To the best of my knowledge, Lewis (2013) is the only study investigating immigrant substitutability in a way that takes into account language skills. Using the U.S data, he investigates whether there exists any co-movement between relative wages of immigrants and the relative supplies in the skill groups, based on education levels and language skills of immigrants. In this study, years of labour market experience are not considered among skill groups. Since theory suggests that an increase in the relative supply of one group decreases the relative wage of that group if the two groups of workers are imperfect substitutes, Lewis (2013) asks whether different language skills affect the degree of substitutability among workers. In other words, he is asking whether the immigrants that have high English proficiency are closer substitutes for native-born workers than immigrants with poor language skills. In order to test the language hypothesis, Lewis (2013) also examines the Puerto Rican labour market where all workers speak Spanish and the labour market is considered as a single market.

Lewis (2013) mainly takes advantage of variations across metropolitan areas and skill groups based on education levels and language skills. To define skill groups, two broad education categories are used for those holding a high school degree or less and those holding a degree higher than high school. Alternative language skill groups are defined for immigrants' self-reports of ability to speak English. Immigrants are also categorized into five-year intervals according to their age-at-arrival years since migration. The study uses the 2000 U.S. Census of Population and American Community Surveys from 2007-2009 and also includes the Puerto Rican Censuses from 1970 to 2000 and 2007-2009 Puerto Rican Community Surveys. The sample is based on individuals between 16 and 65, who have positive labour market experience, and who live in one of the 136 metropolitan areas. Information on hours worked and hourly wages were aggregated to the metropolitan area by year and by broad education group. He defines a wage sample to be the subsample of these workers who are currently employed with positive wage and salary earnings and zero farm or business earnings in the past year. Only the native-born workers who speak English

"only" and "very well" are included in the wage sample. Hourly wages above $200 and below $2 in 1999 dollars are replaced with these thresholds. The main estimating models examined the impact of the immigrant-native hours of work on the immigrant-native wage gap by using the variations across education-metropolitan groups as well as the deviations in immigrants' language groups.

The results suggest that language skills are an essential determinant of imperfect substitutability between native-born and foreign-born immigrants. Lewis (2013) shows that the increase in the relative number of immigrants since 1990 in the average metropolitan areas is associated with about a 6% decline in the hourly wage of immigrants with poor English language skills, but only a 2% decline in the wages of immigrants with strong English language skills conditional on education and residence in the same residential area. In the Puerto Rican labour market, immigration is not associated with a decline in the wages of Puerto Rican immigrants relative to Puerto Rico-born workers in Puerto Rico. It is found that the English-language premium is very small for Spanish-speaking workers and the influx of immigrants only decreases the wages of their counterparts, not the wages of English-speaking workers. In other words, both the wages of Spanish and non-Spanish-speaking workers are affected by the skill ratios in their own language group.

The nested CES structural approach has three advantages. First, the main issue regarding the identification of the impact of immigration on wages is resolved by imposing a structural model on the aggregate economy. Through this national-level production function, native mobility and the endogeneity of immigrants' decisions on destination cities are considered by defining a skill group as the unit of analysis. In this way, dampening effects of market equalizing forces or well-performing labour markets are considered in the analysis. Second it provides an insight into complementarity among workers with different skill levels. Third, it solves the issue regarding dimensionality while estimating the degree of substitutability among various skill groups. In contrast to a translog production function, the nested CES structural model reduces the number of parameters making the estimation of parameters more tractable with limited data (Borjas, 2014). Nevertheless, the nested CES structural approach has its limitations. Borjas (2014) particularly emphasizes that

the unverifiable assumptions of the model are sensitive to alternative identification assumptions that are necessary for parameter estimation. Further, he underlines that the size of the total wage effect depends on the imposed functional form of the aggregate production function. Card and Peri (2016) aggree on the sensitivity of the estimates to alternative identification assumptions, but they emphasize that there is a small but non-negligible degree of substitutability between immigrants and natives regardless of which assumption is imposed to identify the parameters, leading to a small but positive long-run effect of immigration on natives'wages in each case. Borjas (2014) suggests that neither the structural model nor the descriptive analysis should be discarded. The insights from the model should be considered along with the direct estimates of immigration impacts to fully describe the impact of immigration on wages.

A number of other studies examine the impact of immigration on native-born workers in the U.S. using a variety of methodologies. Canadian studies are fewer in number. Aydemir and Borjas (2007), the only study following the national-level analysis by Borjas (2003), analyzes and compares the impact of immigration on wages levels and structures in Canada, the United States and Mexico. Both Canada and the United States receive a large number of immigrants, but they have different immigration policies, which affect the composition of immigrants in these countries. Mexico, on the other hand, provides a good opportunity to analyze the impact of immigrant outflux on the Mexican labour market. The effects of immigrant/emigrant-induced labour supply changes on wages are examined for the time period between 1971 and 2001 in Canada, and between 1960 and 2000 in the U.S. and Mexico.

Their analysis uses the 1971 Canadian Census and all the Canadian Censuses conducted every five years between 1981 and 2001. The 1976 Census did not contain information on incomes, and therefore was not used. The results suggest that immigrants and natives are perfect substitutes, a finding for all three countries. The estimates of the elasticity of substitution between immigrants and natives are imprecise and small for Canada, and range between 62 and 30. The simulated net wage effect of the immigration-induced supply shock between 1980 and 2000 in Canada is estimated as a 3.4% decline

for an average Canadian worker in the short-run, owing to the fact that immigrants and natives are perfect substitutes. The effects vary across education groups, depending on their income shares. The simulation results indicate that the most-educated workers who hold post-graduate degrees are most negatively affected by immigration supply shocks. High school dropouts gain due to immigration. The wage of an average high school dropout in Canada rises 2.8% in the short-run and 7.8% in the long-run.

With regards to other methods to obtain consistent estimates of the impact of immigration on wages in the Canadian context, Tu (2010) uses both a two-step regression method and an instrumental variable approach using the immigrant native ratio in 1991 as the identifying instrument. Using the Canadian Censuses between 1991 and 2001, this study takes advantage of the variation in different types of immigrant workers residing in a particular census metropolitan area (CMA). The sample includes the full-time paid workers that reported at least 50 weeks of work in a given year, between 16 and 64 years of age. A skill group is defined according to education levels. These types include the workers with no high school diploma, with a high school diploma, with a post-secondary degree, and with at least a university degree. Tu (2010) constructs a pseudo-panel data set. First, native individuals' wages are calculated for each year separately. Then, the difference of the adjusted means for skill-area effects are analyzed by controlling for the change in immigrant-native ratio between two censuses. The change in the immigrant-native ratio is instrumented with the immigrant-native ratio in 1991. The results mainly suggest that there is no clear evidence that immigration affects native wages. The impact of immigration-induced supply changes on the average wage level is either statistically insignificant or positive between 1991 and 2001 in Canada.

Grant (1998) uses the same two-step estimation process by taking variations in new immigrant cohorts across occupations into account. Using 1986, 1987, 1991 and 1992 Surveys of Consumer Finances, the analysis includes individuals who are wage workers aged 18 to 64 years. Self-employed workers or workers reporting high incomes are excluded from the sample. A two-step regression analysis first estimates the occupational effects on average wages, and then estimated occupational effects are regressed on the change in the proportion

of new immigrants within an occupation group. For the ratio of new immigrants, Grant (1998) uses various instruments including the intended occupation of new immigrants, the difference between intended occupations and actual occupations, and the skills of potential immigrants. The last instrument is particularly strong, because it represents immigrants' human capital decisions that are made before arrival and dependent on the wage structure of the immigrant's birth country. The findings suggest that immigration has either a very small or no effect at all on native-born workers.

While language skills affect the catch-up rate among immigrants with otherwise observationally comparable characteristics, the question of the role of linguistic diversity on the determination of the average wages of native-born is yet to be answered in the Canadian context. In my analysis, I employ the structure that has been successfully implemented with U.S data to examine the impacts of immigration-induced supply shocks. I focus on the changes in labour supply as a result of immigration, and in terms of its impacts on Canadian wages. On this basis, extending the models suggested by Borjas (2003) and Ottaviano and Peri (2008, 2012), I propose to define new skill groups among immigrant men regarding their differences in linguistic backgrounds instead of aggregating them into a single group. I claim that aggregation might mask the shifts in the characteristics of immigrant cohorts, hide the immigrant-native substitutability and provide misleading results in terms of wage effects of immigration.

2.3 Sample description and data

As mentioned previously, this study is based on the 20% microdata files of the 1981, 1986, 1991, 1996, 2001, and 2006 Canadian Censuses. Following the literature, I construct two analytical samples focusing only on men: (1) the employment sample and (2) the wage sample. The employment sample includes all men between the ages of 21 and 64 (inclusive), not residing in a collective dwelling, and who report working at least one week during the reference year and report not having attended school. These sample selection criteria allow the analysis to uncover the full labour market consequences of immigration, including both wage and employment impacts. The wage sample is more homogeneous and is derived from

the employment sample. Only wage and salary workers who worked full-time and full-year at a positive wage during the reference year are included. I consider full-year workers to be those who reported at least 40 weeks of work during the year. The unweighted number of workers in the employment sample is 1,017,585 in 1981 and 1,378,430 in 2006. The wage sample includes 689,160 workers in 1981 and 914,225 workers in 2006.[7] An immigrant is defined as a naturalized citizen or a non-citizen in Canada. This definition is necessary to describe an immigrant regardless of residency status. It should be noted that non-permanent residents are enumerated in the Canadian Census as of 1991.[8] In earlier censuses, only permanent residents are included in the census. Over time, the number of non-permanent residents had grown substantially and government programs such as employment programs were influenced by their presence. Hence, including these non-permanent residents in the census became necessary to fully document demographic and socio-economic characteristics of those living in Canada and participating in the Canadian labour market. For example, according to Thomas (2010), over 94,000 out of 112,000 enumerated non-permanent residents were holding full-time jobs in 2006.[9]

The labour market outcomes I examine are constructed for different skill groups defined on the basis of education and years of labour market experience for natives, and on education, labour market experience and language skills for immigrants. Education is based on the Census variable referring to the highest degree obtained at the time of the Census. Three groups are defined: (1) no high school diploma or a high school diploma; (2) a vocational training certificate or a diploma below the university level; (3) an undergraduate university degree or above. The choice of categorization is motivated by the findings in Ottaviano and Peri (2008, 2012). They find that high school dropouts and high school graduates are almost perfect substitutes. The years of labour market experience variable is defined as age minus years of schooling minus 6. The 2006 Census did not collect any

[7]The numbers of observations are rounded to a base of 5.

[8] Statistics Canada defines non-permanent residents as the individuals who are neither Canadian citizens by birth nor landed immigrants. They have a valid work or study permit. Alternatively, they might claim a refugee status at the time of the census. Family members living with non-permanent residents in Canada are also considered as non-permanent residents.

[9]The sensitivity of the main estimates to the exclusion of the 1981 and the 1986 censuses is assessed. The results are presented in the section 2.5.1.

information on the years of schooling. For this census I derive years of experience in the labour market by assuming age at the entrance to the labour market based on the highest degree of a person. For example, if a person is a high school dropout, it is assumed that he enters the market at the age of 17. Then, by subtracting this from the actual age of a person, I obtain years of experience in the labour market. I select all individuals with at least one year of labour market experience and no more than 40 years, and group them into five categories: those with 1 to 5 years of experience, those with 6 to 10, those with 11 to 15, those with 16 to 20, and finally those with 21 or higher.

The literature recognizes the fact that Canadian labour market experience may be valued differently by employers than foreign labour market experience. Sweetman and Warman (2014) suggest that although temporary foreign workers receive modestly positive return for their labour market experience, the reward for years of foreign labour market experience becomes either zero or negative for permanent residents. Green and Worswick (2010) find that earning patterns across cohorts differ based on immigrants' arrival years. As compared to immigrants entering Canada during the early 1980s, immigrant cohorts entering Canada during the 1990s and 2000s experienced large declines in their entry earnings due to the exhaustion of rewards by years of foreign experince. The decline is largely explained by the shift in source countries, suggesting immigrants' hardship in transferring human capital acquired from non-traditional source regions. Aydemir and Skuterud (2005) find similar results, suggesting declining returns by years of foreign experience across cohorts, in particular for those speaking non-official languages as mother tongue. Based on the findings in the literature, although it is an extreme assumption, I claim that foreign experience has a zero value in the Canadian labour market. As a result, if I find that the labour market experience variable for an immigrant is higher than years since migration, I define years since migration as the years of experience. Since there is no available information on the entrance year of non-permanent residents to Canada, the Canadian labour market experience of these immigrants cannot be determined by comparing labour market experience in the source country and Canada. As a result their labour market experience is

counted without regard to where it was obtained.[10] The full details regarding the construction of samples and skill categories are provided in Appendix 2.1. As a robustness check, I consider an alternative wage sample that includes both full-time and part-time workers. Appendix 2.2 provides a sample description for these workers. Table 2.1 illustrates the initial samples sizes and the number of observations included after each restriction, and the final sample sizes for the employment and the wage sample.

The reporting of language skills in the Canadian Census is both more and less detailed than in the American Census.[11] There is no direct measure of competency, self-reported or otherwise in the Canadian Census, but there are a host of other questions used to determine an individual's official language status.[12] My analysis focuses on home language and on mother tongue. I make use of the Levenshtein Distance (LD) to further categorize these languages. My categorization of immigrants into finer groups based upon the LD of their home languages and mother tongue is intended to be indicative of their language skills. I assign each immigrant an LD index that represents the "distance" between their reported language and English, or French according to the analysis, and then divide immigrants into groups based on the distance.[13]

Since my analysis focuses on home language and mother tongue, the advantages and drawbacks of each language skill measure should be addressed. Home language is a

[10]Immigrant labour market experience obtained in the source country is counted in two different ways in the literature. First, it is counted as if it was obtained in the host country. Alternatively, it is claimed that the experience in the source country is counted less as compared the experience in the host country. Borjas (2003) suggests a year of labour market experience in the destination country to be equivalent to four years in the source country.

[11]In the American Census, the respondents are asked to report their home language and how well they speak English. The available check boxes for the latter are very well, well, not well, and not at all.

[12]The respondents are asked whether they can conduct a conversation in official languages, and a high percentage of immigrants, around 95%, report their ability to converse in official languages. Hence, the question does not reveal enough information in order to group immigrants based on their language skills. In addition, the information on language at work, language used in performing a major task or job, is also available in the Census, but only for the 2001 and 2006 censuses. Hence, I do not explicitly use it.

[13]In each Census, some immigrants are grouped into residual language categories due to there being few observations. It makes impossible to assign a single distance measure to these immigrants. In this case, a median of LD indexes of all languages, reported by immigrants coming from the same country as the immigrant within a residual language group, is assigned to the immigrant. If this information is not available then an LD index of the national language of his host country is assigned. The national language of the source country, defined as the language of education, of mass media, etc. in the source country, is assigned as a representative index for the language skills of these immigrants. The Ethnologue, an online language catalogue, is used to determine the national language of the country of interest. On this basis, I am able to assign an index number to almost all of the immigrants in our relevant samples in the 1986, 1991, 1996, 2001 and 2006 Censuses.

better approximation for language skills as compared to mother tongue although it has its limitations. A number of immigrants with non-official languages come from countries in which either one of Canadian official languages are used as the national language or they might have simply acquired an official language through schooling in the birth country. In these cases, immigrants likely have better language skills than does the reported home language indicates. Additionally, because immigrants accumulate language skills and adopt official languages as home language, it is a choice variable and might be dependent on other factors affecting both the relative supply and relative wage of immigrants to natives. In contrast to home language, mother tongue provides exogenous shifts for the immigration-induced labour supply because mother tongue does not vary over time. However, as a result of its time-invariant characteristic, mother tongue is a less preferable measure for language skills than home language in capturing accumulated language skills of immigrants.

The LD is a measure of the linguistic distance that can be derived between any two languages using a computer algorithm, called the Automated Similarity Judgement Program, developed by the Max Planck Institute of Evolutionary Anthropology (Isphording and Otten, 2013). The LD quantifies the differences between two languages, and provides a proxy measure in my analysis for the difficulty in learning languages. The more distant the languages are, the higher the LD.

Three language skill groups are defined, which I refer to as high, medium and low. If the Levenshtein Distance is zero, the language is English (or when appropriate in the analysis, French). I consider individuals with a LD of zero to have "high" language skills in English (French). Index values greater than zero indicate lower language skills. I devise a cut-off value based upon the language that has the lowest LD among all non-Indo-European languages reported in the Census. If an individual's reported home language or mother tongue has a LD lower than this value, I consider the individual as having a "medium" level of language skills in English (French). If an individual's reported languages has a higher LD than the LD of the cut-off language, I consider the individual as having a "low" language skills in English (French). Immigrants speaking an Indo-European language

are likely to have higher English skills than those speaking a non-Indo-European language since both English and French are also Indo-European languages.[14,15]

The wage rate within each skill group is measured as the real weekly wage: annual wages and salaries deflated with the 2005 Canadian Consumer Price Index, and divided by the number of weeks worked in the reference year. The mean wage for each skill group is calculated by averaging the log of real weekly wages of corresponding individuals within skill cells. Personal weights provided in each Census are used throughout. Labour supply is measured by summing these weights within skill cells.

Tables 2.2 and 2.3 illustrate the selected home languages and mother tongues reported by male immigrant workers. The LD index is presented in the second and third column of the tables, referring to the "distance" of the reported language to English in the second column, and to French in the third column. Table 2.2 shows that the percentage of immigrants reporting French as home language in Quebec increased by almost a quarter between 1981 and 2006 in Quebec. Moreover, the percentage of immigrants reporting English or French as home language decreased from 68% to 53% in Canada. At the same time, the percentage of English- or French-speaking immigrants only declined by 4 percentage points, from 58% to 54% in Quebec.

Table 2.3 shows that the percentage of immigrants reporting English as mother tongue notably declined between 1981 and 2006 across Canada. The largest decline was found in Quebec where 50% fewer immigrants reported English as mother tongue in 2006 compared to 1981. The percentage of immigrants reporting French as mother tongue went down slightly between 1981 and 2006 across Canada. The decline in the percentage of immigrants reporting French as mother tongue was almost the same in Quebec and the

[14]It turns out that Hungarian is the non-Indo-European language with the smallest LD with a value of 95.22. An immigrant with a home language or mother tongue having an LD index smaller than 95.22 is categorized into medium language skill group. Immigrants speaking Hungarian or languages with LD indexes higher than 95.22 are categorized into low-language group. For French-based categorizations, the cut-off language turns out to be Hebrew with the LD index 93.26. Appendix 2.3 provides the list of languages in language skill groups with respect to English and French.

[15]Since the LD captures subtle differences between two languages, the languages from the same sub-family of the Indo-European language family can be categorized into separate language skill groups. For example, based on the LD to English, Persian, Urdu, and Sinhala belonging to the Indo-Iranian sub-family of Indo-European languages are included in the medium language skill group while Hindi, Bengali, Marathi, Konkani, Gujarati and Pashto from the same sub-family are included in the low language skill group.

rest of Canada. As is well-documented, a decline in the percentage of immigrants speaking European languages other than English and French is also observed. For example, the percentage of Italian-speaking male immigrants fell by almost three-quarters both in Quebec and in Canada. The number of immigrants speaking Chinese, Punjabi and Tagalog increased more than two-fold in Canada. The percentage of immigrants reporting Arabic as mother tongue particularly increased in Quebec, from 2.9% to 11.2%. Another revelation is that the percentage of immigrants with a mother tongue or home language other than those selected in Tables 2.2 and 2.3 made up almost one quarter of the working male population across all regions in 2006. This fact suggests that linguistic skill composition of immigrant male workers in the Canadian labour market substantially changed during the period.

Figures 2.1a and 2.1b categorize immigrants into three language skill groups, and present trends in average log of real weekly wages of immigrants and Canadian born workers between 1981 and 2006. In Figure 2.1a, immigrant language groups are defined with respect to the distance of home language to English outside of Quebec or to French in Quebec. The left-hand side graph is for the immigrants residing in Canada outside of Quebec and the right-hand side graph is for the immigrants residing in Quebec. Figure 2.1b repeats the exercise but for the distance based on mother tongue. There are three notable findings from Figure 2.1a. First, the wages of low language-skilled immigrants were consistently lower than the wages of other immigrants and natives. The pattern follows a distinct V-shape pattern. The wages decreased between 1981 and 1991, then increased. Second, English-speaking immigrants earned more than all other immigrant groups and natives outside of Quebec between 1981 and 2006. Third, the trends of medium and low language-skilled immigrant wages are similar to each other outside of Quebec. Comparing Figure 2.1a with Figure 2.1b, the wages of low language-skilled immigrants were the lowest among all workers in Quebec, parallel to the findings for immigrants in the rest of Canada. In contrast to the findings from Canada outside of Quebec, natives earned more than all immigrant groups in Quebec and the wage gap became larger from 1986 on. Wage trends of the high and medium language-skilled in French show more similarities as compared to that of low language-skilled immigrants in Quebec. This finding is also in contrast with the finding

for the immigrants outside of Quebec, which is likely related with the concentration of immigrants reporting English as home language in the medium language-skilled groups. In both regions, wage trends for language skill categorizations based on home language are consistently lower than those based on mother tongue. This fact suggests that immigrants with high language skills might be incorrectly categorized as the low language-skilled when language skill-based categorizations consider mother tongue instead of home language.

Figures 2.2a and 2.2b plot the wage gap series between immigrants with different language skills as measured by home language and native-born workers outside of Quebec and in Quebec, respectively. In both figures, the left panel presents the immigrant-native log wage gap among the least educated, the ones with at most a high school degree, while the right panel plots the wage gap among the most educated, those with at least a Bachelor degree. Figure 2a reveals that, among the most educated, the immigrant-native wage gap for all immigrant groups widened continuously between 1986 and 2006. The sharpest decline in immigrant wages is observed among the most educated with low language-skilled immigrants. However, in Quebec, it seems that, among the low language-skilled immigrants, the least educated experienced the largest decline in their wages as compared to their Canadian-born counterparts.

To illustrate the variations in labour supply by various skill groups due to immigration, Figures 2.3a and 2.3b present the share of immigrants for the most and the least educated workers by the years of labour market experience for selected census years. Among the most educated with the least years of labour market experience, the plots show that the share of low language-skilled immigrants increased continuously over time in both regions. In contrast, outside of Quebec, the share among the least educated with the least years of labour market experience dropped in the 2000s as compared to the 1990s while it increased steadily between 1981 and 2006 in Quebec. In both regions, the variation in the supply of immigrants across skill groups is substantial. Taking Figures 2.2 and 2.3 together suggests that the most educated immigrants with low language skills experienced large declines in their wages and steep increases in their labour supply share between 1981 and 2006, particularly outside of Quebec.

Table 2.4 presents the wage gaps between immigrants and natives for the years 1981, 1991, 2001 and 2006 outside of Quebec. Table 2.5 shows the ratio of the labour supply of immigrants to that of natives for the same group of immigrants. These tables provide a complete description of skill groups based on education levels, years of labour market experience, and language skills. Further, they present the changes in the main variables used in the analysis. The omission of 1986 and 1996 data from tables is due to space constraints. It seems that low language-skilled immigrants experienced the largest declines in their wages as well as the largest increase in their relative supplies as compared to Canadian born workers. For example, among the most educated with one to five years of labour market experience, the wage gap between the immigrants with low language skills and Canadian born workers declined 0.093 log points between 1981 and 2006 while their relative supply increased 52 percentage points, up from 8% in 1981. Over the period, the immigrants in this group experienced a tremendous increase in their labour supply share. The share of immigrants doubled from 1981 to 1991 and further increased to 19% in 1991. In the next decade, it kept its pace and reached up to 42%. Between 2001 and 2006, it increased more than 40%, and the ratio became 60% in 2006. Table 2.6 and 2.7 present the same data for Quebec. As in all other provinces, low language-skilled immigrants experienced the largest decline in wages in Quebec. The substantial difference in patterns is observed in relative supplies, however. In contrast to the rest of Canada, the largest increase in the relative supplies occurred among the high language-skilled, not among the low language-skilled. For example, the ratio of immigrant-to-native labour supply was 36% for the least experienced most educated high language-skilled immigrants in 2006, up from 7% in 1981. High language-skilled immigrants in Quebec seemed to experience larger declines in their wages than high language-skilled immigrants in other Canadian provinces.

Figure 2.4 plots the relationship between the immigrant-native log wage gap and log ratio of immigrant-to-native labour supply outside of Quebec. Categorizing immigrants into different language skill groups, Figure 2.5 plots the same correlation for each language group separately. Figures 2.6 and 2.7 repeat the same exercise for immigrants in Quebec. Language groups are defined based on the distance of home language to English outside of

Quebec and to French in Quebec. Illustrating the correlation between two main variables provides an insight into the structural model that estimates the elasticity of substitution between immigrant language groups and natives. There are two notable findings from these graphs. First, the relation between the immigrant-native log wage gap and log ratio of immigrant-to-native labour supply outside of Quebec is close to zero. Considering immigrants as a single group indicates that immigrants and natives do not differ from each other substantially. Second, with regards to language skill groups based on home language, the relation between log wage gap and log labour supply ratio is insignificant or significantly positive for all language groups except low language-skilled immigrants residing outside of Quebec. In other words, immigrants and natives might be close substitutes if immigrants have high or medium language skills. For low language-skilled immigrants, the immigrant-native wage gap and the log ratio of immigrant-to-native labour supply do not move together, particularly for immigrants residing outside of Quebec. This suggested negative relationship is consistent with imperfect substitutability between this group of immigrants and natives.

In summary, the data show that low language-skilled immigrants earned less than other immigrants and Canadian-born workers between 1981 and 2006. During the period, the least experienced and the most educated among low language-skilled immigrants experienced the largest declines in their wages as compared to natives. Their supply increased significantly in Quebec and in the rest of Canada. In the following sections, I use the structural model and estimation strategies to examine the impact of immigrant-induced labour supply changes on the wages of different worker groups in Canada, Canada outside of Quebec and Quebec.

2.4 Nested-CES model

I impose a nested CES structural model on the Canadian economy. In what follows I borrow directly from Aydemir and Borjas (2007). A linear homogenous constant elasticity of substitution function (CES) with output Q, capital K, and labour L is assumed for the

aggregate economy at time t as:

$$Q_t = [\lambda_{Kt} K_t^v + \lambda_{Lt} L_t^v]^{1/v} \tag{2.1}$$

with $\sigma_{KL} = 1 - 1/v$, the elasticity of substitution between capital and labour. The productivity shift parameters, λ_{Kt} and λ_{Lt},are time-variant and assumed to sum up to 1. Since the aggregate labour, L_t, is comprised of heterogeneous workers with respect to education and experience, a nested structure is imposed on the production function to incorporate different skill types into the aggregate labour. A second-level nest is defined by education level.

$$L_t = \left[\sum_s \theta_{st} L_{st}^{\rho} \right]^{1/\rho} \tag{2.2}$$

with $\sigma_E = 1 - 1/\rho$, the elasticity of substitution across education groups, and L_{st} the supply of workers with education s at time t . The productivity shift parameters of education groups, θ_{st}, are time-variant and assumed to sum up to 1. Finally, within education groups, the supply of workers with different experience levels is aggregated and, and a third-level CES nest is defined as:

$$L_{st} = \left[\sum_x \alpha_{sxt} L_{sxt}^{\eta} \right]^{1/\eta} \tag{2.3}$$

with $\sigma_X = 1 - 1/\eta$, the elasticity of substitution across experience groups with the same education, L_{sxt} the supply of workers with experience x within education group s at time t . The productivity shift parameters of experience groups, α_{sxt}, are time-variant, and assumed to sum up to 1. Aydemir and Borjas (2007) impose time-invariance on productivity parameters of experience groups as an identification strategy due to data availability concerns.[16]

I extend this model to incorporate variation in the language skills of immigrants. My notation borrows from these studies, and I add another nest at the fourth level to the model. In this regard, each experience-education group incorporates native workers and high, medium and low language-skilled immigrants. Therefore, the fourth-level labour

[16]For identification, time-variant education-experience effects are also considered as an alternative identification assumption in this paper.

supply at time t can be defined as:

$$L_{sxt} = ((1 - \sum \varphi_{sxgt})N_{sxt}^{\psi} + \varphi_{sxgt}M_{sxgt}^{\psi})^{1/\psi} \qquad g = 1, 2, 3 \tag{2.4}$$

where N_{sxt}, within the education-experience group sx, is the labour supplied by natives, M_{sxgt} is the labour supplied by immigrants with g language skill groups. The parameter φ_{sxgt} is the relative productivity parameter across g language groups. The parameter ψ is equal to $1 - 1/\sigma_{EXG}$, where σ_{EXG} is the elasticity of substitution across natives and immigrants, who differ in language skills. The main estimating equation to obtain the estimates of the elasticity of substitution between immigrants and natives with respect to language skills is derived from Equation (2.4).

2.5 The estimates of substitutability for skill groups

2.5.1 The estimates of substitutability for immigrant language groups and natives

Using the CES model, the first order conditions of the production technology with respect to native and immigrant worker groups, as derived from equation (2.4), imply

$$\ln(\frac{w_{sxgt}^{M}}{w_{sxt}^{N}}) = \ln(\frac{\varphi_{sxgt}}{(1 - \sum \varphi_{sxgt})}) - \frac{1}{\sigma_{EXG}} \ln(\frac{M_{sxgt}}{N_{sxt}}) \quad g = 1, 2, 3 \tag{2.5}$$

where $\frac{w_{sxgt}^{M}}{w_{sxt}^{N}}$ is the relative wage of immigrant language groups and natives; $\frac{M_{sxgt}}{N_{sxt}}$ is the relative supply of the two groups; and $\frac{\varphi_{sxgt}}{(1-\sum \varphi_{sxgt})}$ is the demand-shifting relative immigrant-native productivity factor in education group s, experience group x, and language group g. I define an immigrant's language group based on the distance of her to English/French, as described in Section 2.3 above.[17] First, I estimate the average elasticity of substitution between immigrant language groups and natives while taking into account the differences in their education levels and years of labour market experience as well. Then, I provide

[17] As in Chapter 1, I use the distance of home language or mother tongue to English if an immigrant resides outside of Quebec and the distance of home language or mother tongue to French if an immigrant resides in Quebec.

the estimates for each language group separately. In my analysis, the choice of cut-off for the three language skill groups is admittedly arbitrary. As a robustness exercise, I consider two alternative grouping for language skills. The selection rules for alternative groups are provided in Appendix 2.4.

In order to control for demand-shifting factors, I use time-invariant education-experience-language skill, π_{sxg}, and time fixed effects, π_t. I assume that error terms are uncorrelated with the demand-shifting factors and the relative labour supply for immigrants and natives. Hence, the specified equation becomes

$$\ln(\frac{w^M_{sxgt}}{w^N_{sxt}}) = \pi_{sxg} + \pi_t - \frac{1}{\sigma_{EXG}} \ln(\frac{M_{sxgt}}{N_{sxt}}) + \epsilon_{sxgt} \qquad g = 1, 2, 3 \tag{2.6}$$

Since the error terms are likely correlated within skill cells, they are clustered in the estimations. In order to check for the robustness of my results, I use different specifications to control for the demand-shifting factors. First, I assume that the demand-shifting factors are time-invariant. Time effects, π_t, are assumed to be zero. Second, I assume that skill effects are time-variant and consider four alternative strategies to allow for this variation: (1) using time fixed effects; (2) including education fixed effects interacting with time fixed effects; (3) including experience fixed effects interacting with time fixed effects; (4) including language-by-time fixed effects. The first strategy allows for the changes that are common to all individuals but vary over time such as the changes in the census questions on spoken languages. Time effects circumvent the challenges raised on this issue in the previous chapter and minimize their influence on the results. The second strategy allows for the change in the quality in education groups that affect relative productivity of immigrants. The third strategy allows for the change in the quality of recent immigrant cohorts. The fourth strategy allows for the change in the composition of language groups. If the estimate of the inverse of the elasticity of the substitution is close to zero, which means degree of substitutability approaches infinity, immigrants and natives are perfect substitutes. Otherwise, imperfect substitutability exists between the two groups.

For the estimates of the inverse of the elasticity of substitution, $(-1/\sigma_{EXG})$, I have 270 available observations as a result of 45 skill groups (3x5x3) in each Census at the fourth

level.[18] The results for all language skill groups based on both mother tongue and home language are shown in Table 2.8, where each row refers to the language used as a reference for categorization and the region. There are two key findings in this table. First, there is almost perfect substitutability between immigrants and natives when immigrants are categorized based on the distance of their home language to English or French. The only exception occurs for the immigrants outside of Quebec, when skill effects, time effects, and time-variant education and experience effects are controlled for. In this case, the estimate of $(-1/\sigma_{EXG})$ for home language is -0.030 and it is statistically significant at 95% confidence level. Second, categorizing immigrants based on the information on mother tongue reveals a different story. For these regions, the estimates refer to imperfect substitutability, particularly for Canada and Canada outside of Quebec. The estimates of $(-1/\sigma_{EXG})$ range from -0.028 to -0.056 across specifications. All estimates are statistically significant at 95% confidence level. As in the case of home language-based analysis, the estimates of $(-1/\sigma_{EXG})$ are not stable across alternative specifications for Quebec. However, the experience by time fixed effects add substantial information to the model. The estimates of $(-1/\sigma_{EXG})$ range between -0.013 and -0.071 in Quebec, and most coefficients are statistically insignificant although all estimates have the expected negative sign, contrary to when I use home language.

Home language, on the one hand, is a better approximation for language skills relative to mother tongue but a choice of adopting one of the official languages as home language could be endogenous. As a result, a relative supply increase might be positively associated with higher relative wages, in which case the estimated coefficients would underestimate the impact of relative supply on relative wage. Mother tongue, on the other hand, is not a choice variable (as opposed to the language used at home) but combines high and low language skilled immigrants, which introduces a measurement issue. This measurement error might not be "classical": it might be correlated with the language distance. This source of measurement error does not seem to push the estimates toward zero. On the contrary, the estimates from mother tongue-based categorizations are larger in absolute value than those

[18] Following Borjas (2003) and Borjas et al. (2008), the weights at the fourth level are defined as $(n_{imm} * n_{natives})/(n_{imm} + n_{natives})$, where n_{imm} is the number of immigrants within respective language groups and $n_{natives}$ is the number of natives in the relevant skill cell in the wage sample.

from home language-based categorizations, suggesting imperfect substitutability between immigrants and natives. One of the reasons for these large estimates might be the fact that adopting official languages is negatively correlated with the distance of mother tongue. This might contribute to a negative relationship between relative supply and relative wage and enhance the degree of substitutability between immigrants and natives. Alternatively, strong evidence for imperfect substitutability from mother tongue-based specifications might refer to attributes of immigrant groups other than language skills such as the lack of information in the labour market institutions and ethnic social networking.

The difference in the estimates of the inverse elasticity of substitution for Quebec and the rest of Canada should be considered with caution. In chapter 1, I document that immigrant characteristics and labour market performance in Quebec and the rest of Canada differ by language group. The literature also finds evidence for structural differences across the two labour markets in Quebec and the rest of Canada. For example, Nadeau and Seckin (2010) examine immigrants' wage differences in the two regions. They suggest that higher wages of immigrants in all Canadian provinces outside of Quebec relative to immigrants' wages in Quebec cannot be explained by Quebec's autonomy in selecting immigrants. The immigrants in the rest of Canada experience higher wages, because the pay reward for becoming a citizen rather than being a permanent resident has disappeared in the other provinces, but not in Quebec, which generates a dampening effect on wages in Quebec.

Most of the estimates of the inverse elasticity of substitution for Quebec are almost zero, pointing to perfect substitutability between immigrants and natives. These estimates are also unstable, sometimes, with an unexpected positive sign, almost zero and imprecise due to large standard errors. These results raise concerns for a number of reasons. First, dividing the national market into local markets fails to consider selective out-migration. For example, if previous immigrant cohorts with low language skill levels in French prefer to move to other Canadian provinces, this might push the estimates toward zero. Second, Aydemir and Borjas (2011) and Borjas (2014) emphasize that as a result of the sampling of workers in a particular labour market, the estimates are subject to sampling error, even when a large data set such as microdata files of the Canadian census are used. As cell

sizes become smaller, sampling error becomes a serious issue, leading to a specific type of measurement error, attenuation bias. To evaluate the sensitivity of the estimates to small sample sizes, immigrant language groups are incorporated into a single group in Quebec. The size of the estimates slightly increases, but the main conclusions stay the same.

Before presenting the results for separate language groups, I perform two robustness checks. First, I evaluate the sensitivity of the results to sub-periods. As mentioned earlier, non-permanent residents have been included in the census since 1991 due to their increasing presence in the Canadian economic and social life. I include these people in my analysis for a number of reasons due to increasing numbers of temporary foreign workers as full-time workers in the Canadian economy. Additionally, there have been significant changes to the census questions on language background, in particular from 1991 onward. To check the sensitivity of my results to the inclusion of temporary foreign workers and the changes in language-related questions, I exclude the first two censuses and obtain the inverse elasticity of substitution for the 1991-2006 sub-period. The size of estimates for the inverse elasticity of substitution between immigrant language groups and natives become smaller in general for the 1991-2006 period whereas the main conclusions do not change.[19] Home language-based analysis suggests perfect substitutability between immigrants and natives, while mother tongue-based estimates indicate a small degree of imperfect substitutability between immigrants and natives. Thus, the exclusion of earlier years does not change the main conclusions.

Second, I evaluate the sensitivity of the estimates to the change in language definition in Quebec. In the first set of estimates, the immigrants residing in Quebec are divided based on the distance of the reported language from French. However, as previously reported in Table 2.3, the percentage of male immigrants speaking English in Quebec is higher than the percentage of male immigrants speaking French in all the other provinces.[20] In Quebec,

[19]For example, imposing the same identifying assumptions as those in the column (V) in Table 2.8, the estimates for home language-based specifications become -0.003 for Canada, -0.005 for Canada outside Quebec, -0.004 for Quebec. Mother tongue-based estimates become -0.029 for Canada, -0.032 for Canada outside of Quebec, and -0.013 for Quebec. The estimates for Canada and Canada outside of Quebec are statistically significant at the 95 confidence level for mother tongue-based specifications.

[20]The percentage is calculated for male immigrants reporting at least one week of work in the census year and no school attendance, and aged 21 to 64.

almost 17% of male immigrants reported English as mother tongue in 1981. Although this percentage decreased by half over time, 8.5% of immigrants still reported English as their mother tongue in Quebec in 2006. Grenier and Nadeau (2016) find that English-speaking immigrants receive almost no reward for speaking French at work while French-speaking immigrants are rewarded significantly for speaking English at work in the Montreal labour market. In addition, given that there is a higher percentage of jobs requiring English in Quebec as compared to requiring French in the rest of Canada, I categorize immigrants based on the distance of reported languages to English as well. The main results do not change. Most estimates are close to zero, with the expected negative sign. Imposing the same identification assumptions as those in the column (V) in Table 2.8, the estimates are found to be -0.023 for home language-based specification and -0.043, statistically significant at the 95 confidence level, for mother tongue-based specifications.[21,22]

In Table 2.9, I present the inverse elasticity of substitution, $(-1/\sigma_{EXG})$, between immigrants and natives for separate language groups. Instead of estimating an "average" elasticity of substitution for all language groups assuming that immigrant-native subtitutability is to be the same across language groups, I estimate three different elasticities between immigrants with particular language skills -high, medium, and low– and the Canadian-born workers. Each regression has 90 observations, 15 skill groups (3x5) over six Censuses. The first three columns present the estimates of $(-1/\sigma_{EXG})$ under the assumption that the demand-shifting factors are time-invariant. These effects are approximated by only including skill fixed effects in the analysis. The last three columns present the estimates of $(-1/\sigma_{EXG})$, which includes education-experience fixed effects, time fixed effects and time-variant skill effects that allow for time-variant demand-shifting factors.

The estimated coefficients in Table 2.9 suggest that the inverse elasticity of substitution between low language-skilled immigrants and natives is negative and statistically significant for all specifications for Canada and Canada outside of Quebec. This is not the

[21] As another exercise, I defined high language-skilled immigrants in Quebec as those speaking English or French, the estimated coefficient for the specification (V) is -0.012, statistically insignificant. Most estimates are also estimated as being around zero.

[22] Although the immigrants reporting English and residing in Quebec predominantly settle in Montreal, the same analysis cannot be repeated for the immigrants in Montreal and the rest of Quebec due to small sample sizes, which, in some cases, do not satisfy confidentiality requirements of data.

case for Quebec where the estimates are either close to zero or statistically insignificant. For Canada other than Quebec, the estimate of $(-1/\sigma_{EXG})$ is -0.064 if the demand-shifting factors are assumed to be constant over time and -0.040 if the demand-shifting factors are to be time-variant. Parallel to the case of analyzing all language groups together, the estimate becomes larger in size when mother tongue is considered. The differences in the estimates for high and medium language-skilled immigrant groups across specifications vary notably across specifications and require extra attention for interpretation. For example, I estimate statistically significant positive coefficients, i.e. negative elasticity of substitution in some cases, suggesting the CES is no longer supported. Also, for high language-skilled immigrants in Canada, assuming time-invariant demand-shifting factors provides almost zero estimates with a positive sign. These estimates become large with an expected negative sign when the demand-shifting factors are assumed to vary over time. This significant degree of imperfect substitutability for high language-skilled immigrants in Canada suggests that time effects are necessary to control for substantial compositional changes and regional shifts occurred among English-speaking immigrants during the period. Since Borjas (2014) emphasizes that the identification assumptions are not testable or verifiable, it is not possible to justify one assumption over another for high and low language-skilled immigrants but the estimates of the elasticity of substitution for low language-skilled immigrants are stable under alternative identifying assumptions. Hence, an evidence of imperfect substitutability between the group of low language-skilled immigrants and natives.[23]

In summary, extending the models used in Borjas (2003) and Ottaviano-Peri (2008, 2012) to include a fourth level to account for differences in language skills among immigrants has led to four main conclusions. First, most of home language-based estimates point to almost perfect substitutability between immigrants and natives. Second, the estimates of the elasticity of substitution between immigrant language groups and natives are either zero or statistically insignificant across specifications for Quebec. Third, mother tongue-based estimations point to imperfect substitutability between natives and immigrants residing in Canada and in Canada outside of Quebec. Table 2.9 shows that whether home language or

[23] Appendix 2.4 presents the definitions of alternative language skill groups and Appendix Tables A2.3-A2.6 present results for these groups. The main conclusions discussed above are still valid.

mother tongue is considered might matter for high and medium language skilled immigrants residing in Canada outside of Quebec. However, low-language skilled immigrants in English seem imperfectly substitutable regardless of specification and language skill definitions outside of Quebec. Although the findings for Quebec point mainly to perfect substitutability between immigrant language groups and natives, when the signs of the estimates are expectedly negative, how much these results are subjected to sampling error is not known. Overall, considering home language as the more relevant representation of language skills for labour market integration, the estimations suggest that language skills per se might not be the only factor behind imperfect substitutability between immigrants and natives. The stronger results from mother tongue-based estimations might refer to factors other than language skills such as previous exposure to the institutional culture in Canada and networking that might play an important role for immigrant-native substitutability.

2.5.2 Measurement of average wages: Log mean versus mean log?

Borjas et al. (2008) show that one of the main reasons in the differences in findings on the degree of substitutability between immigrants and natives is different wage measures used in the empirical analysis. The studies following Borjas (2003) such as Aydemir and Borjas (2007) and Borjas et al. (2008) suggest following the convention in the literature and using the average of the logarithm of wages corresponding to the log of the geometric mean of wages. These studies find that immigrants and natives are perfect substitutes. Ottaviano and Peri (2008, 2012) use an alternative measure, the logarithm of the arithmetic mean of wages in their analysis to represent the average wage of a skill group, and they find that immigrants and natives are imperfect substitutes. In this study, I use both measures. In Tables 2.8 and 2.9, I present the estimates using the average of the logarithm of wages. I also present the estimates using the logarithm of the arithmetic mean of wages in Tables 2.10 and 2.11. Both tables categorize immigrants into language groups based on the distance of languages to English or French, depending on the region. On this basis, the estimates in these tables are comparable to those in Tables 2.8 and 2.9.

Table 2.10 presents the estimates of the average elasticity of substitution between immigrants and natives while all language groups are analyzed together. In Table 2.11, language groups are examined separately. As shown in Table 2.10, using the logarithm of arithmetic mean of wages provides estimates that point to a degree of imperfect substitutability between immigrant and native workers in the Canadian labour market as compared to using the average log of wages. Although the estimates of $(-1/\sigma_{EXG})$ are still smaller for home language-based categorizations than those for mother tongue-based categorization, they are generally larger in size as compared to the estimates in Table 2.8. For example, under column (V), the estimates of $(-1/\sigma_{EXG})$ for home language-based categorizations of immigrants range between -0.035 and -0.045 in Table 2.10, while the same estimates of $(-1/\sigma_{EXG})$ range between -0.023 and -0.030 in Table 2.8. However, in Table 2.11, the same conclusion is not reached for the estimates of $(-1/\sigma_{EXG})$ for separate language groups. Home language-based estimates for low language-skilled immigrants in Table 2.11 are very similar to those in Table 2.9 for Canada and for Canada outside of Quebec. For example, assuming that the demand-shifting factors are time-variant, $(-1/\sigma_{EXG})$ for Canada is estimated as -0.036 by the log of the arithmetic mean of wages and as -0.034 by the log of the geometric mean of wages. However, I find that the degree of substitutability between low language-skilled immigrants and natives are substantially different for mother tongue-based categorizations in Table 2.11 than those in Table 2.9, specifically for Canada and Canada outside of Quebec. Using the log of the geometric mean of wages measures points to imperfect substitutability between low language-skilled immigrants and natives at the degree of 16, while the log of the arithmetic mean of wages measures lead to perfect substitutability between the two.

These results show that the wage measure used in the analysis is indeed an important decision and might lead to two extreme inferences from the analysis. It should be noted that, as a result of Jensen's inequality, the estimate of the elasticity of substitution between immigrant language groups and natives is expected to be larger if the logarithm of arithmetic mean of wages is used in the analysis instead of the log of geometric means of wages. Hence, comparing the estimates in Tables 2.8 and 2.10, it is indeed found that using

the logarithm of arithmetic mean of wages provides larger estimates in size as compared to using the average log of wages when all language groups are analyzed together. However, estimations for separate language groups do not provide results in the expected direction. Both measures provide almost the same estimates of the elasticity of substitution between low language-skilled immigrants and natives in Canada and in Canada outside of Quebec if home language is considered. On the other hand, mother tongue-based estimations lead to opposite results for low language-skilled immigrants in these regions. These findings show that alternative measures might provide very different results for the estimate of the same parameter. I follow the convention as in Borjas (2003) and use the average of the logarithm of wages.

2.5.3 The estimates of substitutability for experience groups

The estimate of the elasticity of substitution across experience groups in the Canadian labour market is provided in the literature in previous studies such as Aydemir and Borjas (2007). In this section, I aim to provide the estimates for these parameters in a way that is comparable to those found in the literature by using the estimates of $(-1/\sigma_{EXG})$ and demand-shifting technology parameters. Hence, the estimates in Table 2.8 are used in order to obtain the estimate of the elasticity of substitution across experience groups.

Using the relative productivity parameter estimates, $\hat{\pi_{sxg}}$, from Table 2.8, the CES-aggregated labour supply is calculated by multiplying these estimates with corresponding labour supplies within skill groups and instrumented with the labour supplies of immigrants within experience-education skill cells.[24] As implied by the first order conditions of the production technology in equation (2.3), the wage function in the education-experience group can be written as:

$$\ln w_{sxt} = \vartheta_t + \vartheta_{st} + \vartheta_{sxt} - \frac{1}{\sigma_{EX}} \ln L_{sxt} \tag{2.7}$$

[24]Since the relative productivity parameters add up to one in the fourth-level, it is derived that, for an immigrant language group, $\hat{\varphi_{sxgt}} = \frac{\exp(\hat{\pi_{sxg}})}{1+\sum \exp(\hat{\pi_{sxg}})}$, and for natives, $\hat{\varphi_{sxt}} = \frac{1}{1+\sum \exp(\hat{\pi_{sxg}})}$.

where $\vartheta_t = \ln(\alpha_{Lt}) + (1-\rho)\ln(Y_t) + (\rho - \upsilon)\ln(L_t)$ is approximated by time fixed effects; $\vartheta_{st} = \ln(\beta_{st}) + (\upsilon - \mu)\ln(L_{st})$ is approximated by time fixed effects interacted with education fixed effects and, $\vartheta_{sxt} = \ln(\vartheta_{sxt})$ is approximated by education-experience fixed effects. Two assumptions are imposed upon experience-education-time fixed effects to deal with insufficient number of observations in the sample. I first assume that the changes within the education-experience cells are constant over time. Later, I relax this assumption and add education fixed effects interacted with time fixed effects along with experience fixed effects interacted with a linear time trend to the specification. Normalized weights representing the number of workers in the wage sample are used in the regressions, and standard errors are adjusted for clustering within education-experience cells. There are 90 observations available for the regression analysis.

The results from the 3rd level estimation are presented in Table 2.12. The first two columns use the estimates from the column (II) in Table 2.8 while the last two columns use those from the column (V) in Table 2.8. In other words, I first consider the specification assuming that the demand-shifting factors are time invariant at the fourth level. Then, I consider the specification controlling for all the systematic changes across education and experience groups over time. Regardless of the estimates of productivity parameters from the fourth level, the estimates of the elasticity of substitution across experience groups are statistically insignificant when it is assumed that the demand-shifting factors related to experience do not vary over time. If I include the experience fixed effects interacted with a linear time trend, the estimates are statistically significant with a negative sign, as presented in the second and the fourth columns. Aydemir and Borjas (2007) estimate the elasticity of substitution across experience groups to be around 7.5 by interacting experience fixed effects with a linear trend. I also prefer the specification including experience fixed effects interacted with a linear trend.[25] In my analysis, the estimate of the elasticity of substitution across experience groups, is found to be around 5 in Canada. It should be

[25]Following the literature, Aydemir and Borjas (2007) assert that this approximation better represents the data. In my empirical analysis, instead of experience effects interacted with a linear trend, I also consider experience effects interacted with time fixed effects, and I find that the estimates substantially differ across regions, even with unexpected signs of the estimates. As a result, by following Aydemir and Borjas (2007), I prefer to present the estimates from the specifications described above.

noted that Aydemir and Borjas (2007) also provides estimates of $-1/\sigma_X$ based on a model including only education-time fixed effects. These estimates are small, -0.03, statistically significant at the margin, and with an expected negative sign. Henceforth, the estimates presented in Table 2.12 are comparable to those in the literature under the same identifying assumptions for all Canada, and Canada outside Quebec.

2.5.4 The estimates of substitutability for education groups

In the second-level, the wage function for education groups can be written as:

$$\ln w_{st} = \vartheta_t + \vartheta_{st} - \frac{1}{\sigma_E} \ln L_{st} \tag{2.8}$$

where $\vartheta_t = \ln(\alpha_{Lt}) + (1-\rho)\ln(Y_t) + (\rho - \upsilon)\ln(L_t)$ is approximated by time fixed effects or a linear time trend; $\vartheta_{st} = \ln(\beta_{st}) + (\upsilon - \mu)\ln(L_{st})$ is approximated by education fixed effects, and education effects interacted with a linear time trend. The results from the 2nd level are presented in Table 2.13. As in Table 2.12, the same 4th level specifications used to estimate $(-\frac{1}{\sigma_X})$ are also used in these regressions. Also, the preferred estimate of $(-\frac{1}{\sigma_X})$ comes from the identifying assumption that controls for time-variant experience effects in the 3rd level estimations. Using the estimates of $(-\frac{1}{\sigma_{EXG}})$ and $(-\frac{1}{\sigma_X})$, I calculate the CES-aggregated labour supplies of education groups and instrument them with the immigrant labour supply within education cells. Normalized weights for the wage sample are used, and standard errors are adjusted for clustering within education cells. There are 18 observations available for my empirical analysis.

At the 2nd level CES estimation, the estimates of $-1/\sigma_E$ are mostly statistically significant at the 95% confidence level. Education-specific linear time trends are used to control for the systematic changes occurring over time. Additionally, I include education-specific non-linear trends. As previously suggested in the literature, since the underlying trend in Canadian data might not be captured with a linear trend adequately, I define a spline function for 1991, and interact it with education fixed effects. I add these education-specific splines to the specification, and the results are presented in the second and the fourth columns. It seems that alternative trend functions do not matter much in terms of

the size of estimates. Aydemir and Borjas (2007), by constructing a spline function based on decades, estimate the elasticity of substitution across education groups to be around 2.5. Although the estimates in Table 2.13 are comparable to those in the literature at the national level and at regional levels, it should be noted that different identifying assumptions are used to obtain these estimates.

Henceforth, by combining the estimates of the elasticity of substitution parameters at the 4th, 3rd, and 2nd level, I discuss the wage effects of immigration-induced labour supplies in the next section. I present simulation results for Canada outside of Quebec and Quebec separately.

2.6 Total wage effects of immigration

By using the estimates of the elasticity of substitution between different production factors at different CES-levels, I calculate total wage effects of immigrant-induced labour supply changes on the wages of the Canadian native workers and foreign-born workers between 1981 and 2006. In each education - experience - language skill cell, the change in the immigrant-induced labour supply within a skill cell is defined as $\frac{\Delta M_{i,j,g,t}}{M_{i,j,g,t}} = \frac{M_{s,x,g,2006} - M_{s,x,g,1981}}{M_{s,x,g,1981}}$, where s, x and g refer to education experience and language skill of an immigrant group, respectively.

The 4-th level CES function implies that total wage effect of a group of native workers with education s, experience x is as below:

$$\begin{aligned}\triangle \log w_{sx_N t} = \frac{1}{\sigma_{KL}} \sum_{i \epsilon S} \sum_{j \epsilon X} \sum_{g \epsilon G} \gamma^M_{ijg} \frac{\Delta M_{i,j,g,t}}{M_{i,j,g,t}} \\ + (\frac{1}{\sigma_E} - \frac{1}{\sigma_{KL}}) \frac{1}{\gamma_L} \sum_{i \epsilon S} \sum_{j \epsilon X} \sum_{g \epsilon G} \gamma^M_{ijg} \frac{\Delta M_{i,j,g,t}}{M_{i,j,g,t}} + (\frac{1}{\sigma_X} - \frac{1}{\sigma_E}) \frac{1}{\gamma_{L_s}} \sum_{j \epsilon X} \sum_{g \epsilon G} \gamma^M_{s,jg} \frac{\Delta M_{i,j,g,t}}{M_{i,j,g,t}} \\ + (\frac{1}{\sigma_{EXG}} - \frac{1}{\sigma_X}) \frac{1}{\gamma_{L_{sx}}} \sum_{g \epsilon G} \gamma^M_{s,xg} \frac{\Delta M_{s,x,g,t}}{M_{s,x,g,t}} \end{aligned} \tag{2.9}$$

and the expression of the total wage effect of an immigrant groups with education s, experience x, language skill g follows as:

$$\triangle \log w_{sxgt} = \frac{1}{\sigma_{KL}} \sum_{i \epsilon S} \sum_{j \epsilon X} \sum_{g \epsilon G} \gamma^{M}_{ijg} \frac{\Delta M_{i,j,g,t}}{M_{i,j,g,t}}$$
$$+ (\frac{1}{\sigma_E} - \frac{1}{\sigma_{KL}}) \frac{1}{\gamma_L} \sum_{i \epsilon S} \sum_{j \epsilon X} \sum_{g \epsilon G} \gamma^{M}_{ijg} \frac{\Delta M_{i,j,g,t}}{M_{i,j,g,t}} + (\frac{1}{\sigma_X} - \frac{1}{\sigma_E}) \frac{1}{\gamma_{L_s}} \sum_{j \epsilon X} \sum_{g \epsilon G} \gamma^{M}_{s,jg} \frac{\Delta M_{i,j,g,t}}{M_{i,j,g,t}}$$
$$+ (\frac{1}{\sigma_{EXG}} - \frac{1}{\sigma_X}) \frac{1}{\gamma_{L_{sx}}} \sum_{g \epsilon G} \gamma^{M}_{s,xg} \frac{\Delta M_{i,j,g,t}}{M_{i,j,g,t}} (-\frac{1}{\sigma_{EXG}}) \frac{\Delta M_{i,j,g,t}}{M_{i,j,g,t}} \tag{2.10}$$

where γ^{M}_{ijg} refers to the share of wages paid to immigrants with education level i, years of labour market experience j, and language skill g, while γ_L is the share of labour income, γ_{L_s} is the wage share of all workers with education s, and $\gamma_{L_{sx}}$ is the wage share of all workers with education s and experience x. As defined previously, $\sigma' s$ stand for the elasticity of substitution across skill groups at different CES levels. Using total wage effects within skill groups, we can calculate representative group wages by taking the weighted averages of total wage changes belonging to a particular skill cell. For instance, the representative average wage effect of natives is given by: $\bar{w^N}_t = \sum_{i \epsilon S} \sum_{j \epsilon X} \triangle \log w_{sxNt} \gamma^{N}_{ij_N} / \sum_{i \epsilon S} \sum_{j \epsilon X} \gamma^{N}_{ij_N}$. Similarly, that of immigrants belonging to any specific language group g follows by the expression : $\bar{w^M}_{g,t} = \sum_{i \epsilon S} \sum_{j \epsilon X} \triangle \log w_{sxgt} \gamma^{M}_{ijg} / \sum_{i \epsilon S} \sum_{j \epsilon X} \gamma^{M}_{ijg}$. The average wage effect of all immigrants can be expressed as : $\bar{w}^{M}_{t} = \sum_{i \epsilon S} \sum_{j \epsilon X} \sum_{g \epsilon G} \triangle \log w_{sxgt} \gamma^{M}_{ijg} \ / \ \sum_{i \epsilon S} \sum_{j \epsilon X} \sum_{g \epsilon G} \gamma^{M}_{ijg}$.

The simulated wage effects are obtained under the following set of assumptions: (i) the elasticity of substitution between capital and labour, σ_{KL}, is 1; (ii) the share of labour income in the Canadian economy, γ_L, is 0.66. It should be noted that, the degree of substitutability across education groups and across experience groups does not differ across regions. Moreover, I use the estimates provided in the literature, which are also very close to the estimates at the national level from the preferred specifications. Hence, $1/\sigma_E$ and $1/\sigma_X$ are assumed as 0.200 and 0.450, respectively. The estimates of the inverse of the elasticity of substitution across immigrant language groups and native workers are taken from Table 2.8. $1/\sigma_{EXG}$ is assumed to be 0.057 and 0.030 for mother tongue and home language in Canada outside of Quebec. In Quebec, the same inverse elasticities are assumed to be 0.073 and 0.

Table 2.14 presents actual wage changes in all Canadian provinces other than Quebec with respect to mother tongue and home language of immigrants. Tables 2.15 and 2.16 show simulated wage effects outside of Quebec on different groups of workers based on the formulas expressed above. In particular, Tables 2.15 and 2.16 provide simulated wage effects for immigrant groups based on the distance of home languages and mother tongue to English, respectively. I obtain each education group's wage changes by averaging total wage changes within skill cells by weighting them with education groups' shares in the 1981 wage bill. Similarly, the average changes in immigrants' and natives' wages are obtained by weighting their total wage changes by the their share in the 1981 wage bill. The first column presents average wages changes for all education groups in the economy across immigrant language groups, all immigrants, natives, and all workers. The last three columns provide simulated wage effects for separate education groups.

Capital adjustments are taken into consideration in simulations for long-run total wage effects as follows. The elasticity of substitution between capital and labour is assumed to be 1. Hence, the required capital adjustment is approximated with the weighted sum of labour supply changes within skill cells. The weights are defined as the income share of each education-experience-language skill group of immigrants. Then, these weights are multiplied by the income share of capital. Henceforth, capital adjustments are taken into consideration by adding a term to each total wage effect of a skill group and defined as:

$$\Delta K = \gamma_K (\sum \gamma_{s,x,g,t} \frac{\Delta M_{s,x,g,t}}{N_{s,x,g,t}}) \tag{2.11}$$

where γ_K is the income share of capital.

In Table 2.15, simulation results show that immigration reduced native-borns' wages by 4.2% while the wages of immigrants declined by 6.9% due to immigration between 1981 and 2006 outside of Quebec. Overall, average wages declined by 4.9%. The most educated were most severely affected by the immigration-induced supply shock. Also, immigrants with low language-capital levels experienced a 15.1% decline in their wages due to immigration. The largest decline is observed for the most educated with low language-capital. In the long-run, natives' wages increased by 0.6% as a result of immigration while

immigrant wages decreased by 2.1%. In Table 2.16, as expected due to a higher degree of imperfect substitutability, the impact of immigrant inflows is calculated to be smaller on native wages, and larger on immigrants than those in Table 2.15.[26]

Table 2.17 presents actual wage changes in Quebec. In Table 2.18, it is assumed that immigrants are natives are perfect substitutes, depending on home language-based language groups. In this case, wages in Quebec reduced by around 2.2% due to immigration. In Table 2.19, imperfect substitutability between immigrants and natives is assumed, which is based on mother tongue-based language groups. The first panel of Table 2.19 shows that the wage level of natives residing in Quebec decreased by 1.6% in the short-run. The second panel shows that it increased by by 0.6% in the long-run due to immigration. At the same time, immigration reduced immigrant wages by almost 7.2% in the short-run, by 4.9% in the long-run between 1981 and 2006.

There are three notable points to be taken from these simulations. First, simulated total wage effects indicate that the impact of increased immigrant inflows affected immigrants more than natives. Second, simulations show that highly educated immigrants with low language skills were most severely affected. For example, simulations show that the most educated immigrants with low language capital experienced approximately a 40% decline in wages outside of Quebec during the period. Third, comparing simulated wage effects with actual wage changes, immigration-induced supply shocks do not seem to explain wage changes much during the period. It is likely that other factors besides immigration supply shock affect wage trends in the Canadian labour market. However, it is likely that immigration prevented any increases in wages, particularly for previous immigrant cohorts due to large inflows.

2.7 Conclusion

The objective of this chapter is to estimate the impact of immigrant flows on wage rates, while incorporating for the first time immigrant language skills that play an important role in determining their human capital in the Canadian labour market. The previous litera-

[26] Appendix Table 2A.9 shows simulations for Canada.

ture suggests that immigrants and the Canadian-born workers are perfect substitutes in the labour market. This result arises from an assumption that all immigrants and natives can be thought of as a single group given their education and years of labour market experience. This study uses a fairly new linguistic distance measure in the Canadian literature, the Levenshtein Distance. By taking the advantage of this measure, I disaggregate immigrants into language skill groups instead of treating them as a single group. Allowing immigrants skills to further vary with respect to language skill leads to three key findings. First, I find that immigrant inflows have larger effects on the wages of immigrants than on natives. Second, analyzing language groups of immigrants separately suggests that low language-skilled immigrants are more likely to be imperfect substitutes for natives as compared to middle and high language-skilled immigrants. Third, simulations show that immigrant wages drop substantially if highly educated immigrants do not have high levels of language skills as well. For example, according to simulated wage effects, the decline in the average wage of the most educated immigrants with low language skills is 35.2% outside of Quebec between 1981 and 2006 in the long-run. Among education groups, the most educated immigrants are most negatively affected by the immigration-induced labour supply changes. Overall, the findings suggest that other characteristics of immigrants besides education and years of labour market experience might contribute to the degree of imperfect substitutability between immigrants and natives. In this case, the implication is that maintaining high levels of immigrant inflows might create further challenges for recent immigrant cohorts, in particular for the high educated with the most distant languages.

The different results for Quebec and the rest of Canada suggest that regional differences play an essential role for immigrant-native substitutability. In particular, the fact that Quebec has had autonomy in selecting immigrants since 1991, perfect substitutability between immigrant language groups and native might indeed refer to different characteristics of immigrants across the two labour markets. Further, the literature establishes that there are other contributing factors to the gap in immigrants' wages between the two regions such as disappearing pay premiums between citizens and permanent residents in Canada outside of Quebec, as suggested by Nadeau and Seckin (2010). Hence, the difference might

be attributable to the differences in not immigrant characteristics but in labour market characteristics in the two regions. Whatever the factors behind the different set of results, the difference in immigrants' wages in the two regions requires a closer look and will be explored in the future.

Obviously, the study has limitations that lead to more questions. First, the LD is an approximation for language skills. Does a better approximation of language skills provide a different set of results in terms of immigrant-native substitutability? Second, instead of directly estimating the impact of immigration, this study imposes a structure on the aggregate economy. The fact that underlying identification assumptions of this model are not testable or verifiable becomes restrictive in the impact analysis of immigration on wages. The literature suggests that combining the results from the structural model with the estimates from descriptive studies will provide more accurate estimates and policy suggestions. Third, the model defines skills in terms of education and experience although I take the extreme assumption that foreign experience is not valued at all in Canada. Instead of education-experience skill definitions, does the model based on occupation-specific skill definitions shed more light on immigrant-native substitutability? In the literature, it is well-documented that technological innovations destroy certain types of occupations, while rewarding technology-intensive occupations. Further, the literature establishes that language skills are associated with other occupation-specific skills. Instead of education-experience skill definitions, occupation-specific skill definitions combined with language skills might shed more light on what types of skills contribute to immigrant-native substitutability, if any. Most notably, the occupation-specific skill categorization of workers could lead to more practical policy suggestions such as targeting newcomers with a specific skill set or helping immigrants improve those specific skills once they arrive in Canada. Fourth, this study excludes female workers due to the issues related to their labour market participation. However, female workers' labour market attachment has increased over the last decades. Does the inclusion of female workers affect immigrant-native substitutability? Is the immigrant-native substitutability for females the same as the immigrant-native

substitutability for males? In the next chapter, I tackle this particular direction by adding female workers to the analysis.

References 2

[1] **Addison, David A. and Christopher Worswick**. "The impact of immigration on the earnings of natives: Evidence from Australian micro data". *The Economic Record*, 78(1):68–78, 2002.

[2] **Adsera, Alicia and Ana Ferrer**. "The effect of linguistic proximity on the occupational assimilation of immigrant men". IZA Discussion Paper, no. 9499, 2015.

[3] **Adsera, Alicia and Ana Ferrer**. "Occupational skills and labour market progression of married immigrant women in Canada". *Labour Economics*, 39:88–98, 2016.

[4] **Adsera, Alicia and Mariola Pytlikova**. "The role of language in shaping international migration". *The Economic Journal*, 125(586):49–81, 2015.

[5] **Aydemir, Abdurrahman and George J. Borjas**. "Cross-country variations in the impact of international migration: Canada, Mexico, and the United States". *Journal of the European Economic Association*, 5(4):663–708, 2007.

[6] **Aydemir, Abdurrahman and George J. Borjas**. "Attenuation bias in measuring the wage impact of immigration". *Journal of Labor Economics*, 1(4):69–112, 2011.

[7] **Aydemir, Abdurrahman and Mikal Skuterud**. "Explaining the deteriorating entry earnings of Canada's immigrant cohorts: 1966-2000". *Canadian Journal of Economics*, 38(2):641–671, 2005.

[8] **Baker, Michael and Dwayne Benjamin**. "The performance of immigrants in the Canadian labour market". *Journal of Labor Economics*, 13(3):369–405, 1994.

[9] **Beach, Charles, Alan. G. Green, and Christopher Worswick**. *Toward improving Canada's skilled immigration policy: An evaluation approach.* CD Howe Institute, Toronto, 2011.

[10] **Berger, Mark C.** "Demographic cycles, cohort size, and earnings". *Demography*, 26(2):311–321, 1989.

[11] **Bloom, David E., Gilles Grenier, and Morley Gunderson**. "The changing labour market position of Canadian immigrants". *Canadian Journal of Economics*, 28(4b):987–1005, 1995.

[12] **Borjas, George J.** "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labour market". *The Quarterly Journal of Economics*, 118(4):1335–1374, 2003.

[13] **Borjas, George J.** *Immigration Economics*. Harvard University Press, 2014.

[14] **Borjas, George J., Jeffrey Grogger, and Gordon H. Hanson**. "Imperfect substitution between immigrants and natives: A reappraisal". National Bureau of Economic Research Working Paper, no. 13887, 2008.

[15] **Borjas, George J., Richard B. Freeman, Lawrance F. Katz, John DiNardo, and John M. Abowd**. "How much do immigration and trade affect labor market outcomes?" *Brookings Papers on Economic Activity*, 1997(1):1–90, 1997.

[16] **Boyd, Monica and Xingshan Cao**. "Immigrant language proficiency, earnings, and language policies". *Canadian Studies in Population*, 36(1-2):63–86, 2009.

[17] **Butcher, Kristin F. and David Card**. "Immigration and wages: Evidence from the 1980's". *The American Economic Review*, 81(2):292–296, 1991.

[18] **Card, David**. "The impact of Mariel Boatlift on the Miami labor market". *Industrial and Labor Relations Review*, 43(2):245–257, 1990.

[19] **Card, David**. "Is the new immigration really so bad?" *The Economic Journal*, 115(506):F300–F323, 2005.

[20] **Card, David**. "Immigration and inequality". National Bureau of Economic Research Working Paper, no. 14683, 2009.

[21] **Card, David and Giovanni Peri**. "Immigration economics by George J. Borjas: A review essay". *Journal of Economic Literature*, 54(4):1333–1349, 2016.

[22] **Chiswick, Barry R. and Paul W. Miller**. "The endogeneity between language and earnings: International analyses". *Journal of Labor Economics*, 13(2):246–288, 1995.

[23] **Chiswick, Barry R. and Paul W. Miller**. "A model of destination-language acquisition: Application to male immigrants in Canada ". *Demography*, 38(3):391–409, 2001.

[24] **Chiswick, Barry R. and Paul W. Miller**. "Linguistic distance: A quantitative measure of the distance betWeen English and other languages". *Journal of Multilingual and Multicultural Development*, 26(1):1–11, 2005.

[25] **Ferrer, Ana, David A. Green, and William C. Riddell**. "The effect of literacy on immigrant earnings". *Journal of Human Resources*, 41(2):380–410, 2006.

[26] **Ferrer, Ana, Garnett Picot, and William C. Riddell**. "New directions in immigration policy: Canada's evolving approach to the selection of economic immigrants". *International Migration Review*, 48(3):846–867, 2014.

[27] **Frenette, Marc and Renè Morissette**. "Will they ever converge? Earnings of immigrants and Canadian-born workers over the last two decades". Analytical Studies Branch Research Paper Series, no. 215. Statistics Canada Catalogue 11F0019MIE. Ottawa: Statistics Canada, 2003.

[28] **Friedberg, Rachel M. and Jennifer Hunt**. "The impact of immigrants on host country wages, employment and growth". *The Journal of Economic Perspectives*, 9(2):23–44, 1995.

[29] **Grant, Mary L.** "Do immigrants affect the wage of native-born workers? Evidence from Canada", 1998. Institute for Policy Analysis. University of Toronto. `http://hd.handle.net/1807/13124`.

[30] **Grant, Mary L.** "Evidence of new immigrant assimilation in Canada". *Canadian Journal of Economics*, 32(4):930–955, 1999.

[31] **Green, Alan G. and David A. Green**. "The economic goals of Canada's immigration policy: Past and present". *Canadian Public Policy*, 25(4):425–451, 1999.

[32] **Green, David A. and Christopher Worswick**. "Entry earnings of immigrant men in Canada: The roles of labour market entry effects and returns to foreign experience". In Ted McDonald, Elizabeth Ruddick, Arthur Sweetman, and Christopher Worswick, editor, *Canadian Immigration: Economic Evidence for A Dynamic Policy Environment*, pages 77–110. McGill-Queen's University Press, Kingston, ON, 2010.

[33] **Green, David A. and Christopher Worswick**. "Immigrant earnings profiles in the presence of human capital investment: Measuring cohort and macro effects". *Labour Economics*, 19(2):241–259, 2012.

[34] **Grenier, Gilles and Serge Nadeau**. "English as the lingua franca and the economic value of other languages: The case of the language of work in the Montreal labor market". In Michele Gazzola and Bengt-Arne Wickstrom, editor, *The Economics of Language Policy*, pages 267–312. MIT Press, Cambridge, Massachusetts, 2016.

[35] **Hou, Feng and Garnett Picot**. "The deteriorating economic welfare of immigrants and possible causes: Update 2005". Analytical Studies Branch Research Paper Series, no. 356. Statistics Canada Catalogue 11F0019M. Ottawa: Statistics Canada, 2014.

[36] **Isphording, Ingo E. and Sebastian Otten**. "The costs of Babylon–Linguistic distance in applied economics". *Review of International Economics*, 21(2):354–369, 2013.

[37] **Katz, Lawrence F. and Kevin M. Murphy**. "Changes in relative wages, 1963-1987: Supply and demand factors". *The Quarterly Journal of Economics*, 107(1):35–78, 1992.

[38] **Lewis, Ethan**. "Immigrant-native substitutability and role of language". In David Card and Steven Raphael, editor, *Immigration, Poverty, and Socioeconomic Inequality*, pages 60–97. The Russell Sage Foundation, New York, 2013.

[39] **Lewis, Paul, Gary S. Simons, and Charles D. Fenning (Ed.)**. *Ethnologue: Languages of the World*. Seventeenth edition. International, Dallas, Texas, 2013. Online version: `http://www.ethnologue.com`.

[40] **Morin, Louis-Philippe**. "Do men and women respond differently to competition? Evidence from a major education reform". *Journal of Labor Economics*, 33(2):443–491, 2015.

[41] **Nadeau, Serge and Aylin Seckin**. "The immigrant wage gap in Canada: Quebec and the rest of Canada". *Canadian Public Policy*, 36(3):265–285, 2010.

[42] **Okkerse, Liesbet**. "How to measure labour markets effects of immigration: A review". *Journal of Economic Surveys*, 22(1):1–30, 2008.

[43] **Oreopoulos, Philip**. "Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes". *American Economic Journal: Economic Policy*, 3(4):148–171, 2011.

[44] **Ottaviano, Gianmarco I. P. and Giovanni Peri**. "Immigration and national wages: Clarifying the theory and the empirics". National Bureau of Economic Research Working Paper, no. 14188, 2008.

[45] **Ottaviano, Gianmarco I. P. and Giovanni Peri**. "Rethinking the effects of immigration on wages". *Journal of the European Economic Association*, 10(1):152–197, 2012.

[46] **Picot, Garnett and Arthur Sweetman**. "The deteriorating economic welfare of immigrants and possible causes: Update 2005". Analytical Studies Branch Research Paper Series, no. 262. Statistics Canada Catalogue 11F0019MIE. Ottawa: Statistics Canada, 2005.

[47] **Picot, Garnett and Feng Hou**. "The rise in low-income rates among immigrants in Canada". Analytical Studies Branch Research Paper Series, no. 198. Statistics Canada Catalogue 11F0019MIE. Ottawa: Statistics Canada, 2003.

[48] **Picot, Garnett and Feng Hou**. "Chronic low income and low-income dynamics among recent immigrants". Analytical Studies Branch Research Paper Series, no. 364. Statistics Canada Catalogue 11F0019M. Ottawa: Statistics Canada, 2014.

[49] **Picot, Garnett and Patrizio Piraino**. "Immigrant earnings growth: Selection bias or real progress?" Analytical Studies Branch Research Paper Series, no. 340. Statistics Canada Catalogue 11F0019M. Ottawa: Statistics Canada, 2012.

[50] **Picot, Garnett, Feng Hou, and Simon Coulombe**. "Chronic low income and low-income dynamics among recent immigrants". Analytical Studies Branch Research Paper Series, no. 294. Statistics Canada Catalogue 11F0019MIE. Ottawa: Statistics Canada, 2007.

[51] **Skuterud, Mikal**. "Language skills in the new economy and the deteriorating labour market performance of Canada's immigrant workers". Report, Human Resources Development and Skills Canada, 2011.

[52] **Statistics Canada**. "Immigration in Canada: A portrait of the foreign-born population". 2006 Statistics Canada Census Catalogue no:97-557-XIE. Ottawa: Statistics Canada, 2007.

[53] **Statistics Canada**. "Changing labour force". 2006 Statistics Canada Census Catalogue no: 97-559-X. Ottawa: Statistics Canada, 2008.

[54] **Sweetman, Arthur, and Casey Warman**. "Former temporary foreign workers and international students as sources of permanent immigration". *Canadian Social Trends*, 40(4):391–407, 2014.

[55] **Thomas, Derrick**. "Foreign nationals working temporarily in Canada". *Canadian Social Trends*, 90:34–48, 2010.

[56] **Tu, Jiong**. "The impact of immigration on the labour market outcomes of native-born Canadians". IZA Discussion Paper, no. 5120, 2010.

[57] **Warman, Casey and Christopher Worswick**. "Technological change, occupational tasks, and declining immigrant outcomes: Implications for earnings and income inequality in Canada". *Canadian Journal of Economics*, 48(2):736–772, 2015.

[58] **Welch, Finis**. "Effects of cohort size on earnings: The baby boom babies' financial bust". *Journal of Political Economy*, 87(5):65–97, 1979.

[59] **Wooldridge, Jeffrey M.** *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, 2002.

Table 2.1: Sample construction: The number of observations and sample restrictions

	Census Years					
The Employment Sample	**1981**	**1986**	**1991**	**1996**	**2001**	**2006**
Initial Sample Size	5,036,955	5,241,565	5,624,805	6,025,565	6,080,915	6,470,470
Males aged 21-64	1,412,410	1,528,670	1,654,680	1,763,160	1,781,445	1,920,880
Not living in Collective Dwellings	1,358,175	1,472,335	1,601,025	1,713,230	1,738,390	1,875,860
At least a week of work in a Census year	1,261,060	1,321,720	1,431,805	1,468,435	1,512,605	1,632,225
No school attendance	1,225,840	NA*	1,277,785	1,314,450	1,364,055	1,445,910
With 1 to 40 years of labour market experience	1,017,585	1,204,030	1,187,280	1,231,940	1,282,950	1,378,430
With clearly identifiable languages	1,017,585	1,204,030	1,187,280	1,231,940	1,282,950	1,378,430
The Wage Sample						
Work as full-time wage/salary workers	840,765	975,110	979,150	965,065	1,026,980	1,101,070
Report at least 40 weeks of work	700,365	780,970	800,125	789,240	883,200	955,185
Report positive wages	689,160	767,065	786,935	779,105	870,910	914,225

Source: The Canadian Censuses between 1981 and 2006.
Notes: The number of observations are rounded to a base of 5. * In 1986, the information regarding school attendance is not available.

Table 2.2: The percentages of selected home languages spoken by immigrants

	LD Index		Canada		Canada out. of Quebec		Quebec	
	To English	**To French**	**1981**	**2006**	**1981**	**2006**	**1981**	**2006**
English	0.00	91.02	62.32	47.51	67.53	51.85	30.11	19.14
Dutch	63.22	91.06	0.43	0.12	0.47	0.14	0.16	0.03
German	72.21	95.87	2.40	0.64	2.57	0.70	1.37	0.23
Italian	88.94	73.15	9.38	0.64	8.17	0.52	16.90	1.43
Czech	90.98	90.49	0.26	0.10	0.27	0.11	0.19	0.05
French	91.02	0.00	4.39	5.05	0.60	0.57	27.81	34.37
Spanish	93.08	81.07	1.06	2.61	0.90	2.15	2.09	5.65
Polish	93.80	92.89	1.03	1.69	1.04	1.83	0.98	0.77
Portuguese	95.18	74.36	3.29	1.35	3.26	1.34	3.48	1.43
Ukranian	95.18	93.84	0.74	0.23	0.78	0.26	0.52	0.06
Hungarian	95.22	100.65	0.72	0.24	0.72	0.25	0.66	0.16
Hindi	96.17	89.52	0.26	0.66	0.29	0.74	0.10	0.15
Greek	97.15	95.08	2.38	0.40	1.74	0.28	6.37	1.20
Arabic	97.95	97.20	0.64	1.72	0.55	1.24	1.22	4.86
Korean	99.12	102.74	0.33	1.13	0.37	1.27	0.06	0.25
Punjabi	100.58	95.59	1.03	4.11	1.17	4.60	0.15	0.92
Chinese	101.18*	101.88	3.58	10.44	3.90	11.34	1.63	4.56
Tagalog	102.13	102.75	0.65	1.81	0.73	2.00	0.18	0.62
Vietnamese	104.06	101.81	0.40	1.74	0.31	1.67	0.96	2.23
Other languages	-	-	4.70	17.79	4.64	17.17	5.05	21.90

Source: The Canadian Censuses between 1981 and 2006.

Notes: * As of 2006, more options were made available for respondents to report the languages Chinese immigrants speak. The included languages under Chinese category are: Mandarin, Cantonese, Chaochow, Fukien, Hakka, Shanghainese, Taiwanese, Chinese, n.o.s., where n.o.s. refers to not otherwise specified. The LD index for Chinese represents the distance between the Canadian official languages and Mandarin, one of the most frequently reported language among all Chinese languages. "Other" refers to all reported home languages beside the selected ones above. The sample includes male immigrants reporting at least one week of work in the reference year and no school attendance, aged 21 to 64.

Table 2.3: The percentages of selected mother tongues spoken by immigrants

	LD Index		Canada		Canada out. of Quebec		Quebec	
	English	**French**	**1981**	**2006**	**1981**	**2006**	**1981**	**2006**
English	0.00	91.02	35.86	24.47	38.97	26.91	16.67	8.52
Dutch	63.22	91.06	4.62	1.18	5.21	1.32	0.94	0.27
German	72.21	95.87	8.51	2.55	9.18	2.76	4.33	1.15
Italian	88.94	73.15	13.84	3.33	12.43	3.04	22.57	5.27
Czech	90.98	90.49	0.59	0.35	0.62	0.38	0.40	0.18
French	91.02	0.00	3.56	3.14	0.91	0.83	19.96	18.24
Spanish	93.08	81.07	1.52	4.65	1.24	3.88	3.26	9.64
Polish	93.80	92.89	2.29	2.87	2.31	3.10	2.15	1.36
Portuguese	95.18	74.36	4.23	3.14	4.20	3.12	4.44	3.27
Ukranian	95.18	93.84	1.39	0.42	1.48	0.46	0.82	0.17
Hungarian	95.22	100.65	2.02	0.66	2.04	0.68	1.92	0.55
Hindi	96.17	89.52	0.46	1.20	0.49	1.34	0.23	0.28
Greek	97.15	95.08	3.23	0.97	2.53	0.80	7.53	2.14
Arabic	97.95	97.20	1.42	3.65	1.18	2.49	2.90	11.19
Korean	99.12	102.74	0.37	1.34	0.41	1.50	0.09	0.29
Punjabi	100.58	95.59	1.18	5.26	1.34	5.87	0.19	1.22
Chinese	101.18*	101.88	4.22	12.60	4.60	13.72	1.91	5.25
Tagalog	102.13	102.75	0.91	3.32	1.01	3.66	0.28	1.06
Vietnamese	104.06	101.81	0.39	2.19	0.31	2.11	0.93	2.74
Other languages	-	-	9.39	22.71	9.54	22.02	8.46	27.23

Source: The Canadian Censuses between 1981 and 2006.

Notes: * As of 2006, more options were made available for respondents to report the languages Chinese immigrants speak. The included languages under Chinese category are: Mandarin, Cantonese, Chaochow, Fukien, Hakka, Shanghainese, Taiwanese, Chinese, n.o.s., where n.o.s. refers to not otherwise specified. The LD index for Chinese represents the distance between the Canadian official languages and Mandarin, one of the most frequently reported language among all Chinese languages. "Other" refers to all reported mother tongues beside the selected ones above. The sample includes male immigrants reporting at least one week of work in the reference year and no school attendance, aged 21 to 64.

Table 2.4: The difference between the mean log wage of immigrants and natives in Canada outside of Quebec

		1981			1991			2001			2006		
Education	**Experience**	**Language Groups**			**Language Groups**			**Language Groups**			**Language Groups**		
		H	**M**	**L**	**H**	**M**	**L**	**H**	**M**	**L**	**H**	**M**	**L**
High Sch. Dropouts or	1-5	0.040	-0.040	-0.192	0.030	0.087	-0.102	0.084	0.048	-0.119	0.043	0.020	-0.044
Graduates	6-10	0.035	-0.027	-0.055	0.010	-0.034	-0.179	-0.011	-0.068	-0.242	0.014	0.003	-0.162
	11-15	0.054	-0.062	-0.119	-0.018	-0.142	-0.227	-0.055	-0.085	-0.327	-0.010	-0.158	-0.276
	16-20	0.061	-0.134	-0.284	0.006	-0.137	-0.202	-0.065	-0.196	-0.358	-0.038	-0.117	-0.290
	21-40	0.025	-0.123	-0.195	0.029	-0.120	-0.292	0.024	-0.127	-0.296	0.005	-0.148	-0.298
Voc. Cert. or	1-5	0.124	0.018	-0.051	0.040	-0.046	-0.069	0.072	-0.012	-0.156	-0.009	-0.137	-0.202
Below Bachelor Degr.	6-10	0.055	-0.009	-0.037	0.036	-0.106	-0.170	-0.004	-0.122	-0.191	-0.032	-0.099	-0.264
	11-15	0.067	-0.024	-0.047	0.076	-0.029	-0.147	-0.026	-0.136	-0.184	-0.050	-0.161	-0.279
	16-20	0.047	-0.094	-0.030	0.046	-0.111	-0.102	0.002	-0.202	-0.155	-0.046	-0.153	-0.244
	21-40	0.031	-0.094	-0.090	0.038	-0.110	-0.081	0.032	-0.102	-0.131	-0.006	-0.146	-0.206
Bachelor Degr. or	1-5	0.132	0.103	-0.076	0.079	-0.112	-0.107	0.061	-0.079	-0.154	0.051	-0.066	-0.169
Above	6-10	0.028	-0.067	-0.144	0.023	-0.071	-0.196	-0.015	-0.075	-0.250	-0.019	-0.100	-0.223
	11-15	0.019	-0.014	-0.129	0.015	-0.136	-0.188	-0.098	-0.277	-0.329	-0.081	-0.214	-0.354
	16-20	0.007	-0.043	-0.031	-0.016	-0.189	-0.200	-0.004	-0.167	-0.337	-0.077	-0.331	-0.336
	21-40	0.008	-0.151	-0.150	0.000	-0.071	-0.172	-0.009	-0.147	-0.201	-0.032	-0.177	-0.252

Source: The Canadian Censuses between 1981 and 2006.

Notes: The wage sample includes the full-time/full-year male workers residing in all Canadian provinces outside of Quebec, with reported positive wages, aged 21 to 64. The wage gap is the difference between the average of the logarithm of real weekly wages of immigrants and natives. The columns H, M and L refer to wage gaps between high language-skilled, medium language-skilled and low language-skilled immigrants and natives, respectively. The language skill groups are based on the LD referring to distance of home language to English. High language-skilled immigrants report English as home language. Medium language-skilled immigrants have a language whose LD is lower than that of Hungarian. The rest of immigrants are grouped into the low language skill group. A negative value indicates that immigrants have lower wages than natives.

Table 2.5: The ratio of immigrant to native labour supply in Canada outside of Quebec

		1981			1991			2001			2006		
Education	**Experience**	**Language Groups**			**Language Groups**			**Language Groups**			**Language Groups**		
		H	**M**	**L**	**H**	**M**	**L**	**H**	**M**	**L**	**H**	**M**	**L**
High Sch. Dropouts or	1-5	0.15	0.05	0.09	0.22	0.12	0.21	0.16	0.07	0.19	0.15	0.07	0.17
Graduates	6-10	0.15	0.06	0.05	0.14	0.03	0.08	0.17	0.07	0.21	0.13	0.06	0.15
	11-15	0.20	0.11	0.04	0.13	0.03	0.06	0.18	0.07	0.15	0.16	0.06	0.19
	16-20	0.15	0.09	0.02	0.17	0.05	0.05	0.13	0.02	0.07	0.17	0.07	0.14
	21-40	0.16	0.08	0.02	0.18	0.06	0.02	0.14	0.03	0.03	0.13	0.02	0.04
Voc. Cert. or	1-5	0.27	0.06	0.06	0.28	0.10	0.13	0.21	0.10	0.12	0.23	0.09	0.16
Below Bachelor Degr.	6-10	0.30	0.05	0.06	0.17	0.03	0.04	0.22	0.07	0.11	0.20	0.07	0.12
	11-15	0.42	0.08	0.04	0.20	0.02	0.03	0.21	0.06	0.07	0.23	0.06	0.11
	16-20	0.28	0.05	0.01	0.28	0.03	0.04	0.15	0.02	0.03	0.22	0.05	0.07
	21-40	0.36	0.07	0.02	0.33	0.03	0.01	0.22	0.02	0.02	0.19	0.02	0.02
Bachelor Degr. or	1-5	0.27	0.04	0.08	0.31	0.08	0.19	0.40	0.24	0.42	0.53	0.31	0.60
Above	6-10	0.37	0.03	0.11	0.24	0.05	0.08	0.33	0.12	0.21	0.41	0.21	0.39
	11-15	0.43	0.04	0.07	0.27	0.02	0.07	0.29	0.07	0.14	0.33	0.10	0.18
	16-20	0.28	0.02	0.03	0.39	0.02	0.08	0.22	0.03	0.06	0.30	0.06	0.12
	21-40	0.24	0.02	0.01	0.32	0.02	0.03	0.30	0.02	0.05	0.26	0.02	0.05

Source: The Canadian Censuses between 1981 and 2006.

Notes: The employment sample includes male immigrants residing in all Canadian provinces outside of Quebec, reporting at least one week of work in the reference year and no school attendance, aged 21 to 64. The labour supply within skill cells is the sum of the sampling weights of individuals in that skill cell. The columns H, M and L refer to wage gaps between high language-skilled, medium language-skilled and low language-skilled immigrants and natives, respectively. The language skill groups are based on the LD referring to distance of home language to English. High language-skilled immigrants report English as home language. Medium language-skilled immigrants have a language whose LD is lower than that of Hungarian. The rest of immigrants are grouped into the low language skill group.

Table 2.6: The difference between the mean log wage of immigrants and natives in Quebec

		1981			1991			2001			2006		
Education	**Experience**	**Language Groups**			**Language Groups**			**Language Groups**			**Language Groups**		
		H	**M**	**L**	**H**	**M**	**L**	**H**	**M**	**L**	**H**	**M**	**L**
High Sch. Dropouts or	1-5	-0.091	-0.064	-0.159	-0.257	-0.173	-0.262	-0.045	-0.208	-0.110	0.074	-0.067	-0.124
Graduates	6-10	-0.094	-0.089	-0.249	-0.098	-0.218	-0.349	-0.088	-0.192	-0.315	-0.022	-0.136	-0.189
	11-15	-0.004	-0.078	-0.277	-0.088	-0.140	-0.323	-0.203	-0.196	-0.346	-0.125	-0.177	-0.384
	16-20	-0.055	-0.158	-0.312	-0.100	-0.183	-0.604	-0.260	-0.121	-0.372	-0.155	-0.236	-0.364
	21-40	-0.026	-0.118	-0.308	-0.055	-0.145	-0.463	-0.095	-0.172	-0.497	-0.175	-0.135	-0.540
Voc. Cert. or	1-5	-0.048	-0.019	-0.151	-0.083	-0.190	-0.290	-0.029	-0.151	-0.237	-0.064	-0.080	-0.215
Below Bachelor Degr.	6-10	-0.075	-0.033	-0.174	-0.028	-0.109	-0.381	-0.062	-0.161	-0.234	-0.111	-0.154	-0.231
	11-15	-0.037	-0.014	-0.107	-0.106	-0.083	-0.288	-0.212	-0.199	-0.310	-0.275	-0.207	-0.381
	16-20	-0.020	-0.026	-0.134	-0.136	-0.060	-0.358	-0.124	-0.198	-0.377	-0.143	-0.200	-0.469
	21-40	0.005	0.011	-0.128	-0.052	-0.015	-0.254	-0.076	0.006	-0.194	-0.139	-0.130	-0.281
Bachelor Degr. or	1-5	0.057	0.075	-0.073	-0.038	-0.052	-0.262	-0.002	-0.049	-0.181	-0.059	-0.087	-0.271
Above	6-10	-0.022	-0.019	-0.118	-0.016	-0.007	-0.341	-0.095	-0.074	-0.372	-0.094	-0.196	-0.303
	11-15	-0.007	-0.030	-0.033	0.039	-0.043	-0.291	-0.128	-0.171	-0.406	-0.085	-0.104	-0.355
	16-20	-0.001	0.022	-0.158	-0.044	-0.089	-0.154	-0.020	0.021	-0.207	-0.176	-0.068	-0.356
	21-40	-0.044	0.108	-0.042	-0.052	0.091	-0.173	0.021	0.029	-0.129	-0.078	0.000	-0.194

Source: The Canadian Censuses between 1981 and 2006.

Notes: The wage sample includes the full-time/full-year male workers residing in Quebec, with reported positive wages, aged 21 to 64. The wage gap is the difference between the average of the logarithm of real weekly wages of immigrants and natives. The columns H, M and L refer to wage gaps between high language-skilled, medium language-skilled and low language-skilled immigrants and natives, respectively. The language skill groups are based on the LD referring to distance of home language to French. High language-skilled immigrants report French as home language. Medium language-skilled immigrants have a language whose LD is lower than that of Hebrew. The rest of immigrants are grouped into the low language skill group. A negative value indicates that immigrants have lower wages than natives.

Table 2.7: The ratio of immigrant to native labour supply in Quebec

		1981			1991			2001			2006		
Education	**Experience**	**Language Groups**			**Language Groups**			**Language Groups**			**Language Groups**		
		H	**M**	**L**	**H**	**M**	**L**	**H**	**M**	**L**	**H**	**M**	**L**
High Sch. Dropouts or	1-5	0.04	0.06	0.05	0.08	0.14	0.10	0.08	0.10	0.09	0.09	0.10	0.07
Graduates	6-10	0.03	0.06	0.02	0.04	0.06	0.04	0.06	0.08	0.08	0.06	0.07	0.05
	11-15	0.03	0.09	0.03	0.03	0.05	0.03	0.05	0.07	0.05	0.07	0.09	0.08
	16-20	0.03	0.08	0.03	0.03	0.06	0.02	0.03	0.04	0.03	0.06	0.08	0.07
	21-40	0.02	0.06	0.01	0.03	0.07	0.02	0.02	0.04	0.01	0.02	0.04	0.02
Voc. Cert. or	1-5	0.05	0.05	0.02	0.06	0.06	0.03	0.08	0.05	0.03	0.10	0.07	0.04
Below Bachelor Degr.	6-10	0.05	0.06	0.01	0.03	0.04	0.01	0.07	0.06	0.03	0.07	0.05	0.02
	11-15	0.07	0.08	0.01	0.04	0.04	0.01	0.05	0.05	0.02	0.08	0.06	0.03
	16-20	0.05	0.07	0.01	0.05	0.05	0.01	0.03	0.03	0.01	0.05	0.05	0.02
	21-40	0.04	0.07	0.02	0.04	0.06	0.01	0.04	0.04	0.01	0.03	0.03	0.01
Bachelor Degr. or	1-5	0.07	0.11	0.04	0.11	0.12	0.08	0.21	0.15	0.11	0.36	0.32	0.21
Above	6-10	0.08	0.14	0.04	0.07	0.08	0.04	0.12	0.11	0.07	0.20	0.15	0.11
	11-15	0.10	0.14	0.03	0.08	0.08	0.04	0.09	0.08	0.04	0.13	0.11	0.06
	16-20	0.06	0.11	0.02	0.09	0.12	0.03	0.06	0.06	0.02	0.09	0.07	0.04
	21-40	0.04	0.09	0.01	0.07	0.09	0.01	0.07	0.08	0.02	0.06	0.07	0.02

Source: The Canadian Censuses between 1981 and 2006.

Notes: The employment sample includes male immigrants residing in Quebec, reporting at least one week of work in the reference year and no school attendance, aged 21 to 64. The labour supply within respective skill cell is the sum of the sampling weights of individuals in that skill cell. The columns H, M and L refer to wage gaps between high language-skilled, medium language-skilled and low language-skilled immigrants and natives, respectively. The language skill groups are based on the LD referring to distance of home language from French. High language-skilled immigrants report French as home language. Medium language-skilled immigrants have a language whose LD is lower than that of Hebrew. The rest of immigrants are grouped into the low language skill group.

Table 2.8: The estimates of the inverse elasticity of substitution between immigrant language groups and natives

	Immigrant Language Groups vs. Natives					
	(I)	**(II)**	**(III)**	**(IV)**	**(V)**	**(VI)**
Home Language						
Canada-English	**0.085*** **	-0.013	-0.013	-0.014	**-0.022***	-0.009
	(0.012)	(0.011)	(0.008)	(0.010)	(0.011)	(0.018)
Canada out. of Quebec-English	**0.071*** **	-0.015	**-0.016***	**-0.017***	**-0.030*** **	-0.015
	(0.01	(0.012)	(0.008)	(0.009)	(0.008)	(0.017)
Quebec-French	**0.078*** **	-0.006	0.000	0.001	-0.022	-0.030
	(0.025)	(0.016)	(0.014)	(0.012)	(0.021)	(0.025)
Mother Tongue						
Canada-English	-0.012	**-0.028*** **	**-0.030*** **	**-0.033*** **	**-0.052*** **	**-0.046*** **
	(0.026)	(0.009)	(0.008)	(0.009)	(0.007)	(0.013)
Canada out. of Quebec-English	-0.001	**-0.030** **	**-0.031*** **	**-0.034*** **	**-0.056*** **	**-0.050*** **
	(0.025)	(0.011)	(0.010)	(0.010)	(0.006)	(0.012)
Quebec-French	0.012	-0.017	-0.014	-0.013	**-0.071*** **	-0.034
	(0.021)	(0.013)	(0.010)	(0.010)	(0.023)	(0.020)
Skill Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	No	Yes	Yes
LanguageXTime Fixed Effects	No	No	No	No	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives. The independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell. Language skills are defined respect to the distance of reported languages to English in Canada and Canada outside of Quebec, and to French in Quebec. There are 45 skill cells in a given year. Hence, there are 270 observations in the analysis. Each column controls different effects. Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 2.9: The estimates of the inverse elasticity of substitution between immigrants and natives for separate language groups

	Time -Invariant Skill Effects			**Time-Variant Skill Effects**		
	High Lang. Skilled Imm.	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**	**High Lang. Skilled Imm.**	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**
Home Language						
Canada-English	0.008	0.017	**-0.050*** **	**-0.088*** **	**0.121*** **	**-0.034***
	(0.011)	(0.019)	(0.013)	(0.021)	(0.032)	(0.013)
Canada out. of Quebec-English	**0.034*** **	0.007	**-0.064*** **	**-0.069** **	0.083	**-0.040***
	(0.011)	(0.012)	(0.015)	(0.025)	(0.054)	(0.018)
Quebec-French	-0.004	0.006	**-0.040***	-0.074	-0.032	-0.005
	(0.021)	(0.025)	(0.020)	(0.071)	(0.034)	(0.049)
Mother Tongue						
Canada-English	0.009	-0.004	**-0.061*** **	**-0.082*** **	**0.066** **	**-0.062*** **
	(0.013)	(0.012)	(0.013)	(0.031)	(0.028)	(0.015)
Canada out. of Quebec-English	**0.038** **	-0.008	**-0.072*** **	-0.065	**0.079***	**-0.062*** **
	(0.013)	(0.012)	(0.015)	(0.037)	(0.040)	(0.010)
Quebec-French	-0.022	0.024	-0.030	-0.088	-0.069	0.039
	(0.017)	(0.028)	(0.028)	(0.070)	(0.039)	(0.047)
Skill Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	Yes	Yes	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives. The independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell. Language skills are defined respect to the distance of reported languages to English in Canada and in Canada outside of Quebec and to French in Quebec. There are 15 skill cells in a given year. Hence, there are 90 observations in the analysis. The first three columns assumes that skill effects are time invariant. The last three columns assume that time effects vary over time. Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 2.10: The estimates of the inverse elasticity of substitution between immigrants and natives: The logarithm of average real weekly wages - Ottaviano and Peri (2008, 2012)

	Immigrant Language Groups vs. Natives					
	(I)	**(II)**	**(III)**	**(IV)**	**(V)**	**(VI)**
Home Language						
Canada-English	**0.088*** **	**-0.022** **	**-0.022** **	**-0.024** **	**-0.035*** **	-0.011
	(0.012)	(0.008)	(0.008)	(0.009)	(0.010)	(0.014)
Canada out. of Quebec-English	**0.072*** **	**-0.025*** **	**-0.024*** **	**-0.026*** **	**-0.040*** **	**-0.020***
	(0.010)	(0.009)	(0.008)	(0.008)	(0.007)	(0.012)
Quebec-French	**0.082*** **	**-0.026** **	-0.026	**-0.032** **	**-0.045** **	-0.034
	(0.023)	(0.009)	(0.016)	(0.014)	(0.021)	(0.024)
Mother Tongue						
Canada-English	-0.003	**-0.036*** **	**-0.036*** **	**-0.040*** **	**-0.060*** **	**-0.046*** **
	(0.026)	(0.008)	(0.009)	(0.009)	(0.007)	(0.013)
Canada out. of Quebec-English	0.007	**-0.038*** **	**-0.037*** **	**-0.040*** **	**-0.062*** **	**-0.048*** **
	(0.024)	(0.010)	(0.011)	(0.010)	(0.007)	(0.012)
Quebec-French	0.018	**-0.028*** **	**-0.028** **	**-0.034** **	**-0.064** **	-0.034
	(0.018)	(0.009)	(0.011)	(0.010)	(0.019)	(0.021)
Skill Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	No	Yes	Yes
LanguageXTime Fixed Effects	No	No	No	No	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the logarithm of average real weekly wages of immigrants and natives. The independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell. Language skills are defined respect to the distance of reported languages to English and French. There are 45 skill cells in a given year. Hence, there are 270 observations in the analysis. Each column controls different effects. Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 2.11: The estimates of the inverse elasticity of substitution between immigrants and natives for separate language groups: The logarithm of average real weekly wages - Ottaviano and Peri (2008, 2012)

	Time -Invariant Skill Effects			**Time-Variant Skill Effects**		
	High Lang. Skilled Imm.	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**	**High Lang. Skilled Imm.**	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**
Home Language						
Canada-English	-0.019	-0.023	**-0.051*** **	**-0.082*** **	**0.086***	**-0.036***
	(0.015)	(0.014)	(0.015)	(0.025)	(0.040)	(0.018)
Canada out. of Quebec-English	0.004	**-0.027***	**-0.063*** **	**-0.069** **	**0.096***	**-0.035***
	(0.014)	(0.014)	(0.017)	(0.026)	(0.049)	(0.019)
Quebec-French	-0.021	-0.014	**-0.050** **	-0.058	**-0.074***	0.060
	(0.012)	(0.020)	(0.023)	(0.061)	(0.037)	(0.041)
Mother Tongue						
Canada-English	-0.005	-0.006	**-0.044*** **	**-0.061*** **	**0.125*** **	-0.016
	(0.012)	(0.013)	(0.014)	(0.015)	(0.027)	(0.022)
Canada out. of Quebec-English	0.014	-0.017	**-0.057*** **	**-0.049** **	0.066	-0.018
	(0.008)	(0.010)	(0.016)	(0.019)	(0.043)	(0.021)
Quebec-French	-0.005	-0.022	**-0.042** **	-0.064	-0.049	0.029
	(0.014)	(0.014)	(0.015)	(0.061)	(0.032)	(0.044)
Skill Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	Yes	Yes	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the log of average real weekly wages of immigrants and natives. The independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell.Language skills are defined respect to the distance of reported languages to English and French. There are 15 skill cells in a given year. Hence, there are 90 observations in the analysis. The first three columns assumes that skill effects are time invariant. The last three columns assume that time effects vary over time. Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 2.12: The estimates of the inverse elasticity of substitution for experience groups

	[I]		[II]	
	(I.I)	**(I.II)**	**(II.I)**	**(II.II)**
Home Language				
Canada-English	-0.065	**-0.203***	-0.063	**-0.194****
	(0.075)	(0.106)	(0.074)	(0.099)
Canada out. of Quebec-English	-0.105	**-0.229****	-0.102	**-0.219*****
	(0.076)	(0.089)	(0.073)	(0.035)
Quebec-French	0.446	**-0.816***	0.407	**-0.765***
	(0.858)	(0.468)	(0.716)	(0.401)
Mother Tongue				
Canada-English	-0.064	**-0.203***	-0.063	**-0.194****
	(0.075)	(0.105)	(0.074)	(0.099)
Canada out. of Quebec-English	-0.105	**-0.229****	-0.102	**-0.220*****
	(0.076)	(0.089)	(0.074)	(0.084)
Quebec-French	0.261	**-0.748***	0.250	**-0.725****
	(0.324)	(0.365)	(0.300)	(0.343)
Skill Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
EducationXTime Fixed Effects	Yes	Yes	Yes	Yes
ExperienceXLinear Trend	No	Yes	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_X)$. The dependent variable is the average logarithm of real weekly wages of immigrants and natives. The independent variable is the logarithm of CES-aggregated labour supply instrumented with the labour supply of immigrants within the corresponding skill cell. For CES-aggregated labour supplied, the estimates of the relative productivity parameters from the 4th level estimations used. These estimates consider language skills defined according to the distance of reported languages to English. The columns [I] use the estimates in the column (II) in Table 2.8. The columns [II] use the estimates from the column (V) in Table 2.8. Each column above controls different effects. There are 15 education-experience skill cells in a given year. Hence, there are 60 observations in the analysis. Regressions are weighted with appropriate sampling weights and standard errors are clustered within education-experience cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 2.13: The estimates of the inverse elasticity of substitution for education groups

	[I]		[II]	
	(I.I)	**(I.II)**	**(II.I)**	**(II.II)**
Home Language				
Canada-English	**-0.462****	**-0.384***	**-0.434****	**-0.375***
	(0.204)	(0.217)	(0.199)	(0.213)
Canada out. of Quebec-English	-0.212	-0.22	-0.201	-0.216
	(0.151)	(0.153)	(0.150)	(0.153)
Quebec-French	**-0.200*****	**-0.282*****	**-0.200*****	**-0.275*****
	(0.066)	(0.089)	(0.064)	(0.083)
Mother Tongue				
Canada-English	**-0.452****	**-0.382***	**-0.438****	**-0.381***
	(0.204)	(0.217)	(0.214)	(0.225)
Canada out. of Quebec-English	-0.208	-0.218	-0.201	-0.216
	(0.151)	(0.152)	(0.155)	(0.158)
Quebec-French	**-0.200*****	**-0.282*****	**-0.200*****	**-0.278*****
	(0.066)	(0.099)	(0.066)	(0.086)
Skill Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
EducationXtime Fixed Effects	Yes	Yes	Yes	Yes
EducationXTime Splines	No	Yes	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_E)$. The dependent variable is the average logarithm of real weekly wages of immigrants and natives. The independent variable is the logarithm of CES-aggregated labour supply instrumented with the labour supply of immigrants within the corresponding skill cell. For CES-aggregated labour supplied, the estimates of the relative productivity parameters from the 4-th and 3-rd level estimations used. These estimates consider language skills defined according to the distance of reported languages to English. The columns [I] uses the estimates in the column (II) in Table 2.8 and the estimates in the column (I.II) from Table 2.12. The column [II] uses the estimates from the column (V) in Table 2.8 and the estimates in the column (II.II) in Table 2.12. Each column above controls different effects. There are 3 education skill cells in a given year. Hence, there are 18 observations in the analysis. Regressions are weighted with appropriate sampling weights and standard errors are clustered within education-experience cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 2.14: Actual wage changes in Canada outside of Quebec: 1981-2006

Home Language-Based Language Groups of Immigrants	**All Education**	**High Sch. Dropuouts/ High Sch. Grads**	**Trade Cert./ Below Bachelor**	**Bachelor/ Above Bachelor**
Low Language-Skilled Imm.	-9.4%	-15.9%	-17.1%	-12.6%
Medium Language-Skilled Imm.	-1.8%	-14.4%	-14.0%	-17.7%
High Language-Skilled Imm.	-0.6%	-10.1%	-5.7%	-3.3%

Mother Tongue-Based Language Groups of Immigrants	**All Education**	**High Sch. Dropuouts/ High Sch. Grads**	**Trade Cert./ Below Bachelor**	**Bachelor/ Above Bachelor**
Low Language-Skilled Imm.	-8.8%	-17.2%	-15.3%	-11.7%
Medium Language-Skilled Imm.	-0.5%	-8.6%	-9.4%	-12.2%
High Language-Skilled Imm.	2.4%	-8.3%	-2.4%	2.1%

	All Education	**High Sch. Dropuouts/ High Sch. Grads**	**Trade Cert./ Below Bachelor**	**Bachelor/ Above Bachelor**
All Immigrants	-0.4%	-3.4%	-2.3%	-2.5%
Natives	5.2%	-5.8%	2.8%	8.2%
All Workers	0.2%	-1.6%	-0.1%	0.1%

Source: The Canadian Censuses between 1981 and 2006.

Notes: Percentage changes are calculated using the difference between the average log real weekly wages in 1981 and in 2006 for corresponding labour types.

Table 2.15: Simulated total wage effects in Canada outside of Quebec -I-:
Imperfect substitutability between immigrants and natives - 1981-2006

Short-Run Total Wage Effects

	All Education	High Sch. Dropuouts/ High Sch. Grads	Trade Cert./ Below Bachelor	Bachelor/ Above Bachelor
Low Language-Skilled Imm.	-15.1%	-1.8%	-11.3%	-40.0%
Medium Language-Skilled Imm.	-1.8%	4.1%	-3.6%	-40.3%
High Language-Skilled Imm.	-7.1%	3.0%	-3.8%	-28.3%
All Immigrants	-6.9%	2.8%	-4.3%	-30.6%
Natives	-4.2%	2.5%	-3.3%	-24.9%
All Workers	-4.9%	2.5%	-3.6%	-26.5%

Long-Run Total Wage Effects

	All Education	High Sch. Dropuouts/ High Sch. Grads	Trade Cert./ Below Bachelor	Bachelor/ Above Bachelor
Low Language-Skilled Imm.	-10.2%	3.0%	-6.4%	-35.2%
Medium Language-Skilled Imm.	2.9%	8.9%	1.2%	-35.4%
High Language-Skilled Imm.	-2.2%	7.8%	0.9%	-23.5%
All Immigrants	-2.1%	7.7%	0.5%	-25.7%
Natives	0.6%	7.3%	1.5%	-20.0%
All Workers	0.0%	7.4%	1.2%	-21.7%

Source: The Canadian Censuses between 1981 and 2006.
Notes: The simulations assume that $1/\sigma_{EXG}$ is 0.030, $1/\sigma_{EX}$ is 0.200 and $1/\sigma_{E}$ is 0.450. Additionally, the share of labour and capital is assumed to be 0.66 and 0.33, respectively. The elasticity of substitution between capital and labour is 1.

Table 2.16: Simulated total wage effects in Canada outside of Quebec -II-:
Imperfect substitutability between immigrants and natives - 1981-2006

Short-Run Total Wage Effects

	All Education	High Sch. Dropuouts/ High Sch. Grads	Trade Cert./ Below Bachelor	Bachelor/ Above Bachelor
Low Language-Skilled Imm.	-23.1%	-5.2%	-18.1%	-50.4%
Medium Language-Skilled Imm.	-3.4%	4.9%	-3.3%	-37.1%
High Language-Skilled Imm.	-7.3%	3.7%	-4.1%	-26.9%
All Immigrants	-8.1%	3.0%	-5.3%	-34.4%
Natives	-3.9%	2.4%	-3.2%	-22.9%
All Workers	-4.9%	2.5%	-3.8%	-26.2%

Long-Run Total Wage Effects

	All Education	High Sch. Dropuouts/ High Sch. Grads	Trade Cert./ Below Bachelor	Bachelor/ Above Bachelor
Low Language-Skilled Imm.	-18.2%	-0.3%	-13.3%	-45.5%
Medium Language-Skilled Imm.	1.4%	9.7%	1.5%	-32.2%
High Language-Skilled Imm.	-2.4%	8.5%	0.7%	-22.0%
All Immigrants	-3.2%	7.8%	-0.5%	-29.5%
Natives	0.1%	7.3%	1.6%	-18.0%
All Workers	0.0%	7.4%	0.9%	-21.3%

Source: The Canadian Censuses between 1981 and 2006.

Notes: For simulations, it is assumed that $1/\sigma_{EXG}$ is 0.057, $1/\sigma_{EX}$ is 0.200 and $1/\sigma_{E}$ is 0.450. Additionally, the share of labour and capital is assumed to be 0.66 and 0.33, respectively. The elasticity of substitution between capital and labour is 1.

Table 2.17: Actual wage changes in Quebec: 1981-2006

Home Language-Based Language Groups of Immigrants	**All Education**	**High Sch. Dropuouts/ High Sch. Grads**	**Trade Cert./ Below Bachelor**	**Bachelor/ Above Bachelor**
Low Language-Skilled Imm.	-18.9%	-25.1%	-29.6%	-30.0%
Medium Language-Skilled Imm.	-12.8%	-20.3%	-23.7%	-23.9%
High Language-Skilled Imm.	-15.2%	-21.5%	-22.8%	-23.7%

Mother Tongue-Based Language Groups of Immigrants	**All Education**	**High Sch. Dropuouts/ High Sch. Grads**	**Trade Cert./ Below Bachelor**	**Bachelor/ Above Bachelor**
Low Language-Skilled Imm.	-21.1%	-27.9%	-32.6%	-29.0%
Medium Language-Skilled Imm.	-13.5%	-19.9%	-23.1%	-24.7%
High Language-Skilled Imm.	-8.7%	-14.2%	-18.0%	-20.7%

	All Education	**High Sch. Dropuouts/ High Sch. Grads**	**Trade Cert./ Below Bachelor**	**Bachelor/ Above Bachelor**
All Immigrants	-1.1%	-5.1%	-5.5%	-5.8%
Natives	-3.3%	-11.8%	-8.5%	-7.8%
All Workers	-0.3%	-2.8%	-2.1%	-2.4%

Source: The Canadian Censuses between 1981 and 2006.

Notes: Percentage changes are calculated using the difference between the average log real weekly wages in 1981 and in 2006 for corresponding labour types.

Table 2.18: Simulated total wage effects in Quebec -I-:
Perfect substitutability between immigrants and natives - 1981-2006

Short-Run Total Wage Effects

	All Education	High Sch. Dropuouts/ High Sch. Grads	Trade Cert./ Below Bachelor	Bachelor/ Above Bachelor
Low Language-Skilled Imm.	-3.1%	1.8%	3.0%	-12.3%
Medium Language-Skilled Imm.	-3.5%	1.8%	3.0%	-12.2%
High Language-Skilled Imm.	-4.1%	1.8%	-2.9%	-11.7%
All Immigrants	-3.6%	1.8%	-2.9%	-12.1%
Natives	-2.1%	1.8%	-3.0%	-12.4%
All Workers	-2.3%	1.8%	-3.0%	-12.3%

Long-Run Total Wage Effects

	All Education	High Sch. Dropuouts/ High Sch. Grads	Trade Cert./ Below Bachelor	Bachelor/ Above Bachelor
Low Language-Skilled Imm.	-0.9%	4.1%	-0.7%	-10.0%
Medium Language-Skilled Imm.	-1.2%	4.1%	-0.7%	-9.9%
High Language-Skilled Imm.	-1.9%	4.1%	-0.7%	-9.5%
All Immigrants	-1.4%	4.1%	-0.7%	-9.8%
Natives	0.1%	4.1%	-0.7%	-10.1%
All Workers	-0.0%	4.1%	-0.7%	-10.1%

Source: The Canadian Censuses between 1981 and 2006.
Notes: The simulations in the first and second panels assume that $1/\sigma_{EXG}$ is 0, $1/\sigma_{EX}$ is 0.200 and $1/\sigma_{E}$ is 0.450. Additionally, the share of labour and capital is assumed to be 0.66 and 0.33, respectively. The elasticity of substitution between capital and labour is 1.

Table 2.19: Simulated total wage effects in Quebec -II-:
Imperfect substitutability between immigrants and natives - 1981-2006

Short-Run Total Wage Effects

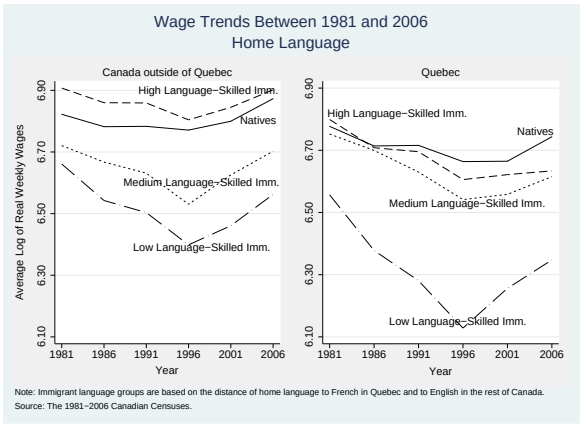
	All Education	High Sch. Dropuouts/ High Sch. Grads	Trade Cert./ Below Bachelor	Bachelor/ Above Bachelor
Low Language-Skilled Imm.	-12.8%	0.3%	-10.3%	-29.6%
Medium Language-Skilled Imm.	-4.6%	4.4%	-7.0%	-18.3%
High Language-Skilled Imm.	-8.4%	3.7%	-6.8%	-21.4%
All Immigrants	-7.2%	3.5%	-7.7%	-21.9%
Natives	-1.6%	1.6%	-2.4%	-10.4%
All Workers	-2.3%	1.8%	-3.0%	-12.5%

Long-Run Total Wage Effects

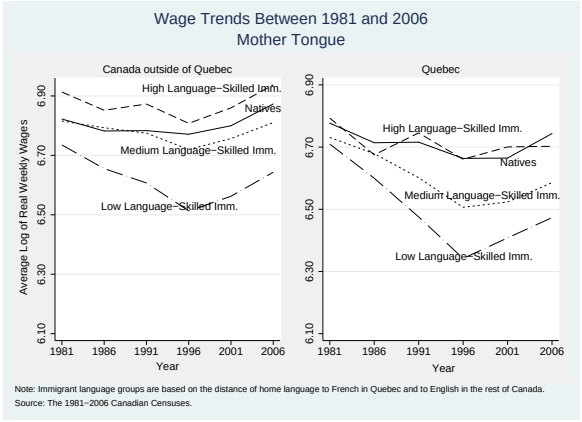
	All Education	High Sch. Dropuouts/ High Sch. Grads	Trade Cert./ Below Bachelor	Bachelor/ Above Bachelor
Low Language-Skilled Imm.	-10.5%	2.6%	-8.0%	-27.3%
Medium Language-Skilled Imm.	-2.3%	6.7%	-4.7%	-16.0%
High Language-Skilled Imm.	-6.1%	6.0%	-4.4%	-19.1%
All Immigrants	-4.9%	5.8%	-5.4%	-19.6%
Natives	0.6%	3.9%	-0.1%	-8.1%
All Workers	0.0%	4.1%	-0.7%	-10.2%

Source: The Canadian Censuses between 1981 and 2006.
Notes: The simulations in the first and second panels assume that $1/\sigma_{EXG}$ is 0.073, $1/\sigma_{EX}$ is 0.200 and $1/\sigma_{E}$ is 0.450. Additionally, the share of labour and capital is assumed to be 0.66 and 0.33, respectively. The elasticity of substitution between capital and labour is 1.

Figure 2.1: Trends in native and immigrant wages: 1981-2006



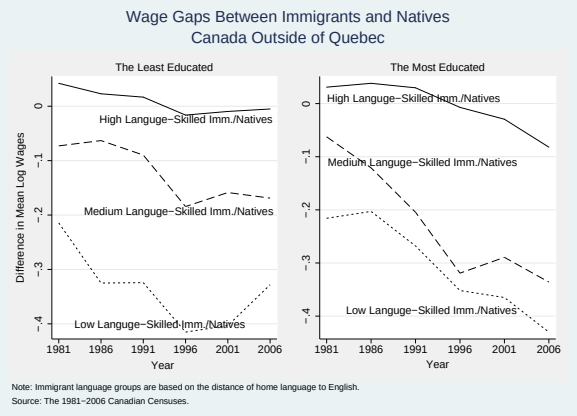
(a)



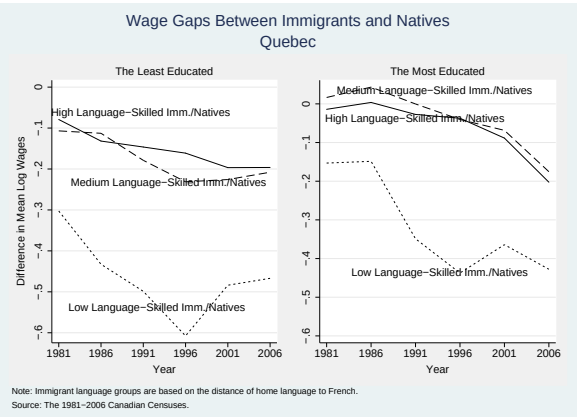
(b)

Source: The Canadian Censuses between 1981 and 2006.
Notes: The sample includes full-time/full-year male immigrants, reporting positive wages, aged 21 to 64. Figure 2.1a and 2.1b present the average log of real weekly wages of these workers residing in Canada outside of Quebec and in Quebec. The average log of real weekly wages are calculated by taking the mean of log of real weekly wages for natives and immigrant groups by using sampling weights. Figure 2.1a and Figure 2.1b categorize immigrants into language skill groups based on home language and mother tongue, respectively. In both figures, high language-skilled immigrants report English if they reside in Canada outside of Quebec, and report French if they reside in Quebec. Cut-off languages to divide immigrants into medium and low skill groups are Hebrew and Hungarian if immigrants reside in Quebec and in the rest of Canada.

Figure 2.2: Immigrant-Native wage gaps for selected education groups



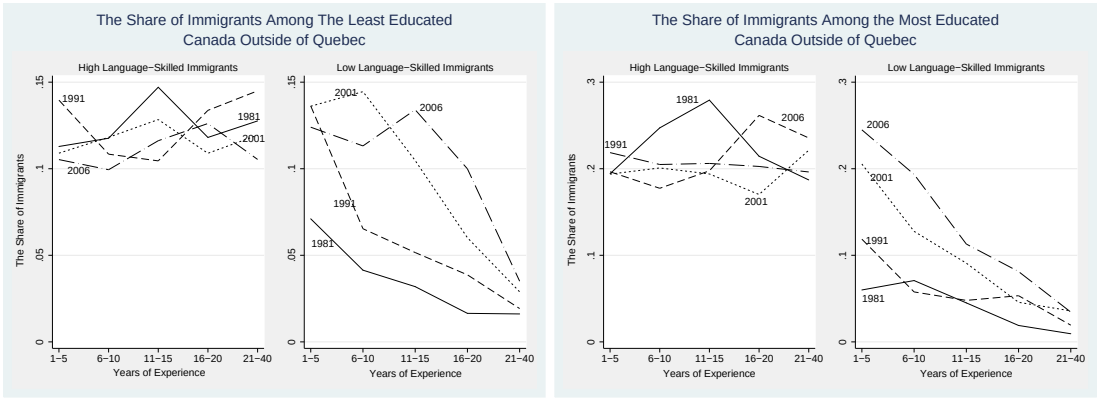
(a)



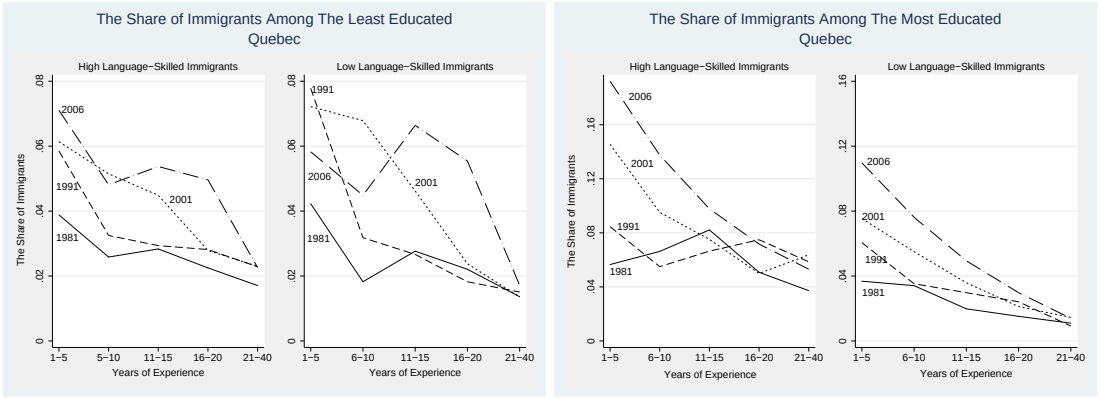
(b)

Source: The Canadian Censuses between 1981 and 2006.
Notes: The sample includes full-time/full-year male immigrants, reporting positive wages, aged 21 to 64. The difference in the average log of real weekly wages of immigrant language groups and natives are presented. The least educated workers have at most a high school diploma while the most educated workers are with a Bachelor degree or more. High language-skilled immigrants report English in Figure 2.2a, and French in Figure 2.2b. Cut-off languages to divide immigrants into medium and low skill groups are Hungarian in Figure 2.2a, and Hebrew in Figure 2.2b.

Figure 2.3: The share of immigrants in selected years



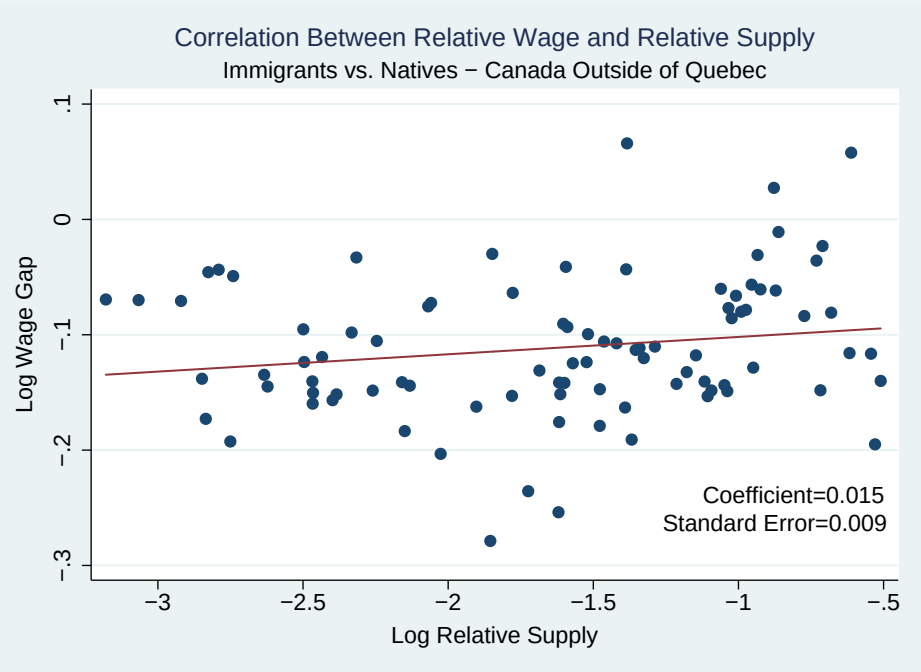
(a)



(b)

Source: The Canadian Censuses between 1981 and 2006.
Notes: The sample includes male immigrants, reporting at least one week of work and no school attendance, aged 21 to 64. The least educated immigrants are with at most a high school degree. The most educated are the ones with at least a Bachelor degree. In Figure 2.3a, language groups are based on the distance of home language to English. High language-skilled immigrants report English and the cut-off language for medium and low language-skilled immigrants is Hungarian. In Figure 2.3b, language groups are based on the distance of home language to French. High language-skilled immigrants report French and the cut-off language for medium and low language-skilled immigrants is Hebrew.

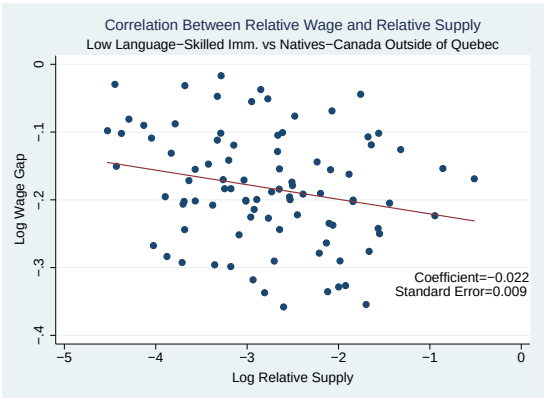
Figure 2.4: The relation between relative wage and relative supply: Education-experience groups in Canada outside of Quebec: 1981-2006



Source: The Canadian Censuses between 1981 and 2006.
Notes: The employment sample includes male immigrants, reporting at least one week of work and no school attendance, aged 21 to 64. The wage sample includes full-time/full-year male immigrants, reporting positive wages, aged 21 to 64. Each observation refers to an education-experience group. Immigrants are considered as a single group, and categorized based on their education and labour market experience.

Figure 2.5: The relation between relative wage and relative supply: Education-experience-language groups in Canada outside of Quebec: 1981-2006

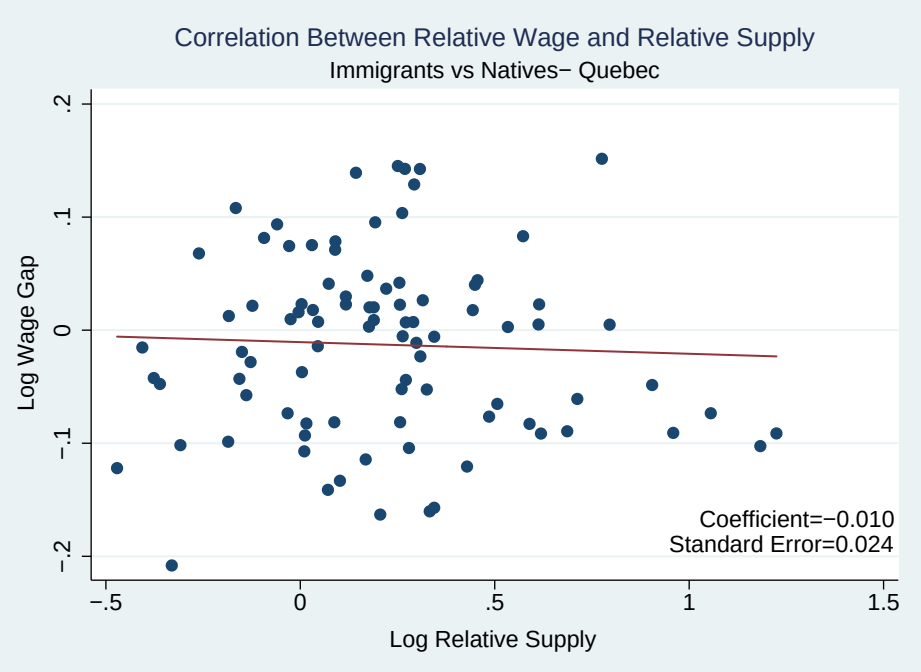
(a) **(b)**



(c)

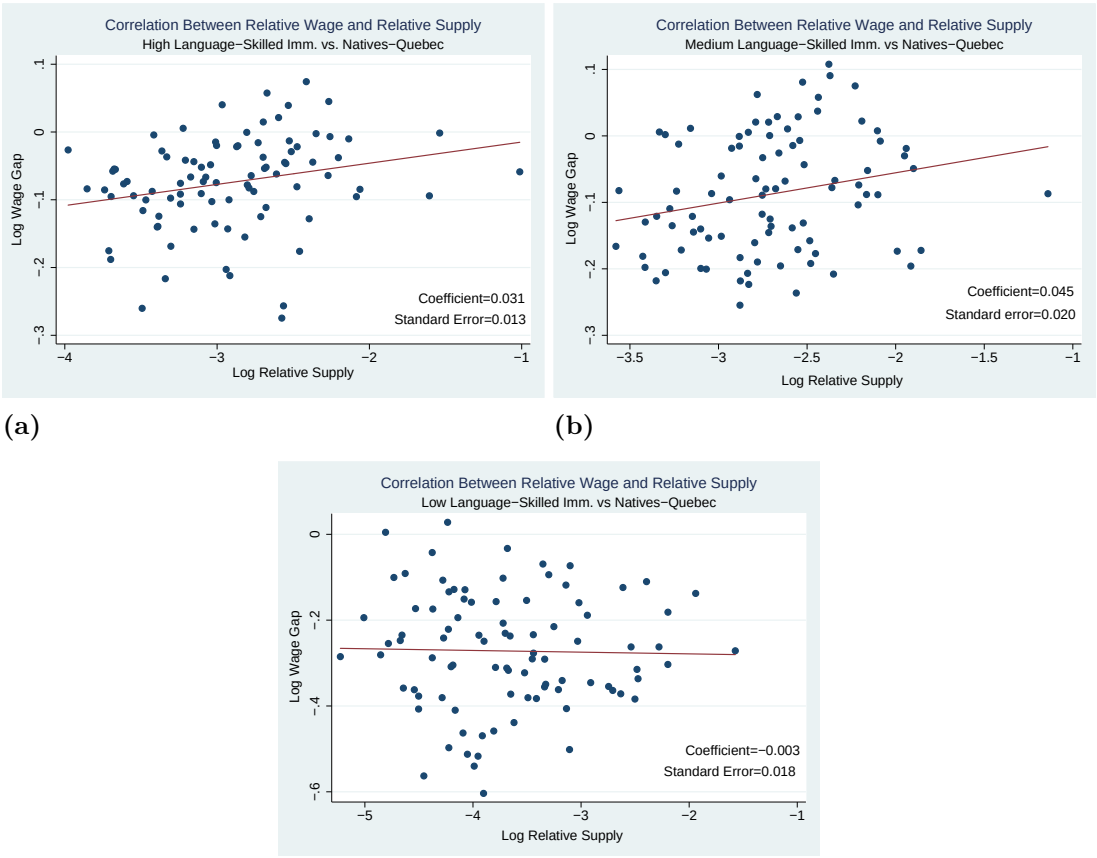
Source: The Canadian Censuses between 1981 and 2006.
Notes: The employment sample includes male immigrants, reporting at least one week of work and no school attendance, aged 21 to 64. The wage sample includes full-time/full-year male immigrants, reporting positive wages, aged 21 to 64. Immigrant groups are defined according to the level of language skills: (a) high language skills, (b) medium language skills, (c) low language skills. Language skill levels are defined with respect to the distance of home language to English.

Figure 2.6: The relation between relative wage and relative supply: Education-experience groups in Quebec: 1981-2006



Source: The Canadian Censuses between 1981 and 2006.
Notes: The employment sample includes male immigrants, reporting at least one week of work and no school attendance, aged 21 to 64. The wage sample includes full-time/full-year male immigrants, reporting positive wages, aged 21 to 64. Each observation refers to an education-experience group. Immigrants are considered as a single group, and categorized based on their education and labour market experience.

Figure 2.7: The relation between relative wage and relative supply: Education-experience-language groups in Quebec: 1981-2006



Source: The Canadian Censuses between 1981 and 2006.
Notes: The employment sample includes male immigrants, reporting at least one week of work and no school attendance, aged 21 to 64. The wage sample includes full-time/full-year male immigrants, reporting positive wages, aged 21 to 64. Immigrant groups are defined according to the level of language skills: (a) high language skills, (b) medium language skills, (c) low language skills. Language skill levels are defined with respect to the distance of home language to French.

Appendix 2.1

Sample construction: The samples are defined separately for the wage and employment sample and the coding is provided according to the names of the variables in the 1981 Census.

The employment sample: The employment sample selection rules follow:

E1: Select the males aged 21-64. (keep if age>=21 and age<=64 and sex==2)

E2: Exclude those living in collective dwellings (doctp! =2)

E3: Select those reporting positive weeks of work in the reference year (weeks>0)

E4: Include non-school attendants (atendr==1 or 2)

E6: Select those with positive years of experience in the labour market

E7: Exclude those ones with no available language information.

In addition to the restrictions above,

The wage sample:

W1: Then, select wage/salary workers (cowd==5)

W2: Select those with positive wages reported (wages>0)

W3: Select those reporting at least 40 weeks of work in the reference year (weeks>=40)

W4: Select those working full-time for most of the year (fptim==1)

The definitions of the variables:

Weekly earnings:

The annual wages are adjusted to 2005 prices by using the Canadian CPI and divided by the number of weeks worked. Then, the simple average of logarithm of real weekly wages is used. The difference in the mean log of real weekly wages is used for the wage gap between immigrants and natives.

Education: The three groups are defined as: (i) less than high school or high school graduates (dgreer==9 or 10 or 5); (ii) vocational training certificate or diploma below B.A. (dgreer==3 or 8 or 11); (ii) B.A. or above (dgreer==1 or 2 or 4 or 6 or 7)

Experience:

Years of experience is defined as (age−years of schooling−6). The variables used to calculate years of schooling: (i) highest grade of elementary or secondary school (hgradr); (ii) years of other non-university (ps_otr); (iii) years of university (ps_uvr). Then, five experience categories are defined: 1-5, 6-10, 11-15, 16-20, 21-40.

Information on years of schooling was not collected in 2006. Hence, the labour market entrance age is assumed to be 17 for high school dropouts, 19 for high school graduates, 20 for post-secondary graduates & (hcdd==3 or 4 or 5), 21 for post-secondary graduates & (hcdd==6), 22 for post-secondary graduates & (hcdd==7 or 8), 23 for those with a bachelor degree, and 25 for those with a degree higher than university level & (hcdd==10 or 11 or 12) and 27 for those with a degree higher than university level & (hcdd==13).

Then, experience is defined as the difference between the age and the entrance age to the labour market.

Appendix 2.2

Sample selection

As another robustness check, I provide the results for both full-time and part-time workers. The employment sample is defined as in Appendix 2.1 The wage sample is defined as below:

The wage sample:

In addition to the restriction imposed on the employment sample,

A1: Select the wage/salary workers (cowd==5 & (fptim==1 or fptim==3)

A2: Select the ones with positive wages reported (wages>0)

A3: We select the ones with years of experience between 1-40

The definitions of the variables:

The definitions of the variables are also the same as in Appendix 2.1.

Appendix Tables A2.1 and A2.2 present the estimates of the elasticity of the substitution between immigrant language groups and natives from the sample described above.

Appendix 2.3

Lists of languages in language skill groups

These lists include the languages reported in the 2006 Census. The order in which I report these languages is by the LD.

1. Language skill groups with respect to English

High language skills: English

Medium language skills: Dutch, Flemish, Norwegian, Swedish, Germanic languages, Dannish, Yiddish, Frisian, German, Icelandic, Romanian, Croatian, Serbian, Serbo-Croatian, Italian, Bosnian, Slovenian, Bulgarian, Romance languages, Byelorussian, Salvic languages, Czech, French, Macedonian, Welsh, Slovak, Spanish, Polish, Russian, Persian, Urdu, Sinhala, Celtic languages, Kurdish, Portuguese, Ukranian.

Low language skills: Hungarian, Amharic, Hindi, Gaelic languages, Bengali, Indo-Iranian languages, Marathi, Konkani, Greek, Latvian, Gujarati, Hebrew, Arabic, Tigrigna, Armenian, Pashto, Malay, Semitic languages, Cantonese, Wolof, Kabyle, Estonian, Ilocano, Korean, Afro-Asiatic languages, Japanese, Azerbaijani, Maltese, Rwanda, Thai, Fukien, Lao, Malayam, Lamayo-Polynesian, Igbo, Swahili, African languages, Panjabi, Kannada, Bantu languages, Shangainese, Tamil, Oromo, Niger-Congo, Tibetan languages, Sino-Tibetan languages, Chinese (other), Lingala, Dravidian, Turkic languages, Turkish, Pampango, Mandarin, Taiwanese, Chaochow, Rundi, Finnish, Shona, Tlingit, Kutenai, Edo, Tagalog, Telugu, Cambodian, Somali, Hakka, Akan, Vietnamese.

2. Language skill groups with respect to French

High language skills: French

Medium language skills: Italian, Portuguese, Romance languages, Spanish, Serbo-Croatian, Croatian, Welsh, Hindi, Macedonian, Serbian, Gujarati, Bosnian, Czech, Frisian, Slovak, Slovenian, English, Flemish, Dutch, Urdu, Slavic languages, Persian, Kurdish, Romanian, Konkani, Bulgarian, Marathi, Russian, Indo-Iranian languages, Polish, Byelorussian, Germanic languages, Danish, Lithuanian.

Low language skills: Hebrew, Icelandic, Bengali, Ukranian, Celtic languages, Yiddish, Swedish, Norwegian, Malayalam, Greek, Panjabi, German, Latvian, Wolof, Armenian, Tigrigna, Oromo, Maltese, Dravidian languages, Tamil, Semitic languages, Pashto, Arabic, Telugu, Pampango, Afro-Asiatic languages, Amharic, Sinhala, Kabyle, Turkic languages, Turkish, Kannada, Cambodian, Finnish, Estonian, Chaochow, Cantonese, Thai, Gaelic languages, Edo, Hakka, Lingala, Fukien, Lao, Chinese (other), Tibetan languages, Sino-Tibetan languages, Somali, Malay, Kutenai, Swahili, Akan, Ilocano, Niger-Congo languages, Hungarian, African languages, Shona, Tlingit, Igbo, Bantu languages, Shangainese, Vietnamese, Mandarin, Japanese, Azerbaijani, Korean, Tagalog, Rwanda, Rundi.

Appendix 2.4

Alternative language definitions

As a robustness check, I provide the results on full-time/full-year workers in which immigrants are categorized into language groups based on two alternative language definitions. First, immigrants are categorized into language skill groups based on the distance of reported languages to English or French *and* their age at migration. If immigrants enter Canada at the age of 12 or younger, then they are considered as high language-skilled immigrants along with English-speaking immigrants in Canada and everywhere in Canada other than Quebec. Other immigrants, who speak language other than English and enter Canada at the age of 13 or older are defined as medium and low language-skilled immigrants based on the distance of home langue or mother tongue to English. The cut-off language is Hungarian. If an immigrant arrives in Canada at the age of 12 or younger and resides in Quebec at the time of the Census, he is grouped into high language skill category along with French-speaking immigrants. Other immigrants that enter Canada at the age of 13 or older and speak a language other than French are categorized into medium and low language skill groups based on the distance of reported languages from French. The cut-off language is Hebrew. Appendix Tables A2.3 and A2.4 present results from this categorization.

Second, a new cut-off language is defined only for English-speaking immigrants in Canada and Canada outside of Quebec. Punjabi, the most distant Indo-European languge to English with the LD 100.58, is used to separate medium and low language-skilled immigrants from each other. The restriction on the enterance age as described above is also imposed upon language groups. The results are presented for Canada and Canada outside of Quebec in Appendix Tables A2.5 and A2.6.

Table A2.1: The estimates of the inverse elasticity of substitution between immigrant language groups and natives: Full- or part-time workers

	Immigrant Language Groups vs. Natives					
	(I)	**(II)**	**(III)**	**(IV)**	**(V)**	**(VI)**
Home Language						
Canada-English	**0.07*** **	-0.012	-0.008	-0.010	-0.021	-0.006
	(0.013)	(0.014)	(0.011)	(0.013)	(0.013)	(0.023)
Canada out. of Quebec-English	**0.053*** **	-0.013	-0.009	-0.011	**-0.029*** **	-0.015
	(0.012)	(0.014)	(0.011)	(0.012)	(0.010)	(0.020)
Quebec-French	**0.068** **	-0.016	-0.000	0.000	-0.027	-0.034
	(0.028)	(0.018)	(0.013)	(0.013)	(0.020)	(0.024)
Mother Tongue						
Canada-English	-0.022	**-0.028** **	**-0.025** **	**-0.030** **	**-0.051*** **	**-0.046*** **
	(0.027)	(0.013)	(0.012)	(0.012)	(0.009)	(0.017)
Canada out. of Quebec-English	-0.016	**-0.027***	**-0.025***	**-0.029** **	**-0.056*** **	**-0.053*** **
	(0.025)	(0.014)	(0.012)	(0.013)	(0.008)	(0.015)
Quebec-French	0.012	-0.017	-0.014	-0.013	**-0.071*** **	-0.034
	(0.021)	(0.013)	(0.010)	(0.010)	(0.025)	(0.020)
Skill Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	No	Yes	Yes
LanguageXTime Fixed Effects	No	No	No	No	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives. The main independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell. Language skills are defined respect to the distance of reported languages to English in Canada and Canada outside of Quebec, and to French in Quebec. There are 45 skill cells in a given year. Hence, there are 270 observations in the analysis. Each column controls for different effects. Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table A2.2: The estimates of the inverse elasticity of substitution between immigrant language groups and natives for separate language groups: Full- or part-time workers

	Time -Invariant Skill Effects			**Time-Variant Skill Effects**		
	High Lang. Skilled Imm.	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**	**High Lang. Skilled Imm.**	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**
Home Language						
Canada-English	0.014	0.021	**-0.054****	**-0.091*****	**0.128*****	**-0.050****
	(0.014)	(0.020)	(0.020)	(0.023)	(0.039)	(0.020)
Canada out. of Quebec-English	**0.043****	0.011	**-0.064*****	**-0.072****	**0.096***	**-0.049***
	(0.014)	(0.013)	(0.021)	(0.032)	(0.049)	(0.024)
Quebec-French	-0.023	0.027	**-0.064***	-0.068	-0.021	-0.049
	(0.013)	(0.033)	(0.034)	(0.054)	(0.043)	(0.067)
Mother Tongue						
Canada-English	0.017	0.001	**-0.065*****	**-0.105*****	**0.087*****	**-0.067*****
	(0.016)	(0.013)	(0.020)	(0.032)	(0.028)	(0.019)
Canada out. of Quebec-English	**0.045****	-0.003	**-0.071*****	**-0.092****	**0.104****	**-0.067****
	(0.017)	(0.013)	(0.022)	(0.042)	(0.037)	(0.023)
Quebec-French	0.007	0.010	**-0.058****	-0.091	0.022	-0.035
	(0.019)	(0.030)	(0.025)	(0.056)	(0.031)	(0.058)
Skill Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	Yes	Yes	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives. The main independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell. Language skills are defined respect to the distance of reported languages to English in Canada and in Canada outside of Quebec and to French in Quebec. There are 15 skill cells in a given year. Hence, there are 90 observations in the analysis. The first three columns assume that skill effects are time invariant. The last three columns assume that time effects vary over time. Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table A2.3: The estimates of the inverse elasticity of substitution between immigrant language groups and natives: Alternative language definition -I

	Immigrant Language Groups vs. Natives					
	(I)	**(II)**	**(III)**	**(IV)**	**(V)**	**(VI)**
Home Language						
Canada-English	**0.076*** **	-0.013	-0.012	-0.013	**-0.019** **	-0.015
	(0.010)	(0.010)	(0.008)	(0.009)	(0.009)	(0.017)
Canada out. of Quebec-English	**0.064*** **	-0.017	**-0.018** **	**-0.018***	**-0.031*** **	**-0.027***
	(0.009)	(0.011)	(0.008)	(0.009)	(0.006)	(0.015)
Quebec-French	**0.076*** **	-0.001	0.006	0.007	-0.011	-0.035
	(0.021)	(0.019)	(0.013)	(0.013)	(0.020)	(0.024)
Mother Tongue						
Canada-English	0.035	-0.018	**-0.018** **	**-0.020** **	**-0.032** **	**-0.025***
	(0.027)	(0.013)	(0.012)	(0.012)	(0.009)	(0.017)
Canada out. of Quebec-English	**0.036***	**-0.021***	**-0.022** **	**-0.024*** **	**-0.040*** **	**-0.033*** **
	(0.019)	(0.011)	(0.009)	(0.008)	(0.006)	(0.011)
Quebec-French	0.039	-0.005	0.000	0.001	**-0.036***	-0.032
	(0.021)	(0.018)	(0.012)	(0.013)	(0.021)	(0.022)
Skill Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	No	Yes	Yes
LanguageXTime Fixed Effects	No	No	No	No	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives. The main independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell. Language skills are defined respect to the distance of reported languages to English in Canada and Canada outside of Quebec, and to French in Quebec. There are 45 skill cells in a given year. Hence, there are 270 observations in the analysis. Each column controls different effects. Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table A2.4: The estimates of the inverse elasticity of substitution between immigrants language groups and natives for separate language groups: Alternative language definition -I

	Time -Invariant Skill Effects			**Time-Variant Skill Effects**		
	High Lang. Skilled Imm.	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**	**High Lang. Skilled Imm.**	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**
Home Language						
Canada-English	0.014	0.010	**-0.050*** **	**-0.078*** **	**0.111*** **	**-0.046*** **
	(0.014)	(0.016)	(0.014)	(0.022)	(0.037)	(0.013)
Canada out. of Quebec-English	**0.038** **	-0.005	**-0.062*** **	**-0.059** **	0.068	**-0.051*** **
	(0.014)	(0.011)	(0.015)	(0.025)	(0.051)	(0.015)
Quebec-French	-0.030	0.033	-0.037	**-0.130*** **	-0.011	0.034
	(0.019)	(0.031)	(0.025)	(0.043)	(0.024)	(0.037)
Mother Tongue						
Canada-English	**0.033***	0.004	**-0.061*** **	**-0.065***	**0.060***	**-0.061*** **
	(0.017)	(0.011)	(0.014)	(0.022)	(0.037)	(0.013)
Canada out. of Quebec-English	**0.057*** **	-0.005	**-0.070*** **	-0.048	0.038	**-0.061*** **
	(0.014)	(0.011)	(0.015)	(0.025)	(0.051)	(0.015)
Quebec-French	0.001	0.025	-0.038	**-0.119*** **	0.026	-0.002
	(0.036)	(0.030)	(0.023)	(0.048)	(0.028)	(0.044)
Skill Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	Yes	Yes	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives. The main independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell. Language skills are defined respect to the distance of reported languages to English in Canada and in Canada outside of Quebec and to French in Quebec. There are 15 skill cells in a given year. Hence, there are 90 observations in the analysis. The first three columns assume that skill effects are time invariant. The last three columns assume that skill effects vary over time. Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table A2.5: The estimates of the inverse elasticity of substitution between immigrants language groups and natives: Alternative language definition -II

	Immigrant Language Groups vs. Natives					
	(I)	**(II)**	**(III)**	**(IV)**	**(V)**	**(VI)**
Home Language						
Canada-English	**0.082*****	-0.012	-0.011	-0.012	**-0.017***	-0.011
	(0.010)	(0.009)	(0.008)	(0.009)	(0.009)	(0.017)
Canada out. of Quebec-English	**0.078*****	-0.012	-0.011	-0.012	**-0.020****	-0.018
	(0.012)	(0.009)	(0.007)	(0.008)	(0.009)	(0.017)
Mother Tongue						
Canada-English	**0.045*****	-0.013	-0.012	**-0.014****	**-0.022****	**-0.023***
	(0.016)	(0.011)	(0.006)	(0.006)	(0.009)	(0.012)
Canada out. of Quebec-English	**0.059*****	-0.013	**-0.013****	**-0.015****	**-0.026*****	**-0.030****
	(0.015)	(0.012)	(0.006)	(0.007)	(0.007)	(0.012)
Skill Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	No	Yes	Yes
LanguageXTime Fixed Effects	No	No	No	No	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives. The main independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell. Language skills are defined respect to the distance of reported languages to English in Canada and Canada outside of Quebec. There are 45 skill cells in a given year. Hence, there are 270 observations in the analysis. Each column controls different effects. Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table A2.6: The estimates of the inverse elasticity of substitution between immigrants language groups and natives for separate language groups:
Alternative language definition -II

	Time -Invariant Skill Effects			**Time-Variant Skill Effects**		
	High Lang. Skilled Imm.	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**	**High Lang. Skilled Imm.**	**Medium Lang. Skilled Imm.**	**Low Lang. Skilled Imm.**
Home Language						
Canada-English	0.014	-0.015	**-0.040****	**-0.078*****	0.028	-0.016
	(0.014)	(0.015)	(0.015)	(0.022)	(0.026)	(0.015)
Canada out. of Quebec-English	**0.038****	-0.026	**-0.045****	**-0.059****	-0.014	-0.007
	(0.014)	(0.015)	(0.017)	(0.025)	(0.036)	(0.017)
Mother Tongue						
Canada-English	**0.033***	-0.011	**-0.047*****	**-0.065***	0.005	**-0.031***
	(0.017)	(0.014)	(0.015)	(0.032)	(0.019)	(0.015)
Canada out. of Quebec-English	**0.057*****	-0.017	**-0.051*****	-0.048	-0.003	-0.024
	(0.015)	(0.017)	(0.016)	(0.036)	(0.022)	(0.017)
Skill Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	Yes	Yes	Yes

Source: The Canadian Censuses between 1981 and 2006.
Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives. The main independent variable is the logarithm of the relative immigrant-native labour supplied within each skill cell. Language skills are defined respect to the distance of reported languages to English in Canada and Canada outside of Quebec. There are 45 skill cells in a given year.Hence, there are 270 observations in the analysis.Each column controls different effects.Regressions are weighted with appropriate sampling weights and standard errors are clustered within skill cells.* indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table A2.7: Simulated total wage effects in Canada: 1981-2006

Short-Run Total Wage Effects Between 1981-2006				
	All Education	**High Sch. Dropuouts/ High Sch. Grads**	**Trade Cert./ Below Bachelor**	**Bachelor/ Above Bachelor**
Low Language-Skilled Imm.	-19.8%	-3.5%	-16.2%	-43.2%
Medium Language-Skilled Imm.	-3.7%	4.5%	-3.6%	-31.5%
High Language-Skilled Imm.	-6.6%	3.5%	-3.9%	-23.6%
All Immigrants	-7.3%	2.9%	-5.1%	-30%
Natives	-3.4%	2.3%	-3.1%	-20.3%
All Workers	-4.3%	2.4%	-3.7%	-22.9%

Long-Run Total Wage Effects Between 1981-2006				
	All workers	**High Sch. Dropuouts/ High Sch. Grads**	**Trade Cert./ Below Bachelor**	**Bachelor/ Above Bachelor**
Low Language-Skilled Imm.	-15.6%	0.6%	-12%	-38.9%
Medium Language-Skilled Imm.	0.5%	8.6%	0.5%	-27.3%
High Language-Skilled Imm.	-2.4%	7.7%	0.3%	-19.4%
All Immigrants	-3.1%	7.1%	-0.9%	-25.8%
Natives	0.8%	6.5%	1.0%	-16.1%
All Workers	0.0%	6.6%	0.5%	-18.7%

Language Skills Based on Home Language				
	All workers	**High Sch. Dropuouts/ High Sch. Grads**	**Trade Cert./ Below Bachelor**	**Bachelor/ Above Bachelor**
All Workers –Short-Run	-4.9%	2.5%	-3.6%	-26.5%
All Workers – Long-Run	-0.1%	7.4%	0.9%	-21.3%

Source: The Canadian Censuses between 1981 and 2006.

Notes: The mother tongue-based simulations assume that $1/\sigma_{EXG}$ is 0.050, $1/\sigma_{EX}$ is 0.200 and $1/\sigma_{E}$ is 0.450. Additionally, the share of labour and capital is assumed to be 0.66 and 0.33, respectively. The elasticity of substitution between capital and labour is 1. Home language-based simulations assume that $1/\sigma_{EXG}$ is 0.

Chapter 3

Gender, Linguistic Diversity, and Labour Market Substitutability

3.1 Introduction

The effect of immigration on host country wage levels depends on the substitutability of immigrants and natives in the labour market. Perfect substitutability between immigrants and natives implies that an increase in labour supply due to immigration decreases wages for both groups, while imperfect substitutability implies different effects on immigrants' and natives' wages. In the case of imperfect substitutability, we would expect larger declines in immigrants' wages as compared to natives' wages. This is of particular policy importance for Canada given the country's high immigration rates and documented increasing wage gap between immigrants and natives in recent decades.[1] According to Statistics Canada, the percentage of foreign-born Canadians increased from 15.3% in 1971 to 20.6% in 2006. The likelihood of hardship experienced by immigrants increased as well. Between 1981 and 2004, the average annual earnings of the Canadian-born increased slightly, but those of newcomers entering Canada as skilled principal applicants declined by 42% (Beach, Green and Worswick, 2011).

[1]See for example Baker and Benjamin (1994), Bloom, Grenier and Gunderson (1995), Grant (1999), Green and Worswick (2004), and Picot, Hou and Coulombe (2007).

The studies that estimate immigrant-native substitutability do not all reach the same conclusion. Imposing an aggregate production function on the national economy and allowing for differences in education and work experience, Aydemir and Borjas (2007) estimate immigrant-native substitutability for Canada, the United States and Mexico. They find perfect substitutability between immigrants and natives in all countries. In contrast, using the same methodology and only US data, Ottaviano and Peri (2008, 2012) conclude that immigrants and natives are imperfect substitutes. Borjas, Grogger and Hanson (2008) challenge Ottaviano and Peri's results, and show that the evidence for imperfect substitutability is sensitive to sample restrictions and empirical methodology.[2] My previous chapter builds on this work by allowing for differences in male immigrants' language skills in addition to differences in education and years of labour market experience in estimating immigrant-native substitutability. Language skills have received significant attention as an important measure of human capital contributing to wage differences between immigrants and natives. In fact, speaking the host country language is one the most important skills predicting labour market success of immigrants.[3] My results from Chapter 2 suggest that male immigrants and natives are not perfect substitutes in Canada when immigrants have low levels of language skills.[4]

Chapter 2, like much of the previous literature, focuses only on male workers. Such a restriction is commonly justified by statistics showing that males comprise the majority

[2]Borjas et al. (2008) describe three issues that bring into question the results of Ottaviano and Peri. First, including young people such as those aged 18 in the sample might create the spurious negative relationship between immigrant-to natives' relative wages and relative labour supplies due to the low-paid high-school students working part-time. Second, Borjas et al. (2008) argue that how wages are measured, the average of log wages versus the log of average wages as in Ottaviano and Peri (2008, 2012), might affect the results. Third, they emphasize the importance of using the appropriate weighting scheme in the analysis.

[3]Chiswick and Miller (1995, 2001 and 2007) describe a detailed model for second language acquisition in the destination country for immigrants. In this model, the factors such as exposure to the host country's official language before migration, the incentive for investment in language-capital and the availability of resources for investment determine capital accumulation. For example, if the national language in the source country is the same as the official language in the host country, the likelihood of accumulating language skills after migration increases. Further, the greater the distance between languages, the higher the cost of return migration. Hence, the likelihood of accumulating language skills in the host country's official language increases as well.

[4]I find that imperfect substitutability is more likely when immigrants are low language-skilled as compared to medium or high language-skilled in English. Language skills are defined according to a linguistic distance measure, the Levenshtein Distance (LD). Based on the LD representing the "distance" between official languages of Canada and the reported language of an immigrant, immigrants are grouped into three language skill groups: the high, medium and low language-skilled.

of the labour force. This is of course less true now than it was decades ago. According to Ferrao (2010), only 37% of the Canadian labour force was female in 1976, but in 2009, female workers made up 48% of the Canadian labour force. Females are included in the studies examining immigrant-native substitutability in the United States labour market and their inclusion does not change the main conclusions regarding substitutability. For example, Borjas et al. (2008) find perfect immigrant-native substitutability for both female and male workers whereas Ottaviano and Peri (2008, 2012) find imperfect substitutability for both groups.

The main focus of this chapter is to estimate the immigrant-native substitutability for female workers in the Canadian labour market taking language skills into account. Considering females separately in this context is potentially interesting as there are reasons to expect differences in language skill accumulation between male and female immigrants. On the one hand, child care might accelerate the language skill accumulation of women as compared to men because women play a larger role in child care and children develop language skills faster than adults, in which case language skills are likely transmitted from children to mothers. On the other hand, children might become interpreters in the outside world and create a disincentive for mothers to accumulate language skills. Although I do not estimate the effect of children on language skills or wages, these channels might help to explain any differences. Additionally, cultural differences might lower the exposure of women to the host country's official languages in the country of origin before they emigrate. Further, labour market prospects might not give enough incentive for immigrant women to invest in language-capital compared to men because of family-related discontinuities in labour market participation or cultural-based barriers within their community. In fact, differences in language skill accumulation between female and male immigrants across ethnic groups are well-understood. For example, female South East Asian refugees are less likely to speak English than their male counterparts, and the difference in language skills is observable even after ten years of living in Canada (Beiser and Hou, 2000). In sum, the differences in the process of language skill accumulation between men and women might create gender-based skill groups among immigrants.

There are two main contributions of this chapter. The research is the first to compare estimates of the elasticity of substitution between immigrants and natives for females and males in the Canadian labour market, using a four-level nested-CES production function, which various skill types of labour are based on educational attainment, labour market experience, and language skills. Second, as an extension, this study is the first to estimate the elasticity of substitution between female and male immigrants by exploiting variations in immigration-induced labour supply across gender groups over time. This is done by extending the model to a fifth level capturing gender. Including this fifth level allows for the possibility that male and female immigrants may not be perfect substitutes in the Canadian labour market, mainly due to the differences in language skill accumulation processes as discussed above. Estimating female-male immigrant substitutability due to language skills provides insight into the differences in immigrant wages between females and males.

The study uses the 20% micro data files of the Canadian Censuses between 1981 and 2006. Since the Canadian Census does not provide direct measures of language ability, the information on home language and mother tongue is used to assign immigrants into language skill groups based on a linguistic distance measure called the Levenshtein Distance Index (LD). This linguistic distance measure, as discussed in previous chapters, is an approximation of the hardships experienced by immigrants in accumulating language skills in the host country. Substitutability is estimated separately for each of three language skill groups (low, medium and high), for two different measures of language skills based on mother tongue and home language. All analyses are carried out for the two regions separately, Canada outside of Quebec and Quebec, to account for linguistic differences across Canada's provinces as explained in Chapter 2.

Using the nested-CES function at the fourth level, I find that the degree of substitutability between immigrant language groups and natives does not change significantly when I separately analyze males and females in all Canadian provinces outside of Quebec. For example, the estimated elasticity of substitution for both males and females ranges from 30 to infinity when calculated using home language as the measure of language skill. The

analysis of the pooled sample of female and male workers, however, reduces the estimates regardless of language definition. For example, using home language-based categorizations, all estimates are statistically or economically insignificant referring to perfect substitutability between immigrants and natives. As a result of the extension of the model to a fifth level, my results indicate that imperfect substitutability exists between male and female immigrant workers outside of Quebec. The estimated elasticity of substitution between female and male immigrants is around 10, suggesting that the increase in labour supply associated with immigration should have had a larger impact on immigrant women's as compared to immigrant men's wages. This corresponds to the simulated decline in female immigrants' wages relative to male immigrants' wages by around 5% between 1981 and 2006. In Quebec, most estimates are imprecise and statistically insignificant even with an unexpected positive sign in some cases, which suggests that sampling error along with small cell sizes might lead to attenuation bias.

This study is organized as follows. Section II provides a detailed discussion of sample selection and definition of skill groups. The third section presents the 4th level structural model and the empirical evidence for the substitutability between the Canadian-born and immigrant female workers in the Canadian labour market. The fourth section extends the model to a 5th level and provides the estimate of the degree of substitutability between female and male immigrants. The last section concludes.

3.2 Data and descriptive statistics

This chapter makes use of the long form censuses between 1981 and 2006 conducted on a random one-in-five sample of the Canadian population.[5] Using the same definition as in Chapter 2, an immigrant is either a naturalized citizen or a non-citizen in Canada.[6] Two

[5]Since the census moved from a mandatory to a voluntary survey in 2011, I do not include the 2011 National Household Survey (NHS) because of the concerns for potentially greater sampling error in NHS than in the previous censuses.

[6]As explained in the previous chapter, non-permanent residents have been enumerated in the censuses only since 1991. As a result of a huge increase in the number of non-permanent residents and their labour market participation as Temporary Foreign Workers (TFW) during the last decades, their enumeration has become necessary to fully capture the characteristics of labour market participants and their labour market performance.

analytic samples are constructed. The employment sample is constructed to represent the aggregate labour supply in skill/year groups. It includes men and women between 21 and 64, not residing in collective dwelling, working at least one week and not attending school during the reference year.[7] This sample is further restricted to construct the wage sample in order to form a homogenous group of workers to represent the average Canadian labour price. The wage sample selects only wage and salary workers with full-time, full-year work during the reference year. The main analysis focuses on Quebec and the rest of Canada separately due to the differences in linguistic characteristics, labour markets, and the composition of immigrants across the two regions. In Canada outside of Quebec, the employment sample includes 754,010 male workers and 539,630 female workers in 1981, and 1,062,110 male workers and 957,460 female workers in 2006. The respective numbers of male and female workers in the wage sample are 514,135 and 271,635 in 1981; 701,890 and 559,645 in 2006. In Quebec, there are 263,550 male workers and 169,795 female workers in the employment sample in 1981. In 2006, the respective numbers become 316,350 and 279,835. The wage sample in Quebec includes 175,015 male workers and 90,985 female workers in 1981 and 212,315 males and 171,100 females in 2006.[8] With regards to the differences in employment and wage samples, women are more likely to work part-time than men. For example, outside of Quebec, the percentage of labour market participants who held a full-time, full-year job was 50% and 58% for female workers in 1981 and 2006, respectively. Both corresponding numbers were around 67% for male workers. In this regard, including both part-time and full-time workers in the employment sample is important to describe the aggregate labour supply more broadly, in particular, for women.

Following the methodology in the previous chapter, all workers' skill types are based on education and years of labour market experience. The education groups are: (1) no high school diploma or a high school diploma; (2) a vocational training certificate or a diploma below the university level; (3) an undergraduate university degree or higher. The

[7] In the literature, labour supply is defined as either the number of the employed workers or total hours of work. I use the number of employed and call it labour supply. Hours of work are not used in this study.

[8] The unweighted number of observations is rounded to a base of 5.

five experience categories are defined as 1 to 5, 6 to 10, 11 to 15, 16 to 20, and 20 and more years of labour market experience.[9]

Immigrant skill types are also defined according to language skills using home language and mother tongue as measures of language skills. Home language is the language spoken most often at home. Mother tongue refers to the first language learned in childhood and that is still understood at the time of the Census. The LD is used to represent the "distance" of an immigrant's reported language to English or French.[10] After assigning an LD index to each immigrant based on the reported language, immigrants are grouped into three language groups: high, medium, and low.

In Canada outside of Quebec, the Levenshtein Distance is zero if the reported language is English, suggesting that immigrants have high language skills in English. An LD with a value greater than zero indicates lower language skills in English. A cut-off language is devised to distinguish immigrants with medium language skills from immigrants with low language skills. The cut-off language is determined as Hungarian with the LD index 95.22. If an immigrant reports a language with an LD index lower than 95.22, the immigrant is considered as medium language-skilled in English. Otherwise, the immigrant is low language-skilled in English. In Quebec, the cut-off language is determined as Hebrew with the LD index 93.26.[11] If an immigrant reports more than one language spoken as mother tongue or home language including English (French), the LD is assigned based on English (French). If an immigrant reports multiple non-official languages other than English or French, the index number for the language spoken most often by the other immigrants from the same birth country is assigned as the LD. In addition to language skill groups,

[9] All details regarding sample selection rules, the definitions of Census variables, and the rules followed for the construction of skill types are presented in Chapter 2, Section 2.3 and Appendix 2.1.

[10] If an immigrant resides in Quebec, the language categorization is based on the distance of the reported language to French. If an immigrant resides in the rest of Canada, the categorization considers the distance of the reported language to English.

[11] The Indo -European languages, a language family including English and French as well, in general has lower distance to English and French compared to the non-Indo-European languages. After arranging the ranking of the LD as lowest to highest, the first non-Indo-European language in ranking, that has the lowest LD among all other non-Indo-European and Indo-European languages, is determined as the cut-off language. Chapter 2 presents alternative rules for language skill groups and discusses the results. Since the main findings are robust to alternative language skill groups, they are not revisited in this study.

females and males within immigrant skill groups are also considered as separate types and added to the analysis later as the final skill level.

The wage rate within each skill group is defined as the average logarithm of real weekly wages using individuals from the wage sample. Weekly wages are derived as the annual wages and salaries deflated with the 2005 Canadian Consumer Price Index, then divided by the number of weeks worked in the reference year. The mean wage for each skill group is calculated by averaging the log of real weekly wages of corresponding individuals within skill cells.[12] The aggregated labour supply for skill groups is measured by summing the sampling weights of individuals within skill groups using individuals from the employment sample.

Since the empirical analysis estimates language-related substitutability using the ratio of immigrant-to-native labour supply by language skills, I present the trends in employment ratios for these skill groups along with their wage trends. Figure 3.1a illustrates the trends in wages for natives and immigrant language groups, while Figure 3.1b presents immigrant-to-native employment ratios between 1981 and 2006. The wage trends use the wage sample, while trends for employment ratios include the employment sample. In both figures, the left and right panels categorize female and male immigrants based on home language in Canada outside of Quebec. Two important facts are highlighted in Figure 3.1a. First, the average real weekly wages of female workers were lower than male workers' wages between 1981 and 2006. High language-skilled male immigrants earned the highest average real weekly wage among all workers. The highest wage-earners among women, the Canadian-born and the high language-skilled, earned as high as the average wage of medium-skilled male workers at its lowest level. Second, female workers' wages followed different patterns than those of male workers. For example, the Canadian-born and high language-skilled female immigrant workers' wages followed steadily increasing patterns, while the wage patterns of corresponding male groups were almost stable between 1981 and 1996. Medium and low language-skilled female immigrants' wage patterns were almost stable between 1981

[12]The cell sizes for the average wages of male workers range from 35 to 49,115 in Quebec and from 110 to 118,330 outside of Quebec. For the average wage of female workers, the number of observations in cells declines further. All cell sizes range from 15 to 38,570 in Quebec and from 30 to 99,055 outside of Quebec. The unweighted numbers of observations in cells are rounded to base 5.

and 1996 and then followed an increasing trend until 2006 whereas the average wages for the corresponding language skill groups for males followed a v-shaped pattern, reaching its lowest point in 1996. In Figure 3.1b, it is illustrated that the labour supply ratios of low language-skilled immigrants to natives increased for both men and women, particularly after 1986. Between 1986 and 1996, the relative labour supply of high language-skilled immigrants to natives declined for men and women although immigrants in the labour market were predominantly high language-skilled. Overall, the strikingly different patterns for male and female wages are not observed for the trends in immigrant-to-native employment ratios. Although there is a suggestive negative relationship between wage trends and labour supply ratios, in particular, for male immigrants with low language skill levels, it requires a detailed analysis which takes into consideration the detailed skill types within aggregated labour supply. Given that there are striking differences in wage patterns by gender but substantial similarities in employment ratio trends, these figures indicate that there might be skill-related differences between female and male immigrants. In light of this, examining language-related substitutability between female and male immigrants as an extension is a worthwhile exercise.

Figures 3.2a and 3.2b present wage and employment ratio trends for female and male workers based on mother tongue. As in Figure 3.1a, female and male wage trends are strikingly different. Further, female wages are lower than those of males in Figure 3.2a, parallel to Figure 3.1a. In Figure 3.2b, employment ratios of high language-skilled immigrant-to-native declined for both males and females whereas employment ratios of low language-skilled immigrant-to-native increased for both groups over time. Mother tongue-based language categorizations reveal some discernible differences for men's and women's employment ratios in all language categories. For example, in regard to the labour supply of medium language-skilled immigrants relative to natives, the decline in the labour supply ratio was larger for male immigrants than for female immigrants between 1981 and 1996. The decrease (increase) in high (low) language-skilled immigrant-to-native ratios refers to the shift in birth countries from traditional to non-traditional countries. Additionally, the halt or reversal of the declining trends in employment ratios of high and medium language-

skilled immigrants after 1996 might be related to increasing the weights assigned to language skills in 1996 and the requirement of a third-party test results as proof of language skills in 2001.

Comparing the wage trends presented in Figures 3.1 and 3.2, the average real weekly wages of female workers are higher if they are grouped based on mother tongue rather than home language in Canada outside of Quebec. This suggests that mother tongue-based categorizations include immigrants with higher language skills than the reported mother tongue implies. In this respect, home language is indeed a better approximation for language skills of immigrants. In addition, trends in employment ratios clearly illustrate that the changes in relative labour supply of immigrant-to-native based on home language is smaller than the changes in relative labour supply based on mother tongue between 1981 and 2006. For example, the decline (increase) in high (low) language-skilled immigrants' relative supply is smaller for home language-based categorizations than those for mother tongue-based categorizations. This implies that the rise in the labour supply of low language-skilled immigrants might not be as large as the increase suggested by mother tongue-based categorizations. Further, mother tongue-based categorizations might reflect other factors such as the lack of information in the labour market institutions and networking within ethnic communities contributing to immigrant-native substitutability rather than language skills per se.

Figures 3.3 and 3.4 illustrate the same trends in Quebec for home language- and mother tongue-based language categorizations of immigrants according to French. The differences in wage trends of female and male workers in Quebec are similar to those in all other Canadian provinces. For example, female wages follow different patterns than male wages regardless of language skill definition, as in the rest of Canada. Also, wages of immigrants by mother tongue-based language groups are higher than those of immigrants by home language-based categorizations. However, different employment patterns are found in Quebec. As an example, according to home language-based categorizations, the ratio of labour supply of low language-skilled immigrants to that of natives did not change in Quebec between 1981 and 2006 while it steadily increased in the rest of Canada after 1986.

Moreover, the supply of high language-skilled immigrants relative to natives in Quebec was the lowest among the other mother tongue-based immigrant language groups almost the whole period. These two figures suggest that the composition of the labour supply of immigrants in Quebec differs from the labour supply in the rest of the country. These illustrations provide further justification to treat Quebec and the rest of Canada separately.

To motivate the extension in which I examine language-related substitutability between female and male immigrants, I present the trends in relative wages and relative labour supply of female-to-male immigrants. Figure 3.5a illustrates trends in female-to-male wage ratios for immigrant language groups outside of Quebec. Immigrant language groups are defined with respect to the distance of home language to English. In Figure 3.5a, female immigrant workers in all language groups fared better over time, as expected from previous figures. Additionally, females with high and medium language skills saw a large improvement in their wages relative to their male counterparts in comparison to low language skill groups of immigrants. Figure 3.5b presents the female-to-male employment ratios for immigrant workers. The figure shows that there are large differences across language skill groups. There was a steep increase in the relative labour supply for high language-skilled female-to-male immigrants, whereas the relative labour supply for low language-skilled female-to-male changed slightly, in particular, between 1981 and 1996. Figure 3.5a and 3.5b indicate that the relation between the relative wage and relative supply of female-to-male immigrants varied across language groups. In this case, controlling for education and years of labour market experience might be necessary to shed light on the degree of substitutability between female and male immigrants regarding language skills if any exists.

Figure 3.6 presents the trends in wage and employment ratios of female-to-male immigrants in Quebec. Female-to-male immigrant wage ratios were the highest for low language-skilled immigrants between 1981 and 2001. In 2006, high language-skilled female-to-male wage ratios exceeded the female-to-male wage ratios of medium language-skilled immigrants. The main difference between Quebec and the rest of Canada is observable for employment ratios between 1981 and 2006. In Quebec, the lowest female-to-male employment ratio was among high language-skilled immigrants, while the highest occurred among

medium language-skilled immigrants. In the rest of Canada, as presented in Figure 3.5, the lowest female-to-male employment ratio was observed among medium language-skilled immigrants. The difference in the two regions might be related to the percentage of English-speaking immigrants in Quebec that are categorized into the medium language-skilled group based on the distance of English from French.

Taken together the figures from this section lead to three main messages with regards to estimating elasticity of substitution. First, the data show that wage trends differ by gender whereas employment ratios are similar. These differences warn against pooling females and males together when examining immigrant-native substitutability. Second, the relative wages for three language groups follow more or less similar patterns, in particular, in Canada outside of Quebec whereas the relative labour supply varies across language groups. In this case, examining gender-based language substitutability might shed light on whether immigration contributes to wage differences between female and male immigrants. Third, the differences in trends in Quebec and the rest of Canada make a separate analysis of two regions essential.

3.3 Immigrant-Native substitutability

3.3.1 Structural model

The model used in this chapter follows the same model in the previous chapter. Both models borrow the notation directly from Borjas (2003) and Ottaviano and Peri (2008). The CES function imposed upon the aggregate economy is assumed as:

$$Q_t = [\lambda_{Kt} K_t^v + \lambda_{Lt} L_t^v]^{1/v} \tag{3.1}$$

where Q is output, K is capital, and L is labour at time t. The elasticity of substitution between capital and labour is $\sigma_{KL} = 1 - 1/v$. Time-variant productivity parameters that sum up to 1 are λ_{Kt} and λ_{Lt}. Following the first-level, a second-level nest is added to the model to divide the aggregate labour, L_t, into different labour groups based on education. The second-level nest is used as follows:

$$L_t = \left[\sum_s \theta_{st} L_{st}^{\rho}\right]^{1/\rho} \tag{3.2}$$

where $\sigma_E = 1 - 1/\rho$ is the elasticity of substitution across education groups. The supply of workers with education s at time t is L_{st}. Time-variant productivity parameters, θ_{st}, sum up to 1. In order to incorporate different experience levels into each education group, a third-level CES nest is defined as:

$$L_{st} = \left[\sum_x \alpha_{sxt} L_{sxt}^{\eta}\right]^{1/\eta} \tag{3.3}$$

where $\sigma_X = 1 - 1/\eta$ is the elasticity of substitution across experience groups with the same education, and L_{sxt} the supply of workers with experience x and education group s at time t . The productivity shift parameters of experience groups that sum up to 1, α_{sxt}, vary over time.

Ottaviano and Peri (2008) allow substitutability between immigrants and natives due to various factors such as language skills and a lack of information in the labour market information by extending the model in Borjas (2003) to a fourth-level. The findings suggest that immigrants and natives are imperfectly substitutable. My model takes this extension one step further and allows for language skill differences between immigrants and natives as a potential source for immigrant-native imperfect substitutability. On this basis, I add:

$$L_{sxt} = [(1 - \sum \beta_{Msxgt}) N_{sxt}^{\gamma} + \beta_{Msxgt} M_{sxgt}^{\gamma}]^{1/\gamma} \qquad g = H, M, L \tag{3.4}$$

where $\sigma_{EXG} = 1 - 1/\gamma$ is the elasticity of substitution between immigrants and natives, and N_{sxt} is the supply of native workers with education-experience sx, at time t, M_{sxgt} is the supply of immigrant workers with education-experience sx, language skill g at time t. I define three language skill groups as High (H), Medium (M), and Low (L). The productivity shift parameters, β, sum up to 1 and they are time-variant. I use Equation (3.4) to estimate the degree of immigrant-native substitutability for female workers and the pooled sample of female and male workers in the Canadian labour market. It should be noted that the

estimate refers to an "average" value, implying that the degree of substitutability is the same for different language groups.

The equations above imply an estimation model:

$$ln(\frac{w_{Msxgt}}{w_{Nsxt}}) = ln(\frac{\beta_{Msxgt}}{1-\beta_{Msxgt}}) - \frac{1}{\sigma_{EXG}} ln(\frac{L_{Msxgt}}{L_{Nsxt}}) + u_{sxgt} \qquad g = H, M, L \tag{3.5}$$

where w_{Msxgt} and w_{Nsxt} are the average wages of immigrants and natives in education-experience group sx and language skill group g at time t. w_{Msxgt}/w_{Nsxt} is the immigrant-to-native relative wage whereas L_{Msxgt}/L_{Nsxt} is the immigrant-to-native relative labour supply. the immigrant-to-native demand-shifting factors are $\beta_{Msxgt}/(1-\beta_{Msxgt})$.

In order to identify these factors and estimate the model, time-variant skill effects are captured by separately representing education-experience-language fixed effects and time fixed effects in the estimated equation. Additionally, it is assumed that error terms are not correlated with relative labour supply for immigrants and natives. In particular, the specified equation becomes

$$ln(\frac{w_{Msxgt}}{w_{Nsxt}}) = \varphi_{sxg} + \varphi_t + \frac{1}{\sigma_{EXG}} ln(\frac{L_{Msxgt}}{L_{Nsxt}}) + u_{sxgt} \qquad g = H, M, L \tag{3.6}$$

where φ_{sxg} are skill fixed effects across education-experience-language groups of sxg. Time effects are captured with φ_t. I use four alternative specifications to identify time-variant demand shifting factors. First, I separate skill effects from time effects. Second, I add education fixed effects interacted with time fixed effects to control for time-variant education effects. Third, I add experience by time fixed effects to control for the changes in the quality of newcomers. Lastly, I include language by time effects to capture the changes in the composition of language skill groups. Error terms are clustered within skill groups in all specifications. I use Equation (3.6) to estimate the inverse elasticity of substitution between immigrants and natives for two different samples. First, I only include female workers and estimate the inverse elasticity of substitution between female immigrant workers and the Canadian-born female workers. Second, I analyze the pooled sample of men and women.

In this literature, endogeneity concerns due to native mobility or immigrants' decisions to settle in cities with thriving labour markets is resolved by the structure imposed on the aggregate economy. The skill levels defined at the national level take into consideration their response to market equalizing forces. In this regard, dividing the national labour market into local markets is an impediment to control for market equalizing forces. However, as presented in the previous chapters, both the differences in labour market performance of immigrants and the changes the composition of immigrants over time make a separate analysis of two regions essential. Since, as documented in Chapter 2, the national level analysis does not change the main conclusions driven from Canada outside of Quebec, I do not provide the estimates for the national market in this chapter. Instead, I focus on the impact of the inclusion of female workers on the estimates. I use immigration-induced labour supply shifts across skill groups to examine the impact of the increase in the aggregate labour supply on wages in the two labour markets.

3.3.2 Estimates of immigrant-native substitutability

This section presents the estimates of the inverse elasticity of substitution between immigrant language groups and natives, $(1/\varphi_{sxg})$, for two samples. First, I present the estimates for female workers. Then, the pooled sample of men and women is analyzed. The inverse elasticity of substitution between female immigrant language groups and female natives is estimated with 270 observations: 3 education groups, 5 years of labour market experience groups, and 3 language skill groups in each of six Censuses. I present the estimates of immigrant-native language substitutability for the two language definitions available: home language and mother tongue.

The estimates of the inverse elasticity of substitution between female immigrant language groups and female natives, $(-1/\varphi_{sxg})$, are expected to be negative. The estimates statistically different from zero imply imperfect substitutability between immigrants and natives. The estimates close to zero imply perfect substitutability. For female workers residing in Canada outside of Quebec, the results from almost all of the specifications using home language-based skill categorizations, as presented in Table 3.1, suggest that female

immigrants and natives are perfect substitutes, similar to the findings for male workers presented in the previous chapter. Except for the specification presented in the column(IV) that controls for time-variant education and time-variant experience effects, the estimates of the elasticity of substitution are statistically insignificant suggesting perfect substitutability between immigrants and natives. All estimates range between -0.001 and -0.031.

Mother tongue-based language skill categorizations provide a different set of estimates for the inverse elasticity of substitution. All estimates range between -0.029 and -0.079, providing approximate values of elasticity of substitution between 12 and 30. The estimates imply that immigrants and natives are imperfect substitutes. Although the estimates vary across alternative specifications, the imperfect substitutability is preserved through different identifying assumptions. These results suggest that a negative association between the likelihood of adopting official languages and the distance of an immigrant's mother tongue to official languages might contribute to the negative relationship between relative supply and relative wage. Additionally, other factors might enhance immigrant-native substitutability. These factors might include the lack of institutional knowledge, the hardship of transferring the skills into the Canadian labour market, networking. In this regard, language skills per se might not be the only contributing factor to immigrant-native substitutability.

Since the average estimate of immigrant-native language substitutability may curtail the differences in the degree of substitutability between immigrants and natives across language groups, Table 3.2 presents the estimates for the inverse of the elasticity of substitution for female immigrants with different language skills and natives. These regressions are estimated with 90 observations: 3 education groups and 5 experience groups for six census years. Time-variant demand-shifting factors are controlled for by including education and experience effects interacted with time effects. For workers residing in Canada outside of Quebec, all of the estimates , $(-1/\sigma_{EXG})$, imply perfect substitutability between immigrants and natives for home language-based skill categorizations. The fact that the hypothesis of the zero value of $(-1/\sigma_{EXG})$ is not rejected even at the 90 percent significance level for each language group suggests female immigrants and their female Canadian

counterparts are perfectly substitutable regardless of female immigrants' language skills with respect to home language. Measuring language skills based on mother tongue, female immigrants with low language skills are imperfectly substitutable with the Canadian-born female workers, with an estimated elasticity around 12. Although the size of the estimates differs, this particular result is again similar to those obtained in the previous chapter which looked only at male workers. Hence, when female and male workers are analyzed separately, the main conclusions for both groups are not so different.

Tables 3.3 and 3.4 present the estimates of the elasticity of substitution between immigrants and natives for the pooled sample of female and male workers. When language skills are constructed based on home language, the estimates of $(-1/\sigma_{EXG})$ are found to be either statistically or economically insignificant. However, when language skills are based on mother tongue, the estimated inverse elasticity is found to be between -0.028 and -0.055, suggesting immigrant and native workers are imperfect substitutes regarding language skills. Compared to the estimates based on female workers alone, when male and female workers are pooled, the estimates of the inverse elasticity of substitution are lower in absolute value in Canada outside of Quebec. The results are not different for separate language groups either. The strong evidence for imperfect substitutability for immigrants and natives in Table 3.4 is only observed for immigrants with low language skills with respect to mother tongue-based analysis.

Turning to the results in Table 3.1 for Quebec, all estimates imply perfect substitutability between immigrant and native female workers if language skills are based on home language. For mother tongue-based categorization, beside the specification in column (IV), the findings imply perfect substitutability as well.[13] In comparison to the findings related to male workers as presented in the previous chapter, there are no striking differences in the degree of immigrant-native language substitutability between female and male workers if language skills are based on home language. However, for mother tongue-based

[13]In Chapter 2, alternative language categorizations are examined for Quebec. For example, instead of the distance of an immigrant's language to French, immigrants are categorized based on the distance of the language to English and the main results stay the same. Since the estimates for female workers in this chapter do not seem to significantly differ from those for male workers in the previous chapter, this research does not revisit those alternative language categorizations for Quebec.

categorization, the size of the estimates declines in absolute value. I find similar results for separate language groups as presented in Table 3.2. The results in Table 3.3 and Table 3.4 present the estimates for the pooled sample. Home language-based specifications imply perfect substitutability in Quebec and the rest of Canada whereas mother tongue-based specifications imply imperfect substitutability outside of Quebec. As discussed in Chapter 2, these results should be considered with caution because smaller cell sizes likely enhance the sampling error problem, which might push the estimates toward zero.

Before summing up the results, it should be noted that this chapter does not consider a more complex analysis needed to address the changes in labour force participation and career progression among women over the period. If the factors such as the number of children and the access to child care are positively associated with the relative supply of women, the estimated coefficients would underestimate the impact of relative supply on relative wage. The potential impact of confounding factors might be particularly in effect for mother tongue-based estimates. Recall that mother tongue is not a choice variable and might be a source of specific type of measurement error related to linguistic distance. In case of the female immigrants' participation, if immigrant women's decision on the number of children based on the opportunities for access to child care is positively associated with linguistic distance, the estimated coefficient for women would be smaller in size as compared to the estimated coefficient for men. Although Adsera and Ferrer (2015) find that there are striking similarities between female immigrants' and natives' participation in the Canadian labour market for immigrants entering Canada between 1986 and 2006, these confounding factors might still play a role in the differences in the size of the estimated coefficients between men and women.

There are three main results in these tables. First, the elasticities based on home language imply perfect substitutability whereas elasticities based on mother tongue suggest imperfect substitutability. This might arise from the fact that mother tongue-based categorizations capture other characteristics of immigrant groups related to source regions and lessen the degree of substitutability between immigrants and natives. With regards to perfect substitutability based on home language, language skills by themselves might not be the

only factor contributing to immigrant-native substitutability. Second, low language-skilled immigrants based on mother tongue are the most likely to be imperfectly substitutable with natives regardless of the sample definition in all Canadian provinces outside of Quebec. Third, the decline in the size of estimates in Quebec is suggestive of attenuation bias, which makes the estimates biased toward zero.

3.4 Extension: Female-Male immigrant substitutability

3.4.1 Motivation

There are two important trends in the Canadian labour market. First, female labour force participation has increased significantly in recent decades.[14] Figure 3.7 illustrates the participation rates of Canadian-born and immigrant women aged 21 to 64 in the Canadian labour force in the census years from 1981 to 2006. The figure shows very different trends for the two groups. Participation rates were more similar between the groups in the earlier part of the sample. Canadian-born women's labour force participation was around 60% and female immigrants' participation was around 63% in 1981. Starting with 1991, the gap in participation between the two groups increased and the differential reached almost 8 percentage points in 2006.

These trends mask some interesting variation among immigrants. Figure 3.8 presents the labour force participation rates of female immigrants from different source regions: North America, Europe, Asia and Africa.[15] The participation rates of female immigrants from Africa and Asia declined dramatically between 1991 and 1996. In 2006, female immigrants from Europe participated in the labour market at a higher rate than women from Asia although European women's labour market participation rate was almost 7 percentage points lower than that of Asian women in 1981. Moreover, the literature establishes that labour force participation of female immigrants varies across ethnic groups. For example, Filipino women participate in the Canadian labour market at a higher rate

[14] According to Statistics Canada (2015), the participation rate for women over 15 years of age was almost 87% of that of men in 2014, up from around 58% in 1976.

[15] South/Central America, Australia and Pacific regions and the residual category are not presented for clarity of the graph.

than Arab, West Asian, Korean and Japanese immigrant women (Preston and Gilles, 2004). These extensive differences in labour supply among female immigrants likely affect the average price of labour across immigrant groups with different language skills.

Second, there has been a narrowing in the female-male wage gap in the Canadian labour market. Figure 3.9 illustrates this narrowing gap in female-to-male wages for natives and immigrants between 1981 and 2006. For both groups, the female-to-male wage ratio increased over time, suggesting that the gap has been narrowing. Moreover, the gap between immigrants' and natives' female-to-male wages has narrowed, starting in 1991. There is a literature attempting to explain the narrowing wage gap between men's and women's wages. Most studies use decomposition techniques in order to identify the contributing factors.[16] A small but related literature exploits variations in the supply of gender-based labour to estimate the female-male substitutability among workers, and hence, its impact on the wages of male and female workers. First, Acemoglu, Autor, and Lyle (2004) exploit the increase in female labour supply associated with the enlistment of men in the United States around the Second World War. The authors find that female and male workers are imperfect substitutes, with the elasticity of substitution around 3. However, the impact of female labour shifts varies across skill groups. Females were found to be closer substitutes for particularly low-educated male workers, which contributed to the expanding earning gap between high and low educated men in the United States after the Second World War. Second, De Giorgi, Paccagnella, and Pelizziari (2014) estimate the elasticity of substitution between female and male workers by exploiting the changes in military enrolment and the number of people conscripted for military service in Italy. The results suggest that an elasticity of substitution between 1.0 and 1.4 between young female and male workers in the Italian labour market.

As an extension, I estimate the degree of substitutability between female and male immigrants by taking the advantage of the increase in female labour supply over time and variations in labour supply across female and male immigrant groups. In this study, I include only immigrant workers and exclude Canadian-born workers from this

[16] Vincent (2013) reviews the studies examining the female/male wage gap in the Canadian labour market using these decomposition methodologies.

analysis. There are two main reasons for this exclusion. First, as a result of increasing educational attainment and labour force participation among women, the gap in human capital factors between males and females has narrowed in recent decades. Second, the main focus of my analysis is the contribution of language skills into workers' substitutability. Language skills, as discussed earlier, have been shown to play an important role in explaining wages of immigrants and there is a reason to think that male and female immigrants might acquire these language skills differently because of the differences in pre- and post-migration experiences. Thus, exploiting variations over time in the relative supply of immigrants, I can estimate the degree of substitutability between female and male immigrants to help explain the wage difference between female and male immigrants.

3.4.2 Structural model

A fifth level extension is used in this study in order to model female-male skill differences among immigrants. Since I concentrate on skill differences by language in this study, I solely focus on the degree of substitutability related to gender among immigrants. In this regard,

$$M_{sxgt} = [\kappa_{fsxgt} f_{sxgt}^{\varphi} + \kappa_{msxgt} m_{sxgt}^{\varphi}]^{1/\varphi} \tag{3.7}$$

where $\sigma_{f-m} = 1 - 1/\varphi$ is the elasticity of substitution between females and males within immigrants, and M_{sxgt} the supply of immigrant workers with education-experience-language skill sxg at time t. The productivity shift parameters of females and males, κ, are time-variant and assumed to sum up to 1.

The wage equations for female and male immigrants imply a common estimation model:

$$\ln(\frac{w_{M_f,sxg,t}}{w_{M_m,sxg,t}}) = \phi_{sxg} + \phi_t - \frac{1}{\sigma_{f-m}} \ln(\frac{L_{M_f,sxg,t}}{L_{M_m,sxg,t}}) + \varepsilon_{sxgt} \tag{3.8}$$

where $w_{M_f,sxg,t}$ and $w_{M_m,sxg,t}$ are the average wages of females and males in group sxg representing 45 education-experience-language skill groups in census year t. $L_{M_f,sxg,t}$ and $L_{M_m,sxg,t}$ are the aggregate labour supply by the corresponding groups. ϕ_{sxg} captures the relative productivity of female versus male workers of similar education, years of labour mar-

ket experience and language skill. Relative productivity captures variations across groups over time with ϕ_t. The remaining time variation ε_{sxgt} is independent of the relative labour supply. Henceforth, the inverse elasticity of substitution between female and male immigrants, $(\frac{1}{\sigma_{f-m}})$, is estimated by using the equation (3.8). The error terms are clustered within skill cells to take into account their correlation within groups.

The main identifying assumption is that education-experience-language skill effects are separable from time effects although they are time-variant. Therefore, time effects are controlled for by including time fixed effects in addition to the education-experience-language fixed effects. Later, all skill effects – education, experience and language skills – are interacted with time effects separately in alternative specifications. In these specifications, education effects interacted with time fixed effects presumably control for the changes in the quality of education within groups over time. Both experience and language skills interacted with time effects take into account the changes in the quality of entering immigrant cohorts over time.[17]

3.4.3 Estimates of female-male immigrant substitutability

For the estimates of the inverse elasticity of substitution between female-male immigrant groups, I have 270 available observations as a result of 45 skill groups (3x5x3) in each Census. The results for all language skill groups based on both mother tongue and home language are shown in Table 3.5. Each column refers to an alternative identification assumption and each row refers to a language measure (home language or mother tongue) and the region. Controlling for time-variant demand-shifting factors as presented in specifications (IV) and (V), imperfect substitutability between male and female immigrants is found regardless how immigrants are categorized with respect to language skills in Canada outside of Quebec. These specifications provide estimates of $-1/\sigma_{f-m}$ for home language between -0.093 and -0.109, and for mother tongue between -0.084 and -0.091. Hence, the elasticity of substitution ranges between 9.2 and 12. Categorizing immigrants based on the distance of

[17]Following Borjas (2003) and Borjas et al. (2008), the weights at the 5th level are defined as $(n_F * n_M)/(n_F + n_M)$,where n_F is the number of female immigrants within respective education-experience-language group and n_M is the number of male immigrant in the relevant skill cell in the wage sample.

home language or mother tongue to French in Quebec provides a different set of results from the rest of Canada. For Quebec, the estimates of $(-1/\sigma_{f-m})$ range from -0.035 to 0.063 across specifications in case language skills are based on home language. Most estimates for mother tongue based categorizations have unexpected positive sign for $(-1/\sigma_{f-m})$.

In Table 3.6, I present the inverse elasticity of substitution between female and male immigrants for separate language groups. These regressions have 90 observations, 15 skill groups (3x5) over six Censuses. The three columns present the estimates of $-1/\sigma_{f-m}$ under the assumption that demand-shifting factors are time-variant. These effects are controlled for by including education-experience fixed effects, time fixed effects, education fixed effects interacted with time fixed effects, and experience fixed effects interacted with time fixed effects.

Although the estimated coefficients in Table 3.6 have the expected negative sign, they are all statistically insignificant for Canada outside of Quebec. The estimate of $-1/\sigma_{f-m}$ is found to be -0.085 for high language-skilled immigrants, -0.105 for medium language-skilled immigrants and -0.048 for low language-skilled immigrants. It should be noted that these estimates are large in size for both mother tongue- and home language-based skill categorizations but they are imprecise. For Quebec, the estimates for all language groups are statistically insignificant, and have the positive sign for some language groups, particularly high language-skilled immigrants regardless of language definition. These imprecise large estimates with unexpected signs again suggest that as cell sizes become smaller, the sampling error problem becomes more serious (Aydemir and Borjas, 2011). Alternatively, it might also suggest that the CES model is not supported for Quebec's labour market.

In conclusion, there are three findings obtained from this extension. First, regardless of whether mother tongue or home language is used to measure language skills, female and male immigrants are found to be imperfectly substitutable for the rest of Canada. Second, for Canada outside of Quebec, similar estimations from home language- and mother tongue-based specifications suggest that language skills per se might not be the only factor behind imperfect substitutability between female and male immigrants. The stronger

results, although imprecise, for high and medium language-skilled immigrant groups might refer to factors other than language skills such as differences in cultural gender roles across different groups. Third, the results for Quebec significantly differ from the results for the rest of Canada. The estimates are imprecise with a positive sign across specifications for Quebec.

3.5 Calculated wage effects for female-male immigrants

In this section I use the elasticity of substitution between female and male immigrants found above to estimate the partial wage effects for female and male immigrants. The estimated elasticity of 10 in the rest of Canada suggests that female immigrants are likely to face heavier economic hardships than do male immigrants. As the focus of this extension is only the elasticity of substitution between female and male immigrants and estimate female-male immigrant substitutability by language group, I will mainly concentrate on partial wage effects instead of total wage effects.

The estimates between female and male immigrants indicate that the degree of substitutability is around 10 regardless of language definition although it varies across skill groups. Using the estimates from home language-based categorizations in Table 3.6, the largest decline occurred among high language-skilled female immigrants, around 2.6%. Medium language-skilled and low language-skilled female immigrants are less severely affected. Their wages dropped by 1.7% and 0.5%, respectively. Between 1981 and 2006, all female immigrants' wages declined by 4.7% outside of Quebec.[18] Using the mother tongue-based estimates, the wages of female immigrants dropped by 5.8%. The wage effects varied across language groups. High and medium language-skilled female immigrants experienced increases between 2.1% and 2.3% while low language-skilled immigrants saw a 1.4% decline in their wages.[19]

[18] According to home language-based categorizations between 1981 and 2006, the log relative supply of high and medium language-skilled female-male immigrants increased by 0.306 and 0.158 units while low language-skilled immigrant increased by 0.098 units outside of Quebec. Hence, by using the estimates from Table 3.6 -0.047=(-0.085*0.306) +(-0.105*0.158) +(-0.048*0.098).

[19] The log relative supply of high and medium language-skilled female-male immigrants decreased by 0.171 and 0.302 units while low language-skilled immigrant increased by 0.250 units outside of Quebec. -0.058=(-0.125*-0.171) +(-0.078*-0.302) +(-0.055* 0.250).

According to the calculations based on the estimates of the inverse elasticity of substitution, female immigrants' wages declined between 1981 and 2006 as compared to the wages of male immigrants. Home language-based calculations suggest that female immigrants with high language skills in English are most severely affected whereas mother tongue-based specifications indicate that female immigrants with the most distant languages make up the only group that experiences a decline in their relative wages. One reason for this difference can be related to the fact that mother tongue-based language skill definitions more or less divide the immigrants with non-official languages into the two groups: European and non-European immigrants. As a result, combined with the evidence that female labour market participation varies across source regions, the geographical concentration in language skill groups can contribute to the decline in the relative wage of female immigrants with low language skills in English. Home language-based definitions spread this effect over language skill groups. Hence, this might lead to the different immigration-induced wage effects for mother tongue- and home language-based definitions. Occupations might play a significant role. If female immigrants from some source regions are concentrated in occupations for which wages stay stagnant or decline over time and/or language skills are not valued as occupation-required skills, this might also contribute to different wage effects of immigration for the two language measures.

3.6 Conclusion

The objective of this chapter is to include female immigrants, while examining the degree of substitutability between immigrants and natives and its impact on wages. Language skills are emphasized as an important determinant of human capital of immigrants. Following the same methodology used in previous chapter, a CES production function is used to obtain the estimate of the substitutability between immigrants and natives. The estimates from the pooled sample of females and males provide smaller coefficients of immigrant-native substitutability than the estimates from the sample of female immigrants in Canada outside of Quebec. This finding suggests that combining female and male immigrants might incorrectly point to perfect substitutability for English-speaking Canada. In Quebec,

pooling female and male workers provides opposite results. The size of estimates increases in absolute value.

As an extension to the main analysis, when female immigrants are considered separately by language skill group, a substantial degree of imperfect substitutability between female and male immigrants, around 10, is found regardless of language definition in Canada outside of Quebec. Based on the results of my estimates, the decline in female immigrants' wages is simulated to be 4.7% and 5.8% for home language-based and mother tongue-based specifications. Female immigrants speaking English as home language experienced the largest decline relative to male immigrants. Female immigrants with the most distant languages are the only language group that experienced a decline in their relative wages between 1981 and 2006 regardless of language measure. In Quebec, the point estimates suggest that they might be subject to sampling error.

The estimation of a substantial amount of substitutability between female and male immigrants suggests that female immigrants may face heavier economic hardships relative to male immigrants in the labour market. However, the occupational distribution of female immigrants by source regions might be important. If female immigrants with the most distant mother tongues but high language skills in English are concentrated in low-paying occupations in which language skills are not valued, an employment program targeting those women based on their skills may help female immigrants to find a job that can take full advantage of their skills.

This chapter has its limitations. First, a range of the estimates from the CES structural model is not verifiable, which is a main shortcoming of this model. Second, small cell sizes raise concerns about the estimates, which might be already biased as a result of attenuation bias due to sampling error. However, it should be noted that the CES structure gives an important insight into worker complementarity. In order to overcome small cell sizes and shed more light on language-related substitutability, skills can be redefined in terms of occupation and language skills required by each occupation to complete the task. This might be particularly interesting for women if the number of the children is also considered in the analysis. Future work will tackle these two issues.

References 3

[1] **Acemoglu, Daron, David H. Autor, and David Lyle**. "Women, war and wages: The effect of female labor supply on the wage structure at midcentury". *Journal of Political Economy*, 112(3):497–551, 2004.

[2] **Addison, David A. and Christopher Worswick**. "The impact of immigration on the earnings of natives: Evidence from Australian micro data". *The Economic Record*, 78(1):68–78, 2002.

[3] **Adsera, Alicia and Ana Ferrer**. "The effect of linguistic proximity on the occupational assimilation of immigrant men". IZA Discussion Paper, no. 9499, 2015.

[4] **Adsera, Alicia and Ana Ferrer**. "Occupational skills and labour market progression of married immigrant women in Canada". *Labour Economics*, 39:88–98, 2016.

[5] **Adsera, Alicia and Mariola Pytlikova**. "The role of language in shaping international migration". *The Economic Journal*, 125(586):49–81, 2015.

[6] **Aydemir, Abdurrahman and George J. Borjas**. "Cross-country variations in the impact of international migration: Canada, Mexico, and the United States". *Journal of the European Economic Association*, 5(4):663–708, 2007.

[7] **Aydemir, Abdurrahman and George J. Borjas**. "Attenuation bias in measuring the wage impact of immigration". *Journal of Labor Economics*, 1(4):69–112, 2011.

[8] **Aydemir, Abdurrahman and Mikal Skuterud**. "Explaining the deteriorating entry earnings of Canada's immigrant cohorts: 1966-2000". *Canadian Journal of Economics*, 38(2):641–671, 2005.

[9] **Baker, Michael and Dwayne Benjamin**. "The performance of immigrants in the Canadian labour market". *Journal of Labor Economics*, 13(3):369–405, 1994.

[10] **Beach, Charles, Alan. G. Green, and Christopher Worswick**. *Toward improving Canada's skilled immigration policy: An evaluation approach.* CD Howe Institute, Toronto, 2011.

[11] **Beiser, Morton and Feng Hou**. "Gender differences in language acquisition and employment consequences among southeast Asian refugees". *Canadian Public Policy*, 25(3):311–330, 2000.

[12] **Berger, Mark C.** "Demographic cycles, cohort size, and earnings". *Demography*, 26(2):311–321, 1989.

[13] **Bloom, David E., Gilles Grenier, and Morley Gunderson**. "The changing labour market position of Canadian immigrants". *Canadian Journal of Economics*, 28(4b):987–1005, 1995.

[14] **Borjas, George J.** "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labour market". *The Quarterly Journal of Economics*, 118(4):1335–1374, 2003.

[15] **Borjas, George J.** *Immigration Economics*. Harvard University Press, 2014.

[16] **Borjas, George J., Jeffrey Grogger, and Gordon H. Hanson**. "Imperfect substitution between immigrants and natives:A reappraisal". National Bureau of Economic Research Working Paper, no. 13887, 2008.

[17] **Borjas, George J., Richard B. Freeman, Lawrance F. Katz, John DiNardo, and John M. Abowd**. "How much do immigration and trade affect labor market outcomes?" *Brookings Papers on Economic Activity*, 1997(1):1–90, 1997.

[18] **Boyd, Monica and Xingshan Cao**. "Immigrant language proficiency, earnings, and language policies". *Canadian Studies in Population*, 36(1-2):63–86, 2009.

[19] **Butcher, Kristin F. and David Card** . "Immigration and wages: Evidence from the 1980's". *The American Economic Review*, 81(2):292–296, 1991.

[20] **Card, David and Giovanni Peri**. "Immigration economics by George J. Borjas: A review essay". *Journal of Economic Literature*, 54(4):1333–1349, 2016.

[21] **Chiswick, Barry R. and Paul W. Miller**. "The endogeneity between language and earnings: International analyses". *Journal of Labor Economics*, 13(2):246–288, 1995.

[22] **Chiswick, Barry R. and Paul W. Miller**. "A model of destination-language acquisition: Application to male immigrants in Canada". *Demography*, 38(3):391–409, 2001.

[23] **Chiswick, Barry R. and Paul W. Miller**. "Linguistic distance: A quantitative measure of the distance between English and other languages". *Journal of Multilingual and Multicultural Development*, 26(1):1–11, 2005.

[24] **De Giorgi, Giacomo, Marco Paccagnella, and Michele Pelizziari**. "Gender complementarities in the labour market". Bank of Italy Occasional Paper, no. 183, 2014.

[25] **Ferrao, Vincent**. "Paid work". Statistics Canada Catalogue no:89-503-X. Ottawa: Statistics Canada, 2010.

[26] **Ferrer, Ana, David A. Green, and William C. Riddell**. "The effect of literacy on immigrant earnings". *Journal of Human Resources*, 41(2):380–410, 2006.

[27] **Ferrer, Ana, Garnett Picot, and William C. Riddell**. "New directions in immigration policy: Canada's evolving approach to the selection of economic immigrants". *International Migration Review*, 48(3):846–867, 2014.

[28] **Frenette, Marc and Renè Morissette**. "Will they ever converge? Earnings of immigrants and Canadian-born workers over the last two decades". Analytical Studies Branch Research Paper Series, no. 215. Statistics Canada Catalogue 11F0019MIE. Ottawa: Statistics Canada, 2003.

[29] **Friedberg, Rachel M. and Jennifer Hunt**. "The impact of immigrants on host country wages, employment and growth". *The Journal of Economic Perspectives*, 9(2):23–44, 1995.

[30] **Grant, Mary L.** "Do immigrants affect the wage of native-born workers? Evidence from Canada". Institute for Policy Analysis. University of Toronto. `http://hd.handle.net/1807/13124`, 1998.

[31] **Grant, Mary L.** "Evidence of new immigrant assimilation in Canada". *Canadian Journal of Economics*, 32(4):930–955, 1999.

[32] **Green, Alan G. and David A. Green**. "The economic goals of Canada's immigration policy: Past and present". *Canadian Public Policy*, 25(4):425–451, 1999.

[33] **Green, David A. and Christopher Worswick**. "Entry earnings of immigrant men in Canada: The roles of labour market entry effects and returns to foreign experience". In Ted McDonald, Elizabeth Ruddick, Arthur Sweetman, and Christopher Worswick, editor, *Canadian Immigration: Economic Evidence for A Dynamic Policy Environment*, pages 77–110. McGill-Queen's University Press, Kingston, ON, 2010.

[34] **Green, David A. and Christopher Worswick**. "Immigrant earnings profiles in the presence of human capital investment: Measuring cohort and macro effects". *Labour Economics*, 19(2):241–259, 2012.

[35] **Grenier, Gilles and Serge Nadeau**. "English as the lingua franca and the economic value of other languages: The case of the language of work in the Montreal labor market". In Michele Gazzola and Bengt-Arne Wickstrom, editor, *The Economics of Language Policy*, pages 267–312. MIT Press, Cambridge, Massachusetts, 2016.

[36] **Hou, Feng and Garnett Picot**. "The deteriorating economic welfare of immigrants and possible causes: Update 2005". Analytical Studies Branch Research Paper Series, no. 356. Statistics Canada Catalogue 11F0019M. Ottawa: Statistics Canada, 2014.

[37] **Isphording, Ingo E. and Sebastian Otten**. "The costs of Babylon–Linguistic distance in applied economics". *Review of International Economics*, 21(2):354–369, 2013.

[38] **Katz, Lawrence F. and Kevin M. Murphy**. "Changes in relative wages, 1963-1987: Supply and demand factors". *The Quarterly Journal of Economics*, 107(1):35–78, 1992.

[39] **Lewis, Ethan**. "Immigrant-native substitutability and role of language". In David Card and Steven Raphael, editor, *Immigration, Poverty, and Socioeconomic Inequality*, pages 60–97. The Russell Sage Foundation, New York, 2013.

[40] **Lewis, Paul, Gary S. Simons, and Charles D. Fenning (Ed.)**. *Ethnologue: Languages of the World*. Seventeenth edition. International, Dallas, Texas, 2013. Online version: `http://www.ethnologue.com`.

[41] **Morin, Louis-Philippe**. "Do men and women respond differently to competition? Evidence from a major education reform". *Journal of Labor Economics*, 33(2):443–491, 2015.

[42] **Nadeau, Serge and Aylin Seckin**. "The immigrant wage gap in Canada: Quebec and the rest of Canada". *Canadian Public Policy*, 36(3):265–285, 2010.

[43] **Okkerse, Liesbet**. "How to measure labour markets effects of immigration: A review". *Journal of Economic Surveys*, 22(1):1–30, 2008.

[44] **Oreopoulos, Philip**. "Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes". *American Economic Journal: Economic Policy*, 3(4):148–171, 2011.

[45] **Ottaviano, Gianmarco I. P. and Giovanni Peri**. "Immigration and national wages: Clarifying the theory and the empirics". National Bureau of Economic Research Working Paper, no. 14188, 2008.

[46] **Ottaviano, Gianmarco I. P. and Giovanni Peri**. "Rethinking the effects of immigration on wages". *Journal of the European Economic Association*, 10(1):152–197, 2012.

[47] **Picot, Garnett and Arthur Sweetman**. "The deteriorating economic welfare of immigrants and possible causes: Update 2005". Analytical Studies Branch Research Paper Series, no. 262. Statistics Canada Catalogue 11F0019MIE. Ottawa: Statistics Canada, 2005.

[48] **Picot, Garnett and Feng Hou**. "The rise in low-income rates among immigrants in Canada". Analytical Studies Branch Research Paper Series, no. 198. Statistics Canada Catalogue 11F0019MIE. Ottawa: Statistics Canada, 2003.

[49] **Picot, Garnett and Feng Hou**. "Chronic low income and low-income dynamics among recent immigrants". Analytical Studies Branch Research Paper Series, no. 364. Statistics Canada Catalogue 11F0019M. Ottawa: Statistics Canada, 2014.

[50] **Picot, Garnett and Patrizio Piraino**. "Immigrant earnings growth: Selection bias or real progress?" Analytical Studies Branch Research Paper Series, no. 340. Statistics Canada Catalogue 11F0019M. Ottawa: Statistics Canada, 2012.

[51] **Picot, Garnett, Feng Hou, and Simon Coulombe**. "Chronic low income and low-income dynamics among recent immigrants". Analytical Studies Branch Research Paper Series, no. 294. Statistics Canada Catalogue 11F0019MIE. Ottawa: Statistics Canada, 2007.

[52] **Skuterud, Mikal**. "Language skills in the new economy and the deteriorating labour market performance of Canada's immigrant workers". Report, Human Resources Development and Skills Canada, 2011.

[53] **Statistics Canada**. "Immigration in Canada: A portrait of the foreign-born population". 2006 Statistics Canada Census Catalogue no:97-557-XIE. Ottawa: Statistics Canada, 2007.

[54] **Statistics Canada**. "Changing labour force". 2006 Statistics Canada Census Catalogue no: 97-559-X. Ottawa: Statistics Canada, 2008.

[55] **Sweetman, Arthur, and Casey Warman**. "Temporary foreign workers and former international students as a source of permanent immigration". Canadian Labour Market and Skills Researcher Network Working Paper, no. 25, 2009.

[56] **Tu, Jiong**. "The impact of immigration on the labour market outcomes of native-born Canadians". IZA Discussion Paper, no. 5120, 2010.

[57] **Vincent, Carole**. "Why do women earn less than men? A syhthesis of findings from Canadian microdata". CRDCN Synthesis Series, 2013.

[58] **Warman, Casey and Christopher Worswick**. "Technological change, occupational tasks and declining immigrant outcomes: Implications for earnings and income inequality in Canada". *Canadian Journal of Economics*, 48(2):736–772, 2015.

[59] **Wooldridge, Jeffrey M.** *Econometric Analysis of Cross Section and Panel Data.* The MIT Press, 2002.

Table 3.1: The estimates of the inverse elasticity of substitution between immigrants language groups and natives for females:1981-2006

	(I)	(II)	(III)	(IV)	(V)
Home Language					
Canada out. of Quebec-English	**0.060*** **	-0.005	-0.001	**-0.031** **	-0.010
	(0.012)	(0.007)	(0.008)	(0.013)	(0.019)
Quebec-French	0.017	0.011	0.013	-0.011	0.011
	(0.019)	(0.011)	(0.010)	(0.018)	(0.025)
Mother Tongue					
Canada out. of Quebec -English	**-0.054** **	**-0.032*** **	**-0.029*** **	**-0.068*** **	**-0.079*** **
	(0.021)	(0.008)	(0.007)	(0.007)	(0.013)
Quebec-French	**-0.047*** **	0.005	0.007	**-0.032***	0.005
	(0.013)	(0.012)	(0.012)	(0.017)	(0.027)
Skill Fixed Effects	No	Yes	Yes	Yes	Yes
Time Fixed Effects	No	Yes	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	Yes	Yes
LanguageXTime Fixed Effects	No	No	No	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of female immigrants and natives while the main independent variable is the logarithm of the relative immigrant-native labour supplied by women within each skill cell. Language skills are defined respect to the distance of reported languages to English in Canada outside of Quebec and French in Quebec. There are 270 observations in each regression. Each column represents alternative specifications under different identifying assumptions. Regressions are weighted and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 3.2: The estimates of the inverse elasticity of substitution for separate language groups for females: 1981- 2006

	High Lang. Skilled Imm.	Medium Lang. Skilled Imm.	Low Lang. Skilled Imm.
Home Language			
Canada out. of Quebec-English	-0.059	-0.024	-0.046
	(0.034)	(0.029)	(0.036)
Quebec-French	0.017	-0.004	0.075
	(0.068)	(0.035)	(0.044)
Mother Tongue			
Canada out. of Quebec -English	-0.037	-0.036	**-0.082****
	(0.065)	(0.026)	(0.033)
Quebec-French	-0.012	0.009	0.054
	(0.059)	(0.041)	(0.053)
Skill Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
EducationXTime Fixed Effects	Yes	Yes	Yes
ExperienceXTime Fixed Effects	Yes	Yes	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of female immigrants and natives while the main independent variable is the logarithm of the relative immigrant-native labour supplied by women within each skill cell. Language skills are defined respect to the distance of reported languages to English (French) in Canada outside of Quebec (Quebec). There are 90 observations in each regression. Each column represents a specification for a selected language group. Regressions are weighted and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 3.3: The estimates of the inverse elasticity of substitution between immigrants language gropus and natives for the pooled sample of female and male Workers: 1981- 2006

	(I)	(II)	(III)	(IV)	(V)
Home Language					
Canada out. of Quebec-English	**0.069*** **	-0.008	-0.008	**-0.023** **	-0.006
	(0.012)	(0.006)	(0.007)	(0.009)	(0.014)
Quebec-French	**0.055** **	-0.002	-0.002	-0.020	-0.021
	(0.025)	(0.013)	(0.010)	(0.019)	(0.021)
Mother Tongue					
Canada out. of Quebec -English	-0.020	**-0.028*** **	**-0.030*** **	**-0.055*** **	**-0.047*** **
	(0.024)	(0.008)	(0.006)	(0.005)	(0.008)
Quebec-French	-0.011	-0.014	-0.014	**-0.061*** **	-0.029
	(0.019)	(0.011)	(0.008)	(0.019)	(0.018)
Skill Fixed Effects	No	Yes	Yes	Yes	Yes
Time Fixed Effects	No	Yes	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	Yes	Yes
LanguageXTime Fixed Effects	No	No	No	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives while the main independent variable is the logarithm of the relative immigrant-native labour supply within each skill cell. Language skills are defined respect to the distance of reported languages to English (French) in Canada outside of Quebec (Quebec). There are 270 observations in each regression. Each column represents alternative specifications under different identifying assumptions. Regressions are weighted and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 3.4: The estimates of the inverse elasticity of substitution for separate language groups of the pooled sample of workers: 1981-2006

	High Lang. Skilled Imm.	Medium Lang. Skilled Imm.	Low Lang. Skilled Imm.
Home Language			
Canada out. of Quebec-English	**-0.044***	0.013	-0.037
	(0.024)	(0.051)	(0.023)
Quebec-French	-0.047	-0.055	0.044
	(0.078)	(0.031)	(0.037)
Mother Tongue			
Canada out. of Quebec -English	-0.037	-0.030	**-0.068*** **
	(0.038)	(0.039)	(0.019)
Quebec-French	-0.066	-0.022	0.010
	(0.072)	(0.025)	(0.037)
Skill Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
EducationXTime Fixed Effects	Yes	Yes	Yes
ExperienceXTime Fixed Effects	Yes	Yes	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of immigrants and natives while the main independent variable is the logarithm of the relative immigrant-native labour supply within each skill cell. Language skills are defined respect to the distance of reported languages to English (French) in Canada outside of Quebec (Quebec). There are 90 observations in each regression. Each column represents a specification for a selected language group. Regressions are weighted and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 3.5: The estimates of the inverse elasticity of substitution between female and male immigrant language groups: 1981- 2006

	(I)	(II)	(III)	(IV)	(V)
Home Language					
Canada out. of Quebec-English	**0.080***	-0.040	0.021	**-0.093****	**-0.109****
	(0.043)	(0.027)	(0.030)	(0.043)	(0.048)
Quebec-French	0.054	-0.012	0.063	-0.035	-0.019
	(0.038)	(0.033)	(0.041)	(0.050)	(0.053)
Mother Tongue					
Canada out. of Quebec -English	0.070	**-0.047***	0.014	**-0.091****	**-0.084***
	(0.047)	(0.025)	(0.020)	(0.036)	(0.043)
Quebec-French	**0.084*****	0.020	**0.101****	0.035	0.041
	(0.024)	(0.033)	(0.044)	(0.054)	(0.051)
Skill Fixed Effects	No	Yes	Yes	Yes	Yes
Time Fixed Effects	No	Yes	Yes	Yes	Yes
EducationXTime Fixed Effects	No	No	Yes	Yes	Yes
ExperienceXTime Fixed Effects	No	No	No	Yes	Yes
LanguageXTime Fixed Effects	No	No	No	No	Yes

Source: The Canadian Censuses between 1981 and 2006.

Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of female and male immigrants while the main independent variable is the logarithm of the relative female-to-male immigrant labour supply within each skill cell. Language skills are defined respect to the distance of reported languages to English and French in Canada outside of Quebec and in Quebec, respectively. There are 270 observations in each regression. Each column represents an alternative specification. Regressions are weighted and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Table 3.6: The estimates of the inverse elasticity of substitution between female and male immigrants for separate language groups: 1981- 2006

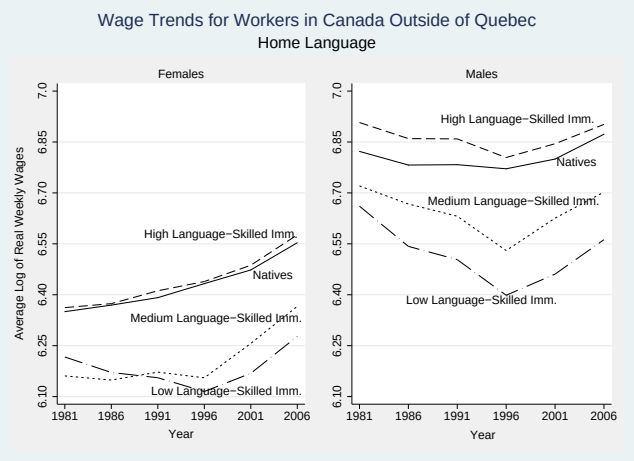
	High Lang. Skilled Imm.	Medium Lang. Skilled Imm.	Low Lang. Skilled Imm.
Home Language			
Canada out. of Quebec-English	-0.085	-0.105	-0.048
	(0.094)	(0.076)	(0.063)
Quebec-French	0.124	-0.140	-0.012
	(0.129)	(0.085)	(0.155)
Mother Tongue			
Canada out. of Quebec -English	-0.125	-0.078	-0.055
	(0.100)	(0.097)	(0.084)
Quebec-French	0.132	-0.083	0.037
	(0.128)	(0.081)	(0.138)
Skill Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
EducationXTime Fixed Effects	Yes	Yes	Yes
ExperienceXTime Fixed Effects	Yes	Yes	Yes

Source: The Canadian Censuses between 1981 and 2006.

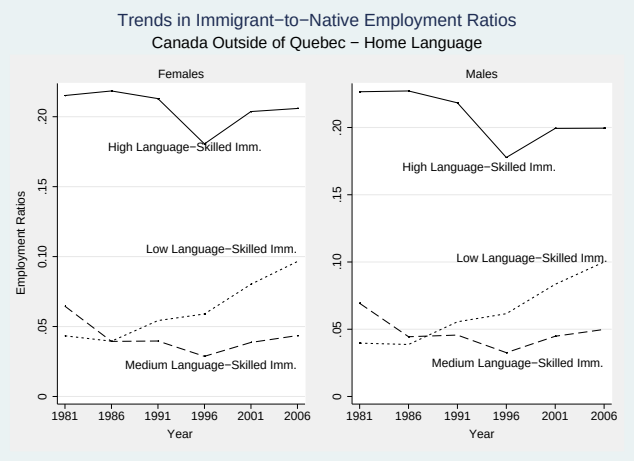
Notes: Each cell reports the estimates of $(-1/\sigma_{EXG})$. The dependent variable is the difference between the average log of real weekly wages of female and male immigrants while the main independent variable is the logarithm of the relative female-to-male immigrant labour supply within each skill cell. Language skills are defined respect to the distance of reported languages to English and French in Canada outside of Quebec and Quebec, respectively. There are 90 observations in each regression. Each column represents a specification for a specific language group. Regressions are weighted and standard errors are clustered within skill cells. * indicates significance at the 10 percent level. ** indicates significance at the 5 percent level. *** indicates significance at the 1 percent level.

Figure 3.1: Wage and employment trends for female and male workers: 1981-2006 Canada outside of Quebec - Home language

(a)



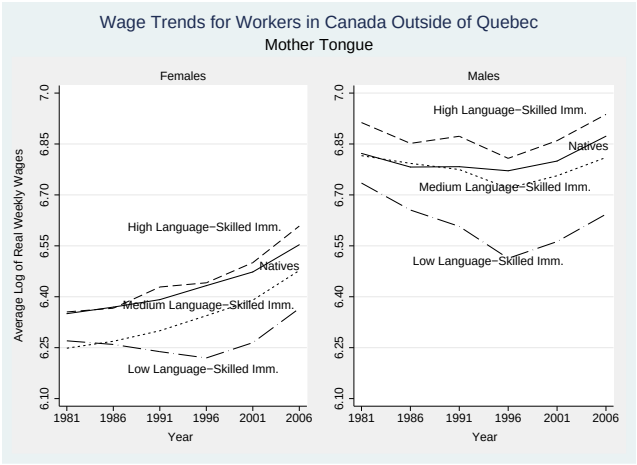
(b)



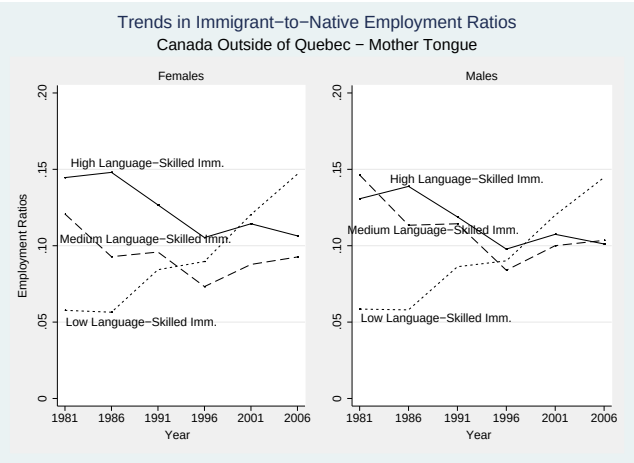
Source: The Canadian Censuses between 1981 and 2006.
Notes: Figure 3.1a includes full-time/full-year female and male workers reporting positive wages, aged 21 to 64 while Figure 3.1b includes both groups reporting at least one week of work and no school attendance, aged 21 to 64 residing in Canada outside of Quebec. The average log of real weekly wages are calculated by taking the mean of log of real weekly wages for natives and immigrant groups by using sampling weights. Figure 3.1b includes workers reporting at least one week of work in the reference year and no school attendance, aged 21 to 64. The relative immigrant-native labour supply is calculated by using the sum of the sampling weights of individuals. Figure 3.1a and Figure 3.1b categorize immigrants into language skill groups based on home language. In both figures, high language-skilled immigrants report English. The cut-off language to group immigrants into medium and low skill groups is Hungarian.

Figure 3.2: Wage and employment trends for female and male workers: 1981-2006 Canada outside of Quebec - Mother tongue

(a)



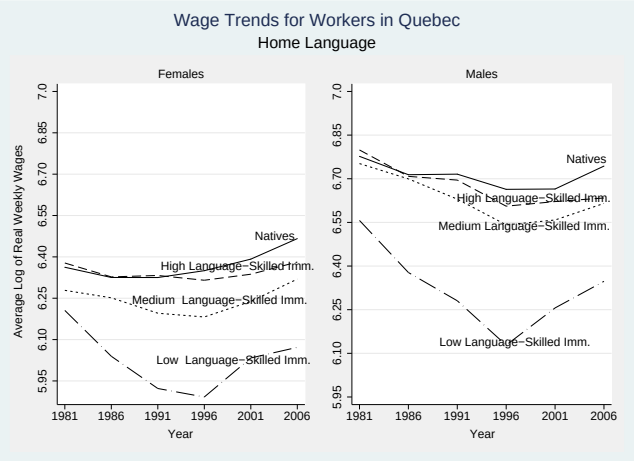
(b)



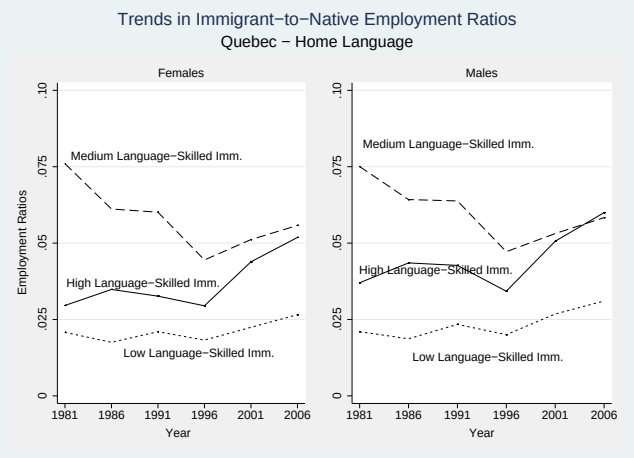
Source: The Canadian Censuses between 1981 and 2006.
Notes: Figure 3.2a includes full-time/full-year female and male workers reporting positive wages, aged 21 to 64 while Figure 3.2b includes both groups reporting at least one week of work and no school attendance, aged 21 to 64 residing in Canadian provinces other than Quebec. The average log of real weekly wages are calculated by taking the mean of log of real weekly wages for natives and immigrant groups by using sampling weights. Figure 3.2b includes workers reporting at least one week of work in the reference year and no school attendance, aged 21 to 64. The relative immigrant-native labour supply is calculated by using the sum of the sampling weights of individuals. Figure 3.2a and Figure 3.2b categorize immigrants into language skill groups based on mother tongue. In both figures, high language-skilled immigrants report English. The cut-off language to group immigrants into medium and low skill groups is Hungarian.

Figure 3.3: Wage and employment trends for female and male workers: 1981-2006 Quebec - Home language

(a)



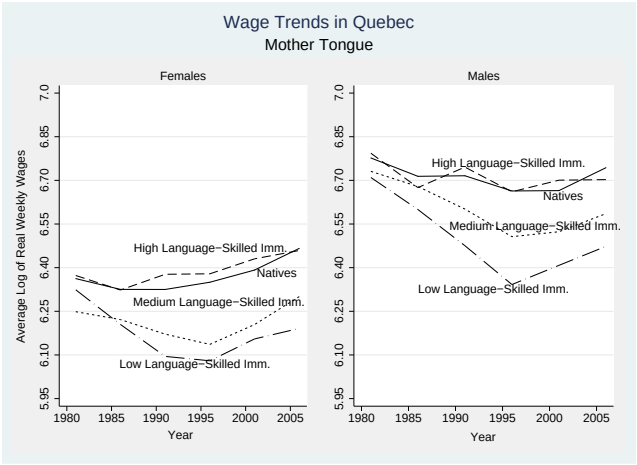
(b)



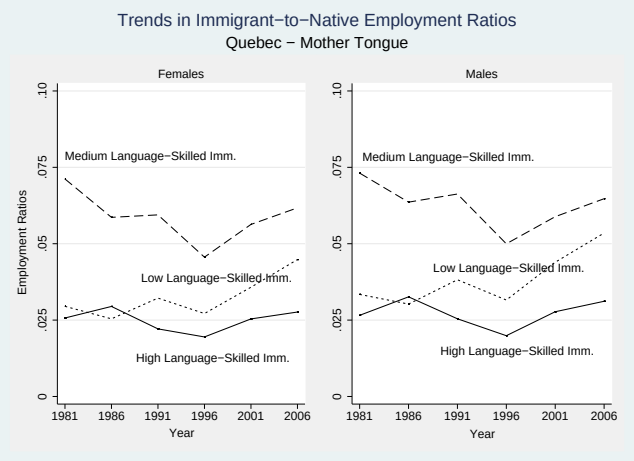
Source: The Canadian Censuses between 1981 and 2006.
Notes: Figure 3.3a includes full-time/full-year female and male workers reporting positive wages, aged 21 to 64 while Figure 3.3b includes both groups reporting at least one week of work and no school attendance, aged 21 to 64 residing in Quebec. The average log of real weekly wages are calculated by taking the mean of log of real weekly wages for natives and immigrant groups by using sampling weights. Figure 3.3b includes workers reporting at least one week of work in the reference year and no school attendance, aged 21 to 64. The relative immigrant-native labour supply is calculated by using the sum of the sampling weights of individuals. Figure 3.3a and Figure 3.3b categorize immigrants into language skill groups based on home language. In both figures, high language-skilled immigrants report French. The cut-off language to group immigrants into medium and low skill groups is Hebrew.
Source: The Canadian Censuses between 1981 and 2006.

Figure 3.4: Wage and employment trends for female and male workers: 1981-2006 Quebec - Mother tongue

(a)



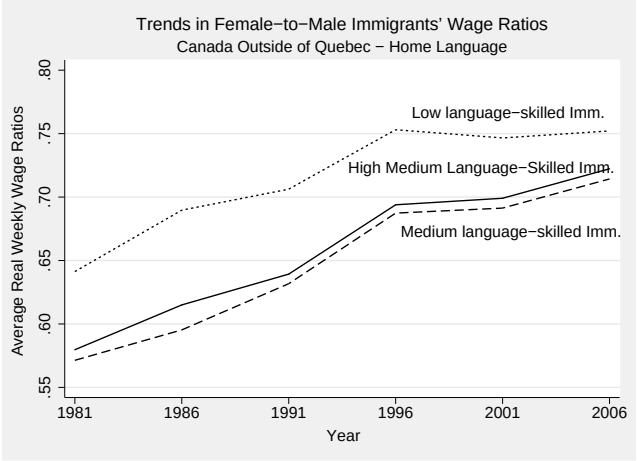
(b)



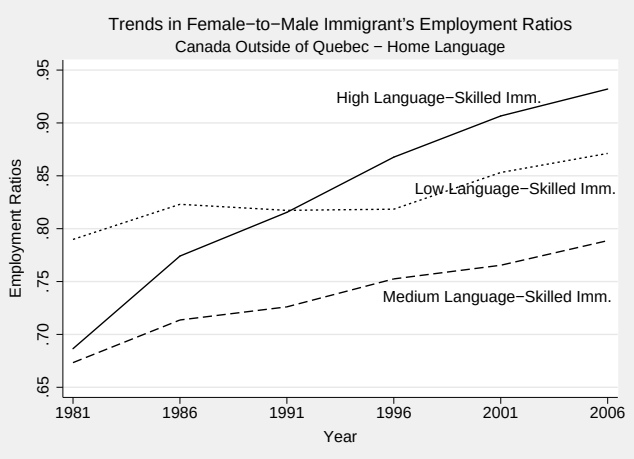
Source: The Canadian Censuses between 1981 and 2006.
Notes: Figure 3.4a includes full-time/full-year female and male workers reporting positive wages, aged 21 to 64 while Figure 3.4b includes both groups reporting at least one week of work and no school attendance, aged 21 to 64 residing in Quebec. The average log of real weekly wages are calculated by taking the mean of log of real weekly wages for natives and immigrant groups by using sampling weights. Figure 3.4b includes workers reporting at least one week of work in the reference year and no school attendance, aged 21 to 64. The relative immigrant-native labour supply is calculated by using the sum of the sampling weights of individuals. Figure 3.4a and Figure 3.4b categorize immigrants into language skill groups based on mother tongue. In both figures, high language-skilled immigrants speak French. The cut-off language to group immigrants into medium and low skill groups is Hebrew.

Figure 3.5: Female-to-Male immigrants' relative wage and employment ratios: 1981-2006 Canada outside of Quebec

(a)



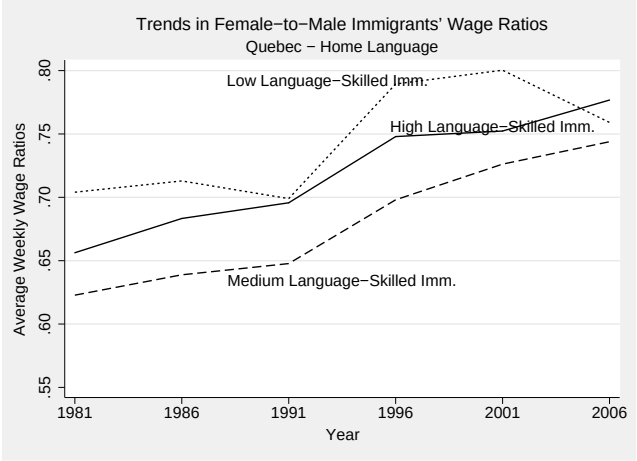
(b)



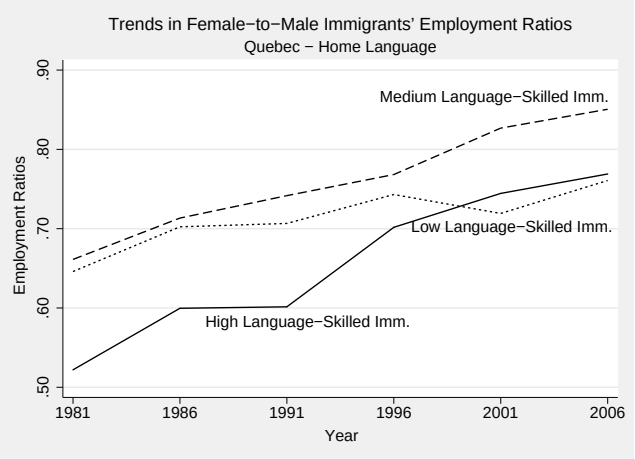
Source: The Canadian Censuses between 1981 and 2006.
Notes: Figure 3.5a includes full-time/full-year female and male immigrants, reporting positive wages, aged 21 to 64. Wage ratios between females and males immigrants are presented. The ratios are calculated as $exp(wage^f - wage^m)$, where $wage^f$ represents the average of the logarithm of real weekly wages of female workers and $wage^m$ stands for the average of the logarithm of real weekly wages of male workers. Figure 3.5b includes female and male immigrant workers working at least one week of work and not attend in school, aged 21 to 64. Language categories are defined as before.

Figure 3.6: Female-to-male immigrants' relative wage and employment ratios: 1981-2006 Quebec

(a)

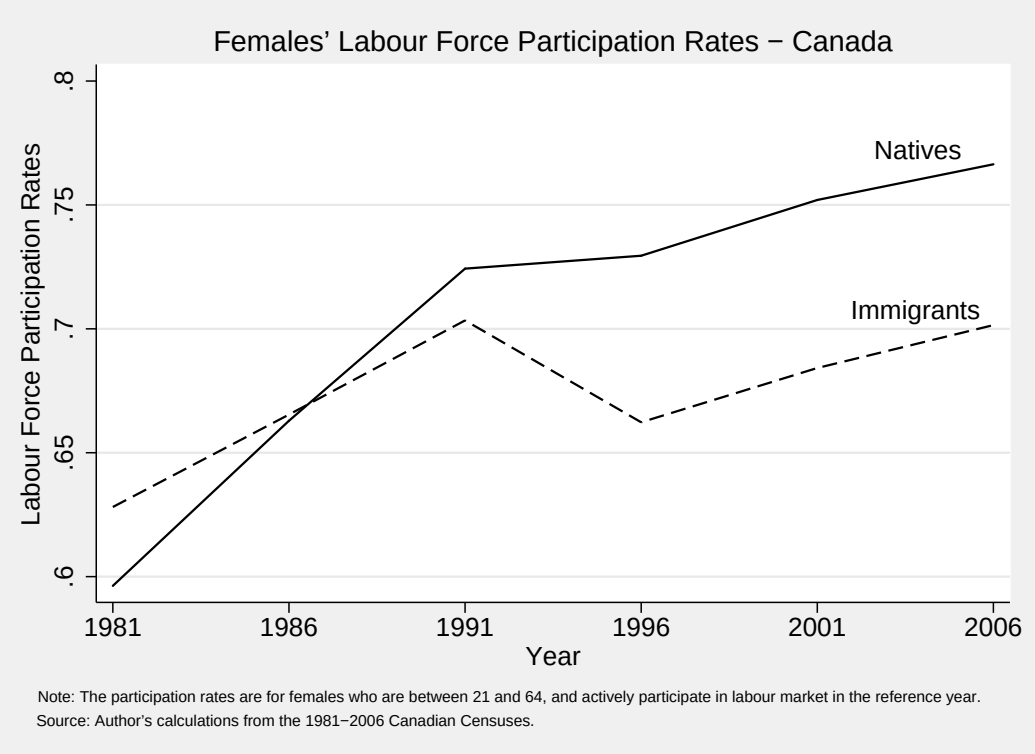


(b)



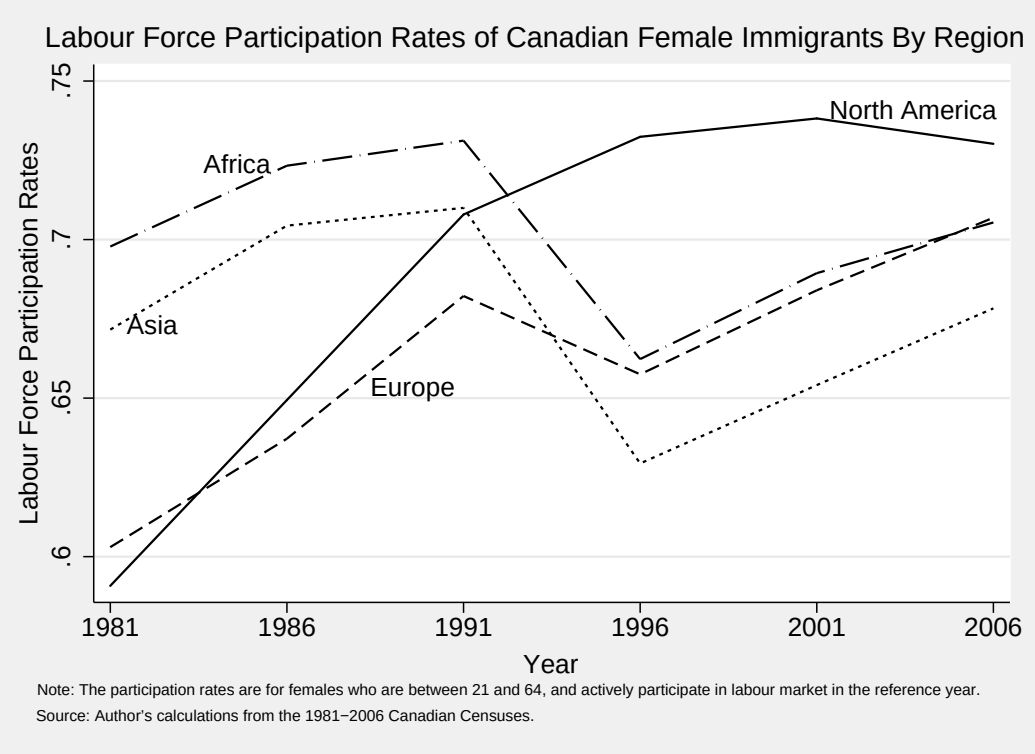
Source: The Canadian Censuses between 1981 and 2006.
Notes: Figure 3.6a includes full-time/full-year female and male immigrants, reporting positive wages, aged 21 to 64. Wage ratios between females and males immigrants are presented. The ratios are calculated as $exp(wage^f - wage^m)$, where $wage^f$ represents the average of the logarithm of real weekly wages of female workers and $wage^m$ stands for the average of the logarithm of real weekly wages of male workers. Figure 3.6b includes female and male immigrant workers working at least one week of work and not attend in school, aged 21 to 64. Language categories are defined as before.

Figure 3.7: Trends of women's labour market participation in Canada: 1981-2006 - Natives and immigrants



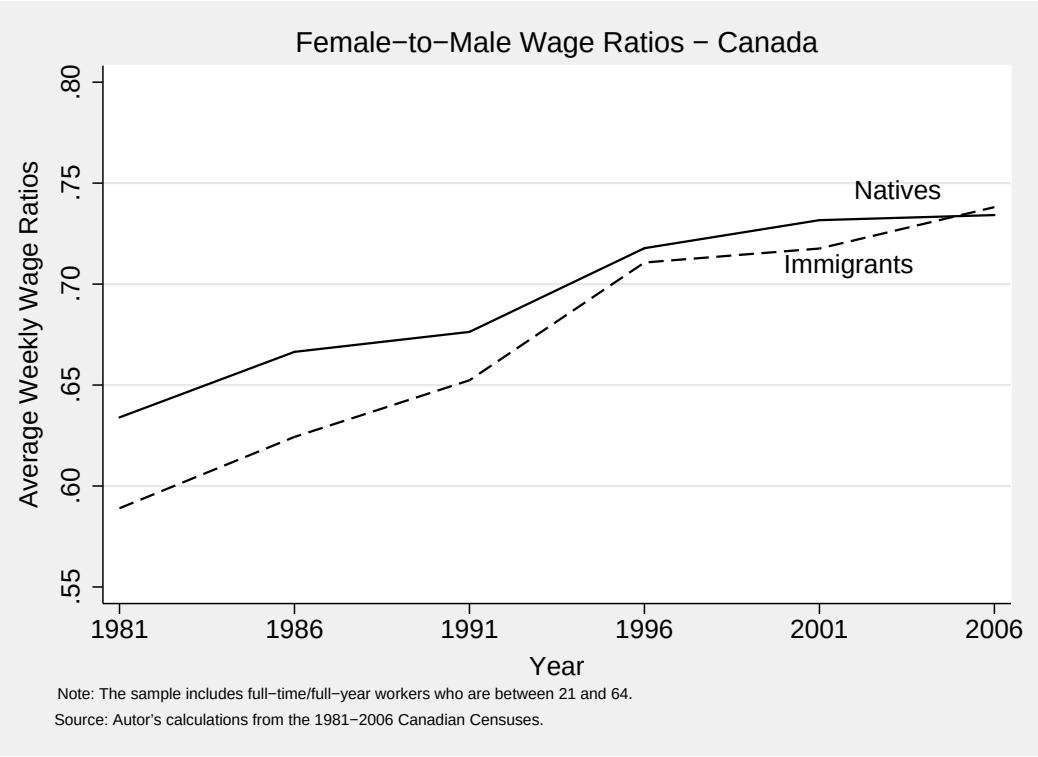
Source: The Canadian Censuses between 1981 and 2006.
Notes: The sample includes female workers aged 21 to 64, reporting active participation in the labour market in the reference week.

Figure 3.8: Trends of female immigrants' labour market participation in Canada: 1981-2006 - By region of birth



Source: The Canadian Censuses between 1981 and 2006.
Notes: The sample includes female workers aged 21 to 64, reporting active participation in the labour market in the reference week.

Figure 3.9: Trends of female-to-male wage ratios in Canada: 1981-2006 - Natives and immigrants



Source: The Canadian Censuses between 1981 and 2006.
Notes: The sample includes full-time/full-year female and male workers reporting positive wages, aged 21 to 64, reporting at least one week of work and no school attendance. The average log of real weekly wages are calculated by taking the mean of log of real weekly wages for natives and immigrant groups by using sampling weights. The trend for natives represents the relative wage of Canadian-born females to males whereas the trend for immigrants represent the relative wage of female immigrants to male immigrants.