Canadian Labour Market and Skills Researcher Network

Working Paper No. 144

The effect of linguistic proximity on the occupational assimilation of immigrant men

Alicia Adsera
Princeton University

Ana Ferrer
University of Waterloo

October 2014

CLSRN is funded by the Social Sciences and Humanities Research Council of Canada (SSHRC) under its Strategic Knowledge Clusters Program. Research activities of CLSRN are carried out with support of Human Resources and Skills Development Canada (HRSDC). All opinions are those of the authors and do not reflect the views of HRSDC or the SSHRC.

The effect of linguistic proximity on the occupational assimilation of immigrant men

Alícia Adserà,
Woodrow Wilson School
Princeton University
437 Wallace Hall,
Princeton NJ 08544
adsera@princeton.edu

Ana Ferrer
Department of Economics
University of Calgary
2500 University Drive NW, SS 432
Calgary, Alberta, Canada T2N 1N4
aferrer@ucalgary.ca

Abstract

This paper contributes to the analysis of the integration of immigrants in the Canadian labour market by focusing in two relatively new dimensions. We combine the large samples of the restricted version of the Canadian Census (1991-2006) with both a novel measure of linguistic proximity of the immigrant's mother tongue to that of the destination country and with information of the occupational skills embodied in the jobs immigrants hold. This allows us to assess the role that language plays in the labour market performance of immigrants and to better study their career progression relative to the native born. Results show that linguistic proximity shapes the evolution of job-skill content of immigrant jobs over time and in some cases affects patterns of wage assimilation of immigrants.

Key words: Linguistic ability, occupational assimilation, immigration

JEL codes: J15, J24

The analysis was conducted at South Western Ontario RDC which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the CRDCN are made possible by the financial or in-kind support of the SSHRC, the CIHR, the CFI, Statistics Canada and participating universities whose support is gratefully acknowledged. The views expressed in this paper do not necessarily represent the CRDCN's or that of its partners'.

1. Introduction

For Western economies that rely on large influxes of immigrants as a way of encouraging economic growth and demographic sustainability, such as Australia and Canada, successful immigrant assimilation is of capital importance. Research highlights the influence of linguistic ability as a key determinant of immigrant economic success. This paper uses a novel measure of linguistic proximity to explore whether language ability greatly influences labour market performance of immigrants to Canada and how it affects their assimilation over time. Besides standard measures of labour market performance, such as wages, we also consider participation in highly skilled jobs and the evolution of the skill involved with the jobs immigrant hold.

Theories of immigrant assimilation consider that, upon immigration, the new entrant's lack of country specific human capital (language fluency, institutional knowledge, recognized credentials) will cause the depreciation of the human capital brought into the country. This depreciation accounts for lower initial earnings when compared to similar native-born workers. Economic theory predicts that, with time in the country, immigrants will experience an increase in earnings associated with the acquisition of country specific skills, bringing immigrants' earnings closer to those of similar native-born workers. There is a long tradition of academic and non-academic studies that document this phenomenon using different measures of labour market performance (Chiswick, 1978 and 1986; Borjas 1985). Since the 1990s, these studies report that new cohorts of immigrants to most major host countries have experienced a significant deterioration in their economic well-being relative to previous arrival cohorts (Borjas 1995; Bell, 1997; Aydemir and Skaterud 2005; Antecol et al, 2006; Clark and Lindley, 2009; Beenstock et al., 2010). In North America, the drop in earnings and employment has largely been attributed to changes in immigrant's background (Friedberg, 2000; Bratsberg and Raaum, 2004; Borjas (), Picott and Hou, 2009). More specifically, language ability has come across different studies as one of the main determinants of successful integration (Borjas, 2013; Bleakley and Chin, 2004; Chiswick et al., 2005).

We build on this literature to assess the influence of linguistic ability on the occupational assimilation of immigrants. This is of particular relevance in the Canadian case because of rising immigrant-native born wage differences in recent years - once education and other observable characteristics have been accounted for - has been particularly important in this country (Aydemir and Skuterud, 2005). The fact that Canada's immigration policy has been targeted at

educated immigrants for quite some time makes this increase in wage differences the more puzzling (Ferrer, Picot and Riddell, 2014). While part of the problem is attributed to the fact that the skills implied by a given level of education might differ between the foreign and Canadian educated (Li and Sweetman (2013); Ferrer et al. 2006), a complementary explanation is that language ability impairs the successful transfer of skills across countries (Chiswick and Miller 2010). Although often acknowledged as a potential cause of differences in outcomes between immigrant and native born workers, this venue has been less explored in the case of Canada because the Canadian census lacks a proper measure of linguistic ability. We add to these studies by incorporating a measure of linguistic proximity to the analysis, which allows us to uncover rich patterns of the role of language on the assimilation of Canadian immigrants.

Lack of language fluency is likely to result in poor labour market outcomes upon arrival. An important issue to consider is whether these poor initial outcomes change with time in the country and, presumably, the picking up of language fluency. To this effect we incorporate data on the occupational skills involved in the jobs immigrants take to assess the extent to which assimilation in skills is influenced by linguistic proximity. Further, we explore whether there are systematic differences in assimilation (in skills and wages) depending on the degree of linguistic proximity between immigrant's mother tongue and English or French.

The next section describes the methodology and data used in the analysis. Section 3 presents the results and section 4 concludes.

2. Methodology and Data

Our analysis of the labour market outcomes of immigrants follows seminal work in the literature by Borjas (1985). We estimate an equation of the following form,

$$Y_{it} = X_{it}\beta_1 + L_i\beta_2 + skill_s\beta_3 + \beta_{jt}\sum_{j=1}^{4} Coh_j * t + t + \epsilon_{it} \quad (1)$$

where the dependent variable (Y_i) is a measure of labour force performance of individual *i*, whom we observe (only) in survey year *t*; X_{jt} is a vector of that individual/household characteristics. L_i is a measure of linguistic proximity (LP) between the individual's mother tongue and the language in the destination; $skill_s$ are dummies for each of the job skills considered here; Coh_j, are indicators designating *j* five-year immigrant arrival cohorts, which we interact with the indicator for survey year (*t*) to track the evolution of different entry cohorts over time. Cohorts are defined over five years (i.e. the 91 cohort includes individuals arriving between

1986 and 1991). This results in ten relevant interactions. The β_{jt} are coefficients for the cohort-time indicators measuring the assimilation of immigrant cohorts. β_2 and β_3 are vectors of coefficients for the LP proximity and job skill indicators. The reference group for the LP indicators are native speakers. The coefficients for the time-cohort indicators correspond, by default to those of immigrant with LP=1 (no distance).

2.1 Linguistic Proximity

A distinctive feature of our analysis is the incorporation of a measure of LP to assess its role on the assimilation of immigrants. As suggested by previous research, both fluency in the language of the destination country and the ability to learn it quickly will influence immigrant's success in destination countries' labor markets (Kossoudji (1988), Chiswick and Miller (2002, 2007, 2010), Dustmann (1994), Dustmann and van Soest (2001 and 2002), and Dustmann and Fabbri (2003). Bleakley and Chin (2004 and 2010) exploit differences on adult English proficiency between immigrants from non-English speaking source countries who arrive as young children versus the others and find that linguistic competence is a key variable to explain immigrant's disparities in terms of educational attainment, earnings and social outcomes. Recent studies (Chiswick and Miller, 2005; Isphording and Otten, 2011) also show that it is easier for a foreigner to acquire a language if her native language is linguistically closer to the language to be learned. Unfortunately, since the Canadian Census has no measure of linguistic fluency it is not possible to replicate the same type of analysis in the sample we employ.

As a proxy of fluency we use a measure of linguistic proximity (LP) between two countries developed in Adserà and Pytliková (2014) and based on information from the encyclopaedia of languages Ethnologue (Lewis, 2009). The LP index ranks languages depending on how many levels of the linguistic family tree the languages of both the destination and the source country share, from zero to five. Languages that do not share any level of the linguistic tree are, for instance English and Chinese, while English and Hindi have only one common branch. Germanic languages share the same first and second- linguistic tree level and Germanic North languages share the same first up to third linguistic tree level. Scandinavian East (Danish, Norwegian and Swedish) languages or South European languages (Italian, Spanish and French) share all four levels of linguistic tree family. A final categorization is used for two languages that are the same. We use the measure of linguistic proximity as separate indicators and label these

categories after the number of branches shared between the languages (0, 1, 2, 3, 4) and "Same language", we will sometimes refer to these immigrants as English speaking immigrants.

The measure provides a far better adjusted and smoother indicator of proximity than the standard dummy for common language used in most of the literature. We match this measure with an indicator variable denoting the individual's country of birth available in the Canadian census. We use three measures of LP: proximity between the first official languages in both origin and destination countries is denoted "*First official language*", the minimum proximity between any official language at origin and destination denoted "*Any official language*" and, the proximity between the most common used in both countries we name "*Most used language"*. Among them, the second measure is likely to produce more flexible results taking into account potential familiarity between languages other than the first official language at origin and either English or French. The third measure is likely to be more accurate as it picks up actual familiarity between the most common language used at origin and English.

In robustness analysis we employ a number of alternative measures of linguistic distance such as the Levenshtein distance index which relies on phonetic dissimilarity of words in two languages for all world languages and was produced at the Max Planck Institute based on phonetic distance and the Dyen index based on the similarity between samples of words among Indo-European languages.

2.2 Wages and Skill Progression

The Census collects information on the wage received and weeks worked the previous year; however information on hours of work is collected over the census reference week. This introduces problems in constructing a reliable measure of hourly wages and we follow other Canadian studies in using weekly wages as our main measure of labour market performance.

A distinctive feature of this paper is a focus on occupational skills involved in the jobs held by immigrants. Occupational skills might track immigrant assimilation better than wages. Skilled immigrants in particular might be switching between unskilled and skilled jobs as they acquire language proficiency, but experiencing little effect on pay, as initially some skilled jobs might be lower paid than some unskilled jobs. Hence in alternative specifications of the model, Y_{it} will become a measure of occupational skills.

This focus on skills rather than on occupation is significant. Typically, most job changes (and their accompanying wage responses) involve substantial adjustments in skill requirements, but might not result in a change in the occupational category, particularly when measured broadly as pointed by Autor and Handel (2013). The occupational matching literature uses detailed information in occupation databases - either the Dictionary of Occupational Titles (DOT) or the Occupational Information Network (O*NET) – to derive a small set of fundamental skill requirements for each job. The skills we consider here come from Imai et al (2011) and are derived from the O*NET. These include two indexes for cognitive skills (social and analytical) and three indexes for manual skills (fine motor skills, physical strength, and visual skills). To facilitate interpretation of the data, the detailed information in the O*NET is summarized by constructing a low-dimensional vector of occupational characteristics using Principal Component Analysis (PCA).[1] The result is a vector of skills necessary to perform the job tasks associated with each 4-digit occupational category using the O*NET. The factor analysis uses as weights the distribution of the skill distribution of the Canadian working population; hence a unit of the skill score (with mean zero) can be interpreted as one standard deviation in the skill distribution of the Canadian population.

2.3 Sample Description

The rest of our data comes from the restricted files of the Canadian Census (1991 through 2006). The confidential files offer detailed information on individual occupation, country of birth, year of arrival and mother tongue that we need to match the Census with the additional variables. The large samples contained in the restricted Census files also help to reduce concerns of attenuation bias (Aydemir and Borjas, 2010). Finally, in order to reduce computing time to reasonable length, from each census we select all immigrants plus a 25 percent random sample of Canadian born individuals and weight the observations accordingly.

In this paper we restrict ourselves to male immigrants who arrived as adults (at age 18 or over), because the behaviour of child immigrants is likely very different from that of adult immigrants (Adsera and Ferrer, 2013; Mayer and Riphalm, 2000; Beck et al. 2012). Our vector of individual characteristics, *X*, contains location (province, and Census Metropolitan Area

[1] In the PCA, factor loadings are calculated so that variation of the data explained by the constructed variable is maximized. A detailed description of the procedure can be found in Imai et al. (2011).

(CMA)), education, experience and experience squared, and indicator for GDP at the country of origin (either when the immigrant immigrated, or at age 18).

Summary statistics are presented in table 1. Immigrants are generally older than the native born, and more likely to have university education than natives. Immigrant men hold jobs that require, on average, similar interpersonal skills and higher analytical skills than the average Canadian-born males. Jobs held by immigrant men also require on average less motor and strength skills than Canadian-born males. Figure 1 shows the full distribution of analytical, social and physical strength skills.[2]

Regarding linguistic proximity (LP) (*first official language*), a large fraction of immigrant men (34%) come from a country where the first official language has no branch in the linguistic tree in common with English. The main countries in this category are China, Philippines and Korea, those where the first official language is Arab (Egypt, United Arab Emirates) and most African countries. These countries remain classified as having no branch in common (LP=0) even if we use the "*most used*" metric. Some countries, particularly former colonies that have English or French as one official language, are no longer classified as having no common branch with English if using the "*any official language metric*" - the more important of these countries in terms of immigration source is Pakistan. An additional 40% of immigrant men come from countries where the first official language has only one branch in common to English. This category includes countries where Spanish is the first official language (Mexico), most of Eastern and Southern Europe (France, Ukraine) and some Eastern Asian countries, such as Afghanistan, Bangladesh, India, and Sri Lanka.[3] Spanish speaking countries are no longer classified as only having one level of the linguistic tree in common when using "*any official language*" since they are linguistically close to French and have at least the 4th level of the linguistic tree in common according to this measure. Countries with the 2nd level in common are Northern European (Denmark, Iceland, Norway and Sweden) and countries with the 3rd level in common are those where the first official language is of Germanic origin (Switzerland, Germany, Belgium, The Netherlands and their colonies). The main countries with the same language as Canada are the US and the UK, when using the first official language and this group

[2] Note that skills here measure the skills involved in performing the job, not the actual skills of the worker.
[3] Iran and Pakistan have a LP=0.1 according to first official language, but a LP=0 according to "most used language" and "any official language"

also includes additional smaller commonwealth countries (such as Mauritius). However, some of them are not classified as such if we use the "*most commonly used*" language criteria instead.

3. Regression results

3.1. Wage and linguistic proximity

Table 2 reports results from equation (1), showing the difference in log wages between immigrant and similar native-born individuals, when controlling for skills and linguistic proximity. We have considered the three measures of linguistic proximity, *"1st official language"* in columns 1 through 3, *"Any official language"* in columns 4 through and 6 and *"Most used language"* in columns 7 through 9. The exercise is performed here with and without controls for place of birth in the analysis, in columns 1, 4 and 7 and in columns 2, 5 and 8 respectively, to disentangle the effect of language from that of area of origin.

Higher wages are generally associated with jobs requiring higher levels of analytical and motor skills and lower levels of physical strength. Not surprisingly, higher wages are also associated with closer linguistic proximity. Relative to immigrants who come from countries where English is the first official language, those from countries where the first official language does not have any branch in common with English receive lower wages, around 33 percent lower. The closer the official language to English the lower the wage penalty is, which is between 18 and 6 percent lower than the native born. The exception of this linear effect with linguistic proximity occurs for countries where their first official language shares up to the 2nd level of the tree with English, which includes most of the Nordic languages. These immigrants show the lowest penalties (9 to 7 percent lower wages than natives). The penalties for those with the closest linguistic proximity (English speakers in some way) vary depending on the definition of LP used. In particular, they are the largest when LP is measured between *any official language* at origin and English/French compared to the other two metrics. This suggests that it is possibly a nosier measure of linguistic ability than the others. In general, controlling for area of origin substantially reduces the estimated wage differentials of linguistic proximity. Particularly, immigrants from Nordic countries (with a common 2nd level with English) no longer experience a wage gap with respect to the native born, and differentials for English speaker immigrants (those with no linguistic distance) are reduced by half or more. Immigrants with the lowest

linguistic proximity (no common branch) also show substantially reduced wage differentials, while those with middle levels of proximity (LP=1, 2 or 3) show little effect.

Finally, we show that the measure of linguistic proximity usedhere produces similar estimates that when using instead other measures of linguistic proximity common in the literature. In columns 3, 6 and 9 in table 2 we show results using the Levenshtein index of linguistic distance. Since the Levenshtein index actually measures linguistic distance rather than linguistic proximity, the coefficient for the dummy variable "immigrant" measures on its own the wage gap for the immigrants who are English speakers (Levenshtein =0). The Levenshtein index produces results that are in line with those presented using the dummies developed through the Adserà and Pytliková index. For instance, according to the estimates obtained with the Levenshtein distance between the *first official languages* (column 3), an English-speaker immigrant has a 9 percent lower wage than that of a native born. An immigrant with an average Levenshtein score (76) faces an additional penalty of 15 percent, whilst an immigrant with a score one standard deviation over the mean (score 114) faces an additional penalty of 23 percent over an English-speaker immigrant. Using the other metrics based on the Levenshtein produces similar results. Estimates with the Levenshtein index that measures the distance between *any official language* in both countries we predict 16 percent lower wages for English speakers and 19 percent lower wages for the average linguistic distance (score of 80) than for natives. With the Levenshtein for the *most used language,* estimates imply 5 percent lower wages for the English speakers and 10 percent lower for the average linguistic distance (score 51).[4]

As mentioned above, the literature on immigrant assimilation has emphasized the role of language fluency as a key determinant of successful labour market integration in various countries (Chiswick and Miller, 2010; Clark and Lindley, 2009; Dustmann and van Soest, 2002). To examine this, we introduce in equation (1) an interaction between the time-cohort indicators and the different levels of linguistic proximity, which allows us to follow over time immigrant cohorts with different levels of linguistic proficiency. Figure 2 shows how the wages of cohorts with different degrees of linguistic proximity fare with time in Canada. The results show that all cohorts of immigrant men from countries with languages that share no branches with English (dotted line) have significantly lower wages at entry (40 to 30 percent lower) than the native

[4] Using the Dyen indexes produce similar results.

born. These cohorts do experience some improvement with time in the country although their performance remains well below that of the Canadian born, at least 20 percent below after 15 to 20 years in Canada. There are no significant differences depending on how LP is measured for this group other than in levels. However, when studying the performance over time in Canada of the English-speaker group, there are important differences in how LP is measured. Under any metric used, immigrant men with English as mother tongue (solid line) fare better, relatively speaking, than those with zero linguistic proximity. The metric used determines the level and shape of wage evolution for this group. Measuring LP by "*any official language"* or "*first official language"*, rather than "*Most used language*", produces estimates of wages that are increasing with time in the country. However, "*any official language"* produces estimates that are closer to those of immigrants with languages with no common branch to English, which supports the idea that this is a nosier proxy for linguistic ability. In contrast, the "*Most used language*" metric generates flat estimates of wages that are indistinguishable from those of the native born. This confirms the idea that this metric is the closest measure of actual knowledge of English rather than simple familiarity with English.

3.2. Does linguistic proximity affect the returns of some skills more than others?

Given the emphasis of Canadian policy during the 1990s and early 2000s on selecting immigrants with high levels of human capital, in particular high levels of formal education, a natural question to ask would be whether or not the transfer of skills embodied in formal education depends greatly on the degree of linguistic proximity.[5] As a first approach we assess the association between wages, LP and job skills for a subsample of university-educated individuals (Table 1 in the appendix) using the same measures of linguistic proximity as in Table 2. Focusing on columns (2), (5) and (8), wage differences between immigrants and native born workers are somewhat smaller among university-educated men than among those with no college, particularly for immigrants with high levels of LP.

To explore this idea further, we introduce an interaction between university education and LP Results shown in Table 3 display the returns to formal education for different levels of LP, relative to those of native born men with no university education. Linguistic proximity is important in determining the returns to education. In particular, penalties for low linguistic

[5] We do not know where post-secondary education has been obtained

proximity are smaller between uneducated immigrant and native-born workers than between educated ones. For instance, according to "*any official language*", a non-university educated immigrant with no -linguistic proximity will have, on average, wages that are 10 percent lower than a non-university educated native born. By comparison, a university-educated immigrant will have wages 20 percent lower than those of an educated native born (0.65-0.45). In general, all metrics produce similar estimates for the uneducated workers at high levels of LP, and more varying estimates at low levels of LP. Among educated immigrants *first official language* produces the more consistent estimates across levels of linguistic proximity, whereas the other two metrics show larger variation for groups with medium linguistic ability.

It is plausible that LP is more important in determining pay when jobs require a specific set of skills. For instance, LP could matter more when associated with jobs requiring social skills than with jobs requiring physical strength. We explore this possibility introducing an interaction between LP and skills in equation (1).

$$Y_{it} = X_{it}\,\beta_1 + L_i\beta_2 + skill_s\beta_3 * \text{NB} + \beta_{ls}\sum_{s=s1}^{s5} \quad \sum_{l=0}^{1} L_l * skill_s + \beta_{jt}\sum_{j=1}^{4} Coh_j * t + t + \epsilon_{it} \tag{2}$$

where s_1 *to* s_5 denote the skills (social, motor, visual, strength and analytical) and *l* denotes the levels of LP (from no common branch to same language). Results from this equation are shown in table 4. The first column shows the vector of coefficients β_3, that indicate the effect of a one-standard deviation increase in a given skill on the wages of native speakers. The rest of the columns show the additional effect on immigrants relative to native born working in jobs that require similar skills ($\beta_2 + \beta_{ls}$). For instance, the first row in the second panel ("*first official language*") indicates that one standard deviation increase in social skills results in a wage increase of 2 percent for the native born. Relative to a native, an immigrant with English as first official in their country of birth, and working a job with similar levels of required skills has a wage 6 percent lower, and an immigrant with null linguistic proximity has a wage that is 30 percent lower. The large differences that arise after controlling for the skill level are likely due to the uni-dimensional nature of the job. It is plausible that immigrants with poor linguistic ability work jobs with different composition of skills than the native born, but we cannot get at these differences. Nevertheless, we can learn several things from Table 4. For instance, the effect of LP seems to be more important for the returns to social skills for immigrant men than for other skills. The difference in the wage gap between immigrants whose first official language in the

country of origin is English and those whose first official language at origin shares not common branch with English is 24 percent points (-30 versus -6), while this difference is lower for other skills such as analytical or strength. The importance of language is significant for motor and strength skills, which again might be related to the skill package associated to jobs. Less intuitive is the relatively little effect that linguistic proximity seems to have on the reward to analytical skills. We attribute this to the differences in probability that an immigrant with high levels of analytical skills to transfer has learned English before arriving. Unfortunately we have no way of measuring this dimension of linguistic ability at the individual level.

Overall we find that linguistic ability is important in determining the rewards to different skills. However, our uni-dimensional skill measures cannot asses the importance of the different skill packages associated with the jobs. We are currently exploring this possibility.

3.3 Does language help in getting better jobs for immigrants?

Another dimension along which to examine the performance of immigrants in the labour market is by looking at the skills imbibed in the jobs they do. To this effect we use the skills index as the dependent variable in a regression that looks into the effect of LP on the skills required for the jobs immigrants hold. A basic regression of the association between skills and language indexes shows that, controlling for different measures of human capital, immigrant men with lower linguistic proximity typically hold jobs requiring more strength skills and less analytical or social skills than the native born (Table 2 in the appendix). This is independent of the metric, although results using "*first official language*" and "*most used language*" metric are larger and rather consistent with one another. Results using linguistic proximity between "*any official languages*" are more muted, consistent with our results with wages. This metric seems less able to pick up differences in the degree of LP.

Previous studies suggest that the ability to learn the language quickly improves the labour market outcomes of immigrants and that this depends on linguistic proximity. If so, we would expect immigrants with close linguistic proximity to improve their job status as measured by the increased use of "high-status" skills (social and analytical skills) and reduced use in "low-status" skills (strength and fine motor skills) faster than immigrants with low linguistic proximity. To examine this possibility we introduce an interaction between the levels of LP and the indicators that follow the evolution of cohorts over time. We show these results graphically in Figure 3 using LP between the most used language and English. Relative to the native speakers, male

immigrants with English as the major language in their country of origin (continuous line) do better than those whose major language has no common branch with English (dashed line). This implies that the former are in jobs using more analytical and less strength skills than the latter, and that the skills they employ are pretty close to those required in the jobs of the native born.[6]

Immigrants with low linguistic proximity work in jobs that require between 0.1 and 0.2 standard deviations higher levels of strength, around 0.35 standard deviations lower levels of social skills and around 0.2 standard deviation lower levels of analytical skills than natives. Those with the highest levels of linguistic proximity work in jobs that are quite close to those of the native born in terms of the skills required – within 0.05 standard deviations. There is slight evidence that the situation of immigrants, in terms of job-skills does not improve with time in the country. Levels of strength rise slightly with time in the country and for successive cohorts, whereas social and analytical skills seem stagnant over time for different cohorts. Interestingly there is not an obvious reduction in the required analytical and social skill levels in the jobs for successive cohorts. This suggests that the decline in immigrant earnings experience by the most recent cohorts might have more to do with the probability of finding a job than with the quality of the jobs held.

Another dimension where higher linguistic proximity could help immigrants get better jobs is in helping them with the transfer of formal education to the host country. This is a key dimension since formal education has been an important component of immigrant selection policies to Canada and the education level of Canadian immigrants did increase substantially during the period of study. We include an interaction between formal education (university) and linguistic proximity, to address this possibility: if higher linguistic ability helps with the transfer of formal education skills, we expect educated immigrants with close linguistic proximity to do considerably better than those without. Results are reported in Table 5, relative to non-university educated native born workers. Skill requirements of the jobs held are the dependent variable in the regression. The first column reports the university education coefficient for the native born. The rest of the columns report the relative difference in skills between immigrants with different degrees of linguistic ability and education relative to the reference group. In Table 5, university-

[6] Curiously this is not the case if using "any official language". Again, this is likely because immigrants from previous English or French colonies may have a LP equal to 1 using this metric, but that might not be an accurate measure of their language ability of the individual or the language skills needed on the job.

educated immigrants with English as the *most used language* in the country of origin work in jobs that require 1.51 standard deviations higher levels of analytical skills than those held by natives. They also work in jobs requiring 0.04 standard deviations higher levels of analytical skills than similarly educated native born (1.51 - 1.47).

We find that, in general, educated immigrants with LP=2 or higher work in higher status jobs, as measured by having higher or similar social and analytical skills and lower strength skills than similarly educated native born workers. This seems to be independent of the metric used, although different metrics produce slightly different results. As it happened before, using the "*any official language*" metric results in a lesser contrast between different levels of linguistic proximity than using other metrics. Interestingly, a high level of linguistic proximity also help less educated immigrants to get better jobs, but not to the same extent than it helps educated immigrants. For instance, according to the most used language metric, for university-educated men, higher linguistic proximity translates into an increase of 0.47 standard deviations in social skills required by the job (from 1.68 for the English speaker to 1.21 for those with no common branch to English) but for the uneducated it is 0.31 standard deviations (from 0.13 for the English speaker to -0.18 for those with no common branch with English). Linguistic proximity is typically enough to erase all differences between immigrant and native-born skills associated with native-born jobs, for both educated and non-educated workers, except when using the "any official language" measure.

4. Conclusion

We combine the large samples of the restricted version of the Canadian Census (1991-2006) with both a novel measure of linguistic proximity of the immigrant's mother tongue to that of the destination country and with information of the occupational skills embodied in the jobs immigrants hold. This allows us to assess the role that language plays in the labour market performance of immigrants and to better study their career progression relative to the native born.

Our results show that linguistic proximity affects the returns to human capital and skills of immigrant men. It certainly plays a role in the level of such returns, but it does not seem to influence how fast wage parity is reached between immigrants and the native born. Linguistic proximity also plays a role in the skills required in the jobs immigrants hold. However, we do not

find evidence that it influences the speed of assimilation into better jobs. We also find that linguistic proximity imposes larger wage penalties to the university-educated men than to non-university-educated, and that it significantly affects the status of the jobs they hold.

References

Antecol H, Kuhn P and Trejo S (2006) Assimilation via Prices or Quantities? Sources of Immigrant Earnings Growth in Australia, Canada and the United States. *Journal of Human Resources* 41: 821-840

Adserà, A. and Ferrer, A. (2014) "Factors influencing the fertility choices of child immigrants in Canada, *Population Studies*, 68(1), pp. 65-79

Adserà and Ferrer, (2014) "The Myth of Immigrant Women as Secondary Workers: Evidence from Canada" (with Alicia Adsera) *AER papers and proceedings*, 104(3), pp.360-364

Adserà, A. and Pytliková, M. (2014) "The Role of Languages in shaping International Migration", Economic Journal (forthcoming).

Aydemir, A. and Borjas, G. (2010) "Attenuation Bias in Measuring the Wage Impact of Immigration", *NBER working Paper 1622906*.

Aydemir, A. and Skuterud, M. (2005) "Explaining the Deteriorating Entry Earnings of Canadian Immigrant Cohorts: 1966-2000", *Canadian Journal of Economics*, Vol. 38(2), pp. 641-72.

Beenstock, M. B. Chiswick and A. Patiel (2010) "Testing the Immigrant Assimilation hypothesis with Longitudinal data" *Review of Economics of the Household,* vol.8(1), pp. 7-27

Bell BD (1997) The Performance of Immigrants in the United Kingdom: Evidence from the GHS. *Economic Journal* 107: 333-345

Bleakley H. and A. Chin (2004) "Language Skills and Earnings: Evidence from Childhood Immigrants." *Review of Economics and Statistics* 84 (2): 481-496.

Borjas,G. (2013) "The Slowdown in the Economic Assimilation of Immigrants: Aging and Cohort Effects Revisited Again" , *NBER Working Papers* 19116,

Borjas, G.(1985) "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980's?" *Journal of Labor Economics* 3, pp. 463-89.

Bratsberg, B. and O. Raaum (2004) Identifying Earnings Assimilation of Immigrants under Changing Macroeconomic Conditions. *Scandinavian Journal of Economics* 106: 1-22

Chiswick, B. R. 1978 The Effect of Americanization on the Earnings of Foreign-Born Men. *Journal of Political Economy* 86(5): 897-922.

_____ 1986, "Ìs the New Immigration Less Skilled Than the Old?" *Journal of Labor Economics*, 4(2):168-192

Chiswick, B. and P. Miller (2002) "Immigrant Earnings: Language Skills, Linguistic Concentrations and the Business Cycle." *Journal of Population Economics* 15(1): 31-57.

Chiswick, B. and P. Miller (2007) "Computer Usage, Destination Language Proficiency and the Earnings of Natives and Immigrants." *Review of the Economics of the Household* 5 (2): 129-157.

Chiswick, B. and P. Miller (2010) "Occupational Language Requirements and the Value of English in the US Labor Market." *Journal of Population Economics* 23(1): 353–372.

Clark, K., and Lindley, J. (2009). Immigrant assimilation pre and post labour market entry: Evidence from the UK labour force survey. *Journal of Population Economics, 22*(1), 175-198

Duleep, H. and M. Regets (1999), “Immigrants and Human-Capital Investment”, *The American Economic Review* , Vol. 89, No. 2, Papers and Proceedings of the One Hundred Eleventh Annual Meeting of the American Economic Association, pp. 186-191

*Duleep, H. and Sanders, S.(1993) “The Decision to Work by Married Immigrant Women” *Industrial and Labor Relations Review* 46: 677-90.

Dustmann, C. (1994) “Speaking Fluency, Writing Fluency and Earnings of Migrants.” *Journal of Population Economics* 7: 133–56.

Dustmann, C. and A. van Soest (2001) “Language Fluency and Earnings: Estimation with Misclassified Language Indicators.” *The Review of Economics and Statistics* 83 (4): 663-674.

---------- (2002) “Language and the Earnings of Immigrants.” *Industrial and Labor Relations Review* 55 (3): 473–492.

Dustmann C and Fabbri F (2003) Language Proficiency and Labour Market Performance of Immigrants in the UK. *Economic Journal* 113: 695-717

Ferrer, A., G. Picott and C. Riddell (2014), “New Directions in Immigration Policy: Canada’s Evolving Approach to Immigration Selection”, *forthcoming International Migration Review*, *CLSRN* working paper #107

Ferrer, A.M., and C. Riddell (2008), “Education, credentials and immigrant eaarnings”, *Canadian Journal of Economics*, vol. 4(1), pp. 186-216

Ferrer, A.M., D. Green, C. Riddell (2006), “The Effect of Literacy on Immigrant Earnings”, The Journal of Human Resources vol. 41(2) p. 380-410

Friedberg R (2000) You Can’t Take it With You: Immigrant Assimilation and the Portability of Human Capital. *Journal of Labor Economics* 18: 221-51

Green, D. and C. Worswick ()

Imai, S., D. Stacey and C. Warman (2011) "From Engineer to Taxi Driver? Occupational Skills and the Economic Outcomes of Immigrants," *Working Papers 1275*, Queen's University,

Ingram, B. and G. Neumann (2006) “The returns to skill.” *Labour Economics* 13 (1):35–59

Kossoudji, S. (1988). “The Impact of English Language Ability on the Labor Market Opportunities of Asian and Hispanic Immigrant Men.” *Journal of Labor Economics* 6 (3): 205-228.

Lewis, P. (ed.) (2009) *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com/.

Qing Li & Arthur Sweetman, 2013. "The Quality of Immigrant Source Country Educational Outcomes: Do they Matter in the Receiving Country?,", CReAM Discussion Paper Series 1332.

Picot, G. and F. Hou (2009) “The Effect of Immigrant Selection and the IT Bust on the Entry Earnings of Immigrants” *CLSRN WP #29*

Sweetman, Arthur & Warman, Casey, 2009. "Temporary Foreign Workers and Former International Students as a Source of Permanent Immigration," *CLSRN Working Papers* 2009-34,

Table 1. Sample Summary Statistics. Census 1991-2006

		Canadian born	Immigrant		
Age		39	46		
Education	HS or less	0.52	0.40		
	Non university	0.33	0.32		
	Bachelor	0.12	0.17		
	Graduate	0.04	0.11		
Labour Force Participation		0.87	0.85		
Skills	Social	-0.08	-0.07		
	Motor	0.33	0.22		
	Strength	0.29	0.16		
	Quantitative	0.05	0.13		
			Any	**First**	**Major**
Linguistic proximity to English or French	None		0.18	0.34	0.44
	1st Level		0.16	0.40	0.37
	2nd Level		0.01	0.00	0.00
	3rd Level		0.05	0.05	0.05
	4th Level		0.17		
	Same Language	1.00	0.43	0.20	0.14
Levenshtein			50.58	75.98	80.49
% Observations (weighted)		80%		20%	

Immigrants are adult at arrival.

"First Language" indicates that linguistic proximity is measured between the first official language at origin and destination (English). "Any language" indicates that linguistic proximity is measured between any official language at origin and official language at destination (English or French). Most used language indicates that distance is measured between the most commonly used language at origin and destination (English).

Table 2. Weekly wages of Immigrant men

		First Language			Any Language			Most used language		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Immigrant				-0.09***			-0. 16***			-0.05***
Levenshtein				-0.002***			-0.001***			-0.003***
Skill index	Social	0.01 **	0.01 **	0.01 **	0.01***	0.01***	0.01 **	0.01 **	0.01***	0.004 **
	Motor	0.10***	0.10***	0.10***	0.10***	0.10***	0.10***	0.10***	0.10***	0.10***
	Visual	0.04***	0.04***	0.04***	0.04***	0.03***	0.04***	0.04***	0.03***	0.04***
	Strength	-0.08***	-0.08***	-0.09***	-0.08***	-0.08***	-0.08***	-0.08***	-0.08***	-0.08***
	Analytical	0.17***	0.17***	0.17***	0.16***	0.16***	0.17***	0.16***	0.16***	0.17***
LP indicators	None	-0.33***	-0.18***		-0.32***	-0.16***		-0.34***	-0.19***	
	1st Level	-0.24***	-0.20***		-0.28***	-0.29***		-0.24***	-0.20***	
	2nd Level	-0.09***	-0.03		-0.09***	-0.01		-0.07***	-0.03	
	3rd Level	-0.16***	-0.10***		-0.21***	-0.27***		-0.21***	-0.18***	
	4th Level				-0.17***	-0.05***				
	Same Language	-0.10***	-0.05***		-0.18***	-0.06***		-0.06***	-0.03***	
POB control [1]		--	Yes	--	--	Yes	--	--	Yes	--

All regressions include controls for experience, experience squared, education, location, GDP at country of origin and survey year. Immigrants are adult at arrival. "First Language" indicates that linguistic proximity is measured between the first official language at origin and destination (English). "Any language" indicates that linguistic proximity is measured between any official language at origin and official language at destination (English or French). Most used language indicates that distance is measured between the most commonly used language at origin and destination (English).

(***, **) indicates the coefficient is significant at 1% or 5% respectively

[1] The omitted POB (place of birth) group is US/Europe

Table 3. Effect of Linguistic Proximity and University education on immigrant wages

	NB	Same Lg	4th level	3rd level	2nd level	1st level	None
LP (Any official language) and Education							
Non university	Ref	-0.03***	-0.06 ***	-0.21***	0.01	-0.28***	-0.10***
University	0.65***	0.54***	0.61 ***	0.33***	0.68***	0.25***	0.45***
LP (First official language) and Education							
Non university	Ref.	-0.02***		-0.12***	-0.01	-0.19***	-0.19***
University	0.65***	0.64***		0.62***	0.54***	0.51***	0.62***
LP (Most used language) and Education							
Non university	Ref.	0.00		-0.18***	-0.02	-0.19***	-0.18***
University	0.65***	0.64***		0.53***	0.71***	0.31***	0.40***

All regressions include controls for experience, experience squared, education, location, GDP at country of origin, place of origin (US/Europe is the omitted group) and survey year. Immigrants are adult at arrival.

Panel 1 shows the effect of university education for different levels of linguistic proximity relative to non-university educated native born, measured by proximity between any official languages at origin and English/French.

Panel 2 shows the effect of university education for different levels of linguistic proximity relative to non-university educated native born, measured by proximity between first official language at origin and English.

Panel 3 shows the effect of university education for different levels of linguistic proximity relative to non-university educated native born, measured by proximity between the most commonly used language at origin and English.

(***) indicates the coefficient is significant at 1%

Table 4. Effect of Linguistic Proximity and Skills on Wages of immigrant men

	NB	Same Lg	4th level	3rd level	2nd level	1st level	None
LP (Any official language) and Job Skills							
Social	0.02***	-0.09***	-0.22***	-0.15***	-0.04	-0.37***	-0.31***
Motor	0.12***	-0.12***	-0.21***	-0.16***	0.00	-0.33***	-0.24**
Visual	0.04***	-0.10***	-0.03	-0.35***	0.00	-0.31***	-0.15***
Strength	-0.09***	0.01	-0.04*	-0.21***	-0.03	-0.29***	-0.20***
Analytical	0.16***	-0.02	0.00	-0.29***	0.00	-0.27***	-0.11***
LP (First official language) and Job Skills							
Social	0.02***	-0.06***		0.02	0.04	-0.32***	-0.29***
Motor	0.12***	-0.08***		-0.05*	-0.01	-0.31***	-0.27**
Visual	0.04***	-0.08***		-0.22***	-0.03	-0.22***	-0.18***
Strength	-0.09***	-0.01		-0.07***	-0.07**	-0.15***	-0.16***
Analytical	0.16***	-0.01		-0.13**	-0.01	-0.17***	-0.12***
LP (Most used language) and Job Skills							
Social	0.02***	-0.06***		-0.03	-0.04	-0.30***	-0.28***
Motor	0.12***	-0.06***		-0.13**	-0.02	-0.30***	-0.26**
Visual	0.04***	-0.08***		-0.28***	-0.04	-0.22***	-0.19***
Strength	-0.09***	-0.02		-0.09***	-0.07***	-0.14***	-0.16***
Analytical	0.16***	0.00		-0.22**	-0.01	-0.17***	-0.12***

All regression includes cohort-time effects for immigrants plus controls for experience, experience squared, education, location, area of origin indicators, GDP at country of origin and survey year. Immigrants are adult at arrival

Coefficients show the effect of an increase of one more SD of skill on wages, for different levels of linguistic proximity relative to a native born working in a job with similar skills. The first panel uses proximity between any official languages at origin and English/French, the second panel uses proximity between first official language at origin and English, and the third panel uses proximity between the most commonly used language at origin and English.

(***) indicates the coefficient is significant at 1%

Table 5. Effect of Linguistic Proximity and University education on immigrant skills

		NB	Same Lg	4th level	3rd level	2nd level	1st level	None
1. LP (Any official language) and Education								
Strength:	Non university	Ref	-0.04***	0.05 ***	-0.02**	0.04*	0.07***	-0.11***
	University	-1.36***	-1.23***	-1.40***	-1.31***	-1.36***	-1.18***	-1.33***
Social	Non university	Ref	-0.04***	-0.17 ***	-0.03***	-0.07***	-0.15***	-0.03***
	University	1.61***	1.40***	1.45***	1.46***	1.47***	1.32***	1.47***
Analytical	Non university	Ref	-0.02***	-0.10 ***	0.06***	0.04*	-0.11***	0.03***
	University	-1.50***	-1.35***	-1.51***	-1.61***	-1.48***	-1.42***	-1.51***
2. LP (First official language) and Education								
Strength:	Non university	Ref.	-0.04***		-0.11***	0.00	0.15***	-0.06***
	University	-1.35***	-1.37***		-1.38***	-1.41***	-1.12***	-1.22***
Social	Non university	Ref	-0.02 ***		0.02**	-0.01	-0.25***	-0.14***
	University	1.60***	1.64***		1.61***	1.51***	1.22***	1.27***
Analytical	Non university	Ref	0.04***		0.08***	0.07*	-0.18***	-0.10***
	University	1.49***	1.49***		1.56***	1.51***	1.30***	1.34***
3. LP (Most used language) and Education								
Strength:	Non university	Ref.	-0.13***		-0.05***	-0.00	0.19***	0.09***
	University	-1.35***	-1.38***		-1.27***	-1.41***	-1.10***	-1.08***
Social	Non university	Ref	0.13 ***		-0.03**	0.02	-0.28***	-0.18***
	University	1.58***	1.68***		1.47***	1.53***	1.17***	1.21***
Analytical	Non university	Ref	0.19***		0.07***	0.12***	-0.18***	-0.08***
	University	1.47***	1.51***		1.43***	1.55***	1.30***	1.34***

All regressions include controls for experience, experience squared, education, location, area of origin, GDP at country of origin, and survey year. Immigrants are adult at arrival.

Panel 1 shows the effect of university education for different levels of linguistic proximity relative to non-university educated native born, measured by proximity between any official languages at origin and English/French. Panel 2 and 3 repeat the exercise using proximity between first official language at origin and English, and proximity between the most commonly used language at origin and English respectively

(***) indicates the coefficient is significant at 1%

Figure 1. Distribution of skills by immigrant status

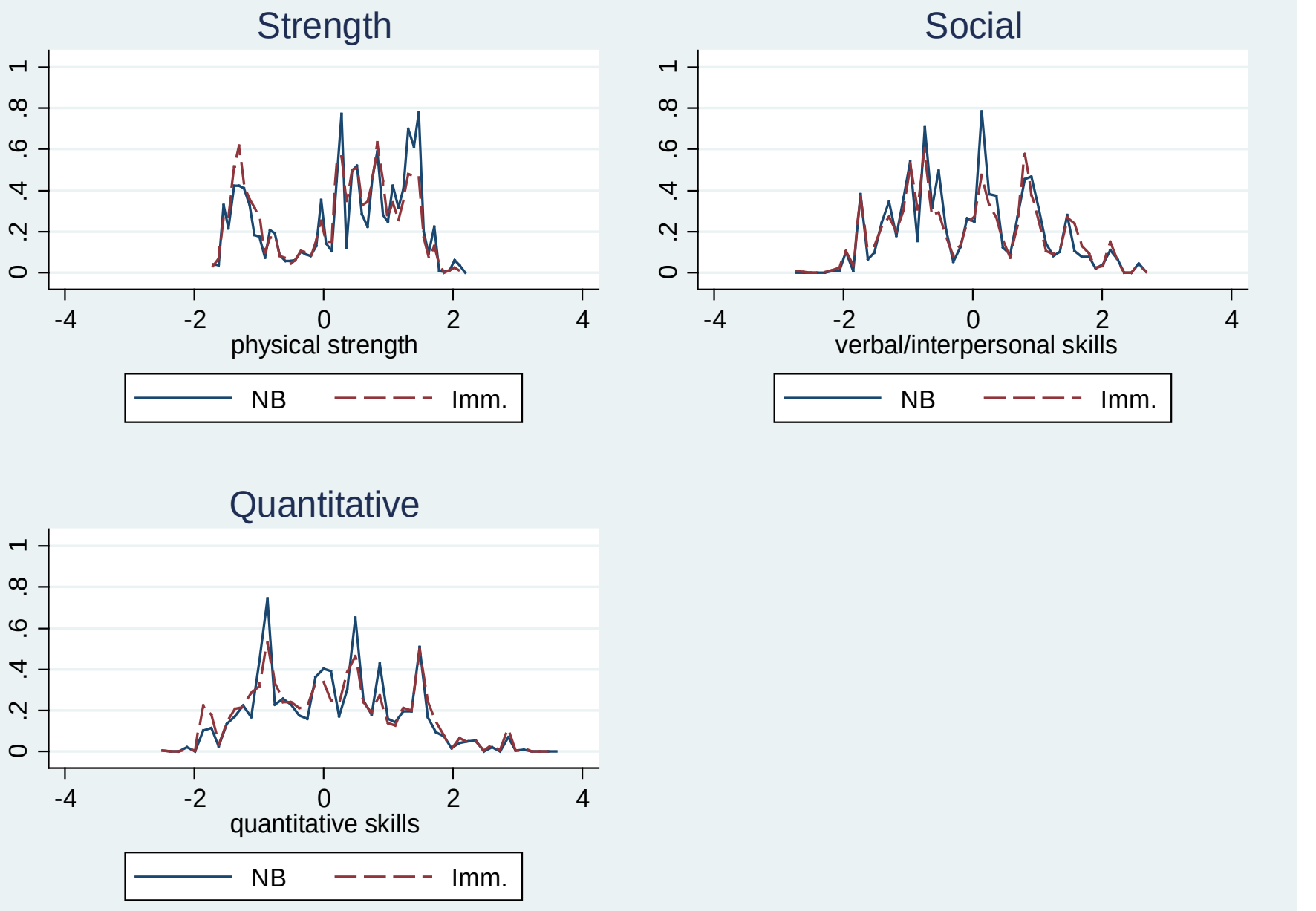
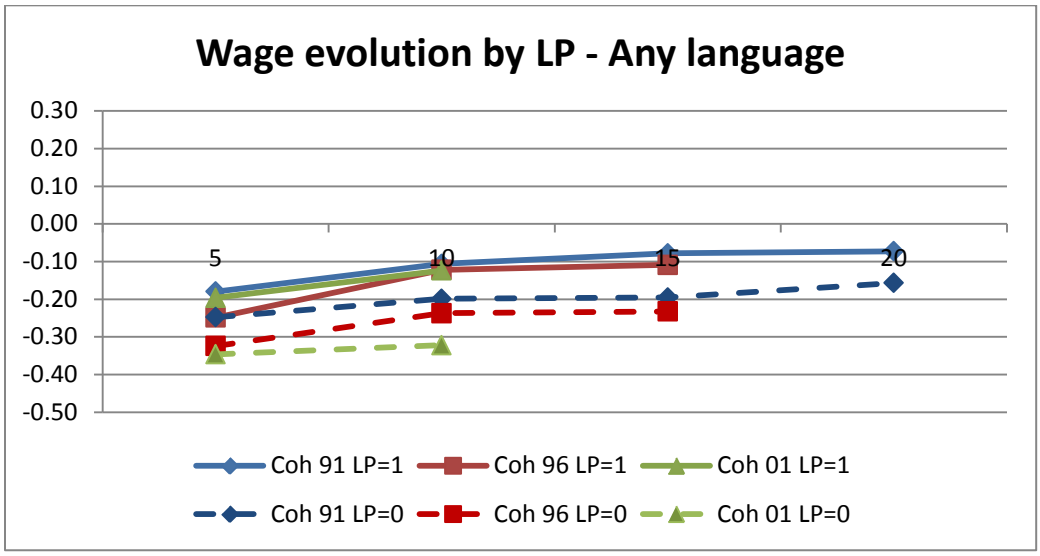
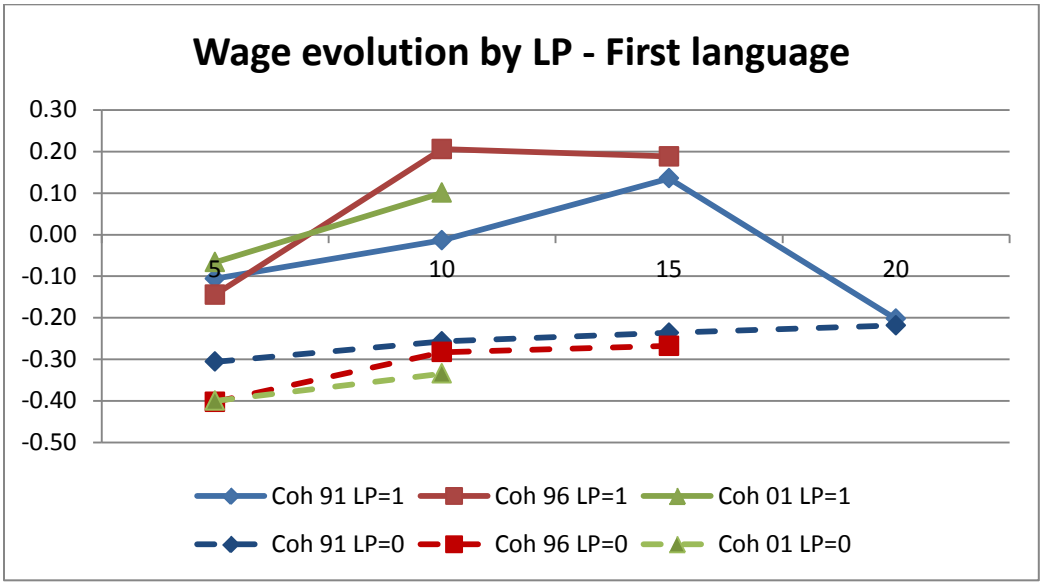


Figure 2. Immigrants' wage assimilation by LP





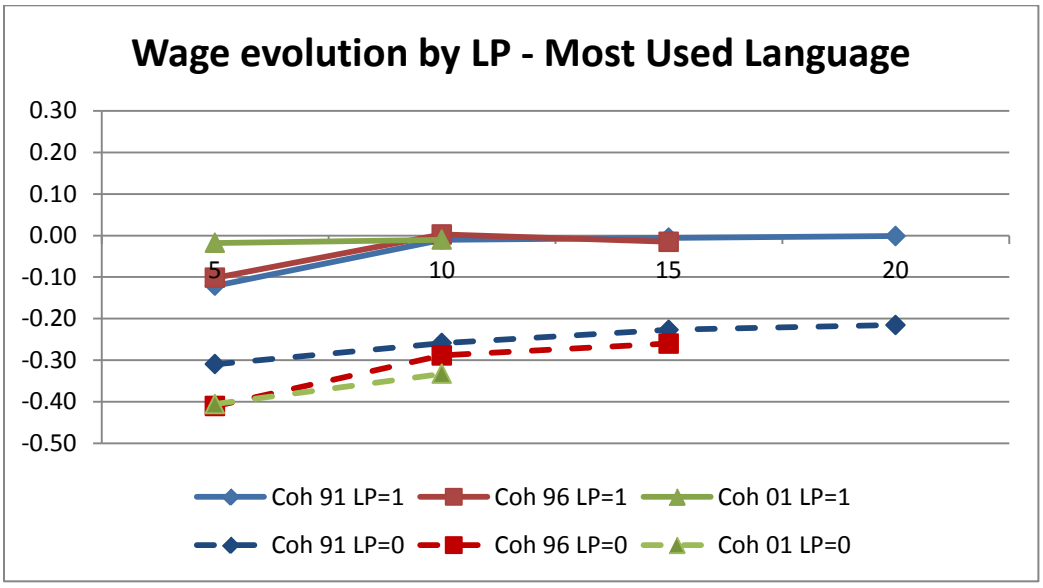
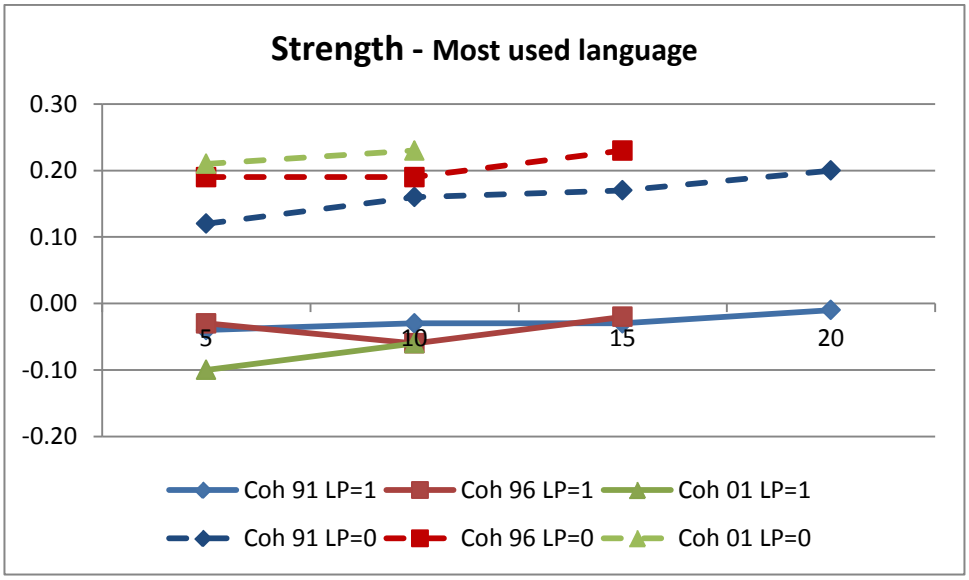
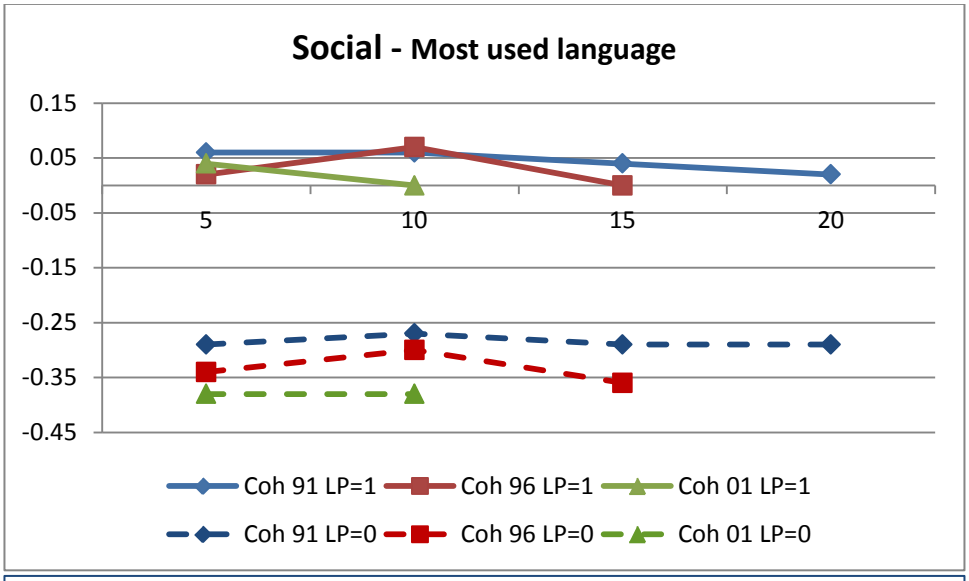
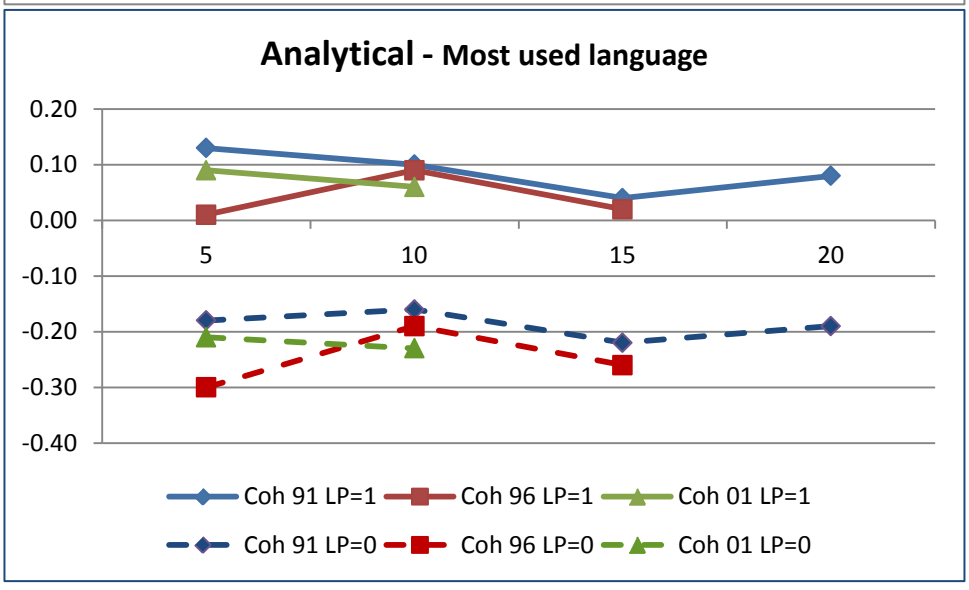


Figure 3. Evolution of skills by Linguistic Proximity (Most used language)







Appendix

Table 1. Weekly wages of University educated immigrant men

		First Language			Any Language			Most used language		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Immigrant				-0.03***			-0. 11***			-0.01***
Levenshtein				-0.003***			-0.001***			-0.003***
Skill index	Social	0.06 ***	0.06 ***	0.06 ***	0.06 ***	0.06 ***	0.06 ***	0.06 ***	0.06 ***	0.06 ***
	Motor	0.08***	0.08***	0.08***	0.08***	0.08***	0.08***	0.08***	0.08***	0.08***
	Visual	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***
	Strength	-0.06***	-0.06***	-0.06***	-0.06***	-0.06***	-0.06***	-0.06***	-0.06***	-0.06***
	Analytical	0.19***	0.19***	0.19***	0.19***	0.19***	0.19***	0.19***	0.19***	0.19***
LP indicators	None	-0.31***	-0.21***		-0.27***	-0.21***		-0.31***	-0.28***	
	1st Level	-0.28***	-0.25***		-0.25***	-0.27***		-0.27***	-0.24***	
	2nd Level	-0.01	0.01		-0.07	-0.01		0.002	0.003	
	3rd Level	-0.02 *	-0.001		-0.17***	-0.24***		-0.13***	-0.13***	
	4th Level				-0.22***	-0.11***				
	Same Language	-0.05***	-0.04***		-0.12***	-0.05***		-0.02***	-0.03***	
POB control [1]		--	Yes	--	--	Yes	--	--	Yes	--

All regressions include controls for experience, experience squared, education, location, region of origin, GDP at country of origin and survey year. Immigrants are adult at arrival. "First Language" indicates that linguistic proximity is measured between the first official language at origin and destination (English). "Any language" indicates that linguistic proximity is measured between any official language at origin and official language at destination (English or French). Most used language indicates that distance is measured between the most commonly used language at origin and destination (English).

(***, **) indicates the coefficient is significant at 1% or 5% respectively

[1] The omitted POB (place of birth) group is US/Europe

Table 2. Effect of language on strength and quantitative skills.

	Any official language [(1)]			First Official language [(2)]			Most used language [(3)]		
LP	Strength	Social	Analytical	Strength	Social	Analytical	Strength	Social	Analytical
None	0.04	-0.06	0.04	0.09	-0.22	-0.13	0.15	-0.25	-0.11
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1st Level	0.13	-0.26	-0.16	0.19	-0.30	-0.19	0.22	-0.33	-0.20
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
2nd Level	0.07	-0.11	0.03	0.01	-0.06	0.07	0.002	-0.05	0.07
	(0.00)	(0.00)	(0.11)	(0.65)	(0.00)	(0.00)	(0.94)	(0.02)	(0.00)
3rd Level	0.02	-0.13	0.02	-0.05	-0.01	0.10	0.001	-0.09	0.01
	(0.01)	(0.00)	(0.00)	(0.00)	(0.25)	(0.00)	(0.87)	(0.00)	(0.19)
4th Level	0.07	-0.14	-0.02						
	(0.00)	(0.00)	(0.00)						
Same Language	0.06	-0.09	-0.03	-0.03	0.02	0.08	-0.08	0.07	0.11
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Regression of human capital and language on skills.
All regressions include controls for experience, experience squared, education, location, region of origin, GDP at country of origin and survey year. Immigrants are adult at arrival. “First Language” indicates that linguistic proximity is measured between the first official language at origin and destination (English). “Any language” indicates that linguistic proximity is measured between any official language at origin and official language at destination (English or French). Most used language indicates that distance is measured between the most commonly used language at origin and destination (English).
(***, **) indicates the coefficient is significant at 1% or 5% respectively
[1] The omitted POB (place of birth) group is US/Europe