THE EFFECTS OF LANGUAGE ABILITY ON LABOR MARKET OUTCOMES

by
AHMET GİRİŞKEN

Submitted to the Graduate School of Social Sciences
in partial fulfilment of
the requirements for the degree of Master of Arts

Sabancı University
July 2022

THE EFFECTS OF LANGUAGE ABILITY ON LABOR MARKET OUTCOMES

Approved by:

Prof. Abdurrahman Bekir Aydemir .....
(Thesis Supervisor)

Asst. Prof. Yusuf Emre Akgündüz .....

Asst. Prof. Selcen Çakır .....

Date of Approval: July 6, 2022

ABSTRACT

THE EFFECTS OF LANGUAGE ABILITY ON LABOR MARKET OUTCOMES

AHMET GIRIŞKEN

ECONOMICS M.A. THESIS, JULY 2022

Thesis Supervisor: Prof. Abdurrahman Bekir Aydemir

Keywords: language ability, immigration, age at arrival, earnings, labor market outcomes

The main purpose of this paper is to analyze the causal impact of language ability on labor market outcomes in developed European countries using objective measure of language ability. To do so, observations from the first round of PIAAC survey conducted in 2012 are used. Using the critical age hypothesis to create an instrumental variable and concentrating on childhood immigrants, this thesis shows that language ability has a causal impact on earnings and occupational prestige, but has no impact on probabilities of labor force participation and employment. Furthermore, the study demonstrates that language ability has a causal effect on several labor market skills. These skills, on the other hand, have not been proved to be mediators between language ability and earnings. Finally, the study demonstrates that improving language skills reduces the experience gap among childhood immigrants.

ÖZET

DİL BECERİSİNİN İŞ PİYASASI ÇIKTILARINA ETKİLERİ

AHMET GİRİŞKEN

EKONOMİ YÜKSEK LİSANS TEZİ, TEMMUZ 2022

Tez Danışmanı: Prof. Dr. Abdurrahman Bekir Aydemir

Anahtar Kelimeler: dil becerisi, göç, göç edilen yaş, gelir, iş piyasası çıktıları

Bu çalışmanın temel amacı, dil yeteneğinin nesnel ölçümünü kullanarak gelişmiş Avrupa ülkelerinde dil becerisinin işgücü piyasası çıktıları üzerindeki nedensel etkisini analiz etmektir. Bunu yapmak için, 2012 yılında gerçekleştirilen PIAAC araştırmasının ilk turundan elde edilen gözlemler kullanılmıştır. Araçsal bir değişken oluşturmak için kritik yaş hipotezini kullanan ve çocukluk çağında göçen bireylere odaklanan bu tez, dil becerisinin kazançlar ve mesleki saygınlık üzerinde nedensel bir etkisi olduğunu, ancak işgücüne katılım ve istihdam olasılıkları üzerinde hiçbir etkisi olmadığını göstermektedir. Ayrıca, dil becerisinin çeşitli iş piyasası becerileri üzerinde nedensel bir etkisi olduğu gösterilmektedir. Bu beceriler ise dil becerisinin kazançlar üzerindeki kanalları olmadığı gösterilmiştir. Son olarak, dil becerilerini geliştirmenin, mesleki deneyim açığını azalttığını gösterilmektedir.

ACKNOWLEDGEMENTS

This paper uses public data from Round 1 of the Survey of Adult Skills PIAAC and Ganzeboom and Treiman's (2010) international stratification and mobility file.

First and foremost, I would like to express my heartfelt and deepest gratitude to my thesis advisor Prof. Abdurrahman Bekir Aydemir for his feedbacks and guidance at every step of my thesis. This thesis would not have been possible without his assisstance.

I would like to thank to my friend Mehmet Akın Şimşek for his help and support during my MA studies. I wish to thank my brother İbrahim Girişken for always being there for me. I would also like to appreciate the support of my friends Abdullah Can Özdemir, Cemilhan Aslan, Muhammed Suha Demirel and Onur Çetinkaya.

I am also grateful to the remaining jury member Yusuf Emre Akgündüz, Selcen Çakır and Ayşe Yeliz Kaçamak for their helpful comments and remarks.

I would like to express my endless gratitude to my parents Mevlüde and Mustafa Girişken, my brother Mehmet Girişken and my sister Eda Kuyucu.

Last but not the least, I would like to thank to my fiancée Doğa Kurt for supporting and believing in me!

In memory of Orhan Seyfi and Saadet Eralp

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION AND LITERATURE REVIEW

Language is a systematic medium of communication that enables people express their emotions and thoughts. It is also a fundamental instrument in the toolbox of human capital. It helps people accumulate more human capital and assists individuals in laying the groundwork for a productive environment (Isphording 2014). In addition to this, proficiency in the language of destination country plays a substantial role in the lives of immigrants, both in terms of integration in the society and the labor market outcomes (Chiswick and Miller 1995; Dustmann and Van Soest 2001). Improved language skills also lead to easier assimilation, higher earnings, and better job opportunities (Bleakley and Chin 2004; Dustmann and Fabbri 2003).

In this thesis, I try to estimate the causal impact of language ability on different labor market outcomes using a multi-country survey of adult skills. To the best of my knowledge, this is the first study using an objective measure of language proficiency to investigate the critical age hypothesis and the causal effect of language ability on earnings in a multi-national European context. Moreover, I investigate the impact of language ability on occupational prestige, a non-pecuniary return. Occupational prestige could be described as the extent to which society appreciates and respects an occupation apart from the pecuniary returns. I also study unique outcomes that have not been investigated in the literature. I analyze how language ability affects an individual's labor market skills such as reading, writing, and influencing others at work. These abilities are crucial since these could be the mechanisms via which an individual's language ability affects their earnings. Finally, I investigate whether language ability influences the time it takes for childhood immigrants to enter the labor market or find a new job after losing one, namely experience gap.

The endogeneity of language ability is a major stumbling block in estimating the influence of language ability on the mentioned labor market outcomes. Language proficiency is linked to several other factors that may influence the outcome of interest. A person with a high degree of ambition or motivation, for example, learns the language very well and may earn more because of her ambition. Furthermore, due to

reverse causality, language proficiency might be endogenous to earnings and occupational choice. For instance, consider an immigrant who works in a job that does not require fluency in a language because she lacks the language of host country. This immigrant who works in less language-intensive occupations will speak less fluently since she does not have the chance to improve the language abilities. Discrimination may also affect the earning outcomes of immigrants. Employers could discriminate immigrants based on their accents. For example, since Australian and Jamaican immigrants speak English with distinct accents, a British employer may treat them differently during a job interview and offer salaries accordingly. These situations could lead to reverse causality in language ability case. As a result, the causal link between language ability and variables of interest cannot be predicted using an OLS estimation. Therefore, I use 2SLS-IV method to alleviate the endogeneity issue.

Several earlier research of the influence of language ability on labor market outcomes, such as Kassoudji (1988) and Chiswick (1991), disregarded the issue of endogeneity and use language ability as a regressor in their models, demonstrating a positive link between language ability and earnings. Chiswick (1991) found that self-reported measure of reading well/very well raises salaries by 0.26 percent using a survey of illegal aliens. On the other hand, Chiswick and Miller (1995) underlines the endogeneity between dominant language proficiency and earnings using IV estimation and benefiting from datasets of several developed countries. In the OLS analysis, they find that the effect of language ability is positive and statistically significant. However, the coefficient exhibits great volatility with varying signs of the effect in the instrumental variables analysis. They note that it is positive and statistically significant in the US, but that the magnitude are surprisingly large compared to the OLS estimation. To alleviate measurement error and endogeneity of linguistic ability, Dustmann and Fabbri (2003), using both self-assessed and interviewer-assessed language ability, combines a matching estimator with an IV estimator. They demonstrated that ethnic minority immigrants in the UK earns less due to a lack of English proficiency. Besides significant earning effects, they also find a higher employment probability of around 20 percent for immigrants who are proficient in English. Dustmann and Van Soest (2003) adresses the issue of endogeneity by using panel data for Germany with self-reported information on immigrants' language ability. First, they utilize a matching type estimator that takes into account the partner's and household's characteristics. Their second strategy is to use parental education as an instrument, which has been questioned and criticized in the consideration of language's effects on earnings. Despite highlighting the instrument's flaws, they continue to use it, stating that the criticism will be less noteworthy in the case of immigrants' language abilities.

Their result shows that language proficiency has a positive causal impact on earnings and the IV result is much higher than the OLS estimate. Bleakley and Chin (2004) finds a strong positive effect of English-language abilities on earnings of individuals who arrived in the US as young individuals using data from the 1990 US Census. Their identification strategy is motivated by the critical age hypothesis from psychobiological literature. Benefiting from the strategy of Bleakley and Chin (2004), Isphording and Sinning (2012) reveals positive returns to language skills in the US for child and adult immigrants. They show that schooling mediates a significant portion of the effect of language skills on earnings of child. This finding is remarkably similar to Bleakley and Chin (2004). However, they find that the returns for adult immigrants are not dependent on schooling. Their instrument for this sample is the interaction term of years since migration and a dummy variable indicating if a person is from a non-English-speaking country . On the other hand, Guven and Islam (2015) shows that schooling is not the mediator of the effect of language ability on earnings for the immigrants in Australia. By concentrating on Germany and using self-reported measures of oral and written fluency in German, Isphording et al. (2014) show that simple OLS regressions underestimate the positive effects of language skills on earnings. Their key findings show that a lack of written and spoken German has a considerable impact on employment probabilities and wages.

Aside from earnings and employment outcomes, the causal link between language proficiency and labor market skills is also worth investigating. Autor and Handel (2013) presents a conceptual model for examining the link between wages, job tasks, and human capital. According to this paper, workers execute various sets of skills in work environment, and these skills are primary drivers of their hourly earnings, both between and within occupational and education groups. They also discovered that job task variation is systematically related to race, gender, and English-language proficiency among workers in the same occupations. Measurement of job tasks at work, on the other hand, is critical for a careful understanding of tasks and their relationship to language ability as well as to earnings. PIAAC data has an advantage of providing the task contents of jobs at the individual level benefiting from questions answered directly by respondents. This advantage helps to measure the task intensity more precisely in contrast to the initial studies such as Autor and Dorn (2011) as well as Acemoglu and Autor (2011). These studies investigated the task intensities at the occupation level utilizing two large occupational datasets derived from questionnaires that provided work content to estimate job tasks. These are the Dictionary of Occupational Titles (DOT) and its successor, the Occupational Information Network (O*NET). However, DOT and O*NET are derived from US surveys

and the task variables are derived from the occupations' characteristics[1]. Despite the similarities of the labor markets in the United States and Europe, it is important to evaluate the task contents within the European settings that is made possible by PIAAC data. Moreover, their method of measuring task intensities could result in the disregard of variation within occupations. In addition to the job skill measures in the PIAAC, parallel to Autor and Handel(2013) and in line with the method of Agasisti et al.(2021) as well as De La Rica and Gortazar (2020), I created task-intensity indices using the PIAAC survey questions implementing principal component analysis. Agasisti et al.(2021) finds that there is a positive association between literacy score and abstract task-intensity index, but there is a negative link between literacy score and manual task-intensity index. They also employed all task-intensity indices, language ability, and other control variables in an extended Mincerian wage regression. They discovered that literacy ability has a positive and significant relationship with earnings. Furthermore, they concluded that the abstract task index has a positive association with earnings, while the manual task index has a negative relationship. However, their results from the OLS estimation can not be given a causal interpretation due to potential endogeneity issues surrounding the various measures. In this thesis, the findings of Agasisti et al.(2021) on the influence of language ability are validated from the causal perspective. Moreover, De La Rica and Gortazar (2020) finds that abstract job skills are positively correlated with earnings while manual job skills are negatively associated with earnings, using OLS estimation. Their results are similar to the findings of Agasisti et al.(2021) but they do not include language ability in their model. Instead, they use numeracy skills.

To predict the causal relationship, I use an instrumental variable (IV) approach proposed by Bleakley and Chin (2004) and modified by Isphording (2014). To construct my instrument, I follow the Critical Age Hypothesis offered by Lenneberg (1967) and Newport (2002). According to this hypothesis, young children acquire a language more readily and successfully than adults, making it easier for them to assimilate into the host country. I also follow the idea that being familiar with the destination language reduces both financial and psychological costs of migrating (Adserà et al. 2016). Therefore, to create the instrument, I use a measure of linguistic distance created by German Max-Planck Institute. The main idea is that having a low linguistic distance score makes it easier for immigrants to assimilate than those who have a larger linguistic distance score. Combining these two ideas, my instrument would be the interaction term of the linguistic distance and age at arrival.

[1] Borelli (2016) proves a strong correlation between and task intensity measures of PIAAC data and O*NET in the context of the US. For example, the correlation between PIAAC writing skill and O*NET writing skill is 0.85, the correlation between PIAAC reading at work and O*NET reading skill is 0.89.

In this thesis, first, I examine the critical age hypothesis for second-language acquisition in a broad multi-national context using an objective measure of language ability and individual level data of the Programme for the International Assessment of Adult Competencies (PIAAC) survey conducted in 2011-2012. According to my findings, the critical age is 12, which is remarkably similar to what has been documented in the literature. Subjective measurements of language ability have been employed in earlier studies evaluating the critical age hypothesis and the causal relationship of language ability and labor market outcomes such as earnings. Individuals were asked how they rate their language skills in surveys utilized in these research' datasets. However, this may lead to a measurement error. For example, an immigrant who generally interacts with individuals from her country of origin may assume that she is fluent in the language of host country, despite the fact that she is not. This issue is not a concern in my study because I use objective measure of language ability. Secondly, previous studies use data of a single country in their analysis. For example, Bleakley and Chin utilize a dataset from the United States. Guven and Islam (2015) use data from an Australian household survey. By combining PIAAC datasets from a homogenous set of developed countries which mainly European countries constitute, I establish a multinational dataset and generalize the findings of the previous studies. The reason for creating this homogenous set is that 27 countries in the PIAAC that include countries such as Kazakhstan, Japan, and Israel have different labor market structures and institutions. Furthermore, the motivations for moving to these countries could differ. An immigrant heading to Kazakhstan, for example, has a different objective than an immigrant moving to Austria. Focusing on this homogenous set of European countries[2], I find that language ability has a positive and significant effect on earnings and educational attainment. In contrast to the findings of Bleakley and Chin(2004), the effect of language ability on earnings does not appeared to be channelled by the effect of language ability on years of schooling. Thirdly, I use a different and stronger instrument than the instrument Bleakley and Chin (2004) uses. I discuss the strenght of each instrument, attempting to explain why one outweighs the other. Additionally, this is the first article that I am aware of that investigates the causal effect of language skill on labor market skills. PIAAC offers several variables of labor market skills in the data. In addition, I also create an index of abstract and manual tasks using principal component analysis. I conclude that literacy ability has a causal and positive impact on labor market skills such as writing, reading as well as influencing others at work. I also conclude that the effect of labor market skills does not appear to mediate any part of the effect of language ability on wages. Moreover, I find

[2]The public use data of US, Canada and Australia are not available. Therefore, it is not possible to make the analysis in the North American context.

that language proficiency has a positive impact on occupational prestige and has a negative effect on the experience gap of childhood immigrants.

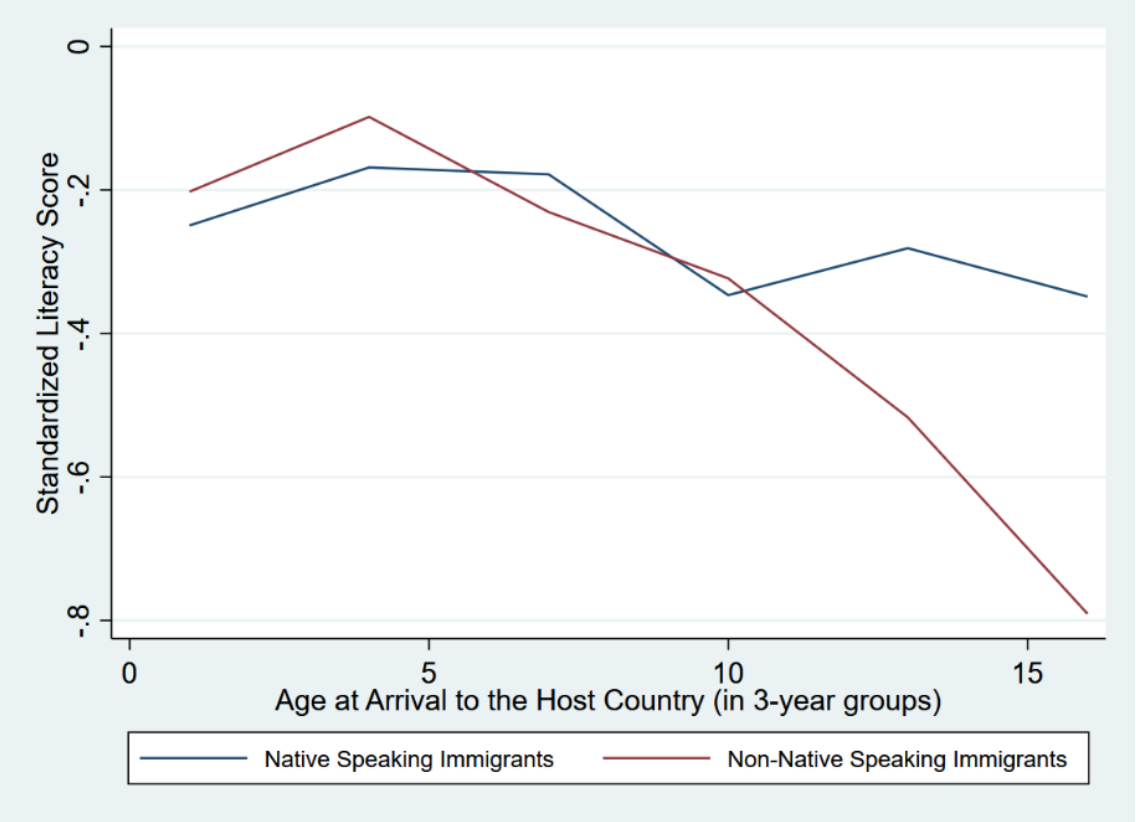
2. THE CRITICAL AGE: AGE AT ARRIVAL AND LANGUAGE PROFICIENCY

According to the critical age hypothesis, initially proposed by Lenneberg (1967), the initial years of life is the crucial time in which a person can acquire language ability at the native level. In other words, it is easier for younger individuals to learn a new language while adults often struggle learning languages. In the context of immigrants, older and non-native speaker immigrants experience greater challenges in language acquisition. As Chiswick and Miller (1995) suggested; the younger the age at arrival, the better the adaptation of immigrants to the host country. The exact time for the critical age is assumed to be around the beginning of adolescence and linked to psychological changes in the brain (Lenneberg 1967). However, there is no consensus over the exact time of the critical age of language acquisition. While Newport (2002) indicated that declines in the second language acquisition might occur as early as age 4–6, Bleakley and Chin (2004) and (2010) found it to be around the age of 10 and 12, respectively. Beck et al. (2012) found the critical age of 8 focusing on education attainments of migrants in their study. Fenoll(2018) finds the critical age of 6 using the New Immigrant Survey (NIS) conducted among US immigrants. Moreover, in their study of occupational achievement, Zhang Ye (2018) employ the critical age of 13.

Figure 2.1 depicts the mean language ability among childhood immigrants from non-native speaking countries and childhood immigrants from native speaking countries of PIAAC dataset for each level of age upon arrival. It is important to note that literacy scores are standardized within each host country using all PIAAC sample including natives and immigrants. Parallel to the findings of previous studies, all immigrants from native speaking countries have a relatively stabile level of standardized literacy scores with a modest reduction after the age of seven. Young immigrants from non-native speaking nations acquire the same levels of literacy as their native-speaking peers. On the other hand, immigrants from non-native speaking countries have considerably lower standardized literacy scores after age at arrival

of twelve compared to immigrants from native speaking countries. The finding supports the critical age hypothesis, which claims that those who came after the critical age had much lower language skills, as well as, the findings of Bleakley and Chin (2004).

Figure 2.1 The Investigation of Critical Age Hypothesis



3. DATA

I use data from the Programme for the International Assessment of Adult Competencies (PIAAC) provided by Organization for Economic Co-operation and Development (OECD) through the first round, which took place in 2011-2012. The survey assesses respondents' skills in areas such as literacy, numeracy, and problem solving, resulting in an internationally comparable skill measure. In addition, it gathers labor market related information such as age, gender, age at immigration, country of birth, languages spoken at home, educational attainment, earnings, skills used at work, and occupation[1]. As recommended by Bleakley and Chin (2004), I restrict my sample to childhood immigrants. Childhood immigrants are individuals who arrived in the host country while they were under the age of 18. The basic rationale behind this restriction is that these individuals did not choose to immigrate at a young age; rather, they followed their parents to the host country. Beck et al. (2012) shows that the age at immigration of child migrants to the US affects educational outcomes and conclude that children and adults have distinct experiences with migration. To alleviate the selection bias issue for language acquisition, I also limit my sample to the first-generation immigrants with both parents born outside of the host country. While doing this, I exclude 1.5 and 2nd generation immigrants because these immigrants have at least one native parent born in the host country. Moreover, the advantage of PIAAC dataset is that it has exact age at arrival while year of arrival in the United States is reported in multiyear intervals in the study of Bleakley and Chin (2004). This helps to test the critical age hypothesis more precisely. I also restrict my sample to individuals aged under 65 during the survey. I utilize hourly earnings in logarithmic form. Hourly earnings including bonuses are converted to US dollars using purchasing power parity exchange rates to makes sure that the results of the analysis are comparable across nations. To exclude outliers in the data, I further restrict my sample to individuals who earn less than 150 US

[1] OECD (2019), Skills Matter: Additional Results from the Survey of Adult Skills, OECD Skills Studies, OECD Publishing, Paris.

dollars PPP per hour. I employ a comprehensive set of 27 countries[2], to investigate the critical age hypothesis. Then, I create a homogeneous group of developed European countries for the analysis of labor market outcomes. These countries are Austria, Belgium, Denmark, Finland, France, Germany, Netherlands, Norway, and United Kingdom. For the language ability, I use literacy score standardized within each host country to eliminate the cross-country differences. In order to calculate the literacy scores, I take the average of all ten plausible values of literacy scores in the PIAAC dataset. According to OECD, literacy is described in the PIAAC assessments as "reading component" skills that enable individuals to derive meaning from written texts: knowledge of vocabulary, capacity to process meaning at the sentence level, and fluency in reading passages of text. In other words, it is the capacity to participate in society, attain one's objectives, and expand one's knowledge and potential through understanding, evaluating, using, and engaging with written materials. Literacy ability is one dimension of the language ability. It may matter more for some tasks compared to other tasks. Although, it is one of the dimensions, it is important to note that there is a positive correlation between reading and speaking abilities according to the findings of Liao et al. (2010) who studies the relationships between test scores.

I divide immigrants in the sample into two categories in terms of language spoken: native speaking immigrants who have spoken the language of host country at home during childhood and non-native speaking immigrants who have not spoken the language of host country at home during childhood. The first group is treatment group whereas the second is the control group. It's important to underline that the comparison group of non-native speaking immigrants does not include all native speakers, including native born individuals and native-speaking immigrants, but solely native-speaking immigrants.

In the IV setting, I employ two separate variables to create instruments. First, the dummy variable "nonnative" representing whether an immigrant speaks the host country's language or not. The dummy equals 1 if the host country's language is the different than the first or second language learnt at home in childhood, 0 otherwise. Second, linguistic distance between the language of the host country and the languages of the birth country. I employ a measure of linguistic distance based on phonetic dissimilarity across languages, carried out by the Automated Similarity Judgment Program (ASJP), a research program by the German Max-Planck Institute of Evolutionary Anthropology. It contains 40-item word lists of different languages. The words are from the 'Swadesh list' (Swadesh 1952), a deductively con-

[2]The list of these countries could be found in the appendix.

structed collection of words that includes fundamental communication words and words that describe ordinary items. The normalized Levenshtein distance provides a continuous measure of linguistic differences. For example, the distance between English and Dutch is 63 whereas the distance between English and Vietnamese is 104. Alternatively, the distance between German and Dutch is 52 whereas the distance between German and Palestinian Arabic is 104.

The PIAAC Survey of Adult Skills develops scale scores of work place skills such as reading at work (READWORK), writing at work (WRITWORK), and influencing others at work(INFLUENCE) using a background questionnaire on real tasks performed at work. READWORK is defined as to what extent one uses reading skills at work while WRITWORK is defined as to what extent one uses writing skills at work. Moreover, INFLUENCE is defined as instructing, teaching, training or advising people, persuading others and negotiating. However, in the calculation of these variables, PIAAC disregarded the answers with "Never". For example, if an immigrant responded "Never" to all questions in the INFLUENCE variable, their INFLUNCE score becomes missing value. However, it is important to take into account these answers for the indices. Therefore, I created the same indices of writing, reading and influence using principal component analysis. I also created measures of task intensity indices of manual and abstract tasks, based on the first principal component analysis following the strategy of Agasisti et al. (2021) and De La Rica et al. (2020). Principal component analysis is mainly used for dimensionality reduction to produce lower-dimensional data while keeping as much variance as feasible.

For labor market indices, combination of variables are reported in the appendix. For manual task index, I use the question of "How frequently working physically for a long period?". Individuals choose one of the following responses to the questions: taking the values 1 through 5, never; less than once a month; less than once a week but at least once a month; at least once a week but not every day; every day.

I also use occupational prestige as a dependent variable in addition to earnings. Occupational prestige refers to how the public perceives an individual's social status based on their professions. The PIAAC data set records immigrants' professions using four-digit level (ISCO08) codes from the International Standard Classification of Jobs (ISCO08). I utilize Ganzeboom and Treiman's (2010) International Stratification and Mobility File to create a measure of occupational prestige. ISCO2C's two-digit level occupational codes are converted into a continuous measure of occupational prestige in this file. To illustrate, "chief executives, senior officials and legislators" have occupational prestige score of 72 while "health professionals" have occupational prestige score of 77. On the other hand, "personal service workers"

such as waiters and hairdressers have occupational prestige score of 28. Workers employed as "mixed crop farmers" have a minimum occupational prestige score of 14.

Finally, I use the difference between potential experience[3] and actual experience to build the experience gap variable. For this study, I excluded participants with a gap of fewer than 15 years in age and experience. The basic premise is that people begin getting paid after the age of fifteen.

Table 3.1 presents the descriptive statistics for the analysis sample of labor market outcomes. For childhood immigrants who arrived before the age of 18, the number of observations and means are presented separately for those who came in the host country as younger (0-11) and older (12-17) individuals from non-native speaking and native speaking countries.

[3]Potential experience is defined as age of an immigrant minus years of schooling minus six

Table 3.1 Descriptive Statistics

	All Immigrants					Non-Native Speakers				Native Speakers			
						AAA (0-11)		AAA (12-17)		AAA (0-11)		AAA (12-17)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
VARIABLES	N	mean	sd	min	max	N	mean	N	mean	N	mean	N	mean
Age	650	34.24	11.93	16	64	295	32.09	213	34.57	121	39.02	21	33.57
Years of Schooling	650	11.84	3.185	5	21	295	11.94	213	11.82	121	11.60	21	12.05
Language Score	650	0.153	0.914	-2.727	2.437	295	0.265	213	-0.0747	121	0.256	21	0.286
Age at Arrival (AAA)	650	8.623	5.328	0	17	295	5.631	213	14.67	121	4.298	21	14.29
Gender (female=1)	650	0.452	0.498	0	1	295	0.437	213	0.451	121	0.463	21	0.619
ln Hourly Earnings	650	2.740	0.513	-0.0739	4.487	295	2.717	213	2.757	121	2.779	21	2.648
Linguistic Distance (LD)	650	65.62	42.39	0	103.4	295	82.68	213	85.75	121	0	21	0
Occupational Prestige	564	39.40	20.26	12.87	88.96	248	39.58	190	39.20	107	39.25	19	39.74
Experience gap	586	3.841	5.067	0	33	254	3.224	203	4.296	110	4.055	19	6
Manual	650	3.560	1.723	1	5	295	3.542	213	3.808	121	3.198	21	3.381
`Manual_D`	650	0.520	0.500	0	1	295	0.512	213	0.577	121	0.446	21	0.476
Abstract	650	0.0206	1.618	-2.051	5.313	295	0.0337	213	-0.0455	121	0.146	21	-0.217
Reading	650	0.0418	1.721	-2.431	4.266	295	0.0626	213	-0.0695	121	0.220	21	-0.146
Influencing	645	-0.00230	1.773	-2.294	4.492	292	-0.0425	211	-0.0829	121	0.245	21	-0.0561
Writing	650	0.0342	1.268	-1.362	4.599	295	0.0220	213	0.00875	121	0.103	21	0.0640

Note: Table includes analysis sample for labor market outcomes (age at immigration 0-17).
AAA denotes age at arrival to the host country. `Manual_D` is a dummy variable which takes the value of one if the question of working physically long is responded with everyday, zero otherwise.

4. EMPIRICAL STRATEGY

Because language ability is endogenous, it is not possible to get an unbiased estimate of the coefficient of interest using an OLS estimation. Following the strategy of Bleakley Chin (2004) and Isphording (2014), I estimate a two-stage least squares (2SLS) instrumental variable regression model to identify the causal impact of language ability on the dependent variables. The second stage equation relates the dependent variable, labor market outcomes, to the endogenous regressor, language proficiency. In the first-stage estimation, the endogenous variable is regressed on the instrument and all exogenous regressors.

The strategy I follow for building the instrument is that the standardized literacy score difference between childhood immigrants from non-native-speaking countries and childhood immigrants from native-speaking countries is zero up to the age of eleven, but then begin to diverge with age at arrival beyond that age. As Bleakley and Chin (2004) point out, the exclusion restriction appears to be difficult to justify with the instrument of age at arrival solely. The reason for this is because younger and older arrivals are likely to differ on non-linguistic aspects (Friedberg 2000), which could have an impact on labor market outcomes. It is also possible to claim that, for reasons unrelated to language abilities, younger immigrants will be better equipped to adapt to the host country's culture. To exemplify, younger arrivals are exposed to the host country's educational system, which may assist them in assimilating and embracing the host country's culture. As underlined by Adserà et al. (2016), being familiar with the destination language decreases both financial and psychological costs of migrating. Therefore, to account for several non-linguistic factors, I use Isphording et al. (2014)'s strategy of employing an interaction term of age at arrival with linguistic distance between the home and host country language rather than a dummy variable that takes the value of one if native language of the immigrant and host country languages are the same (Bleakley and Chin 2004). For native-speaking immigrants, the value is 0. To provide an example for non-native speakers, ceteris paribus, a Dutch immigrant in Germany has a relative advantage over an

Arabic immigrant since German is significantly closer to Dutch in terms of linguistic characteristics, leading to a lower linguistic distance value. In short, immigrants with lower value of linguistic distance suffer less in acquisition costs by a later age at arrival (Isphording 2014). Moreover, linguistic distance has the advantage of being continuous and it creates more variation in the instrument for non-native speaking immigrants. The value of the interaction term for non-native-speaking immigrants varies in contrast to the case of dummy variable of 1. Furthermore, the estimation is conditioned on a comprehensive set of region of origin and host country characteristics to adjust for extra unobservable region of birth and host country heterogeneities, making this identification assumption more plausible. This allows to account for psychological and other types of costs. The tables in the results section show that the instrument satisfies the relevance condition.

As discussed, I use an instrument for the language proficiency to eliminate the endogeneity of language ability. I use the interaction term of the age at arrival and linguistic distance as the instrument:

$$Z_{ijka} = max(0, AAA_{ijka} - 11) * LD \tag{4.1}$$

The regression specification is as follows:

$$Y_{ijka} = \alpha + \beta LANG_{ijka} + \rho W_{ijka} + \delta_a + \gamma_j + \theta_k + u_{ijka} \tag{4.2}$$

for individual i, born in region j, immigrated to the country k at age a. Y_{ijka} is the outcome of interest, and $LANG_{ijka}$ is a measure of language proficiency (the endogenous regressor). δ_a is a set of age-at-arrival fixed effects. This specification is more flexible than including a dummy variable for arriving old (Bleakley and Chin 2004). For robustness, I also employ the dummy of arriving old in a separate specification as in Eq.(4.3):

$$Y_{ijka} = \alpha + \beta LANG_{ijka} + \rho W_{ijka} + \delta * max(0, AAA_{ijka} - 11) + \gamma_j + \theta_k + u_{ijka} \tag{4.3}$$

The vector W_{ijka} contains a set of control variables such as age and gender. I account for differences in age and gender effects in host countries by including interactions

of host country dummies with age and gender. Lastly, u_{ijka} is the error term.

The endogenous regressor is linked to the instrument and all other exogenous variables in the first- stage regression:

$$LANG_{ijka} = \alpha + \pi Z_{ijka} + \rho W_{ijka} + \delta_{1a} + \gamma_{1j} + \theta_{1k} + u_{ijka} \tag{4.4}$$

The coefficient of interest β comes from the second stage regression, which reflects the language proficiency's local average treatment impact (LATE) on labor market outcomes.

5. EMPIRICAL RESULTS: LANGUAGE PROFICIENCY AND LABOR MARKET OUTCOMES

5.1 Labor Force Participation, Employment Probabilities and Experience Gap

With both the OLS and IV estimates, Table 5.1 illustrates the influence of language skills on the probability of labor force participation and the probability of employment conditional on labor force participation. According to the OLS estimates, there is a positive association between language ability and the likelihood of the labor force participation; indicating that a one standard deviation increase in langauge ability increases participation by about 4 percentage points. While the corresponding IV estimates are larger in magnitude, they are not statistically significant. Column 5-6 in Table 5.1 illustrates the effect of language proficiency on the employment probability. OLS estimation demonstrates that there is positive association between these two variables with no significance level. It also shows that language proficiency has no causal effect while the corresponding IV estimates are larger in magnitude. In short, language ability does not have any significant causal effect on the probability of labor force participation and employment for childhood immigrants in the survey year. However, it is important to explore if language ability affects the waiting time for labor market for immigrants who arrived in the host country before reaching the age of 18. It might effect the time it takes to get the first job as well as the time it takes to acquire a job after losing one. The dependent variable is experience gap and results are shown in Table 5.1. IV result in column 10 illustrates that one point increase in the standardized literacy score decreases experience gap by 3.9 year. On the other hand, according to OLS estimate, a one-point standardized literacy score is negatively associated with a 1.3-year of experience gap. Hence, the IV coefficient is three times that of the OLS estimate. The findings in this section suggest that language ability has no causal impact on people's current labor market conditions. However, it is possible that it affected their employment conditions in the early

years of their careers. A childhood immigrant with lower level of language skill, for example, may have chosen to invest more in human capital rather than enter the job market. For the childhood immigrant with weak language skills, this might result in a larger experience gap.

5.2 Pecuniary and Non-Pecuniary Returns to Language Proficiency

Table 5.2 displays both the OLS and IV estimates of the effect of language ability on earnings and occupational prestige score of immigrants who arrived to the host country before the age of 18. These are pecuniary and non-pecuniary returns to language ability.

5.2.1 Pecuniary Return

The estimation with the dummy of arriving old is shown in the first two columns, while the estimation with age-at-arrival fixed effects is shown in columns 3-4. According to the OLS estimate, increase in the literacy score by one standard deviation is associated with a 14 percent increase in earnings whereas the corresponding IV estimate implies a 41 percent rise in earnings. I observe that the magnitudes of the coefficient is identical for the specifications with the dummy of arriving old and with the age at arrival fixed effects, but F-statistic of the fixed effect estimation is greater. In addition, the effect of arriving late is found to be zero after I control for language ability and other variables such as age and gender. To sum up, I find that language ability is significantly positively related to earnings, similar to the findings of Bleakley and Chin (2004), Guven and Islam (2015), and Dustmann and van Soest (2002). Since I use standardized test scores, it is difficult to compare my findings to the findings of these studies because these studies use discrete measures of language skill that varies from 0 to 3. On the other hand, parallel to these papers, the IV coefficients are substantially greater than the OLS coefficients. My findings are in accordance with those of Bleakley and Chin (2004), who found that IV estimates are two to three times greater than OLS estimates. According to Bleakley and Chin (2004), this difference might be explained by two factors: measurement error in the language ability assessment and differences in the weighting methods underlying the OLS and IV estimations. Because the PIAAC measures language ability objectively, I attribute the difference to the latter factor. In other words, while the variation in 2SLS estimation depends on the variation in the instrument, the OLS estimate

reflects all of variation in language skills.

5.2.2 Non-Pecuniary Return

In Table 5.2, OLS estimation demonstrates that there is a positive relationship between language proficiency and occupational prestige scores of childhood immigrants. One point increase in the standardized literacy score is associated with 10 points in the occupational prestige score, which ranges from 12.9 to 89 with a standard deviation of 20.2 in the sample. According to the IV estimation in column 6, language ability has a causal effect on the occupational prestige. It shows that one point increase in the standardized literacy score increases occupational prestige score by 6.5 points. It also shows that arriving old to the host country decreases the score. After the age of 12, each consecutive year reduces the score by 0.6 point. It is worth noting that the IV coefficient is less than the OLS coefficient. This result is not the same as the earnings estimate.

Table 5.1 The Effect of Language Ability on Labor Force Participation and Employment Probabilities, and Experience Gap

Dependent Variable:	LFP	LFP	LFP	LFP	Employed	Employed	Employed	Employed	Experience Gap	Experience Gap	Experience Gap	Experience Gap
Mean:	(0.45)	(0.45)	(0.45)	(0.45)	(0.32)	(0.32)	(0.32)	(0.32)	(3.841)	(3.841)	(3.841)	(3.841)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Language Proficiency	0.035*	0.103	0.038*	0.112	0.021	0.127	0.023	0.151*	-1.278***	-3.865**	-1.180***	-3.576*
	(0.017)	(0.104)	(0.019)	(0.083)	(0.018)	(0.080)	(0.017)	(0.091)	(0.328)	(1.863)	(0.299)	(2.023)
Max(0, AAA-11)	-0.012	-0.006			0.006**	0,014*			0.260	0,089		
	(0.008)	(0.008)			(0.008)	(0.008)			(0.161)	(0.263)		
First Stage												
Max(0, AAA-11)*LD		-0.001***		-0.001***		-0.002***		-0.001***		-0.002***		-0.002***
		(0.000)		(0.001)		(0.000)		(0.000)		(0.001)		(0.000)
t-statistics		-3.088		-2.613		-4.706		-3.193		-2.863		-3.741
Observations	1,208	1,208	1,208	1,208	861	861	861	861	586	586	586	586
R-squared	0.430	0.415	0.441	0.423	0.400	0.341	0.420	0.337	0.632	0.511	0.649	0.556
AAA FE	-	-	YES	YES	-	-	YES	YES	-	-	YES	YES

Note: Each regression includes a gender dummy variable(female = 1), age, and dummy variables for region of birth and host country.
Differences in age and gender effects in host countries are accounted by including interactions of host country dummies with age and gender.
LFP denotes labor force participation. Standard errors are clustered at the host country levels in parentheses *** p<0,01, ** p<0,05, * p<0,1.

Table 5.2 The Effect of Language Ability on Pecuniary and Non-Pecuniary Returns

Dependent Variable:	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Occupational Prestige	Occupational Prestige	Occupational Prestige	Occupational Prestige
Mean:	(2.74)	(2.74)	(2.74)	(2.74)	(39.4)	(39.4)	(39.4)	(39.4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Language Proficiency	0.138***	0.411**	0.137***	0.406**	9.979***	6.542**	10.060***	5.618*
	(0.018)	(0.188)	(0.021)	(0.204)	(0.802)	(3.075)	(0.756)	(3.083)
Max(0, AAA-11)	-0.002	0.017			-0.392**	-0.563***		
	(0.004)	(0.018)			(0.149)	(0.094)		
First Stage								
Max(0, AAA-11)*LD		-0.002***		-0.002***		-0.002***		-0.002***
		(0.001)		(0.001)		(0.001)		(0.001)
t-statistics		-3.393		-4.273		-2.821		-3.560
Observations	650	650	650	650	564	564	564	564
R-squared	0.615	0.483	0.621	0.503	0.611	0.597	0.645	0.624
AAA FE	-	-	YES	YES	-	-	YES	YES

Note: Each regression includes a gender dummy variable(female = 1), age, and dummy variables for region of birth and host country. Differences in age and gender effects in host countries are accounted by including interactions of host country dummies with age and gender. Standard errors are clustered at the host country levels in parentheses *** p<0,01, ** p<0,05, * p<0,1.

5.3 Language Ability and Labor Market Skills

In this part, using the same 2SLS method in equation Eq.(4.3), I investigate the causal impact of language ability on several labor market skills such as reading, writing and influencing people at work. I also create abstract and manual task indices following the definitions of Agasisti et al.(2021) and De La Rica and Gortazar (2020).

According to column one in Table 5.3, one point increase in the standardized literacy score is associated with 0.7 percentage point increase in the score of using reading skill at work. Furthermore, the 2SLS result in column two demonstrates that one-point improvement in the standardized literacy score results in 1 percentage point increase in the employment of reading skills. The F-statistic of the instrument is 9.2. I also observe that the coefficient of 2SLS regression is more than that of the OLS result. This is similar to the findings in the earning results. Another conclusion is that the chance of employing reading skills at work is lower for those where the dummy of arriving old equals one. It means that arriving later than the age of 12 decreases the reading skill.

In the case of writing skill, the main result is similar to the case of reading skills. Language proficiency is positively correlated with the writing skill according to the OLS estimation. The 2SLS result in column three in Table 5.3 demonstrates that a one-point improvement in the standardized literacy score has positive and significant impact on the probability of employing writing skills at work. Arriving in the host country at a later age has no impact on using writing abilities at work.

In the case of influencing others at work, OLS estimation suggests that one point increase in language ability is positively associated with 0.6 point of influencing others at work index. In the 2SLS model, I also find that the skill of influencing others increases with the language ability. One point increase in the standardized literacy score increases the influencing skill by 1.3 percentage point with the significance level of 90 percent. OLS estimate is less than half of the IV estimate as in the case of baseline result. I also observe that there is no age at arrival effect on the skill of influencing after I control for language ability and other observables. The F-statistic for the instrument is 9.2. In addition to skill variables PIAAC offers, I also study the effect of language ability on the abstract and manual job market skills. OLS result shows that standardized literacy score is positively correlated with abstract skill with a coeffificent estimate of 0.7 points. The 2SLS result in column eight in Table 5.3 shows that one standardized score of literacy score increases the abstact

skill index by 0.9 points. The effect of arriving late in the host country is 0 and the F-statistic is 9.24 for the instrument.

In the case of manual job skills which is defined as working physically long, literacy score is negatively correlated in the OLS estimation. Column 9 shows that as the standardized score increases by one point, manual job index decreases 0.6 points. However, column ten shows no significant causal relationship between language ability and manual skill index. One observation in the 2SLS estimation is that as age at arrival increases, the manual index increases by 0.09 points. Since the number of observations is very small, I create a dummy variable for manual index that takes value of 1 if the respondent respond "Everyday" to the question of "How often do you work physically long?", 0 otherwise. Column twelve shows that a unit increase in the standardized literacy test score decreases the likelihood of answering "Everyday" by 0.25 points with the significance level of 90 percent.

Table 5.3 The Effect of Language Ability on Labor Market Skills

Dependent Variable:	Reading	Reading	Writing	Writing	Influence	Influence	Abstract	Abstract	Manual	Manual	Manual_D	Manual_D
Mean:	(2.74)	(2.74)	(2.74)	(2.74)	(39.4)	(39.4)	(39.4)	(39.4)	(3.56)	(3.56)	(0.52)	(0.52)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Language Proficiency	0.674***	1.011**	0.452***	0.509**	0.583***	1.285*	0.705***	0.907*	-0.598***	-0.411	-0.166***	-0.246**
	(0.126)	(0.494)	(0.096)	(0.205)	(0.103)	(0.733)	(0.112)	(0.538)	(0.082)	(0.384)	(0.025)	(0.110)
Max(0, AAA-11)	-0.104	-0.081*	0.001	0.005	-0.093*	-0.045	-0.073***	-0.059	0.077	0.090*	0.016	0.011
	(0.068)	(0.048)	(0.029)	(0.020)	(0.041)	(0.075)	(0.021)	(0.049)	(0.045)	(0.047)	(0.013)	(0.011)
First Stage												
Max(0, AAA-11)*LD		-0.002***		-0.002***		-0.002***		-0.002***		-0.002***		-0.002***
		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)
t-statistics		-3.393		-3.393		-3.381		-3.393		-3.393		-3.393
Observations	650	650	650	650	645	645	650	650	650	650	650	650
R-squared	0.550	0.532	0.507	0.506	0.516	0.442	0.551	0.544	0.515	0.509	0.489	0.477

Note: Each regression includes a gender dummy variable(female = 1), age, and dummy variables for region of birth and host country.
Differences in age and gender effects in host countries are accounted by including interactions of host country dummies with age and gender.
Standard errors are clustered at the host country levels in parentheses *** p<0,01, ** p<0,05, * p<0,1.

6. ROBUSTNESS CHECK

6.1 Earning "Returns" to Language Ability after Controlling for Potential Channels

6.1.1 Schooling as a Channel

Bleakley and Chin (2004) discusses that schooling could be a channel of the effect of language ability on earnings suggesting that individuals with poorer language abilities could suffer from a higher cost of education. To test this idea, I first regress years of schooling on language ability. Again, the OLS estimate of the impact of language ability on educational attainment could be biased as in the case of earnings. Therefore, I use 2SLS model with the Eq.(4.4) with the dependent variable of years of schooling. Table 6.1 shows that language ability has a causal impact on years of schooling among immigrants. According to the IV results, one point increase in the standardized literacy score causes 2.9 additional years of schooling. The 2SLS estimate is nearly twice the OLS estimate of 1.6 additional years of schooling. This finding is in line with the findings of Bleakley and Chin (2004). The results in Table 6.1, however, show that controlling for education does not change the effect of language ability on earnings. This suggests that education is not the channel in the set of European countries, in contrast to the findings of Bleakley and Chin (2004) but parallel to the findings of Guven and Islam (2015). The magnitude of the language ability coefficient does not decrease after controlling years of schooling. Guven and Islam (2015) also finds that the coefficient of language ability remains similar after controlling for schooling levels of immigrants. One possible explanation could be that the returns to schooling differs between European countries and the US.

6.1.2 Labor Market Skills as Channels

The previous part demonstrates that language ability has a causal impact on labor market skills of writing, reading, influencing others at work, as well as abstract and manual skills. In this section, I analyze the idea that these labor market skills may be mediators of language ability, similar to idea of how schooling could be a channel for the effect of language ability on earnings. An individual with a strong command of language ability may tend to work in a position that requires a high level of specific labor market skill such as reading skill at work. This might lead to an increase in earnings for this individual. As a result, it's critical to look into whether there's any kind of channel of language ability from this perspective.

Table 6.2 shows the baseline result with an additional regressors of labor market skills. Columns 1 and 2 show the results with the labor market skill of reading at work. OLS estimation of the effect of language ability on earnings decreases to 11 percent from 14 percent in the main specification. It also shows that reading at work index is positively associated with earnings of immigrants. On the other hand, according to the 2SLS estimation, the effect of language skills is 0.43 percent whereas there is no effect of reading at work on earnings. This implies that even if two immigrants are similar in terms of the intensity of reading at work, they will receive different financial rewards because of differing literacy scores. Furthermore, aside out of its effect from language skills, reading ability has no effect on earnings. In other words, although language ability has a causal impact on reading at work task, this task variable has no impact when it is controlled in the earning analysis. Columns 3 and 4 demonstrate the results with the labor market skill of writing at work. According to the 2SLS estimation, the effect of language ability on earnings is 0.42 percent. Similar to the case of reading, the effect of writing skill is 0 after I control for language ability and other control variables. On the other hand, OLS estimate suggests that writing skill is positively associated with earnings. Similar to the case of reading ability, language proficiency is awarded differently at the same degree of writing ability. Column 5 and 6 illustrate the results with the labor market skill of influencing others at work. According to the OLS estimation, increase in the literacy score by one standard deviation is associated with a 11 percent increase in earnings and a point increase in the influence index is correlated with 0.05 increase in earnings. On the other hand, according to the 2SLS estimate, a unit increase in standardized test score results in a 41 percent rise in earnings whereas there is no impact of influencing others on hourly earnings. Looking at the case of abstract skills, I observe that the causal effect of language ability remains 43 percent and the effect of abstract skills is 0. OLS estimation illustrates very similar pattern. One

standardized score increase is associated with 10 percent increase in earnings and abstract skill is also positively correlated with earnings. The case of manual skill also shows similar trends. To sum up, the results in this part shows that controlling for writing, reading, influencing others at work, abstract and manual labor market skills does not change the effect of language ability on earnings as shown in the 2SLS estimation. However, the effect of these skills disappear. In brief, these skills does not appear to be the channels of the effect of language ability on earnings where language ability has a causal impact on these labor market skills.

Table 6.1 The Effect of Language Ability on Earnings after Controlling for Schooling

Dependent Variable:	Years of Schooling	Years of Schooling	Years of Schooling	Years of Schooling	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings
Mean:	(11.84)	(11.84)	(11.84)	(11.84)	(2.74)	(2.74)	(2.74)	(2.74)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Language Proficiency	1.626***	2.866***	1.645***	2.788***	0.086***	0.531	0.077**	0.512
	(0.223)	(0.664)	(0.205)	(0.600)	(0.020)	(0.415)	(0.023)	(0.413)
Max(0, AAA-11)	-0.123**	-0.038			0.001	0.015		
	(0.041)	(0.077)			(0.006)	(0.018)		
Years of Schooling	-	-	-	-	0.035***	-0.042	0.036***	-0.038
					(0.005)	(0.072)	(0.008)	(0.069)
First Stage								
Max(0, AAA-11)*LD		-0.002***		-0.002***		-0.002***		-0.002***
		(0.001)		(0.001)		(0.001)		(0.001)
t-statistics		-3.393		-4.273		-2.821		-3.560
Observations	650	650	650	650	650	650	650	650
R-squared	0.686	0.615	0.708	0.653	0.631	0.370	0.636	0.415
AAA FE	-	-	YES	YES	-	-	YES	YES

Note: Each regression includes a gender dummy variable(female = 1), age, and dummy variables for region of birth and host country.
Differences in age and gender effects in host countries are accounted by including interactions of host country dummies with age and gender.
Standard errors are clustered at the host country levels in parentheses *** p<0,01, ** p<0,05, * p<0,1.

Table 6.2 The Effect of Language Ability on Earnings after Controlling for Labor Market Skills

Dependent Variable:	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings
Mean:	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Language Proficiency	0.106***	0.430*	0.114***	0.421**	0.111***	0.413*	0.096***	0.425*	0.099***	0.431*	0.116***	0.420**	0.085***	0.438*	0.083***	0.422
	(0.018)	(0.235)	(0.014)	(0.214)	(0.014)	(0.237)	(0.013)	(0.251)	(0.016)	(0.250)	(0.015)	(0.206)	(0.011)	(0.262)	(0.010)	(0.264)
Max(0, AAA-11)	0.003	0.016	-0.002	0.017	0.003	0.017	0.004	0.016	0.003	0.016	0.001	0.017	0.004	0.015	0.005	0.015
	(0.005)	(0.015)	(0.005)	(0.019)	(0.004)	(0.016)	(0.004)	(0.014)	(0.004)	(0.016)	(0.004)	(0.017)	(0.004)	(0.015)	(0.004)	(0.013)
Reading	0.048***	-0.019					0.023	-0.020							0.019	-0.008
	(0.013)	(0.045)					(0.018)	(0.033)							(0.014)	(0.023)
Writing			0.052**	-0.021			0.017	-0.017							0.013	0.005
			(0.017)	(0.054)			(0.022)	(0.039)							(0.028)	(0.025)
Influencing					0.048***	-0.001	0.033***	0.012							0.033**	0.030***
					(0.007)	(0.038)	(0.007)	(0.017)							(0.012)	(0.009)
Abstract									0.055***	-0.023			0.050***	-0.022	0.001	-0.042
									(0.010)	(0.059)			(0.012)	(0.057)	(0.024)	(0.042)
Manual_D											-0.128*	0.040	-0.105	0.032	-0.104	0.022
											(0.059)	(0.083)	(0.062)	(0.068)	(0.060)	(0.066)
First Stage																
Max(0, AAA-11)*LD		-0.002**		-0.002***		-0.002**		-0.001**		-0.002**		-0.002***		-0.001**		-0.001**
		(0.001)		(0.000)		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)
t-statistics		-2.429		-3.667		-2.230		-2.306		-2.300		-2.911		-2.225		-2.136
Observations	650	650	650	650	645	645	645	645	650	650	650	650	650	650	645	645
R-squared	0.627	0.466	0.624	0.474	0.629	0.483	0.633	0.474	0.629	0.465	0.623	0.474	0.634	0.458	0.638	0.479

Note: Each regression includes a gender dummy variable(female = 1), age, and dummy variables for region of birth and host country.
Differences in age and gender effects in host countries are accounted by including interactions of host country dummies with age and gender.
LFP denotes labor force participation. Standard errors are clustered at the host country levels in parentheses *** p<0,01, ** p<0,05, * p<0,1.

6.1.3 The Main Analysis with the Critical Age of 10

In this thesis, critical age is defined as 12 for the estimation in the main analysis. Bleakley and Chin (2010) conduct the study using a critical age of ten instead of the twelve used by Bleakley and Chin (2004). Therefore, it is important to repeat the same analysis using the critical age of 10. The results of this estimation are presented in Table 6.3. It displays both the OLS and IV estimates of the effect of language ability on earnings and occupational prestige score of childhood immigrants. The main findings are in accordance with the initial findings. According to the OLS estimate, a one-standard-deviation increase in literacy is related with a 14 percent increase in hourly earnings, but the IV model predicts a 40 percent increase in hourly earnings. It is important to note, however, that the F-statistics values are substantially larger than the results with the critical age of 12. For example, the F-statistics of the fixed effect model with a critical age of ten is 40, but the F-statistics with a critical age of twelve is 18.3. The IV coefficients are larger than the OLS coefficients once again. The findings for the non-pecuniary return of occupational prestige score are also similar. Language ability has a causal influence on occupational prestige, according to the IV estimation in column 8. With a stronger instrument compared to the case with the critical age of 12, a one-point improvement in the standardized reading score increases occupational prestige score by 6.5 points. The F-statistics are 25.9.

6.1.4 The Main Analysis with no Clustering

In the main specification, standards errors are clustered at the host country levels. For robustness check, Table 6.4 shows the results without any clustering. One can observe that the causal effect of language ability is 41 percent with the same significance levels of 99 percent. On the other hand, the effect of language ability on non-pecuniary return of occupational prestige is not significant when there is no clustering for standard errors.

Table 6.3 The Effect of Language Ability on Pecuniary and Non-Pecuniary Returns

Dependent Variable:	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Occupational Prestige	Occupational Prestige	Occupational Prestige	Occupational Prestige
Mean:	(2.74)	(2.74)	(2.74)	(2.74)	(39.4)	(39.4)	(39.4)	(39.4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Language Proficiency	0.137***	0.394***	0.137***	0.401**	9.992***	6.508*	10.06***	6.426**
	(0.018)	(0.153)	(0.021)	(0.161)	(0.803)	(3.429)	(0.756)	(3.222)
Max(0, AAA-9)	-0.002	0.013			-0.212	-0.356***		
	(0.003)	(0.012)			(0.140)	(0.109)		
First Stage								
Max(0, AAA-9)*LD		-0.001***		-0.001***		-0.001***		-0.001***
		(0.001)		(0.001)		(0.001)		(0.001)
t-statistics		-4.581		-6.323		-3.932		-5.091
Observations	650	650	650	650	564	564	564	564
R-squared	0.615	0.499	0.621	0.508	0.610	0.596	0.645	0.631
AAA FE	-	-	YES	YES	-	-	YES	YES

Note: Each regression includes a gender dummy variable(female = 1), age, and dummy variables for region of birth and host country. Differences in age and gender effects in host countries are accounted by including interactions of host country dummies with age and gender. Standard errors are clustered at the host country levels in parentheses *** p<0,01, ** p<0,05, * p<0,1.

Table 6.4 The Effect of Language Ability on Pecuniary and Non-Pecuniary Returns

Dependent Variable:	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Occupational Prestige	Occupational Prestige	Occupational Prestige	Occupational Prestige
Mean:	(2.74)	(2.74)	(2.74)	(2.74)	(39.4)	(39.4)	(39.4)	(39.4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Language Proficiency	0.138***	0.411***	0.137***	0.406***	9.979***	6.542	10.058***	5.618
	(0.024)	(0.137)	(0.025)	(0.133)	(1.008)	(4.773)	(1.038)	(4.767)
Max(0, AAA-11)	-0.002	0.017			-0.392	-0.563		
	(0.011)	(0.014)			(0.460)	(0.429)		
First Stage								
Max(0, AAA-11)*LD		-0.002***		-0.002***		-0.002***		-0.002***
		(0.001)		(0.001)		(0.001)		(0.001)
t-statistics		-3.040		-3.146		-3.047		-3.024
Observations	650	650	650	650	564	564	564	564
R-squared	0.615	0.483	0.621	0.503	0.611	0.597	0.645	0.624
AAA FE	-	-	YES	YES	-	-	YES	YES

Note: Each regression includes a gender dummy variable(female = 1), age, and dummy variables for region of birth and host country. Differences in age and gender effects in host countries are accounted by including interactions of host country dummies with age and gender. Standard errors are clustered at the person level in parentheses *** p<0,01, ** p<0,05, * p<0,1.

7. CONCLUSION

Focusing on childhood immigrants and utilizing data from the PIAAC first round survey, I analyze whether language proficiency has a significant and causal impact on different labor market outcomes in a multinational European context, some of which have not been studied previously. Especially, labor market skills and experience gap have not been studied in an IV framework before. I find that language ability does not affect the probability of being in the labor force and being employed today. On the other hand, it has a considerable impact on the experience gap among childhood immigrants. Increase in literacy score decreases the time out of the labor market for individuals. Furthermore, the effect of one standardized literacy score is 41 percent increase in earnings for individuals who immigrated to the host country before the age of 18. The effect is much smaller when the endogeneity issue is not taken into account with the OLS estimation. This is similar to the finding of Bleakley and Chin (2004). As a non-pecuniary return, I find that having higher level of literacy score helps individuals ascend the ladder of occupational status score. This paper also provides results on unique labor market skills benefiting the richness of PIAAC dataset. I concluded that language ability has a causal impact on labor market skills of writing, reading and influencing others at work. Moreover, using task variable definitions commonly used in the literature, I find that language proficiency positively affects the employment of abstract skills at work while it has a negative impact on the employment of manual tasks. Finally, Bleakley and Chin (2004) suggest that the effect of language skills appears to be channeled through schooling. On the other hand, my findings reveals that the effect is not mediated by education or labor market skills. Language ability is still rewarded at the same level of education or labor market skill indices.

BIBLIOGRAPHY

Adsera, Alicia, and Barry R. Chiswick. 2007. "Are there gender and country of origin differences in immigrant labor market outcomes across European destinations?" *Journal of population economics* 20(3): 495–526.

Adsera, Alicia, and Mariola Pytlikova. 2015. "The role of language in shaping international migration." *The Economic journal (London)* 125(586): F49–F81.

Adserà, Alícia, and Ana Ferrer. 2021. "Linguistic Proximity and the Labour Market Performance of Immigrant Men in Canada." *Labour (Rome, Italy)* 35(1): 1–23.

AGASISTI, Tommaso, Geraint JOHNES, and Marco PACCAGNELLA. 2021. "Tasks, occupations and wages in OECD countries." *International labour review* 160(1): 85–112.

Bleakley, Hoyt, and Aimee Chin. 2004. "Language skills and earnings: Evidence from childhood immigrants." *The review of economics and statistics* 86(2): 481–496.

Bleakley, Hoyt, and Aimee Chin. 2010. "Age at Arrival, English Proficiency, and Social Assimilation Among US Immigrants." *American economic journal. Applied economics* 2(1): 165–192.

Borelli, Stefania. 2016. "How Workers' Skills Are Used at Work: A Multi-Country Comparison with PIAAC." 16(5).

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." 77(4): 531–553.

Chiswick, Barry R., and Paul W. Miller. 1995. "The Endogeneity between Language and Earnings: International Analyses." *Journal of labor economics* 13(2): 246–288.

Chiswick, Barry R., and Paul W. Miller. 1998. "English language fluency among immigrants in the United States." *Research in labor economics* 17.

Chiswick, Barry R, and Paul W Miller. 2001. "A Model of Destination-Language Acquisition: Application to Male Immigrants in Canada." *Demography* 38(3): 391–409.

Chiswick, Barry R., and Paul W. Miller. 2002. "Immigrant Earnings: Language Skills, Linguistic Concentrations and the Business Cycle." *Journal of population economics* 15(1): 31–57.

Chiswick, Barry R., and Paul W. Miller. 2005. "Linguistic Distance: A Quantitative Measure of the Distance Between English and Other Languages." *Journal of multilingual and multicultural development* 26(1): 1–11.

Chiswick, Barry R, and Paul W Miller. 2010. "Occupational language requirements and the value of English in the US labor market." *Journal of population economics* 23(1): 353–372.

Dustmann, C, and A.H.O van Soest. 2002. "Language and the Earnings of Immigrants." *Industrial labor relations review* 55(3): 473–492.

Dustmann, Christian, and Arthur van Soest. 2001. "Language Fluency and Earnings: Estimation with Misclassified Language Indicators." *The review of economics and statistics* 83(4): 663–674.

Dustmann, Christian, and Francesca Fabbri. 2003. "Language proficiency and labour market performance of immigrants in the UK." *The Economic journal (London)* 113(489): 695–717.

Hlavac, Marek. 2012. "English Language Proficiency and Occupational Choice: Evidence from Childhood Immigrants into the United States.".

Isphording, Clarke, Andrew and Ingo E. 2017. "Language Barriers and Immigrant Health." *Health economics* 26(6): 765–778.

Isphording, I. E., and M Sinning. 2012. "The returns to language skills in the US labor market." *IZA Discussion Paper* (7080).

Isphording, Ingo E. 2013. "Returns to foreign language skills of immigrants in Spain." *Labour (Rome, Italy)* 27(4): 443–461.

Isphording, Ingo E. 2014. "Disadvantages of linguistic origin—Evidence from immigrant literacy scores." *Economics letters* 123(2): 236–239.

Isphording, Ingo E., and Sebastian Otten. 2013. "The Costs of Babylon-Linguistic Distance in Applied Economics." *Review of international economics* 21(2): 354–369.

Isphording, Ingo E, and Sebastian Otten. 2014. "Linguistic barriers in the destination language acquisition of immigrants." *Journal of economic behavior organization* 105: 30–50.

Isphording, Ingo E, Marc Piopiunik, and Núria Rodríguez-Planas. 2016. "Speaking in numbers: The effect of reading performance on math performance among immigrants." *Economics letters* 139: 52–56.

Isphording, Otten. 2011. "Linguistic distance and the language fluency of immigrants." *Ruhr Economic Paper* 274.

Kossoudji, Sherrie A. 1988. "English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men." *Journal of labor economics* 6(2): 205–228.

Liao, C. W., Qu Y. Morgan R. 2010. "The relationships of test scores measured by the TOEIC listening and reading test and TOEIC speaking and writing tests." *Encyclopedia of Cognitive Science* 10(13): 1–15.

Newport, Elissa L. 2002. "Critical Periods in Language Development." *Encyclopedia of Cognitive Science. London: MacMillan Publishing Ltd./Nature Publishing Group* .

Swadesh, Morris. 1952. "Lexico-Statistic Dating of Prehistoric Ethnic Contacts: With Special Reference to North American Indians and Eskimos." *Proceedings of the American Philosophical Society* 96(4): 452–463.

Weinstein, Edwin A. 1967. "An Inquiry into Man's Faculty for Language: Biological Foundations of Language . Eric H. Lenneberg. Wiley, New York, 1967. 507 pp., illus." *Science (American Association for the Advancement of Science)* 156(3782): 1585–1586.

APPENDIX A

Table A.1 The countries in the PIAAC dataset

Countries	Number of Observations for the Analysis of Critical Age	Number of Observations for the Analysis of Labor Market Outcomes
Austria	157	72
Belgium	89	36
Chile	17	-
Czech Republic	50	-
Denmark	286	144
Finland	33	6
France	244	113
Germany	249	147
Greece	126	-
Ireland	87	-
Israel	495	-
Italy	30	-
Japan	1	-
Kazakhstan	142	-
Korea	1	-
Lithuania	26	-
Mexico	2	-
Netherlands	129	20
Norway	95	60
Poland	1	-
Russian Federation	61	-
Slovak Republic	9	-
Slovenia	145	-
Spain	125	-
Sweden	181	-
United Kingdom	148	52
Total	2929	650

APPENDIX B

In this appendix, I report combinations of labor market skill variables with definitions.

Table B.1 Combination of variables for the skill index of Writing

G_Q02a	Write letters memos or mails
G_Q02b	Write articles
G_Q02c	Write reports

Table B.2 Combination of variables for the skill index of Reading

G_Q01a	Read directions or instructions
G_Q01b	Read letters memos or mails
G_Q01c	Read newspapers or magazines
G_Q01d	Read professional journals or publications
G_Q01e	Read books
G_Q01f	Read manuals or reference materials

Table B.3 Combination of variables for the skill index of Influence

F_Q02b	Skill use work - How often - Teaching people
F_Q02c	Skill use work - How often - Presentations
F_Q02d	Skill use work - How often - Selling
F_Q02e	Skill use work - How often - Advising people
F_Q03b	Skill use work - How often - Planning others activities
F_Q04a	Skill use work - How often - Influencing people
F_Q04b	Skill use work - How often - Negotiating with people

Table B.4 Combination of variables for the skill index of Abstract

G_Q03h	Use advanced math or statistics
F_Q05b	Complex problems
G_Q02b	Write articles
G_Q02c	Write reports
G_Q01d	Read professional publications
F_Q04a	Influencing people
F_Q04b	Negotiating with people

For manual task index, I use the question of "How frequently working physically for a long period?". Individuals choose one of the following responses to the questions:

Taking the values 1 through 5, never – less than once a month – less than once a week but at least once a month – at least once a week but not every day – every day.

Table B.5 The Effect of Language Ability on Labor Market Skills with Age at Arrival Fixed Effects

Dependent Variable:	Reading	Reading	Writing	Writing	Influence	Influence	Abstract	Abstract	Manual	Manual	Manual_D	Manual_D
Mean:	(2.74)	(2.74)	(2.74)	(2.74)	(39.4)	(39.4)	(39.4)	(39.4)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Language Proficiency	0.662***	0.894**	0.432***	0.449*	0.576***	1.206	0.689***	0.869	-0.586***	-0.467	-0.170***	-0.253***
	(0.128)	(0.454)	(0.0911)	(0.254)	(0.125)	(0.754)	(0.125)	(0.572)	(0.0900)	(0.340)	(0.0223)	(0.0660)
First Stage												
Max(0, AAA-11)*LD		-0.002***		-0.002***		-0.002***		-0.002***		-0.002***		-0.002***
		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)
t-statistics		-3.393		-3.393		-3.381		-3.393		-3.393		-3.393
Observations	650	650	650	650	645	645	650	650	650	650	650	650
R-squared	0.582	0.574	0.530	0.530	0.545	0.491	0.572	0.566	0.543	0.541	0.516	0.504

Note: Each regression includes a gender dummy variable(female = 1), age, and dummy variables for region of birth and host country.
Differences in age and gender effects in host countries are accounted by including interactions of host country dummies with age and gender.
Standard errors are clustered at the host country levels in parentheses *** p<0,01, ** p<0,05, * p<0,1.

Table B.6 The Effect of Language Ability on Earnings after Controlling for Labor Market Skills

Dependent Variable:	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings	Log Hourly Earnings
Mean:	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)	(2.74)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Language Proficiency	0.107***	0.422*	0.114***	0.412*	0.110***	0.405	0.0969***	0.415	0.0996***	0.422	0.114***	0.414*	0.0847***	0.428	0.0825***	0.415
	(0.0184)	(0.245)	(0.0164)	(0.227)	(0.0150)	(0.249)	(0.0138)	(0.257)	(0.0177)	(0.262)	(0.0137)	(0.225)	(0.0117)	(0.277)	(0.00934)	(0.275)
Reading	0.0447**	-0.0178					0.0194	-0.0234							0.0146	-0.0132
	(0.0185)	(0.0454)					(0.0230)	(0.0371)							(0.0188)	(0.0268)
Writing			0.0523**	-0.0136			0.0188	-0.00885							0.0158	0.0103
			(0.0166)	(0.0495)			(0.0228)	(0.0335)							(0.0312)	(0.0245)
Influencing					0.0471***	0.000948	0.0336***	0.0135							0.0332**	0.0290***
					(0.00854)	(0.0372)	(0.00829)	(0.0173)							(0.0134)	(0.0100)
Abstract									0.0537***	-0.0179			0.0484***	-0.0171	0.000461	-0.0373
									(0.0108)	(0.0585)			(0.0114)	(0.0563)	(0.0277)	(0.0488)
Manual_D											-0.131*	0.0316	-0.110	0.0254	-0.109	0.0152
											(0.0652)	(0.0931)	(0.0670)	(0.0778)	(0.0607)	(0.0740)
First Stage																
Max(0, AAA-11)*LD		-0.00162***		-0.00177***		-0.00159**		-0.00151***		-0.00159**		-0.00170***		-0.00145***		-0.00139**
		(0.000525)		(0.000422)		(0.000644)		(0.000566)		(0.000636)		(0.000389)		(0.000553)		(0.000541)
t-statistics		-3.091		-4.196		-2.473		-2.661		-2.503		-4.365		-2.628		-2.574
Observations	650	650	650	650	645	645	645	645	650	650	650	650	650	650	645	645
R-squared	0.630	0.491	0.629	0.498	0.634	0.506	0.637	0.501	0.633	0.491	0.629	0.497	0.639	0.486	0.642	0.502

Note: Each regression includes a gender dummy variable(female = 1), age, and dummy variables for region of birth and host country.
Differences in age and gender effects in host countries are accounted by including interactions of host country dummies with age and gender.
LFP denotes labor force participation. Standard errors are clustered at the host country levels in parentheses *** p<0,01, ** p<0,05, * p<0,1.