Does Multiculturalism Work? Language, Friendship Homophily, and Well-being of Immigrants in Canada

Qiang Li*

August 24, 2017

Abstract

Multiculturalism is under attack in its birthplace, Canada. I revisit the connection between the two pillars of Canadian multiculturalism, namely acceptance of English and French as the official languages and respect and sharing among all cultural groups. A unique dataset, the 2002 Canadian Ethnic Diversity Survey, is used to analyze how learning more languages affects ethnic composition of people's social networks, which I propose as an indicator for social integration of new immigrants. I find that learning the official languages or learning more languages in general increases the ethnic diversity of a person's network. The economic and social wellbeing of immigrants are related to their language skills and ethnic composition of their social networks. The findings are generally in favor of advocates for multiculturalism.

Keywords: Multiculturalism, Integration of Immigrants, Friendship Homophily, Ethnic Diversity

JEL Classification: H19, J15

*Email: rstlq@nus.edu.sg. Address: Department of Real Estate, School of Design and Environment, National University of Singapore, 4 Architecture Drive, Singapore 117566. This paper is in part based on Chapter 3 of my PhD dissertation submitted to the University of British Columbia. I thank my supervisor Bob Helsley for guidance and helpful comments. I am also grateful for valuable suggestions made by James Brander, David Green, Keith Head, Dean Lacy, Sanghoon Lee, Tsur Somerville, Yuan Ting, and audiences at NARSC 2010, ESPE 2011, WEAI 2016, EALE 2016, and SOLE 2017. All errors are my own.

1 Introduction

In an ethnically diverse society, the ideal of multiculturalism advocates equal treatment of distinct cultural groups and promotes the preservation of cultural diversity. Thirty years after its inception, a backlash against multiculturalism emerged in Europe (Vertovec and Wessendorf, 2010; Malik, 2015). Recent terrorist attacks and the refugee crisis in Europe made multiculturalism increasingly associated with chaos and blamed for failures in assimilation of new immigrants. In Canada, the birthplace of multiculturalism, people also began questioning its viability as a way to build a coherent society. Globe and Mail Editorial (2010) even called for abandoning the term multiculturalism altogether.

When it was originally conceived, multiculturalism in Canada has two pillars: equal respect for all cultural groups and acceptance of English or French as a common language.[1] It is presumed that learning either official language will enable people of all ethnic origins to participate fully in the civil society. On the other hand, respect of all ethnicities will reduce discrimination and ethnic tensions. In turn, elimination of discrimination can generate more inter-cultural sharing. The following quote from the well-known speech by Pierre Trudeau on October 8, 1971 to the House of Commons makes this abundantly clear.

> *A policy of* ***multiculturalism within a bilingual framework*** *commends itself to the government as the most suitable means of assuring the cultural freedom of Canadians. Such a policy should help break down discriminatory attitudes and cultural jealousies. National unity if it is to mean anything in the deeply personal sense, must be founded on* ***confidence in one's own individual identity****; out of this can grow* ***respect for that of others and a willingness to share ideas, attitudes and assumptions****.*

The link between the two pillars seems to have been forgotten over time. In this paper, I revisit this connection of Canadian multiculturalism. Most importantly, if immigrants improve their skills in English or French, are they more likely to become integrated in the host society? If the answer is yes, the government's effort to help immigrants' language learning can be justified. However, acquisition of language skills cannot be forced. What then incentivize immigrants to improve their skills in the official languages and to integrate? Are these actions detrimental or beneficial to preserving individuals' confidence in their cultural identities, enhancing national unity, and building social capital, etc.? Answering those questions enable us to assess whether the many objectives of multiculturalism are achieved.

On the surface, the link between learning the official language and social integration seems to be

[1] Dewing (2009) offers a systematic treatment of the Canadian multiculturalism. He explains the various governmental programs both at the federal and provincial levels. Lupul (2005)'s memoir, on the other hand, provides a unique personal account of the historical development and the politics of this policy. Instead of being planned ahead, multiculturalism was a response to the then established policy of bilingualism, itself a response to Quebec nationalism. The original mission of the Royal Commission on Bilingualism and Biculturalism was to develop bilingualism/biculturalism further. Due to pressures from minority groups from Western Canada, the so-called "Third Force", multiculturalism was proposed as a compromise.

too obvious to warrant a serious investigation. After all, people have to communicate in a common language for relationships to form. However, the cause-and-effect relationship can happen in both ways. It may well be true that more exposure to the wider society helps to improve an immigrant's language skills. This possible reverse causality casts doubt on the usefulness of language policy in improving social integration.

Researchers have long recognized the difficulty in pinning down the causal effects of language on social and economic outcomes (Chiswick and Miller, 1995; Dustmann and van Soest, 2001). Bleakley and Chin (2004, 2010) made significant progress.[2] I adopt an identification strategy similar to theirs. More precisely, I use the interaction between age at immigration and a measure of distance between English and other minority languages (Chiswick and Miller, 2005) as an instrumental variable for language proficiency. Language proficiency, in turn, is measured by language usage at home. This variable, to a large extent, indicates the ability of the individuals in carrying out conversations in the official languages or *Lingua Franca* in Canada, namely English and French. These variables are then augmented by a language distance measure suggested by Chiswick and Miller (2005) using the identified first language.[3] The key idea is similar to Bleakley and Chin (2004). People who immigrate at a younger age is more likely to use the official language as their home language than those who immigrate later. Such differential becomes larger for people who come from a group linguistically farther away from English or French. Furthermore, I restrict the sample to young immigrants arriving before age 25 to address concerns of self-selection of immigrants. This practice is in line with the literature that follows Bleakley and Chin (2004, 2010).

There is another equally serious challenge. How do we measure integration? Many skeptics of multiculturalism argue that the policy itself encourages new immigrants to stay within their comfort zones or ethnic enclaves. This inertia creates problems for immigrants themselves and the society at large. In light of this criticism, a measure of integration must describe how an immigrant interacts with the society. The previous literature has focused on spatial segregation or ethnic enclaves. Many have tested the determinants of ethnic enclaves and their presumed negative impact on socio-economic outcomes, e.g. Borjas (1998), Edin, Fredriksson, and Åslund (2003), Cutler, Glaeser, and Vigdor (2008), Damm (2009), and many others.

In this paper, the ethnic composition of a person's friendship network is proposed as an alternative proxy for immigrant integration. In a sense, it is a measure of segregation along the social dimension. It is well-known that people tend to form friendships with those of the same ethnicity, age group, religion, or social class, etc. This phenomenon has been termed as "friendship homophily". On the other hand, an immigrant, if he interacts more intensively with the society, will achieve a less homophilous friendship

[2] For more details, please refer to the literature section.

[3] The language distance equals to 3 minus a testing score of native English speaker after a period of training in various other foreign languages. Therefore, the higher the measure the farther away the foreign language is to English. With this definition, English is zero distance away from itself. Details are in the Appendix.

network. Ethnic diversity of friendship networks in a society therefore facilitates sharing among diverse cultures. In the most idealistic characterization of a harmonious multicultural society, individuals can preserve their own identities while interacting positively with others.

My findings confirm the intuitive notion that learning official languages helps social integration. It also establishes the link between the two pillars of multiculturalism, i.e. learning official languages and mutual respect and sharing among cultural groups. Such finding is robust to alternative measurements of language proficiency, to different identification strategies, and to different samples. I argue that this relationship is causal. Many programs subsidizing new immigrants to learn the majority language can help achieving the goal of integration, at least as it is measured by their social networks.

This leads naturally to my next question. How to devise such a policy? What incentivize people to learn languages and for people to expand their networks? In other words, do these investments result in better socio-economic outcomes for themselves? Additionally, do these actions generate better results for the society as whole, e.g. better social capital? Answering these questions are essential for sound policy-making about multiculturalism as well as a rigorous evaluation of multiculturalism.

The previous literature do offer many insights.[4] The positive relationship between immigrant earnings and proficiency in the host country's majority language seems to be well established. The effect of friendship networks is mostly unexplored, especially with respect to its impact on social outcomes. However, the endogeneity issue arises again. Friendship networks can be affected by many unobservable factors that determine socio-economic outcomes. Therefore, a simple regression of outcome variables on friendship homophily likely generates inconsistent estimates.

I attempt to resolve the endogeneity of friendship homophily using childhood homophily as its instrumental variable. In my analysis, I restrict the sample to immigrants who arrive after age 15 or after childhood years. Even though immigration dirupts one's friendship networks, it does not replace his friendships completely, resulting in a persistent relationship between childhood and adulthood friendship networks. Since childhood friends of these immigrants are most likely living in the home country, they should not affect the current economic and social well-being of first-generation immigrant directly. The above statement is my key identifying assumption.

I analyze many socio-economic outcomes. They can be thought of as key performance indicators of multiculturalism. As I argue previously, official language skills and friendship homophily are two indicators directly related to the two pillars of Canadian multiculturalism. Therefore, relating language skills and friendship homophily to these socio-economic outcomes tests whether Canadian multiculturalism has achieved its goals. These goals are classified into five categories: economic or labor market well-being, attachment to one's ethnic identity, sense of belonging to Canada and community, social well-being, and civic participation.

[4] Detailed references are discussed in Section 2.

The results on language skills and success of multiculturalism are basically mixed. Language skills do increase immigrants' earnings, consistent with findings in the literature. However, they do not significantly affect employment probability or occupational progression. Perhaps surprisingly, learning official languages do not significantly decrease one's attachement to his ethnic ancetry, but it also has no significant effect on improving one's attachment to the host country. These findings are both good and bad news for multiculturalism. Even though we should not worry about immigrants losing their identities because of learning the official languages, we also do not see evidence for its positive impact on creating a unified national identity. There are some limited evidence for a positive role language plays on improving trust and social wellbeing. No statistically significant effect of language on civic participation is found. Again, these findings cast doubts on the effectiveness of language policy alone in achieving the goals of multiculturalism.

The results on friendship homophily are more encouraging to supporters of multiculturalism. Although friendship homophily does not affect earnings, employment probability, and occupational achievement significantly, it does affect the language one speaks at the workplace. If speaking non-official languages at work is an indication of labor market segmentation or frictions, we should worry about friendship homophily. In terms of social outcomes, friendship homophily significantly reinforces one's ethnic identity as we should expect. Friendship homophily increases some trust and social wellbeing variables and some civic participation indicators significantly. It does not significantly affect sense of belonging even though most estimates are positive.

The findings on friendship homophily refute common criticisms of multiculturalisms. Basically, friendship homophily does not adversely affect civic participation, pro-social behavior, trust of others, and sense of belonging to the country and to the community. The results are consistent with the following narrative. New immigrants find support within their own ethnic community, which reinforces their confidence and attachment to their ethnic identity. At the same time, their sense of belonging to their family and their community provides them with a basis to build a sense of belonging to the host country. At least, it does not adversely affect such attachment. The above narrative can be found repeatedly in the original parliament document on multiculturalism as well as in other writings promoting multiculturalism.

My attempt to assess common assumptions made by proponents and critics of multiculturalism results in a mixed bag. On balance, the findings are more in line with supporters of multiculturalism. First, the link between the two pillars of Canadian multiculturalism seems to be strong, i.e. learning a common language improves inter-cultural sharing and contacts. Second, to the disappointment of believers of Canadian multiculturalism, official language policy is not as effective as people have assumed. Even though language skills seemd to affect one's economic wellbeing significantly, many desirable social outcomes do not arise automatically after people invest in learning the official languages. Third, results on the impact of friendship homophily on performance indicators of multiculturalism undermine criticisms of

multiculturalism. However, they also raise questions about the key assumption of multiculturalism that intercultural sharing is necessary for a successful multicultural society. To really answer this question, new data and research design must be utilized, which is beyond the scope of this study.

The rest of the paper is organized as follows. Section 2 discusses antecedents in the literature, my contributions, and a theoretical framework to understand the relationship between language learning and friendship homophily. Section 3 describes the data, lays out my identification strategy, and explains the interpretation of the results. Section 4 reports the regression results. Section 5 concludes the paper and suggests future avenues of research.

2 Related Literature and Theoretical Background

2.1 Language and Immigrants' Well-being

Many labor economists have studied the relationship between language skills of immigrants and their labor market outcomes. Earlier studies, such as McManus, Gould, and Welch (1983), Kossoudji (1988), McManus (1990), and Chiswick (1991), find that language skills positively affect immigrants' employment probability, earnings, and occupation prestige. They mostly ignore the endogeneity of language skills.

The second wave of studies recognizes the endogeneity of language skills. Dustmann and van Soest (2001) categorize the issues into three types: (1) measurement errors in the language ability measures, which creates downward bias in OLS estimates; (2) unobserved ability that affect language proficiency and socio-economic outcomes in the same direction, which results in an upward bias; (3) unobserved ability that affect language learning costs and economic outcomes in the same direction, which generates a downward bias.

Chiswick and Miller (1995) employ several instruments, including whether an individual got married overseas, the number and ages of his children, and local concentration of people from his own country of origin. Dustmann and van Soest (2001) use parents' education as instruments for language proficiency and exploit the panel structure to address the measurement-error problem. Dustmann and Fabbri (2003) instrument their language measure by the language used during the interview. These studies highlight the difficulties in identifying the causal effect of language on socio-economic outcomes.

Bleakley and Chin (2004) make a significant progress in addressing this endogeneity problem. Their identification strategy is motivated by both our common experience and rigorous studies in psychobiology. It is well-known that young children learn new languages more effectively than adults. This is termed the *"critical period hypothesis"*.[5] To add to the credibility of their instrument, they do not use the age at immigration variable itself but rather an interaction term between age at arrival and whether the immigrant comes from a non-English speaking country. Presumably, age at arrival affects other aspects

[5] See Bleakley and Chin (2004) for further references for these studies.

of assimilation, e.g. adaptation to the education system, and learning of the cultural and social norms, and therefore age at arrival cannot be excluded from a wage equation. In contrast, the interaction term, which represents the additional challenge for immigrants from non-English speaking countries to master the English language, reflects the effect of language proficiency. They also alleviate concerns about the endogenous decision of immigration age by focusing only on childhood immigrants. All in all, Bleakley and Chin (2004) represent a step forward from the previous literature.

Bleakley and Chin (2010) further extend their approach to the analysis of social outcomes. They find that language proficiency significantly affects marriage, fertility, and locational choice of immigrants. In this paper, I expand this line of inquiry by focusing on friendship homophily and a few other social outcomes.[6] The ethnic composition of an immigrant's social networks measures his integration into the host country. Other social outcomes, e.g. immigrants' attitudes toward the society, trust of others, attachment to own ethnicity, etc., gauge the success and failures of multiculturalism from different angles. Utilizing the identification strategy of Bleakley and Chin (2004, 2010), I test an arguably more credible causal link between language policy and the well-being of immigrants themselves and the society at large.

2.2 Language and Friendship Homophily: A Framework

People get both psychological and economic benefits from friendships. On the one hand, a stable long term relationship fosters mutual trust and is an indispensable part of a person's life. On the other hand, friends also are potential agents that bring material benefits, for example, referring the person to good job opportunities and offering information on new business opportunities.

A feature of social networks that received a lot of attention is the tendency for people to form social connections with those of a similar background. This phenomenon has been known as "homophily". Researchers have found "homophily" phenomenon to appear in relationships defined by ethnicity, age, religion, social class, occupation, and other characteristics. McPherson, Smith-Lovin, and Cook (2001) provide an extensive review of the sociology literature.

The economics literature on friendship homophily is relatively thin. A few empirical studies do exist. Marmaros and Sacerdote (2006) study friendship patterns of college students and find that ethnicity and geographic proximity are the most significant determinants. Currarini, Jackson, and Pin (2009) is perhaps the first economic theory of homophily. They emphasize the role of group size and type-sensitive preferences, and thus quite relevant for my analysis.[7]

I briefly explain the basic elements of Currarini et al. (2009). Utilizing their framework, the role language plays in shaping a person's social networks becomes clearer. There are two ethnic groups in a

[6] The studies that relate language skills to social outcomes are too voluminous to review here. Bleakley and Chin (2010) offer an incomplete list relevant for their analysis. I do not know any studies that model friendship networks as a function of language.

[7] The notations and discussions in this section follow Currarini et al. (2009) for the purpose of comparisons.

community.[8] A group i has a population of N_i. The total population is $N = N_1 + N_2$. The share of group i is hence

$$w_i = \frac{N_i}{N}.$$

Suppose that the average number of within-group friendships of a type i agent is denoted s_i. The number of across-group friendships is denoted as d_i.

People of type i have the following preference

$$U(s_i, d_i) = (s_i + \gamma d_i)^{\alpha}, i \in \{1, 2\},$$

where $0 < \alpha < 1$ measures the diminishing returns to more friends. The key parameter is γ. If $\gamma < 1$, people prefer to interact with friends from the same group. Reversely if $\gamma > 1$, people prefer to befriend those who are different.

Currarini et al. (2009) discuss factors affecting the values of γ in information networks, professional networks, pure social networks, and risk sharing networks. The tradeoff between potentially higher benefits from different-type individuals and the higher communication costs with these individuals is highlighted. This tradeoff is also prominent in the literature on the strength of weak ties, e.g. Granovetter (1983) and Zenou (2015). Therefore, we do not know a priori whether $\gamma < 1$. We do know, however, that an immigrant who learns the common language tilts the balance toward a higher γ since such investment reduces his communication barriers with different-type individuals. I interpret language in the more general sense, so it includes not only words, sentence structures, and grammar but also tacit knowledge.

Agents find friends through a pool of potential matches. Once they enter the matching process, they choose the length of time t_i staying in the pool so as to maximize their net utility. They incur a cost of c per unit of time not exiting. Per unit of time, N_i new agents of type i enters the matching process. Therefore, the number of type i agents still in the matching process is $t_i N_i$.

An agent of type i meets another person of type i with probability q_i. He meets people of other groups with probability $1 - q_i$. These probabilities are endogenously determined by the type composition of the potential matches. In general, $q_i \neq w_i$. If the probability q_i equals the share of type i in the matching pool, i.e. $q_i = \frac{t_i N_i}{t_1 N_1 + t_2 N_2}$, we call it an *unbiased matching process*. If that is not the case, the matching process is *biased*.

Currarini et al. (2009) offer an example that allows both unbiased and biased matching processes. The probabilities of meeting same-type and different-type individuals satisfy the following two equations:

$$q_1^{\beta} + q_2^{\beta} = 1,$$

[8] Currarini et al. (2009) actually analyze multiple groups as well. Their simplest example, however, considers the two-group case. This example is sufficient for our purpose.

and

$$N_1 t_1 (1 - q_1) = N_2 t_2 (1 - q_2).$$

The key parameter describing the matching process is β. If $\beta = 1$, $q_1 + q_2 = 1$, and it implies an unbiased matching process. If $\beta > 1$, this creates higher q_1 and q_2 at the same time, with a larger impact on the smaller group.

In terms of immigrants' social interactions, biased matching arises more naturally. Due to language barriers, meetings between people of same-type individuals are more likely. On the other hand, bilingualism or multilingualism can reduce this bias. In addition, spatial segregation of immigrants also leads to more opportunities among same-type individuals. Therefore, investments in learning a common language by immigrants reduce β. In the limit case when no barriers exist, the matching probability approaches the unbiased case.

Before I discuss the key results in Currarini et al. (2009), a few definitions are necessary.

Definition 1 *The homophily index of an individual of type i, or H_i, is defined by*

$$H_i = \frac{s_i}{s_i + d_i}. \tag{1}$$

The homophily index may not be the best measure of social integration, or lack thereof. The bias is due to group size, i.e. people of the larger group tend to have higher homophily index. Such factor is beyond the control of individuals. Coleman (1958) proposes the inbreeding homophily index, which measures the tendency to form intra-group friendships beyond the effect of relative population size.

Definition 2 *The inbreeding homophily of an individual of type i, or IH_i, is*

$$IH_i = \frac{H_i - w_i}{1 - w_i}. \tag{2}$$

The numerator is the difference between the observed homophily and the population share of group i. The denominator $1 - w_i$, measures the maximum that $H_i - w_i$ can attain. IH_i equals zero if there is *baseline homophily*. IH_i equals one if all the friends are intra-group. Of course, IH_i can be negative, too. In this case, there is *heterophily*, or the tendency to form inter-group friendships.

Figure 7 (page 1027) of Currarini et al. (2009) summarizes their key results relevant for my purpose. First, the inbreeding homophily index decreases as γ increases, i.e. as weak ties offer more benefits or involve less communication costs. Consequently, learning the common language, which reduces communication costs for inter-group relationships, decreases a person's inbreeding homophily index. This mechanism works through its impact on preference itself. Second, the inbreeding homophily index increases with β, or the bias in the matching process. Consequently, elimination of language barriers reduces the bias in matching, which in turn decreases a person's inbreeding homophily index. Language

investments alter the opportunities faced by immigrants. Overall, we expect learning the official language or more languages in general reduces an individual's friendship homophily, i.e. she is more integrated in the society.

2.3 Social Networks and Immigrants' Well-being

In comparison with sociologists, economists have become interested in social networks only recently. Two lines of research are closely related to my study. The first literature is interested in how social networks affect labor market outcomes. Ioannides and Loury (2004) and Mouw (2006) provide excellent reviews. Most of these papers do not study immigrants per se, e.g. Bayer, Ross, and Topa (2008), Hellerstein, McInerney, and Neumark (2011). Several researchers specifically study immigrants. Munshi (2003) exploits rainfall variation at communities in Mexico as an instrument for Mexican immigrants' network size and relate it to their employment probability and occupation. Edin et al. (2003), Damm (2009), and Beaman (2012) utilize the exogenous spatial allocation of refugees as natural experiments. These papers on immigration focus primarily on one attribute of social networks, namely the size.

In contrast, this paper looks at the effect of the ethnic composition of friendship networks. This choice is motivated by the distinction between strong ties and weak ties (Granovetter, 1973, 1983; Zenou, 2015). Weak ties are defined as linkage with friends' friends, while strong ties are immediate connections. In light of the friendship homophily literature, connection with a person of a different ethnicity is more likely to create weak ties. Thus, a test of whether ethnic composition of your friendship circle affects labor market outcome also indirectly tests the "strength of weak ties". Patacchini and Zenou (2012) exploit spatial concentration of ethnic groups as an instrument in their analysis of weak ties. They assume that people from the same community are more likely to form strong ties, while those from different neighborhoods are more likely to form weak ties. In a sense, my identification strategy is similar. In this paper, the variation in the number of weak ties is generated by segregation along the social dimension (Zenou, 2013). Xue (2008) analyzes similar issues using a different dataset. She utilizes the panel structure of her longitudinal data. In contrast, I adopt an instrumental variable approach, using ethnic composition of a person's childhood friends as the instrument for her friendship composition during adulthood.

The second literature analyzes the social adjustment of immigrants through social interactions. Most contributions are theoretical. Bisin and Verdier (2000) analyze intergenerational transmission of ethnic identity. They show that minorities/immigrants may not necessarily be assimilated in the majority culture. Kuran and Sandholm (2008) identify two mechanisms through which social integration can happen, namely payoffs from coordination across groups and preference interaction due to inter-group contacts. Kónya (2005, 2007) analyzes assimilation of immigrants in a multicultural society. Brueckner (2006) and Brueckner and Smirnov (2007, 2008) construct explicit models of social networks describing the evolution of identities and attributes of minorities. However, empirical tests of these theories are

rare. In testing whether friendship networks affect social outcomes, this paper fills this literature gap. Moreover, such an analysis is necessary for a comprehensive assessment of multiculturalism.

3 Data and Empirical Strategy

3.1 Data and Descriptive Statistics

The 2002 Ethnic Diversity Survey (abbreviated as EDS) carried out by both Statistics Canada and Canadian Heritage are well-suited to answer my questions. In fact, the objective of this survey is "to better understand how people's backgrounds affect their participation in the social, economic and cultural life of Canada" and "to better understand how Canadians of different ethnic backgrounds interpret and report their ethnicity".[9] It contains much richer information than the census in terms of people's ethnic identity, language usage, social networks, and many socio-economic outcomes. Such variables are essential to provide credible identifications of key relationships.

Table 1: Descriptive Statistics

	All of Canada			Gateway Cities		
	Overall	Immigration Age		Overall	Immigration Age	
		0-14	15-64		0-14	15-64
Panel A: Key Variables of Interest						
Inbreeding Homophily IH	0.4446	0.2616	0.5327	0.5158	0.3313	0.5767
	(0.3995)	(0.4047)	(0.3657)	(0.3509)	(0.3629)	(0.3247)
Childhood Inbreeding Homophily	0.6693	0.3263	0.8346	0.7324	0.394	0.8441
	(0.4304)	(0.4479)	(0.3065)	(0.3674)	(0.3893)	(0.2816)
Homophily Index H	0.5318	0.4079	0.5915	0.5743	0.4271	0.6229
	(0.3038)	(0.2951)	(0.2895)	(0.2951)	(0.2957)	(0.2784)
All Home Language - Official?	0.5054	0.7841	0.3711	0.4213	0.703	0.3283
	(0.5)	(0.4116)	(0.4832)	(0.4939)	(0.4572)	(0.4697)
Most Often Home Language - Official?	0.6085	0.9093	0.4637	0.5314	0.8757	0.4177
	(0.4881)	(0.2873)	(0.4988)	(0.4991)	(0.3302)	(0.4933)
Linguistic Share Change: All Home Languages	0.3184	0.3442	0.306	0.3477	0.4064	0.3284
	(0.3099)	(0.299)	(0.3143)	(0.34)	(0.3187)	(0.3447)
Share Change: Most Often Home Languages	0.2213	0.3138	0.1767	0.2198	0.3589	0.1739
	(0.2953)	(0.2971)	(0.284)	(0.3215)	(0.3292)	(0.3054)
Transition of All Home Languages	0.4631	0.4856	0.4522	0.4683	0.5497	0.4414
	(0.4987)	(0.4999)	(0.4978)	(0.4991)	(0.4979)	(0.4967)
Transition of Most Often Home Languages	0.3131	0.4343	0.2547	0.2883	0.4793	0.2253
	(0.4638)	(0.4958)	(0.4358)	(0.4531)	(0.4999)	(0.4179)
Population Share by Ethnic Ancestry	0.1062	0.1645	0.0781	0.0998	0.126	0.0912
	(0.1538)	(0.1798)	(0.1306)	(0.1082)	(0.1289)	(0.0989)
Distance by First Language	0.4673	0.3956	0.5018	0.4942	0.4116	0.5215
	(0.1451)	(0.0913)	(0.1533)	(0.1545)	(0.1052)	(0.1585)
Panel B: Control Variables						
Visible Minority	0.4364	0.1796	0.5601	0.5578	0.2776	0.6503
	(0.496)	(0.384)	(0.4964)	(0.4967)	(0.4481)	(0.477)
Education	13.7723	13.9337	13.6946	13.7847	14.1533	13.663
	(2.9163)	(2.7424)	(2.9937)	(2.8958)	(2.6962)	(2.9492)
Education in Canada	0.4665	0.97	0.224	0.4169	0.9682	0.2348
	(0.4989)	(0.1707)	(0.417)	(0.4931)	(0.1755)	(0.424)
Age at Immigration below 5	0.1707	0.525	N/A	0.1145	0.4613	N/A
	(0.3763)	(0.4995)		(0.3185)	(0.4988)	

Continued on next page

[9] The full statement about the objective of this survey is the following. "First, the data will help us to better understand how people's backgrounds affect their participation in Canada's social, economic and cultural life. Secondly, the information that is gathered will help us to better understand how Canadians of different ethnic origins interpret and report their ethnicity. The information collected in the survey will also be used to inform policy and program development in the Department of Canadian Heritage."

Table 1 – continued from previous page

	All of Canada			Gateway Cities		
	Overall	Immigration Age		Overall	Immigration Age	
		0-14	15-64		0-14	15-64
Age at Immigration 5-14	0.1544	0.475	N/A	0.1337	0.5387	N/A
	(0.3614)	(0.4995)		(0.3404)	(0.4988)	
Age at Immigration 15-24	0.2828	N/A	0.4191	0.2993	N/A	0.3981
	(0.4504)		(0.4935)	(0.458)		(0.4896)
Age at Immigration 25-44	0.3575	N/A	0.5297	0.4093	N/A	0.5445
	(0.4793)		(0.4992)	(0.4918)		(0.4981)
Age at Immigration 45-64	0.0346	N/A	0.0513	0.0432	N/A	0.0575
	(0.1828)		(0.2205)	(0.2033)		(0.2328)
Age	45.3978	43.3626	46.3782	44.6466	41.9033	45.5522
	(10.2091)	(10.2585)	(10.0406)	(10.0074)	(9.8295)	(9.9021)
Gender	0.5292	0.5131	0.5369	0.5427	0.5318	0.5463
	(0.4992)	(0.5)	(0.4987)	(0.4983)	(0.4993)	(0.498)
Married	0.7284	0.6446	0.7688	0.7192	0.6188	0.7524
	(0.4448)	(0.4788)	(0.4217)	(0.4495)	(0.486)	(0.4317)
Panel C: Labor Market Outcomes						
Employment Status	0.7586	0.8071	0.735	0.7714	0.8492	0.7455
	(0.428)	(0.3947)	(0.4414)	(0.42)	(0.3581)	(0.4357)
Log Hourly Wage	2.9964	3.099	2.9372	3.0005	3.1586	2.9362
	(0.6246)	(0.5535)	(0.655)	(0.6188)	(0.5463)	(0.635)
Occupation Status	3.3906	3.1731	3.4953	3.3459	3.007	3.4575
	(1.2764)	(1.2672)	(1.2678)	(1.2621)	(1.2163)	(1.2573)
Language at Work - Official?	0.8864	0.9759	0.8386	0.8563	0.9681	0.8139
	(0.3173)	(0.1534)	(0.368)	(0.3509)	(0.176)	(0.3893)
Panel D: Social Outcomes						
Retain Ethnic Customs Rated 4+	0.4584	0.3498	0.5109	0.4962	0.413	0.5238
	(0.4983)	(0.4771)	(0.5)	(0.5001)	(0.4927)	(0.4995)
Retain Ethnic Customs Rated 5	0.2984	0.2046	0.3437	0.3286	0.2472	0.3556
	(0.4576)	(0.4036)	(0.475)	(0.4698)	(0.4317)	(0.4788)
Sense of Belonging - Ethnicity	3.6	3.2927	3.7497	3.7429	3.4274	3.8479
	(1.3046)	(1.3555)	(1.2522)	(1.2315)	(1.3049)	(1.188)
Sense of Belonging - Canada	4.4369	4.4307	4.4399	4.4197	4.3634	4.4384
	(0.8817)	(0.9042)	(0.8708)	(0.8862)	(0.9551)	(0.8615)
Sense of Belonging - Family	4.7254	4.6977	4.7388	4.7337	4.7196	4.7384
	(0.6961)	(0.7274)	(0.6801)	(0.6696)	(0.6899)	(0.6628)
Sense of Belonging - City	3.8041	3.5562	3.925	3.8731	3.5853	3.9694
	(1.146)	(1.187)	(1.1057)	(1.1199)	(1.1943)	(1.0771)
Sense of Belonging - Province	3.9027	3.6989	4.0021	3.9632	3.7078	4.0486
	(1.14)	(1.2014)	(1.0953)	(1.1071)	(1.1972)	(1.0619)
Trust - General	0.5395	0.5856	0.5167	0.5062	0.5415	0.4943
	(0.4985)	(0.4928)	(0.4998)	(0.5001)	(0.4986)	(0.5001)
Trust - Family	4.8015	4.7952	4.8045	4.7962	4.7787	4.802
	(0.527)	(0.5392)	(0.5211)	(0.541)	(0.5651)	(0.5328)
Trust - Community	3.7907	3.8157	3.7784	3.6705	3.6606	3.6738
	(1.0051)	(0.9812)	(1.0167)	(1.014)	(0.9671)	(1.0296)
Trust - School and Work	3.8396	3.8936	3.8114	3.7716	3.805	3.7596
	(0.9598)	(0.9471)	(0.9653)	(0.9674)	(0.9748)	(0.9647)
Life Satisfaction	4.2409	4.2673	4.2282	4.1979	4.2476	4.1814
	(0.8541)	(0.8234)	(0.8683)	(0.8734)	(0.8023)	(0.8954)
Volunteered	0.2804	0.3365	0.2534	0.2442	0.2859	0.2304
	(0.4492)	(0.4726)	(0.435)	(0.4297)	(0.4522)	(0.4212)
Voted in Federal	0.8087	0.8258	0.7983	0.7965	0.8197	0.7869
	(0.3934)	(0.3794)	(0.4013)	(0.4027)	(0.3847)	(0.4096)
Voted in Provincial	0.7781	0.792	0.7697	0.7686	0.7873	0.7609
	(0.4156)	(0.406)	(0.4211)	(0.4218)	(0.4096)	(0.4267)
Voted in City	0.6403	0.6401	0.6403	0.6255	0.6144	0.6301
	(0.48)	(0.4801)	(0.48)	(0.4841)	(0.4871)	(0.4829)
Community Activity	0.4068	0.4468	0.3874	0.3796	0.4047	0.3714
	(0.4913)	(0.4973)	(0.4872)	(0.4854)	(0.4912)	(0.4833)

The target population of EDS consists of persons aged 15 and older living in private dwellings in Canada's ten provinces, including Canadian citizens, permanent residents, and non-permanent residents, but excluding aboriginal people and people in remote territories. The full-sample of EDS public use microdata contains 41,695 individuals, representing a target population of 23,092,243. I restrict the data to first-generation, working-age immigrants who age between 25 and 64. Because I analyze the

determinants of their social networks as well as their socio-economic outcomes, working age population are more relevant than children, young students, or retirees. Additionally, my identification strategy relies on variations in language origins and ages of entry into Canada. In this respect, restricting to the first generation is necessary. Furthermore, I require variables on friendship networks, language usage, and key control variables, such as age, age at immigration, etc., to be non-missing. These restrictions lead to a sample size to 4,915. Restricting to gateway cities, i.e. Toronto, Montreal, and Vancouver, further reduces the sample size to 2,917.

Table 1 reports the summary statistics.[10] The *Homophily Index* H_i is constructed from the question about friends' ethnic ancestry. I assign $\{1, 0.75, 0.5, 0.25, 0\}$ to those who report all, most, a half, a few, and none of their students belong to the same ethnic ancestry as theirs.[11] The *Population Share* w_i is constructed based on the full EDS data according to all reported ethnic ancestries.[12] The *Inbreeding Homophily Index* IH_i is calculated using the formula in Equation (2) and the Homophily Index H_i and the Population Share w_i reported above.

EDS does not measure proficiency in English or French directly. I instead rely on language usage at home in combination with one's first language. The former characterizes people's incremental investment in language. The latter describes her language endowment. The distance by first language variable is reported in Table 9, which in turn is based on a measurement strategy proposed by Chiswick and Miller (2005). Two types of reported home languages are used, namely all home languages and home languages most often spoken. Both can have multiple languages, but they are most likely single choices. First, a dummy variable about whether all (or most often) home languages are solely official languages is created. Second, I compare a person's home languages (all or most often) with his first language in terms of the potential pool of same-language speakers. The population measures are constructed from the full EDS data. Lastly, I focus on whether a person's home languages (all or most often) have transitioned from non-official to official languages. In total, I create 6 measures of language skills. The most preferred one is the linguistic share change because it also captures the effort of, or lack thereof, official language speakers in learning new languages. Additionally, these two measures indicate the expansion or contraction of the pool of potential friends to be matched with, whose role in influencing friendship homophily is emphasized by Currarini et al. (2009).

The variables that affect the formation of friendships are listed in the table. The variables *Age, Sex,* and *Marital Status* are self-explanatory. A female individual is assigned 1, while a male is assigned zero. A married person is assigned one, and zero otherwise. *Education* is constructed from the highest level

[10] More details about all the reported variables are explained in Table 8 in the Appendix.

[11] EDS has very detailed information about a person's ethnic ancestry. A person is allowed to report up to 8 ethnic ancestries and to rate those. In the public use microdata, it is reported by variables EAC1-EAC8. I use the highest rated ethnic ancestry EATC1 and EATC2 in identifying a person's ethnic ancestry. The composition of a person's networks is reported by the variable SNQ0201 and SNQ0202, which reports how many of a person's friends belong to the ethnic ancestry reported in EATC1 and EATC2. Details are explained in Table 8.

[12] For the full sample, population shares are calculated for the country as a whole. For the gateway cities sample, population shares are calculated for each CMA respectively.

of schooling variable in the data. If the person's highest academic credential is received in Canada, the dummy variable *Education in Canada* equals one, and zero otherwise. The *Visible Minority Status* indicates whether a person is a visible minority, defined by the Employment Equality Act in Canada as "persons, other than Aboriginal peoples, who are non-Caucasian in race or non-white in colour". The dummy variables about *Age at Immigration* are self-explanatory, too.

Some illustrations of the key relationships help detecting patterns in the data before we delve into the technical details of regressions. Figures 1 and 2 report the relationship between the inbreeding homophily index and two key variables, namely the the population share of an individual's ethnic ancestries and his language skills endowment. Their importance has been highlighted by our discussions in Section 2.2. Both figures are based on the full sample.

Figure 1 fails to reproduce the inverted-U shape relationship between friendship homophily and group shares shown by Currarini et al. (2009). Instead, it shows a convex and decreasing relationship between them. There are potentially two reasons for the different result. First, while Currarini et al. (2009) measure population shares at the schools, I measure them at the national level. The relationship between the two variables can be different at different scales. Second, many people have indicated multiple ethnic ancestries in the data. The population share is the sum of two highest rated ancestries for each individual. These shares are then averaged at ethnic group level. Therefore, these population shares do not share the simple interpretation of group shares, i.e. these population shares do not add up to one even if we have accounted for all groups. In any case, population shares are important determinants of homophily; and they must be accounted for in the regression analysis.

Figure 2 reports some preliminary evidence to support the discussions in Section 2.2. In other words, learning more languages make a person's social network more ethnically diverse. Two measures of language distance are reported: one based on first languages and one based on ethnic ancestries. There are clear positive correlation between language distance and inbreeding homophily index. In the next section, I analyze how individuals' language investments affect their friendship networks beyond what their language endowment predicts.

Many socio-economic outcomes are also reported in Table 1. These variables are included to assess the viability of multicuturalism. In general, five types of variables are included. The first collection of variables reports an individual's labor market outcomes, including employment status, hourly earnings, occupation status, and language at workplace. They are important because they characterize the economic benefits for immigrants to make costly investment in language learning and social networking. These variables have been analyzed extensively in the economics literature, see for example Bleakley and Chin (2004).

The second group describes a person's attachment to his ethnic ancestry. These are crucial in assessing whether ethnic identity is preserved or diluted by language investment and friendship homophily.

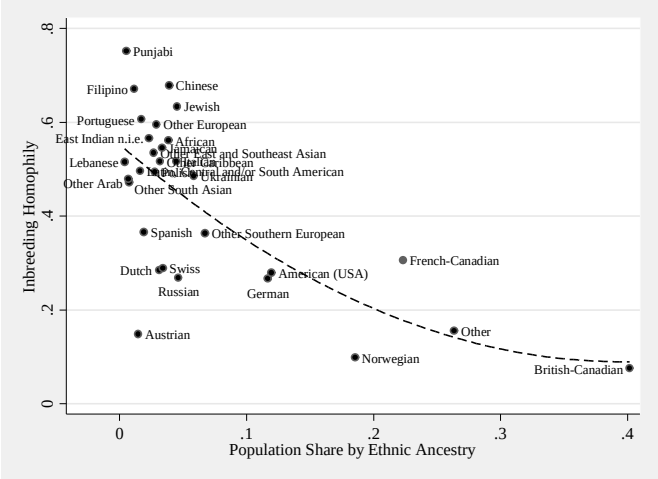

Figure 1: Inbreeding Homophily Index and Ethnic Ancestry Population Share

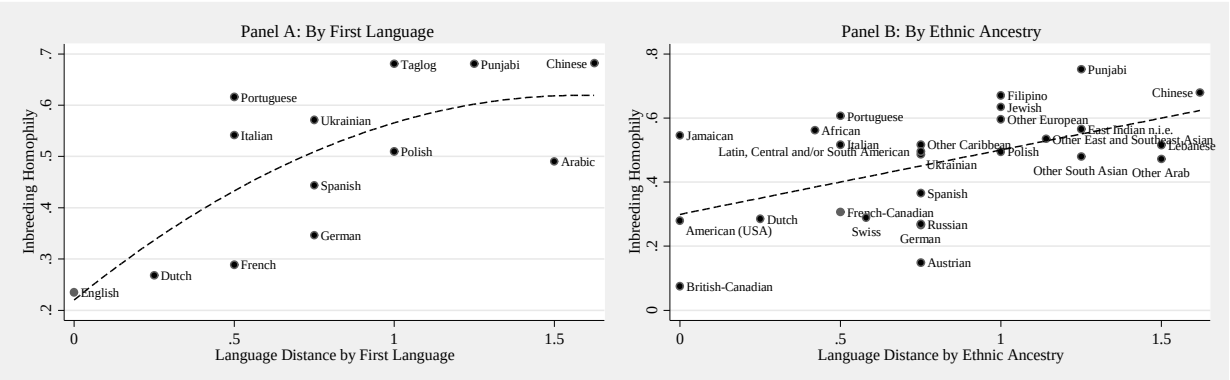

Figure 2: Inbreeding Homophily Index and Language Distance

Proponents of multiculturalism claim that immigrants' confidence in their ethnic identity is crucial for the success of multiculturalism. Additionally, preservation of ethnic identities of immigrants may itself be valuable to the host country in that immigrants can act as a bridge between the host country and their countries of origin.

The next group reports a person's attachment to the host country. They are sense of belonging to Canada, to one's family, to the province of residence, and to the city. Such indicators are crucial in assessing whether the objective of preserving diverse ethnic identities compromises another objective of enhancing a strong sense of belonging to the host country for all. Critics of multiculturalism often raise this issue. They argue that too much emphasis on ethnic labels undermines the coherence of social and political values of the host country.

The fourth group characterizes a person's well-being in terms his interaction with the society. The variables are general trust of others, trust of the people of one's family, of the same community, or of the workplace or school, and general satisfaction with one's life. Similar to labor market outcomes, these indicators describe individuals' private wellbeing. Moreover, they also indicate the well-being of the society as a whole, i.e. they are also measures of social capital.

The last group mainly describes a perons' civic participation, including voting in federal, provincial, and local elections besides participation in volunteering and community activities. Again, these indicators are related to criticisms of multiculturalism that social and spatial isolation of certain ethnic groups hinders their participation in the political process and the civil society generally. In sum, the long list of socio-economic outcomes paint a well-rounded picture of the many objectives of multiculturalism.

3.2 Empirical Strategy: Language and Friendship Homophily

I aim to establish a causal link between language skills and friendship homophily. They represent, as I argue earlier, the two pillars of Canadian multiculturalism. The link between the two can be understood through the model by Currarini et al. (2009). More specifically, learning more languages affect two key parameters of their model, i.e. the relative value of befriending a coethnic person and the probability to encounter a coethnic person. Generally speaking, learning more languages reduces the tendency to form homophilous friendship networks. Currarini et al. (2009) also highlighted the importance of group shares in affecting the matching of friends. Building upon their model, I estimate the following equation

$$IH_{ieg} = \beta_0 + \beta_1 HL_i + \beta_2 LD_g + \beta_3 ES_e + \mathbf{X}_{eg}^{'}\beta_4 + IM_i^{'}\beta_5 + \mathbf{W}_i^{'}\beta_6 + \varepsilon_{ieg}. \tag{3}$$

The dependent variable in Equation (3) is the inbreeding homophily index IH_{ieg} for individual i who belongs to ethnic group e and language group g according to her first language. The key variables of interest are measures of language skills. Unfortunately, the data do not report proficiency in English or

French.[13] Instead, I include language usage at home HL_i in combination with a measure about one's first language LD_g. More specifically, LD_g is the distance between a person's first language and English.[14] It describes a person's language endowment as compared to the majority language. We can think of language distance LD_g as a predictor for a "baseline homophily index".

Language usage at home HL_i, instead, characterizes an individual's incremental investment in language. As mentioned previously, there are six measures available. They are whether all (or most often spoken) languages at home are official languages, whether all (or most often spoken) languages at home have transitioned from non-official to official languages, and the linguisitic share difference between an individual's first languages and all his (or his most often spoken) home languages. Once LD_g is controlled for, HL_i gauges the effort of immigrants to learn and use the official languages. This is exactly what the believers and criticizers of multiculturalism are most concerned about. Moreover, the two measurements of linguistic share changes are preferred because they also capture the effort of, or lack thereof, official language speakers in retaining their bilingual endowment or learning new languages. In sum, I am most interested the coefficient β_1, an estimate of the net effect of language investment over and above the personal endowment of language skills.

To achieve a better prediction of the so-called "baseline homohily index", I also include the population share of a person's ethnic ancestries ES_e in light of Currarini et al. (2009), who finds that the potential pool of coethnic population significantly affects the homophily index.[15] To account for nonlinearity, I include all quadratic terms of both LD_g and ES_e, denoted as $\mathbf{X}_{eg}$. Because the impact of population size may affect friendship formation differently for different language groups, the flexible function form reduces the possibility of misspecification in Equation (3); and hence it provides a more reliable estimate of the "baseline homophily".[16]

The key challenge in identifying β_1 is the fact that HL_i is correlated with unobserved cost/benefits in learning the official languages. These costs can be tuition fees to enroll in a language class. They can be opportunity costs, i.e. the alternative payoff from not learning the official language. They can also be influenced by a person's unobservable "ability". In addition, growing up in a segregated neighborhood or studying in a segregated school may hinder both a person's language ability and his ability to form friendships with those of a different background.

Presumably, high ability individuals have better communication skills, so they tend to have a more diverse network. If high ability individuals are more effective in learning new languages as well, the OLS

[13] Many censuses ask questions about how well a person can speak the official language in the host country. Such measures themselves are not immune to problems as argued by Dustmann and van Soest (2001) and Dustmann and Fabbri (2003).

[14] Details are shown in Table 9 in the Appendix. This option is adopted in all reported tables. An alternative measure is based on the self-reported ethnic ancestry, described in detail in Table 10 in the Appendix. The results associated with this option are discussed briefly in the robustness section.

[15] We do not want to measure this population share at fine geographic units, e.g. census tracts, because neighborhood sorting makes these population measures endogenous. In this paper, I measure population shares both at the national level and at the CMA level.

[16] As a robustness check, I present results for an alternative specification that includes fixed effects of first language groups and their interactions with the ancestry shares in place of LD_g and $\mathbf{X}_{eg}$ in Table 2.

model will overestimate the negative coefficient β_1. If the opportunity cost in language learning is higher for high ability individuals, the OLS model will underestimate the negative coefficient β_1. On the other hand, spatial segregation tends to create negative bias in OLS estimates.

I address the concerns with the endogeneity of HL_i following the strategy of Bleakley and Chin (2004, 2010). Firstly, I add the age at immigration dummy variables IM_i in the regression. Age at immigration dummies are included to address the differentials in constraints faced by immigrants that arrive at different ages. They may not be valid instruments for language skills because they affect other aspects of assimilation, e.g. learning the host culture, the institutions, the education system, etc. In addition, I include a person's age, gender, visible minority status, marital status, education, achievement of highest degree in Canada or not, and birthplace dummies in the regression. These variables are denoted by the vector $\mathbf{W}_i$. Lastly, my main results are based on immigrants who arrived before age 25 to address concerns about the self-selection of immigrants by their ability in English or French above age 25. Age 25 is chosen because Canadian Immigration and Citizenship allows immigrants to bring along their children under the age 22 until August 2014.

Secondly, I construct the following instrumental variable for HL_i. It is the interaction between LD_g and $IMAGE_i$, which is defined as below:

$$IMAGE_i = \max\{0, AgeIM_i - a_i\}, \text{ where } a_i \in \{5, 15\},$$

where $AgeIM_i$ is the self-reported age at immigration and a_i are the cutoff values identifying the "critical period" of language learning. Clearly, $IMAGE_i$ equals 0 if a person immigrated before age $a_i = 5$ or $a_i = 15$. I experiment with the two cutoff values to check robustness of my results. Another reason is the fact that the age at immigration is grouped as 0-4, 5-14, 15-24, 25-44, and 45-64. According to Bleakley and Chin (2004, 2010), the critical period of language learning lasts until age 11. Therefore, experimenting with both age 5 and age 15 seems necessary.

The following equation specifies the first-stage regressions explicitly.

$$HL_{ieg} = \delta_0 + \delta_1 IM_i' + \delta_2 LD_g \times IMAGE_i + \delta_3 LD_g + \delta_4 ES_e + \mathbf{X}_{eg}^{'}\delta_5 + \mathbf{W}_i^{'}\delta_6 + \zeta_{ieg}. \tag{4}$$

People who immigrate at a young age is more likely to invest in learning official languages than those who immigrate later. However, those whose first languages are English can communicate well in English no matter when they enter Canada. On the other hand, immigrating earlier is critical in improving a person's English skills for those whose first language is linguistically farther away from English. Thus, we expect δ_2 to be negative and significant. This interaction term, however, does not enter Equation (3). Such exclusion is justifiable if other aspects of assimilation, such as cultural, educational, and institutional adjustments, do not respond to age at arrival differently across language groups.

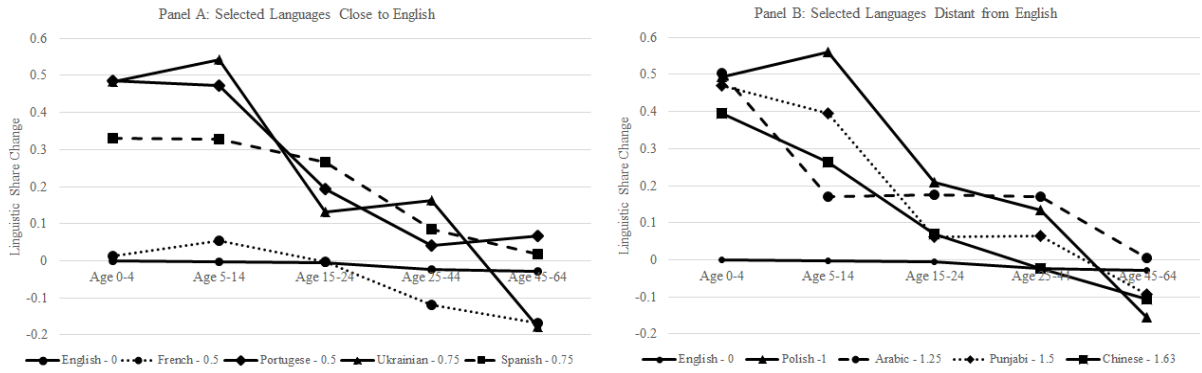

Figure 3: Linguistic Share Change and Age at Immigration by Language Groups

Figure 3 illustrates the key relationship that enables the identification of Equation (3). It illustrates the language skills as a function of age at immigration and first language group. Panel A reports several languages with language distance below 1, while Panel B reports languages more distant from English. Language skills are measured as the linguistic population share changes from a person's first language to his current home language most often spoken. This measure proxies for a person's investment in learning new languages, especially the two official languages as they are the most populous language groups. In the legend, the language distance between English and the respective languages are also shown.

Both panels show that immigrants who arrive during early childhood do increase their linguistic shares. Different language groups differ in their linguistic share changes. As expected, immigrants of the English and French groups do not experience significant drop in language investments at the threshold age of 5 or 15. The linguistic share changes also remain stable throughout all age categories for the two groups. The minority groups, especially the groups with larger language distances from English, experience significant drops in language invesments if they arrive later than age 15 (or age 5). This suggests that the interaction between LD_g and $IMAGE_i$ can be utilized as the instrumental variable for language investment, denoted as HL_{ieg}.

Figure 4 illustrates the reduced form relationships between IH_{ieg} and age at immigration by language groups. For most minority language groups, the inbreeding homophily index increases steadily as age at immigration increases. In particular, at threshold ages of 5 or 15, the jumps in homophily index are more conspicuous. For the English and French groups, no jumps in IH_{ieg} are detected at threshold age 5. Although there are jumps at threshold age 15, such jumps are not as big as those of minority groups. This finding gives another reason for utilizing both age 5 and age 15 as the thresholds in the first stage.

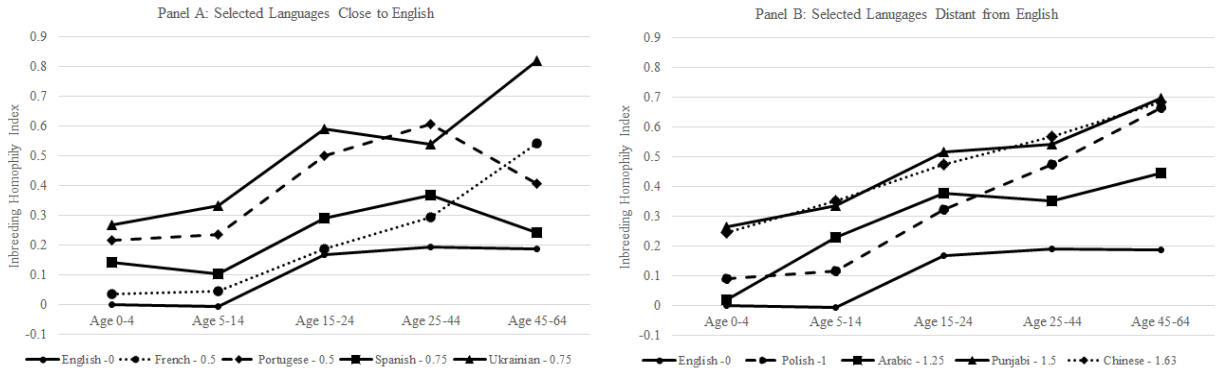

Figure 4: Inbreeding Homophily and Age at Immigration by Language Groups

3.3 Empirical Strategy: Language, Friendship Homophily, and Well-being of Immigrants

My second objective is to assess key performance indicators of multiculturalism. These indicators include labor market outcomes, attachment to one's ethnic ancestry, sense of belonging to the host country, general social well-being, and civic participation. In this section, I relate language skills and friendship homophily to those performance indicators. As discussed previously, the measures of language skills and friendship homophily correspond respectively to the two pillars of Canadian multiculturalism: the official language policy and the prescribed inter-cultural respect and understanding among all ethnicities. The socio-economic indicators, on the other hand, describe the many objectives of muticulturalism. Therefore, my analysis linking the two offers an assessement of the viability of multiculturalism in Canada.

In exploring the the relationship between language skills and socioeconomic outcomes, the estimation equation is similar to Equation (3) except for the change of the dependent variable:

$$y_{ieg} = \alpha_0 + \alpha_1 HL_i + \alpha_2 LD_g + \alpha_3 ES_e + \mathbf{X}_{eg}^{'}\alpha_4 + IM_i^{'}\alpha_5 + \mathbf{W}_i^{'}\alpha_6 + u_{ieg}, \tag{5}$$

where y_{ieg} denotes one of the many socio-economic outcomes. Even though some outcome variables are binary outcomes, I continue to employ a linear specification so that the coefficient estimates have natural interpretations. Again, the challenge in estimating the marginal effect β_1 lies in the endogeneity issue about HL_i. The same solution as that of the previous section is used. 2SLS regressions are estimated using the interaction between $IMAGE_i$ and LD_g as the instrument in a first stage regression expressed in Equation (4). The identifying assumptions are also similar as earlier.

The estimation equation for the relationship between friendship homophily and socio-economic outcomes consists of similar independent variables as Equation (5) except for the new variable of interest IH_{ieg}:

$$y_{ieg} = \gamma_0 + \gamma_1 IH_{ieg} + \gamma_2 LD_g + \gamma_3 ES_e + \mathbf{X}_{eg}^{'}\gamma_4 + IM_i^{\prime}\gamma_5 + \mathbf{W}_i^{'}\gamma_6 + \nu_{ieg}. \tag{6}$$

We are interested in the marginal effect of inbreeding homophily index or IH_{ieg} as it reflects the extent an immigrant experiences inter-cultural contacts. The key challenge in identifying γ_1 is the fact that IH_{ieg} is correlated with unobserved cost/benefits in friendship formation represented by the error term ν_{ieg}. For example, ν_{ieg} may be the unobserved ability to find new friends. It therefore is negatively correlated with IH_{ieg}, but it may be positively related with productivity, trust, and other variables. Additionally, the neighborhood and educational environment in which a person grows up may affect both y_{ieg} and IH_{ieg} in complex ways.

To address this endogeneity issue, I exploit a natural experiment offered by the immigration process itself. Migrating to a new country, especially for children and young adults, disrupts their social networks significantly. As friendship, once formed, needs to be maintained, some of an immigrant's old friendships in the home country may be lost after immigration. On the other hand, some friendships may still be kept, thus leading to a persistent relationship between childhood and adulthood frienship networks. Though friendship networks during adulthood may be affected by attributes of the host country, childhood friendships are not. Therefore, we could exclude all childhood immigrants, i.e. those who arrived below age 15, to sever the link between host country attributes and the two homophily indices.

This strategy becomes even more credible for those young immigrants (below age 25) who migrated along with their parents. The destination country is chosen by their parents, so their childhood frienships are not likely related to their unobserved skills to form friendships in the new country or unobserved characteristics of the environment experienced by them.

The first stage regression is

$$\begin{aligned} IH_{ieg} &= \phi_0 + \phi_1 CIH_{ieg} + \phi_2 CIH_{ieg} \times LD_g + \phi_3 LD_g + \phi_4 ES_e \\ &\quad + \mathbf{X}_{eg}^{'} \phi_5 + IM_i' \phi_6 + \mathbf{W}_i^{'} \phi_7 + \xi_{ieg}. \end{aligned} \tag{7}$$

The identification comes from the exclusion of childhood homophily CIH_{ieg} (before age 15) and its interaction with LD_g from Equation (6). I restrict the sample to those who immigrated after age 15 so that their childhood homophily CIH_{ieg} is the one in the home country. In one specification, I restrict the sample further to those who migrated from age 15 to age 24, i.e. young immigrants only.

Figure 5 depicts the persistence of frienship homophily from childhood to adulthood. In other words, it shows the estimates of ϕ_1 for different language groups and different cohorts by age at arrival. Only age, sex, and education are included in a simpilified version of Equation (7) in obtaining these estimates. In most cases, the persisency parameter ϕ_1 drops significantly at age 15. This confirms our conjecture that immigration disrupts friendships. After age 15, the persisency parameter stablizes for most language groups. Additionally, the drops in ϕ_1 at age 15 differ across language groups. The English and French groups appear to experience the largest drops, followed by those groups distant from English, with the

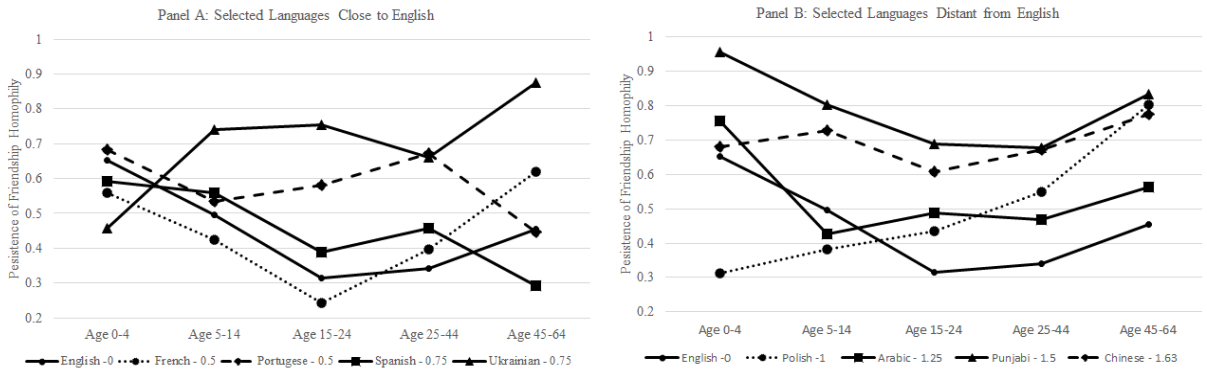

Figure 5: Persistence of Inbreeding Homophily by Age at Immigration and Language Group

middle-range groups experiencing small or no drops. This graph provides support for the utilization of CIH_{ieg} and its interaction with LD_g as instrumental variables.

4 Empirical Results

4.1 Language and Friendship Homophily

Table 2 reports regression results for Equation (3). In particular, language investment HL_{ieg} is measured as the linguistic share difference between a person's first language and her home language most often spoken. First, this measure captures variations in language investments even among official language speakers. For example, English speakers can still learn French and minority languages. On the other hand, people can lose the ability to speak a language if they were raised bilingual. Using this measure, we are testing a more general hypothesis, i.e. does learning an additional language or losing a mother tongue affect the ethnic composition of one's friends. Second, the measurement errors in the reported languages most often spoken are presumably less severe because the frequency of usage is more indicative of language ability.

Two sets of results are reported in Table 2: one for the full sample (OLS 1 and 2SLS 1) and the other for young immigrants only (OLS2, 2SLS 2 - 2SLS 5). For the full sample, a one standard deviation increase in the linguistic share change (about 0.2953) decreases the inbreeding homophily index by 0.05 according to OLS 1 and by 0.25 according to 2SLS1. The 2SLS results are less precisely estimated but much larger. Since the standard deviation of IH_{ieg} is about 0.4, the estimated coefficient of language investment is sizable. The effect of ethnic ancestry share on IH_{ieg} takes a U-shape with a minimum at 0.8 based on OLS 1 (0.65 based on 2SLS 1). Since most ancestry shares are below 0.6, the relationship is strictly decreasing. Because of the nonlinear terms, the marginal effect of LD_g is a bit complicated. It is, however, always positive given reasonal parameter choices. Other coefficients are also sensible. People who are visible minorities are more likely to develop homophilous friendship networks. Education overall

and in Canada in particular both decrease homophily. Age and sex have no detectable impact on IH_{ieg}, while married individuals seem to have more homophilous networks.

For young immigrants, the results are essentially similar to those of the full sample. The exclusion of adult immigrants above age 25 addresses concerns that the immigration process selects those who integrate more easily even though they come from a country linguistically more distant from English or French. Such selection through the immigration process or self-selection of immigrants lead to upward biases. This concern is confirmed by the fact that the full sample estimates before HL_{ieg} are much larger than the young sample. The estimates (OLS 2, 2SLS 2 - 5) also have smaller variances.

Different specifications are shown to check the robustness of the results. 2SLS 2 uses age 15 as the threshold age in the first stage. 2SLS 3 uses age 5 as the threshold age instead. In addition to all control variables in 2SLS 3, 2SLS 4 adds the place of birth dummies to control for differences in age at arrival distributions across countries. 2SLS 5 drops the quardratic function form assumption in predicting the "basline homophily". Instead, I add first language group dummies and their interactions with ethnic ancestry shares. The estimated coefficients before HL_{ieg} vary around 0.4-0.5 across all 2SLS specifications. Thus, one standard deviation increase in HL_{ieg} (about 0.3) lead to an increase of 0.12 - 0.15 in IH_{ieg}. Again, these estimates are sizable numbers. Panel B also reports the first stage results. The interaction term is negative and significant, as expected. The F-statistic and the Stock and Yogo (2005) test statistics do not support the existence of weak instruments.

Table 3 further tests the robustness of the previous findings. I restrict the analysis to young immigrants only, excluding individuals who immigrated after age 25. As explained earlier, the data offer 6 measures of language investments, all of which are based on information about one's home languages and/or first languages. Table 2 analyzes one particluar measure, namely the linguistic share change from one's first languages to his home languages most often spoken, while Table 3 analyzes all of them. Additionally, results for the full sample and immigrants in gateway cities, i.e. Toronto, Montreal, and Vancouver, are both shown. The two analyses differ by their measurement of population shares of a person's ethnic ancestry. For the full sample, population shares are measured nationally. For gateway cities, population shares are measured within each CMA. The control variables, however, remain the same across the two samples.

The results across the six panels are basically similar. They are also similar to those reported by Table 2. Those who speak official languages at home and those who expand their mastery of languages tend to have a more diverse social network. The scales of those estimates are different because their bases of construction differ from each other. All the numbers are also economically significant. The first stage results confirm my presumption that childhood immigrants are more likely to make language investments if they belong to a group linguistically distant from English. The F-stats and Stock and Yogo (2005) tests show no evidence of weak instruments.

Table 2: Linguistic Share Change and Inbreeding Homophily of Friendship Networks

	All Immigrants		Young Immigrants				
	OLS 1	2SLS 1	OLS 2	2SLS 2	2SLS 3	2SLS 4	2SLS 5
Panel A: Dependent Variable - Inbreeding Homophily Index							
Linguistic Share Change by Home	-0.1686^a	-0.8350^a	-0.1668^a	-0.3750^b	-0.4733^a	-0.3918^b	-0.5254^a
Languages Most Often Spoken	(0.0171)	(0.2081)	(0.0240)	(0.1659)	(0.1648)	(0.1642)	(0.1798)
Population Share of Ethnic Ancestry	-1.0005^a	-1.5051^a	-1.3203^a	-1.4763^a	-1.5499^a	-0.9056^a	-0.8262^b
	(0.2685)	(0.3145)	(0.3096)	(0.3299)	(0.3306)	(0.3273)	(0.3278)
Ethnic Share Squared	0.5787	1.1790^c	1.4787^b	1.6413^b	1.7181^b	1.0571	1.0982
	(0.6094)	(0.6358)	(0.6919)	(0.6962)	(0.6970)	(0.6889)	(0.6974)
Language Distance by First Language	0.2285^a	0.5988^a	0.2511^a	0.3901^a	0.4557^a	0.3461^a	
	(0.0345)	(0.1231)	(0.0505)	(0.1231)	(0.1239)	(0.1245)	
Language Distance Squared	-0.0864^a	-0.2938^a	-0.1029^a	-0.1797^a	-0.2160^a	-0.1281^c	
	(0.0199)	(0.0695)	(0.0300)	(0.0687)	(0.0692)	(0.0660)	
Language Distance*Ethnic Share	0.2256	0.2655	-0.0543	-0.0205	-0.0045	-0.2496	
	(0.2182)	(0.2267)	(0.2667)	(0.2662)	(0.2674)	(0.2717)	
Visible Minority	0.0800^a	0.0311	0.0728^a	0.0532^b	0.0439^c	0.0371	0.0707^b
	(0.0140)	(0.0221)	(0.0203)	(0.0251)	(0.0254)	(0.0319)	(0.0307)
Education	-0.0166^a	-0.0091^a	-0.0225^a	-0.0204^a	-0.0194^a	-0.0103^b	-0.0080^c
	(0.0020)	(0.0033)	(0.0040)	(0.0043)	(0.0044)	(0.0046)	(0.0048)
Highest Degree in Canada	-0.2542^a	-0.0794	-0.3304^a	-0.2832^a	-0.2609^a	-0.2042^a	-0.1729^b
	(0.0536)	(0.0810)	(0.0687)	(0.0775)	(0.0789)	(0.0762)	(0.0788)
Education*Degree in Canada	0.0134^a	0.0049	0.0208^a	0.0186^a	0.0176^a	0.0122^b	0.0109^b
	(0.0036)	(0.0048)	(0.0050)	(0.0053)	(0.0054)	(0.0052)	(0.0054)
Age	0.0001	0.0032^a	-0.0001	0.0009	0.0014	0.0011	0.0012
	(0.0006)	(0.0012)	(0.0008)	(0.0011)	(0.0011)	(0.0010)	(0.0010)
Sex	0.0190^b	0.0131	0.0094	0.0062	0.0047	0.0029	0.0020
	(0.0095)	(0.0107)	(0.0129)	(0.0132)	(0.0133)	(0.0128)	(0.0129)
Married	0.0302^a	0.0388^a	0.0375^b	0.0429^a	0.0454^a	0.0319^b	0.0286^c
	(0.0115)	(0.0129)	(0.0148)	(0.0154)	(0.0155)	(0.0152)	(0.0151)
Age at Immigration Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CMA Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Place of Birth Fixed Effects	No	No	No	No	No	Yes	Yes
First Language Fixed Effects	No	No	No	No	No	No	Yes
Language Dummies*Ancestry Share	No	No	No	No	No	No	Yes
Constant	0.6682^a	0.5511^a	0.7278^a	0.6943^a	0.6785^a	0.5359^a	0.4913^a
	(0.0460)	(0.0638)	(0.0679)	(0.0729)	(0.0740)	(0.0862)	(0.0884)
Observations	4915	4915	2988	2988	2988	2988	2988
R-Squared	0.3217	0.1513	0.2624	0.2477	0.2306	0.2908	0.2928
Panel B: First Stage Dependent Variable - Linguistic Share Change by Most Often Home Language							
Language Distance*Immigrate after		-0.0044^a		-0.0351^a	-0.0136^a	-0.0136^a	-0.0127^a
Threshold Age		(0.0006)		(0.0044)	(0.0016)	(0.0016)	(0.0016)
F-Stat		147.03^a		167.27^a	168.62^a	131.20^a	1519.66^a
Stock and Yogo (2005) Test		45.8917^a		75.3624^a	80.4343^a	83.2236^a	75.8467^a

Two sets of results are reported respectively for the full sample and young (below age 25) immigrants. OLS estimates are shown to compare with the more credible 2SLS results. The instruments for 2SLS 1 -2 are the interactions between whether a person immigrated after age 15 and a language distance measure based on his first language. The instruments for 2SLS 3 - 5 are the interactions between whether a person immigrated after age 5 and the same language distance measure. In Panel A, the dependent variable is inbreeding homophily index. The Appendix contains details about the construction of all included variables. All coefficient estimates are shown except for age at immigration, CMA, place of birth, first language dummies and interactions between language dummies and ethnic shares. Panel B reports the first stage statistics. Only the coefficient estimates of the instruments are shown. Tests of existence of weak instruments are also reported. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively, except for Stock and Yogo (2005) tests when ^{a}, ^{b}, and ^{c} represent rejection of presence of weak instruments at maximum rejection rates of 10%, 15%, and 20% respectively.

Table 3: Language Skills and Inbreeding Homophily of Young Immigrants: Other Measures and Samples

	All of Canada			Gateway Cities		
	2SLS 2	2SLS 3	2SLS 4	2SLS 2	2SLS 3	2SLS 4
Panel A: Linguistic Share Change by Most Often Home Languages						
Linguistic Share Change	-0.3750^{b}	-0.4733^{a}	-0.3918^{b}	-0.3989^{b}	-0.4544^{b}	-0.3849^{c}
	(0.1659)	(0.1648)	(0.1642)	(0.1990)	(0.1966)	(0.2190)
Observations	2988	2988	2988	1597	1597	1597
Panel A.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0351^{a}	-0.0136^{a}	-0.0136^{a}	-0.0351^{a}	-0.0137^{a}	-0.0124^{a}
	(0.0044)	(0.0016)	(0.0016)	(0.0059)	(0.0022)	(0.0023)
F-Stat	167.27^{a}	168.62^{a}	131.20^{a}	51.57^{a}	51.68^{a}	77.31^{a}
Stock and Yogo (2005) Test	75.3624^{a}	80.4343^{a}	83.2236^{a}	33.0684^{a}	34.9451^{a}	27.7822^{a}
Panel B: Linguistic Share Change by All Home Languages						
Linguistic Share Change	-0.5105^{b}	-0.6757^{a}	-0.5354^{b}	-0.5943^{c}	-0.7186^{b}	-0.6154
	(0.2365)	(0.2495)	(0.2338)	(0.3155)	(0.3369)	(0.3752)
Observations	2988	2988	2988	1597	1597	1597
Panel B.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0258^{a}	-0.0095^{a}	-0.0100^{a}	-0.0235^{a}	-0.0087^{a}	-0.0077^{a}
	(0.0037)	(0.0013)	(0.0013)	(0.0051)	(0.0019)	(0.0020)
F-Stat	358.09^{a}	359.78^{a}	272.64^{a}	112.74^{a}	112.99^{a}	123.14^{a}
Stock and Yogo (2005) Test	47.6135^{a}	46.1002^{a}	52.7975^{a}	16.9217^{a}	15.8458^{b}	12.3767^{b}
Panel C: Transition to Official Languages of Most Often Home Languages						
Language Transition Dummy: Official=1	-0.2812^{b}	-0.3546^{a}	-0.2918^{b}	-0.2776^{b}	-0.3170^{b}	-0.2493^{c}
	(0.1242)	(0.1240)	(0.1224)	(0.1367)	(0.1356)	(0.1401)
Observations	2988	2988	2988	1597	1597	1597
Panel C.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0469^{a}	-0.0182^{a}	-0.0183^{a}	-0.0504^{a}	-0.0196^{a}	-0.0191^{a}
	(0.0065)	(0.0023)	(0.0024)	(0.0082)	(0.0030)	(0.0031)
F-Stat	262.00^{a}	262.86^{a}	189.49^{a}	119.07^{a}	119.16^{a}	119.48^{a}
Stock and Yogo (2005) Test	61.1929^{a}	65.4289^{a}	68.7889^{a}	39.6198^{a}	41.6708^{a}	39.0302^{a}
Panel D: Transition to Official Languages of All Home Languages						
Language Transition Dummy: Official=1	-0.3936^{b}	-0.5259^{a}	-0.4192^{b}	-0.3793^{c}	-0.4443^{b}	-0.3422^{c}
	(0.1840)	(0.1979)	(0.1851)	(0.1969)	(0.2027)	(0.2007)
Observations	2988	2988	2988	1597	1597	1597
Panel D.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0335^{a}	-0.0122^{a}	-0.0128^{a}	-0.0369^{a}	-0.0140^{a}	-0.0139^{a}
	(0.0055)	(0.0019)	(0.0019)	(0.0070)	(0.0025)	(0.0026)
F-Stat	439.76^{a}	440.76^{a}	322.53^{a}	251.15^{a}	253.13^{a}	230.19^{a}
Stock and Yogo (2005) Test	35.0562^{a}	33.3228^{a}	37.614^{a}	24.1283^{a}	24.0798^{a}	23.4387^{a}
Panel E: Home Languages Most Often Spoken						
Language Dummy: Official = 1	-0.2093^{b}	-0.2550^{a}	-0.2303^{b}	-0.2327^{b}	-0.2569^{b}	-0.2149^{c}
	(0.0913)	(0.0866)	(0.0956)	(0.1136)	(0.1085)	(0.1208)
Observations	2988	2988	2988	1597	1597	1597
Panel E.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0630^{a}	-0.0252^{a}	-0.0232^{a}	-0.0601^{a}	-0.0242^{a}	-0.0222^{a}
	(0.0064)	(0.0023)	(0.0023)	(0.0083)	(0.0031)	(0.0032)
F-Stat	171.58^{a}	172.20^{a}	116.69^{a}	130.16^{a}	130.84^{a}	88.33^{a}
Stock and Yogo (2005) Test	137.287^{a}	158.026^{a}	135.943^{a}	64.5606^{a}	72.937^{a}	61.4054^{a}
Panel F: All Home Languages Spoken						
Language Dummy: Official = 1	-0.3153^{b}	-0.3460^{a}	-0.3133^{b}	-0.3766^{b}	-0.3850^{b}	-0.3477^{c}
	(0.1346)	(0.1146)	(0.1272)	(0.1803)	(0.1583)	(0.1918)
Observations	2988	2988	2988	1597	1597	1597
Panel F.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0418^{a}	-0.0186^{a}	-0.0171^{a}	-0.0372^{a}	-0.0162^{a}	-0.0137^{a}
	(0.0064)	(0.0025)	(0.0025)	(0.0081)	(0.0032)	(0.0032)
F-Stat	335.87^{a}	339.44^{a}	244.61^{a}	224.96^{a}	227.94^{a}	166.91^{a}
Stock and Yogo (2005) Test	52.612^{a}	74.7217^{a}	65.1416^{a}	22.8985^{a}	30.1434^{a}	22.5722^{a}

Results for young (below age 25) immigrants are reported respectively for the full sample and the sample of gateway cities. 2SLS 2 - 3 include the same set of control variables as those in Table 2. 2SLS 2 uses the interaction between whether a person immigrated after age 15 and a language distance measure as the instrument. 2SLS 3 - 4 use the interaction between whether a person immigrated after age 5 and the same language distance measure as instruments. Across all panels, the dependent variable is the inbreeding homophily index. Panel A through Panel E report results for different measurements of langugage skills. Additional details are contained in the Appendix. The first stage regressions and tests of existence of weak instruments are also reported for all panels. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively, except for Stock and Yogo (2005) tests when ^{a}, ^{b}, and ^{c} represent rejection of presence of weak instruments at maximum rejection rates of 10%, 15%, and 20% respectively.

4.2 Language and Well-being of Immigrants

Table 4 reports the labor market outcomes for young and all immigrants. The full sample (column 2 - 4) and the gateway cities (columns 5 - 7) are both shown. 2SLS 1' - 3' differ from 2SLS 1 - 3 of Table 2 in their inclusion of place of birth dummies. 2SLS 1' and 3' use the product of whether a person immigrated after age 15 and a language distance measure as instruments. 2SLS 4 uses the product of whether a person immigrated after age 5 and a language distance measure as the instrument. First stage results are also shown separately.

The relationships between language skills and labor market outcomes have been analyzed extensively in the literature. Most studies focus on earnings and employment status. This paper analyzes two additional outcome variables, namely occupation status and language at work. I follow Green (1999) in defining the occupational achievement of immigrants. They are: senior managers, professionals, white-collar, and blue-collar, assigning numbers from 1 to 4. The language at workplace describes the job nature and an indicator of labor market segement (Li, 2013) of a worker. This variable is binary, with 1 indicating English or French as the language at workplace and zero otherwise. The key independent variable measuring investment in language skills is the linguistic share change from one's first languages to his home languages most often spoken.

In terms of employment probability, the effects of language investments are ambiguous across all specifications. Language investments increase workers' earnings significantly. A one standard deviation increase in HL_{ieg} (about 0.3) increases hourly earnings by 30 - 40 percent (based on 2SLS4). With respect to occupation status, language investments are found to improve occupation prestige. However, the identification is not as credible as that for young immigrants. For young immigrants, no statistically significant effects of language investments are found. In terms of language at workplace, langugage investments make a person more likely to speak official languages at work. According to the estimates based on young immigrants, one standard deviation increase in HL_{ieg} increases the probability of using official languages by approximately 20 percent.

Table 5 expands the analysis to non-economic outcomes. These variables are equally important in assessing the successes and failures of multiculturalism. The first set of indicators describe immigrants' attachment to their ethnic heritage. They are the importance of carrying on heritages rated above 4 and above 5 and sense of belonging to own ethnic ancestry. All coefficients in Panel A are not statistically significant. It shows that more investments in language learning do not necessarily cause immigrants to lose their ethnic identity.

The second group of indicators report immigrants' attachment to the host country. Naturally, they are important performance indicators of multiculturalism. These indicators include their sense of belongings to Canada, their family, their city of residence, and their province. Again, all estimates are not statistically significant. This finding is somewhat surprising. Learning more languages does not

Table 4: Language Skills and Labor Market Outcomes

	All of Canada			Gateway Cities		
	All	Young		All	Young	
	2SLS 1'	2SLS 3'	2SLS 4	2SLS 1'	2SLS 3'	2SLS 4
Panel 1. Employment Probability						
Linguistic Share Change by Home Languages	0.0905	-0.3110	-0.2963	0.3449	-0.4091	-0.4337
	(0.2350)	(0.2056)	(0.2021)	(0.3228)	(0.2777)	(0.2806)
Observations	4735	2894	2894	2821	1547	1547
Panel 1.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0048^a	-0.0342^a	-0.0131^a	-0.0041^a	-0.0309^a	-0.0117^a
	(0.0006)	(0.0045)	(0.0016)	(0.0007)	(0.0062)	(0.0023)
F-Stat	117.83^a	128.89^a	129.47^a	77.22^a	76.00^a	74.75^a
Stock and Yogo (2005) Test	52.2533^a	70.9098^a	73.3898^a	19.6293^a	24.5894^a	23.837^a
Panel 2. Log Hourly Wage						
Linguistic Share Change by Home Languages	1.7790^a	1.1320^a	1.0590^a	1.0429^b	1.2677^a	1.2633^a
	(0.4376)	(0.343)	(0.3179)	(0.4969)	(0.44)	(0.4502)
Observations	2985	1889	1889	1769	1018	1018
Panel 2.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0051^a	-0.0358^a	-0.0134^a	-0.0047^a	-0.0348^a	-0.0125^a
	(0.0008)	(0.0057)	(0.0019)	(0.0010)	(0.0077)	(0.0028)
F-Stat	84.03^a	97.49^a	98.01^a	86.23^a	66.36^a	63.89^a
Stock and Yogo (2005) Test	37.0275^a	52.1195^a	53.8943^a	14.9543^b	21.062^a	18.8831^a
Panel 3. Occupation Status						
Linguistic Share Change by Home Languages	-1.9898^a	-0.0063	-0.0653	-2.4560^b	0.7425	0.9004
	(0.7517)	(0.6146)	(0.6048)	(1.1464)	(0.8685)	(0.8684)
Observations	4836	2944	2944	2865	1570	1570
Panel 3.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0047^a	-0.0355^a	-0.0135^a	-0.0039^a	-0.0314^a	-0.0119^a
	(0.0006)	(0.0045)	(0.0016)	(0.0007)	(0.0061)	(0.0023)
F-Stat	117.41^a	128.39^a	129.09^a	78.45^a	75.43^a	74.14^a
Stock and Yogo (2005) Test	52.949^a	78.1923^a	80.7124^a	18.0014^a	25.5902^a	25.0034^a
Panel 4. Language at Workplace						
Linguistic Share Change by Home Languages	1.1824^a	0.5842^a	0.5328^a	1.5111^a	0.6810^a	0.6565^a
	(0.2236)	(0.1459)	(0.1393)	(0.3693)	(0.2262)	(0.2471)
Observations	3575	2232	2232	2164	1232	1232
Panel 4.1: First Stage Regression						
Language Distance*Immigrate after Threshold	-0.0048^a	-0.0323^a	-0.0124^a	-0.0042^a	-0.0301^a	-0.0108^a
	(0.0007)	(0.0052)	(0.0018)	(0.0009)	(0.0070)	(0.0026)
F-Stat	93.20^a	106.69^a	107.25^a	130.94^a	101.26^a	94.83^a
Stock and Yogo (2005) Test	39.2962^a	51.0792^a	53.3383^a	15.048^b	19.3146^a	16.949^a

Panel 1 to Panel 4 report the regression results for four labor market outcomes. Language skills are measured by linguistic share change from a person's first language to the home language most often spoken. Results are reported respectively for the full sample and gateway cities and for young (below age 25) and all immigrants. 2SLS 1' - 3' differ from 2SLS 1 - 3 of Table 2 in their inclusion of place of birth dummies. 2SLS 1' and 3' use the product of whether a person immigrated after age 15 and a language distance measure as instruments. 2SLS 4 uses the product of whether a person immigrated after age 5 and a language distance measure as the instrument. The first stage regressions and tests of existence of weak instruments are also reported for all panels. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively, except for Stock and Yogo (2005) tests when ^{a}, ^{b}, and ^{c} represent rejection of presence of weak instruments at maximum rejection rates of 10%, 15%, and 20% respectively.

necessarily improve immigrants' sense of belonging to the host country.

The next set of indicators measure individuals' general social well-being. They gauge people's wellbeing in terms of their interactions with the society. Among the four trust variables and general satisfaction about their lives, some statistically significant coefficients are found. For example, trust of coworkers and classmates is found to be increased by language investments. Generally speaking, learning more languages can help building social capital. Lastly, civic participation is measured by volunteering activity, participation in community activity, and voting in federal, city, and provincial elections. The results are mixed, though some estimated coefficients are positive and significant. This finding is again surprising. It seems that learning the official languages does not necessarily improve civic participation, contradicting one of the key assumptions of Canadian multiculturalism.

Table 5: Language Skills and Social Outcomes

	All of Canada			Gateway Cities		
	All	Young		All	Young	
	2SLS 1'	2SLS 3'	2SLS 4	2SLS 1'	2SLS 3'	2SLS 4
Panel A: Importance of Ethnic Ancestry						
1. Retain Customs - At Least One Ethnic Ancestry Rated Above 4						
Linguistic Share Change by Home Language	0.1374	-0.0042	-0.0326	-0.0064	0.0923	0.1547
	(0.2673)	(0.2309)	(0.2279)	(0.3945)	(0.3252)	(0.3290)
Observations	4900	2982	2982	2906	1593	1593
2. Retain Customs - At Least One Ethnic Ancestry Rated Above 5						
Linguistic Share Change by Home Language	-0.1305	-0.2210	-0.2945	-0.4926	-0.0507	-0.0411
	(0.2363)	(0.2040)	(0.1976)	(0.3599)	(0.2830)	(0.2758)
Observations	4900	2982	2982	2906	1593	1593
3. Sense of Belonging to Ethnic Ancestry						
Linguistic Share Change by Home Language	0.2737	0.1808	0.1218	0.5061	0.3544	0.3112
	(0.6808)	(0.6058)	(0.6022)	(0.9604)	(0.8302)	(0.8549)
Observations	4807	2928	2928	2839	1559	1559
Panel B: Sense of Belonging						
4. Sense of Belonging to Canada						
Linguistic Share Change by Home Language	-0.5424	-0.3879	-0.3545	-0.1490	-0.2776	-0.2489
	(0.5050)	(0.4126)	(0.4265)	(0.7059)	(0.5882)	(0.6152)
Observations	4875	2966	2966	2888	1582	1582
5. Sense of Belonging to Family						
Linguistic Share Change by Home Language	0.0204	0.2646	0.4440	0.1714	-0.4001	-0.2024
	(0.4321)	(0.3532)	(0.3353)	(0.5650)	(0.5031)	(0.4846)
Observations	4887	2977	2977	2895	1590	1590
6. Sense of Belonging to City						
Linguistic Share Change by Home Language	-0.7974	-0.2958	-0.2160	-1.3808	-0.1058	0.1032
	(0.6242)	(0.5295)	(0.5310)	(0.9215)	(0.7349)	(0.7672)
Observations	4860	2963	2963	2877	1579	1579
7. Sense of Belonging to Province						
Linguistic Share Change by Home Language	-0.7901	-0.4314	-0.4686	-0.6092	-0.1312	-0.1174
	(0.6262)	(0.5286)	(0.5386)	(0.8790)	(0.7225)	(0.7686)
Observations	4860	2957	2957	2883	1581	1581

Continued on next page

Table 5 – continued from previous page

	All of Canada			Gateway Cities		
	All	Young		All	Young	
	2SLS 1'	2SLS 3'	2SLS 4	2SLS 1'	2SLS 3'	2SLS 4
Panel C: Social Well-being						
8. General Trust						
Linguistic Share Change by Home Language	0.0458	0.3131	0.2915	0.2096	0.2742	0.1979
	(0.2814)	(0.2382)	(0.2318)	(0.4067)	(0.3427)	(0.3387)
Observations	4775	2927	2927	2817	1553	1553
9. Trust Family Members						
Linguistic Share Change by Home Language	0.2495	0.4818^{b}	0.4873^{b}	0.7278^{c}	-0.0010	0.0363
	(0.2755)	(0.2336)	(0.2274)	(0.4200)	(0.3637)	(0.3556)
Observations	4891	2976	2976	2900	1589	1589
10. Trust Community						
Linguistic Share Change by Home Language	0.5493	1.3204^{a}	1.0636^{b}	0.5076	0.8237	0.5139
	(0.5506)	(0.4600)	(0.4498)	(0.8304)	(0.6411)	(0.6411)
Observations	4774	2930	2930	2816	1557	1557
11. Trust School and Workplace						
Linguistic Share Change by Home Language	0.7731	1.1869^{b}	1.0193^{b}	1.0147	1.3237^{b}	1.1918^{c}
	(0.6075)	(0.5087)	(0.4922)	(0.9313)	(0.6735)	(0.6632)
Observations	4301	2675	2675	2579	1456	1456
12. Satisfaction						
Linguistic Share Change by Home Language	0.9372^{c}	0.3008	0.3987	0.8595	0.0283	0.0286
	(0.4931)	(0.3914)	(0.3801)	(0.7708)	(0.5687)	(0.5589)
Observations	4890	2978	2978	2901	1593	1593
Panel D: Civic Participation						
13. Volunteer Activity						
Linguistic Share Change by Home Language	0.3545	0.3560	0.3930^{c}	0.4451	0.2864	0.3210
	(0.2620)	(0.2281)	(0.2276)	(0.3771)	(0.3067)	(0.3114)
Observations	4911	2985	2985	2916	1597	1597
14. Voted in Federal Elections						
Linguistic Share Change by Home Language	0.9065^{a}	-0.1785	-0.0728	0.8957^{c}	0.0526	0.1145
	(0.3212)	(0.2417)	(0.2303)	(0.4767)	(0.3579)	(0.3576)
Observations	3811	2551	2551	2221	1353	1353
15. Voted in Provincial Elections						
Linguistic Share Change by Home Language	0.8518^{b}	-0.2742	-0.1586	1.0340^{c}	0.0466	0.1196
	(0.3419)	(0.2592)	(0.2502)	(0.5403)	(0.3846)	(0.3862)
Observations	3781	2526	2526	2204	1337	1337
16. Voted in City Elections						
Linguistic Share Change by Home Language	0.1510	-0.5404^{b}	-0.5215^{b}	0.0382	-0.2884	-0.3273
	(0.3593)	(0.2703)	(0.2563)	(0.5265)	(0.3751)	(0.3732)
Observations	3761	2519	2519	2187	1328	1328
17. Community Activity						
Linguistic Share Change by Home Language	0.1146	0.1369	0.1283	0.2981	0.0969	0.0002
	(0.2750)	(0.2323)	(0.2305)	(0.3989)	(0.3197)	(0.3208)
Observations	4912	2987	2987	2916	1597	1597

Panels A - D report the regression results for four sets of social outcomes related to the objectives of multiculturalism. Language skills are measured by linguistic share change from a person's first language to the home language most often spoken. Results are reported respectively for the full sample and gateway cities and for young (below age 25) and all immigrants. 2SLS 1' - 3' differ from 2SLS 1 - 3 of Table 2 in their inclusion of place of birth dummies. 2SLS 1' and 3' use the product of whether a person immigrated after age 15 and a language distance measure as instruments. 2SLS 4 uses the product of whether a person immigrated after age 5 and a language distance measure as the instrument. No first stage results are reported to save space. All specifications reject the presence of weak instruments. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively.

4.3 Friendship Homophily and Well-being of Immigrants

Table 6 reports how inbreeding homophily affects labor market outcomes. Results are reported respectively for the full sample and gateway cities and for young (below age 25) and all immigrants. 2SLS 6 - 7 use childhood (before age 15) homophily index as the instrument. 2SLS 8 uses both childhood homophily and its interaction with language distance as instruments. First stage results are also reported.

Again, the effects of social networks on employment probability and earnings have been analyzed extensively. However, most papers focus on the size of a person's network rather than its composition. The results in Table 4 are new because they test the strength of weak ties in helping immigrants to improve their labor market outcomes. In addition, it would be very interesting to link friendship homophily, a measure of social integration, to economic well-being of immigrants.

First, although the homophily index seems to affect employment negatively across all specifications, the estimated coefficients are not statistically significant. The results on earnings are also not significant. In terms of occupation status, people with more homogeneous social networks are less likely to get more prestigious jobs, but the effect is not statistically significant. Lastly, if a person's network is homogeneous, he is less likely to speak official languages at work. This finding is robust in all specifications.

In summary, I find some evidence that the lack of integration or that of language skills do adversely affect an individual's labor market outcomes. Though not all of these estimates are significant, they are of the expected signs. There are possibly a lot of heterogeneity among immigrants' skills beyond our measured variables. More in-depth analysis with better measurement of those individual characteristics is definitely required to further test the effect of friendship homophily.

Table 7 expands the analysis to non-economic outcomes. Instead of analyzing the determinants of a person's social networks, we analyze how social networks affect these outcomes. The ethnic diversity of an immigrant's friendship networks is both a performance indicator of and a pillar that supports multiculturalism. This point has been made by both proponents and critics of multiculturalism.

First, people's attachment to their own ethnic heritage increases with their tendency to form homophilous relationships. These findings are not simply correlations but rather causal effects if our identification assumption is true. In other words, having more friends from the same heritage help reinforce a person's attachment to that heritage.

Next, I analyze a series of variables that ask the respondents about their sense of belongings to Canada, their family, their city of residence, and their province. All these subjective evaluations increase with inbreeding homophily index. These results are somewhat surprising especially to those who are skeptical of multiculturalism. They often presume that lack of integration or social isolation can hinder a person's attachment to the host country, to the community, etc. On the other hand, these findings corroborate the belief of many proponents of multiculturalism. They claim that people can only find their sense of belongings to the host country if they find confidence in their own heritage.

Third, social wellbeing as measured by their interactions with the society is found to be enhanced by friendship homophily. Friendship homophily increases general trust, trust of family members and of community significantly in some specifications. Its effect on trust of people in the same school and workplace turns out to be insignicant. Life satisfaction is not improved by friendship homophily. These results offer partial support for proponents of multiculturalism. Immigrants, through their strong ties

Table 6: Inbreeding Homophily and Labor Market Outcomes

	All of Canada			Gateway Cities		
	All	Young		All	Young	
	2SLS 6	2SLS 7	2SLS 8	2SLS 6	2SLS 7	2SLS 8
Panel 1. Employment Probability						
Inbreeding Homophily Index	-0.1749^{c}	-0.0587	-0.0604	-0.1071	-0.0399	-0.0235
	(0.0920)	(0.1256)	(0.1255)	(0.1014)	(0.1550)	(0.1549)
Observations	3185	1344	1344	2118	844	844
Panel 1.1: First Stage Regression						
Childhood Inbreeding Homophily	0.2865^{a}	0.2907^{a}	0.2836^{a}	0.3270^{a}	0.2993^{a}	0.3186^{a}
	(0.0223)	(0.0321)	(0.0439)	(0.0279)	(0.0382)	(0.0533)
Language Distance*Childhood Friendship			0.0165			-0.0367
Homophily			(0.0465)			(0.0535)
F-Stat	51.32^{a}	22.51^{a}	22.30^{a}	32.72^{a}	187.76^{a}	182.36^{a}
Stock and Yogo (2005) Test	241.509^{a}	117.449^{a}	58.7472^{a}	211.618^{a}	80.5281^{a}	40.4782^{a}
Panel 2. Log Hourly Wage						
Inbreeding Homophily Index	-0.0372	0.0792	0.0791	-0.1621	0.0086	0.0169
	(0.1717)	(0.2410)	(0.2409)	(0.1837)	(0.2853)	(0.2823)
Observations	1894	798	798	1258	507	507
Panel 2.1: First Stage Regression						
Childhood Inbreeding Homophily	0.2649^{a}	0.2639^{a}	0.2501^{a}	0.3081^{a}	0.2758^{a}	0.2889^{a}
	(0.0288)	(0.0438)	(0.0615)	(0.0364)	(0.0517)	(0.0754)
Language Distance*Childhood Friendship			0.0303			-0.0223
Homophily			(0.0607)			(0.0706)
F-Stat	28.47^{a}	13.76^{a}	13.92^{a}	354.31^{a}	7.64^{a}	7.43^{a}
Stock and Yogo (2005) Test	127.003^{a}	57.3344^{a}	28.7708^{a}	113.811^{a}	39.8935^{a}	19.9647^{a}
Panel 3. Occupation Status						
Inbreeding Homophily Index	0.1991	-0.0989	-0.0980	0.3286	0.2960	0.2544
	(0.2683)	(0.3875)	(0.3864)	(0.3019)	(0.4665)	(0.4715)
Observations	3265	1373	1373	2155	860	860
Panel 3.1: First Stage Regression						
Childhood Inbreeding Homophily	0.2948^{a}	0.2947^{a}	0.2917^{a}	0.3413^{a}	0.3220^{a}	0.3496^{a}
	(0.0219)	(0.0311)	(0.0423)	(0.0277)	(0.0378)	(0.0530)
Language Distance*Childhood Friendship			0.0070			-0.0501
Homophily			(0.0453)			(0.0527)
F-Stat	53.03^{a}	24.04^{a}	23.64^{a}	33.25^{a}	193.88^{a}	188.19^{a}
Stock and Yogo (2005) Test	265.091^{a}	125.801^{a}	62.8662^{a}	230.53^{a}	92.4617^{a}	46.6791^{a}
Panel 4. Language at Workplace						
Inbreeding Homophily Index	-0.2172^{a}	-0.2297^{a}	-0.2347^{a}	-0.2296^{a}	-0.2421^{b}	-0.2031^{b}
	(0.0665)	(0.0767)	(0.0776)	(0.0792)	(0.1062)	(0.0972)
Observations	2330	987	987	1569	637	637
Panel 4.1: First Stage Regression						
Childhood Inbreeding Homophily	0.2693^{a}	0.2867^{a}	0.2832^{a}	0.3120^{a}	0.3052^{a}	0.3381^{a}
	(0.0257)	(0.0376)	(0.0516)	(0.0322)	(0.0440)	(0.0627)
Language Distance*Childhood Friendship			0.0082			-0.0608
Homophily			(0.0533)			(0.0607)
F-Stat	34.58^{a}	16.81^{a}	16.81^{a}	465.29^{a}	10.23^{a}	9.84^{a}
Stock and Yogo (2005) Test	162.704^{a}	87.4818^{a}	43.708^{a}	149.307^{a}	66.0597^{a}	33.5612^{a}

Panel 1 to Panel 4 report the regression results for four labor market outcomes. Results are reported respectively for the full sample and gateway cities and for young (below age 25) and all immigrants. 2SLS 6 - 7 use childhood (before age 15) homophily index as the instrument. 2SLS 8 use both childhood homophily and its interaction with language distance as instruments. The set of control variables are the same as those in 2SLS 4 of Table 2. The first stage regressions and tests of existence of weak instruments are also reported for all panels. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively, except for Stock and Yogo (2005) tests when ^{a}, ^{b}, and ^{c} represent rejection of presence of weak instruments at maximum rejection rates of 10%, 15%, and 20% respectively.

with coethnic friends, develop their trust of people around them, thus contributing to the accumulation of social capital.

Lastly, I analyze civic participation variables, which are important performance indicators. The findings are not definitive. The signs are mostly positive for friendship homophily. Some coefficients are statistically significant, e.g. volunteering activity, community activities, and voting in city elections. These findings lend further support to the proponents of multiculturalism. At least, homophilous networks do not decrease the tendency for people to vote and to volunteer. Rather, more homophily networks may lead people to join community groups and to volunteer, which in turn helps immigrants to build their sense of belonging to the country.

Table 7: Inbreeding Homophily and Social Outcomes

	All of Canada			Gateway Cities		
	All	Young		All	Young	
	2SLS 6	2SLS 7	2SLS 8	2SLS 6	2SLS 7	2SLS 8
Panel A: Importance of Ethnic Ancestry						
1. Retain Customs - At Least One Ethnic Ancestry Rated Above 4						
Inbreeding Homophily Index	0.5865^{a}	0.4919^{a}	0.4955^{a}	0.5054^{a}	0.5256^{a}	0.5244^{a}
	(0.0969)	(0.1330)	(0.1330)	(0.1139)	(0.1686)	(0.1672)
Observations	3302	1384	1384	2182	869	869
2. Retain Customs - At Least One Ethnic Ancestry Rated Above 5						
Inbreeding Homophily Index	0.5438^{a}	0.4855^{a}	0.4900^{a}	0.5416^{a}	0.5432^{a}	0.5397^{a}
	(0.0877)	(0.1254)	(0.1254)	(0.1004)	(0.1508)	(0.1500)
Observations	3302	1384	1384	2182	869	869
3. Sense of Belonging to Ethnic Ancestry						
Inbreeding Homophily Index	1.0442^{a}	0.7995^{b}	0.8094^{b}	1.0761^{a}	1.0082^{b}	0.9831^{b}
	(0.2835)	(0.3964)	(0.3954)	(0.3177)	(0.4899)	(0.4943)
Observations	3232	1353	1353	2130	850	850
Panel B: Sense of Belonging						
4. Sense of Belonging to Canada						
Inbreeding Homophily Index	0.3145	0.0014	0.0057	0.3957^{c}	0.2232	0.2099
	(0.1982)	(0.2457)	(0.2455)	(0.2235)	(0.3115)	(0.3120)
Observations	3280	1371	1371	2167	861	861
5. Sense of Belonging to Family						
Inbreeding Homophily Index	0.5302^{a}	0.1425	0.1452	0.6972^{a}	0.2751	0.2561
	(0.1801)	(0.2157)	(0.2152)	(0.2214)	(0.3164)	(0.3166)
Observations	3289	1379	1379	2171	866	866
6. Sense of Belonging to City						
Inbreeding Homophily Index	0.6770^{a}	0.4710	0.4694	0.6775^{b}	0.1837	0.1860
	(0.2536)	(0.3445)	(0.3433)	(0.2799)	(0.4102)	(0.4106)
Observations	3267	1370	1370	2156	858	858
7. Sense of Belonging to Province						
Inbreeding Homophily Index	0.4731^{c}	0.1988	0.1991	0.5800^{b}	0.0218	0.0326
	(0.2469)	(0.3344)	(0.3326)	(0.2797)	(0.4069)	(0.4122)
Observations	3266	1363	1363	2161	859	859
Panel C: Social Well-being						
8. General Trust						
Inbreeding Homophily Index	0.0875	0.0981	0.0994	0.2316^{c}	0.3732^{b}	0.3564^{b}
	(0.1027)	(0.1400)	(0.1400)	(0.1239)	(0.1827)	(0.1816)
Observations	3197	1349	1349	2106	842	842
9. Trust Family Members						
Inbreeding Homophily Index	0.3952^{a}	0.2256	0.2314	0.5104^{a}	0.3777	0.3530
	(0.1349)	(0.1708)	(0.1711)	(0.1722)	(0.2641)	(0.2630)
Observations	3294	1379	1379	2177	866	866
10. Trust Community						
Inbreeding Homophily Index	0.2204	0.2526	0.2570	0.4788^{c}	0.8069^{b}	0.7927^{b}
	(0.2129)	(0.2930)	(0.2935)	(0.2541)	(0.3620)	(0.3605)
Observations	3190	1346	1346	2103	844	844
11. Trust School and Workplace						
Inbreeding Homophily Index	0.1264	0.1358	0.1461	0.1490	0.2447	0.2031
	(0.2197)	(0.2930)	(0.2937)	(0.2630)	(0.3795)	(0.3785)
Observations	2826	1200	1200	1897	774	774

Continued on next page

Table 7 – continued from previous page

	All of Canada			Gateway Cities		
	All	Young		All	Young	
	2SLS 6	2SLS 7	2SLS 8	2SLS 6	2SLS 7	2SLS 8
12. Satisfaction						
Inbreeding Homophily Index	0.2334	-0.0101	-0.0108	0.0887	-0.2116	-0.1994
	(0.1878)	(0.2736)	(0.2730)	(0.2085)	(0.3199)	(0.3220)
Observations	3296	1384	1384	2178	870	870
Panel D: Civic Participation						
13. Volunteer Activity						
Inbreeding Homophily Index	0.0196	0.1284	0.1291	0.1192	0.2632^{c}	0.2615^{c}
	(0.0945)	(0.1317)	(0.1316)	(0.1021)	(0.1522)	(0.1518)
Observations	3315	1389	1389	2192	873	873
14. Voted in Federal Elections						
Inbreeding Homophily Index	-0.0130	-0.0681	-0.0703	0.0128	-0.0747	-0.0661
	(0.1037)	(0.1414)	(0.1410)	(0.1148)	(0.1732)	(0.1708)
Observations	2370	1110	1110	1572	704	704
15. Voted in Provincial Elections						
Inbreeding Homophily Index	0.1069	0.1339	0.1385	0.0699	0.0671	0.0365
	(0.1128)	(0.1529)	(0.1531)	(0.1219)	(0.1861)	(0.1814)
Observations	2353	1098	1098	1560	693	693
16. Voted in City Elections						
Inbreeding Homophily Index	0.1985	0.3505^{b}	0.3591^{b}	0.2016	0.2985	0.2497
	(0.1249)	(0.1769)	(0.1767)	(0.1388)	(0.2201)	(0.2191)
Observations	2341	1099	1099	1549	690	690
17. Community Activity						
Inbreeding Homophily Index	0.1016	0.1531	0.1561	0.2619^{b}	0.3872^{b}	0.3763^{b}
	(0.1021)	(0.1411)	(0.1409)	(0.1159)	(0.1775)	(0.1776)
Observations	3314	1389	1389	2192	873	873

Panels A - D report the regression results for four sets of social outcomes related to the objectives of multiculturalism. Results are reported respectively for the full sample and gateway cities and for young (below age 25) and all immigrants. 2SLS 6 - 7 use childhood (before age 15) homophily index as the instrument. 2SLS 8 use both childhood homophily and its interaction with language distance as instruments. The set of control variables are the same as those in 2SLS 4 of Table 2. No first stage results are reported to save space. Most specifications reject the presence of weak instruments. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively.

4.4 Robustness

There are two key types of variables, namely population shares of a person's ethnic ancestry and measures of language skills. Since my identification hinges upon the validity of these measures. My robustness checks focus on these variables. In EDS, there are many alternative indicators we could utilize. I discuss these alternatives sequentially.

In my previous regressions, I use people's first languages to predict a baseline homophily index. After that, people's transition from first language to home language is included to explain additional variations in friendship composition. First, I substitute language distances of individuals' ethnic ancestries to English for those of their first languages in predicting the baseline homophily. The results are essentially the same.

With respect to the population shares, there is also another option. In the first step, I rely on all reported ethnic ancestries in calculating population shares of each ethnicity previously. As an alternative, I use the two highest rated ethnic ancestries of individuals in the full sample to calculate the shares of all ethnic ancestries. In the second step, I use the sum of population shares of a person's two highest rated

ethnicities to represent his ethnic population share in my previous calculation. In the robustness check, I use only the highest-rated ancestry share. The results continue to hold after adopting the alternative method. All the robustness results are omitted to save space. They are available from the author upon request.

5 Conclusion

This paper sets out to test a few presumptions made by advocates and denouncers of multiculturalism. I employ an instrumental variable approach to test the causal effect of language skills on ethnic composition of an individual's friends, measuring both integration of new immigrants and inter-cultural sharing among them. I argue that this is a direct test of one of the implicit premise of Canadian multiculturalism. My findings confirm this intuitive notion. I also explore whether common criticisms of multiculturalism are valid or not. I find no evidence that lack of social integration of new immigrants hinders their participation in the host society or their social wellbeing. In other words, inter-cultural sharing and contacts are not necessary for achieving other goals of multiculturalism. Rather, inter-group sharing should be considered an independent goal by itself if it is still valued by the society. Language skills affect economic variables positively but have no statistically significant effect on social outcomes. These are disappointing for proponents of multiculturalism, who often assume that language policy can achieve many goals of multiculturalism.

This paper has its limitations. Firstly, the study is based on a cross-section of Canadian immigrants in 2001. Conditions may have changed in Canada over the last 10 to 20 years. It will be interesting to analyze such evolution over this period. It will be also interesting to explore the same issues in other countries. Secondly, more direct evaluation of the many programs carried out in Canada and other countries may yield more insighful policy prescriptions. As many public resources are spent on these programs, such studies are really needed to inform the public and the agencies carrying out these programs. Lastly, the paper leaves many questions unanswered. For example, language skills and friendship homophily are found to significantly affect only a few socio-economic goals of multiculturalism. The determination of these outcomes remains a mystery for other researchers to resolve.

References

P. Bayer, S. L. Ross, and G. Topa. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, 116(6):1150–1196, December 2008.

L. A. Beaman. Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the u.s. *The Review of Economic Studies*, 79(1):128–161, 2012.

A. Bisin and T. Verdier. Beyond The Melting Pot: Cultural Transmission, Marriage, And The Evolution Of Ethnic And Religious Traits. *The Quarterly Journal of Economics*, 115(3):955–988, August 2000.

H. Bleakley and A. Chin. Language Skills and Earnings: Evidence from Childhood Immigrants. *The Review of Economics and Statistics*, 86(2):481–496, May 2004.

H. Bleakley and A. Chin. Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants. *American Economic Journal: Applied Economics*, 2(1):165–92, January 2010.

G. J. Borjas. To ghetto or not to ghetto: Ethnicity and residential segregation. *Journal of Urban Economics*, 44(2):228 – 253, 1998. ISSN 0094-1190.

J. K. Brueckner. Friendship networks. *Journal of Regional Science*, 46(5):847–865, 2006.

J. K. Brueckner and O. Smirnov. Workings of the melting pot: Social networks and the evolution of population attributes. *Journal of Regional Science*, 47(2):209–228, 2007.

J. K. Brueckner and O. Smirnov. Social networks and the convergence of population attributes: A generalization. *Journal of Regional Science*, 48(2):359–365, 2008. ISSN 1467-9787.

B. R. Chiswick. Speaking, Reading, and Earnings among Low-Skilled Immigrants. *Journal of Labor Economics*, 9(2):149–70, April 1991.

B. R. Chiswick and P. W. Miller. The Endogeneity between Language and Earnings: International Analyses. *Journal of Labor Economics*, 13(2):246–88, April 1995.

B. R. Chiswick and P. 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.

J. Coleman. Relational analysis: The study of social organizations with survey methods. *Human Organization*, 17(4):28–36, 1958.

S. Currarini, M. O. Jackson, and P. Pin. An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045, 07 2009.

D. M. Cutler, E. L. Glaeser, and J. L. Vigdor. Is the melting pot still hot? explaining the resurgence of immigrant segregation. *The Review of Economics and Statistics*, 90(3):478–497, August 2008.

A. P. Damm. Ethnic enclaves and immigrant labor market outcomes: Quasi-Experimental evidence. *Journal of Labor Economics*, 27(2):281–314, 2009.

M. Dewing. *Canadian Multiculturalism*. PRB (Canada. Library of Parliament. Parliamentary Information and Research Service). Parliamentary Information and Research Service, Library of Parliament, 2009.

C. Dustmann and F. Fabbri. Language proficiency and labour market performance of immigrants in the UK. *Economic Journal*, 113(489):695–717, 07 2003.

C. Dustmann and A. van Soest. Language Fluency And Earnings: Estimation With Misclassified Language Indicators. *The Review of Economics and Statistics*, 83(4):663–674, November 2001.

P.-A. Edin, P. Fredriksson, and O. Åslund. Ethnic enclaves and the economic success of immigrants - Evidence from a natural experiment. *Quarterly Journal of Economics*, 118(1):329–357, 2003.

Globe and Mail Editorial. Strike Multiculturalism from the National Vocabulary. *Globe and Mail*, (October 8), 2010.

M. Granovetter. The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, 1: 201–233, 1983.

M. S. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973. ISSN 00029602.

D. A. Green. Immigrant Occupational Attainment: Assimilation and Mobility over Time. *Journal of Labor Economics*, 17(1):49–79, January 1999.

J. K. Hellerstein, M. McInerney, and D. Neumark. Neighbors and Coworkers: The Importance of Residential Labor Market Networks. *Journal of Labor Economics*, 29(4):659 – 695, 2011.

Y. M. Ioannides and L. D. Loury. Job Information Networks, Neighborhood Effects, and Inequality. *Journal of Economic Literature*, 42(4):1056–1093, December 2004.

S. A. Kossoudji. English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men. *Journal of Labor Economics*, 6(2):205–28, April 1988.

T. Kuran and W. H. Sandholm. Cultural integration and its discontents. *The Review of Economic Studies*, 75(1):201–228, 2008.

I. Kónya. Minorities and majorities: a dynamic model of assimilation. *Canadian Journal of Economics*, 38(4):1431–1452, 2005. ISSN 1540-5982.

I. Kónya. Optimal immigration and cultural assimilation. *Journal of Labor Economics*, 25(2):367–391, 2007. ISSN 0734306X, 15375307.

Q. Li. Language and urban labor market segmentation: Theory and evidence. *Journal of Urban Economics*, 74(C):27–46, 2013.

M. R. Lupul. *The Politics Of Multiculturalism: A Ukrainian-Canadian Memoir*. Canadian Institute of Ukrainian Studies Press, Edmonton, June 2005.

K. Malik. The Failure of Multiculturalism: Community Versus Society in Europe. *Foreign Affairs*, 94 (2):21–32, March/April 2015.

D. Marmaros and B. Sacerdote. How do friendships form? *The Quarterly Journal of Economics*, 121 (1):79–119, 02 2006.

W. McManus. Labor market effects of language enclaves: Hispanic men in the united states. *Journal of Human Resources*, 25(2):228–252, 1990.

W. McManus, W. Gould, and F. Welch. Earnings of hispanic men: The role of english language proficiency. *Journal of Labor Economics*, 1(2):101–30, 1983.

M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444, 2001.

T. Mouw. Estimating the causal effect of social capital: A review of recent research. *Annual Review of Sociology*, 32:79–102, 2006. ISSN 0360-0572.

K. Munshi. Networks in the modern economy: Mexican migrants in the u. s. labor market. *The Quarterly Journal of Economics*, 118(2):549–599, 2003.

E. Patacchini and Y. Zenou. Ethnic networks and employment outcomes. *Regional Science and Urban Economics*, 42(6):938 – 949, 2012. ISSN 0166-0462.

J. Stock and M. Yogo. *Testing for Weak Instruments in Linear IV Regression*, pages 80–108. Cambridge University Press, New York, 2005.

S. Vertovec and S. Wessendorf, editors. *The Multiculturalism Backlash: European Discourses, Policies and Practices*. Routledge, Taylor and Fancis Group, 2010.

L. Xue. *Social capital and labour market outcomes of recent immigrants to Canada: Employment entry, wages and duration of access to the first job in intended occupation.* Ph.D. dissertation, University of Ottawa, 2008.

Y. Zenou. Spatial versus social mismatch. *Journal of Urban Economics*, 74(C):113–132, 2013.

Y. Zenou. A Dynamic Model of Weak and Strong Ties in the Labor Market. *Journal of Labor Economics*, 33(4):891 – 932, 2015.

A Appendix

A.1 Construction of Key Variables

I briefly explain the construction of key variables in this Appendix. Table 8 provides as many details as possible. Most explanations are sufficient on their own. In carrying out robustness checks, I need to modify a few key variables. First, I use EATC1 and EATC2, namely the two highest rating ancestries, to calculate the population shares using the full sample. Next, I match individual ethnic ancestries with these estimated population shares. In the specifications shown in the main text, I use the sum of population shares calculated by both EATC1 and EATC2. In one robustness check, I use only EATC1.

Table 9 lists the language distance measures by first languages, which are used in my baseline models. Note that language distance is defined as

$$LD_i = 3 - LScore_i,$$

so English is zero distance away from itself. Some robustness checks involve changing the basis upon which we measure language distances. Table 10 lists the language distance measures by ethnic ancestry as an alternative method. Again, language scores shown in the table equal 3− language distances. Basically, EAC1 - EAC8, EATC1, and EATC2 correspond closely to the ethnicities listed in this table. I aggregate a few groups into British - Canadian and French - Canadian. Such aggregations are also explained in the table.

Table 8: Construction of Variables

Variable Names	Variables in EDS	Construction Method
Population Share by Ethnic Ancestry	EAC1-EAC8, EATC1, EATC2	EAC1, $\cdots$, EAC8 indicate up to 8 ethnic ancestries. EATC1 and EATC2 are up to two highest rated ethnic ancestry among EAC1-EAC8. First, EAC1-EAC8 are used to construct the population shares of ethnic ancestries. Second, EATC1 and EATC2 are used to find the corresponding population shares as calculated previously. Finally, the sum of the two shares based on EATC1 and EATC2 equals the finalized population share. In the sample of gateway cities, population shares of EAC1-EAC8 are calcuated within Toronto, Montreal, and Vancouver respectively.
Homophily Index H	SNQ0201, SNQ0202	I assign 1,0.75,0.5,0.25,0 to those who report all, most, a half, a few, and none of their friends belong to the same ethnic ancestry as reported in SNQ0201 and SNQ0202. The variable is defined as the sum of the two values. It is censored at 1.
Inbreeding Homophily IH	N/A	Constructed from Homophily Index H and Population Share according to Equation (2).
Childhood IH	SNQ0301, SNQ0302	First, a friendship homophily index during childhood is constructed in a similar fashion as Homophily Index H using SNQ0301 and SNQ0302. Second, this index is combined with population share according to Equation (2).
All Home Languages	LGHO	1 if all home languages LGHO are English only, French only, or a mix of English and French, and 0 if non-official languages or a mix of non-official languages and English or French.
Most Often Home Languages	LGHMO	1 if home languages most often spoken LGHMO are English only, French only, or a mix of English and French, and 0 if non-official languages or a mix of non-official languages and English or French.
Linguistic Share Change: All Home Languages	L1_ENG, L1_FRE, $\cdots$, L1_UKR, LHA_ENG, LHA_FRE, $\cdots$, LHA_UKR	L1_ENG, L1_FRE, $\cdots$, L1_UKR are dummy variables indicating whether the respondent reports English, French, $\cdots$, and Ukranian a first language respectively. Analogously, LHA_ENG, $\cdots$, LHA_UKR report for all home languages. Population shares of language groups are calculated using the full sample of EDS for which the above variables are available. The population share change is the difference between the population share of a person's home language (all) and that measured by a person's first language. For gateway cities, the population shares and the corresponding changes are calculated for Toronto, Montreal, and Vancouver individually.

Continued on next page

Table 8 – continued from previous page

Variable Name	Variables in EDS	Construction Method
Linguistic Share Change: Most Often Home Languages	L1_ENG, L1_FRE, ⋯, L1_UKR, LHM_ENG, LHM_FRE, ⋯, LHM_UKR	LHM_ENG, ⋯, LHM_UKR report for home languages most often spoken. Population shares of language groups population share change are calculated analogously as the all home languages. For gateway cities, the population shares and the corresponding changes are calculated for Toronto, Montreal, and Vancouver individually.
Transition All Languages	L1_ENG, L1_FRE, LHA_ENG, LHA_FRE	This variable calculates LHA_ENG - L1_ENG and LHA_FRE - L1_FRE. If either of these two is one, the dummy variable is one. If otherwise, the dummy variable equals zero.
Transition Most Often Lan.	L1_ENG, L1_FRE, LHM_ENG, LHM_FRE	This variable calculates LHM_ENG - L1_ENG and LHM_FRE - L1_FRE. If either of these two is one, the dummy variable is one. If otherwise, the dummy variable equals zero.
Distance by First Language	L1_ENG-L1_UKR	L1_ENG-L1_UKR are a list of dummy variables indicating whether the respondent reports the respective language as a first language. The distance measure is explained in Table 9 in the appendix.
Visible Minority	VISMIND	1 if a visible minority, 0 otherwise.
Education	HLOS, HLOSCAN	I assign 19, 16, 15, 14, 13, 12, and 9 for HLOS values of 1, 2, 3, 4, 5, 6, and 7. I assign value 0 for people with HLOSCAN=7.
Education in Canada	HLOSCAN	1 if HLOSCAN=1, 0 if HLOSCAN=2.
Age at Immigration	GENAGE	Dummies are defined by GENAGE, including ranges: 0-5, 6-14, 15-24, 25-44, 45-64, and 65+.
Place of Birth	PBREG7, PBSLCT	Dummies are defined by PBREG7, augmented by PBSLCT. The final categories are: Canada, US, UK, Germany, Netherlands, Poland, Italy, Portugal, China, Hong Kong, Philippines, India, NA, Other Unknown, Latin America, Other European, Africa, Other Asian, Oceania and Other.
CMA Fixed Effects	CMA3	Same as CMA3, including Toronto, Montreal, Vancouver, other CMAs, and non-CMAs.
Age	AGES	Average within the bands of AGES.
Gender	SEX	1 if female, 0 if male, based on SEX.
Married	MARSTAT	1 if married, 0 if otherwise, based on MARSTAT.
Employment Status	MAINACTS	1 if employed, 0 if unemployed or homemaking.
Log Hourly Wage	INCP20N, WKWEEKS, WKHOURS	Logarithm of hourly wages, based on continuous measurement of INCP20N/(WKWEEKS*WKHOURS). They are the averages within the bands identified by the previous variables.
Occupation Status	SOCSUM	1, 2, ⋯, 5 for managers (1), professionals (3, ⋯,6), other white collar (2, 7), blue collar (8, 9, 10), and not working (97). Numbers of SOCSUM are in parentheses.

Continued on next page

Table 8 – continued from previous page

Variable Name	Variables in EDS	Construction Method
Language at Work	LGWO	1 if speaking English, French, or a mix of the two only at work; 0 if speaking non-official languages or mixture with official languages.
Ethnic Customs Rated 4+	CUSTOM45	1 if at least 1 ethnic tradition is rated 4 or above, 0 otherwise.
Ethnic Customs Rated 5	CUSTOM05	1 if at least 1 ethnic tradition is rated 5, 0 otherwise.
Sense Belonging - Ethnicity	AT_Q020	1-5 ratings, 5 means the strongest sense of belongings.
Sense Belonging - Canada	AT_Q050	1-5 ratings, 5 means the strongest sense of belongings.
Sense Belonging - Family	AT_Q010	1-5 ratings, 5 means the strongest sense of belongings.
Sense Belonging - City	AT_Q030	1-5 ratings, 5 means the strongest sense of belongings.
Sense Belonging - Province	AT_Q040	1-5 ratings, 5 means the strongest sense of belongings.
Trust - General	TS_Q020	1 if trust people generally, 0 if no.
Trust - Family	TS_Q030	1-5 ratings, 5 means people can be trusted a lot.
Trust - Community	TS_Q040	1-5 ratings, 5 means people can be trusted a lot.
Trust - School and Work	TS_Q050	1-5 ratings, 5 means people can be trusted a lot.
Life Satisfaction	TS_Q010	1-5 ratings, 5 means very satisfied.
Volunteered	VOLUNT	1 if volunteered in community/club activities, 0 otherwise.
Voted in Federal	PC_Q110	1 if voted in Federal election, 0 if no, missing otherwise.
Voted in Provincial	PC_Q120	1 if voted in Provincial election, 0 if no, missing otherwise.
Voted in City	PC_Q130	1 if voted in Municipal election, 0 if no, missing otherwise.
Community Activity	PC_Q020	1 if participated in community/club activities, 0 otherwise.

Table 9: Language Distance and First Language Group

First Language	Language Score	Language Distance	Note
English	3.00	0.00	
French	2.50	0.50	
Arabic	1.50	1.50	
Chinese	1.38	1.63	Average of Mandarin and Cantonese
Dutch	2.75	0.25	
German	2.25	0.75	
Italian	2.50	0.50	
Polish	2.00	1.00	
Portuguese	2.50	0.50	
Punjabi	1.75	1.25	Substituted by Hindi
Spanish	2.25	0.75	
Taglog	2.00	1.00	
Ukrainian	2.25	0.75	Substituted by Russian

Source: Chiswick and Miller (2005) and Canadian Ethnic Diversity Survey

Table 10: Language Distance and Ethnic Ancestry (EATC1/EATC2)

Ethnic Ancestry	Score	Distance	Note
No response	N/A	N/A	
French Canadian	2.50	0.50	Includes: French, French-Canadian, and Quebecois in the data
British Canadian	3.00	0.00	Includes: Other provincial and regional groups, English, Irish, Scottish, Welsh, Other British, and Hyphenated response with Canadian
Austrian	2.25	0.75	German
Belgian	2.63	0.38	Dutch + French
Dutch	2.75	0.25	
German	2.25	0.75	
Swiss	2.42	0.58	German + French + Italian
Other Western European	2.50	0.50	German + French + Dutch
Danish	2.25	0.75	
Finnish	2.00	1.00	
Norwegian	3.00	0.00	
Swedish	3.00	0.00	
Other Northern European	3.00	0.00	Icelandic 617
Hungarian	2.00	1.00	
Polish	2.00	1.00	
Romanian	3.00	0.00	
Russian	2.25	0.75	
Ukrainian	2.25	0.75	Russian
Other Eastern European	2.00	1.00	Czech
Greek	1.75	1.25	
Italian	2.50	0.50	
Spanish	2.25	0.75	
Portuguese	2.50	0.50	
Other Southern European	2.00	1.00	Serbo-Croatian
Jewish	2.00	1.00	Herbrew
Other European	2.00	1.00	Slav
African	2.58	0.42	Afrikaans + Swahili + Amharic
Lebanese	1.50	1.50	Arabic
Other Arab	1.50	1.50	
West Asian	2.00	1.00	Farsi + Turkish + Dari
Punjabi	1.75	1.25	Hindi
East Indian n.i.e.	1.75	1.25	Hindi
Other South Asian	1.75	1.25	Bengali + Sinhala + Nepali
Chinese	1.38	1.63	Mandarin + Catonese
Filipino	2.00	1.00	
Japanese	1.00	2.00	
Vietnamese	1.50	1.50	
Other East and Southeast Asian	1.86	1.14	Indonesian + Malay + Burmese + Lao+ Cambodian + Thai + Korean
Jamaican	3.00	0.00	
Other Caribbean	2.25	0.75	Spanish
Latin, Central and/or South American	2.25	0.75	Spanish + Portuguese
American (USA)	3.00	0.00	highest score
Other	N/A	N/A	
Hyphenated response without Canadian	N/A	N/A	
Unknown	N/A	N/A	

Source: Chiswick and Miller (2005) and Canadian Ethnic Diversity Survey