Three Essays on Skills in the Labour Market

by

Nick Manuel

A Thesis

Presented To

The University of Guelph

In partial fulfilment of requirements

for the degree of

Doctor of Philosophy

in

Economics

Guelph, Ontario, Canada

ABSTRACT

THREE ESSAYS ON SKILLS IN THE LABOUR MARKET

Nick Manuel Advisor: Dr. Miana Plesca

University of Guelph, 2020

This thesis contains three chapters, each of which explore a different research question that pertains to the role of skills in the labour market. In the first chapter, we find that immigrants educated in Canada receive higher returns to their communication skills than immigrants educated abroad. This gap in skill returns is able to explain the entirety of Canadian educated immigrant's 10% earnings advantage. Our results are robust to controlling for the quality of universities in the immigrant's country of study, and for occupation and industry choice. The gaps are stable across time and across quantiles of the immigrant earnings distribution.

In the second chapter, we find evidence that approximately half of the gap in self-employment rates between immigrants and natives occurs within occupations. Furthermore, this within occupation gap is concentrated in high skill occupations, despite immigrants' earnings advantage from self-employment being in low skill occupations. To explain this, we propose a theoretical model which predicts that migration costs are less likely to deter migration for those individuals who intend to pursue high earning occupations in self-employment. Empirical results are consistent with this model. Furthermore, similar results are found when using natives who have migrated from one state to another as a proxy for immigrants.

In the third chapter, we find that post-secondary graduates of all major fields tend to be employed in occupations that are good matches for their skills. However, there is considerable disparity in the degree of matching across fields of study. We find that graduates of STEM fields are, on average, more poorly matched than graduates of other post-secondary fields of study. This is the result of STEM graduates being more likely to be very poorly matched on the basis of the skills that they use in their jobs. We also find important differences in matching across levels of post-secondary education, and between men and women. Even after conditioning for the skills that are associated with an individual's occupation, we continue to find that those who studied fields that use dissimilar skills to those that are associated with their occupation receive a small earnings penalty.

Acknowledgements

First of all, I wish to thank my advisor, Dr. Miana Plesca, for her invaluable assistance in the completion of this thesis. Without her, this body of work would not have been possible. I also wish to extend my sincere thanks to Dr. Michael Hoy for serving as a member of my advisory committee and for providing immensely helpful advice on this thesis.I would also like to thank the numerous faculty members and graduate students in the Department of Economics and Finance who have each contributed to my development as an economist over these past few years. Lastly, I wish to thank my grandparents, Tony & Marjorie Manuel, and my partner Nadia Nassar. They have provided endless love and support throughout this journey. I never would have made it through without them.

Contents

List of Figures

List of Tables

Chapter 1: Skill Transferability and the Earnings of Immigrants

1.1 Introduction

Each year, well over 100,000[1] immigrants are admitted to Canada as part of the economic class, with the expectation that they will make significant contributions to the Canadian economy. Despite this objective, there has been a marked deterioration in the earnings and employment outcomes of immigrants since the 1980s, with successive cohorts of immigrants continuing to perform worse relative to native born Canadians entering the labour market at the same time (Baker and Benjamin 1994; Aydemir and Skuterud 2005). Furthermore, there is substantial heterogeneity in labour market outcomes across groups of immigrants, with immigration class (Sweetman and Warman 2013) and English/French language ability (Aydemir and Skuterud 2005, Clarke and Skuterud 2016) being among the well established predictors of immigrants' labour market success. Considerable research effort has been devoted towards identifying why some immigrants perform better than others, with one of the primary objectives being to improve our knowledge of how to identify immigrants who will be highly valued in the Canadian labour market.

Recent literature has noted that those immigrants who earned a degree from a Canadian post-secondary institution have higher labour market earnings than immigrants who obtained their degree outside of Canada (Sweetman and Warman 2014; Fortin et al. 2016; Hou and Lu 2017; Chen and Skuterud 2017). However, we need to have a better understanding of why this gap exists, since education is a major criteria upon which prospective immigrants are evaluated in the Canadian system.

One proposed explanation relates to the quality of education in an immigrant's source country. Li and Sweetman (2014) find that immigrants who were born in countries with higher quality education, as measured by standardized test scores, receive greater returns to their education in the Canadian labour market compared to those immigrants who were born in countries with lower quality education.

[1] Based on recent admissions reported in the 2017 Annual Report to Parliament on Immigration (Immigration and Canada 2017), and the federal government's current projections for 2019-2021 immigration levels (Immigration and Canada 2018).

Related literature proposes that foreign educated immigrants are at a disadvantage in the Canadian labour market because their otherwise equivalent foreign credentials are not recognized by Canadian employers or Canadian professional associations (Augustine 2015, Sweetman et al. 2015).

We propose here another explanation, related to the transferability of skills. We create a measure of *ex-ante* skills starting from field of study information, from which we map into seven skill categories: communication (or social expression), comprehension, logical, physical, executive, technical and hard science (STEM). We show that, compared to immigrants educated in Canada, immigrants educated abroad find it more difficult to transfer communication skills, and, to a lesser extent, logical skills, to the Canadian labour market. This explains away their wage penalty relative to immigrants educated in Canada.

By focusing on skills, our work is related to (Fortin et al., 2016), who find that the gap in earnings between immigrants who studied in Canada and immigrants who studied abroad differs across post-secondary fields of study, and to (Ferrer et al., 2006) who provide evidence that Canadian educated immigrants have superior basic English literacy skills compared to foreign educated immigrants. Our contribution is to define a measure of *ex-ante* skill and estimate the heterogeneity in the returns to that skill between immigrants educated in Canada and abroad, conditional on quality of schooling and on any existing occupational licensing requirements. We describe the resulting differential in skill returns as transferability of skill.

Recently, researchers have measured skills at the individual level using data from the O*NET database (Autor et al. 2003; Acemoglu and Autor 2011). This is typically accomplished by taking the skills of a given individual's occupation and assuming that these are also the skills of that individual. This approach was used by (Imai et al., 2019) and (Orlov, 2018) to generate a measure of skills associated with an immigrant's last source country occupation. However, this established methodology poses a problem when attempting to measure transferability of skill on the basis of location of study. Since an immigrant's occupation in Canada is a reflection of both their *ex-ante* skill levels and the portion of their skills that can be transferred to the Canadian labour market, transferability itself can not be measured without an estimate of *ex-ante* skills.[2]

[2] This can be seen clearly when comparing two prospective immigrants from country Z who have comparably poor English/French language abilities, however one of them possesses strong expression skills when communicating in language Z, while the other one does not. If expression skills are rendered non-transferable to Canada as a result of the language difference between Z and Canada, we would expect that both immigrants would end up employed in occupations that do not require strong social skills. Therefore, we would be unable to separate

We propose to address this by measuring individuals' skills using their post-secondary field of study as opposed to their occupation. The U.S. based National Center for Education Statistics (NCES) provides a concordance between fields of study and the occupations for which those fields prepare an individual. We then use O*NET to assign the skills of the associated occupations back to the field of study and from there to the individuals who studied this field. Since this measure of skills does not depend on one's eventual labour market outcomes, we are able to use it as an estimate of *ex-ante* skill levels. In order to address concerns that the instructional content associated with a given field of study may differ across countries in ways that are not related to location-specific content, we include a measure of university quality in the immigrant's location of study in our analysis.

We are aware of one other paper where the skill content of a field of study is inferred using the O*NET link and used as a measure of skill supply. (Altonji et al., 2014) construct the abstract, routine, and manual task composition for each field of study by linking each major to the occupations that are typically associated with it. The authors document changes across time in the relative demand and supply of skill, and argue that changes in the price of skill account for about two-thirds of the rise in inequality in the USA. We use a similar approach of inferring skill from the supply side. Even though there is no heterogeneity across individuals within a major in this measure, regression analysis remains appropriate as long as this heterogeneity balances within the major.

By constructing *ex-ante* skills this way, we provide the first estimates of the extent to which foreign educated immigrants are able to transfer their skills to the Canadian labour market, relative to Canadian educated immigrants. We find that immigrants who received a bachelor's degree in their home country receive significantly lower returns to their expression skills[3] (and, to a lesser extent, to their logical and technical skills) than immigrants who received a bachelor's degree in Canada. We make the case that this is the result of differences in skill transferability, since our results hold after controlling for years of potential domestic and foreign labour market experience, quality of universities in the immigrant's country of study, whether the individual is employed in an occupation that is regulated or licensed in their province, as well as one-digit NAICS industry codes and two-digit NOC occupation codes.

their skill levels from their skill transferability without some measure of their ex ante skills.

[3]An individual's expression skills consist of their written and oral communication abilities. This is in contrast to comprehension skills, which consist of their ability to *understand* written and oral communication.

While our findings may not be surprising, the main contribution comes from showing that a Canadian university education is associated with more highly valued communication skills, even after accounting for potential confounding factors such as quality of education. Furthermore, our finding that the advantage to having studied in Canada is strongest for immigrants from non-English/French speaking countries brings empirical evidence in support of claims that a Canadian university education facilitates stronger language skill transferability.

We note that discrimination and selection issues are less of a concern in our analysis because our entire sample consists of immigrants. To address potential concerns about immigrants selecting into receiving a Canadian education, we use propensity score weights to account for observed differences between the immigrants educated abroad or in Canada. We also perform sensitivity analysis on a sample of refugees, whose immigration decisions are not based on their skills or expected returns in the labour market, and we document a similar result: that immigrant refugees educated abroad have lower returns to certain skills, especially those related to communication, compared with immigrants educated in Canada.

The rest of the paper is organized as follows. Section 1.2 describes our data, Section 1.3 describes how we measure skills, Section 1.4 estimates the contribution of skill transferability to the earnings gap, Section 1.5 includes the discussion of robustness checks, and Section 3.5 concludes. Robustness analysis results are presented in the appendix.

1.2 Data and descriptive statistics

1.2.1 Basic Statistics

We use a combined data set consisting of the 2006 Census, the 2011 National Household Survey (NHS), and the 2016 Census Master Files, obtained from a Statistics Canada Research Data Centre (RDC).

We restrict our sample to include only immigrants who became landed immigrants when they were 18 years of age or older and were employed in the year prior to the survey. We exclude immigrants who landed while they were under the age of 18 since they are relatively more likely to have studied in Canada, and are expected to possess unobservable characteristics that are closer to those of natives than the unobservable characteristics of a typical international student. We further restrict the sample to individuals who have a bachelor's degree as their highest level of education. We exclude those who hold post-graduate and professional degrees, because in the data we are unable to observe where individuals with professional or graduate degrees received their undergraduate education.[4] We also exclude immigrants with a post-secondary degree or certificate below a bachelor's degree, due to concerns that differences in education quality both within and across countries may vary substantially for these programs in comparison to a bachelor's degree program. Nevertheless, we provide sensitivity analysis from estimating the model on a subsample of immigrants who have completed college programs, for whom we find that the location of programs has little effect (due possibly to the shorter duration of college programs), as well as subsample of immigrants who have completed graduate degrees, for whom we find similar results as for the sample of bachelor graduates. We further exclude immigrants who studied in a country other than Canada or their country of birth.[5] Individuals who reported that they are an immigrant but were born in Canada are also excluded, as are those rare individuals whose location of study is unclear.

Table 1.1 contains a set of basic descriptive statistics of our weighted sample of 946,955 observations, 23.2% of whom received their bachelor's degree in Canada.[6] It is immediately apparent that there is a large gap in the unconditional earnings between Canadian and foreign educated immigrants, with Canadian educated immigrants earning approximately 22% more per week than immigrants educated in their home country. There is also a notable difference in the age of immigration to Canada, with immigrants who studied in Canada being on average nearly 8 years younger when they became landed immigrants. This is most likely the result of many of these individuals having been international students who subsequently became landed immigrants in the years immediately following completion of their education. However, there is little difference in current age between these two groups of immigrants, which results from more

[4] An immigrant who obtained both their graduate and their undergraduate degrees from Canadian institutions would have received a different level of exposure to Canadian education than an immigrant who obtained a graduate degree in Canada after having completing an undergraduate degree abroad.

[5] For instance, an immigrant who was born in China but studied in the United States could have been raised in China, and later chose to attend university in the United States. However, it is also possible that they were raised in the United States after having migrated with their family at a very young age, and later moved to Canada as an adult. In our data, we are unable to tell the difference between these two types of immigrants.

[6] As per Statistics Canada requirements, all sample sizes reported in this paper are estimated using survey weights. These are estimates of the total size of the population, and do not necessarily correspond to the number of individual observations in the subsample. While we are unable to report unweighted sample sizes, we encourage readers to be aware of the fact that the 2006 Census and 2016 Census are 20% samples of the Canadian population, while the 2011 National Household Survey was sent to a 33% sample of the Canadian population, with a response rate of 68.6%.

Table. 1.1. Mean (Standard Deviation) Levels of Characteristics For Immigrants with a Canadian Degree or a Foreign Degree

Variable	Studied in Canada	Studied Abroad
Ln(Weekly Earnings)	7.10	6.88
	(0.92)	(0.96)
Age	42.65	44.40
	(11.28)	(9.84)
Age at Immigration	25.32	33.08
	(6.01)	(7.40)
Domestic Experience (years)	16.52	11.18
	(11.09)	(8.76)
Foreign Experience (years)	4.13	11.16
	(5.24)	(7.31)

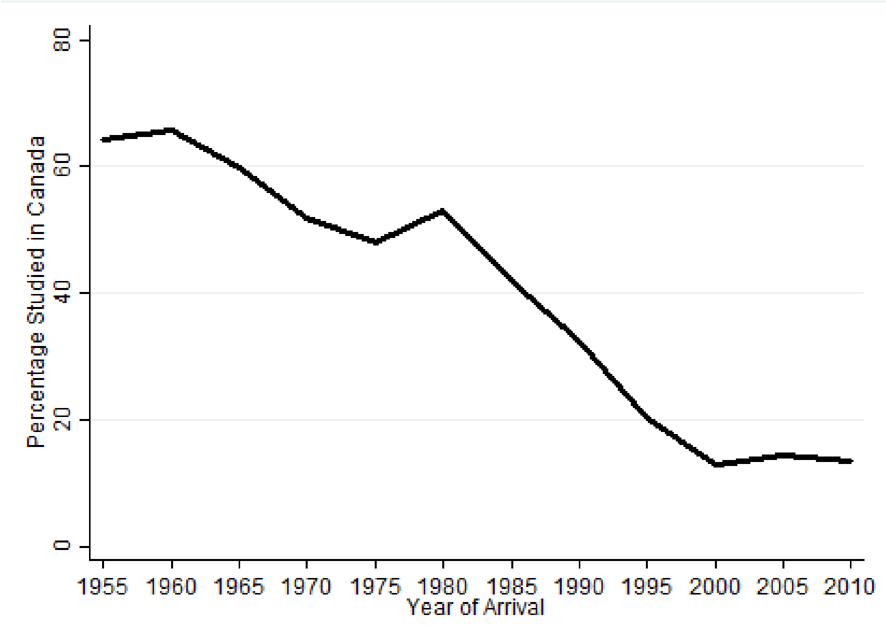
Source: Census 2006, Census 2016 and NHS 2011.
Weighted Sample of 946,955 immigrants who had landed after age 18, employed, with a Bachelor's education. The proportion who studied in Canada is 23.2%.
Actual sample size is about 20% of weighted sample size, see footnote 6.

recent immigrants being less likely to have studied in Canada.[7] We impute potential years of domestic and foreign labour market experience as, respectively, age minus age at immigration, and age at immigration minus 22. For immigrants who landed below the age of 22, their foreign experience is zero and their domestic experience is current age minus 22.

In sensitivity analysis, we illustrate the percentage of immigrants from each entry cohort who obtained their bachelor's degree in Canada. We do two types of exercises: (i) we look back at broad cohorts of immigrants by half-decade arrivals, as well as (ii) immigrants who have arrived more recently in a given year: 1990, 1995 or 2000. In the first exercise, entry cohorts are separated into half decades: early sixties indicate immigrants who landed from 1960-1964 inclusive, while late sixties indicate 1965-1969 inclusive; other decades follow the same pattern, and the earliest cohort consists of immigrants who landed in the late 1950s. We note that there is a clear downward trend in the percentage of immigrants who received their degree in Canada, as shown in Figure 1.1. However, Table 1.2 shows that for a given year of arrival, the percentage of immigrants with a Canadian Bachelor's degree increases across surveys during the first decade after landing. However, this percentage is more stable for

[7] While we will return to a more formal discussion of differences between those educated in Canada and abroad, the age and cohort implications are reflected in the probit coefficients for participation in Table A1.1 of the Appendix, showing a pattern of more recent cohorts being less likely to have received their degree in Canada.

Fig. 1.1. Percentage of Immigrants Who Studied in Canada: By Year of Arrival Cohort



earlier arrival cohorts[8]. This indicates that the downward trend observed across cohorts from the late 1950s until the late 1990s is likely to be capturing a decline in the likelihood that long-term immigrants obtained their degree in Canada, as opposed to differences in emigration rates of these immigrants, or differences in pursuing further education in Canada.

Table. 1.2. Percentage of Immigrants Who Studied in Canada, By Cohort and Survey Year

Year of Arrival	2006 Census	2011 NHS	2016 Census
1990	32.1%	33%	32.7%
1995	22.8%	25.5%	25%
2000	13.2%	18.3%	15.3%
2002	9.4%	17.1%	14.5%

The distribution of birth places for immigrants, as well as the percentage of immigrants from a given region who studied in Canada, can be found in Table 1.3. By far the largest group of immigrants comes from South and South-East Asia (India, Pakistan, Bangladesh, Indonesia, Philippines and other nearby countries), with East Asia (China, Mongolia, Japan, South Korea, Taiwan) also being a major source region. The percentage of immigrants who studied in Canada differs substantially across regions of origin. More immigrants from Sub-Saharan Africa received their degree in Canada than in their country of origin, the only region

[8] It is noteworthy that the percentage of immigrants who studied in Canada is persistently highest in the 2011 NHS subsample. This is particularly notable for cohorts who arrived in years 2000 or 2002. It is probable that this is a reflection of higher response rates to the voluntary NHS survey among immigrants who studied in Canada.

Table. 1.3. Birthplaces of Immigrants, Percentages by Location of Study

Location of Birth	Percentage Born in Location	Percentage Studied in Canada
Oceania	0.6%	18.5%
East Asia	19.3%	34.9%
South and South-East Asia	37.3%	11%
Sub-Saharan Africa	4.1%	51.5%
USA	0.7%	26.5%
Latin America	8.5%	32.4%
Middle East	10.55%	26.4%
Eastern Europe	9.7%	16.8%
Western Europe	6.6%	31.4%

Source: Census 2006, Census 2016 and NHS 2011.
In the first row, 0.6% indicates that 0.6% of the immigrants in our sample were born in Oceania; 18.5% indicates that, out of those immigrants who were born in Oceania, 18.5% of them received their Bachelor's degree from a Canadian university.

for which this is true. At the other end of the spectrum, immigrants from South/South-East Asia are the least likely to have received their degree in Canada, with immigrants from this region being approximately 9 times more likely to have received their degree in their home country as opposed to Canada.

1.2.2 Other Covariates

We include $RegulatedOccupation_i$ as an indicator of whether an individual's occupation is regulated (requires a licence to practice) in their province of residence. Data on provincial licensing requirements is obtained from The Canadian Information Centre for International Credentials[9]. We include this control in order to mitigate any bias that may result from a combination of regulated occupations carrying an earnings premium (Kleiner and Krueger 2010, Plesca et al. 2015), and immigrants with foreign credentials experiencing greater difficulty in obtaining the necessary licenses (Zietsma et al., 2010). 24.7% of the immigrants in our sample who studied in Canada are employed in a regulated occupation, as opposed to 20.4% of immigrants who studied in their home country.

We furthermore include $Citizen_i$ as a dummy variable indicating whether the individual is a Canadian citizen, since this may proxy for labour market attachment, and is related to earnings growth (Bratsberg et al. 2002). In our data, 81.8% of immigrants who studied in Canada are

[9] For more information on which occupations require a licence to practice in each province, visit https://www.cicic.ca/935/perform˙an˙advanced˙search˙in˙the˙directory˙of˙occupational˙profiles.canada

Canadian citizens, as opposed to 65.3% of immigrants who studied in their home country.

1.2.3 University quality

We construct $\mathit{UniversityQuality}_i$ as a measure of university quality for each country consisting of its number of top 1000 universities per 100,000 post-secondary graduates[10], and we then assign this value to each individual in our sample who studied in that country. Our data on the top 1000 universities is obtained from the 2016 edition of The Center for World University Rankings (CWUR), which ranks universities on the basis of both the quality of research coming from the institution and the outcomes of its prominent alumni.[11] Higher values of this variable correspond to the individual having studied in a country with more high quality institutions, relative to the total size of the post-secondary educated population. We include this measure in order to address concerns that differences in the quality of university education (reflected in different skill levels) across countries affect our results. We use this measure instead of other potential education quality measures based on test scores (such as PISA or PIAAC), since these are more likely to reflect differences in the quality of primary or secondary education, as opposed to differences in the quality of a university education. For immigrants who received their bachelor's degree outside of Canada, the mean level of this university quality score is 0.21, with a standard deviation of 0.62. The value of the university quality score for immigrants who studied in Canada is 0.57.[12]

1.3 Measuring skills

We construct a measure of an individual's skills using data from O*NET, a U.S. Department of Labor database which is a continuation of the Dictionary of Occupational Titles. O*NET contains information on the expected individual characteristics of workers for each of the 840 occupations that are classified by the Standard Occupational Classification (SOC). Most researchers use O*NET to infer an individual's skills directly from their occupation, under the

[10] We estimate the number of post-secondary graduates in a given country by taking World Bank data consisting of the percentage of individuals aged 25+ who have tertiary education, and multiplying it by the estimated population aged 25+, taken from United Nations data.

[11] For a detailed description of CWUR's methodology, see http://cwur.org/methodology/world-university-rankings.php

[12] Note that it does not vary among Canadian educated immigrants since there is a only a single value of the score per country

assumption that workers are sorted across occupations according to their skills. While this is an appropriate methodology for many research questions, it is not appropriate for ours, since we would expect workers to be sorted across occupations according to their levels of skills as they are valued in the Canadian labour market. This makes it impossible to separate their *ex-ante* skills from the portion of their skills that are transferable to Canada. As a result of this problem, we would be unable to measure the transferability of skills for immigrants who studied abroad compared to those who studied in Canada.

To circumvent this issue, we match an immigrant's field of study to those occupations for which the field prepares an individual, even if they do not necessarily correspond to the individual's actual occupation. Since our focus is on estimating the degree to which a Canadian university education affects the value of skills in the Canadian labour market (relative to a university education in one's home country) these skills that are associated with graduates of an individual's program are the relevant skills that we are looking to capture.

This mapping of fields of study to skills is accomplished by matching each field of study to one or more occupations for which it prepares an individual by using the CIP-SOC crosswalk developed by the National Center for Education Statistics (NCES) and the Bureau of Labor Statistics (BLS). The skills of this occupation are then taken to be the skills of an individual who studied the field in question. For cases in which the CIP-SOC crosswalk matches a field to more than one occupation, we take the mean of each skill measure across these occupations to be the skills of someone who studied that field. As shown in Table 1.4, nearly half of all fields of study are linked to a single occupation, while over 90% of fields are linked to 4 occupations or fewer. We take our skill measures from the *Skills, Knowledge, and Abilities* content data contained in the O*NET database (hereafter collectively referred to as skills for simplicity). Each of these dimensions contain both *importance* and *level* scores, which we place on scales of 0-10 followed by taking an equally average of these two scores as our measure of that particular skill.[13]

We are left with 109 different skill measures. It is clearly impractical to use each of them in our analysis. Therefore, we create seven different aggregate skill measures: *expression, comprehension, logical, executive, hard science, technical,* and *physical*. The individual skills

[13]Our choice of the Skills, Knowledge, and Abilities measures, and the choice to take an equally weighted average of level and importance, are largely influenced by Allen et al. 2011, who find that this method leads to scores that are most closely related to an individual's capacity to perform a job.

Table. 1.4. Number of Field-Occupation Matches

(1)	(2)	(3)	(4)
1	792	46.05%	46.05%
2	366	21.28%	67.33%
3	301	17.5%	84.83%
4	136	7.91%	92.73%
5	72	4.19%	96.92%
6	28	1.63%	98.55%
7	10	0.58%	99.13%
8	5	0.29%	99.42%
9	3	0.17%	99.59%
10	3	0.17%	99.77%
12	2	0.12%	99.88%
13	1	0.06%	99.94%
15	1	0.06%	100%

Column (2) indicates the number of fields of study that match to the number of occupations in Column (1). For example, the first row indicates that 792 post-secondary fields of study are matched to a single occupation. Column (3) indicates the percentage of fields of study that match with the number of occupations given in Column (1). Column (4) indicates the percentage of fields of study that match to the number of occupations given in Column (1) (or fewer)

used for each of these aggregate skills can be found in Table 1.5. In each case, these individual skills are added up, and the aggregated skill measure is converted to a 0-10 scale, with 0 representing the lowest level of the score in the data and 10 representing the highest level of the score in the data.

Expression and comprehension are our two measures of social skills, with the former comprising an individual's capacity to express themselves through oral or written communication, while the latter is their capacity to understand information presented orally or through writing. Although researchers have traditionally combined these two measures into a single measure of "social skills", we make a distinction between these two types of skills since it is not clear that an immigrant's ability to transfer their expression skills to the Canadian labour market will be the same as their ability to transfer their comprehension skills. Linguists researching second language acquisition frequently make a distinction between these two capacities, and the existing evidence suggests that practicing language comprehension results in limited language expression abilities (Swain 1985, DeKeyser and Sokalski 1996, Wen 2018). Furthermore, formal second language instruction (which in many cases will be the main form of English/French experience that foreign educated immigrants will have) generally emphasizes comprehension over expression (Swain 1985, Wen 2018). Since studying in Canada will afford individuals greater opportunities at practicing expression in English and/or French than they would have obtained studying abroad, we anticipate that these immigrants will have a greater advantage over foreign educated immigrants in transferring their expression skills to Canada, relative to their comprehension skills.

Logical skills primarily capture an individual's reasoning skills, which is a form of fluid intelligence, as opposed to hard science which contains crystallized knowledge of sciences such as chemistry or physics. Technical skills are an individual's capacity to understand and utilize machines and other forms of technology. Physical skills relate to an individual's capacity to perform manual tasks. Our final skill measure, executive skills[14], primarily comprises one's ability to manage time and resources.

We use these seven skills primarily because they have clear interpretations, which is an advantage over using skills derived from methods such as principal components. Principal Components are, however, used in sensitivity analysis as a robustness check. Furthermore, we utilize these seven skills over the commonly used cognitive/non-cognitive or routine/non-routine classifications since we anticipate that skill transferability will vary across more specific skills within each of these categories, which would reduce the clarity of our results. For example, cognitive non-routine skills would incorporate components of both expression and logical skills, however we anticipate that immigrants who studied in their home country will have greater difficulty in transferring their expression skills.

One potential concern with our methodology is that fields of study may be different in Canada in comparison to elsewhere. However, many of these differences likely arise from location specific content (including language of instruction) rather than an entirely different set of skills. It is these differences that are associated with difficulty transferring skills to Canada, and we are looking to capture these differences rather than eliminate them. However, should there be cross-country differences in the *types* of skills that are associated with a given field, this would create bias in our estimates of skill transferability. We attempt to address this concern by including the university quality measure previously described in the data section. If fields of study across countries with institutions of comparable quality involve similar broad sets of skills (even if location specific content may differ), then this variable should address such concerns. We find no notable change in our results when including this control.

[14]We think of executive skills as being short for *executive functioning skills*, rather than referring to executive as an occupation.

Table. 1.5. Individual Skills from O*NET That Comprise Our 7 Skill Measures

Hard Science	Logical	Executive
Chemistry	Deductive Reasoning	Critical Thinking
Physics	Programming	Judgment and Decision Making
Engineering and Technology	Inductive Reasoning	Systems Analysis
Computers and Electronics	Mathematical Reasoning	Systems Evaluation
Science	Information Ordering	Time Management
	Flexibility of Closure	Management of Financial Resources
	Number Facility	Management of Material Resources
	Mathematics	Management of Personnel Resources
	Biology	Monitoring
Comprehension	**Physical**	**Technical**
Written Comprehension	Dynamic Flexibility	Equipment Maintenance
Active Listening	Dynamic Strength	Equipment Selection
Oral Comprehension	Explosive Strength	Installation
Reading Comprehension	Extent Flexibility	Operation and Control
	Gross Body Coordination	Operation Monitoring
	Gross Body Equilibrium	Quality Control Analysis
	Stamina	Repairing
	Static Strength	Technology Design
	Trunk Strength	Troubleshooting
	Arm-hand Steadiness	
	Control Precision	
	Finger Dexterity	
	Manual Dexterity	
	Multilimb Coordination	
	Rate Control	
	Reaction Time	
	Response Orientation	
	Speed of Limb Movement	
	Wrist-finger Speed	
	Expression	
	Written Expression	
	Oral Expression	
	Instructing	
	Persuasion	
	Speech Clarity	
	Speaking	
	Negotiation	
	Sales and Marketing	
	Customer and Personal Service	

For more information on the O*NET Content, please see https://www.onetcenter.org/content.html#cm1 and various subpages under the heading of "Abilities", "Skills" and "Knowledge"

1.4 Estimating the Skill Transferability of Immigrants

1.4.1 Propensity Score Re-Weighting

To mitigate concerns that our results are driven by selection into education in Canada or abroad, and to address covariate imbalance, we estimate all of our specifications under the Conditional Independence Assumption that individuals select into obtaining their degree abroad or in Canada based on characteristics observable in the data[15]. All our main results in Table 1.6 are obtained using selection weights, while in the Appendix Table A1.2 we provide the unweighted results for comparison.

We estimate the propensity that a given immigrant studied abroad using a probit estimation of the form $Pr(Abroad_i = 1|X_i)$, where the treatment indicator is "obtaining a Bachelor's degree abroad" and X_i consists of the full set of controls discussed above, given these X_i are likely to determine both the outcome (earnings) and the selection into treatment (being educated abroad). Table A1.1 in the appendix provides the probit coefficients from the propensity score estimation. As expected from the discussion in the data section, individuals with a Canadian degree are more likely to be of Asian descent, Canadian citizens, from countries with lower quality of education, from earlier cohorts, and working in regulated occupations. Individuals with degrees from abroad are more likely to be married and live in large urban centres (with the exception of Ottawa, which is associated with an increased likelihood of having a Canadian degree).

We re-weight the observations in our sample by the inverse of the probability that individuals were assigned to their observed treatment, and drop observations outside of common support by trimming estimated probabilities below 2% or above 98%. We trim these probabilities for both groups, the one educated in Canada and the one educated abroad. We further investigate the sensitivity of our results to the particular cut-off probabilities of 2% and 98% by re-estimating our main model using three other sets of trimming cut-off probabilities: below 0.5% or above 99.5%, below 1% or above 90%, and below 5% or above 95%. This analysis is reported in Table A1.3 and indicates that the main story on skill transferability is not sensitive to particular

[15] With regards to selection on unobservables, our analysis on a subsample of refugees in Table 1.7 mitigates the possibility that it is driving our results.

cut-off values.[16]

As long as the observed covariates capture the selection into the treatment mechanism, the propensity score re-weighting is correcting the main estimates for selection into treatment. We find evidence that the re-weighting procedure successfully balances covariates across the two groups, the $Abroad_i = 0$ and $Abroad_i = 1$ subsamples, as shown in Table A1.4. Propensity score re-weighting reduces the gap in the weighted means between the covariates in the $Abroad_i = 0$ and $Abroad_i = 1$ subsamples, confirming that propensity score re-weighting makes the observed characteristics of the two groups comparable[17]

1.4.2 Main Estimates

To assess the contribution of differences in skill transferability to the earnings gap, we estimate the following model using OLS:

$$\begin{aligned} log(WeeklyEarnings_i) = \beta_0 + \beta_1 \quad X_i + \beta_2 RegulatedOccupation_i \\ + \beta_3 UniversityQuality_i + \beta_4 Citizen_i + \beta_5 Abroad_i + \beta 6 \quad Skill_i \\ + \beta 7 \quad Abroad_i \quad Skill_i + \quad _i \end{aligned} \tag{1.1}$$

The dependent variable is log of average weekly earnings, with weekly earnings being calculated as annual earnings divided by weeks worked in a year. X_i is a vector of baseline control variables. Controls include indicators for ethnicity and region of birth in order to mitigate concerns that our results can be attributed to discrimination; we include the 9 regions of birth as shown in Table 1.3, with immigrants born in the USA being the omitted region and "arab" the omitted ethnicity. The remaining covariates include dummy variables indicating whether the individual is; female, working part-time, has a child between the ages of 0 and 5 (and it's interaction with female), has a child between the ages of 6-14 (and it's interaction with female), married or in a common law relationship, a native English speaker, or a native French

[16] The population size estimates reported for this sensitivity analysis are not indicative for the fraction of the sample trimmed, because the observations discarded tend to have smaller survey weights.

[17] As a further check for the correction procedure, we re-estimate the original probit model from Table A1.1 using the propensity score weights, and report this estimation in the last two columns of the same table. We find, as expected, that the pseudo R^2 of the model drops from 0.2712 to 0.0075, indicating that after balancing the covariates between the two immigrant groups, these covariates no longer have much explanatory power in predicting which subsample a given immigrant is in.

speaker. Also included are sets of dummy variables for; province of residence, major Census Metropolitan Area of residence, and survey year.Years of domestic and foreign labour market experience, along with their quadratic and cubic terms, are also included in X_i.

$Abroad_i$ is an indicator taking on the value of 1 if the individual received their bachelor's degree in their home country and 0 if they received the degree in Canada. One of our primary parameter of interest is the coefficient β_5 on $Abroad_i$.

Columns 1 and 2 of Table 1.6 report results from estimating a reduced form of this model. Both of these specifications account for the demographic characteristics in X_i, along with $RegulatedOccupation_i$, $UniversityQuality_i$, and $Citizen_i$. The second specification also adds our skill measures, but these two specifications do not contain the interaction terms between abroad and skill, so no heterogeneity in the returns to skills is allowed. We report in Table 1.6 only the main coefficients of interest for our story, while full estimation results are available in Table A1.5 of the Appendix.

The coefficient on $Abroad_i$ in column 1 shows that immigrants who studied abroad continue to earn about 9.5% less than their Canadian educated counterparts even after accounting for a rich set of controls, including university quality, different rates of employment in regulated occupations, differences in domestic experience, etc. Further analysis focuses on the role that skills play in explaining this remaining portion of the earnings gap.

The remaining coefficients reported in Appendix Table A1.5 are consistent with similar findings in the literature. We find that the earnings of immigrant cohorts arriving since the late 1960s have been deteriorating. However, we also find early evidence of a recent reversal of this trend, since our omitted group of immigrants who arrived between 2010-2014 outperform their counterparts who arrived between 2005-2009, although they are still receive lower earnings than immigrants from all pre-2005 cohorts. Furthermore, also consistent with existing literature (Aydemir and Skuterud 2005, Green and Worswick 2017), we find that years of domestic labour market experience are more highly valued in the labour market than years of foreign labour market experience.

The coefficient on "University Quality" is positive and statistically significant. It indicates that an individual is expected to receive an earnings premium of approximately 9.2% for each additional top 1000 university per 100,000 post-secondary graduates in their country of study.

Table. 1.6. Returns to Skill for Immigrants Educated Abroad or in Canada

	(1) Basic	(2) With Skill	(3) With Interact.	(4) Includes Occup.
Abroad	-0.095***	-0.105***	0.241**	0.161
	(0.008)	(0.008)	(0.103)	(0.098)
Expression		0.006	0.034**	0.027**
		(0.008)	(0.013)	(0.013)
Comprehension		-0.02***	-0.03**	-0.023*
		(0.007)	(0.013)	(0.0125)
Logical		0.025***	0.04***	0.022***
		(0.005)	(0.009)	(0.008)
Hard Science		0.004	-0.001	0.008
		(0.003)	(0.005)	(0.005)
Executive		-0.01	-0.014	-0.015
		(0.006)	(0.011)	(0.011)
Technical		0.043***	0.047***	0.033***
		(0.005)	(0.008)	(0.008)
Physical		-0.042***	-0.048***	-0.039***
		(0.005)	(0.008)	(0.008)
Abroad*Expression			-0.059***	-0.052***
			(0.016)	(0.016)
Abroad*Comprehension			0.02	0.022
			(0.014)	(0.014)
Abroad*Logical			-0.029***	-0.022**
			(0.011)	(0.01)
Abroad*Hard Science			0.01*	0.004
			(0.006)	(0.006)
Abroad*Executive			0.012	0.015
			(0.013)	(0.012)
Abroad*Technical			-0.007	-0.001
			(0.01)	(0.009)
Abroad*Physical			0.01**	0.013
			(0.01)	(0.009)
Weighted N	936,835	936,835	936,835	936,835
R^2	0.2233	0.2294	0.23	0.2859
Nb. covariates	68	75	83	146

Sample after trimming 2% at the top and bottom of the propensity score distribution. Actual sample size is about 20% of weighted N sample size, see footnote 6.
Standard errors in parentheses. Each of these models includes the full set of demographic controls that are included in Table A1.4 of the Appendix. Column 4 includes a set of indicators for two digit NOC occupations. Asterisks denote significance at the 1%, 5% or 10% level.

We also find that immigrants who are employed in regulated occupations earn approximately 14% more than immigrants who are employed in other occupations.[18] Immigrants who are Canadian citizens receive an earnings premium of approximately 8.4% relative to non-citizen immigrants.

The returns to skills are in many cases what one would anticipate them to be. STEM oriented logical and technical skills carry positive and significant returns, while physical skills carry negative returns. For language based expression and comprehension skills, we find that they have opposite signs, with expression skills carrying a premium, while comprehension skills carry a penalty. We find that hard science and executive skills are not valued in the Canadian labour market, conditional on other skills.

Regarding our primary parameter of interest, controlling for *ex-ante* skill levels produces little change in the estimated earnings gap between immigrants who received their bachelor's degree in Canada and immigrants who received their bachelor's degree in their home country. We take this as evidence that there is little difference between Canadian educated immigrants and foreign educated immigrants in terms of the types of skills that they possess. This indicates that Canadian educated immigrants do not receive higher earnings because they were more likely to have studied fields that are more highly valued in the Canadian labour market. However, even though these two groups of immigrants possess similar *ex-ante* skills, this does not preclude differences in skill transferability.

In the last 2 specifications of Table 1.6 we investigate the extent to which the returns to various skills differ between immigrants educated in the source country versus Canada, which we attribute to skill transferability. Immigrants who studied in their home country receive lower returns to expression and logical skills than immigrants who studied in Canada. Most notable is the gap in returns to expression skills, which indicates that it is the skill that foreign educated immigrants have the most difficulty with transferring to Canada, potentially because Canadian educated immigrants have stronger English/French language abilities that allow to them to effectively use their expression skills. Being educated abroad becomes positive and statistically significant (albeit with a large standard error). This indicates that immigrants who studied abroad receive a premium over their Canadian educated counterparts on the basis of charac-

[18] When controlling for industry using one-digit NAICS codes, this earnings advantage drops, but remains consistent with the notion that workers in regulated occupations earn rents.

teristics other than these skills, but that this premium is masked by lower returns to logical and especially expression skills in a model that does not allow for skill return heterogeneity.

We note a large difference in transferability between expression and comprehension skills. This is consistent with the hypothesis that foreign educated immigrants have language comprehension capabilities that are comparable to their Canadian educated counterparts (as evidenced by the lack of statistical significance on the abroad*comprehension coefficient), but have much weaker language expression capabilities. To the best of our knowledge, this is the first instance in the immigration literature in which the distinction between these two types of skills has been made.

Previous work has found that immigrants pursue different labour supply strategies than natives (Duleep and Sanders 1993, Baker and Benjamin 1997). To address the possibility that differences in occupation/industry composition between Canadian educated and foreign educated immigrants may be driving our results, due to differential selection into occupations, we include in the 4^{th} specification one-digit NAICS industry codes and two-digit NOC occupation codes. Nevertheless, the results do not change substantively, possibly because we already control for regulated occupations.

For sensitivity purposes, we replicate our analysis on a subsample that includes only refugees, since this subsample is least likely to suffer from issues related to selection into education abroad or in Canada, or from issues of return migration. This is only possible for the 2016 Census, since the earlier surveys do not contain information on immigration class. Furthermore, only immigrants who landed after 1980 are categorized by immigration class. Due to the availability of only a single wave of survey data for refugees, we are unable to control for cohort effects in this model, as we do in all other models. These results can be found in Table 1.7 and confirm the findings regarding skill transferability: that refugees who studied abroad receive significantly lower returns to their expression skills than refugees who studied in Canada.

1.5 Robustness Checks

In what follows, we present sensitivity results which help us better understand the mechanism behind our transferability findings. While this discussion is in the main text of the paper, the results are in the appendix.

Table. 1.7. Results For Refugees

	Coefficient	Standard Error
Abroad	-0.524	0.422
Expression	0.129**	0.062
Comprehension	-0.078	0.048
Logical	-0.03	0.038
Hard Science	0.011	0.022
Executive	-0.079	0.053
Technical	0.119***	0.04
Physical	-0.002	0.036
Abroad*Expression	-0.215***	0.077
Abroad*Comprehension	0.154***	0.058
Abroad*Logical	0.021	0.045
Abroad*Hard Science	-0.011	0.027
Abroad*Executive	0.143**	0.062
Abroad*Technical	-0.096*	0.051
Abroad*Physical	0.013	0.045
Weighted N	21,455	
R^2	0.2151	

This subsample consists only of immigrants who reported arriving in Canada as refugees. Since immigration class is available only in the 2016 Census, we are unable to separate cohort effects from domestic experience effects, and therefore cohort dummies are omitted from this specification. Otherwise, this model is equivalent to the main specification from column 3 of Table 1.6.
Actual sample size is about 20% of weighted N sample size, see footnote 6.
Asterisks denote significance at the 1%, 5% or 10% level.

1.5.1 Language Transferability

Our finding that immigrants who studied abroad receive lower returns to expression skills than immigrants who studied in Canada is consistent with the hypothesis that Canadian educated immigrants possess stronger English/French language abilities than their foreign educated counterparts. However, if this is the case, the difference in transferability should be most pronounced among immigrants from countries in which neither English nor French is a predominant language. Indeed, we should expect that immigrants whose home country is English or French speaking will receive comparable returns to expression skills whether they studied in their home country or in Canada. Should the penalty associated with having studied abroad be similar for immigrants from English/French speaking countries as it is for immigrants from non-English/French speaking countries, this would be indicative that factors other than language (perhaps networking opportunities) are driving our results. As a check to ensure that our claim that our results are driven by language differences is true, we include a set of interactions between our skill measures and whether the immigrant is from a country that is primarily English or French speaking. The full set of countries that we consider to be English/French speaking can be found in Table A1.6 in the appendix.

The first column of Table A1.7 in the appendix reports the results from this estimation. We find immigrants who studied in their non-English/French speaking home country experience returns to their expression skills that are approximately 6.8% lower than the returns of their Canadian educated compatriots. However, for immigrants who are from English/French speaking countries, those who received their degree at home experience returns to expression skills that are not statistically different from zero (F-Statistic on the interaction term is 0.23 with a corresponding p-value 0.6309), with a point estimate indicating the returns as being approximately 2.5% lower than those of their Canadian education counterparts. This is consistent with our hypothesis that immigrants who received their bachelor's degree from a Canadian university receive higher returns to their expression skills as a result of the greater exposure to an English/French language environment.

1.5.2 Heterogeneity By University Quality

While our earlier models have controlled for differences in the distribution of university quality, the results in column 2 of Appendix Table A1.7 add interactions between our 7 skill measures and university quality. We add these out of concern that a higher quality education may be associated with a stronger ability to transfer skills to Canada, on top of the unconditional earnings advantage of a higher quality education. Including these interactions fails to explain the differences in skill returns between Canadian educated immigrants and foreign educated immigrants.

1.5.3 Heterogeneity Over Time

Since assimilation has been capturing attention in the immigration literature, we include in column 3 of Appendix Table A1.7 a set of interactions of domestic labour market experience with our 7 skill measures. We find that having spent more time in Canada fails to account for Canadian educated immigrant's advantage in skill transferability. Furthermore, in column 4 we add interactions between our skills, domestic labour market experience, and educated abroad, in order to assess whether the gap in returns to skills dissipates over time. Our results show that the gaps in returns to skills between Canadian educated and foreign educated immigrants are persistent over the lifecycle. This suggests that any extra pre-migration time that Canadian educated immigrants have spent in the country is unlikely to be the source of their earnings advantage. Rather, it appears as though attending a Canadian university in and of itself results in their advantage with transferring their skills to the Canadian labour market.

We include interactions between our skill measures and age at immigration in order to address concerns that individuals who immigrate at an older age experience greater difficulty with transferring their skills to the Canadian labour market. These results are in column 5 of Table A1.7. We find evidence of a modest reduction in the penalty on expression skills for immigrants educated abroad, and a small penalty on expression skills for immigrants who arrived at older ages. This indicates that a small portion of the lower returns to expression skills for immigrants educated in their home country can be attributed to the fact that these immigrants tend to arrive at an older age. Nevertheless, a substantial penalty to expression skills remains even after accounting for this.

1.5.4 Education Levels Other Than Bachelor's

We estimate our model on two additional subsamples of immigrants with different levels of education: *college* consists of immigrants who completed college programs of one or two years in length, while *grad* consists of immigrants who have completed a post-graduate degree (Master's or PhD). The results for these models can be found in Table A1.8 in the appendix. We find that immigrants who received a graduate degree in their home country receive lower returns to their social expression skills than immigrants with a Canadian graduate degree, consistent with our results for Bachelor's degree graduates. With regards to College graduates, we do not find any difference in the returns to expression skills for immigrants educated in Canada relative to immigrants educated abroad. We argue that this is not inconsistent with expectations, primarily because we expect that College graduates will have completed programs that are much shorter in duration than the 4-5 years typically required in order to obtain a Bachelor's degree. If exposure to an English/French language environment facilitates transferability of expression skills, program length should be an important determinant of the gap in skill returns between immigrants educated in Canada and immigrants educated abroad.

Our ability to estimate the skill transferability of immigrants is confined to those immigrants who have some post-secondary education, since there is no field of study for individuals with high school education or below. However, in trying to infer the skill transferability at lower levels of education, we estimate a version of our model that includes immigrants with an age of arrival between 13-17 years old, since these immigrants are likely to have completed at least part of their high school education in Canada. We compare immigrants in this subpopulation who went on to complete a bachelor's degree in Canada with those immigrants who arrived aged 18+ and also completed a bachelor's degree in Canada. The results of this analysis are reported in the bottom panel of Table A1.8, and indicate very similar results between the two groups in terms of skill transferability, suggesting that it is the highest level of education which matters for skill transferability.

1.5.5 Lived in Canada For More Than 10 Years

Since we do not follow immigrants across time, we we cannot bring direct evidence as to who from our sample will return to their home country or emigrate to a third country. We can nevertheless bring suggestive evidence by looking at a subsample of immigrants who are still in Canada ten years after originally immigrating.[19] While the proportion of immigrants who studied in Canada increases substantively with recently arrived cohorts, this is most pronounced in the first years after arriving in Canada, as reflected by the smaller increases for earlier years of arrival found in Table 1.2[20].This indicates that, should there be any differences in emigration rates (or return migration rates) between Canadian educated immigrants and immigrants educated in their home country, these differences should occur primarily in the first years post-arrival. When we re-estimate the main specification using a subsample of immigrants who have been in Canada for at least a decade (results in Table A1.9 in the appendix), we find that immigrants who studied abroad receive similarly low returns to their expression skills as the overall group of immigrants who studied abroad.[21]

Moreover, we can use these results to address a slightly different issue related to labour force attachment. Immigrants educated in Canada have a lower unemployment rate compared to immigrants educated in their home country (6.5% compared to 7.2% in the 2016 Census). However, this difference is largely confined to recent immigrants. Among immigrants who have been landed immigrants to Canada for at least 10 years, those who were educated in Canada have an unemployment rate of 5.5% (labour force participation rate 72.4%), while those who were educated in their home country have an unemployment rate of 5.6% (participation rate 72.3%). While there are differences in unemployment and participation in the overall immigrant sample, these differences disappear for the subsample of immigrants who have been in the country for a decade or more. Because the sensitivity analysis reported in Table A1.9 indicates that the skill transferability conclusions are similar no matter whether we use the entire sample or the subsample who have been in the country longer, it suggests that, despite potential differences in unemployment and participation rates between the two groups of immigrants, these differences are not driving our results with regards to skill transferability.

[19] We are grateful to an anonymous referee for raising these questions.

[20] The persistently higher percentage of Canadian educated immigrants in the voluntary NHS survey is part of the motivation for analyzing the results without the NHS survey as a robustness check. These results are found in Table A1.10. The results reported in this Appendix table are similar to those from our full sample

[21] Another argument supporting our main transferability conclusions can be made from the analysis on refugees reported in Table 1.7. Refugees are potentially less susceptible to return migration, yet we find similar results on the sample of refugees.

1.5.6 Differences Across Quantiles

We use quantile regressions in order to assess whether our results are confined to a particular part of the earnings distribution. The coefficients from quantile regressions at the 20th, 40th, 60th, and 80th quantiles are in Appendix Table A1.11. A similar representation is in Figure A1.1 in the Appendix, which shows the coefficients and 95% confidence bands for $\text{Abroad}_i \quad \text{Skill}_i$ at each decile of the conditional earnings distribution. We find that foreign educated immigrants receive a penalty to their expression and logical skills across the entirety of the conditional earnings distribution, with these penalty being larger towards the tails of the distribution. We also find evidence that immigrants who studied abroad receive a premium over their Canadian educated counterparts at higher quantiles of the conditional earnings distribution, but there is no evidence of such a premium at lower quantiles of the distribution.

1.5.7 Principal Components In Defining Skill Measures

As an additional method of aggregating skills, we take the first 3 principal components of the 109 O*NET skills[22]. The eigenvectors for these principal components can be found in Table A1.12 of the appendix. The components are sorted on the basis of the weights for the first principal component, which is primarily language based. We report the results using principal components as skills in Table A1.13 of the appendix. Immigrants who received their bachelor's degree abroad receive substantially lower returns to the 1^{st} principal component, and modestly lower returns to the 3^{rd} principal component, than immigrants who studied in Canada. This is consistent with our other results, since it implies lower returns to language based skills for immigrants who studied abroad.

1.6 Conclusion

Looking at Canadian immigrants, we have shown that immigrants who obtain a Bachelor's degree in Canada are better positioned to transfer their social expression skills to the Canadian labour market compared to immigrants educated abroad. We have proposed a novel methodology to infer skill from a link between field of study and a matrix of competencies in O*NET. As such, our identification of skill and skill transferability only works for individuals with an

[22] In addition, we have estimated the model using 14 principal components, which collectively account for 90% of the variation in skills.

education above high school who report a field of study. Especially for countries such as Canada and Australia, this is the main immigrant population of interest, given how the point system selects the economic class principal applicant immigrants. Our findings are consistent with Canadian educated immigrants receiving an earnings advantage as a result of greater exposure to English/French environment. However, further work is necessary in order to disentangle any language effects from potential networking effects.

To date, literature on the selection of immigrants has focused on selection differences between immigrants and natives (Chiswick 1999), and differences between immigrants and non-migrants from the same source country (Chiquiar and Hanson 2005). Our work is filling a gap in this literature by focusing on the skill transferability of immigrants, as well as their selection into Canadian education. Evidence from propensity score re-weighting suggests that immigrants educated abroad are positively selected compared to immigrants educated in Canada [23].

Despite our finding that immigrants educated abroad are positively selected, this does not compensate for their disadvantage in transferring their skills to the Canadian labour market, which is reflected in their overall lower earnings. Therefore, we would expect that further prioritizing former international students for immigration to Canada would improve the overall labour market performance of immigrants.

[23] This is evident from the fact that in Table 1.6, the coe□cient on $Abroad_i$ is positive and statistically significant after accounting for skill transferability, as in column (3). It remains positive (albeit no longer statistically significant) after further accounting for 2-digit occupations in column (4). The coe□cient on $Abroad_i$ also remains positive and significant in the majority of robustness checks in which skill transferability is accounted for (see Tables A1.7-A1.11).

Chapter 2: Self-Employment Rates of Immigrants and Natives: Occupations and Skills

2.1 Introduction

Researchers have long noted that immigrants are more likely to be self-employed than natives, a finding that dates back as early as Borjas (1986), and has persisted in papers using more recent data (Li 2001, Hunt 2011, Fairlie and Lofstrom 2015, Kerr and Kerr 2016, Green et al. 2016). The academic debate surrounding the source of this immigrant-native gap in self-employment rates has primarily been framed in terms of the "push" and "pull" theories of self-employment[24] (Clark and Drinkwater 2000,Abada et al. 2014 Fisher and Lewin 2018). According to the push theory of self-employment, immigrants choose to become self-employed since they have poor opportunities in wage employment, either because they face discrimination or because they possess skills that are not highly valued by employers (Volery 2007, Abada et al. 2014, ,Fisher and Lewin 2018). However, proponents of the pull theory argue that immigrants tend to be more entrepreneurial, either due to their cultural backgrounds, or because they have self-selected into migrating to the U.S. in part on the basis of their entrepreneurial abilities (Yuengert 1995, Hammarstedt and Shukur 2009, Vandor and Franke 2016, Vinogradov and Jørgensen 2017). The empirical research that has attempted to separate between these two potential explanations has produced mixed results, with each of the aforementioned papers finding results supporting the theory in question.

It has also been postulated that immigrants may be more likely to be self-employed than natives as a result of stronger preferences for self-employment, either due to strong self-employment traditions in their country of origin (Yuengert 1995), or because immigrants have a higher tolerance for risk (Andersson and Hammarstedt 2010, Batista and Umblijs 2014). However, the evidence for these theories has also been mixed.Fairlie and Meyer 1996 find that immigrants who come from countries with higher self-employment rates are no more likely to be self-employed after migrating to the U.S. than immigrants from countries with lower self-employment rates. However, Hammarstedt and Shukur (2009) find the opposite result among immigrants to Sweden[25]. Although there have been no empirical studies that attempt to ex-

[24] The terms "necessity" and "opportunity" are often used in place of push and pull.

[25] It should also be noted that these two papers use different methodologies. The difference in their results is

plain this gap in self-employment rates using risk preferences, the literature comparing the risk tolerances of immigrants and natives has itself produced mixed results. Consistent with expectations, Jaeger et al. (2010) and Bauernschuster et al. (2014) find that individuals who migrate across countries or across cultural barriers have higher tolerances for risk. However, in contrast, Bonin et al. (2009) find that immigrants to Germany are *more* risk averse than native born Germans.

I contribute to this literature on immigrant self-employment by examining immigrant self-employment patterns using both occupation and skill based perspectives. I find that the higher self-employment rate among immigrants is not entirely the result of immigrants being more likely to enter occupations that are traditionally associated with self-employment. Rather, immigrants are on average more likely to be self-employed than natives who work in the same occupation. This evidence suggests that the immigrant-native gap in self-employment rates is not being driven by differences in occupational choice patterns between immigrants and natives, or by immigrants facing higher barriers to entry into select occupations.

Taking inspiration from the burgeoning literature on occupational skills (Autor et al. 2003, Acemoglu and Autor 2011), I demonstrate that while the aforementioned within occupation gaps in self-employment rates occur across the entirety of the skill distribution, they are most pronounced in higher skill occupations. However, it is in these high skill occupations where an immigrant's earnings advantage from choosing self-employment is the smallest, relative to a native's earnings advantage from choosing self-employment. If the choice to be self-employed or wage employed is based primarily on which sector is expected to result in higher earnings, then we should expect to observe a pattern of self-employment across the skill distribution that is the opposite of what occurs in the data. Furthermore, I fail to find evidence that immigrants are more likely to be self-employed in occupations that typically require higher levels of preparation or training. This is inconsistent with another hypothesis surrounding immigrant self-employment, which is that it may be a response to potential employers discounting the training and experience of immigrants.

In order to provide clarity on these findings, I develop a theoretical model that allows an individual to choose either the self-employment sector or the wage employment sector in any give occupation, and in any given region. Their choice of an occupation, sector, and region is

not necessarily due to differences between immigrants in the U.S. and Sweden.

made based on a combination of earnings potential and non-pecuniary preferences. This model intuitively predicts that individuals are more likely to enter a given sector when their earnings potential is higher in that sector. However, it also predicts that in the presence of a monetary migration cost, earnings potential in self-employment has a stronger effect on the decision of whether to become self-employed than does wage employment earnings potential. Migration costs act a deterrent to migrating, but so long as self-employment earnings are more variable than wage employment earnings, they will act as a weaker deterrent to migrating in order to choose self-employment in high skill/high earnings occupations. This generates the prediction that migrants, who by definition face such a migration cost, should be relatively more likely to be self-employed in these high skill/high earnings occupations. This pattern of self-employment is consistent with the idea that the higher self-employment rate among immigrants results from non-random selection into migrating, rather than any post-arrival response to labour market conditions.

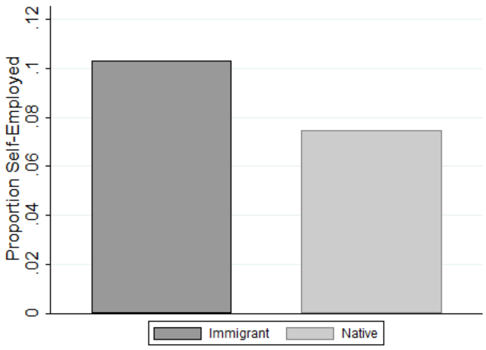
As a robustness check, I estimate the empirical results while substituting immigrants with natives who have moved from one state to another, and I find remarkably similar results. Since these natives have demonstrated a willingness to migrate, but otherwise are expected to possess characteristics that are similar to those of other natives who have not migrated, this supports the hypothesis that the higher self-employment rate among immigrants is a general characteristic of individuals who migrate, rather than any other characteristic of immigrants.

The remainder of this paper is divided as follows; Section 2.2 describes the data and presents some basic descriptive statistics, Section 2.3 introduces the theoretical model of self-employment entry, Section 2.4 describes the empirical framework, Section 2.5 presents the main results, Section 2.6 presents results that compare native born cross-state migrants to natives who remain in their state birth, Section 2.7 consists of a series of robustness checks, and Section 2.8 concludes.

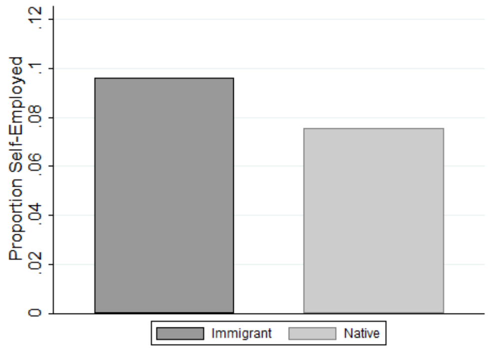
2.2 Data and Descriptive Statistics

The data consists of the American Community Survey 2016 1-year Person Files. My primary population of interest consists of individuals who earn a living working full-time in either self-employment or wage employment. Therefore, I restrict the sample to eliminate individuals who report working fewer than 20 hours per week on average. I also exclude individuals with an average hourly income of less than \$7.25[26] or more than \$100[27].Furthermore, I exclude individuals without a recorded occupation, as well as immigrants who arrived in the United States prior to the age of 18, since they are expected to possess characteristics that are similar to those of natives, relative to the characteristics of other immigrants. After imposing these restrictions, I find a substantial gap in self-employment rates between immigrants and natives that is consistent with earlier work in the literature, with 7.5% of natives being self-employed as opposed to 10.3% of immigrants (see panel a) of Figure 2.1).

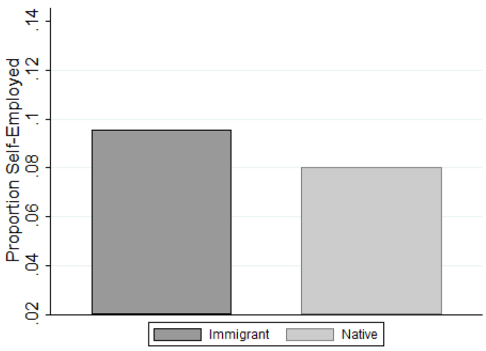
Fig. 2.1. Estimated Margins of Self-Employment Rates for Immigrants and Natives



a) Raw Self-Employment Rates



b) Controlling for Demographic Characteristics



c) Controlling for Demographics, Industry, and Occupation

Panel a) contains the estimated marginal effects for a simple probit in which self-employment status is the dependent variable, and immigrant status is the only independent variable. Panel b) includes as independent variables the demographic characteristics that are listed in 2.1. Panel c) includes further controls for industry , and detailed occupations.

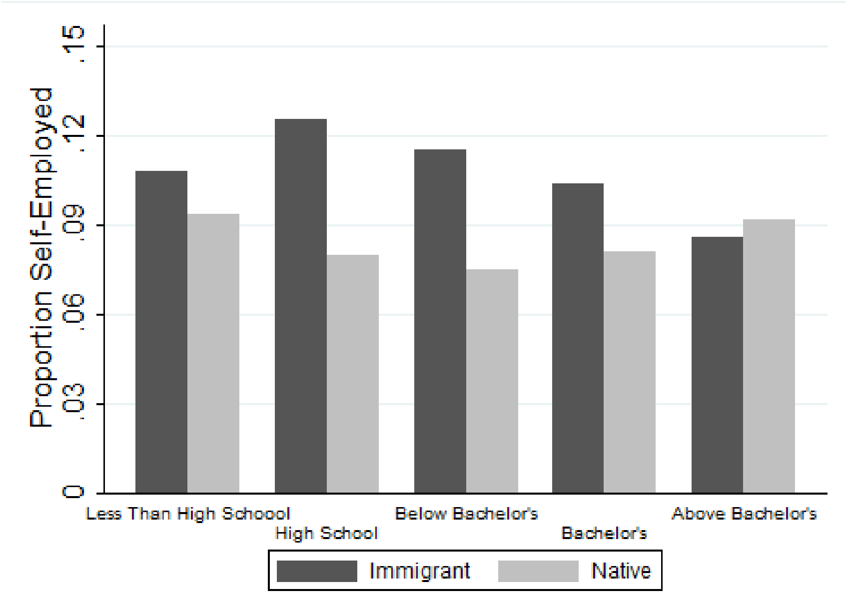
While there exists a large literature examining self-employment rates across demographic groups such as age (Blanchflower 2000, Zissimopoulos and Karoly 2007), education (Blanchflower 2000, Parker 2018), and in particular ethnicity (Fairlie and Meyer 1996,Fairlie and Meyer 2000, Bogan and Darity Jr 2008), it is worthwhile noting that the immigrant-native differences in self-employment rates are largely separate from differences in these characteristics. While

[26] The U.S. federal minimum wage

[27] I exclude very low and very high earners in order to ensure that the results are not driven by outliers.

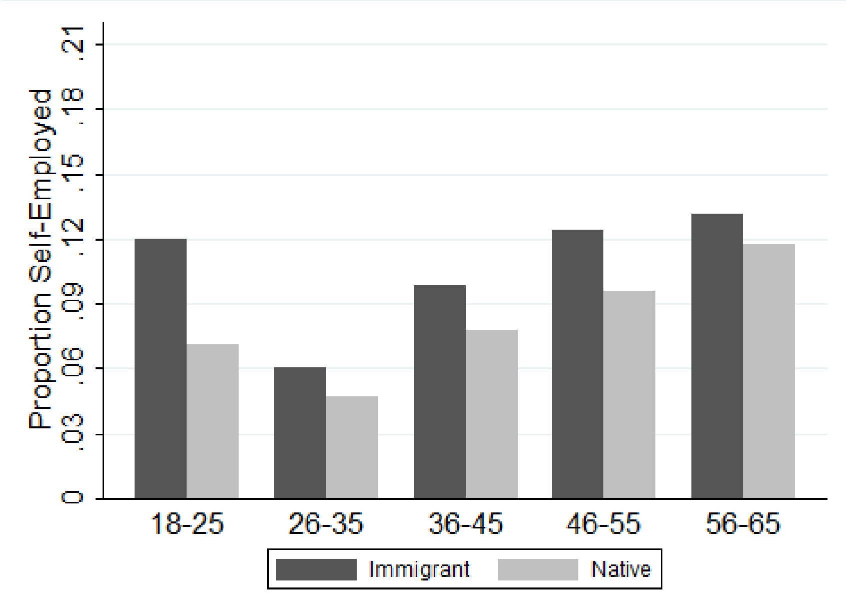
I observe striking differences in self-employment rates across demographic characteristics (as shown in Figures 2.2, 2.3, 2.4 & 2.5), it is equally striking that immigrants are more likely to be self-employed than natives *within* each ethnicity, age category, gender, and level of education (with the exception of individuals with graduate or professional degrees).

Fig. 2.2. Self-Employment Rates by Education and Immigration Status



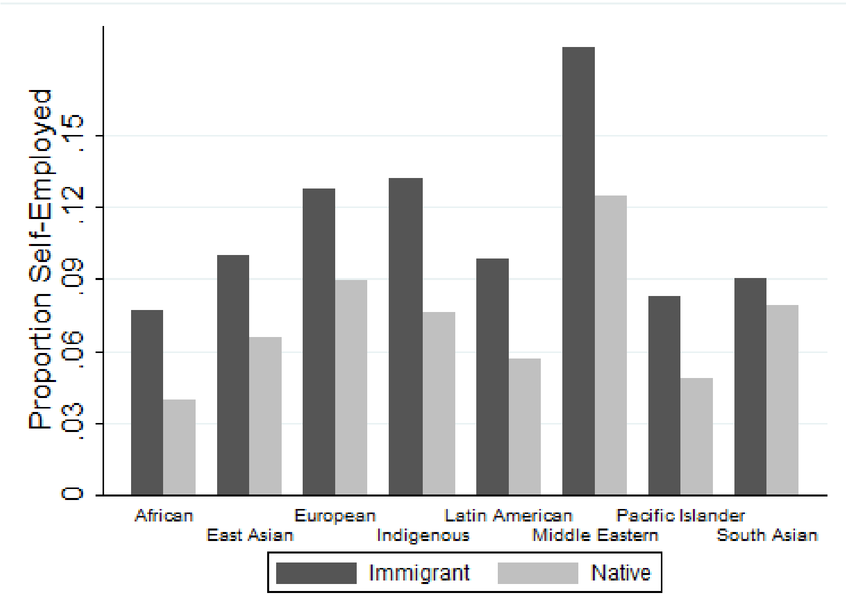
# of Natives with Less Than High School=120,135, # of Immigrants with Less Than High School=42,245, # of Natives with High School=267,690, # of Immigrants with High School=28.775, # of Natives with Below Bachelor's=444,375, # of Immigrants with Below Bachelor's=31,845, # of Natives with Bachelor's=287,440, # of Immigrants with Bachelor's=34,115, # of Natives with Above Bachelor's=171,400, # of Immigrants with Above Bachelor's=31,425

Fig. 2.3. Self-Employment Rates by Age and Immigration Status



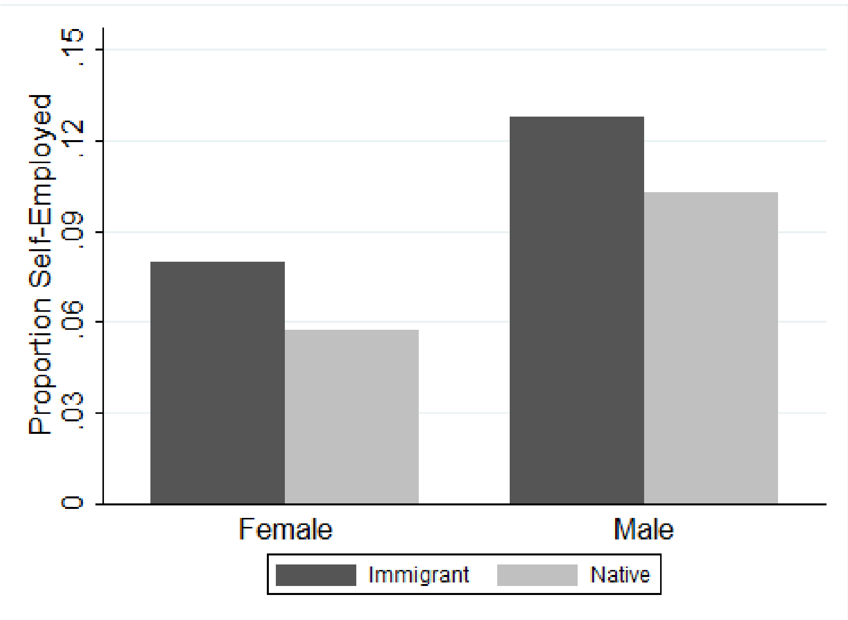
# of Natives 18-25=199,670, # of Immigrants 18-25=6075, # of Natives 26-35=248,185, # of Immigrants 26-35=28,045, # of Natives 36-45=220,940, # of Immigrants 36-45=42,870, # of Natives 46-55=259,485, # of Immigrants 46-55=44,910, # of Natives 56-65=256,170, # of Immigrants 56-65=34,020

Fig. 2.4. Self-Employment Rates by Ethnicity and Immigration Status



# of Natives with African Ethnicity=112,760, # of Immigrants with African Ethnicity=7340, # of Natives with East Asian Ethnicity=23,365, # of Immigrants with East Asian Ethnicity=34,425, # of Natives with European Ethnicity=1,002,195, # of Immigrants with European Ethnicity=41,330, # of Natives with Indigenous Ethnicity=47,365, # of Immigrants with Indigenous Ethnicity=390, # of Natives with Latin American Ethnicity=102,435, # of Immigrants with Latin American Ethnicity=63,625, # of Natives with Middle Eastern Ethnicity=6955, # of Immigrants with Middle Eastern Ethnicity=6500, # of Natives with Pacific Islander Ethnicity=3320, # of Immigrants with Pacific Islander Ethnicity=905, # of Natives with South Asian Ethnicity=2850, # of Immigrants with South Asian Ethnicity=15,090

Fig. 2.5. Self-Employment Rates by Gender and Immigration Status



# of Male Natives=658,865, # of Male Immigrants=91,425, # of Female Natives=632,170, # of Female Immigrants=76,980

I control for a rich array of demographic characteristics[28] in a probit model of the form:

$$\Phi \quad = \alpha_1 + \beta_x \quad X_i + \beta_{imm} \quad Imm_i + \quad_i \tag{2.1}$$

Where the dependent variable is an indicator that takes on a value of 1 if the individual is self-employed, and 0 if they are wage employed. Individuals choose self-employment if $\Phi \quad > 0$, or wage employment if $\Phi \quad \leq 0$. X_i consists of a rich set of demographic controls from Table 2.1. "Female" is an indicator that takes on the value of 1 if the individual is female, or 0 otherwise. "Married" takes on a value of 1 if the individual is married or in a common law relationship, or 0 otherwise. Education levels "Less Than High School", "High School", "Below Bachelors", "Bachelors", and "Above Bachelors" indicate the highest level of education that the individual has completed[29]. Ethnicities of "Indigenous", "Latin American", "Middle Eastern", "European", "African", "East Asian", "South Asian" and "Pacific Islander" are indicators that take on a value of 1 if the individual reports an ancestry that belongs in that broad ethnic group[30]. Including these demographic controls in the probit model fails to account for a large portion of the raw gap in self-employment rates between immigrants and natives. The estimated margins from this model, shown in panel b) of Figure 2.1, yield predicted self-employment rates of 7.5% and 9.6% for natives and immigrants respectively.

What is perhaps even more striking is that even though there have been well documented differences in the occupational distribution of immigrants relative to natives, differences in industry and occupation are only able to explain approximately half of the remaining gap in self-employment rates between immigrants and natives[31]. As shown in panel c) of Figure 2.1, adding controls for two digit NAICS industry codes and detailed Standard Occupational Classification (SOC) codes[32] to Equation 2.1 explains slightly more than half of the immigrant-

[28] The full set of characteristics, along with differences in these characteristics between immigrants and natives, can be found in Table 2.1.

[29] Note that Below Bachelors indicates the individual has obtained some post-secondary education, but less than that which is needed to complete a bachelor's degree. Above Bachelors indicates that the individual has completed more post-secondary education than what is needed in order to complete a bachelor's degree. Examples include professional and graduate degrees. It is worthwhile noting that immigrants are over-represented at both extremes of the education distribution, as shown in Table 2.1.

[30] Individuals may report up to 2 different ancestries.

[31] Contrary to what may be expected, narrow occupational groups are usually not strongly linked to either the self-employment or wage employment sector. Of the 455 occupations in the data, 133 of them have self-employment rates of at least 10%. Of these, only one has a self-employment rate above 80%, "Farmers, Ranchers, and Other Agricultural Managers". The full list of occupations in the data, along with their within occupation self-employment rates, can be found in Table A2.1.

[32] In addition to controlling for SOC codes, I also estimate a model that controls for the OCC codes that are

Table. 2.1. Demographic Characteristics of Natives and Immigrants

Characteristic	% of Natives	% of Immigrants
Female	49%	45.7%
Married	53.5%	70.1%
Less Than High School	9.3%	25.1%
High School	20.7%	17.1%
Below Bachelors	34.4%	18.9%
Bachelors	22.3%	20.3%
Above Bachelors	13.3%	18.7%
Indigenous	3.7%	0.2%
Latin American	27%	48.5%
Middle Eastern	0.5%	3.9%
European	52.9%	6.4%
African	8.7%	4.4%
East Asian	1.8%	20.4%
South Asian	0.2%	9%
Pacific Islander	5.1%	7.3%
Homeowner	62.7%	50.9%
Characteristic	Native Mean	Immigrant Mean
Age	43.974	47.08

# of Natives=1,291,040, # of Immigrants=168,405.Each value indicates the percentage of individuals in the subsample (Natives or Immigrants) for whom the covariate is equal to 1. The exception to this is age, with the values consisting of the average age for natives and immigrants.

native gap in self-employment rates that remains after accounting for demographics. Even within the same industry and occupation, my results show that approximately 7.5% of natives are expected to be self-employed, in comparison to 8.4% of immigrants.

These results suggest that the common idea that immigrants to OECD countries are more likely to be self-employed because they disproportionately work in occupations that are related to businesses such as corner stores, restaurants, nail salons, or motels (Patel and Vella 2013, Kerr and Mandorff 2015) provides only a partial picture of immigrant self-employment in the United States. While immigrants are relatively more likely to be working in occupations that are associated with high self-employment rates [33(34)], this does not explain the entire gap in self-employment rates between immigrants and natives. Rather, immigrants are on average more likely to be self-employed than natives who are working in the same occupation.

included in the ACS data. These two models generate similar results.

[33]See Figure A2.1, which is constructed by first estimating the self-employment rate for each occupation in the sample. This is followed by plotting the estimated density of both immigrants and natives by occupational self-employment rate.

[34]Notable examples include Miscellaneous Personal Appearance Workers (57.4% immigrants), Tailors, Dressmakers and Sewers (44.6% immigrants), Taxi Drivers and Chauffeurs (36.5% immigrants), Laundry and Dry Cleaning Workers (30.3% immigrants) and Chefs and Head Cooks (28.2% immigrants).

2.2.1 Skills

The U.S. Bureau of Labor Statistic's O*NET database contains information on the characteristics that are expected of workers for each of the 840 occupations classified by the Standard Occupational Classification (SOC). I utilize the increasingly common approach among researchers[35] of using the O*NET data in order to estimate an ***individual's*** skill levels, under the assumption that workers are on average sorted into occupations that are a good match for their skills. The skill measures are constructed using the ***Skills, Knowledge, and Abilities*** content data contained in the O*NET database (hereafter collectively referred to as skills for simplicity). Each of these dimensions contain both ***importance*** and ***level*** scores[36], which I then place on scales of 0-10. Since many of the occupations that are included in the ACS data are aggregations of 2 or more of the detailed SOC occupations, I re-classify these aggregated occupations as the detailed occupation which is most common or representative of the group[37]. In some cases the ACS data classifies an occupation as a miscellaneous member of a broader occupational category. For many of these cases, the aggregation was judged to be too broad to assign the skills of any one occupation, therefore these groups are omitted from the estimation sample[38]. A total of 455 distinct occupations remain.

The ONET database contains measures of 109 different types of skills. However, it is clearly impractical to use them all in the analysis. Therefore, I create a series of orthogonal aggregations of these skill measures (including both the importance and level scores) using Principal Component Analysis, and combine these principal components into a single measure by weighing each of the first 19 principal components by their respective eigenvalues[39]. The generated measure of skill captures the degree of complexity or specialization that is associated with a particular occupation. Occupations with high skill requirements can include either occupations that are high in cognitive skill requirements, or occupations that are high in manual

[35]Notable examples include Acemoglu and Autor (2011), Frey and Osborne (2017), and the pioneering work of Autor et al. (2003) using O*NET's predecessor, the Dictionary of Occupational Titles.

[36]"Importance" is a measure of the frequency with which that attribute is required on the job, while "Level" is a measure of the complexity of the attribute as it is required on the job.

[37]A complete list of these re-classifications is available upon request.

[38]An example of such a broad category is "Miscellaneous managers, including funeral service managers and postmasters and mail superintendents". With the exception of the aforementioned occupation, representing 3.23% of the total data, each of these occupations comprise less than 0.1% of the total observations. A list of these omitted occupational groups is available upon request.

[39]I retain only those principal components for which the corresponding eigenvalue is greater than 1, which consists of the top 19 principal components.

skill requirements (or both). Skill in this context should not be confused with educational requirements, since this captures only one aspect of the skill measure. The top 10 and bottom 10 occupations in terms of their skill requirements are presented in Table 2.2.[40].

Table. 2.2. Top 10 and Bottom 10 Occupations in Terms of Their Skill Requirements

Bottom 10	
Occupation	Skill
Textile, Garment, And Related Materials Pressers	0
Graders And Sorters, Agricultural Products	0.0924891
Cleaners Of Vehicles And Equipment	0.4969215
Dishwashers	0.8813386
Sewing Machine Operators	0.9439152
Janitors And Building Cleaners	0.9683249
Hand Packers And Packagers	1.060512
Miscellaneous Food Preparation And Serving Related	1.070117
Crossing Guards	1.102751
Telemarketers	1.183745
Top 10	
Occupation	Skill
First-Line Supervisors Of Fire Fighting And Prevention	9.416777
Natural Sciences Managers	9.454381
Producers And Directors	9.471089
Registered Nurses	9.492625
Environmental Engineers	9.546906
Industrial Production Managers	9.569338
Compliance Officers	9.620062
All Other Computer Occupations	9.688899
Miscellaneous Engineers, Including Nuclear Engineers	9.957451
Chief Executives And Legislators	10

Skills are estimated using Principal Component Analysis, as described in Section 2.2.1. This table presents the 10 occupations with the lowest estimated skill levels, as well as the 10 occupations with the highest estimated skill levels.

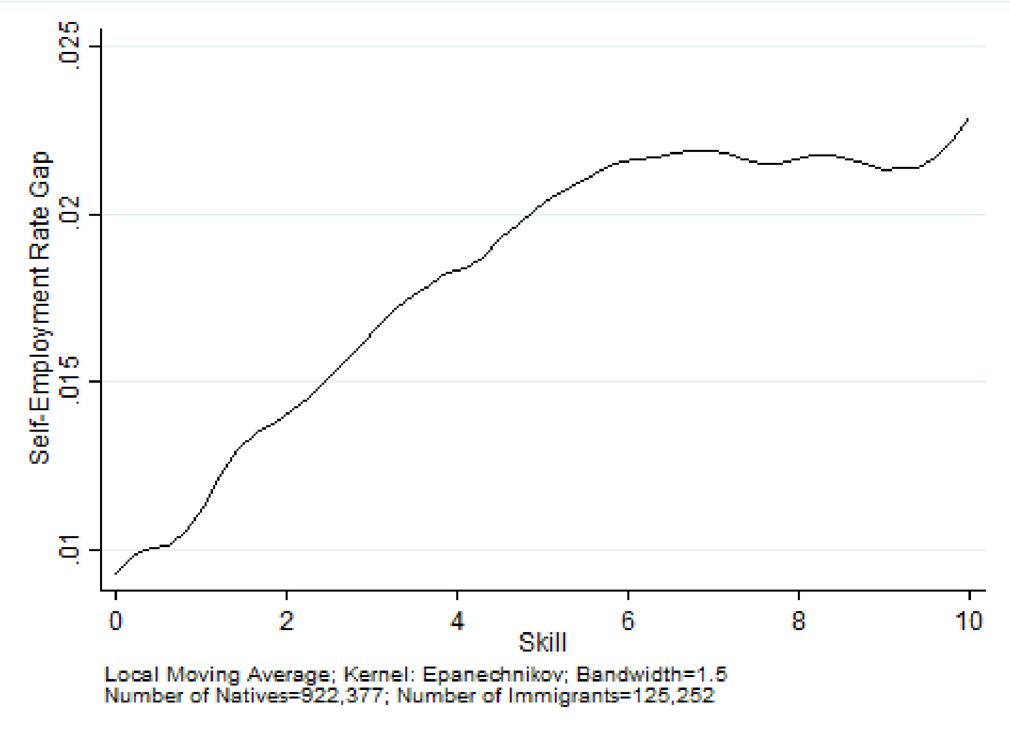
Figure 2.6[41] demonstrates that the gap in self-employment rates between immigrants and natives is primarily a high skill occupation phenomenon. Although immigrants are more likely to be self-employed than natives in similarly skilled occupations across the entirety of the skill distribution, the average gap increases from approximately 1 percentage point in the lowest skill occupations, to more than 2 percentage in high skill occupations. The purpose of the

[40] The complete set of skill levels for different occupations can be found at this link

[41] Note that a handful of occupations have been eliminated in creating this graph. These are occupations for which the gap in self-employment rates between immigrants and natives exceeds 20 percentage points in absolute value. Only 15 out of 455 occupations are eliminated, which represents approximately 3.5% of the sample. When no occupations are eliminated, the graph appears as in Figure A2.2. The decline in the self-employment rate gap that is observed above a skill level of 6 is driven by only 3 different high skilled occupations in which natives are more likely to be self-employed. These occupations are: Farmers, Ranchers, and Other Agricultural Managers; Food Service Managers; Morticians, Undertakers, and Funeral Directors.

remainder of this paper is to provide an explanation for why this is the case.

Fig. 2.6. Gap in Self-Employment Rates Between Immigrants and Natives Conditional On Occupational Skill Level



This figure is produced by estimating, for each occupation, the difference between the self-employment rate for immigrants in that occupation and the self-employment rate of natives in that occupation. A Local Moving Average with a bandwidth of 1.5 is used in order to generate the smooth trend. All occupations for which the gap in self-employment rates between immigrants and natives exceeds 20% are eliminated from this graph. Only 15 out of 455 occupation are eliminated. These occupations are: "Farmers, Ranchers, and Other Agricultural Managers", "Food Service Managers", "Announcers","Motion Picture Projectionists","Morticians, Undertakers, and Funeral Directors","First-Line Supervisors of Retail Sales Workers","Proofreaders and Copy Makers","Fishing and Hunting Workers", "Logging Workers","Locksmiths and Safe Repairers","Miscellaneous Woodworkers, Including Model Makers and Patternmakers","Molders, Shapers and Casters, Except Metal and Plastic","Taxi Drivers and Chauffeurs","Dredge, Excavating and Loading Machine Operators","Refuse and Recyclable Material Collectors"

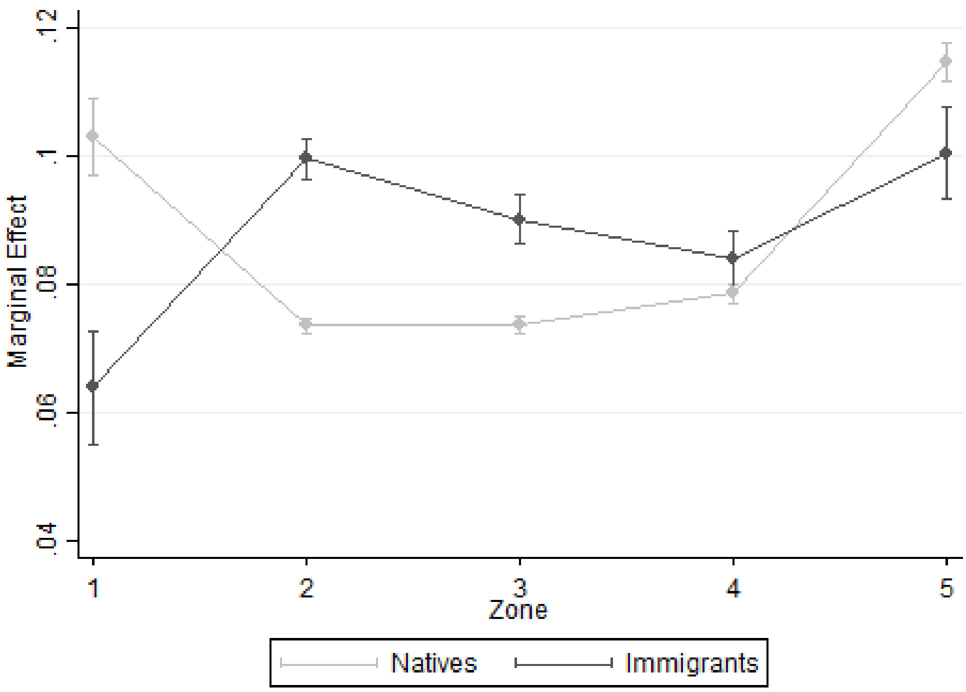
2.2.2 Job Zones and Mean Log Wage Gap

Curiously, despite being more likely to be self-employed in high skilled occupations, there is no evidence that immigrants are more likely to be self-employed in occupations that require high levels of preparation in the form of education and/or experience. Figure 2.7 shows the estimated marginal effect of each O*NET Job Zone[42] on the self-employment rate, when including these zones in a probit model that also includes as independent variables the demographic characteristics in Table 2.1, as well as controls two digit NAICS industries. The 5 Job Zones that are outlined by O*NET indicate the level of preparation that is expected of individu-

[42] Detailed information on Job Zones can be found at https://www.onetonline.org/help/online/zoneszone1

als in the occupation, Zone 1 indicates “Little or No Preparation Needed”, Zone 2 indicates “Some Preparation Needed”, Zone 3 indicates “Medium Preparation Needed”, Zone 4 indicates “Considerable Preparation Needed”, Zone 5 indicates “Extensive Preparation Needed”.

Fig. 2.7. Marginal Effects of Job Zone on Self-Employment Rates



Marginal effects estimated from probit in which self-employment status is the dependent variable, demographic characteristics in Table 2.1, controls for two digit NAICS industries, and Job Zones as described in Section 2.2.2 are the independent variables.
# of Natives in Zone 1=26,615, # of Immigrants in Zone 1=5640, # of Natives in Zone 2=518,810, # of Immigrants in Zone 2=79,555, # of Natives in Zone 3=330,720, # of Immigrants in Zone 3=34,025, # of Natives in Zone 4=288,915, # of Immigrants in Zone 4=32,510, # of Natives in Zone 5=11,810, # of Immigrants in Zone 5=1355

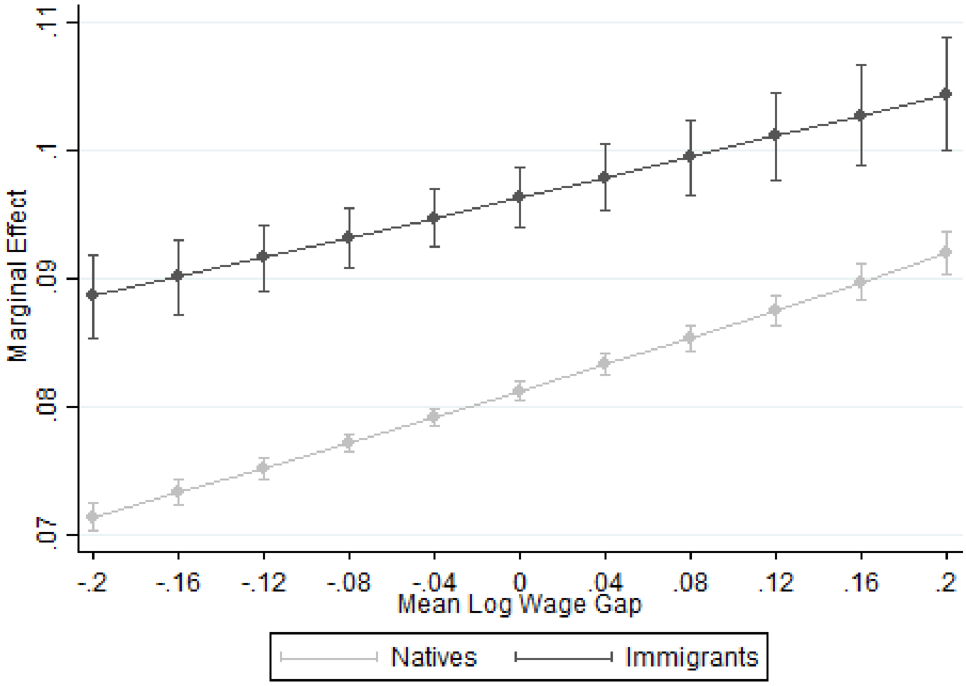
Furthermore, immigrants do not appear to be more likely to be self-employed in those occupations in which immigrants tend to receive lower wage sector earnings than their native born counterparts. The difference in the mean log wage between natives and immigrants is calculated for each occupation, as in Equation 2.2:

$$Gap_d = \hat{Y}^{Nat}_{d,we} - \hat{Y}^{Imm}_{d,we} \tag{2.2}$$

Where

$$\hat{Y}^{Imm}_{d,we} = \sum_i \frac{Y^i_{d,we}}{N_{Imm,d,we}} \; if \; Imm = 1 \qquad \hat{Y}^{Nat}_{d,we} = \sum_i \frac{Y^i_{d,we}}{N_{Nat,d,we}} \; if \; Imm = 0 \tag{2.3}$$

Fig. 2.8. Marginal Effects of Mean Log Wage Gap Between Natives and Immigrants on Self-Employment Rates



Marginal effects estimated from probit in which self-employment status is the dependent variable, demographic characteristics in Table 2.1, two digit NAICS industry codes, and the mean log wage gap between natives and immigrants (as described in Section 2.2.2) are the independent variables.

With d indicating the occupation, we indicating the wage employment sector, Imm indicating that an individual is an immigrant, and Nat indicating that they are a native. Higher values of the gap in Equation 2.2 indicate that natives on average earn more than immigrants in occupation d[43]. Figure 2.8 shows the estimated marginal effects of this earnings difference on self-employment, with it being clear that immigrants tend not to be self-employed in those occupations in which they are expected to earn less than natives.

2.3 Theoretical Model of Self-Employment vs. Wage Employment Choice

Lucas Jr (1978) developed an early model of self-employment entry, by allowing self-employment earnings to vary based on an individual's managerial ability. This model predicts that individuals with higher levels of managerial ability are more likely to be self-employed and to manage larger firms conditional on being self-employed. Evans and Jovanovic (1989) expand upon the

[43] It is interesting to note that the mean of this value is -.048 , which indicates that on average, immigrants tend to earn nearly 5% more than natives who work in the same occupation.

work of Lucas by introducing credit constraints, and by allowing wage sector earnings to vary based on an individual's characteristics. Their model predicts, similar to Lucas Jr (1978), that on average individuals with higher levels of entrepreneurial ability will be more likely to be self-employed. However, rather than a strict cutoff level of managerial ability at which individuals choose to become self-employed, variations in access to credit and potential wage sector earnings allow for differences in the likelihood of self-employment entry conditional on a given level of entrepreneurial ability. More recent papers by Hamilton et al. (2019) and Humphries (2016) introduce non-pecuniary benefits into the decision of whether to become self-employed. The latter work by Humphries (2016) consists of a dynamic model that allows for differences in self-employment rates across the lifecycle.

In this paper, I develop a model that is based upon the work of the above authors, with the introduction of a pair of substantial modifications. The first modification is to assign a skill level to each occupation, where the skill's value in each sector affects the probability that an individual chooses self-employment. Second, since I am interested in gaps in self-employment rates within occupations, I explicitly allow for individuals to enter any given occupation in either the self-employment sector or the wage employment sector.

Perhaps in contrast to expectations, the vast majority of occupations in the data are not exclusive to either the self-employment or wage employment sector. Table A2.1 shows that 444 of the 455 occupations in the ACS contain at least one individual who is self-employed and at least one individual who is wage employed. Furthermore, 340 of them (which represents nearly 75% of occupations) have at least two percent of their workers in self-employment and at least two percent in wage employment. This indicates that the assumption that individuals may pursue occupations in either the self-employment or wage employment sector is broadly reflective of the reality that is observed in the data.

The model is intended to demonstrate why immigrants are more likely to be self-employed than native born Americans at the upper end of the skill distribution. It is able to generate this prediction, even in the simple case in which differences in earnings capability, non-pecuniary preferences, and access to specific occupations do not vary between immigrants and natives. While differences in such characteristics between immigrants and natives do certainly exist, they are not necessary in order to generate the pattern of immigrants being more likely to be self-employed in high skilled occupations. Furthermore, the idea that these differences are driving this pattern of self-employment appears to be inconsistent with the data. As described in the previous section, immigrants are not more likely to be self-employed in occupations in which they have a relative earnings advantage (as in Figure 2.7), nor are they more likely to be self-employed in occupations that require higher levels of preparation (as in Figure 2.8). Furthermore, it is doubtful whether differences in preferences for self-employment between immigrants and natives in and of itself is driving the fact that self-employment among immigrants is concentrated in high skilled occupations. Such preference differences would be insufficient to explain why there are differences in the gap in self-employment rates across the skill distribution.

2.3.1 Model Outline

Individual i chooses from a set of occupations $d = (1, ..., D)$, from sectors $j = (se, we)$, and from regions $r = (1, 2)$. se denotes the self-employment sector, while we denotes the wage employment sector. D indicates the total number of possible occupations. Individual i's region of birth is denoted by $br = (1, 2)$. Each occupation may be pursued in either self-employment or wage employment, and may be pursued in either Region 1 or Region 2. If the individual chooses occupation d and chooses to be wage employed, their income will be:

$$Y^{i}_{d,we,r} = W_r(X_i, \eta_d) - c_{i,r} + \tag{2.4}$$

Where X_i denotes a set of individual characteristics, η_d is the skill level that is associated with occupation d, and W_r is the region specific earnings function for wage employment. represents the error term, distributed according to $F^{we,r}(\ )$. $c_{i,r}$ is a monetary migration cost incurred by individual i, which is given by:

$$c_{i,r} = \begin{cases} c_r & \text{if } br_i = r \\ 0 & \text{otherwise} \end{cases} \tag{2.5}$$

Where $c_r > 0$. This indicates that the individual has to pay a migration cost only if they choose to emigrate from their region of birth. If the individual is self-employed, their earnings are given by:

$$Y^i_{d,se,r} = \Theta_r(X_i, \eta_d) - c_{i,r} + \quad (2.6)$$

Where   represents the error term, distributed according to $F^{se,r}(\ )$. θ_r is the region specific self-employment earnings function. Components of earnings other than the migration cost $c_{i,r}$ are independent of the individual's region of birth. When employed in occupation d, sector j, and region r, individual i receives a utility of:

$$U^i_{d,j,r} = u(Y^i_{d,j,r}) + v^i_{d,j,r} \quad (2.7)$$

Where $Y^i_{d,j,r}$ is their earnings from choosing d, j, r, and $v^i_{d,j,r}$ denotes their non-pecuniary benefits from this choice. $u(Y^i_{d,j,r})$ is the portion of utility that is derived from the earnings of individual i. It is assumed that $u\ \ > 0$ and $u\ \ < 0$, that is agents are risk averse, and have utility that is increasing in income. Non-pecuniary benefits are distributed with unconditional density $f_{d,j,r}(v_{d,j,r})$ and joint density $f_{d,j,r;p,q,z}(v_{d,j,r}, v_{p,q,z})\quad \{p, q, z\} = \{d, j, r\}$. It is assumed that the limits of the unconditional and joint density functions as $v_{d,j,r} \to \{-\infty, \infty\}$ and $v_{p,q,z} \to \{-\infty, \infty\}$ are equal to zero.

The individual chooses the occupation/sector/region combination that maximizes their expected utility according to:

$$\underset{d,j,r}{\text{argmax}}\ E[U^i_{d,j,r}] \quad (2.8)$$

Where

$$E[U^i_{d,j,r}] = E[u(Y^i_{d,j,r}) + v^i_{d,j,r}] = E[u(Y^i_{d,j,r})] + E[v^i_{d,j,r}] = E[u(Y^i_{d,j,r})] + v^i_{d,j,r} \quad (2.9)$$

and

$$E[u(Y^i_{d,se,r})] = \int_a^b u(\theta_r(X_i, \eta_d) - c_{i,r} + \ )f^{d,se,r}(\ )d \quad (2.10)$$

$$E[u(Y^i_{d,we,r})] = \int_a^b u(W_r(X_i, \eta_d) - c_{i,r} + \ )f^{d,we,r}(\ )d \quad (2.11)$$

Where a, b are the lower and upper limits of the error distribution, respectively[44]. I restrict the distribution of the error term to be such that $\int_{-\infty}^{\infty} f(\ )d\ = 0$ and $f(0+\Pi) = f(0-\Pi) \quad \Pi$. In other words the density of the error term is symmetric and has a mean of 0. From the point of view of the econometrician, the probability that an individual with characteristics X_i chooses combination d, j, r is given by:

$$Prob(E[u(Y^i_{d,j,r})] + v^i_{d,j,r} > E[u(Y^i_{p,q,z})] + v^i_{p,q,z} \quad \{p, q, z\} = \{d, j, r\}) \tag{2.12}$$

That is; the probability that d, j, r is chosen is equal to the probability that d, j, r yields a higher utility than all other possible occupation/sector pairs. This probability can be written as:

$$\begin{aligned} Pr(\{d, j, r\} = \underset{p,q,z}{\operatorname{argmax}} \ E[U^i_{p,q,z}]) = & \int_{-\infty}^{\infty} Prob(v^i_{d,j,r} = X - E[u(Y^i_{d,j,r})]) \\ & \prod_{p,q,z=\{d,j,r\}} Prob(v^i_{p,q,z} < X - E[u(Y^i_{p,q,z})] \mid v^i_{d,j,r} = X - E[u(Y^i_{d,j,r})])dX \end{aligned} \tag{2.13}$$

Where p, q, z represents an arbitrary occupation/sector/region triplet. Equation 2.13 can also be re-written as:

$$\begin{aligned} & \int_{-\infty}^{\infty} f_{d,j,r}(X - E[u(Y^i_{d,j,r} - c_{i,r} + \ )]) \\ & \prod_{p,q,z=\{d,j,r\}} \frac{\frac{\partial F_{d,j,r;p,q,z}(X - E[u(Y^i_{d,j,r} - c_{i,r} + \ ), X - E[u(Y^i_{p,q,z} - c_{i,z} + \ )]}{\partial(X - E[u(Y^i_{d,j,r} - c_{i,r} + \ )]}}{f_{d,j,r}(X - E[u(Y^i_{d,j,r} - c_{i,r} + \ )]} dX \end{aligned} \tag{2.14}$$

Where $\frac{\frac{\partial F_{d,j,r;p,q,z}(X - E[u(Y^i_{d,j,r} - c_{i,r} + \ ), X - E[u(Y^i_{p,q,z} - c_{i,z} + \ )]}{\partial(X - E[u(Y^i_{d,j,r} - c_{i,r} + \ )]})}{f_{d,j,r}(X - E[u(Y^i_{d,j,r} - c_{i,r} + \ )])}$ is the probability that combination p, q, z yields a utility smaller than X conditional on the utility of choice d, j, r being equal to X. For notational simplicity, $f_{d,j,r}(X - E[u(Y^i_{d,j,r} - c_{i,r} + \ )])$, which represents the probability

[44] Note the distinction in notation between $f^{d,j,r}(\ )$ and $f_{d,j,r}(v_{d,j,r})$. The former is the density function of the random component of earnings for choice d, j, r, while the latter is the density of non-pecuniary benefits from this choice.

that choice d, j, r yields a utility equal to X, shall be written as $f_{d,j,r}$ from this point forward. $F_{d,j,r;p,q,z}(X - E[u(Y^i_{d,j,r} - c_{i,r} + \;)], X - E[u(Y^i_{p,q,z} - c_{i,z} + \;)]$ shall be written as $F_{d,j,r;p,q,z}$. This represents the probability that choice d, j, r yields a utility that is less than or equal to X given that choice p, q, z yields a utility that is equal to X. Corresponding distribution and density functions shall be written similarly. The expected pecuniary utility from choice d, j, r, $E[u(Y^i_{d,j,r} - c_{i,r} + \;)]$ shall be written as $E^i_{d,j,r}$.

The probability that an individual chooses to become self-employed conditional on choosing to live in region 2 is given by:

$$\frac{Pr(j = se \,\&\, r = 2)}{Pr(r = 2)} \tag{2.15}$$

Where

$$Pr(j = se \,\&\, r = 2) = \sum_{d=1}^{D} \int_{-\infty}^{\infty} f_{d,se,2} \prod_{p,q,z=\{d,se,2\}} \frac{\frac{\partial F_{d,se,2;p,q,z}}{\partial (X - E^i_{d,se,2})}}{f_{d,se,2}} dX \tag{2.16}$$

and

$$Pr(r = 2) = \sum_{j=\{se,we\}} \sum_{d=1}^{D} \int_{-\infty}^{\infty} f_{d,j,2} \prod_{p,q,z=\{d,j,2\}} \frac{\frac{\partial F_{d,j,2;p,q,z}}{\partial (X - E^i_{d,j,2})}}{f_{d,j,2}} dX \tag{2.17}$$

This conditional probability is analogous to the probability that an individual is self-employed conditional on living in region 2. Among the subpopulation of individuals who are originally from region 1, this is the probability of being self-employed conditional on having migrated.

If we consider a given occupation d, and imagine that the skill level η_d of this occupation increases (but non-pecuniary benefits remain constant) this probability of self-employment conditional on residing in region 2 changes according to:

$$\frac{\frac{\partial Pr(j=se \,\&\, r=2)}{\partial \eta_{d'}} Pr(r = 2) - \frac{\partial Pr(r=2)}{\partial \eta_{d'}} Pr(j = se \,\&\, r = 2)}{Pr(r = 2)^2} \tag{2.18}$$

Equation 2.18 is ambiguous in sign[45]. Following an increase in the skill level of a given

[45] The full expression is available in Equation A2.5 of the appendix, with the necessary steps to derive this

occupation, the probability of being self-employed conditional on residing in region 2 may increase or decrease, depending upon how skills are rewarded in each sector/region pair, as well as how non-pecuniary preferences are distributed.

It is, however, clear that the when self-employment skill values in region 2 increase, the conditional probability of being self-employed in a high skilled occupation increases (as shown in Equation 2.19). Likewise, a similar effect occurs in response to wage employment skill values, as shown in Equation 2.20, with the conditional probability of self-employment being lower in high skilled occupations when wage employment skill values increase. These results are fairly intuitive, and indicate that individuals are relatively more likely to choose a given sector j in high skilled occupations when skills are more highly rewarded in that sector.

$$\frac{\partial}{(\frac{\partial \theta_2}{\partial \eta_d})}\frac{\partial}{\partial \eta_d}\frac{Pr(j=se\ \&\ r=2)}{Pr(r=2)} = \square_{-\infty}^{\infty}\square_{a}^{b}[u\,(\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ]\ \square_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}$$
$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^{i}_{d,se,2})}}{f_{d,se,2}}dX\ \ [\frac{Pr(r=2)-Pr(j=se\ \&\ r=2)}{Pr(r=2)^2}] \quad (2.19)$$

$$\frac{\partial}{(\frac{\partial W_2}{\partial \eta_d})}\frac{\partial}{\partial \eta_d}\frac{Pr(j=se\ \&\ r=2)}{Pr(r=2)} = \square_{-\infty}^{\infty}\square_{a}^{b}[u\,(W_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,we,2}(\ )d\ ]\ \square_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z}$$
$$\square_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial(X-E^{i}_{d,we,2})}}{f_{d,we,2}}dX\ \ [\frac{Pr(j=se\ \&\ r=2)}{Pr(r=2)^2}] \quad (2.20)$$

However, the magnitude of this effect is not the same for individuals who are originally from region 1 and individuals who are originally from region 2. Differentiating Equations 2.19 and 2.20 with respect to the migration cost (shown in Equations 2.21 and 2.22) indicates that this effect of occupational skill values on the probability of being self-employed in region 2 is stronger in the presence of a higher migration cost, so long as individuals are risk averse. Since individuals from region 1 are the only ones who face a migration cost for occupation/sector choices in region 2, we can also think of this as indicating that individuals who are originally from region 1 will be expected to have a higher probability of working in occupation/sector

expression being found in Equations A2.1-A2.4.

pairs in which they are expected to have higher earnings.

$$\frac{\partial}{\partial c_{i,2}}\frac{\partial}{\partial(\frac{\partial\theta_2}{\partial\eta_d})}\frac{\partial}{\partial\eta_d}\frac{Pr(j=se\ \&\ r=2)}{Pr(r=2)} = -\int_{-\infty}^{\infty}\int_a^b [u\ (\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ] \sum_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z} \sum_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}} dX \quad [\frac{Pr(r=2)-Pr(j=se\ \&\ r=2)}{Pr(r=2)^2}] \tag{2.21}$$

$$\frac{\partial}{\partial c_{i,2}}\frac{\partial}{\partial(\frac{\partial W_2}{\partial\eta_d})}\frac{\partial}{\partial\eta_d}\frac{Pr(j=se\ \&\ r=2)}{Pr(r=2)} = -\int_{-\infty}^{\infty}\int_a^b [u\ (W_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,we,2}(\ )d\ ] \sum_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z} \sum_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,we,2})}}{f_{d,we,2}} dX \quad [\frac{Pr(j=se\ \&\ r=2)}{Pr(r=2)^2}] \tag{2.22}$$

The intuition behind this result can be thought of as follows. Consider two individuals, one from region 1 and the other from region 2, both of whom have a strong preference for working as a wage employed investment banker in region 2, a choice that is associated with high expected earnings. As a result of these high earnings, a migration cost causes only a small decrease in utility for the individual from region 1 relative to the individual from region 2, since utility costs are lower at high income levels so long as individuals are risk averse. Therefore, the migration cost would only serve as a small deterrent to migrating to region 2 and becoming a wage employed investment banker. However, consider two individuals from region 1 and region 2 who have preferences for working as a wage employed short order cook in region 2. Since one would expect relatively low earnings from this choice, the migration cost imposes a larger decrease in utility for the prospective migrant from region 1, making them far less likely to choose this option than the individual from region 2 with similar preferences. This occurs because the migration cost has a stronger effect on utility at low income levels.

Therefore, we should expect to observe that, conditional on being in region 2, migrants who are originally from region 1 will be more likely to be employed in the sector in which they have higher earnings potential, whether this high earnings potential is due to the skill level of the occupation, as shown here, or their idiosyncratic characteristics X_i. Observed empirically, this would give the appearance that migrants place a greater value on earnings potential when choosing an occupation/sector pair than non-migrants as a result of different preferences. However, it is important to note that this does not need to be the case, and the results from Equations 2.21-2.22 do not require any differences in pecuniary or non-pecuniary preferences between individuals from either region. Rather, migration costs are deterring individuals from migrating and pursuing professions that would yield them low earnings in expectation.

This difference in the response to skill values between migrants and non-migrants also depends upon the degree of variability of the error term . The migration cost will result in a larger decrease in utility for the riskier self-employment sector. However, this will be stronger at the lower end of the expected earnings distribution if it is assumed that individuals possess a utility function that has the property of decreasing absolute risk aversion. When expected earnings are low, the migration cost will cause a larger increase in the probability of a "very low" earnings outcome when earnings are more variable. Therefore, the migration cost is expected to act as a more significant deterrent to migrating in order to pursue self-employment (vis a vis wage employment) in lower earning occupation/sector pairs than in higher earning occupation/sector pairs. Empirically, we should expect it to appear as though migrants have a stronger response to skill values in self-employment than in wage employment, even though this may not be the case.

The following section empirically tests the following predictions of the model:

1. Conditional on being in the United States, individuals are more likely to be employed in a sector j when they have higher estimated earnings potential in that sector.

2. The above effect is stronger for immigrants than it is for natives.

3. For immigrants, this effect is stronger for the self-employmet sector than the wage employment sector.

2.4 Empirical Framework

I use a Heckman Correction procedure in order to estimate the expected contribution of an occupation's skill level to an individual's earnings in both the self-employment and wage employment sectors. An individual's unobserved earnings potential in occupation d and sector we[46] are given by:

$$Log(Y^{i}_{d,we}) = \zeta_{we} + \pi_{we} X_i + g_{we}\eta_d + \rho_{we}\lambda_{we} + e^{i}_{we} \tag{2.23}$$

With their corresponding unobserved earnings potential in self-employment being:

$$Log(Y^{i}_{d,se}) = \zeta_{se} + \pi_{se} X_i + g_{se}\eta_d + \rho_{se}\lambda_{se} + e^{i}_{se} \tag{2.24}$$

The equation for selection into wage employment is given by:

$$\begin{aligned}\Phi_{we} = \alpha_{we} + \beta_{x,we} \quad X_i + \beta_{imm,we} \quad Imm_i + \kappa_{we}\eta_d \\ + \tau_{we} HousingResidual_i + \Pi_{we} \quad Homeowner_i \quad HousingResidual_i + \quad {}_{i,we}\end{aligned} \tag{2.25}$$

Similarly, the equation for selection into self-employment is given by:

$$\begin{aligned}\Phi_{se} = \alpha_{se} + \beta_{x,se} \quad X_i + \beta_{imm,se} \quad Imm_i + \kappa_{se}\eta_d \\ + \tau_{se} HousingResidual_i + \Pi_{se} \quad Homeowner_i \quad HousingResidual_i + \quad {}_{i,se}\end{aligned} \tag{2.26}$$

An individual's observed self-employment earnings are given by:

$$Log(Y^{i}_{d,se}) = \begin{cases} Log(Y^{i}_{d,se}) & if \quad \Phi_{se} > 0 \\ 0 \quad if \quad \Phi_{se} \leq 0 \end{cases} \tag{2.27}$$

While their observed wage employment earnings are given by:

[46] Note that region r is omitted from the notation in the empirical portion of this paper since all individuals reside in the same region, the United States.

$$Log(Y^i_{d,we}) = \begin{cases} Log(Y^i_{d,we}) & if \ \Phi_{we} > 0 \\ 0 \ if \ \Phi_{we} \leq 0 \end{cases} \tag{2.28}$$

g_{we} and g_{se} represent the skill prices that are associated with wage employment and self-employment respectively. λ_{we} and λ_{se} are the standard Heckman Correction lambdas. I account for selection into self or wage employment by including the local area's unanticipated housing price shock, as well as it's interaction with homeowner status, as exclusion restrictions in the selection equations. Equations 2.24 and 2.26 are jointly estimated using the Maximum Likelihood Heckman procedure, with Equations 2.23 and 2.25 being estimated using the same method.

I further estimate a similar set of equations to 2.23-2.26 with the addition of the term $Imm_i \ \ \eta_d$ to each equation in order to allow for a difference in skill prices between immigrants and natives.

2.4.1 Local Housing Price Shocks

Housing collateral has been found to be an important source of small business financing (Adelino et al. 2015, Schmalz et al. 2017), and increases in housing value are anticipated to result in a higher probability of being self-employed (Disney et al. 2010). However, there are also concerns that changes in housing prices are related to local economic conditions, and therefore may possibly also be related to earnings potential in self or wage employment for reasons other than shocks to collateral availability. Previous work in the literature has used the *residual* component of housing price changes as a source of exogenous variation in the self-employment decision, with these residuals being cleaned of the effects of local macroeconomic characteristics such as the unemployment rate (Disney et al. 2010). I use this approach in this paper.

I use county level data on housing prices,unemployment rates, and the size of the labour force[47]. Furthermore, I use GDP per capita data at the Metropolitan Area level for urban areas, and the state level for rural areas[48]. I convert these variables to the Public Use Microdata Area

[47] The source for housing price data is the Federal Housing Finance Agency, with this data being available thanks to the work of Bogin et al. (2019). The U.S. Bureau of Labor Statistics is the source for the unemployment and labour force data.

[48] Local GDP per capita data made available by the St. Louis Federal Reserve.

(PUMA) level[49]. PUMAs are geographic areas that are defined by the U.S. Census Bureau, with each PUMA containing a minimum population of 100,000 and being entirely contained within a single state[50]. Combining these variables results in a panel of local housing prices and macroeconomics characteristics from 2001 to 2016. The following fixed effects model is estimated in order to predict local housing price changes:

$$\begin{aligned}\%\Delta HousingPrice_{pu,t} = & \sum_{k=0}^{6} \beta_{k1}\ UnemploymentRate_{pu,t-k} + \sum_{k=0}^{6} \beta_{k2}\ UnemploymentRate^{2}_{pu,t-k} \\ & + \sum_{k=0}^{6} \beta_{k3} GDPPerCapita_{pu,t-k} + \sum_{k=0}^{6} \beta_{k4} GDPPerCapita^{2}_{pu,t-k} + \\ & \sum_{k=0}^{6} \beta_{k5} LabourForce_{pu,t-k} + \sum_{k=0}^{6} \beta_{k6} LabourForce^{2}_{pu,t-k} + \alpha_{pu} + \gamma_t + \epsilon_{jt} \end{aligned} \tag{2.29}$$

Where pu represents the PUMA and t represents the year. The dependent variable is the *percentage change* in local housing prices from the previous year, while the right hand side of the equation consists of lagged macroeconomic characteristics and their quadratic terms over the preceding 6 years[51].α_{pu} and γ_t capture PUMA and time specific fixed effects. ϵ_{jt} is the error term, with the corresponding residuals $\hat{\epsilon}_{jt}$ being the primary parameters of interest from this equation. The results from this model can be found in Table A2.2.

I include the residual component of the lagged annual change in housing prices over the course of the previous year in an individual's PUMA as a source of exogenous variation in the Heckman equation. Furthermore, I include the interaction of these residual housing price shocks with a dummy variable indicating whether the individual has owned their current place of residence for at least the previous 2 years[52]. If the residual portion of housing price shocks effects the probability that one is self-employed through the channel of increasing the value of

[49] Conversion is conducted using the "MABLE/Geocorr14: Geographic Correspondence Engine" from the Missouri Census Data Center

[50] For more details on PUMAs, see https://www.census.gov/programs-surveys/geography/guidance/geo-areas/pumas.html

[51] When using fewer than 6 lags, the residuals from this model are found to be positively related to wage sector earnings, indicating that these residuals do not represent housing price shocks that are independent of local macroeconomic characteristics. My choice of 6 lags is motivated by this finding. Furthermore, use of the Bayesian Information Criterion demonstrates that a quadratic model fits the data better than a linear model. A Hausman test shows a similar advantage for using fixed effects over random effects.

[52] I include only those individuals who have owned their residence for at least 2 years rather than using a contemporaneous homeownership variable only since the decision to purchase a home in an area in an area that has experienced a recent housing price shock is clearly not independent of one's earnings. Therefore I exclude recent purchases in order to avoid this source of endogeneity.

collateral, the results should show a positive effect of residual housing prices on self-employment probability for homeowners, and a smaller effect of ambiguous sign for non-homeowners[53].

2.5 Results

The results in Table 2.3 show the main results of interest from the Heckman procedure. The full results are available in Tables A2.3-A2.6 for the models that allow no difference in the return to skills between immigrants and natives, and Tables A2.7-A2.10 for the models that include the interaction between skill and immigrant status. The two exclusion restrictions containing residual housing price shocks show that for non-homeowners, residual housing price shocks contain no explanatory power in terms of predicting an individual's propensity towards self-employment. However, for homeowners of at least 2 years, an increase in residual housing prices is associated with a large and significant increase in self-employment probability[54]. The marginal effects of the housing price shock on the probability of being self-employed that correspond to column (1) of Table 2.3, are found in Figure 2.9. The results in Table 2.3 also show that the housing price shocks have no effect on the probability that an individual is wage employed. Also noteworthy is the fact that the negative coefficient on λ from the self-employment model in Table 2.3 indicates that individuals who are more likely to be self-employed on average receive higher earnings conditional on being self-employed. The estimated degree of sample selection bias in the wage employment model is weaker than that from the self-employment model.

The earnings components of these models show that while being employed in a higher skilled occupation is associated with a significant positive return in both the self and wage employment sectors, this return is higher in wage employment. I also find that the skill premium is larger for immigrants than it is for natives in both sectors[55], which reflects the poor earnings opportunities of immigrants in low skilled occupations. While immigrants in low skilled occupations receive

[53]I believe the sign of the effect of residual housing prices on the self-employment propensity of non-homeowners to be ambiguous ex ante. Increases in local housing prices are expected to result in higher rental costs for non-homeowners, potentially reducing the quantity of assets that can reasonably be allocated to a business. However, since family members are a frequent source of small business financing, an increase in the value of family member's homes may increase one's probably of being self-employed. Regardless of which one of these effect dominates, the effect for for non-homeowners should be smaller than that for homeowners.

[54]I test for the joint strength of the exclusion restrictions, and find a Chi-Square statistic of 38.67 (p-value=0.000), indicating that these exclusion restrictions have significant power in explaining self-employment propensity.

[55]This difference is statistically significant at the 1% level

Table. 2.3. Main Results From Heckman Estimation

Earnings Equation Results				
	(1)	(2)	(3)	(4)
Immigrant	-0.098***	-0.088***	-0.303***	-0.379***
	(0.01)	(0.002)	(0.053)	(0.012)
Skill	0.055***	0.07***	0.052***	0.067***
	(0.002)	(3.703e-04)	(0.002)	(3.845e-04)
Immigrant*Skill	-	-	0.029***	0.041***
	-	-	(0.007)	(0.002)
λ	-0.0765**	-0.023***	-0.0752**	-0.023***
	(0.03)	(0.001)	(0.031)	(0.001)
Selection Equation Results				
	(1)	(2)	(3)	(4)
Housing Residual	0.273	-0.005	0.277	-0.009
	(0.187)	(0.215)	(0.186)	(0.215)
Homeowner*Housing Residual	0.464**	0.009	0.466**	0.098
	(0.223)	(0.265)	(0.222)	(0.264)

(1)- Results from Equation 2.26 (Self-Employment Heckman Equation)
(2)- Results from Equation 2.25 (Wage Employment Heckman Equation)
(3)-Results from Equation 2.26. Allowing different returns to skills for immigrants and natives.
(4)-Results from Equation 2.25. Allowing different returns to skills for immigrants and natives.

a substantial earnings penalty relative to natives in both sectors, this evidence suggests that they fare relatively better in self-employment. The reverse is true for immigrants in high skilled occupations, who appear to fare better in wage employment.

I use Equations 2.30 and 2.31 to convert the estimated returns to skills from these log models into the expected dollar value of skills in both sectors.

$$SkillValue_{i,d,j} = e^{Log(\hat{Y}^i_{d,j})} - e^{Log(\hat{Y}^i_{d,j}) - g_j S_d} \tag{2.30}$$

$$SkillValueInteraction_{i,d,j} = e^{Log(\hat{Y}^i_{d,j})} - e^{Log(\hat{Y}^i_{d,j}) - g2_j S_d - h_j S_d Imm_i} \tag{2.31}$$

Where the individual's expected dollar earnings in an occupation with a skill level of zero are subtracted from their overall expected dollar earnings[56]. These values are to be interpreted as the earnings premium that an individual receives relative to working in an occupation at the bottom of the skill distribution, while accounting for their individual characteristics. Based on Equation 2.30, the average immigrant in the sample is expected to earn \$5.65 per hour

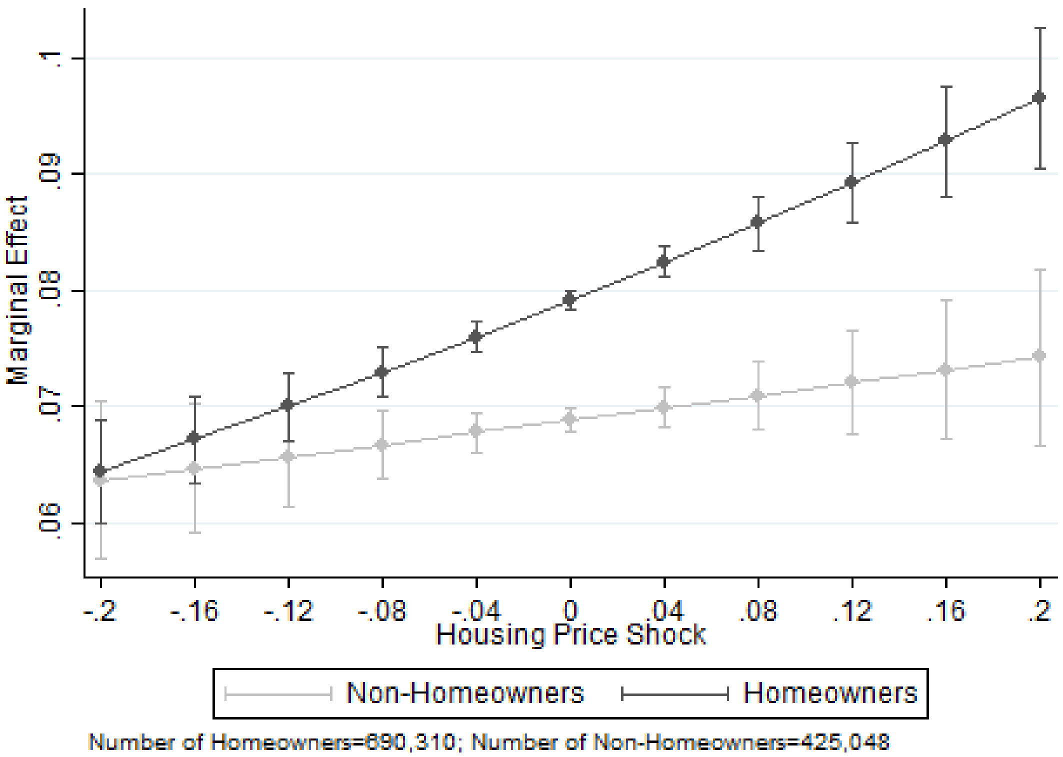
[56] The distributions of skill values for immigrants and natives in both sectors are presented in Figures A2.3-A2.6.

as a result of the skills of their occupation, compared to \$6.43 for the average native. The corresponding figures for the wage sector are \$6.34 for immigrants and \$6.82 for natives.

When allowing skill returns to vary between immigrants and natives, as per Equation 2.31, the average native in self-employment is estimated as earning \$6.09 per hour more than they would earn from an occupation with a skill level of 0. However, since immigrants are more heavily penalized when working in lower skilled occupations, the average immigrant is expected to earn \$8.72 per hour more than they would earn from an occupation of skill level 0. The corresponding values for the wage sector are \$6.55 for natives and \$10.03 for immigrants.

I modify the simple probit in Equation 2.1 by adding the expected dollar values of skills in both the self-employment and wage employment sectors, as calculated in Equations 2.30 and 2.31. By estimating the dollar value of skills separately for both the self-employment and wage employment sectors, I am able to obtain estimates of whether the earnings that an individual is expected to derive from their skills in each sector influences their decision to become self-employed or wage employed.

Fig. 2.9. Estimated Self-Employment Margins Across Residual Housing Price Shocks, Split by Homeowner Status



Marginal effects of unanticipated housing price shocks on the probability of being self-employed. Results from probit estimation found in Table 2.3. Housing price shocks are estimated as the residuals from the fixed effects model found in Equation 2.29 (results in Table A2.2)

$$\Phi = \alpha_1 + \beta_x X_i + \beta_{imm} Imm_i + \tau_{se} SkillValue_{i,d,se} + \tau_{we} SkillValue_{i,d,we} + {}_i \quad (2.32)$$

Where τ_{se} captures the responsiveness of self-employment probability to the value of one's skills in self-employment, while τ_{we} is the corresponding response for wage employment. It is worthwhile noting the interpretation of these coefficients. Conditional on a given skill level, the earnings that are expected to be derived from skills will differ for the same individual between the self-employment and wage employment sectors. These values, as calculated in Equations 2.30 or 2.31, will depend upon the other observed characteristics of the individual, and are not linear combinations of $Skill_i$. τ_{se} can be interpreted as the effect of an additional dollar in expected hourly earnings that is derived from skills in self-employment on the probability of choosing self-employment, having accounted for the equivalent dollar value of skills in wage employment. τ_{we} can be similarly interpreted as the effect of an additional dollar in expected hourly earnings that is derived from skills in wage employment on the probability of choosing self-employment, having accounted for the equivalent dollar value of skills in self-employment.

The probit coefficients for the model in Equation 2.32 can be found in Table 2.4. Unsurprisingly, conditional on how skills are valued in the other sector, high skill values in self-employment are associated with higher predicted self-employment rates, while higher skill values in wage employment are associated with a decrease in the probability of self-employment that is nearly identical in magnitude.

However, consistent with the earlier finding in this chapter that there exists a large within occupation gap in self-employment rates, controlling for skill values does not help to explain the higher self-employment rates of immigrants. The estimated marginal effects for this model are 7.4% for natives and 8.4% for immigrants, nearly identical to the estimated marginal effects of 7.4% and 8.5% in an equivalent model without skill values. This result demonstrates that it is not the case that immigrants are more likely to be self-employed because they are in occupations that carry relatively higher values in self-employment.

For immigrants, working in a low skilled occupation carries a higher penalty in wage employment than it does in self-employment, with this gap in penalties being larger for immigrants

Table. 2.4. Probit Coefficients with Dollar Values of Skills in Self and Wage Employment

Variable	Coefficient	Standard Error
Agriculture	0.621***	0.04
Mining	0.439***	0.06
Utilities	-0.082	0.074
Construction	1.145***	0.027
Manufacturing	0.378***	0.028
Wholesale	0.775***	0.03
Retail	0.733***	0.027
Transportation	0.986***	0.028
Cultural	0.715***	0.032
Finance	0.544***	0.03
Real Estate	0.639***	0.032
Professional	1.18***	0.027
Management	-0.006	0.134
Support	1.132***	0.028
Education	0.239***	0.03
Health	0.643***	0.028
Arts	1.025***	0.031
Food	0.862***	0.029
Other	1.107***	0.028
Female	-0.164***	0.006
Married	0.132***	0.006
High School	0.022*	0.011
Below Bachelor's	0.11***	0.011
Bachelor's	0.247***	0.013
Above Bachelor's	0.302***	0.016
Immigrant	0.096***	0.011
Age	0.054***	0.002
Age2	-3.57e-04***	1.94e-05
Indigenous	-0.007	0.017
Latin American	-0.123***	0.011
Middle Eastern	0.285***	0.025
African	-0.144***	0.013
South Asian	0.113***	0.027
East Asian	-0.014	0.016
Pacific Islander	0.056***	0.056
Occupation Self-Employment Rate	3.881***	0.025
Homeowner	0.117***	0.007
$SkillValue_{i,d,se}$	0.095***	0.009
$SkillValue_{i,d,we}$	-0.097***	0.008
Constant	-4.607***	0.056

N=922,060; χ^2=66902; P-value=0.000, $Pseudo\ R^2$=0.273

Probit coefficients as estimated from Equation 2.32.

Table. 2.5. Probit Coefficients with Dollar Values of Skills in Self and Wage Employment, Allowing Differences in Returns to Skills

Variable	Coefficient	Standard Error
Agriculture	0.616***	0.04
Mining	0.434***	0.06
Utilities	-0.083	0.074
Construction	1.152***	0.027
Manufacturing	0.38***	0.028
Wholesale	0.774***	0.03
Retail	0.732***	0.027
Transportation	0.989***	0.028
Cultural	0.714***	0.032
Finance	0.535***	0.03
Real Estate	0.631***	0.032
Professional	1.176***	0.027
Management	-0.034	0.135
Support	1.134***	0.028
Education	0.236***	0.03
Health	0.635***	0.028
Arts	1.026***	0.031
Food	0.86***	0.029
Other	1.107***	0.028
Female	-0.16***	0.006
Married	0.13***	0.006
High School	0.025**	0.011
Below Bachelors	0.114***	0.011
Bachelors	0.248***	0.013
Above Bachelors	0.299***	0.016
Immigrant	0.151***	0.011
Age	0.054***	0.002
Age2	-3.574e-04***	1.92e-05
Indigenous	-0.004	0.017
Latin American	-0.122***	0.011
Middle Eastern	0.285***	0.025
African	-0.14***	0.013
South Asian	0.138***	0.027
East Asian	-0.008	0.016
Pacific Islander	-0.252***	0.056
Occupation Self-Employment Rate	3.888***	0.025
Homeowner	0.118***	0.007
$SkillValueInteraction_{i,d,se}$	0.009***	0.009
$SkillValueInteraction_{i,d,we}$	0.007***	0.007
Constant	-4.633***	0.056

N=922,060; χ^2=66668; P-value=0.000, $Pseudo\ R^2$=0.273

Probit coefficients as estimated from Equation 2.33.

than it is for natives. This can be seen by comparing columns (3) and (4) of Table 2.3[57]. Consistent with the push theory of self-employment, we may expect that immigrants who work in low skilled occupations will choose self-employment more frequently than natives in the same occupation, as a result of their larger gap in skill values between self-employment and wage employment. In order to investigate this possibility, I estimate Equation 2.33, which includes the skill values that have been estimated while allowing for differences in skill returns between immigrants and natives, as in Equation 2.31. The full set of results from this model can be found in Table 2.5.

$$\begin{aligned}\Phi \quad = \alpha_1 + \beta_x \quad X_i + \beta_{imm} \quad Imm_i + \tau 2_{se} \quad SkillValueInteraction_{i,d,se} \\ + \tau 2_{we} \quad SkillValueInteraction_{i,d,we} + \quad_i\end{aligned} \tag{2.33}$$

However, what I find is not consistent with the aforementioned theory. The estimated marginal effects from this model, found in column b) of Table 2.3, show that in fact the estimated gap in self-employment rates between immigrants and natives expands after allowing for different skill returns for immigrants. This is a reflection of the fact that the occupations in which immigrants are more likely to be self-employed are predominantly not the lower skilled occupations in which their earnings capacity is lower, but rather occupations in the mid to high skill range in which immigrants have a smaller earnings advantage from self-employment compared to wage employment.

This finding stands in contrast with the push theory of self-employment. While immigrants in low skilled occupations do indeed have relatively better earnings opportunities in self-employment, these are not the immigrants who are most likely to actually become self-employed. Rather, it is the immigrants in more highly skilled occupations with better wage sector opportunities who are the most likely to become self-employed, as shown in Figure 2.6.

I estimate additional models where instead of restricting the estimated coefficients in the probit model to be the same for immigrant and natives, I include interaction terms which allow the responsiveness to earnings potential in each sector to differ between immigrants and natives. These are shown in Equations 2.34 and 2.35.

[57] The corresponding full set of earnings results can be found in Tables A2.7 and A2.8. Results from the selection equations can be found in Tables A2.9 and A2.10

$$\Phi = \alpha_1 + \beta_x \quad X_i + \beta_{imm} \quad Imm_i + \tau_{se} \quad SkillValue_{i,d,se} + \tau_{we} \quad SkillValue_{i,d,we} \\ +\varphi_{se} \quad Imm_i \quad SkillValue_{i,d,se} + \varphi_{we} \quad Imm_i \quad SkillValue_{i,d,we} + \quad {}_i \tag{2.34}$$

$$\Phi = \alpha_1 + \beta_x \quad X_i + \beta_{imm} \quad Imm_i + \tau 2_{se} \quad SkillValueInteraction_{i,d,se} \\ +\tau 2_{we} \quad SkillValueInteraction_{i,d,we} + \varphi 2_{se} \quad Imm \quad SkillValueInteraction_{i,d,se} \\ +\varphi 2_{we} \quad Imm \quad SkillValueInteraction_{i,d,we} + \quad {}_i \tag{2.35}$$

It is noted from the results in Tables 2.7 and 2.8 that immigrants have a stronger response to skill values than natives. Conditional on the skill value in wage employment, a higher value of skills in self-employment is expected to increase an immigrant's propensity towards self-employment by more than is the case for natives. Likewise, conditional on the skill value in self-employment, a higher value of skills in wage employment is expected to decrease the immigrants propensity towards self-employment by more than is the case for natives. This is consistent with the theoretical predictions from Section 2.3, which predict that prospective immigrants will be more likely to be deterred from migrating in order to pursue a given occupation/sector pair if that choice is expected to yield low earnings.

However, the most noteworthy result in Tables 2.7 and 2.8 is that unlike natives, immigrants' responsiveness to skill values in each sector is asymmetric. Higher skill values in self-employment increase their propensity towards self-employment by more than higher skill values in wage employment decrease their self-employment propensity. This indicates that it is not the gap in earnings potential between the two sectors that is predominantly driving immigrant self-employment. Rather, immigrants are more likely to be self-employed when they have strong self-employment earnings opportunities, even though their skills may be just as highly valued in wage employment. This finding is also consistent with the theoretical predictions, and supports the idea that out of all sector and earnings level combinations, immigrants are least likely to be deterred from migrating in order to pursue self-employment in high earnings occupations.

A similar result is found regardless of whether the returns to skills differs between immigrants and natives, as can be seen by comparing the coefficients on the skill value terms (and their interactions with the immigrant variable) in Tables 2.7 and 2.8, the latter of which includes skill values calculated from the Heckman Correction that allows for differences in returns to skills between immigrants and natives. Controlling for this difference in the responsiveness to skill values is able to explain the entire gap in self-employment rates between immigrants and natives, as can be seen from the coefficient on "immigrant" in Tables 2.7 and 2.8, and the marginal effects in columns 3 and 4 of Table 2.6. In other words, no gap in self-employment rates between immigrants and natives is expected when skill values are equal to zero. Furthermore, it is noteworthy that the estimated margins are nearly identical regardless of whether the returns to skill varies between immigrants and natives.

These results stand in contrast with some of the existing literature on immigrant self-employment. Although I find evidence that immigrants in low skilled occupations face relatively weaker earnings opportunities in wage employment, I do not find any evidence to support the idea that this penalty pushes immigrants into self-employment. In fact, those immigrants who may be expected to be pushed into self-employment are relatively *less* likely to be self-employed than their counterparts with stronger wage employment opportunities. Nor do these results support the idea that immigrants are more entrepreneurial and have better earnings opportunities in self-employment. Despite receiving slightly higher returns to skills in self-employment than natives, immigrants have expected self-employment earnings that are well below those of comparable natives across the entirety of the skill distribution.

Instead, these results show that even though immigrants do not have relatively better opportunities in self-employment, they are more likely to be self-employed when their expected earnings from self-employment are high, even in cases where they also have high estimated

Table. 2.6. Estimated Margins of Self-Employment Rates

	(1)	(2)	(3)	(4)
Natives	.0737	.0728	.0756	.07577
	(3.43e-04)	(3.4e-04)	(4.38e-04)	(4.65e-04)
Immigrants	.0841	.0895	.0746	.0738
	(1.06e-03)	(1.14e-03)	(1.5e-03)	(1.62e-03)

(1)-Controlling for Dollar Value of Skill, as in Equation 2.32.
(2)-Controlling for Dollar Value of Skills, as in Equation 2.33.
(3)-Controlling for Dollar Value of Skills, as in Equation 2.32, interacted with immigrant status.
(4)-Controlling for Dollar Value of Skills, as in Equation 2.33, interacted with immigrant status.

Table. 2.7. Probit Coefficients with Dollar Values of Skills in Self and Wage Employment, Interaction with Immigrant Status

Variable	Coefficient	Standard Error
Agriculture	0.612***	0.04
Mining	0.441***	0.06
Utilities	-0.08	0.074
Construction	1.154***	0.027
Manufacturing	0.379***	0.028
Wholesale	0.772***	0.03
Retail	0.728***	0.027
Transportation	0.989***	0.028
Cultural	0.717***	0.032
Finance	0.537***	0.03
Real Estate	0.625***	0.032
Professional	1.182***	0.027
Management	-0.063	0.14
Support	1.131***	0.028
Education	0.232***	0.03
Health	0.633***	0.028
Arts	1.021***	0.031
Food	0.857***	0.029
Other	1.105***	0.028
Female	-0.163***	0.006
Married	0.133***	0.006
High School	0.019*	0.011
Below Bachelor's	0.109***	0.011
Bachelor's	0.252***	0.013
Above Bachelor's	0.306***	0.016
Immigrant	-0.009	0.017
Age	0.056***	0.002
Age2	-3.676e-04***	1.95e-05
Indigenous	-0.008	0.017
Latin American	-0.106***	0.011
Middle Eastern	0.268***	0.025
African	-0.145***	0.013
South Asian	0.138***	0.028
East Asian	-0.011	0.016
Pacific Islander	-0.273***	0.057
Occupation Self-Employment Rate	3.902***	0.025
Homeowner	0.118***	0.007
$SkillValue_{i,dq,se}$	0.088***	0.009
$SkillValue_{i,dq,we}$	-0.092***	0.008
Immigrant*$SkillValue_{i,d,se}$	0.15***	0.013
Immigrant*$SkillValue_{i,d,we}$	-0.115***	0.011
Constant	-4.625***	0.057

N=922,060; χ^2=66933; P-value=0.000, $Pseudo\ R^2$=0.273

Probit coefficients as estimated from Equation 2.34.

Table. 2.8. Probit Coefficients with Dollar Values of Skills in Self and Wage Employment, Allowing Differences in Returns to Skills, Interacted with Immigrant Status

Variable	Coefficient	Standard Error
Agriculture	0.605***	0.04
Mining	0.438***	0.06
Utilities	-0.077	0.074
Construction	1.159***	0.027
Manufacturing	0.378***	0.028
Wholesale	0.768***	0.03
Retail	0.721***	0.027
Transportation	0.985***	0.028
Cultural	0.716***	0.032
Finance	0.526***	0.03
Real Estate	0.611***	0.032
Professional	1.178***	0.027
Management	-0.095	0.141
Support	1.128***	0.028
Education	0.226***	0.03
Health	0.624***	0.028
Arts	1.016***	0.031
Food	0.853***	0.029
Other	1.101***	0.028
Female	-0.162***	0.006
Married	0.133***	0.006
High School	0.019*	0.011
Below Bachelor's	0.111***	0.011
Bachelor's	0.257***	0.013
Above Bachelor's	0.308***	0.016
Immigrant	-0.019	0.019
Age	0.057***	0.002
Age2	-3.784e-04***	1.98e-05
Indigenous	-0.008	0.017
Latin American	-0.102***	0.011
Middle Eastern	0.265***	0.025
African	-0.143***	0.013
South Asian	0.115***	0.027
East Asian	-0.012	0.016
Pacific Islander	-0.279***	0.057
Occupation Self-Employment Rate	3.913***	0.025
Homeowner	0.119***	0.007
$SkillValueInteraction_{i,d,se}$	0.101***	0.01
$SkillValueInteraction_{i,d,we}$	-0.103***	0.009
Immigrant*$SkillValueInteraction_{i,d,se}$	0.066***	0.01
Immigrant*$SkillValueInteraction_{i,d,we}$	-0.038***	0.008
Constant	-4.655***	0.058

N=922,060; χ^2=66866; P-value=0.000, $Pseudo\ R^2$=0.2734

Probit coefficients as estimated from Equation 2.35.

wage employment earnings potential. This is consistent with the results from the theoretical model, which predict that in the presence of a migration cost, migrants will tend to appear to be more responsive to earnings potential than non-migrants. It also predicts that this stronger response to earnings potential should be especially pronounced for expected earnings in the self-employment sector. The results in this section are consistent with these predictions, and suggest that the higher self-employment rates among immigrants arise as a result of the non-random selection into migrating that would be induced by a migration cost.

2.6 Results with Internal Migrants

As outlined in Section 2.3, one of the potential explanations for the higher rates of self-employment among immigrants in high skilled occupations is that those individuals who choose to migrate are also relatively more likely to choose to pursue self-employment in high skilled occupations. The empirical results in the previous section are consistent with this hypothesis. However, if this is indeed the case, then we should expect qualitatively similar results (albeit perhaps of a smaller magnitude due to smaller migration costs) when we compare native born Americans who have migrated from one state to another within the U.S. to those natives who remain in their state of birth[58].

The subsample of internal migrants consists of individuals who reside in a state that is not their state of birth, but who are married to someone who was born in the same state that they were[59]. All immigrants, and natives who are not married or in a common law relationship, are excluded from the analysis in this section. The total subsample consists of 689,500 married non-immigrants, approxiately 14.7% of whom are married to someone from their state of birth, but now reside in another state.

Estimating the models from the previous section with this subsample of natives in place of immigrants has the advantage of preserving the fact that these individuals have demonstrated a willingness to relocate. However, in terms of factors such as language, culture, and recognition

[58] The results from the model that was outlined in Section 2.3 are expected to hold regardless of whether the regions r comprise different countries, or whether they comprise different parts of the same country.

[59] Since the immigrant population of interest consists of those immigrants who migrated to the U.S. as adults, a proper comparison would use only native born individuals who relocated to another state as an adult. Unfortunately, no information on when an individual relocated states is present in the ACS data. In order to address this address, I restrict the internal migrants subsample to consist only of individuals who are married to someone from their state of birth. While this would not capture the full subsample of individuals who migrated as adults, it is unlikely that an individual in this subsample would have moved states during their childhood.

of credentials, they are otherwise similar to the native born population as a whole. I find that these internal migrants are equally likely to be self-employed as native non-migrants, with the estimated margins from a model that controls for demographics and industry being approximately 8.8% and 9% respectively[60], as shown in Table 2.9. After accounting for the skill value of their occupations in each sector, with the margins shown in column 2), I fail to find any evidence that internal migrants are predominantly found in occupations for which skills are relatively more highly valued in self-employment. However, when allowing the effect of skill values on self-employment probability to vary between internal migrants and native non-migrants, I find similar results as those from the immigrant-native comparison. Internal migrants have a stronger response to how their skills are valued, as shown in Table 2.11, and in particular have a stronger response to skill values in the self-employment sector. Accounting for this difference in the responsiveness to skill values creates a substantial gap in self-employment rates between these two subsamples, with internal migrants being significantly less likely to be self-employed after accounting for this differential, as shown in column 3) of 2.9. Since natives who have exhibited a willingness to relocate exhibit similar self-employment choice patterns as immigrants, this strengthens the case the the high self-employment rates among immigrants are a characteristic of individuals who have migrated, as opposed to a characteristic that is specific to immigrants coming from another country.

To mitigate the possibility that self-employment patterns among native migrants are a response to a lack of local knowledge or state specific occupational requirements, rather than a characteristic of individuals who have moved, I further restrict the subsample to consist only of individuals who have resided at their current address for at least 10 years. Since such state specific characteristics are expected to affect recent arrivals the most, the gap should be reduced following this restriction. Although self-employment rates are higher among these individuals (as a result of them being older), there is no discernible difference between the gaps in self-employment rates that arise from this estimation, as compared to the models that consisted of all of the native migrants and native non-migrants. The marginal effects from models estimated with this subsample are found in Table 2.10.

[60] Note that the high self-employment rates are the result of these population consisting only of married individuals.

Table. 2.9. Estimated Margins of Self-Employment Rates for Internal Migrants and Native Non-Migrants

	(1)	(2)	(3)
Native Non-Migrants	.09	.09	.091
	(4.87e-04)	(4.87e-04)	(6.43e-04)
Internal Migrants	.088	.088	.083
	(1.2e-03)	(1.2e-03)	(2.42e-03)

(1)-Demographic and Industry Controls Only.
(2)- Controlling for Dollar Value of Skills, Same Skill Price for Both Subsamples.
(3)- Controlling for Dollar Value of Skills, Same Skill Price for Both Subsamples, Interaction with Internal Migrant variable.

Table. 2.10. Estimated Margins of Self-Employment Rates for Internal Migrants and Native Non-Migrants Who Have Resided at Their Current Address for at Least 10 Years

	(1)	(2)	(3)
Native Non-Migrants	.101	.102	.102
	(6.15e-04)	(6.17e-0.4)	(7.92e-04)
Internal Migrants	.101	.099	.095
	(1.57e-0.3)	(1.55e-0.3)	(3.14e-03)

(1)-Demographic and Industry Controls Only.
(2)- Controlling for Dollar Value of Skills, Same Skill Price for Both Subsamples.
(3)- Controlling for Dollar Value of Skills, Same Skill Price for Both Subsamples, Interaction with Internal Migrant variable.

2.7 Robustness Checks

2.7.1 Results with Immigrants Who Arrived As Children

The results presented in this paper thus far have restricted the subsample of immigrants to consist only of those immigrants who arrived in the United States aged 18 or older. For the purposes of comparison, in this subsection I include an additional subsample of immigrants who were between the ages of 4 and 14 inclusive at the time of immigration[61]. After accounting for demographic characteristics[62], I find that immigrants who arrived as children are significantly more likely to be self-employed than natives, and have self-employment rates that are similar

[61] Immigrants who arrived under the age of 4 are excluded since they would have completed the entirety of their formal education in the United States, and their adjustment to the U.S. would be expected to have been easier than that of their older peers. Furthermore, there are a disproportionate number of immigrants who arrived in their mid to late teens relative to younger ages. It is plausible that a significant portion of these individuals were not tied movers, but rather made the decision to migrate to the United States, similar to their counterparts who were adults at the time of immigration.

[62] It is noteworthy that the raw self-employment rates among those who immigrated as children is similar to that of natives (7.7% vs. 7.5%), however immigrants who arrived as children are on average 4 years younger than natives at the time of the survey.

Table. 2.11. Probit Coefficients with Dollar Values of Skills in Self and Wage Employment, Interacted with Internal Migrant Status

Variable	Coefficient	Standard Error
Agriculture	0.781***	0.052
Mining	0.413***	0.077
Utilities	0.02	0.097
Construction	1.346***	0.038
Manufacturing	0.566***	0.038
Wholesale	0.851***	0.04
Retail	0.855***	0.037
Transportation	0.999***	0.039
Cultural	0.751***	0.044
Finance	0.69***	0.038
Real Estate	0.752***	0.042
Professional	1.359***	0.036
Management	0.237	0.157
Support	1.29***	0.039
Education	0.198***	0.04
Health	0.712***	0.037
Arts	1.075***	0.043
Food	1.051***	0.04
Other	1.197***	0.038
Female	-0.175***	0.009
High School	-8.107e-04	0.017
Below Bachelor's	0.107***	0.017
Bachelor's	0.276***	0.02
Above Bachelor's	0.389***	0.023
Internal Migrant	-0.066***	0.025
Age	0.049***	0.003
Age2	-2.925e-04***	3.11e-05
Indigenous	-0.011	0.022
Latin American	-0.106***	0.018
Middle Eastern	0.21***	0.049
African	-0.184***	0.02
South Asian	0.254**	0.099
East Asian	-0.066*	0.035
Pacific Islander	-0.225***	0.084
Occupation Self-Employment Rate	4.056***	0.038
Homeowner	0.134***	0.01
$Skillvalue_{I,d,se}$	0.091***	0.01
$Skillvalue_{I,d,we}$	-0.103***	0.009
$InternalMigrant\quad SkillValue_{i,d,se}$	0.046***	0.012
$InternalMigrant\quad SkillValue_{i,d,we}$	-0.037***	0.01
Constant	-4.48***	0.094

N=456,955; χ^2=38,044; P-value=0.000, $Pseudo\ R^2$=0.2788

Probit coefficients as estimated from Equation 2.35 with internal migrants in place of immigrants.

to those of adult immigrants. However, despite a similar gap in self-employment rates, I find no evidence that this gap can be accounted for by how these immigrants respond to how their skills are valued in each sector when estimating Equation 2.36[63]. The results in appendix Table A2.11 show that immigrants who arrived as children respond more strongly to the value of their skills in the self-employment sector than in the wage employment sector, however this difference is not as strong as it is for adult immigrants[64]. Furthermore, accounting for the difference in the responsiveness to skill values fails to account for a notable portion of the gap in self-employment rates between immigrants who arrived as children and natives, with the estimated marginal effects from this model being approximately 8.6% and 7.5% respectively.

$$\begin{aligned}\Phi = \alpha_1 + \beta_x \quad X_i + \beta_{imm} \quad Imm_i + \tau 2_{se} \quad SkillValueInteraction_{i,d,se} \\ + \tau 2_{we} \quad SkillValueInteraction_{i,d,we} + \varphi 2_{se} \quad AdultImm \quad SkillValueInteraction_{i,d,se} \\ + \varphi 2_{we} \quad AdultImm \quad SkillValueInteraction_{i,d,we} + \pi 2_{se} \quad ChildImm \quad SkillValueInteraction_{i,d,se} \\ + \pi 2_{we} \quad ChildImm \quad SkillValueInteraction_{i,d,we} + \quad _i\end{aligned} \tag{2.36}$$

These results are also consistent with the idea that the stronger response to earnings potential in self-employment than wage employment among adult immigrants is the result of characteristics of individuals who choose to migrate. Since child immigrants were predominantly tied movers, they were not self-selected in the same way that adult immigrants were. I do not, however, rule out the possibility that child immigrants have different preferences than natives that result from their family's characteristics. The fact that immigrants who arrived as children have a higher self-employment rate than their native born counterparts could be explained by patterns of intergenerational transmission of self-employment, such as those observed by Hout and Rosen (2000) and Andersson and Hammarstedt (2010).

[63] Adult Imm in this equation is equivalent to Imm in earlier expressions, the word Adult is included in order to clarify the distinction between them and immigrants who arrived as children

[64] The gap in the response to skill values in the self-employment sector relative to the wage employment sector among adult immigrants, given by $\tau 2_{se} + \tau 2_{we} + \varphi 2_{se} + \varphi 2_{we}$, is statistically larger than the equivalent gap among child immigrants $\tau 2_{se} + \tau 2_{we} + \pi 2_{se} + \pi 2_{we}$ at the 1% level.

2.7.2 Wealth Levels

A further potential concern is that these results could be driven by the behaviour of wealthy immigrant entrepreneurs. In order to mitigate these concerns, I re-estimate Equation 2.35 while including the value of one's home measured in dollars (the best proxy for wealth available in the ACS data) along with it's interactions with immigrant status, dollar value of skills, and a three way interaction between immigrant status, skill value and housing value. By necessity, this subsample is restricted to consist only of homeowners. While the results from this model (found in Table A2.12) do suggest that individuals who own more valuable homes are more likely to be self-employed, there is no indication that such wealthy individuals, whether they are immigrants or natives, respond more strongly to the value of their skills in self-employment. Indeed, there is weak evidence that they may be slightly less responsive. These results indicate that the stronger response to skill values in self-employment among immigrants is not being driven by wealthy immigrants.

2.7.3 Undocumented Immigrants

Another potential explanation for higher rates of self-employment among immigrants is the potential that undocumented immigrants are more likely to pursue self-employment. The ACS data does not contain a measure of whether an immigrant is documented or not, however since the overwhelming majority of undocumented immigrants are originally from Mexico (Bratsberg 1995), I estimate Equation 2.35 while separating immigrants from Mexico from the rest of the immigrant population. First of all, it is worthwhile noting that immigrants from Mexico form an exception to the rule of immigrants being more likely to be self-employed than natives. After accounting for demographic characteristics, only 6.1% of immigrants from Mexico are estimated as being self-employed, compared to 7.3% of native born Americans and 8.8% of immigrants from countries other than Mexico. When estimating the responsiveness of immigrants from Mexico and immigrants from other countries to the value of their skills, as in Table A2.13, I find that both subgroups of immigrants show a stronger response to earnings potential in self-employment than in wage employment, with no such gap occurring among the native born population. For both immigrants from Mexico and immigrants from other countries, this differential response to earnings potential is able to explain a portion of their propensity towards self-employment, as shown in Table A2.14. However, this is not the case for natives. Since the response of self-employment rates to skill values is present and significant for those immigrants who are not from Mexico, it is unlikely that the main set of results is being driven by undocumented immigrants.

2.8 Conclusion

The literature on self-employment of immigrants has tended to focus on their relative earnings disadvantage in wage employment as being the primary cause of their higher self-employment rates. However, the evidence presented in this chapter suggests that the story is more complex. I do find that in comparison to natives, immigrants in low skilled occupations have weaker earnings opportunities in wage employment relative to self-employment. However, these are not the occupations in which the immigrant-native gap in self-employment rates is most pronounced.

Instead, I find evidence that the self-employment patterns of immigrants are similar to those of cross-state internal migrants within the United States. This suggests that these patterns are largely characteristics of individuals who migrate, rather than being exclusive to migrants from another country. This is further supported by the theoretical model introduced in this paper, which suggests that migration costs are more likely to deter prospective migrants from migrating in order to pursue low earnings occupations in self-employment.

A key takeaway from this paper is that immigrants tend to choose the self-employment sector precisely in those occupations in which it confers the smallest earnings advantage. Although the theoretical model in this paper provides an explanation for why this is the case, it is still the case that this results in a larger earnings disadvantage for immigrants relative to natives. A suggestion for future research would be to quantify the macroeconomic impact of these occupation/sector choice patterns among immigrants, and to assess whether policy changes could reduce this impact.

Another key implication of this research is that it suggests that even in the absence of discrimination, skill differentials, or occupational licensing barriers, we should continue to see higher self-employment rates among immigrants in high skill occupations. This suggests the higher self-employment rates among immigrants should be thought of as a by-product that results from individuals having actually made the choice to migrate to the United States.

Chapter 3: The Skill Distance Between Occupations and Post-Secondary Fields of Study

3.1 Introduction

The question of whether post-secondary graduates tend to become employed in an occupation that is a good match for their education is one that captures extensive attention among policy makers and prospective post-secondary students. Field of study choices are highly dependent upon a student's beliefs on the likelihood that they will end up with a job that they will enjoy (Zafar 2013), and that will pay well (Arcidiacono et al. 2012, Wiswall and Zafar 2015). Furthermore, encouraging more post-secondary students to pursue STEM fields has increasingly become a major public policy objective (Rask 2010, Chen 2013), with the end goal of increasing the number of workers who are using essential STEM (Science, Technology, Engineering and Math) related skills in their jobs (Chen 2013). Given the high degree of importance that both students and policy makers place on the eventual occupational outcomes of post-secondary graduates, there has been a growing academic literature on the topic of field of study-occupation matching.

This literature has divided occupation-field of study mismatch into two broad categories, known as vertical mismatch and horizontal mismatch (Nordin et al. 2010). An individual is vertically mismatched if they are employed in an occupation that typically requires less (or occasionally more) education than the individual actually possesses. Horizontal mismatch indicates that the *types* of skills that individuals use in their job differ from those that we would expect on the basis of their field of study. The example that Sloane (2003) uses for horizontal mismatch is an English major being employed as a Statistician.

The most commonly used methods of measuring mismatch include; self-reported mismatch (Robst 2007a), job analysis mismatch (Nordin et al. 2010, Montt 2017), and realized mismatch (Verhaest and Omey 2010). Job analysis mismatch occurs when an individual is not employed in a given occupation (or narrow range of occupations) that is associated with their field of study. Realized mismatch occurs when an individual is employed in an occupation that is unusual for graduates of their field, relative to the population as a whole. The prevalence of mismatch, and to a lesser extent the differences in mismatch across fields of study and demographics, differ depending upon which type of measure is used. The estimated prevalance of mismatch tends to be higher when job analysis measures are used (Sellami et al. 2018). Science, Technology, Engineering and Math (STEM) graduates are consistently found to be among the most likely to be well matched, along with Business graduates (Robst 2007a, Nordin et al. 2010, Lemieux 2014). It is also found that individuals with graduate degrees are less likely to be mismatched than individuals with lower levels of post-secondary education. Gender differences are less clear, with Robst (2007b) finding that men are less likely to be mismatched than women, while Yuen et al. (2010) finds similar rates of mismatch for both men and women.

Recent work by McGuinness et al. (2018) has noted that mismatch on the basis of an individual's level of education does not necessarily imply that they are mismatched on the basis of the types of skills that are associated with their occupation. Following this work, Reis (2018) introduces a novel method of measuring mismatch based on the distance between the tasks that an individual performs in their job, and the tasks that we should expect them to use in their job based on their post-secondary field of study. Using Brazilian data, Reis (2018) finds that this distance based measure captures an important element of mismatch that is missing from the more traditional binary measure, since there are significant differences in the *degree* to which one is mismatched.

In this paper, we use a measure of mismatch similar to that of Reis (2018)[65] and apply it to the Canadian context. This fills a significant gap in the field of study-occupation mismatch literature, as little is known about the degree or severity of this mismatch in Canada. We explore these differences in occupation-field of study distance across fields of study, levels of post-secondary education, and gender. Overall, we find that post-secondary graduates tend to be employed in occupations with skill requirements that are relatively close to the skills that their field of study trains individuals for. However, in contrast to the literature using other methods of measuring mismatch, we find that graduates of STEM fields end up, on average, in occupations that use skills that are the furthest away from what we would expect based on their field. This hitherto unknown element of occupation-field of study mismatch indicates that despite the fact that STEM graduates are frequently employed in occupations that are strong

[65] While our measure of skill mismatch is similar to that of Reis (2018), it is not identical. Reis measures occupational skills by the list of tasks that are considered as being important to a given occupation. Our use of O*NET provides a continuous measure of the level and importance of a broad range of skills to the occupation. We believe that the use of O*NET is more appropriate for the Canadian context.

matches for their field, those who are not tend to be employed in occupations that use a highly dissimilar set of skills. This dissimilarity in skills for those who are not well matched is stronger for STEM graduates than it is for graduates of other post-secondary fields. This element of field-occupation mismatch is not captured using traditional binary measures of mismatch, and is one of the main contributions of this paper.

We also find, consistent with the literature, that individuals with graduate degrees are usually well matched in comparison to their counterparts with lower levels of post-secondary education. However, this gap closes over the course of the lifecycle. We also find that among college graduates, women tend to be employed in occupations that are closer to their field of study than men. No such difference is found among bachelor's and graduate degree holders. However, we find that while women tend to study fields that are more commonly associated with better field-occupation matches, men tend to be better matched conditional on a field of study. After re-weighing to assign a common field of study distribution for men and women, we find that men are better matched than women at the college and bachelor's degree level, with no discernable difference at the graduate degree level.

In addition, we find that even conditional on the skills of their occupation,individuals who studied fields that are more dissimilar to their occupation receive a modest earnings penalty. This penalty is larger for women than it is for men, and is also larger for individuals with a college degree compared to those with bachelor's or graduate degrees. While we do find evidence of an earnings penalty, this penalty is modest compared to that which has been found in the majority of the literature on mismatch and earnings (Robst 2007a, Nordin et al. 2010).

The remainder of this paper is divided into 4 sections. Section 3.2 describes the data, Section 3.3 shows the prevalence of mismatch, Section 3.4 estimates the earnings effects of mismatch, and Section 3.5 concludes.

3.2 Data

Our data consists of Panels 4 and 5 of the Survey of Labour and Income Dynamics (SLID). These panels cover the years 1999-2004 and 2002-2007 respectively. SLID panels prior to 4 and after 5 contain field of study only for those individuals who were post-secondary students at the time of the survey. Using these waves of SLID would limit our scope of analysis to recent graduates only, therefore we do not make use of these panels. We exclude individuals with no post-secondary education as well as those who do not report a field of study. We also restrict our analysis to individuals who are employed at the time of the survey, and we further exclude individuals who were post-secondary students at the time of the survey. This leaves us with 24,170 observations on 6830 individuals.

3.2.1 O*NET

The O*NET database from the U.S. Department of Labor contains measures of the degree to which various skills[66] are required for each of the 960 occupations that are classified using Standard Occupational Classification (SOC) codes. There are 109 such skills, with these skills being as diverse as Oral Expression, Finger Dexterity, and Mathematical Reasoning. We match each occupation NOC code in the SLID data to it's SOC counterpart, with these occupations being subsequently matched to their skill requirements.

In addition, using the National Center for Educational Statistics crosswalk between CIP field of study codes and NOC occupation codes, we link each field of study to a set of occupations for which that field prepares an individual[67]. The skills of those occupations are then linked to graduates of the field in question. For example, graduates of a bachelor's of education program are trained to be a teacher, and therefore we link the skills that are associated with a teacher to graduates of a bachelor's of education program. In cases where a field is linked to multiple occupations, we assign an equally weighted average of these occupation's skills to be the skills of a graduate of the field in question. As shown in Table 3.1, nearly half of all fields of study are linked to a single occupation, while over 90% of fields are linked to 4 occupations or fewer.

We aggregate the expected skill levels that are described above using Principal Component Analysis on the skills that are derived solely from individual's occupations. The top 5 components are used in further analysis, with the eigenvectors of these components being found in Table A3.1. These components are denoted $Comp1occ_i$, $Comp2occ_i$, $Comp3occ_i$, $Comp4occ_i$, $Comp5occ_i$. By applying the eigenvectors in Table A3.1 to the skills as they are derived from fields of study, we are able to create similarly aggregated skills that associated

[66] We include the Skill, Abilities, and Knowledge requirements for each occupation, as classified by O*NET. However, for simplicity we simply refer to these collectively as skills.

[67] For more information on the CIP-SOC crosswalk, see https://nces.ed.gov/ipeds/cipcode/resources.aspx?y=55

Table. 3.1. Number of Field-Occupation Matches

(1)	(2)	(3)	(4)
1	792	46.05%	46.05%
2	366	21.28%	67.33%
3	301	17.5%	84.83%
4	136	7.91%	92.73%
5	72	4.19%	96.92%
6	28	1.63%	98.55%
7	10	0.58%	99.13%
8	5	0.29%	99.42%
9	3	0.17%	99.59%
10	3	0.17%	99.77%
12	2	0.12%	99.88%
13	1	0.06%	99.94%
15	1	0.06%	100%

Column (2) indicates the number of fields of study that match to the number of occupations in Column (1). For example, the first row indicates that 792 post-secondary fields of study are matched to a single occupation. Column (3) indicates the percentage of fields of study that match with the number of occupations given in Column (1). Column (4) indicates the percentage of fields of study that match to the number of occupations given in Column (1) (or fewer)

with one's field of study[68].

We use the eigenvectors from this estimation in order to create similar aggregated skills using the field of study approach, denoted as $Comp1f_i, Comp2f_i, Comp3f_i, Comp4f_i, Comp5f_i$.

In order to estimate the match between an individual's occupation and their field of study, we calculate the Euclidean distance between the skills that are associated with their field of study and the skills that are associated with their occupation.

$$\begin{aligned} Distance_i = &|Comp1occ_i - Comp1f_i| + |Comp2occ_i - Comp2f_i| + |Comp3occ_i - Comp3f_i| \\ &+ |Comp4occ_i - Comp4f_i| + |Comp5occ_i - Comp5f_i| \end{aligned} \quad (3.1)$$

The larger that the above distance measure is, the more dissimilar the skills that an individual uses in their job are from the skills that we would have expected them to use in their job. If an individual studies a given field that is associated with high levels of analytical skills and communication skills, but ends up working in an occupation that uses little of both of this

[68] The eigenvectors that result from performing Principal Component Analysis on occupational skills are utilized in order to aggregate the skills that derived from field of study. We use this approach, as opposed to performing Principal Component Analysis on the skills that are derived from fields of study, in order to ensure that the two skill aggregations are capturing the same type of skill.

Table. 3.2. Distance Scores for Economics Graduates

Percentile	Distance	Occupation
Minimum	3.097472	Mathematicians, Statisticians and Actuaries
25th	14.80047	Post-Secondary TAs and RAs
Median	20.4524	Cooks
75th	26.39334	Carpenters
Maximum	36.45193	Maîtres d'hôtel and Hosts / Hostesses

This table shows the occupations that have the lowest and highest distance score for graduates of Economics programs, along with the occupations at the 25th, 50th, and 75th percentiles of the distance distribution.

Table. 3.3. Distance Scores for Acting Graduates

Percentile	Distance	Occupation
Minimum	6.152482	Actors and Comedians
25th	19.9924	Travel Counsellors
Median	25.32885	Logging and Forestry Labourers
75th	30.77629	Other Products Machine Operators
Maximum	51.94054	Architects

This table shows the occupations that have the lowest and highest distance score for graduates of Acting programs, along with the occupations at the 25th, 50th, and 75th percentiles of the distance distribution.

types of skills, the distance measure will be large. Although level of education does not factor in to the distance score, this characteristic of the distance measure is broadly similar to the measures of vertical mismatch that exist in the literature.

However, it is worthwhile noting that higher distances indicate dissimilar skills, but not necessarily low skills. If an individual studies a field that is associated with high levels of analytical skills and low levels of communication skills, but ends up in an occupation with the opposite characteristics, they will be deemed poorly matched. This is broadly similar to horizontal mismatch as it has been defined in the literature. An individual will be mismatched if they are using the wrong types of skills, even if they are still in an occupation that is highly skilled overall.

Therefore, our distance measures incorporates elements of both vertical and horizontal mismatch. It will therefore be different for individuals who studied different fields, even conditional on working in the same occupation. Tables 3.2 and 3.3 show the closest and furthest occupations, as well as the occupation at each quartile of the distance distribution, for graduates of Economics or Acting programs.

3.3 Average Matches Between Fields of Study and Occupations

3.3.1 Results with Random Field/Occupation Assignments

We first examine the distance scores that would arise if post-secondary graduates were to be randomly assigned to occupations such that the probability of working in a given occupation does not depend upon one's field of study, and is always equal to the proportion of the overall population that works in that occupation, as shown in Equation 3.2.

$$Prob_i(d = d \,|\, f = f\ ) = Prob_i(d = d\ ) \quad d, f \tag{3.2}$$

Where d indicates an individual's occupation and f indicates their field of study, with d' and f' representing a given occupation and a given field of study respectively. Figure 3.1 shows that with this random assignment, the mean distance score for each broad field of study is within the 65-73 range. This is to be anticipated, since this is approximately the midpoint of the distance distribution, with the highest possible distance score being 128. It is also worthwhile noting that there is only minimal heterogeneity in the average distance score across broad fields of study. This indicates that the skills of all fields of study are similarly applicable across occupations. That is, we do not observe fields of study that train individuals for a highly specialized set of skills that are outliers within the occupation distribution.

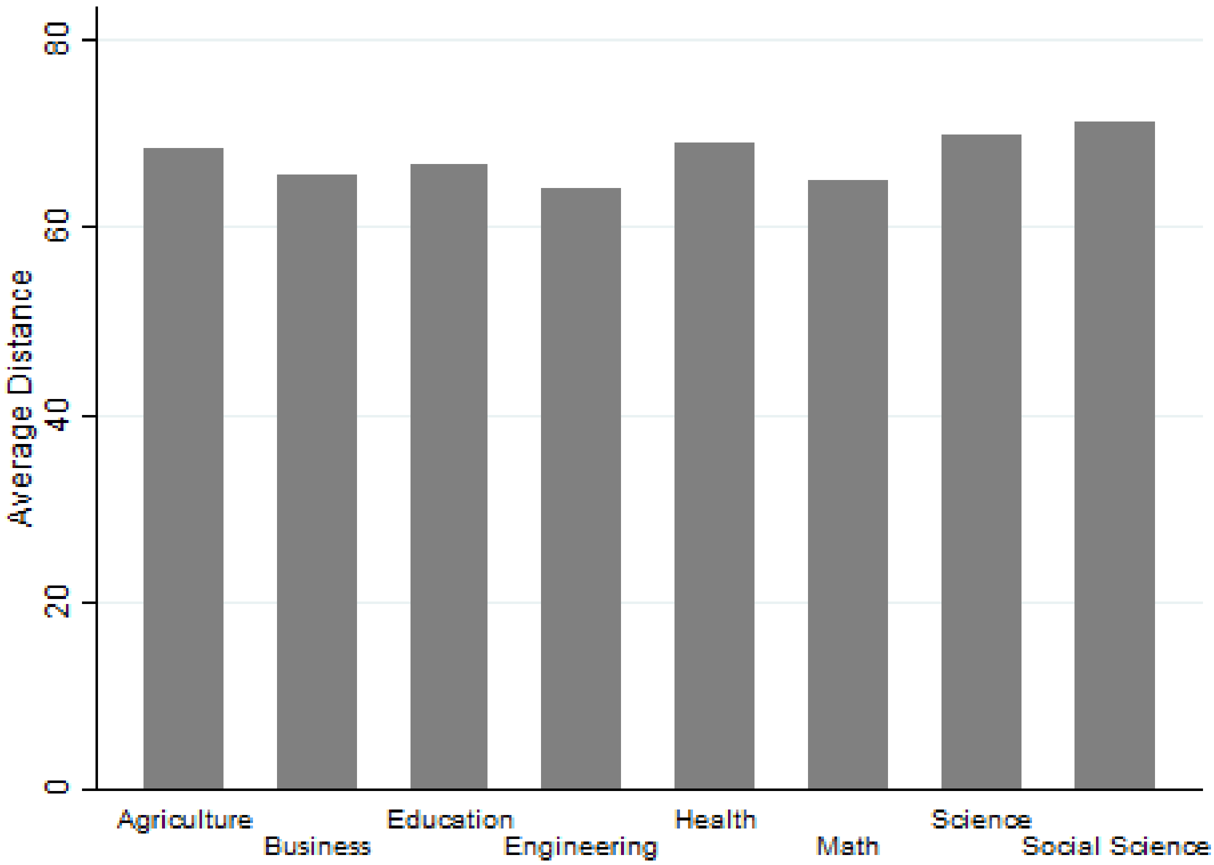
3.3.2 Lifecycle Analysis of Actual Distances

We examine the expected path that distance will take over the course of the lifecycle. In doing so, we choose not to compare the average distance of individuals who, at the time of the survey, completed their degree k years ago with the average distance of individuals who completed their degree more recently. This is because using this approach would conflate years since degree effects on distance with any cohort specific effects. Instead, we estimate the expected *transition* in distance from k years post-degree to $k+1$ years post-degree. The following equation is used:

$$Transition_{i,t,k} = Distance_{i,t,k} - Distance_{i,t-1,k-1} \tag{3.3}$$

Where i denotes an individual, k denotes the number of years since the individual obtained their degree, and t denotes the survey year. We estimate the expected transition for a member

Fig. 3.1. Average Distance for Each Field of Study: Random Assignment to Occupations



This figure shows the mean distance scores that would occur for graduates of all post-secondary fields of study if graduates were assigned randomly to occupations according to Equation 3.2

of subpopulation j[69] who graduated k years ago as:

$$Transition_{k,j} = \sum_{t=2000}^{2007} \sum_{i \in j} \Big[\frac{Transition_{i,t,k-2}}{N_{t,j,k-2}} + \frac{Transition_{i,t,k-1}}{N_{t,j,k-1}} + \frac{Transition_{i,t,k}}{N_{t,j,k}} + \frac{Transition_{i,t,k+1}}{N_{t,j,k+1}} + \frac{Transition_{i,t,k+2}}{N_{t,j,k+2}}\Big] \quad (3.4)$$

That is the expected transition in distance for an individual who obtained their degree k years ago is a weighted average of the transitions for members of subpopulation j who completed their degree between $k-2$ and $k+2$ years ago. To obtain the expected distance for a member of subpopulation j who graduated m years ago, we add up the expected transitions for each year post-degree up until m to the average distance at the start of one's career.

$$Distance_{m,j} = \sum_{i \in j} \frac{\left(\frac{Distance_{i,t,1}}{N_{t,j,1}} + \frac{Distance_{i,t,2}}{N_{t,j,2}} + \frac{Distance_{i,t,3}}{N_{t,j,3}}\right)}{3} + \sum_{k=1}^{m} Transition_{k,j} \quad (3.5)$$

Where $Distance_{i,t,k}$ refers to the estimated distance for individual i, k years after they have completed their degree.

Figure 3.2 shows both the early career distance scores for graduates of 8 broad fields of study[70], as well as the expected lifecycle path of distance for these graduates[71]. First of all, it is noteworthy that the distance scores that arise in reality are always far lower than those from the random assignment case. This indicates that post-secondary graduates of all fields of study tend to be sorted into occupations that require similar skills as those required of the occupations that their field of study trains for.

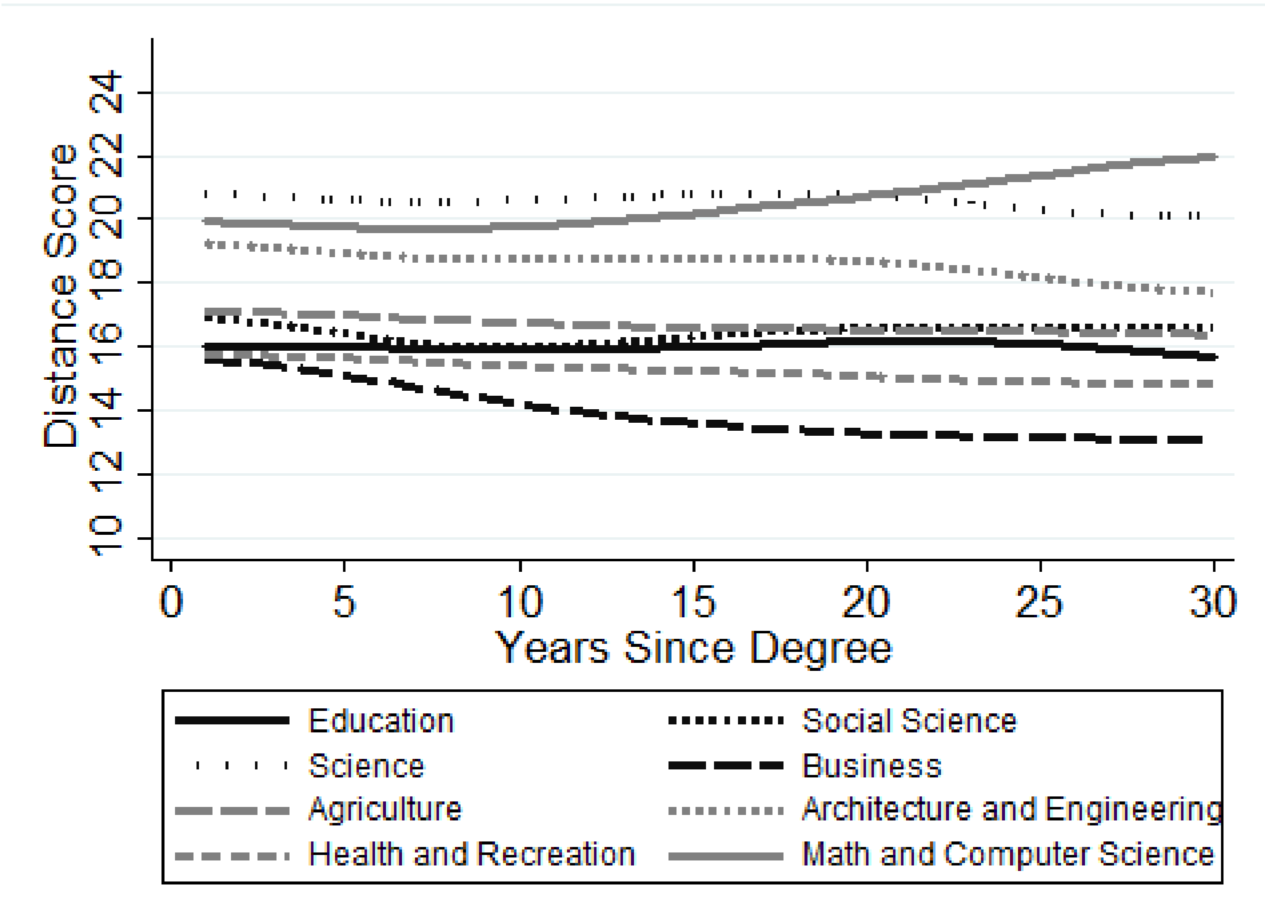
Furthermore, Figure 3.2, show that across most fields of study, there is little change in distance over the course of the lifecycle. Somewhat notable exceptions include Business, as well as Architecture and Engineering. Graduates of these fields tend to move into occupations that are closer matches for their field in the years after they graduate. It also noteworthy that Math and Computer Science graduates tend to move into occupations that are further away from

[69] Subpopulations may include graduates of a given field of study, individuals with a given level of post-secondary education, or individuals of a given gender.

[70] We omit Visual and Performing Arts, and Humanities because of small sample sizes.

[71] Individual figures for each field of study can be foud in the appendix, in Figures A3.1-A3.8

Fig. 3.2. Distance Scores For All Fields of Study Over the Lifecycle



those that we would expect them to be in. Due to the analytical nature of Math and Computer Science occupations, it is probable that this increase in distance arises as a result of movements into managerial roles later in one's career, as described by Bender and Heywood (2011). This would place these graduates further away from the types of skills that are associated with their training.

We also note that at the outset of their careers, graduates of STEM fields tend to be the furthest away from the skills that we would expect them to use on the job. Agriculture and Social Science graduates are noticably better matched, while Education, Health and Recreation, and Business graduates are in occupations that are closer to what we would expect for graduates of these fields. This ordering of early career matches is perhaps surprising, since STEM graduates are often found to be well matched. We contend that this arises as a result of STEM graduates using a far more diverse range of skills in their jobs than graduates of other fields of study. That is, despite oftentimes being very well matched, STEM graduates are also frequently very poorly matched. The perception of close matching for STEM graduates is likely to be driven in part by the fact that they are more likely to be very well matched than graduates of most post-secondary fields.

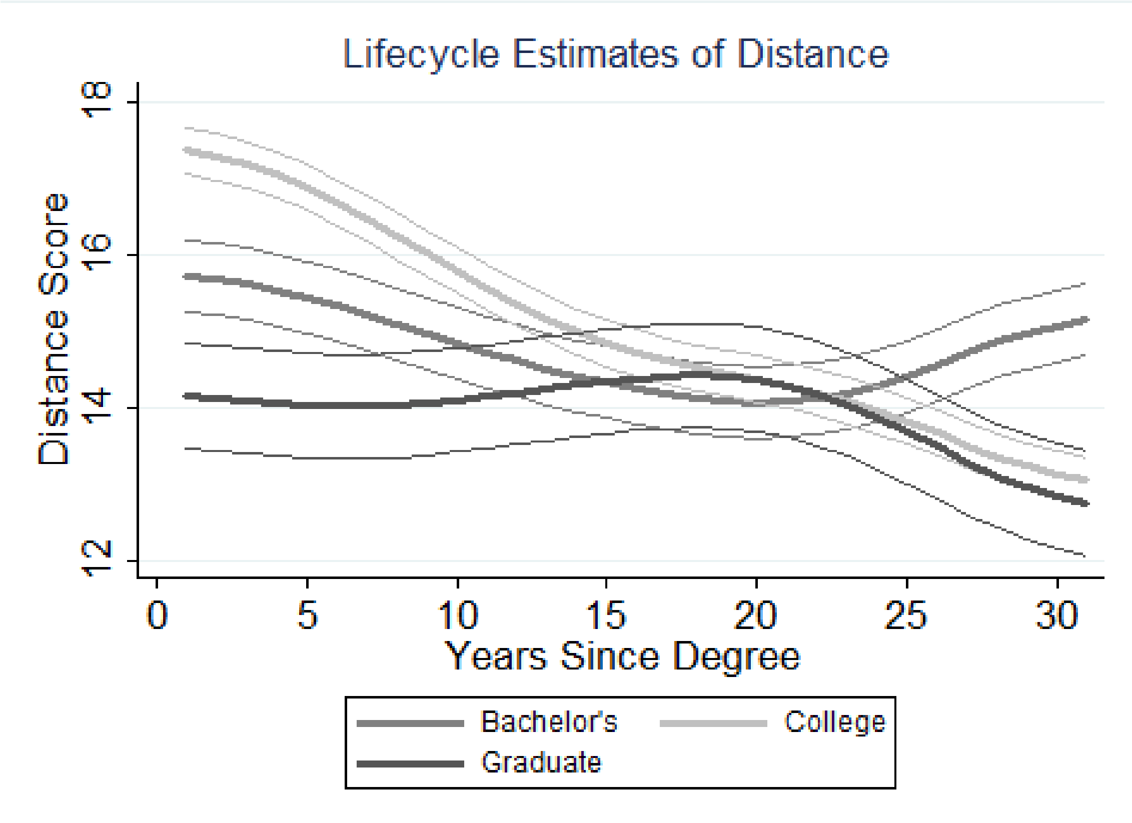
In Table 3.4, we confirm that STEM graduates are often very closely matched by examining the probability that the field of study that is demanded for a given occupation matches an individual's actual field of study. This table is constructed using an alternative measure of the degree of match between a field of study and and occupation. Career Handbook, from Employment and Social Development Canada[72], lists the employment requirements that are associated with each occupation, including any necessary programs of study. A field-occupation match occurs if an individual studied a field that is associated with their occupation. Table 3.4 contains results from a probit model in which this binary measure of field-occupation match is the dependent variable. The independent variables are listed in the table, and include several post-secondary fields of study. This model is estimated for the entire set of post-secondary graduates in Column 1, the first occupation that an individual is working in over the course of the panel in Column 2, college graduates in Column 3, university graduates in Column 4, men in Column 5, and women in Column 6.

[72]For more information on Career Handbook, visit https://noc.esdc.gc.ca/CareerHandbook/ChWelcome/87c0ba81da80431495d3347ded560212

Table. 3.4. SLID: Match probability between occupation-demanded field of study (FS) and workers' actual field of study (CIP)

	(1)	(2) Start occ.	(3) CC	(4) UNI	(5) Men	(6) Women
Education	**0.006**	**0.005**	0.003	0.001	**-0.014**	-0.003
	(0.002)	**(0.002)**	(0.004)	(0.004)	**(0.004)**	(0.003)
Visual and	**-0.006**	-0.005	-0.005	-0.010	-0.007	-0.006
performing arts	**(0.003)**	(0.004)	(0.005)	(0.009)	(0.005)	(0.004)
Humanities	0.004	0.005	**0.008**	-0.006	-0.004	-0.003
	(0.003)	(0.003)	**(0.004)**	(0.006)	(0.005)	(0.004)
Social sciences	**0.012**	**0.013**	**0.009**	**0.012**	0.005	**0.006**
and law	**(0.002)**	**(0.002)**	**(0.003)**	**(0.005)**	(0.004)	**(0.002)**
Business, mangm.,	**0.090**	**0.094**	**0.078**	**0.123**	**0.092**	**0.080**
public admin	**(0.001)**	**(0.002)**	**(0.002)**	**(0.004)**	**(0.002)**	**(0.002)**
Phys and life sci.,	-0.003	-0.003	-0.003	-0.010	**-0.016**	-0.007
technology	(0.004)	(0.005)	(0.007)	(0.009)	**(0.007)**	(0.006)
Math, comp sci.,	**0.011**	**0.011**	**0.013**	**0.008**	**0.011**	0.004
information sys.	**(0.002)**	**(0.002)**	**(0.003)**	**(0.005)**	**(0.003)**	(0.002)
Archit., engineer.,	**0.009**	**0.011**	**0.013**	**0.013**	**0.013**	**0.013**
and related	**(0.002)**	**(0.002)**	**(0.002)**	**(0.005)**	**(0.002)**	**(0.003)**
Agric., natl. res.,	-0.003	-0.002	0.000	-0.008	-0.005	-0.001
conservation	(0.003)	(0.003)	(0.004)	(0.007)	(0.004)	(0.004)
Health, parks,	**0.003**	**0.004**	0.002	0.002	**-0.008**	0.001
rec. and fitness	**(0.002)**	**(0.002)**	(0.002)	(0.004)	**(0.003)**	(0.002)
Personal, protect.	**-0.005**	**-0.005**	-0.001	-0.010	-0.002	-0.002
and transport	**(0.002)**	**(0.002)**	(0.003)	(0.006)	(0.002)	(0.003)
Female	**-0.008**	**-0.009**	**-0.004**	**-0.014**		
	(0.001)	**(0.001)**	**(0.001)**	**(0.002)**		
Bachelor's				**0.032**	**0.016**	
					(0.002)	**(0.001)**
Post-Grad					**0.029**	**0.026**
					(0.002)	**(0.002)**
constant	**0.019**	**0.020**	**0.012**	**0.035**	0.003	0.001
	(0.002)	**(0.002)**	**(0.003)**	**(0.005)**	(0.002)	(0.001)
N. Obs.	71378	57189	29309	18709	35199	35024

Fig. 3.3. Distance Scores Over the Lifecycle By Degree Type

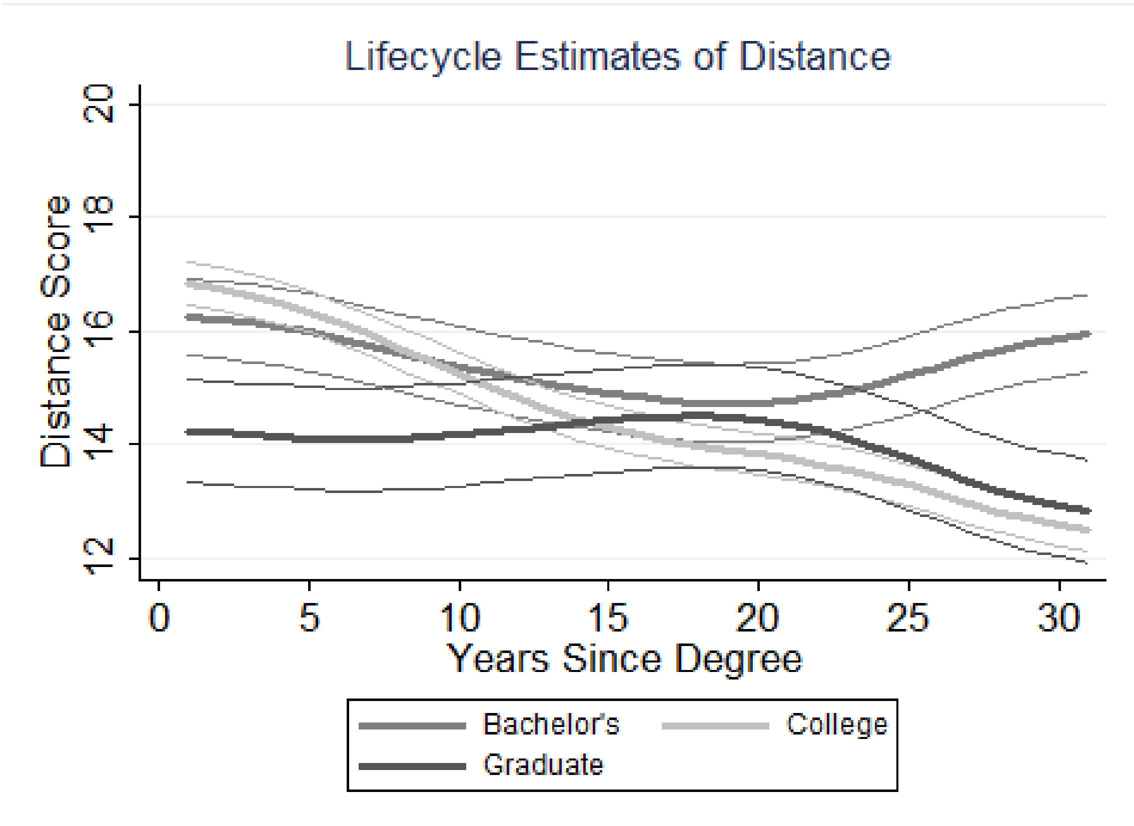


Business graduates are the most closely matched using both this approach and our distance based approach, however Architecture and Engineering and Math and Computer Science graduates are more likely to be in occupations that demand their field than graduates of all non-Business fields. The high distance scores for these fields (shown in Figure 3.2) is not driven by a low probability of a very good match, but rather a high probability of a very poor match. This element of mismatch is not captured using traditional approach of measuring mismatch.

3.3.3 Levels of Education

We explore the differences in our distance measure for individuals with 3 different levels of post-secondary education; college degrees, bachelor's degrees, and graduate degrees. The results in Figure 3.3 show a very clear early career stratification of distance across degree types. Recipients of graduate degrees are employed in occupations that are closer matches to their field of study than bachelor's degree recipients, who are in turn better matched than college degree recipients. Over the course of the lifecycle, this dispersion narrows as college and bachelor's degree recipients tend to move into occupations that are closer matches to their field at a higher rate than those with graduate degrees.

Fig. 3.4. Distance Scores Over the Lifecycle By Degree Type: Re-Weighted to Assign a Common Field of Study Distribution



However, since the field of study distribution is not equivalent across the 3 different levels of post-secondary education, we re-weigh the observations to assign each of them a common field of study distribution.

$$Distance_{m,l} = \sum_{f=1}^{F} \sum_{i \in l,f} \frac{1}{N_{f,l}} \left(\frac{N_f/N}{N_{f,l}/N_l}\right)\left[\left(\frac{Distance_{i,t,1}}{N_{t,1}} + \frac{Distance_{i,t,2}}{N_{t,2}} + \frac{Distance_{i,t,3}}{N_{t,3}}\right) + \sum_{k=1}^{m} Transition_{k,f,l}\right] \tag{3.6}$$

Where f denotes the field of study, and l denotes the level of education. N_f/N is the proportion of individuals in the total population of post-secondary graduates who studied field f, while $N_{f,l}/N_l$ is the equivalent proportion among those individuals with a level of education l. This approach equalizes the proportion of each field of study across levels of education by placing more weight on individuals who studied a field that is less common among graduates with their level of education than among post-secondary graduates as a whole. Figure 3.4 shows the estimated distance scores over the lifecycle after completing this re-weighting. In this case, the initial career distances are similar, although bachelor's degree recipients are no longer as well matched relative to college graduates. Individuals with bachelor's degree also appear to transition into occupations that are further away from their field of study over the course of the lifecycle[73], while college graduates have a strong tendency to move into better matching occupations. By mid-career college graduates and those with graduate degrees are similarly well matched, while individuals with bachelor's degree are the furthest away from the occupations that they are trained for.

3.3.4 Gender

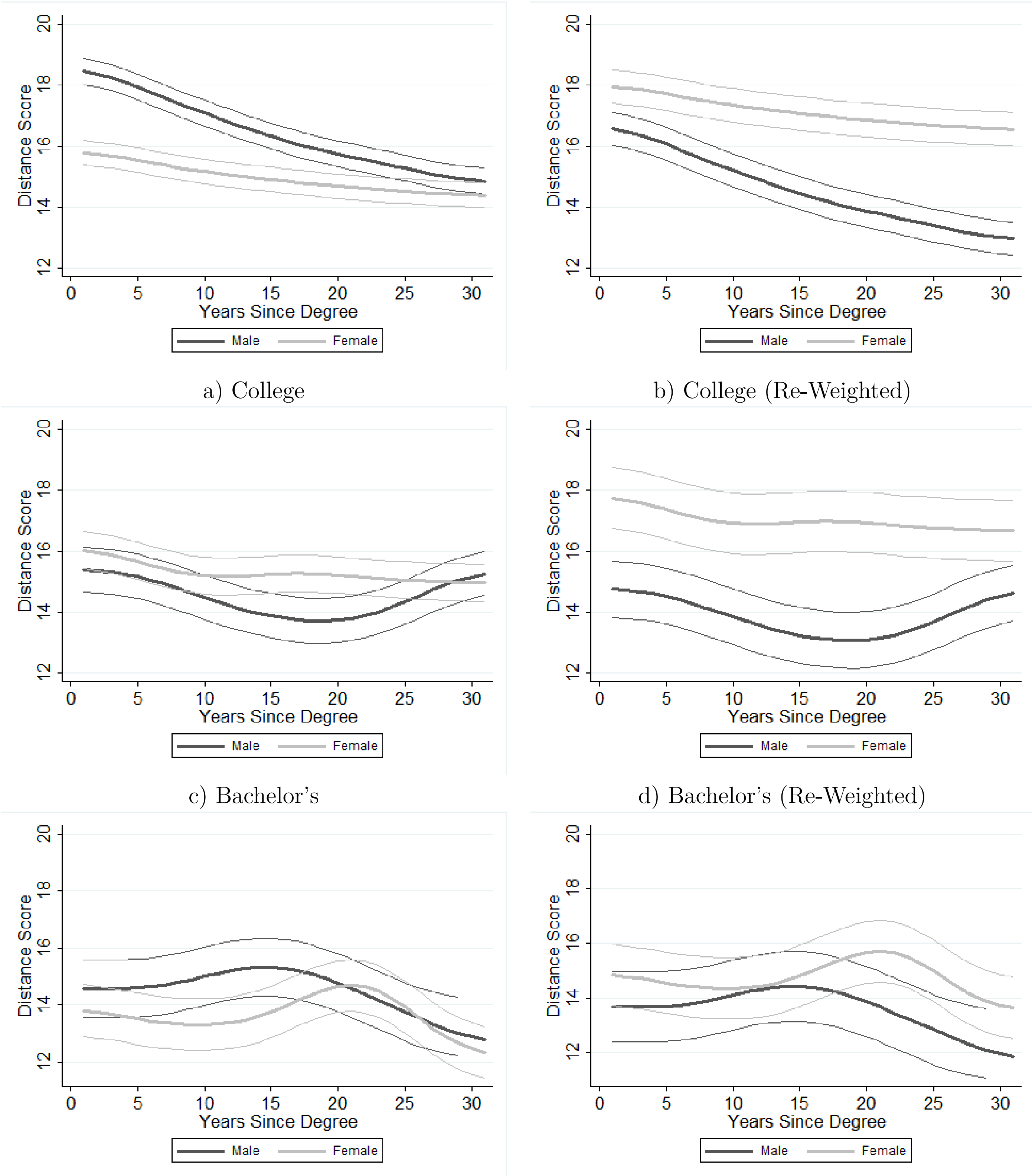
We furthermore examine distance separately for men and women, for each of the 3 levels of education. These results are found in Figure 3.5, with the non-reweighted results being found in the figures shown in the first column, while the rewighted results are found in the figures shown in the second column. It is noteworthy that with the exception of bachelor's degree holders, women are on average in occupations that are more similar to what we would expect based on their field. However, this is largely because women are more likely to have studied a field for which distance tends to be lower, as opposed to being better matched conditional on field of study. With a common field of study distribution, men are on average better matched than women. This is especially true for bachelor's degree holders, and for college degree holders later in their careers.

3.4 Wage Analysis

We expand upon the literature on field of study mismatch and earnings by estimating whether individuals who are employed in occupations that use the skills that are associated with their field of study have higher earnings than their more highly mismatched counterparts. We estimate the fixed effects models in Equations 3.7-3.10, where X_{it} denotes a set of non-constant individuals characteristics. These characteristics are; marital status, labour union status, province of residence, and year controls.

[73] We would like to note that this does not necessarily indicate that individuals with Bachelor's degrees are moving into low skill, low paying occupations.

Fig. 3.5. Distance Scores for Men and Women, By Level of Post-Secondary Education, Reweighted and Non-Reweighted



a) College

b) College (Re-Weighted)

c) Bachelor's

d) Bachelor's (Re-Weighted)

e) Graduate

f) Graduate (Re-Weighted)

$$W_i = \alpha_{1i} + \beta_1 X_{it} + \gamma_{11} Comp1occ_{it} + \gamma_{12} Comp2occ_{it} + \gamma_{13} Comp3occ_{it} + \gamma_{14} Comp4occ_{it} + \gamma_{15} Comp5occ_{it} + \lambda_1 Distance_{it} + u_{it} \quad (3.7)$$

$$W_i = \alpha_{2i} + \beta_2 X_{it} + \gamma_{21} Comp1occ_{it} + \gamma_{22} Comp2occ_{it} + \gamma_{23} Comp3occ_{it} + \gamma_{24} Comp4occ_{it} + \gamma_{25} Comp5occ_{it} + \lambda_{21} Distance_{it} + \lambda_{22} College_i \quad Distance_{it} + \lambda_{23} Bachelor\,s_i \quad Distance_{it} + u_{it} \quad (3.8)$$

$$W_i = \alpha_{3i} + \beta_3 X_{it} + \gamma_{31} Comp1occ_{it} + \gamma_{32} Comp2occ_{it} + \gamma_{33} Comp3occ_{it} + \gamma_{34} Comp4occ_{it} + \gamma_{35} Comp5occ_{it} + \lambda_{31} Distance_{it} + \lambda_{34} YearsSinceDegree_{it} \quad Distance_{it} + u_{it} \quad (3.9)$$

$$W_i = \alpha_{4i} + \beta_4 X_{it} + \gamma_{41} Comp1occ_{it} + \gamma_{42} Comp2occ_{it} + \gamma_{43} Comp3occ_{it} + \gamma_{44} Comp4occ_{it} + \gamma_{45} Comp5occ_{it} + \lambda_{41} Distance_{it} + \lambda_{45} Female_i \quad Distance_{it} + u_{it} \quad (3.10)$$

We control for changes in the skill composition of the occupation in order to ensure that our earnings results are not driven by movements into higher skilled occupations. For example, consider someone who studied a field that trains an individual for an occupation that has high analytical skill requirements. If this individual moves into an occupation that has higher analytical skill requirements than their previous job, their distance score would be reduced, but without controlling for skill levels, we would be unable to tell whether an increase in their pay reflects the higher analytical skill level of the occupation or whether it can be attributed to the match improvement.

The main results of interest from estimating the models in Equations 3.7-3.10 are found in Table 3.5 (full results in Tables A3.2-A3.4). We find that the types of skills that are used on the job are the strongest predictor of earnings, however our distance score also matters.

Table. 3.5. Fixed Effects Results: Effect of Distance on Log Hourly Earnings

Variable	(1)	(2)	(3)	(4)
Comp1occ	0.012***	0.012***	0.012***	0.012***
	(0.002)	(0.002)	(0.002)	(0.002)
Comp2occ	0.011***	0.011***	0.011***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
Comp3occ	-0.002	-0.002	-0.002	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)
Comp4occ	-0.001	-9.819e-04	-9.991e-04	-9.388e-04
	(0.003)	(0.003)	(0.003)	(0.003)
Comp5occ	-0.001	-0.001	-0.001	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)
Distance	-0.002*	-8.553e-04	-0.002*	-9.874e-04
	(0.001)	(0.003)	(0.001) (0.002)	
College*Distance		-0.002		
		(0.003)		
Bachelor's*Distance		-0.001		
		(0.002)		
Years Since Degree*Distance			2.24e-05	
			(4.96e-05)	
Female*Distance				-0.003*
				(0.002)

$N = 24,170, \; \# \; of \; Groups = 6830$

Results in each column in this table consist of the primary results of interest for Equations 3.7-3.10
The full set of results for each column in this Table can be found in Appendix Tables A3.2-A3.5

Conditional on a given change in occupational skills, if this change results in one being further from the skills that are associated with their degree, this carries a small earnings penalty. To place this earnings penalty into perspective, since men with bachelor's degrees tend to be 3 units of distance closer to their field of study than otherwise similar women at the outset of their career (as per panel d) in Figure 3.5), this would correspond to an earnings premium of approximately 0.6% for men purely driven by men having a tendency to be in occupations that use skills that are similar to those that are associated with their degree, independent of what those skills are. Since the average distance scores for all fields of study and demographic groups are below 25, our results suggest that mismatch should at most account for approximately a 5% earnings penalty relative to being perfectly matched.

. Table 3.6 contains the estimated effect of distance on log hourly earnings for the subpopulations of women, graduates of college programs, and graduate of bachelor's programs. The results in this table simply combine relevant coefficients in the bottom 5 rows of Table 3.5 to present the overall effect of distance for these subpopulations of interest.

Table. 3.6. Coefficients for Women, College Graduates, and Bachelor's Graduates

College Graduates ($\lambda_{21} + \lambda_{22}$)	-0.003*
Bachelor's Graduates ($\lambda_{21} + \lambda_{23}$)	-0.002
Graduate Degree Holders (λ_{21})	-8.553e-04
Women ($\lambda_{41} + \lambda_{45}$)	-0.004***

$N = 24,170$, *# of Groups* = 6830

Results in rows 1-3 are obtained from the results in column (2) in Table 3.5

The result in row 4 is obtained from the results in column (4) in Table 3.5

When allowing the effect of distance to vary based on the level of post-secondary education, we find that the effect of being further away from one's degree is insignificant for those with graduate or bachelor's degrees, but is significant for individuals with college degrees (the estimated coefficients for women, bachelor's graduates, and college graduates can be found in Table 3.6). This indicates that, especially early in their careers when college graduate tend to be poorly matched, we should expect their greater degree of mismatch to result in a wage penalty. When splitting the wage effects of distance by gender, we find that the penalty is significant for women, but not for men. Surprisingly, we do not find any evidence that the earnings penalty for mismatch dissipates over the course of one's career.

3.5 Conclusion

Traditional binary measures of occupation-field of study mismatch have become well established. By using a continuous measure of mismatch based on skill distance between fields and occupations, we find evidence that STEM graduates are not as well matched as popular perception or binary measures of mismatch would make it appear.STEM graduates are often employed in occupations that use highly dissimilar skills from those of occupations that are traditionally associated with their field. This raises questions surrounding the benefits of public policies directed at increasing the number of STEM majors graduating from post-secondary institutions. We believe that to be an important avenue for future research.

We also find, consistent with the literature on mismatch and earnings, that individuals who are mismatched receive lower earnings than their well matched counterparts in occupations that use similar skills. However, this earnings penalty is modest, with the skills that are used on the job being a stronger predictor of earnings than the strength of the match. This raises important questions surrounding the importance of post-secondary programs in matching graduates to highly skilled jobs vs. matching graduates to similarly skilled jobs.

References

Teresa Abada, Feng Hou, and Yuqian Lu. Choice or necessity: do immigrants and their children choose self-employment for the same reasons? *Work, employment and society*, 28(1):78–94, 2014.

Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, volume 4, pages 1043–1171. Elsevier, 2011.

Manuel Adelino, Antoinette Schoar, and Felipe Severino. House prices, collateral, and self-employment. *Journal of Financial Economics*, 117(2):288–306, 2015.

Matthew T Allen, Suzanne Tsacoumis, and Rodney A McCloy. Updating occupational ability profiles with o* net® content model descriptors. *Raleigh: National Center for O* NET Development*, 2011.

Joseph G Altonji, Lisa B Kahn, and Jamin D Speer. Trends in earnings differentials across college majors and the changing task composition of jobs. *American Economic Review*, 104 (5):387–93, 2014.

Lina Andersson and Mats Hammarstedt. Intergenerational transmissions in immigrant self-employment: Evidence from three generations. *Small Business Economics*, 34(3):261–276, 2010.

Peter Arcidiacono, V Joseph Hotz, and Songman Kang. Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16, 2012.

Hon Jean Augustine. Employment match rates in the regulated professions: trends and policy implications. *Canadian Public Policy*, 41(Supplement 1):S28–S47, 2015.

David H Autor, Frank Levy, and Richard J Murnane. The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4):1279–1333, 2003.

Abdurrahman Aydemir and Mikal Skuterud. Explaining the deteriorating entry earnings of canada's immigrant cohorts, 1966–2000. *Canadian Journal of Economics/Revue canadienne d'économique*, 38(2):641–672, 2005.

Michael Baker and Dwayne Benjamin. The performance of immigrants in the canadian labor market. *Journal of labor economics*, 12(3):369–405, 1994.

Michael Baker and Dwayne Benjamin. The role of the family in immigrants' labor-market activity: an evaluation of alternative explanations. *The American Economic Review*, pages 705–727, 1997.

Catia Batista and Janis Umblijs. Migration, risk attitudes, and entrepreneurship: evidence from a representative immigrant survey. *IZA Journal of Migration*, 3(1):17, 2014.

Stefan Bauernschuster, Oliver Falck, Stephan Heblich, Jens Suedekum, and Alfred Lameli. Why are educated and risk-loving persons more mobile across regions? *Journal of Economic Behavior & Organization*, 98:56–69, 2014.

Keith A Bender and John S Heywood. Educational mismatch and the careers of scientists. *Education Economics*, 19(3):253–274, 2011.

David G Blanchflower. Self-employment in oecd countries. *Labour economics*, 7(5):471–505, 2000.

Vicki Bogan and William Darity Jr. Culture and entrepreneurship? african american and immigrant self-employment in the united states. *The Journal of Socio-Economics*, 37(5): 1999–2019, 2008.

Alexander Bogin, William Doerner, and William Larson. Local house price dynamics: New indices and stylized facts. *Real Estate Economics*, 47(2):365–398, 2019.

Holger Bonin, Amelie Constant, Konstantinos Tatsiramos, and Klaus F Zimmermann. Native-migrant differences in risk attitudes. *Applied Economics Letters*, 16(15):1581–1586, 2009.

George J Borjas. The self-employment experience of immigrants. Technical report, National Bureau of Economic Research, 1986.

Bernt Bratsberg. Legal versus illegal us immigration and source country characteristics. *Southern Economic Journal*, pages 715–727, 1995.

Bernt Bratsberg, James F Ragan, Jr, and Zafar M Nasir. The effect of naturalization on wage growth: A panel study of young male immigrants. *Journal of Labor Economics*, 20(3): 568–597, 2002.

Xianglei Chen. Stem attrition: College students' paths into and out of stem fields. statistical analysis report. nces 2014-001. *National Center for Education Statistics*, 2013.

Zong Chen and Mikal Skuterud. The relative labour market performance of former international students: Evidence from the canadian national graduates survey. 2017.

Daniel Chiquiar and Gordon H Hanson. International migration, self-selection, and the distribution of wages: Evidence from mexico and the united states. *Journal of political Economy*, 113(2):239–281, 2005.

Barry Chiswick. Are immigrants favorably self-selected? *American Economic Review*, 89(2): 181–185, 1999.

Kenneth Clark and Stephen Drinkwater. Pushed out or pulled in? self-employment among ethnic minorities in england and wales. *Labour Economics*, 7(5):603–628, 2000.

Andrew Clarke and Mikal Skuterud. A comparative analysis of immigrant skills and their utilization in australia, canada, and the usa. *Journal of Population Economics*, 29(3):849–882, 2016.

Robert M DeKeyser and Karl J Sokalski. The differential role of comprehension and production practice. *Language Learning*, 46(4):613–642, 1996.

Richard Disney, John Gathergood, and Andrew Henley. House price shocks, negative equity, and household consumption in the united kingdom. *Journal of the European Economic Association*, 8(6):1179–1207, 2010.

Harriet Orcutt Duleep and Seth Sanders. The decision to work by married immigrant women. *ILR Review*, 46(4):677–690, 1993.

David S Evans and Boyan Jovanovic. An estimated model of entrepreneurial choice under liquidity constraints. *Journal of political economy*, 97(4):808–827, 1989.

Robert W Fairlie and Magnus Lofstrom. Immigration and entrepreneurship. In *Handbook of the economics of international migration*, volume 1, pages 877–911. Elsevier, 2015.

Robert W Fairlie and Bruce D Meyer. Ethnic and racial self-employment differences and possible explanations. *Journal of human resources*, pages 757–793, 1996.

Robert W Fairlie and Bruce D Meyer. Trends in self-employment among white and black men during the twentieth century. *Journal of human resources*, pages 643–669, 2000.

Ana Ferrer, David A Green, and W Craig Riddell. The effect of literacy on immigrant earnings. *Journal of Human Resources*, 41(2):380–410, 2006.

Monica Fisher and Paul A Lewin. Push and pull factors and hispanic self-employment in the usa. *Small Business Economics*, 51(4):1055–1070, 2018.

Nicole Fortin, Thomas Lemieux, and Javier Torres. Foreign human capital and the earnings gap between immigrants and canadian-born workers. *Labour Economics*, 41:104–119, 2016.

Carl Benedikt Frey and Michael A Osborne. The future of employment: How susceptible are jobs to computerisation? *Technological forecasting and social change*, 114:254–280, 2017.

David A Green and Christopher Worswick. Canadian economics research on immigration through the lens of theories of justice. *Canadian Journal of Economics/Revue canadienne d'économique*, 50(5):1262–1303, 2017.

David Alan Green, Huju Liu, Yuri Ostrovski, and W Garnett Picot. *Immigration, business ownership and employment in Canada*. Statistics Canada, Analytical Studies Branch Ottawa, 2016.

Barton H Hamilton, Nicholas W Papageorge, and Nidhi Pande. The right stuff? personality and entrepreneurship. *Quantitative Economics*, 10(2):643–691, 2019.

Mats Hammarstedt and Ghazi Shukur. Testing the home-country self-employment hypothesis on immigrants in sweden. *Applied Economics Letters*, 16(7):745–748, 2009.

Feng Hou and Yuqian Lu. International students, immigration and earnings growth: the effect of a pre-immigration host-country university education. *IZA Journal of Development and Migration*, 7(1):5, 2017.

Michael Hout and Harvey Rosen. Self-employment, family background, and race. *Journal of Human Resources*, pages 670–692, 2000.

John Eric Humphries. The causes and consequences of self-employment over the life cycle. Technical report, Technical Report, Mimeo, University of Chicago, 2016.

Jennifer Hunt. Which immigrants are most innovative and entrepreneurial? distinctions by entry visa. *Journal of Labor Economics*, 29(3):417–457, 2011.

Susumu Imai, Derek Stacey, and Casey Warman. From engineer to taxi driver? language proficiency and the occupational skills of immigrants. *Canadian Journal of Economics/Revue canadienne d'économique*, 52(3):914–953, 2019.

Refugees Immigration and Citizenship Canada. 2017 annual report to parliament on immigration. 2017.

Refugees Immigration and Citizenship Canada. Notice–supplementary information 2019-2021 immigration levels plan. 2018.

David A Jaeger, Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin. Direct evidence on risk attitudes and migration. *The Review of Economics and Statistics*, 92 (3):684–689, 2010.

Sari Pekkala Kerr and William R Kerr. Immigrant entrepreneurship. Technical report, National Bureau of Economic Research, 2016.

William R Kerr and Martin Mandorff. Social networks, ethnicity, and entrepreneurship. Technical report, National Bureau of Economic Research, 2015.

Morris M Kleiner and Alan B Krueger. The prevalence and effects of occupational licensing. *British Journal of Industrial Relations*, 48(4):676–687, 2010.

Thomas Lemieux. Occupations, fields of study and returns to education. *Canadian Journal of Economics/Revue canadienne d'économique*, 47(4):1047–1077, 2014.

Peter S Li. Immigrants' propensity to self-employment: evidence from canada. *International Migration Review*, 35(4):1106–1128, 2001.

Qing Li and Arthur Sweetman. The quality of immigrant source country educational outcomes: Do they matter in the receiving country? *Labour Economics*, 26:81–93, 2014.

Robert E Lucas Jr. On the size distribution of business firms. *The Bell Journal of Economics*, pages 508–523, 1978.

Seamus McGuinness, Konstantinos Pouliakas, and Paul Redmond. Skills mismatch: Concepts, measurement and policy approaches. *Journal of Economic Surveys*, 32(4):985–1015, 2018.

Guillermo Montt. Field-of-study mismatch and overqualification: labour market correlates and their wage penalty. *IZA Journal of Labor Economics*, 6(1):2, 2017.

Martin Nordin, Inga Persson, and Dan-Olof Rooth. Education–occupation mismatch: Is there an income penalty? *Economics of education review*, 29(6):1047–1059, 2010.

George Orlov. The impact of language training on the transfer of pre-immigration skills and the wages of immigrants. *Available at SSRN 3412195*, 2018.

Simon C Parker. *The economics of entrepreneurship*. Cambridge University Press, 2018.

Krishna Patel and Francis Vella. Immigrant networks and their implications for occupational choice and wages. *Review of Economics and Statistics*, 95(4):1249–1277, 2013.

Miana Plesca et al. The impact of introducing mandatory occupational licensing. *Modern Economy*, 6(12):1309, 2015.

Kevin Rask. Attrition in stem fields at a liberal arts college: The importance of grades and pre-collegiate preferences. *Economics of Education Review*, 29(6):892–900, 2010.

Mauricio Reis. Measuring the mismatch between field of study and occupation using a task-based approach. *Journal for labour market research*, 52(1):9, 2018.

John Robst. Education and job match: The relatedness of college major and work. *Economics of Education Review*, 26(4):397–407, 2007a.

John Robst. Education, college major, and job match: Gender differences in reasons for mismatch. *Education Economics*, 15(2):159–175, 2007b.

Martin C Schmalz, David A Sraer, and David Thesmar. Housing collateral and entrepreneurship. *The Journal of Finance*, 72(1):99–132, 2017.

Sana Sellami, Dieter Verhaest, and Walter Van Trier. How to measure field-of-study mismatch? a comparative analysis of the different methods. *LABOUR*, 32(4):141–173, 2018.

Peter J Sloane. Much ado about nothing? what does the overeducation literature really tell us. *Overeducation in Europe*, pages 11–45, 2003.

Merrill Swain. Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in second language acquisition*, 15:165–179, 1985.

Arthur Sweetman and Casey Warman. Canada's immigration selection system and labour market outcomes. *Canadian Public Policy*, 39(Supplement 1):S141–S164, 2013.

Arthur Sweetman and Casey Warman. Former temporary foreign workers and international students as sources of permanent immigration. *Canadian Public Policy*, 40(4):392–407, 2014.

Arthur Sweetman, James Ted McDonald, and Lesleyanne Hawthorne. Occupational regulation and foreign qualification recognition: an overview. *Canadian Public Policy*, 41(Supplement 1):S1–S13, 2015.

Peter Vandor and Nikolaus Franke. See paris and… found a business? the impact of cross-cultural experience on opportunity recognition capabilities. *Journal of Business Venturing*, 31(4):388–407, 2016.

Dieter Verhaest and Eddy Omey. The determinants of overeducation: different measures, different outcomes? *International Journal of Manpower*, 2010.

Evgueni Vinogradov and Eva Jenny Benedikte Jørgensen. Differences in international opportunity identification between native and immigrant entrepreneurs. *Journal of International Entrepreneurship*, 15(2):207–228, 2017.

Thierry Volery. Ethnic entrepreneurship: a theoretical framework. *Handbook of research on ethnic minority entrepreneurship*, 1:30–41, 2007.

Qiufang Wen. The production-oriented approach to teaching university students english in china. *Language Teaching*, 51(4):526–540, 2018.

Matthew Wiswall and Basit Zafar. Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824, 2015.

Jennifer Yuen et al. *Job-education match and mismatch: Wage di erentials*. Statistics Canada, 2010.

Andrew M Yuengert. Testing hypotheses of immigrant self-employment. *Journal of human resources*, pages 194–204, 1995.

Basit Zafar. College major choice and the gender gap. *Journal of Human Resources*, 48(3): 545–595, 2013.

Danielle Zietsma et al. Immigrants working in regulated occupations. *Population*, 1(1,207.2): 208–7, 2010.

Julie M Zissimopoulos and Lynn A Karoly. Transitions to self-employment at older ages: The role of wealth, health, health insurance and other factors. *Labour Economics*, 14(2):269–295, 2007.

Chapter 1 Appendix

Table A1.1. Probit Model Coefficients

	Survey weights		P-score weights	
	Coef.	Std.err.	Coef.	Std.Err.
Female	0.005	0.011	-0.004	0.014
Part-Time	0.12***	0.014	-0.009	0.017
Citizen	-0.354***	0.014	0.027	0.017
Kid 0-5	-0.171***	0.012	-0.018	0.015
Kid 6-14	0.116***	0.012	0.043***	0.015
Female*Kid 0-5	-0.016	0.017	0.006	0.021
Female*Kid 6-14	-0.048***	0.017	-0.054**	0.022
University Quality	-0.607***	0.013	0.041***	0.01
White	0.283***	0.025	0.06**	0.025
Black	-1.31***	0.031	-0.022	0.032
Latin American	0.119***	0.04	0.113**	0.045
Chinese	-1.12***	0.036	-0.036	0.036
Japanese	-0.62***	0.075	-0.213***	0.067
Korean	-0.688***	0.049	0.043	0.047
South Asian	-0.777***	0.019	0.114***	0.024
South-East Asian	-1.267***	0.032	-0.065**	0.032
West Asian	-0.048*	0.028	0.003	0.032
Regulated Occupation	-0.079***	0.009	-0.02*	0.012
Married	0.447***	0.01	0.011	0.012
Anglophone	0.026	0.017	-0.105***	0.02
Francophone	-0.444***	0.025	-0.017	0.027
Domestic Experience	0.055***	0.006	0.013***	0.005
$DomesticExperience^2$	0.001***	2.749e-04	-5.772e-04**	2.756e-04
$DomesticExperience^3$	-2.69e-05***	4.8e-06	1.83e-05***	4.17e-06
Western Europe	0.299***	0.029	-0.183***	0.032

Eastern Europe	0.036	0.031	0.015	0.036
Latin America	0.221***	0.036	-0.039	0.039
Subsaharan Africa	0.998***	0.044	-0.016	0.051
Middle East	-0.021	0.034	0.01	0.037
South Asia	1.042***	0.037	0.042	
East Asia	0.777***	0.045	-0.016	0.052
Oceania	0.485***	0.062	-0.211***	0.071
Quebec City	0.06	0.054	0.092	0.057
Montreal	0.205***	0.036	0.044	0.038
Toronto	0.109***	0.019	0.03	0.023
Ottawa	-0.233***	0.024	-0.011	0.028
Hamilton	0.18***	0.04	0.03	0.047
Kitchener	0.172***	0.039	0.037	0.049
Winnipeg	0.065	0.086	0.081	0.097
Calgary	-0.072*	0.043	-0.003	0.053
Edmonton	-0.131***	0.045	-0.034	0.055
Vancouver	0.194***	0.032	-0.057	0.038
Newfoundland	-0.259**	0.114	0.147	0.111
PEI	-0.305**	0.154	0.193	0.195
Nova Scotia	-0.334***	0.052	-0.061	0.054
New Brunswick	-0.358***	0.069	0.102	0.086
Quebec	-0.29***	0.036	-0.036	0.037
Manitoba	-0.177**	0.085	-0.009	0.093
Saskatchewan	0.072*	0.041	-0.002	0.054
Alberta	0.152***	0.044	0.024	0.053
BC	0.048	0.035	0.044	0.041
2011	-0.454***	0.021	-0.071***	0.019
2016	-0.772***	0.037	-0.13***	0.029
Late Fifties	-4.771***	0.344	-1.682***	0.222
Early Sixties	-4.475***	0.245	-1.211***	0.176

Late Sixties	-4.265***	0.184	-1.019***	0.124
Early Seventies	-3.843***	0.151	-0.838***	0.107
Late Seventies	-3.464***	0.127	-0.684***	0.094
Early Eighties	-3.094***	0.104	-0.588***	0.081
Late Eighties	-2.522***	0.081	-0.413***	0.065
Early Nineties	-1.935***	0.065	-0.273***	0.053
Late Nineties	-1.12***	0.05	-0.289***	0.041
Early Two Thousands	-0.439***	0.034	-0.248***	0.029
Late Two Thousands	-0.228***	0.019	-0.12***	0.02
Constant	1.389***	0.05	0.236***	0.053
Weighted N	946,955		946,955	
Pseudo R^2	0.2712		0.0075	

Table. A1.2. Returns to Skill for Immigrants Educated Abroad or in Canada; Not Weighted by Propensity Score Weights

	(1) Basic	(2) With Skill	(3) With Interact.	(4) Includes Occup.
Abroad	-0.107***	-0.114***	0.09	0.033
	(0.007)	(0.007)	(0.082)	(0.079)
Expression		-0.006	0.052***	0.046***
		(0.005)	(0.011)	(0.010)
Comprehension		-0.008*	-0.042***	-0.033***
		(0.005)	(0.010)	(0.010)
Logical		0.02***	0.045***	0.029***
		(0.003)	(0.007)	(0.007)
Hard Science		0.006***	-5.83E-04	0.005
		(0.002)	(0.004)	(0.004)
Executive		-0.01**	-0.034***	-0.03***
		(0.004)	(0.009)	(0.009)
Technical		0.043***	0.057***	0.033***
		(0.004)	(0.007)	(0.007)
Physical		-0.035***	-0.05***	-0.034***
		(0.003)	(0.008)	(0.008)
Abroad*Expression			-0.077***	-0.063***
			(0.012)	(0.012)
Abroad*Comprehension			0.04***	0.04***
			(0.011)	(0.011)
Abroad*Logical			-0.033***	-0.025***
			(0.008)	(0.008)
Abroad*Hard Science			0.009*	0.003
			(0.005)	(0.005)
Abroad*Executive			0.032***	0.031***
			(0.010)	(0.010)
Abroad*Technical			-0.019**	-0.01
			(0.008)	(0.008)
Abroad*Physical			0.019**	0.022***
			(0.009)	(0.009)
Weighted N	946,955	946,955	946,955	946,955
R^2	0.2310	0.2377	0.2380	0.2958

Same model as in Table 1.6 except the estimation is using survey weights as opposed to propensity score weights. Standard errors in parentheses. Actual sample size is about 20% of weighted N sample size, see footnote 6. Asterisks denote significance at the 1%, 5% or 10% level.

Table. A1.3. Sensitivity Analysis to Different Cut-off Probabilities for Having Studied Abroad

	Less than 0.5% or more than 0.995%		Less than 1% or more than 99%		Less than 5% or more than 95%	
	Coef.	Std.err.	Coef.	Std.Err.	Coef.	Std.Err.
Abroad	0.203	0.171	0.372**	0.146	0.228**	0.091
Expression	0.032*	0.017	0.032**	0.015	0.025**	0.012
Comprehension	-0.033*	0.017	-0.026*	0.014	-0.025**	0.011
Logical	0.03***	0.011	0.036***	0.01	0.039***	0.008
Hard Science	0.011	0.007	0.003	0.006	-0.004	0.005
Executive	-0.016	0.015	-0.013	0.011	-0.01	0.01
Technical	0.036***	0.013	0.047***	0.012	0.047***	0.008
Physical	-0.049***	0.012	-0.056***	0.01	-0.036***	0.007
Abroad*Expression	-0.048**	0.023	-0.043**	0.018	-0.06***	0.015
Abroad*Comprehension	-0.005	0.027	-0.029	0.025	0.022*	0.013
Abroad*Logical	-0.006	0.015	-0.013	0.012	-0.026***	0.009
Abroad*Hard Science	0.001	0.01	0.011	0.008	0.012**	0.005
Abroad*Executive	0.018	0.019	0.018	0.014	0.011	0.012
Abroad*Technical	-0.006	0.017	-0.022	0.016	-0.007	0.009
Abroad*Physical	0.015	0.015	0.023*	0.012	-0.002	0.008
Weighted N	1,027,585		990,820		812,280	
R^2	0.2333		0.2333		0.2333	

Sensitivity relative to the main results in Table 1.6 which use a cut-off of 2%. All of the above results are estimated using the model from specification (3) in Table 1.6. Actual sample size is about 20% of weighted sample size, see footnote 6. Asterisks denote significance at the 1%, 5% or 10% level.

Table A1.4. Means of Covariates For The Educated in Canada and Educated Abroad Groups with Survey Weights (Raw Data) or Propensity Score Weights (Balanced Data)

	Survey weights (Trimmed,Not balanced)		P-score weights (Trimmed,Balanced)	
	Canada	Abroad	Canada	Abroad
Female	0.485	0.497	0.503	0.493
Part-Time	0.117	0.133	0.134	0.13
Citizen	0.813	0.642	0.712	0.71
Kid 0-5	0.3	0.276	0.287	0.289
Kid 6-14	0.259	0.325	0.287	0.295
Female*Kid 0-5	0.134	0.124	0.131	0.128
Female*Kid 6-14	0.127	0.16	0.148	0.146
White	0.248	0.236	0.323	0.313
Black	0.121	0.027	0.055	0.05
Latin American	0.011	0.017	0.014	0.016
Chinese	0.293	0.143	0.192	0.177
Japanese	0.006	0.007	0.011	0.007
Korean	0.017	0.022	0.021	0.022
South Asian	0.128	0.211	0.158	0.201
South-East Asian	0.039	0.014	0.02	0.018
West Asian	0.03	0.027	0.026	0.027
Regulated Occupation	0.253	0.201	0.231	0.219
Married	0.677	0.825	0.768	0.78
Anglophone	0.194	0.127	0.21	0.184

Francophone	0.063	0.018	0.038	0.034
Western Europe	0.091	0.056	0.108	0.08
Eastern Europe	0.071	0.101	0.089	0.096
Latin America	0.118	0.073	0.09	0.085
Subsaharan Africa	0.09	0.025	0.081	0.09
Middle East	0.118	0.101	0.098	0.106
South Asia	0.179	0.443	0.271	0.303
East Asia	0.29	0.164	0.212	0.195
Oceania	0.0054	0.0065	0.0101	0.007
Quebec City	0.0078	0.0048	0.0059	0.0062
Montreal	0.156	0.107	0.127	0.126
Toronto	0.392	0.416	0.389	0.403
Ottawa	0.064	0.029	0.041	0.039
Hamilton	0.013	0.015	0.015	0.015
Kitchener	0.011	0.013	0.014	0.013
Winnipeg	0.02	0.03	0.021	0.022
Calgary	0.056	0.071	0.064	0.064
Edmonton	0.037	0.043	0.038	0.038
Vancouver	0.122	0.154	0.154	0.144
Newfoundland	0.0016	0.001	0.0011	0.0012
PEI	0.0012	0.0006	0.0008	0.001
Nova Scotia	0.0093	0.0053	0.0075	0.0065
New Brunswick	0.0049	0.0024	0.0033	0.0038
Quebec	0.184	0.12	0.146	0.145

Manitoba	0.022	0.033	0.024	0.026
Saskatchewan	0.01	0.015	0.012	0.012
Alberta	0.102	0.13	0.116	0.114
BC	0.147	0.177	0.184	0.176
2011	0.369	0.351	0.355	0.3
2016	0.352	0.396	0.407	0.43
Late Fifties	0.0019	0.0003	0.0008	0.0009
Early Sixties	0.0037	0.0007	0.0018	0.0015
Late Sixties	0.024	0.005	0.011	0.01
Early Seventies	0.058	0.016	0.028	0.025
Late Seventies	0.06	0.019	0.032	0.028
Early Eighties	0.074	0.022	0.038	0.032
Late Eighties	0.104	0.042	0.062	0.059
Early Nineties	0.155	0.092	0.108	0.12
Late Nineties	0.141	0.151	0.161	0.157
Early Two Thousands	0.178	0.307	0.287	0.268
Late Two Thousands	0.144	0.242	0.196	0.205
Weighted N	936,835	936,835	936,835	936,835

Weighted N=936,835

The columns under Survey Weights contain the mean of each variable as calculated using thhe survey weights. Under P-Score Weights, these means are re-calculated using the propensity score weights that are described in Section 1.4.1.

Table A1.5. Earnings Gap Between Immigrants Who Studied in Canada and Immigrants Who Studied Abroad with the Inclusion of Skills (Full Results)

X	Coef.	(Std.er.)	X	Coef.	(Std.er.)
Female	-0.099***	(0.009)	Calgary	-0.007	(0.037)
Part Time	-0.603***	(0.015)	Edmonton	-0.082**	(0.04)
Citizen	0.084***	(0.011)	Vancouver	0.018	(0.028)
Kid Age 0-5	0.024**	(0.01)	Newfoundland	0.04	(0.061)
Kid Age 6-14	0.049***	(0.01)	Prince Edward Island	-0.412***	(0.1)
Female*Kid Age 0-5	-0.084***	(0.015)	Nova Scotia	-0.124***	(0.034)
Female*Kid Age 6-14	-0.041***	(0.015)	New Brunswick	-0.224***	(0.045)
University Quality	0.092***	(0.007)	Quebec	-0.168***	(0.024)
White	0.025	(0.016)	Manitoba	0.01	(0.049)
Black	-0.088***	(0.018)	Saskatchewan	0.008	(0.037)
Latin American	-0.019	(0.03)	Alberta	0.184***	(0.038)
Chinese	0.077***	(0.018)	British Columbia	-0.049*	(0.03)
Japanese	0.315***	(0.041)	2011 NHS	0.283***	(0.013)
Korean	0.047*	(0.027)	2016 Census	0.435***	(0.02)
South Asian	-0.031**	(0.014)	1955-1959	0.1	(0.189)
SouthEast Asian	-7.238e-04	(0.019)	1960-1964	0.204	(0.141)
West Asian	-0.118***	(0.019)	1965-1969	0.319***	(0.088)
Regulated Occupation	0.14***	(0.008)	1970-1974	0.352***	(0.074)
Married	0.064***	(0.008)	1975-1979	0.312***	(0.066)
Anglophone	0.125***	(0.013)	1980-1984	0.243***	(0.056)
Francophone	0.071***	(0.017)	1985-1989	0.158***	(0.044)
Domestic Experience	0.07***	(0.004)	1990-1994	0.071**	(0.036)
$DomesticExperience^2$	-0.002***	(0.0002)	1995-1999	0.046*	(0.027)
$DomesticExperience^3$	2.42e-05***	(0.000000325)	2000-2004	0.019	(0.02)
Foreign Experience	0.019***	(0.002)	2005-2009	-0.031**	(0.013)
$ForeignExperience^2$	-0.001***	(0.0002)	Abroad	0.241**	(0.103)
$ForeignExperience^3$	2.14e-05***	(0.00000327)	Expression	0.034**	(0.013)
Western Europe	0.111***	(0.022)	Comprehension	-0.03**	(0.013)
Eastern Europe	-0.102***	(0.024)	Logical	0.039***	(0.009)

Latin America	-0.095***	(0.025)	Hard Science	-0.001	(0.005)
SubSaharan Africa	-0.12***	(0.03)	Executive	-0.014	(0.011)
Middle East	-0.125***	(0.024)	Technical	0.047***	(0.008)
South Asia	-0.159***	(0.027)	Physical	-0.048***	(0.008)
East Asia	-0.266***	(0.03)	Abroad*Expression	-0.059***	(0.016)
Oceania	-0.004	(0.055)	Abroad*Comprehension	0.02	(0.014)
Quebec City	0.031	(0.035)	Abroad*Logical	-0.029***	(0.011)
Montreal	0.026	(0.023)	Abroad*Hard Science	0.01*	(0.006)
Toronto	0.053***	(0.017)	Abroad*Executive	0.012	(0.013)
Ottawa	0.1***	(0.019)	Abroad*Technical	-0.007	(0.01)
Hamilton	0.091***	(0.032)	Abroad*Physical	0.01	(0.01)
Kitchener/Waterloo	0.05*	(0.03)	Constant	6.028***	(0.095)
Winnipeg	-0.129**	(0.05)			
Weighted N	936,835		R^2	0.23	

Table. A1.6. English/French Speaking Countries

Antigua and Barbuda	Haiti	Senegal
Australia	India	Seychelles
Barbados	Ireland	Sierre Leone
Belize	Jamaica	Sri Lanka
Benin	Kenya	Swaziland
Botswana	Kiribati	The Bahamas
Brunei	Laos	The Gambia
Burkina Faso	Madagascar	Togo
Central African Republic	Malawi	Tonga
Dominica	Mali	Trinidad and Tobago
Fiji	Mauritius	Tuvalu
France	Nauru	Uganda
Gabon	New Zealand	United Kingdom
Ghana	Niger	United States of America
Grenada	Nigeria	Vanuatu
Guinea	Saint Kitts and Nevis	Zambia
Guyana	Saint Lucia	Zimbabwe
	Saint Vincent and the Grenadines	

Alongside the United Kingdom and France, this list consists of their former colonies in which English or French are likely to be a primary language of instruction at post-secondary institutions today. We have excluded a number of former colonies in which English or French are less commonly used today. The excluded countries are:Afghanistan, Algeria, Bahrain, Cambodia, Chad, Comoros, Cyprus, Djibouti, Jordan, Iraq, Israel, Lebanon, Libya, Lesotho, Kuwait, Malta, Maldives, Malaysia, Mauritania, Mynamar, Pakistan, Qatar, Solomon Islands, Somalia, Sudan, Tanzania, Tunisia, United Arab Emirates, and Yemen.

Table A1.7. Interactions of Skills with University Quality, Domestic Experience, Age at Immigration, and English/French Speaking Country

	(1) English/French Language	(2) University Quality	(3) Domestic Experience.	(4) Abroad Domestic	(5) Age at Immigration
Abroad	0.263**	0.232**	0.236**	0.236**	0.355***
	(0.125)	(0.1)	(0.104)	(0.104)	(0.116)
Expression	0.03*	0.027*	0.029	0.028	0.113***
	(0.015)	(0.15)	(0.19)	(0.23)	(0.37)
Comprehension	-0.027*	-0.019	-0.035**	-0.025	-0.112***
	(0.016)	(0.14)	(0.17)	(0.21)	(0.31)
Logical	0.037***	0.039***	0.022*	0.022	0.08***
	(0.01)	(0.01)	(0.013)	(0.015)	(0.023)
Hard Science	-8.063e-04	-0.002	0.008	0.007	0.013
	(0.006)	(0.006)	(0.007)	(0.009)	(0.013)
Executive	-0.015	-0.01	-0.007	-0.015	-0.084***
	(0.013)	(0.013)	(0.015)	(0.019)	(0.029)
Technical	0.049***	0.053***	0.05***	0.046***	0.087***
	(0.01)	(0.009)	(0.012)	(0.014)	(0.021)
Physical	-0.046***	-0.048***	-0.054***	-0.056***	-0.086***
	(0.01)	(0.01)	(0.012)	(0.015)	(0.021)
A*Expression	-0.068***	-0.058***	-0.06***	-0.058**	-0.047**
	(0.019)	(0.015)	(0.016)	(0.027)	(0.019)
A*Comprehension	0.022	0.019	0.021	0.005	0.003
	(0.018)	(0.014)	(0.014)	(0.022)	(0.018)
A*Logical	-0.035***	-0.028***	-0.03***	-0.032*	-0.032**
	(0.013)	(0.01)	(0.011)	(0.018)	(0.014)
A*Hard Science	0.008	0.01*	0.011*	0.011	0.022***
	(0.007)	(0.006)	(0.006)	(0.01)	(0.008)
A*Executive	0.022	0.011	0.013	0.026	-0.001
	(0.016)	(0.013)	(0.013)	(0.022)	(0.016)

A*Technical	-0.016	-0.007	-0.008	0.002	-0.011
	(0.012)	(0.009)	(0.01)	(0.016)	(0.012)
A*Physical	0.013	0.01	0.01	0.012	-0.006
	(0.012)	(0.01)	(0.01)	(0.017)	(0.012)
Eng/Fr	-0.043				
	(0.189)				
A*Eng/Fr	-0.037				
	(0.216)				
Eng/Fr*Expression	0.016				
	(0.029)				
Eng/Fr*Comprehension	-0.01				
	(0.026)				
Eng/Fr*Logical	0.006				
	(0.019)				
Eng/Fr*Hard Science	-1.457e-04				
	(0.011)				
Eng/Fr*Executive	0.001				
	(0.023)				
Eng/Fr*Technical	-0.01				
	(0.018)				
Eng/Fr*Physical	-0.003				
	(0.017)				
A*Eng/Fr*Expression	0.043				
	(0.034)				
A*Eng/Fr*Comprehension	-0.02				
	(0.029)				
A*Eng/Fr*Logical	0.022				
	(0.022)				
A*Eng/Fr*Hard Science	0.013				
	(0.013)				
A*Eng/Fr*Executive	-0.046*				
	(0.027)				

A*Eng/Fr*Technical	0.038*			
	(0.021)			
A*Eng/Fr*Physical	-0.021			
	(0.02)			
Quality*Expression		0.013		
		(0.013)		
Quality*Comprehension		-0.019*		
		(0.01)		
Quality*Logical		-4.94e-06		
		(0.009)		
Quality*Hard Science		0.001		
		(0.005)		
Quality*Executive		-0.008		
		(0.011)		
Quality*Technical		-0.011		
		(0.008)		
Quality*Physical		2.668e-04		
		(0.009)		
Domestic*Expression			3.914e-04	4.531e-04
			(7.92E-04)	(0.001)
Domestic*Comprehension			3.649e-04	-3.207e-04
			(6.38E-04)	(9.60E-04)
Domestic*Logical			0.001***	0.001*
			(5.20E-04)	(7.31E-04)
Domestic*Hard Science			-7.73e-04**	-6.878e-04*
			(2.99E-04)	(4.10E-04)
Domestic*Executive			-5.651e-04	3.46e-05
			(6.27E-04)	(9.69E-04)
Domestic*Technical			-2.448e-04	4.26e-05
			(4.92E-04)	(6.66E-04)
Domestic*Physical			5.285e-04	6.551e-04
			(4.96E-04)	(7.30E-04)

A*Domestic*Expression				-8.85e-05	
				(0.002)	
A*Domestic*Comprehension				0.001	
				(0.001)	
A*Domestic*Logical				2.835e-04	
				(9.84E-04)	
A*Domestic*Hard Science				-5.67e-05	
				(5.54E-04)	
A*Domestic*Executive				-0.001	
				(0.001)	
A*Domestic*Technical				-9.208e-04	
				(9.09E-04)	
A*Domestic*Physical				-6.46e-05	
				(9.79E-04)	
Age Imm*Expression					-0.002*
					(0.001)
Age Imm*Comprehension					0.003**
					(0.001)
Age Imm*Logical					-9.728e-04
					(8.79E-04)
Age Imm*Hard Science					-7.802e-04
					(4.78E-04)
Age Imm*Executive					0.002**
					(0.001)
Age Imm*Technical					-0.001
					(8.01E-04)
Age Imm*Physical					0.001*
					(7.63E-04)
R^2	0.2311	0.2305	0.2302	0.2303	0.2101
Weighted N	936,835	936,835	936,835	936,835	936,835

Table. A1.8. Results For Immigrants with Other Type of Secondary Education

	Post-graduate		College	
	Coef.	Std.err.	Coef.	Std.Err.
Abroad	0.043	0.132	0.114*	0.064
Expression	0.067***	0.012	-0.039***	0.005
Comprehension	-0.09***	0.016	0.049***	0.006
Logical	0.07***	0.008	0.022***	0.006
Hard Science	-0.031***	0.005	-0.005	0.004
Executive	0.01	0.009	0.006	0.004
Technical	0.054***	0.008	0.013***	0.004
Physical	-0.067***	0.008	0.01***	0.003
Abroad*Expression	-0.067***	0.015	0.003	0.009
Abroad*Comprehension	0.049***	0.017	-0.02*	0.011
Abroad*Logical	-0.052***	0.011	-0.015	0.012
Abroad*Hard Science	0.039***	0.007	0.008	0.009
Abroad*Executive	0.006	0.012	0.006	0.008
Abroad*Technical	-0.006	0.01	7.42e-04	0.007
Abroad*Physical	0.023**	0.009	-0.016***	0.006
R^2	0.2393		0.1739	
Weighted N	836,810		600,750	

The first subsample consists of individuals who report a post-graduate degree (either a Master's or a Doctorate) as their highest level of education. The second subsample consists of individuals with one or two years College degrees. Actual sample size is about 20% of weighted sample size, see footnote 6. Asterisks denote significance at the 1%, 5% or 10% level.

Table. A1.9. Excluding Immigrants who Landed in the Past Decade Results

Variable	Coefficient	Standard Error
Abroad	0.355***	0.128
Expression	0.033**	0.016
Comprehension	-0.006	0.015
Logical	0.045***	0.01
Hard Science	-0.008	0.006
Executive	-0.016	0.013
Technical	0.051***	0.009
Physical	-0.026***	0.009
Abroad*Expression	-0.056***	0.02
Abroad*Comprehension	0.01	0.018
Abroad*Logical	-0.03**	0.014
Abroad*Hard Science	0.014*	0.008
Abroad*Executive	0.009	0.016
Abroad*Technical	-0.011	0.012
Abroad*Physical	-0.008	0.013
Weighted N	503,805	
R^2	0.1565	

Subsample of immigrants who have been landed for ten years or more.
2006 and 2016 Census and 2011 NHS data.
Actual sample size is about 20% of weighted sample size, see footnote 6.
Similar model to the main analysis reported in Table 1.6.

Table. A1.10. 2006 and 2016 Census Data Excluding 2011 National Household Survey

Variable	Coefficient	Standard Error
Abroad	0.042	0.145
Expression	0.012	0.018
Comprehension	-0.035*	0.018
Logical	0.032***	0.011
Hard Science	0.007	0.007
Executive	-0.001	0.015
Technical	0.031**	0.013
Physical	-0.045***	0.013
Abroad*Expression	-0.051**	0.021
Abroad*Comprehension	0.035*	0.02
Abroad*Logical	-0.021	0.013
Abroad*Hard Science	-0.007	0.009
Abroad*Executive	0.009	0.017
Abroad*Technical	0.027*	0.015
Abroad*Physical	0.007	0.014
Weighted N	615,385	
R^2	0.2548	

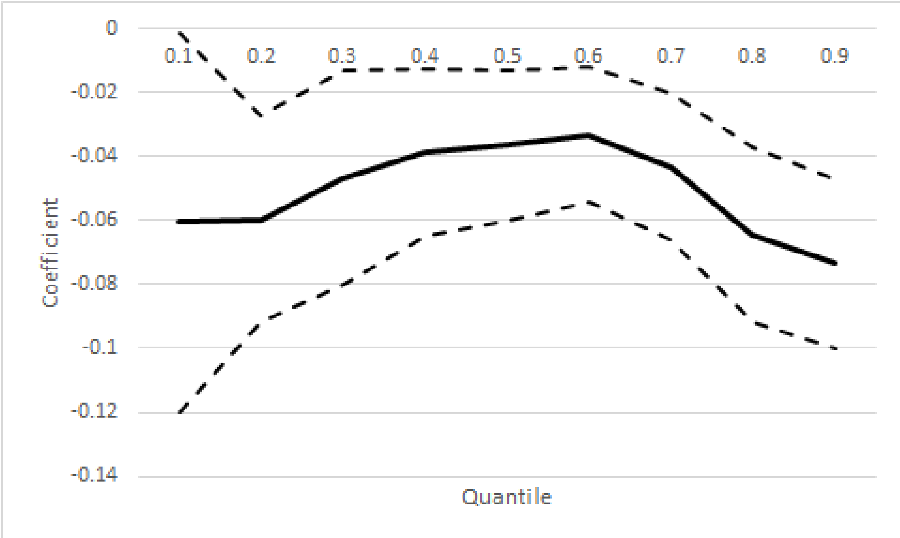
Sensitivity to excluding NHS survey data. Actual sample size is about 20% of weighted sample size, see footnote 6. Asterisks denote significance at the 1%, 5% or 10% level.

Table A1.11. Quantile Regression Estimation Main Results

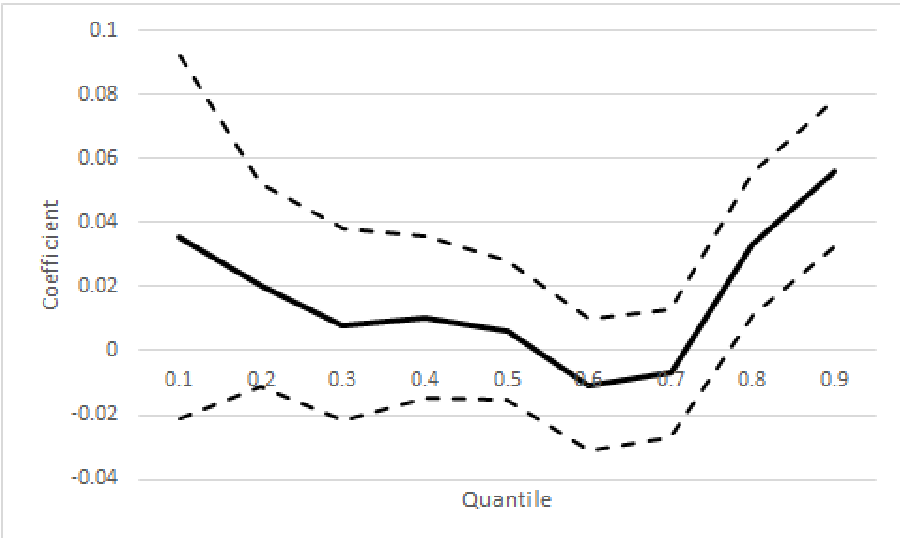
	20th Quantile	40th Quantile	60th Quantile	80th Quantile
Abroad	0.157	0.176**	0.263***	0.212**
	(0.114)	(0.088)	(0.072)	(0.085)
Expression	0.034**	0.011	0.006	0.034***
	(0.013)	(0.011)	(0.009)	(0.011)
Comprehension	-0.033**	-0.026**	-0.004	-0.041***
	(0.014)	(0.011)	(0.009)	(0.009)
Logical	0.039***	0.029***	0.03***	0.046***
	(0.01)	(0.008)	(0.007)	(0.007)
Hard Science	-0.007	0.0004291	0.002	0.003
	(0.006)	(0.004)	(0.003)	(0.004)
Executive	-0.024**	-0.006	-0.005	-0.013
	(0.01)	(0.009)	(0.007)	(0.009)
Technical	0.065***	0.05***	0.041***	0.029***
	(0.01)	(0.007)	(0.005)	(0.006)
Physical	-0.076***	-0.047***	-0.035***	-0.032***
	(0.01)	(0.008)	(0.006)	(0.007)
Abroad*Expression	-0.06***	-0.039***	-0.033***	-0.064***
	(0.017)	(0.013)	(0.011)	(0.014)
Abroad*Comprehension	0.02	0.01	-0.011	0.033***
	(0.016)	(0.013)	(0.01)	(0.011)
Abroad*Logical	-0.035***	-0.024**	-0.014*	-0.031***
	(0.012)	(0.009)	(0.008)	(0.009)
Abroad*Hard Science	0.017**	0.009*	0.007*	0.013**
	(0.007)	(0.005)	(0.004)	(0.005)
Abroad*Executive	0.023*	0.003	0.001	0.006
	(0.013)	(0.011)	(0.009)	(0.011)
Abroad*Technical	-0.018	-0.000476	0.005	0.005
	(0.012)	(0.009)	(0.007)	(0.008)
Abroad*Physical	0.025**	-0.003	-0.012*	0.001

	(0.011)	(0.009)	(0.007)	(0.009)
R^2	0.1626	0.1621	0.1578	0.1373
Weighted N	936,835	936,835	936,835	936,835

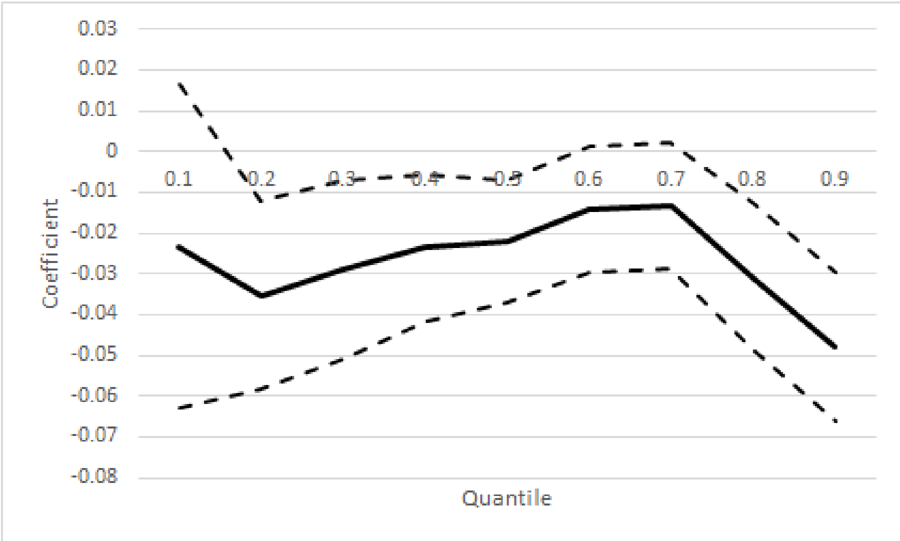
Fig. A1.1. Quantile Analysis



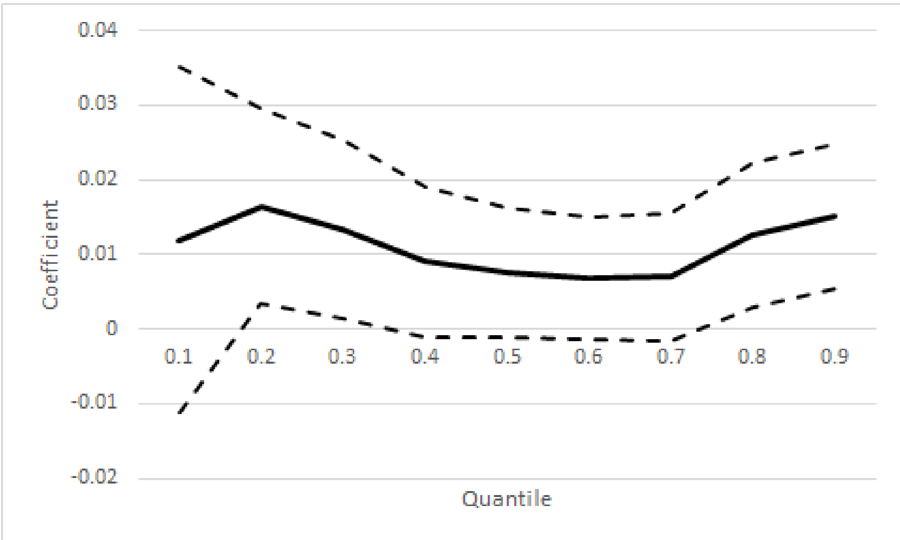
(a) Abroad*Expression



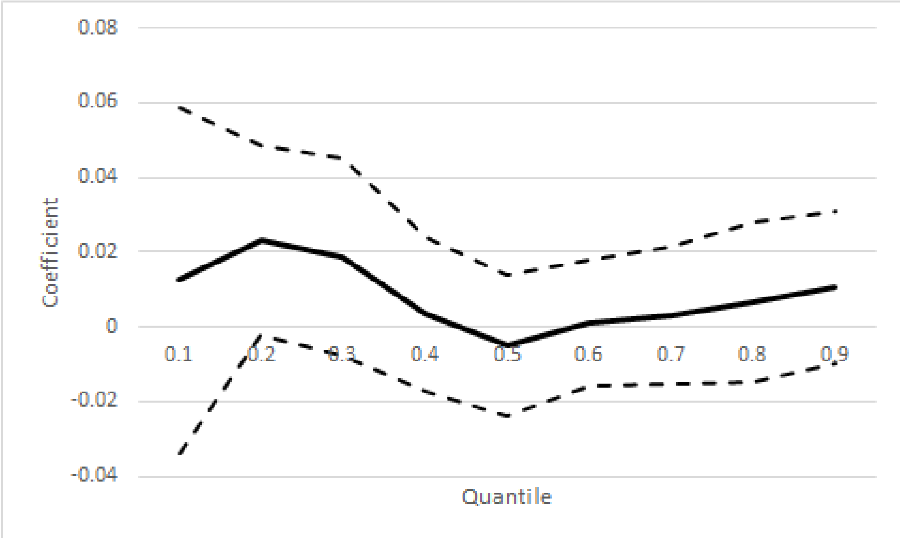
(b) Abroad*Comprehension



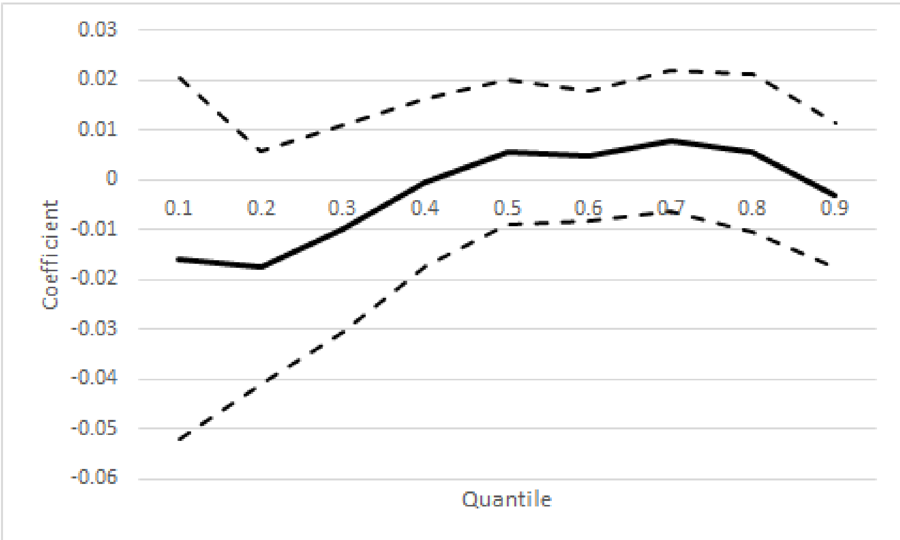
(c) Abroad*Logical



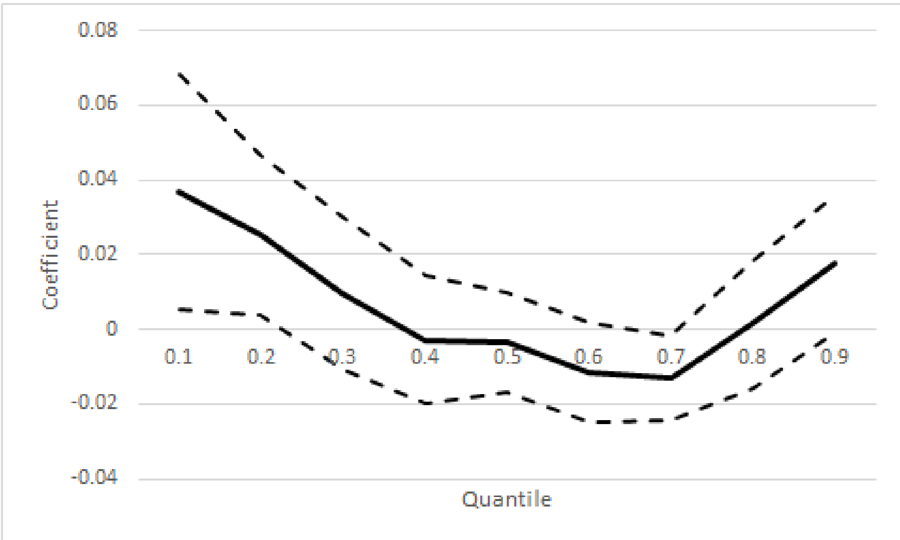
(d) Abroad*Hard Science



(e) Abroad*Executive



(f) Abroad*Technical



(g) Abroad*Physical

Quantile regressions were estimated at each decile of the conditional earnings distribution. These 7 figures demonstrate the estimates of the degree to which individuals who studied abroad are able to transfer their skills to the Canadian labour market. The dark solid line indicates the coe□cients from these quantile regressions, with the dotted lines indicating a 95% confidence band.

Table A1.12. Eigenvectors of Principal Components

Skill	Component 1	Component 2	Component 3
Written Expression	0.1568	0.0133	0.0149
Writing	0.1534	-0.0077	0.0279
Written Comprehension	0.1522	0.031	-0.0072
Oral Expression	0.1503	-0.0055	0.0659
Reading Comprehension	0.15	0.0261	0.0027
Oral Comprehension	0.149	0.0155	0.05
Active Listening	0.148	-0.028	0.0616
Active Learning	0.1457	0.0319	0.0606
Speaking	0.145	-0.0526	0.0859
Critical Thinking	0.1404	0.0694	0.0248
English Language	0.1376	-0.0561	0.0379
Judgement and Decision Making	0.137	0.0794	0.0594
Speech Clarity	0.1369	-0.0581	0.0983
Complex Problem Solving	0.1296	0.1018	0.0155
Deductive Reasoning	0.1247	0.1089	-0.0039
Learning Strategies	0.1212	-0.0139	0.1147
Fluency of Ideas	0.1209	0.0964	0.0205
Instructing	0.1178	-0.0013	0.1231
Systems Analysis	0.1166	0.1192	-0.0128
Persuasion	0.1161	0.0458	0.0903
Inductive Reasoning	0.1148	0.1102	0.0183
Systems Evaluation	0.1138	0.1129	0.0093
Originality	0.1131	0.0943	0.0368
Negotiation	0.1099	0.0416	0.0912
Time Management	0.1099	0.0198	0.1009
Monitoring	0.1086	0.042	0.1576
Communications and Media	0.1073	-0.043	0.0299
Law and Government	0.1072	0.0002	0.0485
Management of Personnel Resources	0.1031	0.0671	0.1211

Personnel and Human Resources	0.1011	0.0365	0.0981
Memorization	0.1003	0.0235	0.0514
Social Perception	0.0966	-0.0341	0.1719
Education and Training	0.0924	-0.028	0.142
Administration and Management	0.0922	0.0683	0.0449
Sociology and Anthropology	0.0921	-0.1156	0.1315
Geography	0.0873	0.0279	0.0405
Category Flexibility	0.0854	0.1573	-0.0257
Coordination	0.0837	0.036	0.1596
Operations Analysis	0.0804	0.1094	-0.0628
Psychology	0.0802	-0.0756	0.1813
History and Archeology	0.0748	-0.0958	0.0663
Problem Sensitivity	0.0705	0.1303	0.0802
Philosophy	0.0699	-0.1262	0.1266
Mathematical Reasoning	0.0621	0.1516	-0.094
Information Ordering	0.0618	0.1714	-0.0463
Management of Financial Resources	0.0608	0.1155	0.0166
Biology	0.0584	0.0585	0.1106
Mathematics	0.0569	0.1569	-0.0768
Management of Material Resources	0.0552	0.1276	0.071
Science	0.0549	0.1259	0.011
Clerical	0.0545	-0.0898	0.0093
Number Facility	0.0535	0.143	-0.0897
Therapy and Counseling	0.0535	-0.0876	0.1836
Economics and Accounting	0.0491	0.0304	-0.0456
Sales and Marketing	0.0477	0.0316	0.0658
Service Orientation	0.0475	-0.0733	0.1479
Foreign Language	0.0472	-0.0773	0.0939
Speed of Closure	0.0403	0.1403	0.037
Computers and Electronics	0.0372	0.1031	-0.1538
Medicine and Dentistry	0.03	-0.0129	0.1611
Customer and Personal Services	0.0258	-0.0492	0.1069

Fine Arts	0.0225	-0.064	0.0475
Near Vision	0.0221	0.0589	-0.0577
Chemistry	0.0198	0.1459	0.0701
Physics	0.012	0.179	-0.005
Transportation	0.0061	0.1207	0.0665
Engineering and Technology	0.0017	0.1924	-0.0593
Public Safety	0.0009	0.1162	0.1473
Building and Construction	0.0004	0.1613	0.0158
Telecommunications	0.0003	0.1154	-0.0837
Far Vision	-0.0017	0.149	0.1178
Flexibility of Closure	-0.002	0.178	-0.0095
Design	-0.0039	0.1782	-0.0582
Food Production	-0.0053	0.0322	0.1068
Time Sharing	-0.0073	0.0432	0.1849
Technology	-0.0088	0.158	-0.0997
Production and Processing	-0.0174	0.1579	0.0096
Selective Attention	-0.0177	0.1255	0.0229
Visualizaton	-0.0336	0.1825	-0.0001
Mechanical	-0.0394	0.1808	0.0243
Perceptual Speed	-0.0466	0.1658	-0.0109
Dynamic Flexibility	-0.0728	-0.013	0.1007
Depth Perception	-0.0747	0.1647	0.0947
Quality Control	-0.0808	0.1609	-0.0053
Explosive Strength	-0.0849	0.0064	0.1313
Hearing Sensitivity	-0.0851	0.1066	0.0923
Trunk Strength	-0.0881	-0.0409	0.1871
Spatial Orientation	-0.0885	0.0817	0.1136
Night Vision	-0.0938	0.0423	0.1075
Sound Localization	-0.0982	0.0376	0.109
Troubleshooting	-0.0992	0.1389	-0.0017
Repairing	-0.1057	0.0992	0.0175
Equipment Selection	-0.1096	0.1112	0.0004

Stamina	-0.1106	-0.0058	0.1876
Gross Body Coordination	-0.1134	-0.0008	0.1838
Gross Body Equilibrium	-0.1163	0.0111	0.1726
Speed of Limb Movement	-0.1178	0.0075	0.1604
Response Orientation	-0.123	0.0501	0.1326
Rate Control	-0.1231	0.0649	0.1096
Multilimb Coordination	-0.1246	0.0612	0.1489
Dynamic Strength	-0.1249	0.0042	0.1572
Reaction Time	-0.1252	0.0607	0.1215
Finger Dexterity	-0.1272	0.0878	0.0103
Control Precision	-0.1279	0.087	0.0568
Wrist-Finger Speed	-0.1283	0.0415	0.0468
Extent Flexibility	-0.1286	0.0097	0.1538
Static Strength	-0.1307	0.0116	0.1528
Arm-Hand Steadiness	-0.1399	0.0492	0.0658
Manual Dexterity	-0.1421	0.0488	0.0483

Table. A1.13. Model Estimated with Skills Derived From Principal Components

	Coefficient	Standard Error
Abroad	0.245***	0.075
Component 1	0.015**	0.007
Component 2	0.043***	0.004
Component 3	-0.033***	0.005
Abroad*Component 1	-0.034***	0.008
Abroad*Component 2	-0.005	0.004
Abroad*Component 3	-0.011*	0.006
Weighted N	936,835	
R^2	0.2296	

Skills grouped according to principal components.
2006 and 2016 Census and 2011 NHS data.
Actual sample size is about 20% of weighted sample size, see footnote 6.
Similar model to the main analysis reported in Table 1.6.
Asterisks denote significance at the 1%, 5% or 10% level.

Chapter 2 Appendix

2.1 Theoretical Model

As the skill level of a given occupation d increases, the probability that an individual chooses self-employment and region 2 changes according to:

$$
\begin{aligned}
\frac{\partial Pr(j = se \,\&\, r = 2)}{\partial \eta_d} &= \int_{\infty}^{\infty} \frac{\partial f_{d,se,2}}{\partial X}\left(-\frac{\partial E^i_{d,se,2}}{\partial \eta_d}\right) \prod_{p,q,z=\{d,se,2\}} \frac{\frac{\partial F_{d,se,2;p,q,z}}{\partial (X - E^i_{d,se,2})}}{f_{d,se,2}} dX \\
&+ \int_{-\infty}^{\infty} \sum_{d=1}^{D} \sum_{q,z=\{se,2\}} \left[\frac{\partial^2 F_{d,se,2;d,q,z}}{\partial^2 (X - E^i_{d,se,2})}\left(-\frac{\partial E^i_{d,se,2}}{\partial \eta_d}\right) - \frac{\frac{\partial F_{d,se,2;d,q,z}}{\partial (X - E^i_{d,se,2})}\left(-\frac{\partial E^i_{d,se,2}}{\partial \eta_d}\right)\frac{\partial f_{d,se,2}}{\partial X}}{(f_{d,se,2})}\right] \\
&\prod_{\alpha,\beta,\gamma=\{d,q,z\},\{d,se,2\}} \frac{\frac{\partial F_{d,se,2;d,q,z}}{\partial (X - E^i_{d,se,2})}}{f_{d,se,2}} dX + \sum_{d=1}^{D} \int_{\infty}^{\infty} f_{d,se,2} \sum_{q,z=\{se,2\}} \frac{f_{d,se,2;d,q,z}}{f_{d,se,2}}\left(-\frac{\partial E^i_{d,q,z}}{\partial \eta_d}\right) \\
&\prod_{\alpha,\beta,\gamma=\{d,se,2\},\{d,q,z\}} \frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial (X - E^i_{d,se,2})}}{f_{d,se,2}} dX \qquad \text{(A2.1)}
\end{aligned}
$$

Using integration by parts, the above expression becomes:

$$\int_{-\infty}^{\infty}\frac{\partial E^i_{d,se,2}}{\partial\eta_d}\prod_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}\,dX\prod_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX$$

$$-\sum_{d=1}^{D}\int_{\infty}^{\infty} f_{d,se,2}\prod_{q,z=\{se,2\}}\frac{f_{d,se,2;d,q,z}}{f_{d,se,2}}\left(\frac{\partial E^i_{d,q,z}}{\partial\eta_d}\right)\prod_{\alpha,\beta,\gamma=\{d,se,2\},\{d,q,z\}}\frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX \quad \text{(A2.2)}$$

Where $\frac{\partial E^i_{d,se,r}}{\partial\eta_d}=\int_a^b u\,(\theta_2(X_i,\eta_d)-c_{i,r}+\;)\frac{\partial\theta_r}{\partial\eta_d}f^{d,se,r}(\;)d$

and $\frac{\partial E^i_{d,we,r}}{\partial\eta_d}=\int_a^b u\,(W_2(X_i,\eta_d)-c_{i,r}+\;)\frac{\partial W_r}{\partial\eta_d}f^{d,we,r}(\;)d$

$$\frac{\partial Pr(r=2)}{\partial\eta_d}=\sum_{j=\{se,we\}}\int_{\infty}^{\infty}\frac{\partial f_{d,j,2}}{\partial X}\left(-\frac{\partial E^i_{d,j,2}}{\partial\eta_d}\right)\prod_{p,q,z=\{d,j,2\}}\frac{\frac{\partial F_{d,j,2;p,q,z}}{\partial(X-E^i_{d,j,2})}}{f_{d,j,2}}dX$$

$$+\int_{-\infty}^{\infty}\sum_{j=\{se,we\}}\sum_{d=1}^{D}\prod_{q,z=\{j,2\}}\left[\frac{\partial^2 F_{d,j,2;d,q,z}}{\partial^2(X-E^i_{d,j,2})}\left(-\frac{\partial E^i_{d,j,2}}{\partial\eta_d}\right)-\frac{\frac{\partial F_{d,j,2;d,q,z}}{\partial(X-E^i_{d,j,2})}\left(-\frac{\partial E^i_{d,j,2}}{\partial\eta_d}\right)\frac{\partial f_{d,j,2}}{\partial X}}{(f_{d,j,2})}\right]$$

$$\prod_{\alpha,\beta,\gamma=\{d,q,z\},\{d,j,2\}}\frac{\frac{\partial F_{d,j,2;d,q,z}}{\partial(X-E^i_{d,j,2})}}{f_{d,j,2}}dX+\sum_{j=\{se,we\}}\sum_{d=1}^{D}\int_{\infty}^{\infty} f_{d,j,2}\prod_{q,z=\{j,2\}}\frac{f_{d,j,2;d,q,z}}{f_{d,j,2}}\left(-\frac{\partial E^i_{d,q,z}}{\partial\eta_d}\right)$$

$$\prod_{\alpha,\beta,\gamma=\{d,j,2\},\{d,q,z\}}\frac{\frac{\partial F_{d,j,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,j,2})}}{f_{d,j,2}}dX \quad \text{(A2.3)}$$

Using integration by parts this becomes:

$$\sum_{j=\{se,we\}}\int_{-\infty}^{\infty}\frac{\partial E^i_{d,j,2}}{\partial\eta_d}\prod_{p,q,z=\{d,j,2\}} f_{d,j,2;p,q,z}\,dX\prod_{\alpha,\beta,\gamma=\{d,j,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,j,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,j,2})}}{f_{d,j,2}}dX$$

$$-\sum_{j=\{se,we\}}\sum_{d=1}^{D}\int_{\infty}^{\infty} f_{d,j,2}\prod_{q,z=\{j,2\}}\frac{f_{d,j,2;d,q,z}}{f_{d,j,2}}\left(\frac{\partial E^i_{d,q,z}}{\partial\eta_d}\right)\prod_{\alpha,\beta,\gamma=\{d,j,2\},\{d,q,z\}}\frac{\frac{\partial F_{d,j,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,j,2})}}{f_{d,j,2}}dX \quad \text{(A2.4)}$$

The change in the conditional probability of self-employment following an increase in the skill level is therefore given by:

$$\frac{\partial}{\partial \eta_d}\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)} = \Big[\int_{-\infty}^{\infty} \frac{\partial E^i_{d,se,2}}{\partial \eta_d} \square_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}\, dX$$

$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial (X-E^i_{d,se,2})}}{f_{d,se,2}} dX - \square_{d=1}^{D} \int_{\infty}^{\infty} f_{d,se,2} \square_{q,z=\{se,2\}} \frac{f_{d,se,2;d,q,z}}{f_{d,se,2}} \Big(\frac{\partial E^i_{d,q,z}}{\partial \eta_d}\Big)$$

$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{d,q,z\}} \frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial (X-E^i_{d,se,2})}}{f_{d,se,2}} dX\Big] \quad Pr(r=2)$$

$$-\Big[\square_{j=\{se,we\}} \int_{-\infty}^{\infty} \frac{\partial E^i_{d,j,2}}{\partial \eta_d} \square_{p,q,z=\{d,j,2\}} f_{d,j,2;p,q,z}\, dX \quad \square_{\alpha,\beta,\gamma=\{d,j,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,j,2;\alpha,\beta,\gamma}}{\partial (X-E^i_{d,j,2})}}{f_{d,j,2}} dX$$

$$- \square_{j=\{se,we\}} \square_{d=1}^{D} \int_{\infty}^{\infty} f_{d,j,2} \square_{q,z=\{j,2\}} \frac{f_{d,j,2;d,q,z}}{f_{d,j,2}} \Big(\frac{\partial E^i_{d,q,z}}{\partial \eta_d}\Big)$$

$$\square_{\alpha,\beta,\gamma=\{d,j,2\},\{d,q,z\}} \frac{\frac{\partial F_{d,j,2;\alpha,\beta,\gamma}}{\partial (X-E^i_{d,j,2})}}{f_{d,j,2}} dX\Big] \quad Pr(j=se\,\&\,r=2)/(Pr(r=2)^2) \qquad \text{(A2.5)}$$

When differentiated with respect to the self-employment skill price in region 2, this becomes:

$$\frac{\partial}{(\frac{\partial \theta_2}{\partial \eta_d})}\frac{\partial}{\partial \eta_d}\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)} = \int_{-\infty}^{\infty} \int_a^b [u\,(\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ] \square_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}$$

$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial (X-E^i_{d,se,2})}}{f_{d,se,2}} dX \quad \Big[\frac{Pr(r=2)-Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}\Big] \qquad \text{(A2.6)}$$

The equivalent differential with respect to the wage employment skill price is:

$$\frac{\partial}{(\frac{\partial W_2}{\partial \eta_d})}\frac{\partial}{\partial \eta_d}\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)} = \int_{-\infty}^{\infty} \int_a^b [u\,(W_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,we,2}(\ )d\ ] \square_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z}$$

$$\square_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial (X-E^i_{d,we,2})}}{f_{d,we,2}} dX \quad \Big[\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}\Big] \qquad \text{(A2.7)}$$

Differentiating with respect to the cost of migrating to region 2 gives us:

$$\frac{\partial}{\partial c_{i,2}}\frac{\partial}{(\frac{\partial\theta_2}{\partial\eta_d})}\frac{\partial}{\partial\eta_d}\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)}=-\square_{-\infty}^{\infty}\square_{a}^{b}[u\ (\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ]$$

$$\square_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}\ \square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX\quad[\frac{Pr(r=2)-Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]$$

$$+\square_{-\infty}^{\infty}\square_{a}^{b}[u\,(\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ]\ \square_{p,q,z=\{d,se,2\}}\frac{\partial f_{d,se,2;p,q,z}}{\partial(X-E^i_{d,se,2})}(-\frac{\partial E^i_{d,se,2}}{\partial c_{i,2}})$$

$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX\quad[\frac{Pr(r=2)-Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]$$

$$+\square_{-\infty}^{\infty}\square_{a}^{b}[u\,(\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ]\ \square_{p,q,z=\{d,se,2\}}\frac{\partial f_{d,se,2;p,q,2}}{\partial(X-E^i_{p,q,2})}(-\frac{\partial E^i_{p,q,2}}{\partial c_{i,2}})$$

$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,2\}}\frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX\quad[\frac{Pr(r=2)-Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]$$

$$+\square_{-\infty}^{\infty}\square_{a}^{b}[u\,(\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ]\ \square_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}$$

$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial^2 F_{d,se,2;\alpha,\beta,\gamma}}{\partial^2(X-E^i_{d,se,2})}f_{d,se,2}}{f_{d,se,2}}(-\frac{\partial E^i_{d,se,2}}{\partial c_{i,2}})\ \square_{\tau,\varphi,\delta=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,se,2;\tau,\varphi,\delta}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX$$

$$[\frac{Pr(r=2)-Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]+\square_{-\infty}^{\infty}\square_{a}^{b}[u\,(\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ]\ \square_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}$$

$$\square_{\alpha,\beta=\{d,se\},\{p,q\}}\frac{f_{d,se,2;\alpha,\beta,2}}{f_{d,se,2}}(-\frac{\partial E^i_{\alpha,\beta,2}}{\partial c_{i,2}})\ \square_{\tau,\varphi,\delta=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,se,2;\tau,\varphi,\delta}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX$$

$$[\frac{Pr(r=2)-Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]+\square_{-\infty}^{\infty}\square_{a}^{b}[u\,(\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ]\ \square_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}$$

$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}\frac{\partial f_{d,se,2}}{\partial X}}{f_{d,se,2}}(-\frac{\partial E^i_{d,se,2}}{\partial c_{i,2}})\ \square_{\tau,\varphi,\delta=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,se,2;\tau,\varphi,\delta}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX$$

$$[\frac{Pr(r=2)-Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]+\square_{-\infty}^{\infty}\square_{a}^{b}[u\,(\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ]\ \square_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}$$

$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX\quad[\frac{Pr(r=2)-Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]$$

$$+\square_{-\infty}^{\infty}\square_{a}^{b}[u\,(\theta_2(X_i,\eta_d)-c_{i,2}+\ )f^{d,se,2}(\ )d\ ]\ \square_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z}$$

$$\square_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}}\frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX\quad[(\frac{\partial Pr(r=2)}{\partial c_{i,2}}-\frac{\partial Pr(j=se\,\&\,r=2)}{\partial c_{i,2}})\quad Pr(r=2)^2$$

$$-(Pr(r=2)-Pr(j=se\,\&\,r=2))\quad 2\frac{\partial Pr(r=2)}{\partial c_{i,2}}]\quad\text{(A2.8)}$$

Which simplifies down to:

$$\frac{\partial}{\partial c_{i,2}}\frac{\partial}{(\frac{\partial \theta_2}{\partial \eta_d})}\frac{\partial}{\partial \eta_d}\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)} = -\int_{-\infty}^{\infty}\int_a^b [u\ (\theta_2(X_i,\eta_d) - c_{i,2} + \ )f^{d,se,2}(\ )d\ ]$$

$$\prod_{p,q,z=\{d,se,2\}} f_{d,se,2;p,q,z} \sum_{\alpha,\beta,\gamma=\{d,se,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,se,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,se,2})}}{f_{d,se,2}}dX \quad [\frac{Pr(r=2) - Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]$$

(A2.9)

The equivalent differential for wage employment is:

$$\frac{\partial}{\partial c_{i,2}}\frac{\partial}{(\frac{\partial W_2}{\partial \eta_d})}\frac{\partial}{\partial \eta_d}\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)} = -\int_{-\infty}^{\infty}\int_a^b [u\ (W_2(X_i,\eta_d) - c_{i,2} + \ )f^{d,we,2}(\ )d\ ]$$

$$\prod_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z} \sum_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,we,2})}}{f_{d,we,2}}dX \quad [\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]$$

$$+\int_{-\infty}^{\infty}\int_a^b [u\ (W_2(X_i,\eta_d) - c_{i,2} + \ )f^{d,we,2}(\ )d\ ] \prod_{p,q,z=\{d,we,2\}} \frac{\partial f_{d,we,2;p,q,z}}{\partial(X - E^i_{d,we,2})}(-\frac{\partial E^i_{d,we,2}}{\partial c_{i,2}})$$

$$\sum_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,we,2})}}{f_{d,we,2}}dX \quad [\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]$$

$$+\int_{-\infty}^{\infty}\int_a^b [u\ (W_2(X_i,\eta_d) - c_{i,2} + \ )f^{d,we,2}(\ )d\ ] \prod_{p,q,z=\{d,we,2\}} \frac{\partial f_{d,we,2;p,q,2}}{\partial(X - E^i_{p,q,2})}(-\frac{\partial E^i_{p,q,2}}{\partial c_{i,2}})$$

$$\sum_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,2\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial(X-E^i_{d,we,2})}}{f_{d,we,2}}dX \quad [\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}]$$

$$+\int_{-\infty}^{\infty}\int_a^b [u\ (W_2(X_i,\eta_d) - c_{i,2} + \ )f^{d,we,2}(\ )d\ ] \prod_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z}$$

$$\sum_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial^2 F_{d,we,2;\alpha,\beta,\gamma}}{\partial^2(X-E^i_{d,we,2})}f_{d,we,2}}{f_{d,we,2}}(-\frac{\partial E^i_{d,we,2}}{\partial c_{i,2}}) \sum_{\tau,\varphi,\delta=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\tau,\varphi,\delta}}{\partial(X-E^i_{d,we,2})}}{f_{d,we,2}}dX$$

$$[\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}] + \int_{-\infty}^{\infty}\int_a^b [u\ (W_2(X_i,\eta_d) - c_{i,2} + \ )f^{d,we,2}(\ )d\ ] \prod_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z}$$

$$\sum_{\alpha,\beta=\{d,we\},\{p,q\}} \frac{f_{d,we,2;\alpha,\beta,2}}{f_{d,we,2}}\left(-\frac{\partial E^{i}_{\alpha,\beta,2}}{\partial c_{i,2}}\right) \sum_{\tau,\varphi,\delta=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\tau,\varphi,\delta}}{\partial(X-E^{i}_{d,we,2})}}{f_{d,we,2}}dX \quad \left[\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}\right]$$

$$+\int_{-\infty}^{\infty}\int_{a}^{b}[u\,(W_2(X_i,\eta_d)-c_{i,2}+\;)f^{d,we,2}(\;)d\;] \sum_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z}$$

$$\sum_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial(X-E^{i}_{d,we,2})}\frac{\partial f_{d,we,2}}{\partial X}}{f_{d,we,2}}\left(-\frac{\partial E^{i}_{d,we,2}}{\partial c_{i,2}}\right) \sum_{\tau,\varphi,\delta=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\tau,\varphi,\delta}}{\partial(X-E^{i}_{d,we,2})}}{f_{d,we,2}}dX$$

$$\left[\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}\right]+\int_{-\infty}^{\infty}\int_{a}^{b}[u\,(W_2(X_i,\eta_d)-c_{i,2}+\;)f^{d,we,2}(\;)d\;] \sum_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z}$$

$$\sum_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial(X-E^{i}_{d,we,2})}}{f_{d,we,2}}dX \quad \left[\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}\right]$$

$$+\int_{-\infty}^{\infty}\int_{a}^{b}[u\,(W_2(X_i,\eta_d)-c_{i,2}+\;)f^{d,we,2}(\;)d\;] \sum_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z}$$

$$\sum_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial(X-E^{i}_{d,we,2})}}{f_{d,we,2}}dX \quad \left[\left(\frac{\partial Pr(r=2)}{\partial c_{i,2}}-\frac{\partial Pr(j=se\,\&\,r=2)}{\partial c_{i,2}}\right)\quad Pr(r=2)^2\right.$$

$$\left.-\left(Pr(r=2)-Pr(j=se\,\&\,r=2)\right)\quad 2\frac{\partial Pr(r=2)}{\partial c_{i,2}}\right] \quad \text{(A2.10)}$$

Which simplifies down to:

$$\frac{\partial}{\partial c_{i,2}}\frac{\partial}{(\frac{\partial W_2}{\partial \eta_d})}\frac{\partial}{\partial \eta_d}\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)}=-\int_{-\infty}^{\infty}\int_{a}^{b}[u\;(W_2(X_i,\eta_d)-c_{i,2}+\;)f^{d,we,2}(\;)d\;]$$

$$\sum_{p,q,z=\{d,we,2\}} f_{d,we,2;p,q,z} \sum_{\alpha,\beta,\gamma=\{d,we,2\},\{p,q,z\}} \frac{\frac{\partial F_{d,we,2;\alpha,\beta,\gamma}}{\partial(X-E^{i}_{d,we,2})}}{f_{d,we,2}}dX \quad \left[\frac{Pr(j=se\,\&\,r=2)}{Pr(r=2)^2}\right]$$

(A2.11)

2.2 Tables and Figures

Table A2.1. Occupation List With Self-Employment Rates

Occupations	Obs	% Self-Employed
Farmers, Ranchers, And Other Agricultural Managers	6600	80.3
Health Diagnosing And Treating Practitioners	350	77.1
Chiropractors	540	74.3
Fishing And Hunting Workers	500	60.9
Dentists	1790	60.2
Artists And Related Workers	2590	59.7
Photographers	2130	59.4
Real Estate Brokers And Sales Agents	9360	59
Podiatrists	100	58.9
Barbers	920	56.2
Door To Door Sales Workers, News And Street Vendors	1680	52.3
Jewelers And Precious Stone And Metal Workers	420	51.8
Hairdressers, Hairstylists, And Cosmetologists	8200	50.3
Animal Trainers	460	49.1
Optometrists	410	48.3
Massage Therapists	1920	48.2
Musicians, Singers, And Related Workers	2360	45.9
First-Line Supervisors Of Landscaping, Lawn Service And Groundskeeping Workers	1600	45.3
Writers And Authors	2930	44.9

Tailors, Dressmakers, And Sewers	830	43.2
Appraisers And Assessors Of Real Estate	1000	41.3
Miscellaneous Woodworkers, Including Model Makers And Patternmakers	350	40.7
Shoe And Leather Workers	120	40.2
Carpet, Floor, And Tile Installers And Finishers	1490	39.3
Psychologists	2300	39
Upholsterers	320	38.9
Firstline Supervisors Of Personal Service Workers	890	37.9
Painters And Paperhangers	5550	36.8
Actors	740	36.8
Carpenters	11640	36
Logging Workers	570	36
Management Analysts	9930	35.8
Taxi Drivers And Chauffeurs	4830	34.9
Veterinarians	890	34.6
Television, Video, And Motion Picture Camera Operators	730	34.4
Construction Managers	6690	34.2
Announcers	520	33.1
Home Appliance Repairers	390	32.8
Dancers And Choreographers	220	32.6
Childcare Workers	14990	32.4
Miscellaneous Personal Appearance Workers	3200	31.8
Miscellaneous Media And Communication Workers	1250	30.8

Furniture Finishers	130	30.3
Lawyers, And Judges, Magistrates, And Other Judicial Workers	12410	30.2
Automotive Glass Installers And Repairers	190	29.9
Nonfarm Animal Caretakers	2720	29.8
Chief Executives And Legislators	13750	28.9
Agents And Business Managers Of Artists, Performers	470	28.6
Locksmiths And Safe Repairers	270	28.5
Dredge, Excavating, And Loading Machine Operators	370	28.2
Cabinetmakers And Bench Carpenters	570	28
Property, Real Estate, And Community Association Managers	6340	27.8
First-Line Supervisors Of Housekeeping And Janitorial Workers	2190	27.5
Architects, Except Naval	2090	27.4
Designers	9980	27.3
Grounds Maintenance Workers	14050	26.8
Drywall Installers, Ceiling Tile Installers	1240	26.8
Insurance Sales Agents	5660	26.5
Travel Agents	780	25.8
Fence Erectors	260	25.4
Etchers And Engravers	110	25
Molders, Shapers, And Casters, Except Metal And Plastic	300	24.9
Construction Laborers	17540	24.8

Small Engine Mechanics	480	24.5
Producers And Directors	1760	24.4
First Line Supervisors Of Construction Trades And Extraction Workers	8250	23.6
Tax Preparers	1580	23.4
Physicians And Surgeons	9780	23
Proofreaders And Copy Markers	150	23
Brickmasons, Blockmasons, Stonemasons, And Reinforcing Iron And Rebar Workers	1380	22.9
Personal Financial Advisors	3870	22.5
Electronic Home Entertainment Equipment Installers	360	22.1
Motion Picture Projectionists	30	21.9
Editors	2230	21.8
Other Teachers And Instructors	10600	21.7
Food Service Managers	8660	21.6
Textile Knitting And Weaving Machine Setters	100	20.8
Morticians, Undertakers, And Funeral Directors	380	20.1
Other Therapists, Including Exercise Physiologists	1710	19.7
Medical Transcriptionists	540	19.7
Firstline Supervisors Of Non-Retail Sales Workers	12580	19.6
Automotive Service Technicians And Mechanics	8160	19.6
Plasterers And Stucco Masons	280	19.6
Roofers	1950	19.2

Automotive Body And Related Repairers	1320	19.1
Construction And Building Inspectors	1060	17.7
Miscellaneous Managers, Including Funeral Service	46460	17.4
Athletes, Coaches, Umpires, And Related Workers	3700	17.3
Miscellaneous Legal Support Workers	1820	16.9
Models, Demonstrators, And Product Promoters	790	16.9
Pest Control Workers	770	16.9
Firstline Supervisors Of Retail Sales Workers	30260	16.7
Broadcast And Sound Engineering	1170	16.4
Couriers And Messengers	2330	15.9
Miscellaneous Textile, Apparel, And Furnishings Workers, Except Upholsterers	210	15.8
Lodging Managers	1330	15.7
Recreation And Fitness Workers	5760	15
Pipelayers, Plumbers, Pipefitters, And Steamfitters	5430	14.9
Heating, Air Conditioning, And Refrigeration Mechanics	3760	14.8
Miscellaneous Construction Workers, Including Solar Photovoltaic Installers, Septic Tank Servicers And Sewer Pipe Cleaners	790	14.5
Private Detectives And Investigators	900	14.4
Tour And Travel Guides	810	14.3
Securities, Commodities, And Financial Services Sales	2400	14

Market Research Analysts And Marketing Specialists	3260	13.5
Computer, Automated Teller, And Office Machine Repairers	1940	13.5
Cement Masons, Concrete Finishers, And Terrazzo Workers	600	13.3
Electronic Equipment Installers And Repairers, Motor	80	13.3
Coin, Vending, And Amusement Machine Servicers And Repairers	370	13.2
Meeting Convention, And Event Planners	1580	13.1
Earth Drillers, Except Oil And Gas	240	13
Sewing Machine Operators	2050	12.8
Ship And Boat Captains And Operators	480	12.8
Refuse And Recyclable Material Collectors	1080	12.1
Medical, Dental, And Ophthalmic Laboratory Technicians	820	12
Other Education, Training, And Library Workers	1490	11.9
News Analysts, Reporters And Correspondents	840	11.9
Nurse Anesthetists	340	11.8
Advertising And Promotions Managers	490	11.4
Driver/Sales Workers And Truck Drivers	34360	11.3
Surveyors, Cartographers , And Photogrammetrists	380	11.3
Laundry And Dry Cleaning Workers	1960	11.2
Electricians	7850	11

Conservation Scientists And Foresters	240	11
Audiologists	160	11
Forest And Conservation Workers	170	10.9
Sales Representatives, Wholesale And Manufacturing	14380	10.8
Bookkeeping, Accounting, And Auditing Clerks	13950	10.8
Personal Care Aides	14200	10.7
Accountants And Auditors	21340	10.4
Public Relations Specialists	1480	10.4
Security And Fire Alarm Systems Installers	590	10.3
Furnace, Kiln, Oven, Drier, And Kettle Operators And Tenders	110	10.3
Sales Representatives, Services, All Other	6800	10.1
Advertising Sales Agents	1890	10.1
Cost Estimators	1510	10.1
Correspondence Clerks And Order Clerks	1550	9.9
Technical Writers	760	9.9
Surveying And Mapping Technicians	660	9.8
Miscellaneous Social Scientists, Including Survey Researchers And Sociologists	450	9.7
Wholesale And Retail Buyers, Except Farm Products	2180	9.5
Financial Specialists, All Other	540	9.5
Glaziers	400	9.5
Pumping Station Operators	270	9.4

Helpers Installation, Maintenance, And Repair Workers	210	9.4
Speech-Language Pathologists	1630	9.3
Other Healthcare Practitioners And Technical Occupations	1340	9.1
Drafters	1890	9
Chefs And Head Cooks	4060	8.9
Painting Workers	1360	8.9
Agricultural And Food Scientists	250	8.9
Physical Therapists	2490	8.8
Food And Tobacco Roasting, Baking, And Drying Machine Operators And Tenders	100	8.8
Counselors	8460	8.7
First Line Supervisors Of Farming, Fishing And Forestry Workers	610	8.7
Buyers And Purchasing Agents, Farm Products	100	8.7
Cleaners Of Vehicles And Equipment	3780	8.6
Bakers	2290	8.2
Civil Engineers	3800	8.1
Dietitians And Nutritionists	1100	8.1
Photographic Process Workers And Processing Machine Operators	310	7.9
Retail Salespersons	37340	7.7
Occupational Therapists	1170	7.7
Clergy	4890	7.6
Firstline Supervisors Of Food Preparation And Serving Workers	5460	7.5
Electric Motor, Power Tool, And	270	7.4

Related Repairers		
Janitors And Building Cleaners	26910	7.3
Miscellaneous Agricultural Workers,	8220	7.3
Miscellaneous Vehicle And Mobile Equipment Mechanics, Installers And Repairers	790	7.1
Structural Metal Fabricators And Fitters	250	7.1
Construction Equipment Operators, Except Paving, Surfacing And Tamping Equipment Operators	3560	7
Environmental Scientists And Geoscientists	880	7
General And Operations Managers	9210	6.7
Heavy Vehicle And Mobile Equipment Service Technicians	2100	6.5
Astronomers And Physicists	140	6.5
Welding, Soldering, And Brazing Workers	5750	6.4
Computer Systems Analysts	5600	6.4
Business Operations Specialists, All Other	3150	6.4
Archivists, Curators, And Museum Technicians	640	6.4
Printing Press Operators	1830	6.3
Ambulance Drivers And Attendants, Except Emergency Medical Technicians	130	6.3
Financial Analysts	2300	6.2
Gaming Managers	160	6.2
Sheet Metal Workers	1260	6.1
Counter And Rental Clerks	900	6.1

Sawing Machine Setters, Operators, And Tenders, Wood	300	6.1
Human Resources Managers	4930	6
Bus And Truck Mechanics And Diesel Engine Specialists	3040	6
Helpers, Construction Trades	490	5.6
Marketing And Sales Managers	10180	5.5
Supervisors Of Transportation And Material Moving Workers	2320	5.4
Insulation Workers	410	5.4
Miscellaneous Engineers, Including Nuclear Engineers	5940	5.3
Opticians, Dispensing	660	5.3
Urban And Regional Planners	270	5.3
Structural Iron And Steel Workers	600	5.2
Radio And Telecommunications Equipment Installers And Repairers	1740	5.1
Training And Development Specialists	1530	5
Transportation Inspectors	420	5
Medical And Health Services Managers	6910	4.9
Fundraisers	1200	4.8
Computer Hardware Engineers	520	4.8
Embalmers And Funeral Attendants	170	4.8
Maintenance And Repair Workers, General	5390	4.7
Sales Engineers	380	4.7
Secretaries And Administrative Assistants	38990	4.6
Computer Support Specialists	6970	4.6
Derrick, Rotary Drill, And Service	260	4.6

Unit Operators, And Roustabouts , Oil, Gas And Mining		
Cargo And Freight Agents	220	4.6
Word Processors And Typists	4240	4.5
Paving, Surfacing, And Tamping Equipment Operators	130	4.5
Economists	320	4.4
Telemarketers	950	4.3
Miscellaneous Office And Administrative Support	6530	4.2
Pharmacists	3250	4.2
Telecommunications Line Installers And Repairers	1420	4.2
Miscellaneous Extraction Workers, Including Roof Bolters And Helpers	580	4.2
Prepress Technicians And Workers	290	4.2
Financial Managers	12290	4.1
Nurse Practitioners And Nurse Midwives	1700	4.1
Software Developers,Applications And Systems Software	13310	4
Human Resources Workers	8490	4
Helpers Of Production Workers	510	4
Petroleum, Mining And Geological Engineers, Including Mining Safety Engineers	400	4
Firstline Supervisors Of Office And Administrative Support Workers	15080	3.9
Firstline Supervisors Of Production And Operating Workers	9160	3.9
Dental Hygienists	1830	3.9

Miscellaneous Health Technologists And Technicians	1260	3.9
Claims Adjusters, Appraisers, Examiners, And Investigators	3120	3.8
Computer Network Architects	1050	3.8
Aircraft Pilots And Flight Engineers	1850	3.7
Court, Municipal, And License Clerks	820	3.7
Natural Sciences Managers	250	3.7
Education Administrators	10450	3.6
Animal Control Workers	140	3.6
Machinists	3410	3.5
Computer Operators	1050	3.5
Gaming Services Workers	1010	3.5
Food Batchmakers	910	3.5
Pressers, Textile, Garment, And Related Materials	400	3.5
Agricultural And Food Science Technicians	370	3.5
Office Clerks, General	15220	3.4
Preschool And Kindergarten Teachers	6280	3.4
Billing And Posting Clerks	5130	3.4
Purchasing Agents, Except Wholesale, Retail, And Farm Products	3050	3.4
Electrical And Electronics Engineers	2320	3.4
Computer And Information Research Scientists	240	3.4
Woodworking Machine Setters, Operators, And Tenders, Except Sawing	180	3.4
Riggers	120	3.4

Paralegals And Legal Assistants	4420	3.3
Industrial And Refractory Machinery	3980	3.3
Biological Scientists	870	3.3
Engine And Other Machine Assemblers	120	3.3
Miscellaneous Community And Social Service Specialists, Including Health Educators And Community Health Workers	1020	3.2
Environmental Engineers	350	3.2
Marine Engineers And Naval Architects	160	3.2
Social And Community Service Managers	4080	3.1
Firstline Supervisors Of Mechanics And Installers	2980	3.1
Directors, Religious Activities And Education	870	3.1
Miscellaneous Mathematical Science Occupations, Including Mathematicians And Statisticians	710	3.1
Fire Inspectors	230	3.1
Food Cooking Machine Operators And Tenders	130	3.1
Textile Bleaching And Dyeing, And Cutting Machine Setters, Operators And Tenders	100	3.1
Nursing, Psychiatric, And Home Health Aides	19080	3
Computer Occupations, All Other	6660	3
Reservation And Transportation Ticket Agents And Travel Clerks	1380	3
Geological And Petroleum Technicians,	200	3

And Nuclear Technicians		
Control And Valve Installers And Repairers	200	3
Industrial Production Managers	2460	2.9
Payroll And Timekeeping Clerks	1730	2.9
Chemists And Materials Scientists	900	2.9
Millwrights	480	2.9
Cooks	22680	2.8
File Clerks	3000	2.8
Dispatchers	2840	2.8
Bill And Account Collectors	1580	2.8
Model Makers, Patternmakers, And Molding Machine-Setters, Metal And Plastic	430	2.8
Laborers And Freight, Stock, And Material Movers, Hand	23510	2.7
Social Workers	8700	2.7
Data Entry Keyers	4060	2.7
Miscellaneous Life, Physical, And Social Science Research Assistants	2460	2.7
Transportation, Storage, And Distribution Managers	2450	2.7
Butchers And Other Meat, Poultry, And Fish Processing Workers	2330	2.7
Architectural And Engineering Managers	1760	2.7
Parts Salespersons	1080	2.7
Public Relations And Fundraising Managers	670	2.7
Veterinary Assistants And Laboratory	510	2.7

Animal Caretakers		
Boilermakers	190	2.7
Emergency Management Directors	110	2.7
Inspectors, Testers, Sorters, Samplers, And Weighers	8530	2.5
Computer And Information Systems Managers	6770	2.5
Production, Planning, And Expediting	3610	2.5
Credit Counselors And Loan Officers	3180	2.5
Mechanical Engineers	2730	2.5
Network And Computer Systems Administrators	2250	2.5
Database Administrators	1250	2.5
Miscellaneous Assemblers And Fabricators	9970	2.4
Diagnostic Related Technologists And Technicians	3670	2.4
Physical Scientists, All Other	2620	2.4
Miscellaneous Entertainment Attendants And Related Workers	2600	2.4
Administrative Services Managers	1760	2.4
Compensation, Benefits, And Job Analysis Specialists	580	2.4
Chemical Processing Machine Setters, Operators, And Tenders	550	2.4
Miscellaneous Plant And System Operators	430	2.4
Biological Technicians	250	2.4
Financial Examiners	170	2.4
Bartenders	4120	2.3

Machine Tool Cutting Setters, Operators, And Tenders, Metal And Plastic	1400	2.3
Loan Interviewers And Clerks	1230	2.3
Physician Assistants	1000	2.3
Training And Development Managers	700	2.3
Crane And Tower Operators	620	2.3
Machine Feeders And Offbearers	300	2.3
Social And Human Service Assistants	2060	2.2
Interviewers, Except Eligibility And Loan	1610	2.2
Highway Maintenance Workers	1130	2.2
Automotive And Watercraft Service Attendants	910	2.2
Baggage Porters, Bellhops , And Concierges	820	2.2
Materials Engineers	410	2.2
Hazardous Materials Removal Workers	320	2.2
Occupational Therapy Assistants And Aides	230	2.2
Miscellaneous Production Workers, Including Semiconductor Processors	11920	2.1
Compensation And Benefits Managers	190	2.1
Conveyor Operators And Tenders, And Hoist And Winch Operators	150	2.1
Recreational Therapists	140	2.1
Counter Attendants, Cafeteria, Food Concession, And Coffee Shop	1980	2
Electrical Power-Line Installers And	1140	2

Repairers		
Physical Therapist Assistants And Aides	1030	2
Mining Machine Operators	510	2
Print Binding And Finishing Workers	210	2
Logisticians	1460	1.9
Healthcare Support Workers, All Other, Including Medical Equipment Preparers	1450	1.8
Insurance Underwriters	1180	1.8
Procurement Clerks	390	1.8
Licensed Practical And Licensed Vocational Nurses	8520	1.7
Miscellaneous Metal Workers And Plastic Workers, Including Multiple Machine Tool Setters	4000	1.7
Medical Scientists, And Life Scientists, All Other	1580	1.7
Extruding And Drawing Machine Setters, Operators, And Tenders, Metal And Plastic	120	1.7
Customer Service Representatives	28190	1.6
Postsecondary Teachers	18300	1.6
Dental Assistants	2880	1.6
Purchasing Managers	2220	1.6
Operations Research Analysts	1520	1.6
Aerospace Engineers	1440	1.6
Electrical, Electronics, And Electromechanical Assemblers	1380	1.6
Information Security Analysts	850	1.6
Pharmacy Aides	380	1.6
Actuaries	310	1.6

Paper Goods Machine Setters, Operators, And Tenders	250	1.6
Electrical And Electronics Repairers, Transportation	190	1.6
Cashiers	35820	1.5
Receptionists And Information Clerks	12700	1.5
Bus Drivers	6440	1.5
Tool And Die Makers	650	1.5
Office Machine Operators , Except Computer	470	1.5
Sailors And Marine Oilers, And Ship Engineers	340	1.5
Avionics Technicians	260	1.5
Food Preparation Workers	10140	1.4
Engineering Technicians , Except Drafters	4140	1.4
Compliance Officers	2750	1.4
Industrial Engineers, Including Health And Safety	2130	1.4
Medical Records And Health Information Technicians	1890	1.4
Parking Lot Attendants	940	1.4
Crushing, Grinding, Polishing, Mixing, And Blending Workers	840	1.4
Telephone Operators	420	1.4
Firstline Supervisors Of Gaming Workers	210	1.4
Forging Machine Setters, Operators, And Tenders, Metal	70	1.4
Shipping, Receiving, And Traffic Clerks	5940	1.3

Computer Control Programmers And Operators	900	1.3
Chemical Technicians	720	1.3
Security Guards And Gaming Surveillance Officers	9800	1.2
Packers And Packagers, Hand	5050	1.2
Insurance Claims And Policy Processing Clerks	4010	1.2
Combined Food Preparation And Serving Workers, Including Fast Food	4000	1.2
Aircraft Mechanics And Service Technicians	1850	1.2
Brokerage Clerks	80	1.2
Registered Nurses	32750	1.1
Hosts And Hostesses, Restaurant, Lounge, And Coffee Shops	3170	1.1
Packaging And Filling Machine Operators And Tenders	2590	1.1
Weighers, Measurers, Checkers, And Samplers, Recordkeeping	820	1.1
Credit Authorizers, Checkers, And Clerks	370	1.1
Maintenance Workers, Machinery	280	1.1
Explosives Workers, Ordnance Handling Experts, And Blasters	180	1.1
Postal Service Mail Carriers	3690	1
Locomotive Engineers And Operators	500	1
Extruding, Forming, Pressing, And Compacting Machine Setters	300	1

, Operators And Tenders		
Textile Winding, Twisting, And Drawing Out Machine Setters, Operators And Tenders	100	1
Waiters And Waitresses	21260	0.9
Stock Clerks And Order Fillers	17620	0.9
Medical Assistants	4880	0.9
Clinical Laboratory Technologists And Technicians	3270	0.9
Hotel, Motel, And Resort Desk Clerks	1410	0.9
Phlebotomists	990	0.9
Stationary Engineers And Boiler Operators	940	0.9
Chemical Engineers	700	0.9
Statistical Assistants	220	0.9
Miscellaneous Law Enforcement Workers	110	0.9
Rail Track Laying And Maintenance	110	0.9
Elementary And Middle School Teachers	42930	0.8
Secondary School Teachers	8670	0.8
Special Education Teachers	2860	0.8
Budget Analysts	600	0.8
Meter Readers, Utilities	270	0.8
Elevator Installers And Repairers	260	0.8
Food Servers, Nonrestaurant	1830	0.7
Detectives And Criminal Investigators	1300	0.7
Water And Wastewater Treatment Plant And System Operators	940	0.7
Graders And Sorters, Agricultural Products	690	0.7
Atmospheric And Space Scientists	140	0.7
Radiation Therapists	140	0.7

Health Practitioner Support Technologists And Technicians	6030	0.6
Industrial Truck And Tractor Operators	5610	0.6
Miscellaneous Food Preparation And Serving Related Workers, Including Dining Room And Cafeteria Attendants And Bartender Helpers	3480	0.6
Dishwashers	3440	0.6
Lifeguards And Other Recreational, And All Other	2340	0.6
Emergency Medical Technicians And Paramedics	1950	0.6
Mail Clerks And Mail Machine Operators, Except Postal	930	0.6
Miscellaneous Material Moving Workers, Including Mine Shuttle Car Operators, And Tank Car, Truck And Ship Loaders	460	0.6
Agricultural Inspectors	170	0.6
Biomedical And Agricultural Engineers	170	0.6
Postal Service Clerks	1390	0.5
Air Traffic Controllers And Airfield Operations Specialists	410	0.5
Subway, Streetcar, And Other Rail Transportation Workers	210	0.5
Librarians	2290	0.4
Human Resources Assistants, Except Payroll And Timekeeping	550	0.4
Power Plant Operators, Distributors, And Dispatchers	510	0.4
Switchboard Operators, Including	280	0.4

Answering Service		
Credit Analysts	280	0.4
Tellers	3400	0.3
Respiratory Therapists	1150	0.3
Postal Service Mail Sorters, Processors, And Processing Machine Operators	770	0.3
Crossing Guards	750	0.3
Transportation Attendants, Except Flight Attendants	360	0.3
Miscellaneous Transportation Workers, Including Bridge And Lock Tenders And Traffic Technicians	310	0.3
Library Assistants, Clerical	1310	0.2
Residential Advisors	1230	0.2
Flight Attendants	1120	0.2
Tax Examiners And Collectors, And Revenue Agents	590	0.2
First Line Supervisors Of Fire Fighting And Prevention Workers	550	0.2
Library Technicians	540	0.2
Ushers, Lobby Attendants, And Ticket Takers	540	0.2
Railroad Conductors And Yardmasters	470	0.2
Transportation Security Screeners	420	0.2
Police Officers	7210	0.1
Eligibility Interviewers, Government Programs	900	0.1
Probation Officers And Correctional Treatment Specialists	870	0.1

Bailiffs, Correctional Officers, And Jailers	3890	0
Firefighters	2950	0
First Line Supervisors Of Police And Detectives	1240	0
First Line Supervisors Of Correctional Officers	660	0
Metal Furnace Operators, Tenders, Pourers, And Casters	210	0
New Accounts Clerks	140	0
Tire Builders	130	0
Gaming Cage Workers	90	0
Rolling Machine Setters, Operators, And Tenders, Metal And Plastic	90	0
Adhesive Bonding Machine Operators And Tenders	90	0
Aircraft Structure, Surfaces, Rigging, And Systems Assemblers	70	0

1st column gives the name of the occupation.

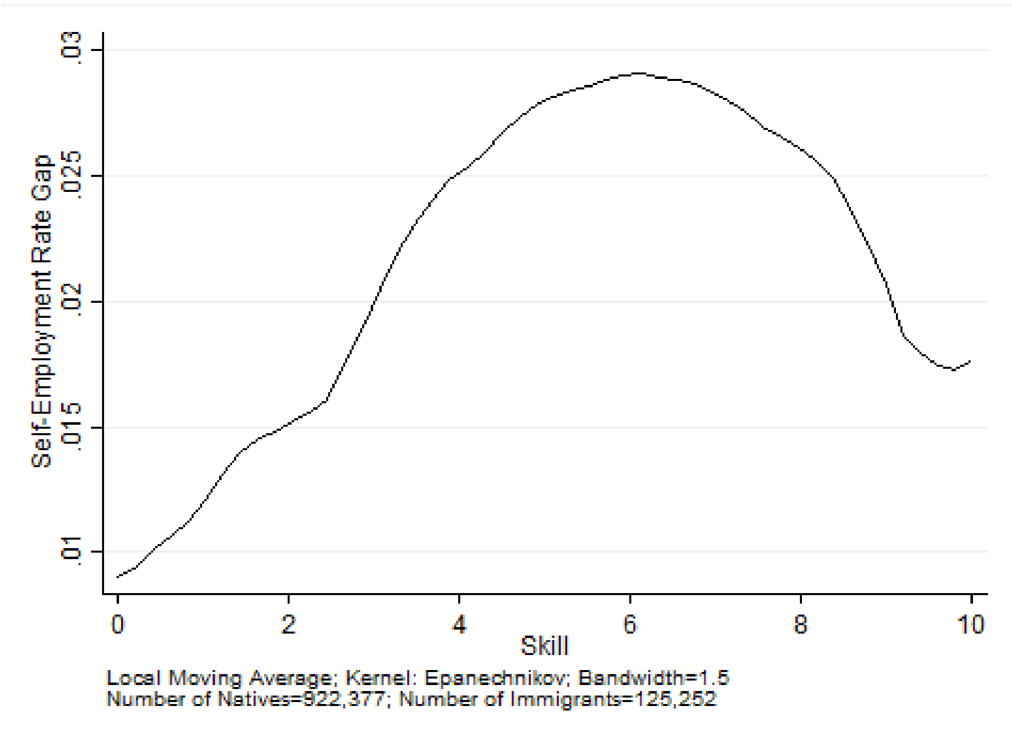
2nd column the number of observations (rounded to the nearest 10).

3rd column gives the percentage of individuals working in that occupation who are self-employed

Fig. A2.1. Density of Immigrants and Natives By Self-Employment Rate of Occupation

X-axis consists of the proportion of individuals within a given occupation who are self-employed. Y-axis consists of the separate densities of immigrants and natives in these occupations.

Fig. A2.2. Gaps in Self-Employment Rates Between Immigrants and Natives Conditional on Skill Level (No Outliers Eliminated)



This graph is equivalent to Figure 2.6, however no outliers are eliminated.

Table A2.2. Fixed Effects Model with Local Housing Prices and Macroeconomic Characteristics

Variable	Coefficient	Standard Error
2008	-0.034***	(0.002)
2009	0.004	(0.003)
2010	0.031***	(0.003)
2011	-0.017***	(0.003)
2012	-0.04***	(0.003)
2013	-0.016***	(0.003)
2014	-0.022***	(0.003)
2015	-0.048***	(0.003)
2016	-0.034***	(0.003)
Unemployment Rate	-0.024***	(0.001)
GDP	-1.2e-06	(7.42e-07)
Labor Force	1.06e-06***	(5.54e-08)
Unemployment Rate2	7.321e-04***	(6.79e-05)
GDP2	2.61e-12***	(5.74e-12)
Labor Force2	-2.68e-13***	(1.33e-14)
Lag˙Unemployment Rate	-0.017***	(0.002)
Lag˙GDP	3.43e-06***	(1.07e-06)
Lag˙Labor Force	-5.69e-07***	(7.55e-08)
Lag˙Unemployment Rate2	8.049e-04***	(8.57e-05)
Lag˙GDP2	-4.83e-12***	(9.05e-12)
Lag˙Labor Force2	8.94e-13***	(1.67e-14)
Lag2˙Unemployment Rate	0.02***	(0.002)
Lag2˙GDP	1.72e-07	(1.09e-06)
Lag2˙Labor Force	2.14e-07***	(7.26e-08)
Lag2˙Unemployment Rate2	-6.139e-04***	(8.74e-05)
Lag2˙GDP2	-2.10e-12	(9.35e-12)
Lag2˙Labor Force2	-2.97e-14**	(1.54e-14)

Lag3˙Unemployment Rate	0.003	(0.002)
Lag3˙GDP	-8.36e-06***	(1.15e-06)
Lag3˙Labor Force	-7.4e-07***	(6.84e-08)
Lag3˙Unemployment Rate2	2.572e-04***	(8.62e-05)
Lag3˙GDP2	5.91e-12***	(1e-09)
Lag3˙Labor Force2	1.79e-13***	(1.47e-14)
Lag4˙Unemployment Rate	-0.005***	(0.002)
Lag4˙GDP	-5.41e-06***	(1.16e-06)
Lag4˙Labor Force	10e-07***	(6.89e-08)
Lag4˙Unemployment Rate2	5.281e-04***	(8.66e-05)
Lag4˙GDP2	2.63e-12**	(1.02e-11)
Lag4˙Labor Force2	-2.30e-13***	(1.79e-14)
Lag5˙Unemployment Rate	0.005***	(0.002)
Lag5˙GDP	-1.82e-06*	(1.08e-06)
Lag5˙Labor Force	-7.28e-07***	(6.7e-08)
Lag5˙Unemployment Rate2	-1.075e-04	(8.43e-05)
Lag5˙GDP2	2.54e-12***	(2.54e-11)
Lag5˙Labor Force2	1.88e-13***	(1.88e-13)
Lag6˙Unemployment Rate	0.005***	(0.001)
Lag6˙GDP	4.56e-06***	(8.07e-07)
Lag6˙Labor Force	4.73e-07***	(4.37e-08)
Lag6˙Unemployment Rate2	-5.58e-05	(6.28e-05)
Lag6˙GDP2	-2.52e-12	(0.007e-09)
Lag6˙Labor Force2	-2.78e-14**	(1.15e-14)
Constant	-0.062**	(0.029)
ρ	0.98	
$\sigma^2_{\alpha_{pu}}$	0.29	
σ^2_{i}	0.041	

N=21,440 Numb. of Groups=2146 F(51,19243)=758.24, Within R^2 = 0.6677

Dependent Variable: Local Housing Price Index (Normalized to 1 in base year 2001)

Table. A2.3. Full Results for Self-Employment Earnings Equation (No Skill*Immigrant Interaction)

Variable	Coefficient	Standard Error
Agriculture	-0.159***	0.041
Mining	0.187**	0.087
Utilities	0.146*	0.081
Construction	-0.095**	0.045
Manufacturing	0.03	0.039
Wholesale	0.049	0.045
Retail	-0.077*	0.042
Transportation	0.005	0.045
Cultural	0.001	0.044
Finance	0.175***	0.042
Real Estate	0.145***	0.042
Professional	0.071	0.046
Management	0.446***	0.125
Support	-0.107**	0.047
Education	-0.129***	0.04
Health	0.016	0.041
Arts	-0.105**	0.046
Food	-0.125***	0.044
Other	-0.143***	0.046
Female	-0.129***	0.007
Married	0.067***	0.007
High School	0.034***	0.01
Below Bachelor's	0.06***	0.01
Bachelor's	0.184***	0.011
Above Bachelor's	0.354***	0.013
Immigrant	-0.098***	0.01
Age	0.017***	0.002
Age2	-1.587e-04***	1.64e-05
Native	-0.036**	0.016
Latin American	-0.071***	0.01
Middle Eastern	0.011	0.025
African	-0.081***	0.014
South Asian	0.022	0.028
East Asian	0.005	0.016
Pacific Islander	0.13**	0.052
Occupation Self Employment Rate	-0.298***	0.083
Own	0.035***	0.007
Skill	0.055***	0.002
Constant	2.447***	0.132

N=918,914; Uncensored N=75,244; χ^2= 9570.07; P-value=0.000, λ=-0.0765**
Results from Heckman Correction estimation of Equation 2.24.

Table. A2.4. Full Results for Wage Employment Earnings Equation (No Skill*Immigrant Interaction)

Variable	Coefficient	Standard Error
Agriculture	-0.216***	0.011
Mining	0.221***	0.01
Utilities	0.213***	0.007
Construction	0.017***	0.004
Manufacturing	0.043***	0.003
Wholesale	0.013***	0.004
Retail	-0.155***	0.003
Transportation	0.029***	0.004
Cultural	0.042***	0.005
Finance	0.072***	0.003
Real Estate	-0.016***	0.006
Professional	0.087***	0.003
Management	0.132***	0.016
Support	-0.12***	0.004
Education	-0.215***	0.003
Health	-0.09***	0.003
Arts	-0.147***	0.006
Food	-0.18***	0.003
Other	-0.179***	0.004
Female	-0.129***	0.001
Married	0.072***	0.001
High School	0.058***	0.003
Below Bachelor's	0.125***	0.002
Bachelor's	0.304***	0.003
Above Bachelor's	0.474***	0.003
Immigrant	-0.088***	0.002
Age	0.039***	3.232e-04
Age2	-3.623e-04***	3.66e-06
Native	-0.048***	0.004
Latin American	-0.033***	0.002
Middle Eastern	0.02***	0.007
African	-0.049***	0.002
South Asian	0.103***	0.007
East Asian	0.045***	0.004
Pacific Islander	0.032***	0.011
Occupation Self Employment Rate	-0.069***	0.008
Own	0.053***	0.001
Skill	0.07***	3.703e-04
Constant	1.604***	0.007

N=918,914; Uncensored N=880,561; χ^2 = 318,268; P-value=0.000, λ=-0.023***
Results from Heckman Correction estimation of Equation 2.23.

Table. A2.5. Full Results for Self-Employment Selection Equation (No Skill*Immigrant Interaction)

Variable	Coefficient	Standard Error
Agriculture	0.662***	0.04
Mining	0.392***	0.06
Utilities	-0.157**	0.074
Construction	1.092***	0.027
Manufacturing	0.367***	0.028
Wholesale	0.798***	0.03
Retail	0.789***	0.027
Transportation	0.984***	0.028
Cultural	0.659***	0.032
Finance	0.618***	0.028
Real Estate	0.756***	0.03
Professional	1.145***	0.027
Management	0.305**	0.133
Support	1.149***	0.028
Education	0.327***	0.029
Health	0.72***	0.027
Arts	1.061***	0.031
Food	0.899***	0.029
Other	1.143***	0.028
Female	-0.152***	0.006
Married	0.125***	0.006
High School	0.01	0.011
Below Bachelor's	0.074***	0.011
Bachelor's	0.16***	0.012
Above Bachelor's	0.179***	0.013
Immigrant	0.103***	0.011
Age	0.042***	0.002
Age2	-2.403e-04***	1.66e-05
Native	0.001	0.017
Latin American	-0.138***	0.011
Middle Eastern	0.28***	0.025
African	-0.158***	0.013
South Asian	0.038	0.026
East Asian	-0.045***	0.016
Pacific Islander	-0.189***	0.056
Occupation Self Employment Rate	3.761***	0.021
Skill	-0.019***	0.002
Own	0.1***	0.007
Housing Residual	0.273	0.187
Homeowner*Housing Residual	0.464**	0.223
Constant	-4.208***	0.046

N=918,914; Uncensored N=75,244; χ^2 = 318,729; P-value=0.000
Results from Heckman Correction estimation of Equation 2.24.

Table. A2.6. Full Results for Wage Employment Selection Equation (No Skill*Immigrant Interaction)

Variable	Coefficient	Standard Error
Agriculture	-0.486***	0.042
Mining	-0.267***	0.084
Utilities	0.176*	0.091
Construction	-0.982***	0.032
Manufacturing	-0.221***	0.035
Wholesale	-0.549***	0.037
Retail	-0.564***	0.033
Transportation	-0.835***	0.034
Cultural	-0.587***	0.04
Finance	-0.476***	0.035
Real Estate	-0.543***	0.036
Professional	-0.95***	0.032
Management	-0.341**	0.161
Support	-0.994***	0.034
Education	-0.353***	0.035
Health	-0.637***	0.032
Arts	-0.95***	0.036
Food	-0.537***	0.036
Other	-1.036***	0.033
Female	0.078***	0.008
Married	-0.027***	0.008
High School	0.032**	0.013
Below Bachelor's	0.023*	0.012
Bachelor's	0.009	0.014
Above Bachelor's	0.038**	0.017
Immigrant	-0.099***	0.012
Age	-0.033***	0.002
Age2	1.91e-04***	1.95e-05
Native	-0.041**	0.02
Latin American	0.042***	0.013
Middle Eastern	-0.122***	0.033
African	0.106***	0.015
South Asian	0.054	0.035
East Asian	0.043**	0.019
Pacific Islander	0.073	0.068
Occupation Self Employment Rate	-3.167***	0.023
Skill	0.043***	0.002
Own	-0.019**	0.008
Housing Residual	-0.005	0.215
Homeowner*Housing Residual	0.099	0.265
Constant	3.771***	0.053

N=918,915; Uncensored N=880,560; χ^2=318,269; P-value=0.000
Results from Heckman Correction estimation of Equation 2.23.

Table. A2.7. Full Results for Self-Employment Earnings Equation (with Skill*Immigrant Interaction)

Variable	Coefficient	Standard Error
Agriculture	-0.154***	0.041
Mining	0.19**	0.088
Utilities	0.145*	0.082
Construction	-0.093**	0.046
Manufacturing	0.03	0.039
Wholesale	0.05	0.046
Retail	-0.076*	0.042
Transportation	0.01	0.046
Cultural	0.004	0.044
Finance	0.177***	0.042
Real Estate	0.146***	0.042
Professional	0.074	0.047
Management	0.449***	0.124
Support	-0.105**	0.048
Education	-0.127***	0.041
Health	0.018	0.041
Arts	-0.104**	0.047
Food	-0.125***	0.045
Other	-0.141***	0.047
Female	-0.129***	0.007
Married	0.068***	0.007
High School	0.032***	0.01
Below Bachelor's	0.057***	0.01
Bachelor's	0.183***	0.011
Above Bachelor's	0.353***	0.013
Immigrant	-0.303***	0.053
Age	0.017***	0.002
Age2	-1.599e-04***	1.65e-05
Indigenous	-0.037**	0.016
Latin American	-0.068***	0.01
Middle Eastern	0.008	0.025
African	-0.082***	0.015
South Asian	0.016	0.028
East Asian	0.003	0.016
Pacific Islander	0.128**	0.052
Occupation Self Employment Rate	-0.292***	0.086
Own	0.035***	0.007
Skill	0.052***	0.002
Immigrant*Skill	0.029***	0.007
Constant	2.456***	0.135

N=918,914; Uncensored N=75,244; $\chi^2 =$ 9596.17; P-value=0.000, λ=-0.0752***

Results from Heckman Correction estimation of Equation 2.24, with the addition of an interaction between skill and immigrants status.

Table. A2.8. Full Results for Wage Employment Earnings Equation (with Skill*Immigrant Interaction)

Variable	Coefficient	Standard Error
Agriculture	-0.216***	0.011
Mining	0.218***	0.01
Utilities	0.211***	0.007
Construction	0.016***	0.004
Manufacturing	0.042***	0.003
Wholesale	0.012***	0.004
Retail	-0.157***	0.003
Transportation	0.028***	0.004
Cultural	0.041***	0.005
Finance	0.071***	0.003
Real Estate	-0.019***	0.006
Professional	0.085***	0.003
Management	0.131***	0.016
Support	-0.121***	0.004
Education	-0.215***	0.003
Health	-0.091***	0.003
Arts	-0.149***	0.006
Food	-0.181***	0.003
Other	-0.179***	0.004
Female	-0.128***	0.001
Married	0.073***	0.001
High School	0.053***	0.003
Below Bachelor's	0.119***	0.002
Bachelor's	0.3***	0.003
Above Bachelor's	0.467***	0.003
Immigrant	-0.379***	0.012
Age	0.039***	3.234e-04
Age2	-3.661e-04***	3.66e-06
Indigenous	-0.049***	0.004
Latin American	-0.028***	0.002
Middle Eastern	0.015**	0.007
African	-0.051***	0.002
South Asian	0.082***	0.007
East Asian	0.04***	0.004
Pacific Islander	0.033***	0.011
Occupation Self Employment Rate	-0.069***	0.008
Own	0.052***	0.001
Skill	0.067***	3.845e-04
Immigrant*Skill	0.041***	0.002
Constant	1.618***	0.007

N=918,914; Uncensored N=880,561; χ^2= 318,729; P-value=0.000, λ=-0.023***

Results from Heckman Correction estimation of Equation 2.23, with the addition of an interaction between skill and immigrants status.

Table. A2.9. Full Results for Self-Employment Selection Equation (with Skill*Immigrant Interaction)

Variable	Coefficient	Standard Error
Agriculture	0.661***	0.04
Mining	0.389***	0.06
Utilities	-0.16**	0.074
Construction	1.091***	0.027
Manufacturing	0.365***	0.028
Wholesale	0.796***	0.03
Retail	0.787***	0.027
Transportation	0.983***	0.028
Cultural	0.658***	0.032
Finance	0.617***	0.028
Real Estate	0.754***	0.03
Professional	1.143***	0.027
Management	0.303**	0.133
Support	1.149***	0.028
Education	0.328***	0.029
Health	0.719***	0.027
Arts	1.059***	0.031
Food	0.898***	0.029
Other	1.144***	0.028
Female	-0.152***	0.006
Married	0.126***	0.006
High School	0.004	0.011
Below Bachelor's	0.068***	0.011
Bachelor's	0.156***	0.012
Above Bachelor's	0.172***	0.013
Immigrant	-0.3***	0.051
Age	0.042***	0.002
Age2	-2.439e-04***	1.66e-05
Indigenous	-8.174e-04	0.017
Latin American	-0.129***	0.011
Middle Eastern	0.273***	0.025
African	-0.161***	0.013
South Asian	0.013	0.027
East Asian	-0.05***	0.016
Pacific Islander	-0.19***	0.056
Occupation Self Employment Rate	3.764***	0.021
Skill	-0.024***	0.002
Immigrant*Skill	0.057***	0.007
Own	0.099***	0.007
Housing Residual	0.277	0.186
Homeowner*Housing Residual	0.466**	0.222
Constant	-4.181***	0.046

N=918,915; Uncensored N=75,244; χ^2=9596; P-value=0.000

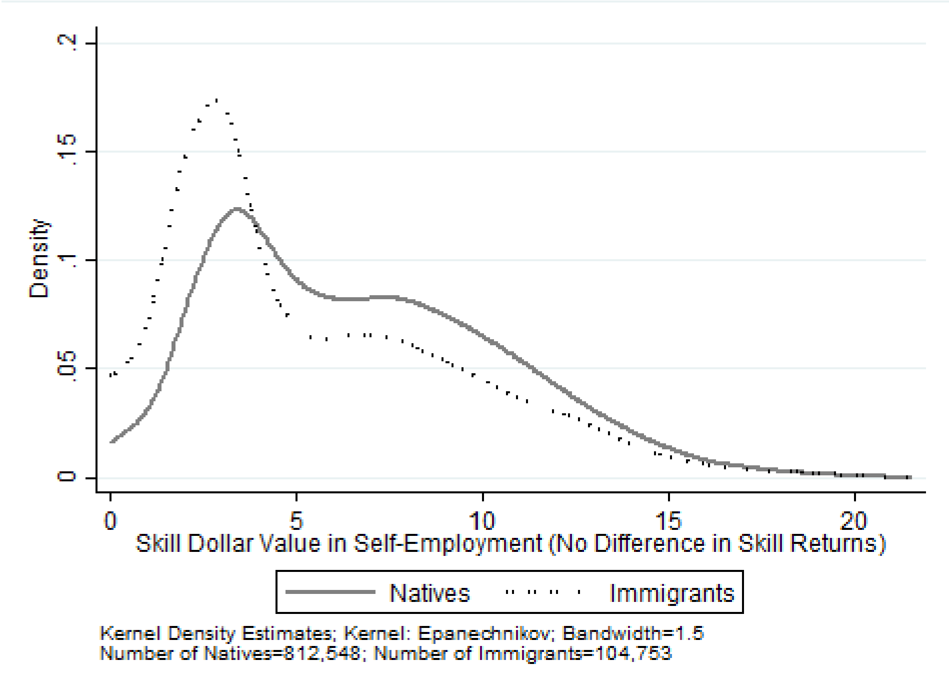
Results from Heckman Correction estimation of Equation 2.26, with the addition of an interaction between skill and immigrants status.

Table. A2.10. Full Results for Wage Employment Selection Equation (with Skill*Immigrant Interaction)

Variable	Coefficient	Standard Error
Agriculture	-0.488***	0.042
Mining	-0.265***	0.084
Utilities	0.179*	0.091
Construction	-0.981***	0.032
Manufacturing	-0.221***	0.035
Wholesale	-0.548***	0.037
Retail	-0.562***	0.033
Transportation	-0.835***	0.034
Cultural	-0.587***	0.04
Finance	-0.476***	0.035
Real Estate	-0.541***	0.036
Professional	-0.949***	0.032
Management	-0.34**	0.161
Support	-0.994***	0.034
Education	-0.354***	0.035
Health	-0.636***	0.032
Arts	-0.948***	0.036
Food	-0.536***	0.036
Other	-1.036***	0.033
Female	0.078***	0.008
Married	-0.028***	0.008
High School	0.036***	0.013
Below Bachelor's	0.027**	0.012
Bachelor's	0.012	0.014
Above Bachelor's	0.043**	0.017
Immigrant	0.23***	0.061
Age	-0.034***	0.002
Age2	1.939e-04***	1.95e-05
Indigenous	-0.039*	0.02
Latin American	0.036***	0.013
Middle Eastern	-0.116***	0.033
African	0.109***	0.015
South Asian	0.075**	0.035
East Asian	0.047**	0.019
Pacific Islander	0.074	0.068
Occupation Self Employment Rate	-3.169***	0.023
Skill	0.048***	0.002
Immigrant*Skill	-0.047***	0.009
Own	-0.018**	0.008
Housing Residual	-0.009	0.215
Homeowner*Housing Residual	0.098	0.264
Constant	3.751***	0.053

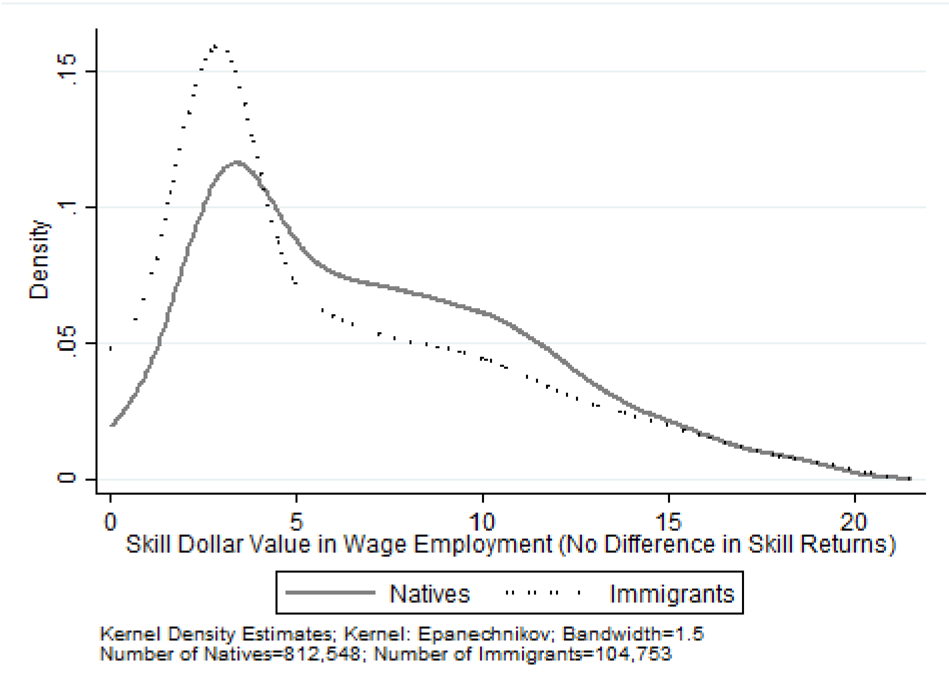
N=918,915; Uncensored N=880,560 χ^2=318,729; P-value=0.000, $\lambda = -0.0225$ Results from Heckman Correction estimation of Equation 2.25, with the addition of an interaction between skill and immigrants status.

Fig. A2.3. Density of Self-Employment Dollar Values



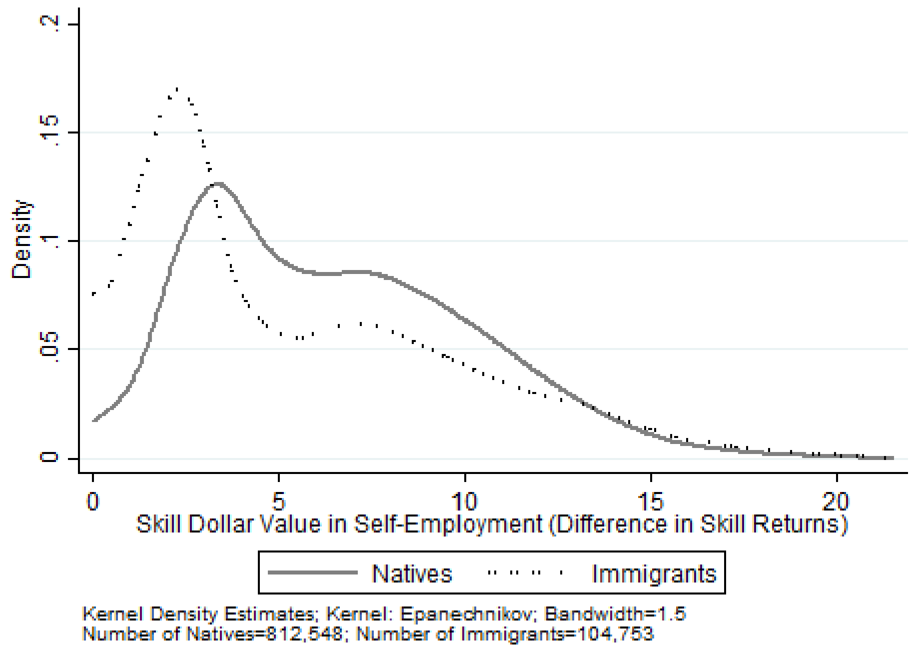
Kernel Density Estimation of the dollar value of skills in self-employment, separately for immigrants and natives, as calculated in Equation 2.30.

Fig. A2.4. Density of Wage Employment Dollar Values



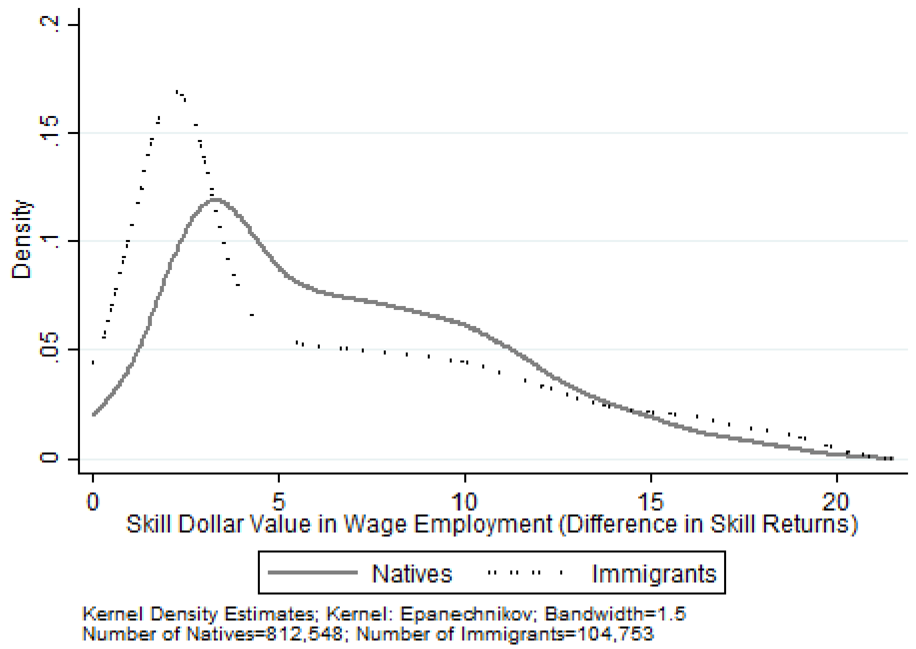
Kernel Density Estimation of the dollar value of skills in wage employment, separately for immigrants and natives, as calculated in Equation 2.30.

Fig. A2.5. Density of Self-Employment Dollar Values (With Skill/Immigrant Interaction)



Kernel Density Estimation of the dollar value of skills in self-employment, separately for immigrants and natives, as calculated in Equation 2.31.

Fig. A2.6. Density of Wage Employment Dollar Values (With Skill/Immigrant Interaction)



Kernel Density Estimation of the dollar value of skills in wage employment, separately for immigrants and natives, as calculated in Equation 2.31.

Table A2.11. Probit Coefficients with Dollar Values of Skills in Self and Wage Employment, Interacted with Adult Immigrant and Child Immigrant Status

Variable	Coe□cient	Standard Error
Agriculture	0.588***	0.039
Mining	0.413***	0.059
Utilities	-0.044	0.077
Construction	1.13***	0.027
Manufacturing	0.368***	0.028
Wholesale	0.758***	0.029
Retail	0.723***	0.027
Transportation	0.978***	0.027
Cultural	0.694***	0.031
Finance	0.534***	0.029
Real Estate	0.629***	0.031
Professional	1.161***	0.026
Management	-0.142	0.143
Support	1.114***	0.028
Education	0.226***	0.029
Health	0.62***	0.027
Arts	1.003***	0.03
Food	0.846***	0.028
Other	1.093***	0.027
Female	-0.164***	0.006
Married	0.135***	0.006
High School	0.017	0.011
Below Bachelor's	0.106***	0.011
Bachelor's	0.241***	0.013
Above Bachelor's	0.292***	0.015
Adult Immigrant	-0.003	0.017
Child Immigrant	0.103***	0.033
Age	0.054***	0.002
Age2	-3.522e-04***	1.91e-05
Indigenous	-0.008	0.017
Latin American	-0.112***	0.011
Middle Eastern	0.255***	0.024
African	-0.15***	0.013

South Asian	0.144***	0.026
East Asian	-0.011	0.015
Pacific Islander	-0.279***	0.055
Occupation Self-Employment Rate	3.874***	0.024
Homeowners	0.114***	0.007
$SkillValue_{i,d',se}$	0.076***	0.009
$SkillValue_{i,d',we}$	-0.082***	0.008
Adult Immigrant*$SkillValue_{i,d',se}$	0.144***	0.013
Adult Immigrant*$SkillValue_{i,d',we}$	-0.112***	0.011
Child Immigrant*$SkillValue_{i,d',se}$	0.115***	0.021
Child Immigrant*$SkillValue_{i,d',we}$	-0.099***	0.018
Constant	-4.557***	0.055

Number of Observations=981,370, Pseudo R^2=0.269, χ^2=70724

Results from estimation of Equation 2.36.

Table A2.12. Probit Coefficients with Dollar Values of Skills in Self and Wage Employment, With Housing Value Interactions

Variable	Coe□cient	Standard Error
Agriculture	0.746***	0.046
Mining	0.472***	0.066
Utilities	-0.014	0.075
Construction	1.254***	0.031
Manufacturing	0.43***	0.032
Wholesale	0.773***	0.033
Retail	0.758***	0.031
Transportation	0.99***	0.032
Cultural	0.622***	0.038
Finance	0.502***	0.034
Real Estate	0.59***	0.037
Professional	1.24***	0.03
Management	-0.313**	0.143
Support	1.188***	0.033
Education	0.15***	0.034
Health	0.614***	0.032
Arts	0.941***	0.037
Food	1.002***	0.033
Other	1.123***	0.032
Female	-0.179***	0.007
Married	0.135***	0.008
High School	0.008	0.014
Below Bachelor's	0.092***	0.014
Bachelor's	0.243***	0.016
Above Bachelor's	0.32***	0.019
Immigrant	0.04	0.027
Age	0.068***	0.002
Age2	-4.74e-04***	2.4e-05
Indigenous	0.018	0.021
Latin American	-0.103***	0.014
Middle Eastern	0.254***	0.03
African	-0.128***	0.018
South Asian	0.213***	0.032
East Asian	-0.04**	0.019

Pacific Islander	-0.377***	0.073
Occupation Self-Employment Rate	4.079***	0.031
Homeowner 2+	0.054***	0.011
Housing Value	2.9e-07***	1.97e-08
Immigrant*Housing Value	-5.15e-08	4.36e-08
$SkillValue_{i,d^{\square},se}$	0.162***	0.012
$SkillValue_{i,d^{\square},we}$	-0.16***	0.01
Housing Value*$SkillValue_{i,d^{\square},se}$	-1.78e-08*	1.06e-08
Housing Value*$SkillValue_{i,d^{\square},se}$	1.62e-08*	8.85e-09
Immigrant*$SkillValue_{i,d^{\square},we}$	0.16***	0.021
Immigrant*$SkillValue_{i,d^{\square},we}$	-0.122***	0.018
Immigrant*Housing Value*$SkillValue_{i,d^{\square},se}$	-5.02e-08*	2.78e-08
Immigrant*Housing Value*$SkillValue_{i,d^{\square},we}$	3.76e-08*	2.27e-08
Constant	-5.029***	0.072

N=648,120; χ^2=52722; P-value=0.000, $Pseudo\ R^2$=0.2891

Probit coefficients as estimated from Equation 2.35, with the addition of housing value interactions as described in Section 2.7.2.

Note that since this entire subsample consists of homeowners, this dummy variable for homeowner indicates whether the individual has been a homeowners for at least 2 years, consist with the use of this variable in other models.

Table A2.13. Probit Coefficients with Dollar Values of Skills in Self and Wage Employment, Estimating Immigrants from Mexico Separately (Including Interactions with Skill Values)

Variable	Coefficient	Standard Error
Agriculture	0.62***	0.04
Mining	0.444***	0.06
Utilities	-0.08	0.074
Construction	1.169***	0.027
Manufacturing	0.381***	0.028
Wholesale	0.773***	0.03
Retail	0.727***	0.027
Transportation	0.987***	0.028
Cultural	0.719***	0.032
Finance	0.538***	0.03
Real Estate	0.622***	0.032
Professional	1.185***	0.027
Management	-0.075	0.142
Support	1.138***	0.028
Education	0.232***	0.03
Health	0.633***	0.028
Arts	1.022***	0.031
Food	0.858***	0.029
Other	1.104***	0.028
Female	-0.164***	0.006
Married	0.135***	0.006
High School	0.008	0.011
Below Bachelor's	0.099***	0.011
Bachelor's	0.243***	0.013
Above Bachelor's	0.3***	0.016
Immigrant	0.07***	0.018
Mexico	-0.337***	0.037
Age	0.056***	0.002
Age2	-3.737e-04***	1.96e-05
Indigenous	-0.008	0.017
Latin American	-0.089***	0.011
Middle Eastern	0.26***	0.025
African	-0.145***	0.013

South Asian	0.11***	0.028
East Asian	-0.03*	0.016
Pacific Islander	-0.268***	0.056
Occupation Self-Employment Rate	3.914***	0.026
Homeowner	0.119***	0.007
$SkillValue_{i,d^*,se}$	0.092***	0.009
$SkillValue_{i,d^*,we}$	-0.096***	0.008
Immigrant*$SkillValue_{i,d^*,se}$	0.1***	0.014
Immigrant*$SkillValue_{i,d^*,we}$	-0.079***	0.012
Mexico*$SkillValue_{i,d^*,se}$	0.319***	0.043
Mexico*$SkillValue_{i,d^*,we}$	-0.245***	0.039
Constant	-4.64***	0.057

N=922,060; χ^2=67014; P-value=0.000, $Pseudo\ R^2$=0.2738

Probit coe□cients as estimated from Equation 2.35, with a variable added to indicating whether an immigrant is from Mexico.

Table. A2.14. Estimated Marginal Effects with Immigrants from Mexico Included Separately From Other Immigrants

	(1)	(2)	(3)
Immigrants From Mexico	0.061	0.059	0.047
	(1.76e-03)	(1.72e-03)	(2.5e-03)
All Other Immigrants	0.088	0.088	0.082
	(1.13e-04)	(1.15e-03)	(1.74e-03)
Natives	0.073	0.073	0.074
	(3.42e-04)	(3.4e-04)	(4.34e-04)

(1)- Demographic and Industry Controls Only.
(2)- Controlling for Dollar Value of Skills, Same Skill Price For All Subsamples.
(3)- Controlling for Dollar Values of Skills, Same Skill Price For All Subsamples, Interactions with Immigrants and Mexico Variables.

Chapter 3 Appendix

Table A3.1. Eigenvectors From Principal Component Analysis

Skill Type	Comp1occ	Comp2occ	Comp3occ	Comp4occ	Comp5occ
Oral Comprehension I	0.0886	0.0009	0.0834	-0.0163	-0.0151
Written Comprehension I	0.1004	0.014	0.0439	0.0044	-0.0442
Oral Expression I	0.0811	-0.0078	0.1134	-0.0286	0.0036
Written Expression I	0.1008	0.0108	0.0598	0.0058	-0.0281
Fluency of Ideas I	0.079	0.0468	0.0825	-0.0092	-0.0469
Originality I	0.0732	0.0483	0.0887	0.0047	-0.0433
Problem Sensitivity I	0.0595	0.0976	0.0428	-0.048	-0.0013
Deductive Reasoning I	0.0911	0.0698	0.0134	-0.0547	-0.0244
Inductive Reasoning I	0.0852	0.0693	0.0104	-0.045	-0.0311
Information Ordering I	0.0733	0.0756	-0.0283	-0.0161	-0.061
Category Flexibility I	0.0771	0.0756	-0.007	0.0263	-0.0545
Mathematical Reasoning I	0.066	0.0245	-0.1274	0.0028	-0.0345
Number Facility I	0.0577	0.0214	-0.141	-0.0014	-0.025
Memorization I	0.062	0.0096	0.0734	0.0026	-0.0722
Speed of Closure I	0.0343	0.0779	0.0449	-0.0409	-0.0055
Flexibility of Closure I	0.0216	0.1174	0.0058	-0.0018	-0.06
Perceptual Speed I	-0.003	0.0992	-0.0494	0.0096	-0.0395
Spatial Orientation I	-0.0622	0.0997	-0.0019	-0.0264	0.0981
Visualization I	-0.0094	0.1268	0.0062	0.02	-0.0787
Selective Attention I	0.0063	0.0637	0.0474	0.021	-0.053
Time Sharing I	0.0162	0.0533	0.1318	-0.016	-0.011
Arm Hand Steadiness I	-0.0936	0.0656	0.0058	0.0191	-0.0214
Manual Dexterity I	-0.0941	0.0598	-0.0125	0.0085	-0.0119
Finger Dexterity I	-0.0732	0.0585	-0.0105	0.0463	-0.0495
Control Precision I	-0.0838	0.0818	-0.033	-0.0113	0.004
Multi-Limb Coordination I	-0.091	0.0785	0.0289	-0.0064	0.0424
Response Orientation I	-0.0811	0.0817	0.0399	-0.0475	0.0393
Rate Control I	-0.0879	0.0845	-0.0034	-0.0309	0.0255
Reaction Time I	-0.0805	0.0896	0.0259	-0.0501	0.0286
Wrist Finger SpeedI	-0.0771	0.0566	-0.0118	-0.0241	-0.0543
Speed of Limb Movement I	-0.0892	0.0574	0.0359	-0.0056	0.0667
Static Strength I	-0.0894	0.0554	0.0608	-0.0048	0.0533
Dynamic Strength I	-0.0888	0.0556	0.0457	-0.0205	0.054

Trunk Strength I	-0.0814	0.0377	0.0923	0.0113	0.0128
Stamina I	-0.0849	0.0457	0.0816	-0.0009	0.0528
Extent Flexibility I	-0.0945	0.0569	0.0376	-0.0042	0.035
Dynamic Flexibility I	-0.0561	0.0215	0.0054	-0.0226	0.0576
Gross Body Coordination I	-0.086	0.0496	0.0769	-0.0059	0.0566
Gross Body Equilibrium I	-0.0842	0.0636	0.0652	-0.0237	0.0603
Near Vision I	0.0383	0.06	-0.0663	0.0428	-0.0286
Far Vision I	-0.0185	0.1201	0.0606	0.0004	0.0169
Visual Color Discrimination I	-0.0513	0.1109	0.0127	0.0525	-0.0657
Night Vision I	-0.0647	0.0854	0.001	-0.0501	0.1041
Peripheral Vision I	-0.0688	0.0824	0.005	-0.0521	0.1085
Depth Perception I	-0.0559	0.1209	-0.0146	-0.0247	0.0315
Glare Sensitivity I	-0.0667	0.0882	-0.0039	-0.0497	0.1015
Hearing Sensitivity I	-0.0573	0.0775	0.0623	0.0075	-0.0352
Auditory Attention I	-0.0558	0.0852	0.0748	-0.0062	0.0072
Sound Localization I	-0.0724	0.0822	0.0101	-0.0401	0.0928
Speech Recognition I	0.0698	-0.0112	0.1059	-0.0228	0.0428
Speech Clarity I	0.0747	-0.0109	0.1261	-0.0176	0.0446
Reading Comprehension I	0.102	0.0217	0.0372	-0.0101	-0.0403
Active Listening I	0.0905	-0.0021	0.0819	-0.0519	0.001
Writing I	0.1008	0.0153	0.0477	-0.0055	-0.0296
Speaking I	0.0856	-0.0025	0.1125	-0.0401	0.0147
Mathematics I	0.062	0.0289	-0.1265	0.0079	-0.0388
Science I	0.0159	0.0963	-0.0037	0.0665	-0.091
Critical Thinking I	0.0892	0.0685	0.0223	-0.0615	-0.0055
Active Learning I	0.0936	0.0407	0.064	-0.0226	-0.0104
Learning Strategies I	0.0514	0.0186	0.1219	0.0084	-0.0533
Monitoring I	0.0472	0.0607	0.1053	-0.0291	0.0096
Social Perceptiveness I	0.0484	-0.0064	0.1499	-0.0584	0.0529
Coordination I	0.0553	0.0401	0.0984	-0.0741	0.0688
Persuasion I	0.0684	0.0099	0.0594	-0.0931	0.0944
Negotiation I	0.0703	0.0085	0.0623	-0.099	0.0829
Instructing I	0.0426	0.0162	0.1296	0.0325	-0.0423
Service Orientation I	0.0353	-0.0268	0.126	-0.0098	0.0558
Complex Problem Solving I	0.0884	0.0771	0.0065	-0.0446	-0.0184
Operations Analysis I	0.0595	0.0716	-0.0526	-0.0232	-0.0246
Technology Design I	0.0268	0.0946	-0.0286	0.0059	-0.1548

Equipment Selection I	-0.0559	0.0931	-0.0289	0.0002	-0.1331
Installation I	-0.041	0.0635	-0.0436	-0.0078	-0.0857
Programming I	0.0423	0.0477	-0.0724	-0.0016	-0.1176
Operation Monitoring I	-0.0605	0.1178	-0.0367	-0.0136	-0.059
Operation And Control I	-0.0781	0.0999	-0.0129	-0.03	-0.0024
Equipment Maintanence I	-0.064	0.077	-0.0393	-0.0166	-0.0943
Troubleshooting I	-0.0614	0.1088	-0.0198	0.0078	-0.0876
Repairing I	-0.0587	0.0761	-0.043	-0.0193	-0.105
Quality Control I	-0.0412	0.1128	-0.0086	0.025	-0.0775
Judgment and Decision Making I	0.0914	0.0609	0.0333	-0.0637	0.0118
Systems Analysis I	0.0889	0.0802	-0.0136	-0.045	-0.0336
Systems Evaluation I	0.0872	0.0734	0.0202	-0.0463	-0.034
Time Management I	0.067	0.039	0.0563	-0.0569	0.0565
Management of Financial Resources I	0.0544	0.0492	-0.0958	-0.0751	0.0867
Management of Material Resources I	0.0348	0.0851	-0.0363	-0.0325	0.0668
Management of Personnel Resources I	0.0626	0.0729	0.0355	-0.0612	0.0735
Administration and Management I	0.0584	0.0424	-0.0484	-0.0656	0.135
Clerical I	-0.0013	0.013	0.0506	0.1266	0.0617
Economics and Accounting I	0.0381	0.0335	-0.097	0.0395	0.1195
Sales and Marketing I	0.0654	0.0332	-0.0046	-0.0421	0.1013
Customer and Personal Service I	0.0218	0.0084	0.0287	0.1213	0.1016
Personnel and Human Resources I	0.0198	0.0417	0.0557	0.1783	0.0009
Production and Processing I	0.0206	0.0239	0.0951	0.1659	-0.0145
Food Production I	0.0446	0.0307	-0.1207	0.056	0.0642
Computers and Electronics I	0.022	-0.0006	0.024	0.1174	0.0783
Engineering and Technology I	0.046	0.0451	-0.1283	0.0355	0.0857
Design I	0.0279	0.0312	0.0012	0.117	0.1009
Building and Construction I	0.07	0.0568	-0.0732	-0.0293	0.0817
Mechanical I	0.0092	0.0546	0.0304	0.2085	0.0005
Mathematics Knowledge I	0.0392	0.0449	-0.0159	0.1601	0.0491
Physics I	0.0281	0.029	0.1021	0.1583	-0.0419
Chemistry I	0.0523	0.0643	-0.037	0.034	0.0977
Biology I	0.0756	0.0493	-0.0709	-0.0708	0.1004
Psychology I	0.0268	0.0054	0.136	0.1265	-0.0038
Sociology and AnthropologyI	0.0645	0.0295	0.0025	-0.0494	0.0989
Geography I	0.0357	0.0393	-0.0915	0.1101	0.0664
Medicine and Dentistry I	0.0112	0.0523	0.0478	0.1958	-0.0014

Therapy and Counselling I	0.0628	0.0154	0.0793	0.0466	0.0418
Education and Training I	0.0458	0.0366	-0.1225	0.057	0.1256
Foreign Language I	0.0343	0.0436	-0.123	0.084	0.0825
Fine Arts I	0.0448	0.0343	-0.1243	0.0604	0.0657
History and Archaeology I	0.0397	0.0413	-0.1091	0.1013	0.0702
Philosophy and Theology I	0.0213	0.0407	0.0603	0.1683	-0.0211
Public Safety I	0.0137	0.0641	0.0812	0.033	0.0883
Law and Government I	0.0407	0.0387	-0.072	0.1377	0.0494
Telecommunications I	0.0249	0.0555	0.0389	0.0452	0.0538
Communications and Media I	0.0221	0.0072	0.0251	0.1193	0.0758
Transportation I	-0.0282	0.0744	0.0138	-0.0151	0.097
Oral Comprehension L	0.0887	0.0475	0.0316	-0.0015	-0.0757
Written Comprehension L	0.098	0.0406	0.002	-0.0054	-0.0675
Oral Expression L	0.0971	0.0294	0.0348	-0.0121	-0.0647
Written Expression L	0.1014	0.0292	0.0432	-0.0091	-0.0488
Fluency of Ideas L	0.0914	0.0648	0.0345	-0.0275	-0.0355
Originality L	0.0851	0.0668	0.0609	-0.0087	-0.0489
Problem Sensitivity L	0.0799	0.0847	0.0199	-0.0375	-0.0274
Deductive Reasoning L	0.0943	0.067	-0.0228	-0.0345	-0.0203
Inductive Reasoning L	0.0849	0.0734	0.0003	-0.026	-0.0458
Information Ordering L	0.0758	0.0897	-0.0498	-0.0194	-0.0706
Category Flexibility L	0.0739	0.079	-0.0485	0.0313	-0.0522
Mathematical Reasoning L	0.0766	0.0476	-0.1131	-0.0008	-0.0344
Number Facility L	0.0715	0.0353	-0.126	-0.0053	-0.017
Memorization L	0.0798	0.0312	0.0576	-0.0215	-0.0547
Speed of Closure L	0.059	0.0906	0.0166	-0.0239	-0.0195
Flexibility of Closure L	0.0307	0.1128	-0.0387	-0.014	-0.044
Perceptual Speed L	0.0022	0.1131	-0.0535	-0.014	-0.053
Spatial Orientation L	-0.0604	0.1026	-0.0013	-0.02	0.0938
Visualization L	-0.0005	0.1316	-0.0135	0.0123	-0.0744
Selective Attention L	0.0008	0.1053	0.0382	-0.0062	-0.0596
Time Sharing L	0.0042	0.0794	0.1165	-0.0212	0.0097
Arm Hand Steadiness L	-0.0886	0.0695	0.0085	0.0301	-0.0323
Manual Dexterity L	-0.0898	0.0668	-0.009	0.0145	-0.0211
Finger Dexterity L	-0.0671	0.0759	-0.0151	0.0264	-0.0569
Control Precision L	-0.0791	0.0879	-0.0312	-0.0036	-0.0051
Multi-Limb Coordination L	-0.0891	0.0814	0.03	0.0019	0.0423

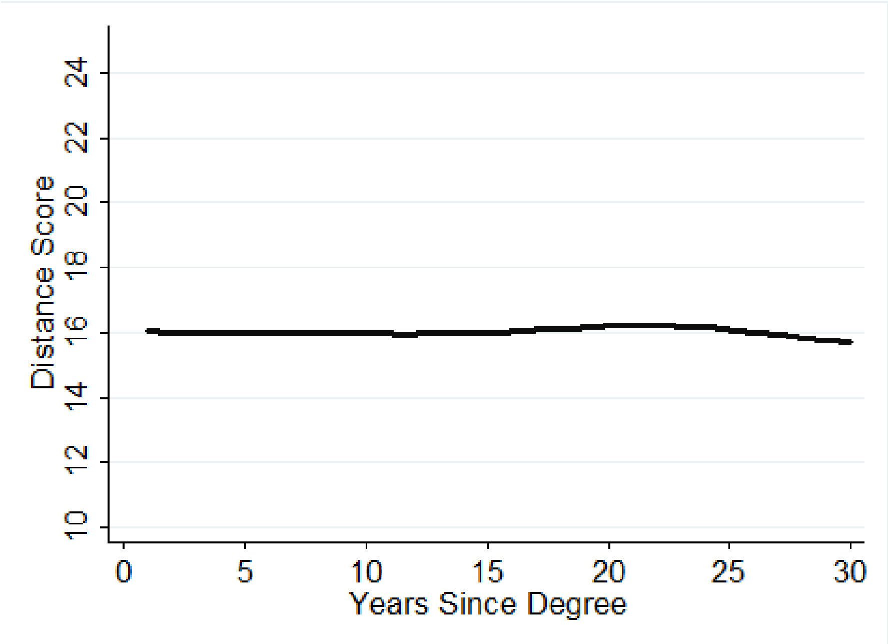
Response Orientation L	-0.0808	0.0815	0.033	-0.0444	0.046
Rate Control L	-0.0855	0.088	-0.006	-0.0318	0.0314
Reaction Time L	-0.0778	0.0909	0.0234	-0.0487	0.0326
Wrist Finger Spee L	-0.076	0.0544	-0.0181	-0.025	-0.0593
Speed of Limb Movement L	-0.0882	0.0613	0.0345	-0.0138	0.0746
Static Strength L	-0.0882	0.0571	0.0648	-0.0129	0.0572
Explosive Strength L	-0.033	0.046	0.0683	-0.0646	0.0973
Dynamic Strength L	-0.0881	0.0599	0.0468	-0.0222	0.058
Trunk Strength L	-0.0802	0.041	0.0941	0.0103	0.017
Stamina L	-0.0833	0.0514	0.0792	-0.0073	0.0575
Extent Flexibility L	-0.0948	0.0559	0.0363	-0.009	0.0295
Dynamic Flexibility L	-0.053	0.0205	0.0083	-0.0211	0.0559
Gross Body Coordination L	-0.0854	0.0522	0.0755	-0.0099	0.0554
Gross Body Equilibrium L	-0.0827	0.0662	0.0631	-0.0285	0.0635
Near Vision L	0.0626	0.0307	-0.0613	-0.0113	-0.0006
Far Vision L	0.0039	0.1119	0.0602	0.0018	-0.0082
Visual Color Discrimination L	-0.0402	0.114	-0.0087	0.0327	-0.0704
Night Vision L	-0.0621	0.086	0.0025	-0.0459	0.1073
Peripheral Vision L	-0.0658	0.0864	0.0101	-0.054	0.1106
Depth Perception L	-0.0547	0.1243	-0.0085	-0.0174	0.0168
Glare Sensitivity L	-0.0671	0.0878	-0.0046	-0.0488	0.1026
Hearing Sensitivity L	-0.0541	0.0911	0.0413	0.0028	-0.0392
Auditory Attention L	-0.0541	0.1004	0.0589	-0.0173	-0.0008
Sound Localization L	-0.0718	0.0833	0.0091	-0.0352	0.0967
Speech Recognition L	0.0808	-0.0066	0.0682	-0.0283	0.0272
Speech Clarity L	0.0753	-0.0092	0.1053	-0.0035	-0.0019
Reading Comprehension L	0.095	0.0451	0.0249	0.0162	-0.0579
Active Listening L	0.0972	0.0161	0.0513	-0.0264	-0.0082
Writing L	0.1003	0.0311	0.0416	-0.0081	-0.034
Speaking L	0.0981	0.0209	0.0513	-0.0433	0.0064
Mathematics L	0.0661	0.0536	-0.1174	0.0128	-0.0507
Science L	0.0182	0.0947	-0.0063	0.0641	-0.1016
Critical Thinking L	0.0951	0.0589	-0.0043	-0.0361	-0.0016
Active Learning L	0.1019	0.0561	0.0177	-0.0237	-0.0343
Learning Strategies L	0.0664	0.0324	0.1133	0.0114	-0.0486
Monitoring L	0.0851	0.0693	0.0493	-0.0488	0.0042
Social Perceptiveness L	0.0683	0.0137	0.1226	-0.0655	0.032

Coordination L	0.0699	0.0716	0.0655	-0.0695	0.0503
Persuasion L	0.0831	0.0271	0.0332	-0.0971	0.0751
Negotiation L	0.0808	0.0178	0.045	-0.0958	0.0707
Instructing L	0.0759	0.0451	0.0784	-0.0061	-0.0508
Service Orientation L	0.0599	-0.0183	0.1002	-0.045	0.043
Complex Problem Solving L	0.0874	0.0834	-0.0267	-0.0471	-0.0252
Operations Analysis L	0.0637	0.0717	-0.0587	-0.031	-0.0199
Technology Design L	0.019	0.0946	-0.0256	0.0293	-0.1586
Equipment Selection L	-0.047	0.0974	-0.0332	-0.0013	-0.1454
Installation L	-0.039	0.0682	-0.0467	-0.0045	-0.0959
Programming L	0.0459	0.0539	-0.0832	0.0126	-0.1262
Operation Monitoring L	-0.0504	0.1243	-0.0336	-0.0041	-0.0716
Operation And Control L	-0.072	0.1067	-0.012	-0.0298	-0.011
Equipment Maintanence L	-0.061	0.0808	-0.0434	-0.0153	-0.0991
Troubleshooting L	-0.0555	0.1109	-0.0221	0.0074	-0.104
Repairing L	-0.0558	0.081	-0.0449	-0.0166	-0.1086
Quality Control L	-0.0327	0.1217	-0.0235	0.0215	-0.0713
Judgment and Decision Making L	0.0974	0.0626	-0.0068	-0.0481	0.0045
Systems Analysis L	0.0915	0.0702	-0.0385	-0.0447	-0.019
Systems Evaluation L	0.0948	0.0685	-0.0069	-0.0562	-0.0287
Time Management L	0.0819	0.0654	0.0347	-0.0631	0.0358
Management of Financial Resources L	0.0596	0.0519	-0.09	-0.0672	0.0794
Management of Material Resources L	0.0467	0.0859	-0.0371	-0.037	0.0487
Management of Personnel Resources L	0.0744	0.0752	0.0165	-0.06	0.0527
Administration and Management L	0.0754	0.0484	-0.0739	-0.0686	0.1015
Clerical L	0.0222	0.0101	0.0256	0.1196	0.0756
Economics and Accounting L	0.0506	0.0327	-0.122	0.0527	0.1171
Sales and Marketing L	0.0612	0.0373	-0.0038	-0.0411	0.1065
Customer and Personal Service L	0.0195	0.0244	0.0291	0.115	0.0799
Personnel and Human Resources L	0.0171	0.0459	0.0565	0.1818	-0.0149
Production and Processing L	0.0225	0.0274	0.0986	0.1552	-0.0034
Food Production L	0.0406	0.0359	-0.1091	0.0481	0.0678
Computers and Electronics L	0.0212	0.0034	0.0305	0.1138	0.0855
Engineering and Technology L	0.0419	0.0486	-0.1259	0.0529	0.0708
Design L	0.029	0.0276	-0.0052	0.1019	0.1162
Building and Construction L	0.0663	0.0603	-0.0712	-0.0158	0.0775
Mechanical L	0.0126	0.0547	0.0453	0.1973	-0.0035

Mathematics Knowledge L	0.0162	0.0502	0.0205	0.1977	0.0365
Physics L	0.0203	0.0251	0.0978	0.1636	-0.0094
Chemistry L	0.0457	0.0622	-0.0196	0.0513	0.0997
Biology L	0.0763	0.0533	-0.0697	-0.0656	0.0954
Psychology L	0.0336	0.0167	0.1126	0.1436	0.0031
Sociology and Anthropology L	0.0626	0.0334	-0.0013	-0.0402	0.1042
Geography L	0.0134	0.0453	0.0364	0.2072	0.0138
Medicine and Dentistry L	0.0629	0.016	0.0725	0.047	0.0468
Therapy and Counselling L	0.0524	0.0391	-0.1247	0.0198	0.0989
Education and Training L	0.0449	0.0336	-0.127	0.0627	0.0611
Foreign Language L	0.0347	0.0444	-0.1167	0.0919	0.0854
Fine Arts L	0.0448	0.0309	-0.1202	0.055	0.0708
History and Archaeology L	0.0406	0.0379	-0.0773	0.1297	0.0524
Philosophy and Theology L	0.0247	0.0288	0.0988	0.1647	-0.0152
Public Safety L	0.0259	0.068	0.0761	0.0336	0.1016
Law and Government L	0.0408	0.044	-0.0685	0.139	0.0481
Telecommunications L	0.0296	0.0626	0.0161	0.0403	0.0529
Communications and Media L	0.02	-0.0013	0.0276	0.1202	0.0828
Transportation L	-0.0183	0.0903	0.003	-0.0097	0.0843

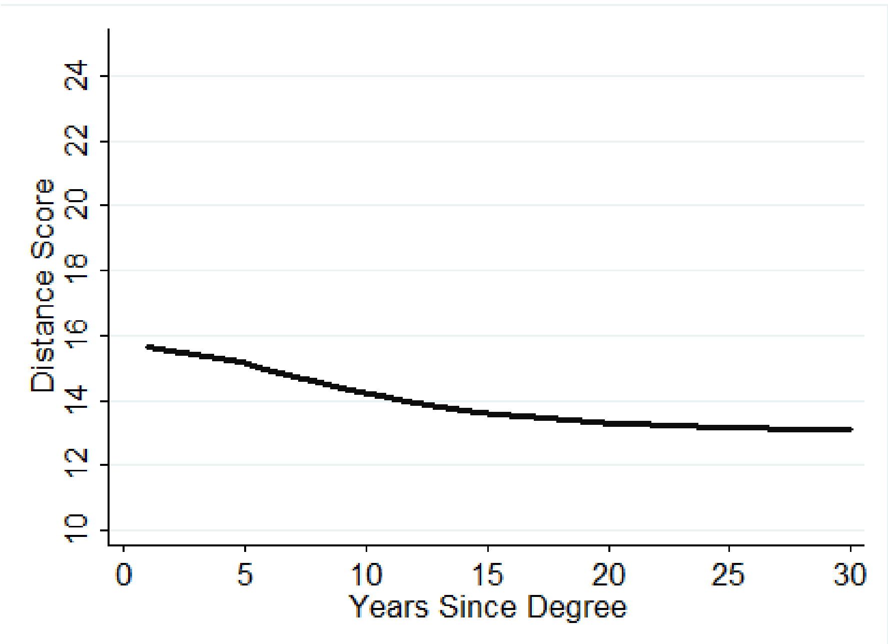
Eigenvectors for each of the top 5 Principal Components that Result From the Principal Component Analysis as described in Section 3.2.1. An "I" following the individual skill indicates that it is the importance value, while "L" indicates that it is the level value.

Fig. A3.1. Distance Scores for Education Graduates Over the Lifecycle



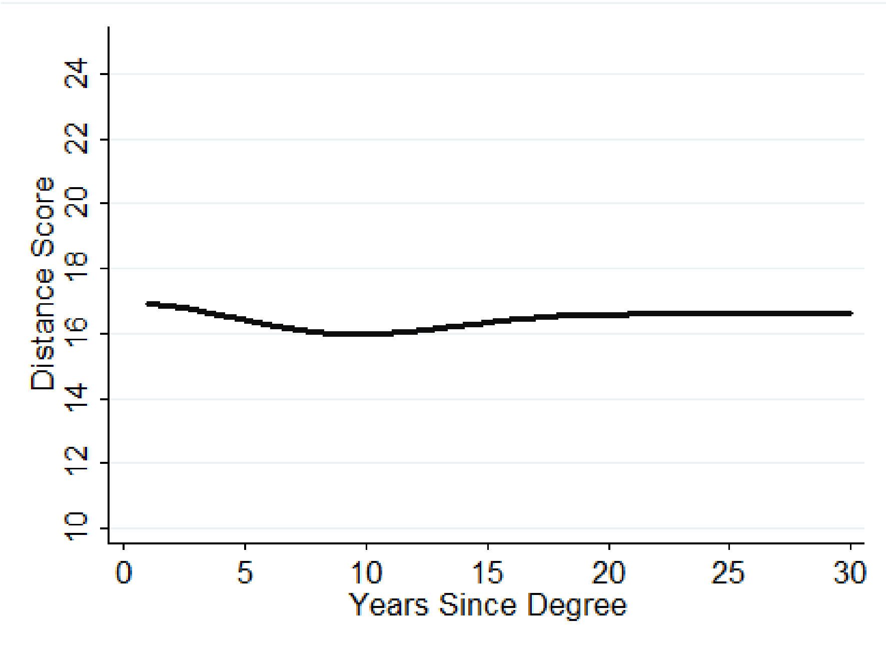
Lifecycle estimates of distance for education graduates only, calculated using Equation 3.5.

Fig. A3.2. Distance Scores for Business Graduates Over the Lifecycle



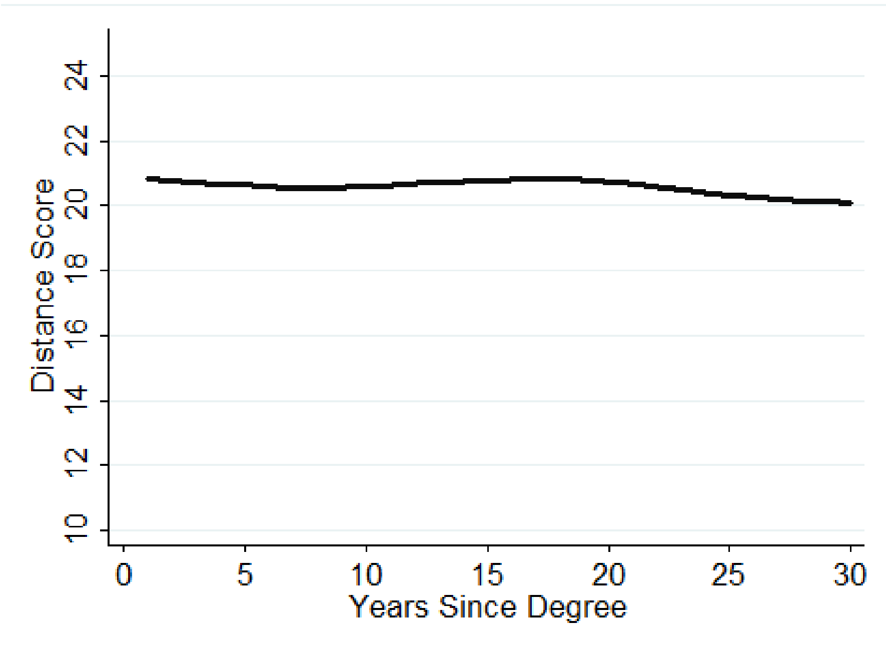
Lifecycle estimates of distance for business graduates only, calculated using Equation 3.5.

Fig. A3.3. Distance Scores for Social Science Graduates Over the Lifecycle



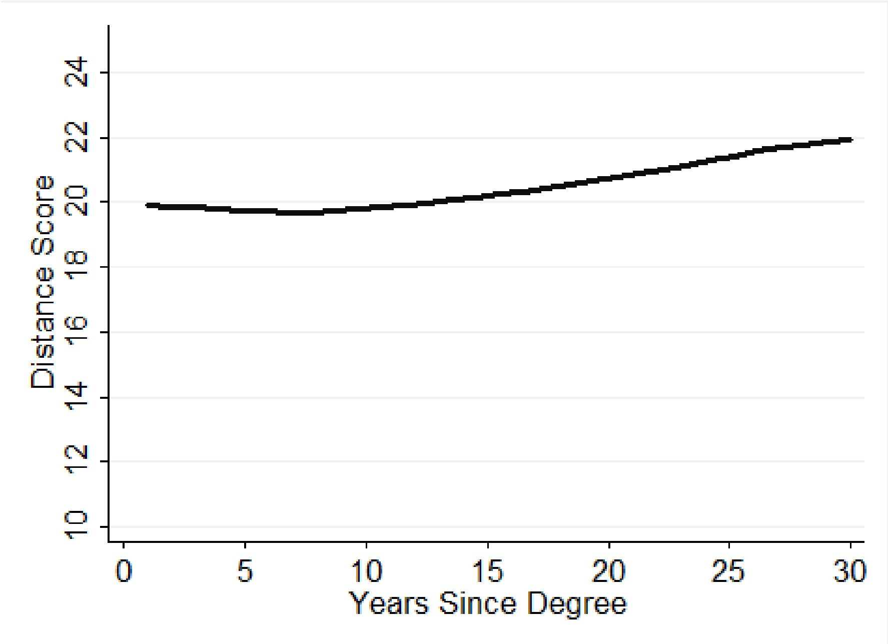
Lifecycle estimates of distance for social science graduates only, calculated using Equation 3.5.

Fig. A3.4. Distance Scores for Science Graduates Over the Lifecycle



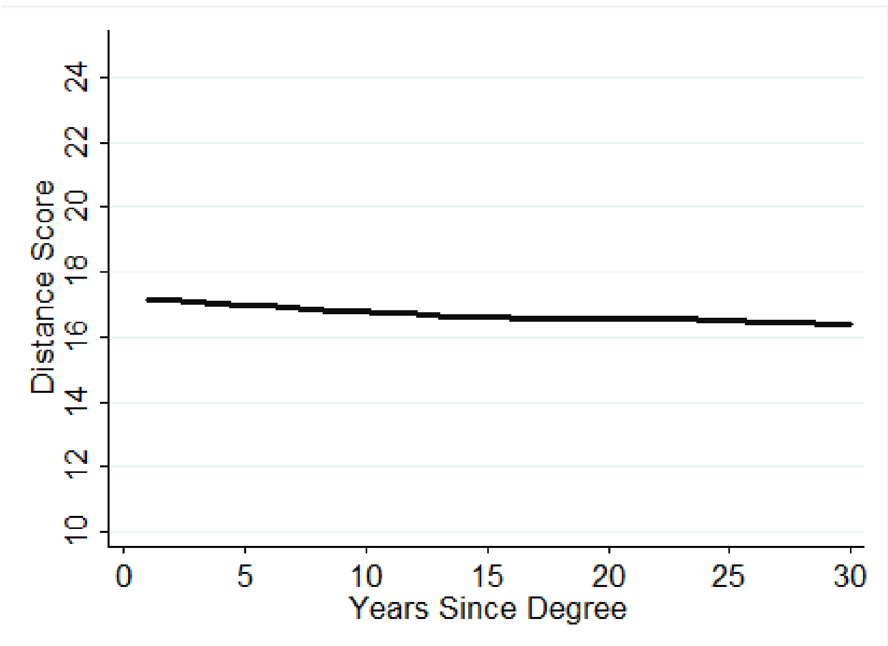
Lifecycle estimates of distance for science graduates only, calculated using Equation 3.5.

Fig. A3.5. Distance Scores for Math and Computer Science Graduates Over the Lifecycle



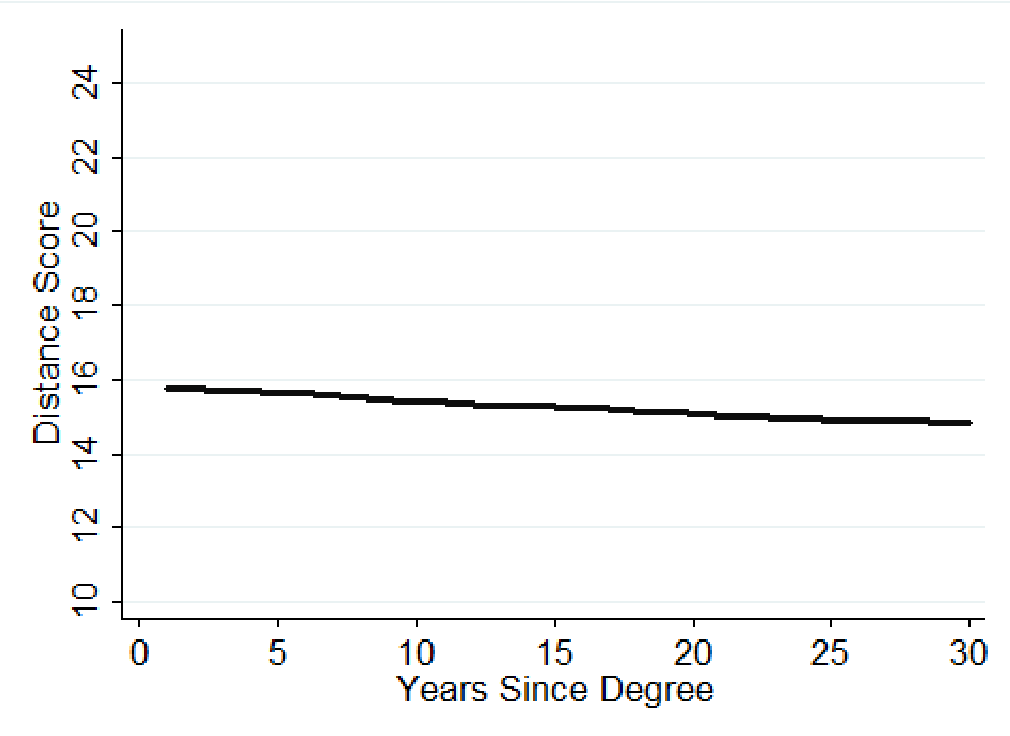
Lifecycle estimates of distance for math and computer science graduates only, calculated using Equation 3.5.

Fig. A3.6. Distance Scores for Agriculture Graduates Over the Lifecycle



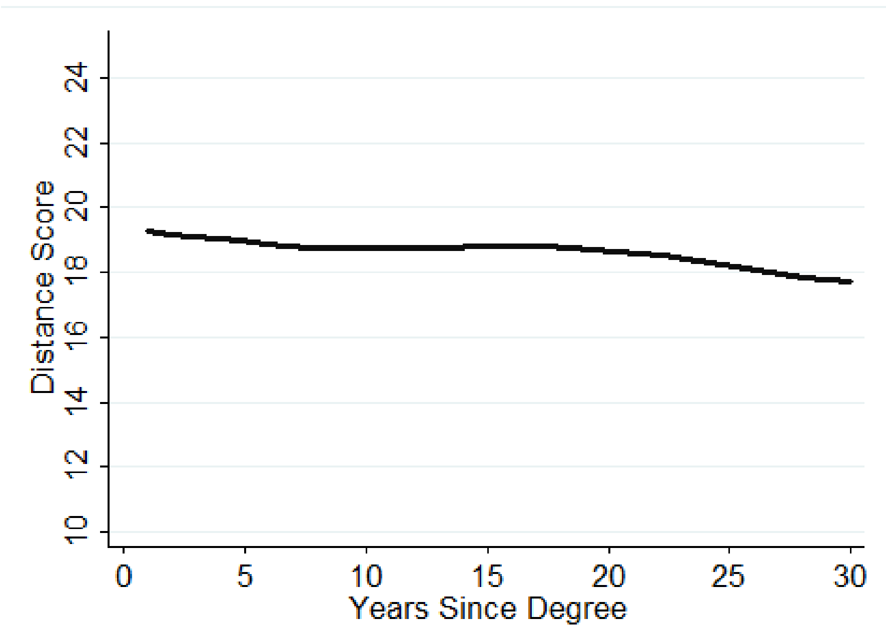
Lifecycle estimates of distance for agriculture graduates only, calculated using Equation 3.5.

Fig. A3.7. Distance Scores for Health Graduates Over the Lifecycle



Lifecycle estimates of distance for health graduates only, calculated using Equation 3.5.

Fig. A3.8. Distance Scores for Architecture and Engineering Graduates Over the Lifecycle



Lifecycle estimates of distance for architecture and engineering graduates only, calculated using Equation 3.5.

Table. A3.2. Full Results for Column (1) in Table 3.5

	Coefficient	Standard Error
Union	0.071***	0.017
Married	0.02	0.016
British Columbia	0.044	0.052
Alberta	0.055	0.046
Manitoba	0.086	0.07
Saskatchewan	-0.023	0.063
Quebec	0.108*	0.064
New Brunswick	-0.174**	0.078
Nova Scotia	-0.024	0.048
PEI	-0.16***	0.062
Newfoundland	-0.158*	0.081
2000	0.053***	0.012
2001	0.087***	0.014
2002	0.162***	0.014
2003	0.204***	0.014
2004	0.247***	0.014
2005	0.309***	0.016
2006	0.365***	0.016
2007	0.415***	0.016
Comp1occ	0.012***	0.002
Comp2occ	0.011***	0.003
Comp3occ	-0.002	0.003
Comp4occ	-0.001	0.003
Comp5occ	-0.001	0.004
Distance	-0.002*	0.001
Constant	2.297***	0.103

$Within\ R^2 = 0.19$, $Between\ R^2 = 0.176$, $Overall\ R^2 = 0.172$

$N = 24,170$, $\#\ of\ groups = 6830$, $\sigma_\alpha = 0.2$, $\sigma_u = 0.426$

Results estimated from Fixed Effects Model in Equation 3.7

Table. A3.3. Full Results for Column (2) in Table 3.5

	Coefficient	Standard Error
Union	0.071***	0.017
Married	0.019	0.016
British Columbia	0.044	0.052
Alberta	0.055	0.046
Manitoba	0.088	0.071
Saskatchewan	-0.022	0.063
Quebec	0.11*	0.064
New Brunswick	-0.174**	0.078
Nova Scotia	-0.022	0.048
PEI	-0.16**	0.062
Newfoundland	-0.157*	0.081
2000	0.053***	0.012
2001	0.087***	0.014
2002	0.162***	0.014
2003	0.203***	0.014
2004	0.246***	0.014
2005	0.309***	0.016
2006	0.365***	0.016
2007	0.415***	0.016
Comp1occ	0.012***	0.002
Comp2occ	0.011***	0.003
Comp3occ	-0.002	0.003
Comp4occ	-9.819e-04	0.003
Comp5occ	-0.001	0.004
Distance	-8.553e-04	0.003
College*Distance	0.003	0.003
Bachelor's*Distance	-0.001	0.002
Constant	2.299***	0.103

Within R^2 = 0.19 , *Between* R^2 = 0.183, *Overall* R^2 = 0.179

N = 24,170, *# of groups* = 6830, σ_α = 0.2, σ_u = 0.424

Results estimated from Fixed Effects Model in Equation 3.8

Table. A3.4. Full Results for Column (3) in Table 3.5

	Coefficient	Standard Error
Union	0.071***	0.017
Married	0.02	0.016
British Columbia	0.044	0.052
Alberta	0.055	0.046
Manitoba	0.086	0.069
Saskatchewan	-0.023	0.063
Quebec	0.108*	0.064
New Brunswick	-0.173**	0.078
Nova Scotia	-0.022	0.049
PEI	-0.159**	0.062
Newfoundland	-0.157*	0.081
2000	0.053***	0.012
2001	0.087***	0.014
2002	0.161***	0.014
2003	0.203***	0.014
2004	0.245***	0.014
2005	0.307***	0.016
2006	0.363***	0.016
2007	0.413***	0.016
Comp1occ	0.012***	0.002
Comp2occ	0.011***	0.003
Comp3occ	-0.002	0.003
Comp4occ	-9.991e-04	0.003
Comp5occ	-0.001	0.004
Distance	-0.002*	0.001
Years Since Degree*Distance	2.24e-05	4.96e-05
Constant	2.299***	0.103

Within R^2 = 0.19 , *Between* R^2 = 0.18, *Overall* R^2 = 0.175

N = 24,170, *# of groups* = 6830, σ_α = 0.2, σ_u = 0.425

Results estimated from Fixed Effects Model in Equation 3.9

Table. A3.5. Full Results for Column (4) in Table 3.5

	Coefficient	Standard Error
Union	0.07***	0.017
Married	0.02	0.016
British Columbia	0.043	0.052
Alberta	0.054	0.046
Manitoba	0.079	0.07
Saskatchewan	-0.028	0.064
Quebec	0.108*	0.064
New Brunswick	-0.176**	0.077
Nova Scotia	-0.03	0.049
PEI	-0.164***	0.061
Newfoundland	-0.161*	0.083
2000	0.053***	0.012
2001	0.087***	0.014
2002	0.161***	0.014
2003	0.203***	0.014
2004	0.246***	0.014
2005	0.308***	0.016
2006	0.364***	0.016
2007	0.415***	0.016
Comp1occ	0.012***	0.002
Comp2occ	0.011***	0.003
Comp3occ	-0.003	0.003
Comp4occ	-9.388e-04	0.003
Comp5occ	-0.001	0.004
Distance	-9.874e-04	0.002
Female*Distance	-0.003*	0.002
Constant	2.298***	0.104

Within R^2 = 0.191 , *Between* R^2 = 0.208, *Overall* R^2 = 0.2
N = 24,170, # *of Groups* = 6830, σ_α = 0.199, σ_u = 0.419
Results estimated from Fixed Effects Model in Equation 3.10