



Munich Personal RePEc Archive

# **The effects of language skills on immigrant employment and wages in Italy**

Pieroni, Luca and d'Agostino, Giorgio and Lanari, Donatella

Department of Political Science, University of Perugia, Department  
of Economics, Roma Tre University, Department of Medicine,  
University of Perugia

25 January 2019

Online at <https://mpra.ub.uni-muenchen.de/91725/>

MPRA Paper No. 91725, posted 29 Jan 2019 16:22 UTC

# The effects of language skills on immigrant employment and wages in Italy

Luca Pieroni<sup>a,\*</sup>, Giorgio d'Agostino<sup>b</sup>, Donatella Lanari<sup>c</sup>

<sup>a</sup>*Department of Political Science, University of Perugia*

<sup>b</sup>*Department of Economics, Roma Tre University*

<sup>c</sup>*Department of Medicine, University of Perugia*

---

## Abstract

In this paper, we examine how the Italian language problems of immigrants affect their labour market performance using two hitherto unexploited immigration surveys recently published by the Italian Institute of Statistics. With respect to immigrants with good Italian proficiency, our empirical findings suggest that language problems reduce the employment rate by about 30%, and point estimates are even larger when evaluating job discrimination. Italian language skills also significantly affect the wages of immigrants. The point estimates suggest a wage gap of about 20% between immigrants with Italian proficiency and those without Italian proficiency, a magnitude that increases to 25% for male immigrants. Robustness checks confirmed our estimates.

**Keywords:** Immigrants, Language skills, Employment, Wages

**JEL codes:** J15, J20, J31

---

---

\*Corresponding author: Luca Pieroni, Department of Political Science, University of Perugia. Via Pascoli 20, Perugia. Email: luca.pieroni@unipg.it.

## 1. Introduction

Immigration has become an important socioeconomic and public policy issue in Southern Europe with the recent increase in migration. It is a fact that immigrants have often a weak labour market position. The evidence shows that the employment rate of immigrants is typically lower than that of natives and that their job opportunities mainly consist of low-skilled (and often precarious) work (Adserà and Pytliková, 2016).

Language skills are certainly an important aspect of the individual human capital of immigrants and a determinant of employment success. Job selection on employment (and/or job satisfaction) may be based on a lack of proficiency in the destination country's language, inducing the migrant to work in jobs that require a lower education level than the level achieved. This can lead to lower job performance and, in turn, amplifies differences in the employment rate and wages.

In this paper, we examine how the language problems of immigrants affect their labour market performance in terms of employment, wages and job discrimination. Our laboratory is Italy, and different from English or Spanish, immigrants' knowledge of Italian is much more limited<sup>1</sup>. The sparse estimates we found for employment outcomes are also controversial. Dustmann and Fabbri (2003) found a decrease of 22% in employment probability in the UK for immigrants who lack English skills. Gonzales (2010) found significant negative effects of the lack of host country language skills on employment in Spain, while Yao and van Ours (2015) found that language problems affect wages but not the employment probability of immigrants in the Netherlands. However, it is undeniable that immigrants' lack of language proficiency can foster discrimination by natives, and contributes to ghettoisation and social isolation in the workplace (Zschirnt and Ruedin, 2016). Although the language fluency of immigrants is just one component of a larger integration process for migrants, a lack of proficiency is generally correlated to perceived job (non)satisfaction as a result of perceived job discrimination. Therefore, significant effects are expected. However, the consequences of language proficiency that have received the most attention in the labour market relate to the effect on wages. Chiswick and Miller (2015) motivated it for the role of wages in summarizing economic status, as well as the historical availability of data on wages and language proficiency of immigrants in several developed countries. Empirical evidence has unanimously shown a significant relationship between the language proficiency of immigrants and wages, varying from 10% – 20% percent for the US, the UK and Australia and slightly less for Germany and Spain, although these point estimates increased when unobserved heterogeneity and measurement errors were included<sup>2</sup> (Yao and van Ours, 2015; Guven and Islam, 2015).

We contribute to the debate, given the recent availability of Italian data on labour market outcomes including immigrants' language proficiency. We base our analysis on data from two unexploited Italian surveys on immigrants published by the Italian Institute of Statistics (IIS): the Conditions and Social Integration of Foreign Citizens that was conducted between 2011 and 2012 - and published in 2014 (CSIFC 2011-2012), and the Income and Living Conditions of Households with Foreigners that was conducted in 2009 and refers to the year

---

<sup>1</sup>See Chiswick (2008) for a discussion on the importance to study less common languages.

<sup>2</sup>See Table 5.5. of Chiswick and Miller (2015) for extended analysis of the empirical estimates.

2008 for wages (ILCHF 2009). Both surveys include questions on the social and economic conditions of interviewees and measures of language proficiency. The data collected from both surveys are partly overlapping and complementary. For instance, the CSIFC 2011-2012 survey includes relevant information about immigrants' household members in the country of origin and in Italy, but limits to employment information about the labour market outcomes and individual perceptions of job discrimination. We also exploit the ILCHF 2009, the first nation-wide survey on the socioeconomic conditions of the foreign population living in Italy<sup>3</sup>, which contains information about the employment and wages of immigrants, although it does not include items relating to job discrimination. Both datasets include self-reported language proficiency, even if the information from interviewees does not overlap completely. In the CSIFC 2011-2012, the self-reported language proficiency is measured using ordered modalities that records writing and reading problems, whereas the ILCHF 2009 registers a general perception of knowledge of the Italian language. Despite this limitation, a plausible aggregation over modalities of language proficiency provides interesting insights into the relationship between language and labour market performance, comparing the evolution of these effects on employment probability between 2009 and 2012.

Our empirical strategy exploits the assumption that the language proficiency of migrants is known to be age-different. Adult immigrants typically make the decision to migrate after they have obtained their education in the country of origin and are mainly motivated by the absence of employment opportunities, while younger immigrants come to destination countries for different reasons and typically learn the language of the host country quickly. This difference in acquiring language knowledge determines a different measurement bias of the host country language, which overestimates the effective ability of adult immigrants in terms of labour market performance. Therefore, the key challenge for causal estimates is that age at arrival may conceivably correlate with language acquisition, and it explains why the literature generally uses age at arrival as an instrumental variable (IV) to investigate the relationship between the language skills of immigrants and labour market performance. We extend our analysis to account for unobserved heterogeneity due to the endogenous choice of language acquisition and propose propensity score-matching IV estimator (PSM-IV) that helps in estimating the 'true effects' of language skills on labour market performance.

Based on these considerations, we first show that a lack of Italian language proficiency is associated with a significantly large decrease in employment probability, irrespective of the dataset used. Our findings mirror those of existing studies, with employment reduction ranging from 20% – 30% for immigrants with language problems. Second, we conduct a conceptually similar exercise analysing whether a lack of language proficiency reflects workplace discrimination. We measure this outcome as self-perceived job discrimination. Our findings support the evidence that improving host country language proficiency helps to alleviate workplace discrimination by more than 40%. Third, we focus on the effects of language proficiency on wages. We find that a good knowledge of the host country

---

<sup>3</sup>It is worth noting that the definitions of immigrant and foreigner do not completely overlap because foreigners include individuals who are born in Italy but do not have Italian citizenship, while some immigrants are individuals who are born abroad and have moved to Italy, which may also be the case for some Italian citizens. Although the survey refers to foreigners, in this paper we use the two terms, immigrants and foreigners, interchangeably.

language significantly increases the expected wages of immigrants (23% in the IV baseline estimates.) However, some aspects of the wage gap are persistent, even considering a scenario in which all immigrants speak good Italian. This means there are other factors that explain the difference in wages between Italians and immigrants. We also extend our analysis to investigate factors influencing immigrants' knowledge of the host country language, such as gender. While there is a large body of literature showing the different effects of gender on the relationship between language skills and labour market performance (e.g., Yao and van Ours (2015)), we emphasise that sexual selection into occupations helps to explain the different effects of this relationship. Robustness analyses validated our main findings.

The remainder of this paper is organised as follows. Section 2 describes the background of immigration in Italy and the evolution of immigrants' employment and wages in Italy. Section 3 discusses the data and presents descriptive statistics of the surveys used in the empirical section, while Section 4 illustrates the estimation strategy. Baseline results are discussed in Section 5. Estimates of robustness are presented in Section 6, and Section 7 concludes.

## 2. Background

### *2.1. The patterns of immigration in Italy*

For almost a century after its unification, Italy was one of the leading European countries in terms of emigration. About twenty-six million Italians went abroad to overcome the poverty that changes in the demographic, economic and social structure had produced. The long period of mass emigration stopped in the second half of the 1970s, mainly because of the restrictive policies implemented in the traditional receiving countries after the international oil crisis. This marked the moment of Italy's passage from an emigration country to an immigration country. Even though, at the beginning of this change in migration net flow, national returnees who came back to Italy characterized the positive immigration flows, the arrival of foreigners steadily increased during the 1980s. This immigration was justified by the growing push factors in the origin countries (e.g., economic and social disparities, conflicts, poverty, discrimination and persecution) and by pull factors (e.g., sustained economic growth of Italy) that attracted individuals seeking for better economic opportunities, more jobs and, in turn, the promise of a better life.

The public's initial tolerant attitude towards immigration and the weak response to governing immigration helped to consolidate regular immigration flows based on labour market quotas<sup>4</sup>(Bonifazi, 2009). Starting from the Turco-Napolitano law in 1998, the Italian government was called on set a maximum quota of foreigners allowed to enter for work reasons. For example, the number of residence permits had to be proportional to the needs of the Italian labour market and to those already issued for family reunification or for reasons of temporary social protection. Once the quotas were set, a date was announced for employers to start filing applications to sponsor an immigrant. The recruitment of foreign workers within this legal framework never worked properly and mainly focused on immigrants who were already

---

<sup>4</sup>Italy based immigration control on labour market quotas that were introduced in 1995 and applied extensively starting from 1998.

living and working in Italy, implying that the number of applications was higher than the quotas (Finotelli and Arango, 2011). The rationing of residence permits, which arose due to increasing rhetoric against immigration, encouraged illegal immigration and, in turn, the recurring need for regularisation programs (Mastrobuoni and Pinotti, 2015).

The first attempt to regulate the entry, residence and employment procedures for immigrants occurred in 1986 (Law 943). Successive regularisations in 1990, 1995 and 1998 had the same aim of regularising workers and all immigrants who could prove they were living in Italy before the law came into force. Differently, the regularisation introduced in 2002 (Law 189/2002) applied more restrictive eligibility rules linking permits to stay with work contracts and made procedures for renewals more expensive. In fact, only those working in families or working in companies could apply, thus excluding the self-employed, the unemployed and family members. More than 705,000 applications were received and nearly 647,000 were accepted, which contributed to the growth in regular immigration from 1.8 million to almost 4.0 in the five-year period 2003-2008<sup>5</sup>.

Currently, Italy is one of the main immigration countries in Europe, with many foreign residents. The majority of foreign nationals residing in Italy (more than five million in 2017) are extra-European citizens, although immigrants arriving from Central and Eastern Europe (e.g. Romania) from the second half of the 1990s onwards represent the most important group. The immigration flow was encouraged when Romania entered the European Union (January 2007), so that Romanian immigrants, along with Bulgarian, acquired the right to reside and work in Italy; thus, 'irregular' immigrants became 'regular' immigrants. Now, Romanians represent the 23.3% of the total immigrants in Italy, while Albanians are the second most important immigrant group representing 10% of all immigrants. Emigration from Northern Africa, particularly Morocco, represents the third largest group at 8%. Moreover, data on immigrants in the last 20 years shows that emigration from Asia (China, India, Philippines) and Latin America (Peru, Ecuador) also increased sharply (IIS, various years).

## *2.2. Economic outcomes for immigrants in the Italian labour market: employment and wages*

Italy constitutes an ideal case study for our research question to evaluate the effect of the language proficiency on the labour market performance of immigrants because of the unique structural and institutional characteristics of its labour market based on segmentation and rigidity.

Table 1 lists the employment prevalence and wage means of immigrants and Italian workers by sector and, within manufacturing, by high-tech/low-tech sectors using the Labour Force Survey (LFS), which is redacted by the Italian Institute of Statistics (IIS). Compared to Italian workers, there is a higher employment rate among immigrants in household services

---

<sup>5</sup>There have been successive regularisation procedures. In 2006, the Italian government decided to issue a second decree law on immigration flows, which allowed all the 540,000 foreigners who had presented an application to enter the country for economic reasons to be regularly employed (e.g. second Prodi decree law on immigration flows). In addition, two additional extraordinary regularisations were implemented in 2009 and 2012. During these regularisations, migration businesses often played the role of intermediary by providing immigrants with fake contracts so they could obtain a residence permit and gain access to the formal labour market (e.g. Avallone 2017).

Table 1: Differences in labour market performance by sectoral employment (Mean of years 2015-2016)

	Agriculture	Manufacturing		Construction	Commerce	Services	Household services	Mean
		Low-tech	High-tech					
<b>Foreign</b>								
Employment (%)	4.97	13.67	6.52	9.13	9.22	37.50	18.96	
Wage (euros)	918	1208	1382	1249	1104	1089	738	1069
<b>Italian</b>								
Employment (%)	3.82	11.43	9.32	6.03	14.46	53.99	0.92	
Wage (euros)	1022	1359	1539	1353	1193	1385	572	1348
Wage gap (%)								20.67
Hourly wage gap (%)	12.81	18.14	22.28	14.08	14.12	22.08	26.06	18.55

*Notes:* Data are extracted by the Labour Force Survey and redacted by the Italian Institute of Statistics (IIS). The number of observations are 344,416 for Italian workers and 53,688 for foreign workers. The ATECO2007 classification is used to obtain disaggregation by economic sectors, and the OECD classification, to distinguish high-tech and low-tech sectors (Pieroni and Pompei, 2008).

(e.g. domestic services and personal assistants)<sup>6</sup> (18.96%), construction (9.13%) and low-skill sectors (13.67%). In many of these sectors, immigrants benefit from the seasonal nature of some activities, namely tourism and agriculture. Data analyses confirm that newly arrived migrants are mainly absorbed into specific segments of the labour market, those in which natives often refuse to work (Ponzo et al., 2015).

The data also show a stable differential wage in favour of native workers, irrespective of the employment sector. On average, Italian workers earned 278 euros more than immigrants in 2015-2016 (wages of 1368 euros for Italian workers against 1069 euros for immigrants). We also estimate overall and hourly wage gaps: Italian workers had wages and hourly wage that were, respectively, 20% and 18.5% higher than those of immigrants. The hourly wage gap was close to the 18.6% estimated by the Italian Institute of Statistics for 2014 (IIS, 2016)<sup>7</sup>. It is worth noting that in the high-tech manufacturing sector, we find a higher level of wages for natives (1539 euros), 155 euros more each month than immigrant wages. This may suggest that the tasks and responsibilities performed by foreign workers, on average, are not completely equivalent to those performed by natives, although they may have equivalent educational backgrounds.

Table 2 shows the differences in labour market performance by educational level for 2015-2016, and for 2011-2012 and 2008-2009, corresponding to the years the surveys were conducted. The recent wave shows that Italian workers with high education have more than 10 percentage points above the employment rate mean (68.4% vs. 57.9%). In relative terms, the employment rate for immigrants with high education is above the mean, even if the prevalence decreased 2.5 percentage points in six years. As expected, Italians with low education had an employment rate of 43% in 2015-2016 compared to 51.5% for immigrants. In addition, the data suggests a mismatch between the potential associated with the educational level of immigrants and their economic achievement. Higher educated immigrants have similar wages and lower hourly wages than Italian low-educated workers, although they

<sup>6</sup>The regularisation of Law 189/2002 accounted for a significant prevalence of women applicants from Romania, Ukraine, Moldavia, Poland and Ecuador working in domestic help and carers.

<sup>7</sup>The indicator is calculated as: (wages natives-wages immigrants)\*100/wages immigrants.

Table 2: Differences in labour market performance by educational level

	2015-2016			2011-2012			2008-2009		
	Low Education	High Education	Mean	Low Education	High Education	Mean	Low Education	High Education	Mean
<b>Foreign</b>									
Employment (%)	51.5	64.2	58.1	52.2	66.1	59.3	54.2	68.4	61.4
Wage (euros)	999	1,103	1,058	973	1,062	1024	967	1,060	1,019
<b>Italian</b>									
Employment (%)	43.0	68.4	57.9	41.6	67.2	55.3	43.1	69.3	56.2
Wage (euros)	1,186	1,432	1360	1,120	1,378	1292	1,079	1,325	1,235
Wage gap (%)	15.76	23.00	22.22	13.12	23.00	20.74	10.38	20.00	17.49
Hourly wage gap (%)	13.49	23.56	21.89	13.57	23.33	21.57	12.07	22.64	20.26

*Notes:* The table reports mean values of the labour market indicators by the Labour Force Survey (IIS, various years). It is worth noting that the column labelled "2008-2009" lists the value for wages at the end of the year 2008 and employment for the year 2009, respectively. People are classified as having low education when the highest education degree is pre-primary, primary or lower secondary and as having high education when the highest education degree is upper secondary, lower tertiary or upper tertiary.

are largely higher than the wages of low-educated immigrants.

The observed employment and wage gaps between natives and immigrants show some specificity when investigated by gender (Table 3). The employment rate by gender has remained stable over the years, although it is known that the unemployment rate increased with the downturn in the Italian economy<sup>8</sup>. We can show that the wage gap mean has increased compared to previous years, reaching 18% among men and 25% among women.

The data listed support the general idea that immigrant inequalities in the Italian labour market are a consequence of the selective allocation into low-status jobs, the mismatch between immigrant education and job achievement, and gender pay gaps. Some theories help to explain this disadvantage relative to natives and its persistence over time. For example, the segmented labour market theory suggests that the institutional framework differs greatly between high- and low-skill sectors. The first sector offering stable jobs, relatively high salaries, acceptable work conditions and upward mobility, while the low-skill is often characterised by unstable jobs with low pay, poor working conditions and limited prospects for promotion (McGovern, 2007), such that shortages of labour supplied by natives in the secondary sector are eliminated - at least in the short term - by immigrant workers.

However, the lack of country-specific skills on arrival can largely explain the differences in labour market performance. Limited knowledge about the functioning of the labour market or, more importantly, the lack of fluency in the host country language may represent an obstacle for immigrants in the economic assimilation process and to finding better job opportunities<sup>9</sup> (Chiswick and Miller, 2003; Dustmann and Fabbri, 2003). As argued above,

<sup>8</sup>Immigrant men experienced a doubling of unemployment (e.g. from 7.8% in 2009 to 15.1% in 2015-2016).

<sup>9</sup>This disadvantage has been attributed to the difficulties of immigrants, upon their arrival in the host country, in transferring formal schooling, experience, and training obtained overseas (Chiswick and Miller,



Table 3: Differences in labour market performance by gender

	2015-2016			2011-2012			2008-2009		
	Men	Women	Mean	Men	Women	Mean	Men	Women	Mean
<b>Foreign</b>									
Employment rate (%)	65.8	49.2	58.1	66.8	48.3	59.3	67.9	48.1	61.4
Wage (euros)	1,220	898	1,058	1,177	853	1024	1,156	853	1,019
<b>Italian</b>									
Employment rate (%)	66.2	50.0	57.9	65.3	45.7	55.3	67.3	45.5	56.2
Wage (euros)	1,491	1,211	1,360	1,413	1,148	1,292	1,349	1,092	1,235
Wage gap (%)	18.17	25.91	22.22	16.70	25.69	20.74	14.31	21.88	17.48
Hourly wage gap(%)	18.26	25.92	22.04	17.48	26.12	21.81	17.01	24.17	20.57

*Notes:* The table reports mean values of the labour market indicators for the three periods. It is worth noting that the column labelled "2008-2009" lists the value for wages at the end of the year 2008 and employment for the year 2009, respectively. The data are extracted by the Labour Force Survey redacted by the Italian Institute of Statistics (IIS).

this central determinant is reinforced in Italy by the cyclical regularisation and establishment of temporary resident permits that could have led skilled migrants to decide not to invest in a long search for higher-status jobs or in acquiring the language proficiency relevant to the host country, as they were unlikely to remain in Italy long enough to enjoy a return on these investments<sup>10</sup> (Kalter and Kogan, 2006).

Clearly, Italian language fluency can explain not only the differences in the economic success of immigrants but also discrimination in the labour market. There is substantial literature documenting the extent of labour discrimination against immigrants and ethnic minorities in different countries (Heath and Cheung, 2007). Individuals from different cultures may suffer a different treatment than natives because employers may prefer applicants with the same culture and language or with a higher affinity or lower social distance (Ebner and Helbling, 2016).

To empirically test the hypothesis that the language abilities of immigrants in Italy helps explaining the differences in labour market outcomes, we detail in the next paragraphs the sources and representativeness of our datasets, which include unexploited data of Italian language skills.

2009; Clark and Drinkwater, 2008; Friedberg, 2000).

<sup>10</sup>The absence of any particular rules controlling the entry of foreign workers can be explained by the fact that immigration, in its initial phase, was considered a temporary phenomenon that would not involve large numbers of immigrants since Italy was seen as only a stage on the journey on their way towards the final destination to the traditional European immigration countries.

### 3. Data

#### 3.1. Data sources

Two national surveys were used to estimate the effect of Italian language skills on the labour market outcomes of immigrants. The first survey is the CSIFC 2011-2012, published by the IIS in 2014. The reference population of this survey were immigrants permanently or temporarily living in Italy between May 2011 and December 2012<sup>11</sup>. A sample of 12,000 households living in about 800 Italian cities was used for the interviews. The second is the ILCHF 2009, financed by the Italian Minister of Labour and Social Policies and conducted by the IIS in 2009. This survey made use of the methodological framework of the survey on Income and Living Conditions, carried out yearly in 27 EU countries (plus Norway and Iceland), and coordinated by Eurostat (e.g., EU-SILC). The questionnaire, data collection and correction procedures, as implemented and improved by EU-SILC, have been adapted to the specific needs implied by the foreigners' survey.

The final datasets were adjusted for some data issues. First, we considered only immigrants aged 15-64, excluding those who retired in 2011 and 2008, respectively, for the CSIFC and ILCHF. Second, although the survey included information on Italians living in households with immigrants, we excluded them from the successive analysis because they are not representative of the Italian population. Using this strategy, we obtained a CSIFC dataset of 17,298 immigrants, 14,990 of which were first-generation immigrants, whom we mainly considered in our investigation<sup>12</sup>. The ILCHF dataset included 8877 first-generation immigrants.

Table 4 lists the composition of samples by groups of immigrants' countries of birth compared with the Population Census (2011) and adjusted for changes in 2012 by official demographic statistics (e.g., the average of resident immigrants between 2011 and 2012). It appears that both surveys have similar group composition as the 2011 Census. Only the American immigrant group appears to be slightly over-represented in the ILCHF survey (16.8% vs. the 7.9% of the Population Census).

#### 3.2. Variables

Both datasets provide information about the sociodemographic characteristics of immigrant household members' in the country of origin and in Italy, outcomes of employment and, for our interest, language skills. In addition, the surveys list other interesting labour market outcomes. The CSIFC reports self-perceived work discrimination of immigrants, asking "*Have you ever been discriminated (treated differently with respect to the other workers) during your last job?*". The ILCHF reports individual monthly wages [*log wages* and *log hourly wages*] and, as in the EU-SILC, the amount refers to the previous calendar year (i.e. 2008). Table 5 shows the descriptive statistics for the outcomes of interest included in our analysis by gender and Italian language knowledge.

---

<sup>11</sup>The interviews were extended until February 2013 extending the sample for some big cities (e.g., Milan, Rome and Naples).

<sup>12</sup>Here, second-generation immigrants are defined as those born in Italy of at least one foreign citizen or those born abroad but who completed a cycle of study in Italy.

Table 4: Composition of samples by country of birth

Immigrant groups	CSIFC survey (2011-2012)		ILCHF survey (2009)		Pop. Census (2011)	
	N.	%	N.	%	N.	%
F-Y*, Albany, Romania	5649	38.80	2930	33.00	1,613,015	37.40
Other-Europe	2961	20.30	1557	17.50	693,388	16.20
Asia	2118	14.50	1161	13.10	742,994	17.20
China & India	(774)	(5.30)	(490)	(5.50)	(323,221)	(7.50)
Africa	2840	19.50	1733	19.50	918,515	21.30
Tunisia & Morocco	(1785)	(12.30)	(1030)	(11.60)	(525,189)	(12.20)
America	976	6.70	1,496	16.8	342,718	7.90
Total	14,544		8,877		4,569,317	

*Notes:* The table compares the percentage of immigrants by country of origin according to the CSIFC survey of 2011-2012, the ILCHF survey of 2009 and the average of foreign population as extracted by the population census of 2011 and population flows in the 2012. Subsamples of immigrants by continents in parentheses. \*F-Y: Former Yugoslavia

Table 5: Descriptive statistics, outcome variables and predictors

	CSIFC survey (2011-2012)				ILCHF survey (2009)			
	Men		Women		Men		Women	
	No	Yes	No	Yes	No	Yes	No	Yes
<b>Labour Market Outcomes</b>								
Employment rate (%)	65.40	63.20	57.40	52.40	67.40	58.60	56.30	48.50
Observations	4922	1720	6383	1922	2633	1717	2787	1741
Job discrimination (%)	78.10	86.20	77.80	83.9				
Observations	4692	1615	5123	1041				
Log wages					7.295	7.083	6.827	6.709
Observations					1799	1157	1455	815
Log hourly wages					2.205	2.038	1.963	1.854
Observations					1799	1157	1455	815
<b>Italian Language Knowledge</b>								
Difficulty in speaking Italian (often or sometimes)		26.95		24.05				
Difficulty in reading Italian (often or sometimes)		36.45		32.58				
Language skills (Poor)						8.78		8.50
Language skills (Sufficient)						30.69		29.95
Language problems (LP)		26.95		24.05		39.40		38.40
Observations		1720		1922		1104		1738

*Notes:* In this table, we compare the labour market outcomes of the CSIFC 2011-2012 and the ILCHF 2009. In the CSIFC survey (2011-2012), language problems (LP) is defined as a dummy variable, which equals one if the individual answered that he/she often or sometimes had problems either in speaking or reading and zero otherwise (few times, never). In the ILCHF survey, language problems (LP) is defined as a dummy variable, which equals one if the individual answered that he/she had poor or sufficient language abilities and zero for those with discrete or good language abilities.

Concerning language abilities, for the CSIFC survey the household member answered the following questions: "*Is it difficult for you to read in Italian?*" and "*Is it difficult for you to speak in Italian?*". The answer choices were: often, sometimes, few times, never. Following Yao and van Ours (2015), we defined a dummy variable that equals one if the individual answered that he/she often or sometimes had problems either in speaking or reading and zero otherwise (few times, never). A four-modality framework for language skills is also used in the ILCHF survey, although it more generally asked respondents to evaluate whether their Italian is *Poor*, *Sufficient*, *Discrete* or *Good*. Also in this case, we obtained a binary outcome for language problems (LP) by collecting respondents who answered that they have poor or sufficient Italian language abilities (LP=1) and those who answered they have discrete or good language abilities.

Some important differences arise for employment outcomes among gender and language problem groups. Immigrant women have a lower probability of being employed than men. In particular, when the performance is measured by employment rate, the probability of women without linguistic problems compared to men without linguistic problems is 8 percentage points lower (65.4% for men and 57.4% for women). Within gender groups, employment differences can be also extended for women with and without linguistic problems. The probability of a woman with linguistic problems being employed is reduced by a further 5 percentage points (from 57.4% to 52.4%). These patterns are even more striking (e.g. 8 percentage points) using the ILCHF (2009) survey. We also found significant perceptions of job discrimination, irrespective of the language difficulties and gender differences. About 80% of the responses suggest job discrimination against immigrants.

Table 5 also reports ILCHF estimates of the hourly wages and log monthly wages (absolute values in parentheses). By restricting the sample to the first-generation immigrants, we may underestimate mean wages with respect to the labour force statistics. However, if we calculate the gender pay gap, the representativeness of our sample is maintained; the estimates of the pay gap using the ILCHF show that men earn 35% more than women, a figure close to that obtained using the labour force statistics (i.e. 34.6%)<sup>13</sup>.

Clearly, the groups of immigrants with language problems differ in terms of observable characteristics from those without language problems. These individual and household differences concern the immigrants' actual and past conditions, if they never studied in the country of origin and the reason they immigrated to Italy. In addition, other control variables may differ. We recorded the most common control variables, such as age, marital status, education degree, macro-region and area of residence; a more complete set of background variables including childcare in the household, and whether the woman in the household is an homemaker are included. To account for individual (unobserved) abilities obtained in the country of origin, which may affect the possibility of a good employment and wage in the destination country, we include a proxy for the immigrant has never worked in the country of origin. Descriptive statistics of these covariates by gender and linguistic problems are listed in Appendix A.

---

<sup>13</sup>The gender wage gap is estimated from the following data in columns 7-8 of Table 3:  $[(0.2188 - 0.1431)/0.2188]$ .

### 3.3. Descriptive statistics of the key variables

Our dataset is completed by creating the variable 'age at arrival'. This is calculated as the difference between the year in which the person immigrated to Italy and the year of birth. The kernel density plot in Figure 1 (a) shows the age at arrival of the first-generation immigrants for the CSIFC survey. The density appears to be different by gender, estimating a greater concentration for younger men, which is overcome by women after age 35. A statistical distribution that follows that of the CSIFC survey is also obtained using the ILCHF survey (2009) (panel b).

Figure 1: Kernel density plots of age at arrival

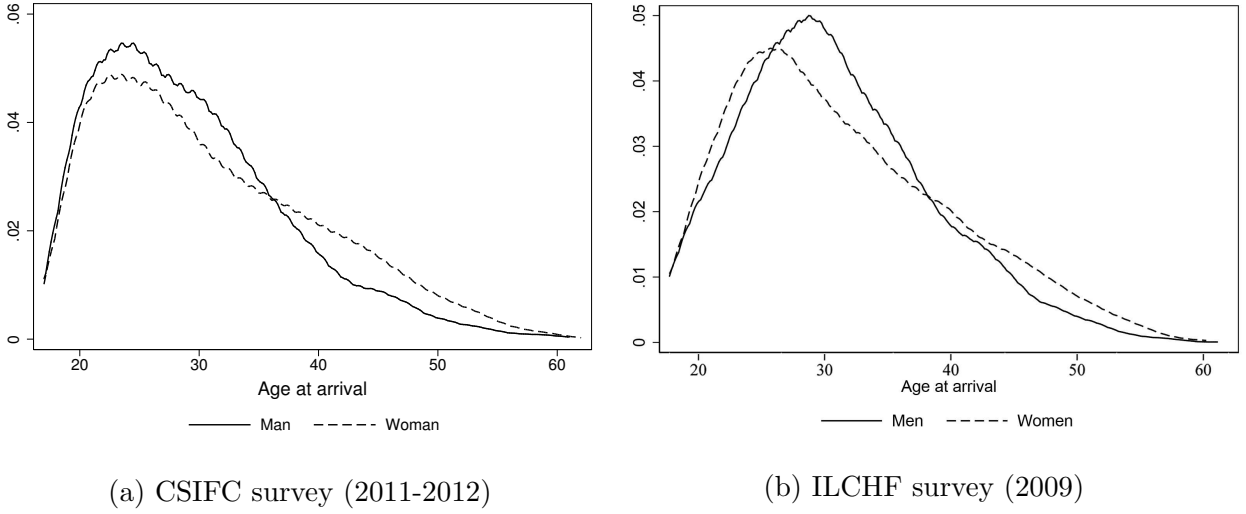
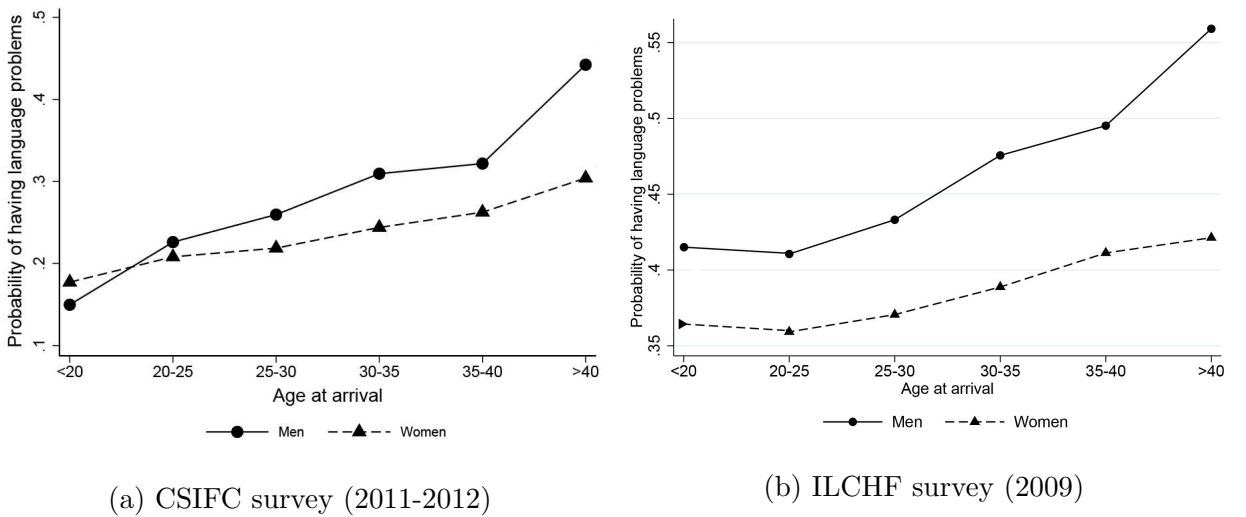
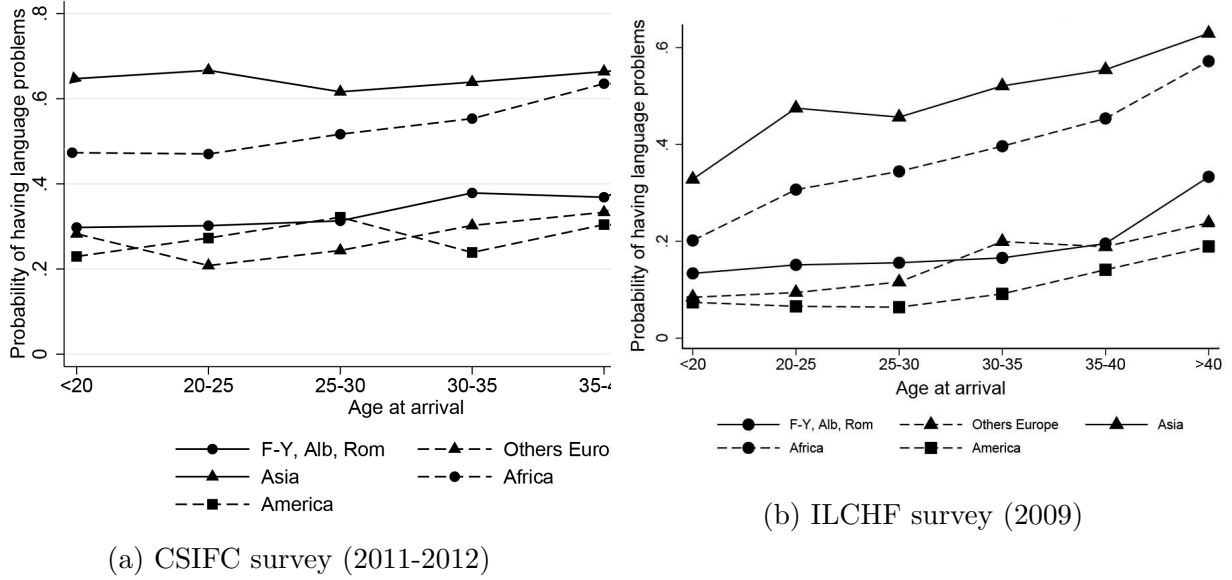


Figure 2: Probability of having language problems and age at arrival



This evidence confirms the argument that women generally arrive later in host countries for family reunification, while men mainly emigrate earlier for work opportunities. However,

Figure 3: Probability of having language problems based on age at arrival, heterogeneity



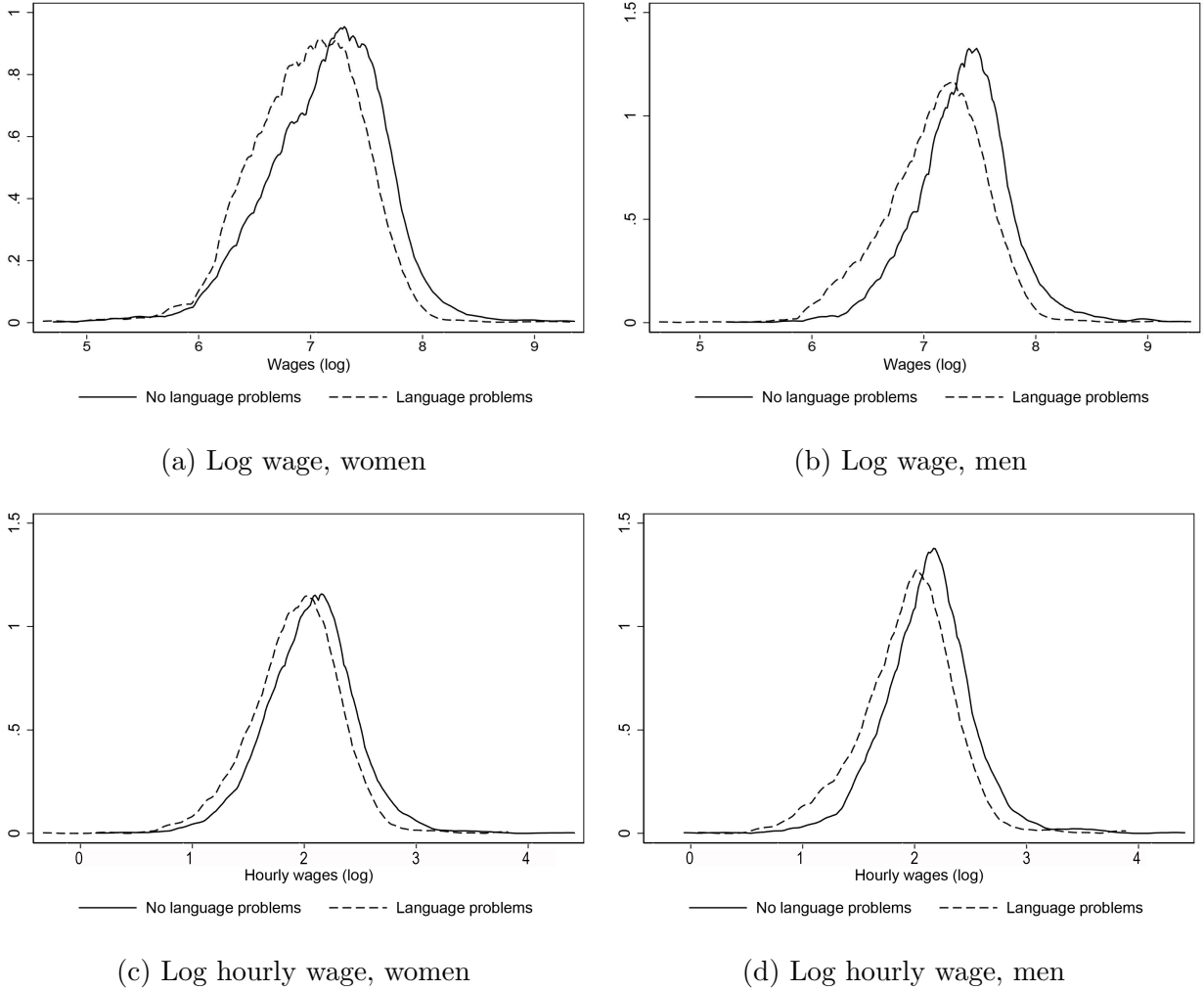
it is worth noting that in Italy this delay may be overestimated due to the cyclical regularisation discussed in Section 2. Indeed, the increasing demand of immigrant women working in household service sector as caregivers, and often employed without permits to stay in Italy, generated the expectation of illegal immigrants for successive regularisations (Salmasi and Pieroni, 2015).

Figure 2 shows the probability of having language problems based on age classes at arrival in both surveys. This figure shows increasing trends of having language problems steeper for men, irrespective of which we use. This evidence supports the hypothesis of a significant relationship between age at arrival and the language problems of immigrants and suggests that the latter variable may be a good instrument for estimating the causal relationship between language proficiency and labour market outcomes for the Italian case.

Clearly, the goodness of the instrument 'age at arrival' may be affected by the heterogeneous effect on the probability of having language problems. As argued by Hermansen (2017) Different groups of immigrants may also follow different patterns, which may be hidden if we use the aggregate variable of age at arrival of all immigrants, as an instrument to obtain estimates for these groups. Figure 3 suggests that the different immigrant groups mimic the patterns shown in Figure 2 of a positive relationship between age at arrival and Italian language problems.

Figure 4 shows kernel density plots of (log) wages and hourly wages conditionally to the language problems. As expected, immigrants with language problems have lower average hourly wages (and wages) than immigrants without language problems. Although this is true for both men and women, the gap is larger for men than women.

Figure 4: Kernel density, ILCHF survey (2009)



#### 4. Estimating the effect of differential language proficiency on labour market outcomes

The main premise of our analysis is that the labour market is less accessible to immigrants and less profitable based on low Italian language proficiency. This section provides evidence to support this premise.

The simplest way to formalise the relationship between language problems and the labour market performance of immigrants is through a regression model:

$$y_i = \alpha_1 + \alpha_2 LP_i + X' \alpha_3 + \epsilon_i \quad (1)$$

where  $i$  denotes each immigrant in the dataset,  $y_i$  denotes an indicator of employment, discrimination in the workplace or wage performances,  $LP_i$  is a dummy variable that is equal to one when individual  $i$  has language problems and  $X$  is the vector that includes all individual and household characteristics listed in Appendix A. The parameter of interest  $\alpha_2$  should measure how much less or more likely are labour market outcomes in immigrants

with language problems with respect to immigrants with no-language problems.

Table 6 illustrates the ordinary least squares (OLS) estimated coefficients for the two surveys separately and for men and women. The coefficients of employment in the CSIFC 2011-2012 survey are of the expected sign, although they are, in the great part, close to zero in both surveys. Even when  $\alpha_2$  is significant at the usual 5% level, the magnitude of the differences in Italian language skills is up to 2.1 percentage points. More importantly, the inclusion of the conditional variables corrects for the unobserved heterogeneity. For example, in the sample women with language problems is evident a reduction of the upward bias passing from the unconditional differences on employment rate in sample women with language problems to the point estimates in the conditional equation 1, where the effect of Italian language problems is smaller (i.e. the difference is 10 percentage points, see Table 5).

In addition, when we consider all samples in the ILCHF survey (2009), the coefficient for employment outcome is smaller with respect to the unconditional estimate in Table 5, although significant. Finally, the conditional OLS point estimates for log wages and log hourly wages suggest that the Italian language problems of immigrants may cause a reduction of 3.8% and 4.9% in log wages and log hourly wages, respectively. The relationship between language problems and wages appears to be driven by the men, with point estimates 1.5%-2% higher than the average of the sample.

Table 6: Estimation results, conditional OLS

	CSIFC survey (2011-2012)			ILCHF survey (2009)		
	Total sample	Women	Men	Total sample	Women	Men
<b>Employment rate</b>	-0.012	-0.021**	-0.001	-0.035***	-0.026*	-0.053***
	(0.009)	(0.009)	(0.013)	(0.009)	(0.012)	(0.013)
<i>Observations</i>	13,091	5297	5637	7316	4041	3275
<b>Job discrimination</b>	0.007	0.022*	0.001			
	(0.009)	(0.013)	(0.013)			
<i>Observations</i>	11,056	7491	5637			
<b>Log wages</b>				-0.038***	-0.036*	-0.053***
				(0.011)	(0.015)	(0.015)
<i>Observations</i>				4432	2086	2346
<b>Log hourly wages</b>				-0.049***	-0.046*	-0.078***
				(0.014)	(0.020)	(0.017)
<i>Observations</i>				4432	2086	2346

*Notes:* Robust clustered standard errors are in parentheses. The asterisks stand for the *p*-value significance levels (\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ ).

Overall, these correlations suggest that language abilities may have a role in determining the labour market outcomes of immigrants. However, estimates from equation 1 cannot identify the causal effect of language skills on labour performances due to some sources of endogeneity that could decrease or increase the bias of the estimated coefficient. Below, we discuss the strategy to account for these estimation biases.



#### 4.1. The proposed empirical model

The set of conditioning variables available includes indicators that are likely to be correlated with unobserved heterogeneity in the relationship between language skills and employment that sorts individuals into groups of those who do and those who do not acquire the host country language, such as education, ability tests and partner information. If these variables do not fully account for immigrant selection into the proficient and non-proficient Italian language groups, language effects may still be upwardly biased. We proved this in the previous section.

To account more efficiently for the language problem selection, we use a propensity score-matching (PSM) estimator, which balances the covariates of immigrants with or without language problems. Using a nearest-neighbour method without replacement, the PSM estimator ensures that all individuals in the treatment group are compared with their counterparts in the comparison group who are similar in terms of the observable characteristics. In Figure 5 we present the extent of balancing the covariates between the groups of immigrants. A quick comparison between the distribution of the covariates (dots) reveals that, after the application of the PSM estimator both CSIFC and ILCHF surveys have largely improved balancing of covariates. The standardised bias (%) is approximately zero and the variance ratio of the residuals lie within the usual confidence intervals (Austin, 2009).

Therefore to eliminate the bias induced by differences in observable characteristics, we will estimate the coefficient  $\alpha_2$  using OLS which includes the individual weight of the observation attributed to matched individual  $j$  when compared with treated individual  $i$  (i.e. OLS-PSM).

The identification issue merits further explanation because self-reported language skill measures are subject to substantial measurement error. The immigrants interviewed tend generally to over-estimate their language abilities either for an incomplete knowledge of the language skills of the destination country or for a positive propensity through the language abilities to show that they are integrated in the immigration country. This propensity leads to downward bias, which is alleviated when the language evaluation is carried out by the interviewer directly (Dustmann and van Soest, 2001).

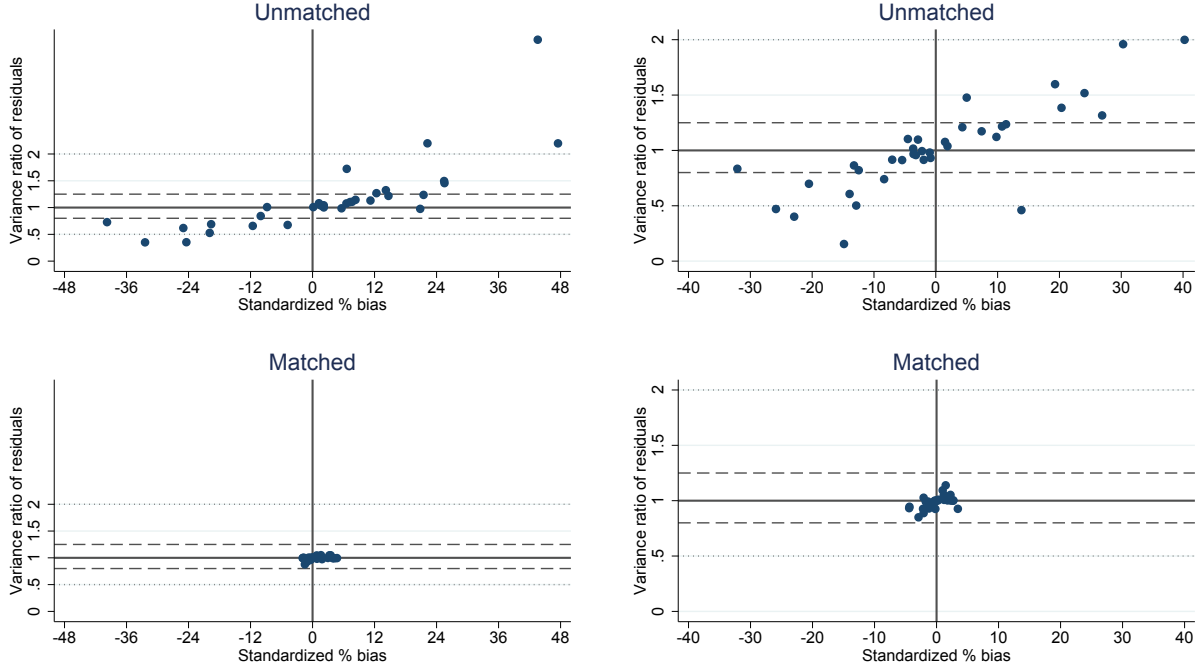
To deal with this source of bias, we use an IV estimator. After applying PSM, an IV approach (IV-PSM) addresses to interpret causally the language skill coefficient. A typical IV proposed in the literature is age at arrival in the host country (Bleakley and Chin, 2004, 2010; Miranda and Zhu, 2013; Sweetman and van Ours, 2008; Yao and van Ours, 2015; Guven and Islam, 2015). The main argument is that people who are exposed to a new language early are likely to have good language skills at adulthood, whereas immigrants arriving at a later age have much more difficult acquiring language skills. Figures 2 and 3 empirically confirm the assumption of the validity of the instrument for our dataset. While other IVs have been used in the literature<sup>14</sup>, the limited diffusion of Italian language outside Italy excludes immigrants who grew up speaking Italian, limiting the threat to affect the relationship between age at arrival and language proficiency. Estimates from the CSIFC survey suggest that the proportion of immigrants who state they spoke or read Italian during

---

<sup>14</sup>For example, minority concentration in the area where the immigrant lives, linguistic distance between the immigrant's mother tongue and the language of the host country, language spoken at home, number of children, overseas marriage and parental education. For a discussion see, Yao and van Ours (2015).

childhood is small, about 0.5%.

Figure 5: Unmatching and matching of covariates



(a) CSIFC survey (2011-2012)

(b) ILCHF survey (2009)

## 5. Results

### 5.1. Baseline estimates

We first use the OLS-PSM estimator to evaluate the effects of the immigrants' differences in Italian language proficiency on labour market performance. Table 7 lists the estimated coefficients for employment and job discrimination. Using the CSIFC 2011-2012 survey, the point estimates for employment are close to zero in the full sample. This suggests that if estimations would be corrected, the employment probability of immigrants with or without language problems may be equivalent. Conclusions are similar when the outcome used is job discrimination. For the ILCHF survey (2009), the aggregate point estimate of employment is significant although small in magnitude, suggesting that the probability to be employed for immigrants with language problems decreases by 2.4 percentage points (s.e. 0.009). In addition, the point estimates are significant only for men in driving the difference in the relationship between language proficiency and unemployment ( $\alpha_2=0.019$ ; s.e. 0.010).

Table 8 lists the parameter estimates controlled by PSM for the effect of language problems on wage outcomes. The results show significant negative effects; the magnitude ranges from 4%-7.5% and is consistent with OLS results (Table 6), irrespective of gender differences.

We proceed by addressing the concern that the effect of language skills on labour market outcomes could potentially be driven by measurement error related to the immigrants' over-estimation of their perceived Italian language abilities, as discussed in Section 4. As shown

Table 7: Estimation results by OLS matching, employment rate and job discrimination

	Employment rate			Job discrimination		
	Total sample	Women	Men	Total sample	Women	Men
<b>CSIFC survey (2011-2012)</b>						
Language problems	-0.011 (0.008)	-0.019** (0.009)	-0.001 (0.012)	0.003 (0.011)	0.007 (0.015)	0.002 (0.015)
Observations	13,809	7489	5600	11,055	5636	5419
<b>ILCHF survey (2009)</b>						
Language problems	-0.024** (0.009)	-0.032* (0.014)	-0.019** (0.010)			
Observations	6099	2933	3136			

*Notes:* Language problems are defined as having either speaking and reading problems. Robust clustered standard errors are in parentheses. The asterisks stand for the *p-value* significance levels (\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ ).

Table 8: Estimation results by OLS matching, wages by ILCHF survey (2009)

	Log wages			Log hourly wages		
	Total sample	Women	Men	Total sample	Women	Men
Language problems	-0.044*** (0.012)	-0.045** (0.015)	-0.057*** (0.017)	-0.049** (0.015)	-0.048** (0.020)	-0.075** (0.018)
Observations	4432	2086	2346	3332	2086	2346

*Notes:* Language problems are defined as having either speaking or reading problems. Robust clustered standard errors are in parentheses. The asterisks stand for the *p-value* significance levels (\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ ).

in Table 9, the results of the IV matching estimates for employment are larger and statistically significant, irrespective if estimated using the CSIFC survey (2011-2012) or the ILCHF survey (2009). In both samples, we find that immigrants with language problems had a significantly decreased employment rate, with a magnitude of the point estimates around 16-20 percentage points. This suggests that measurement errors accounting for Italian language abilities in employment are very important and that they do not vary systematically across sample surveys. In terms of percentage, we calculate that these effects vary from 27%-33% in the two samples<sup>15</sup>.

<sup>15</sup>We calculate this percentage change using the immigrants' employment rate of 0.6.

Table 9: Estimation results by IV matching, employment rate and job discrimination

	Employment rate			Job discrimination		
	Total sample	Women	Men	Total sample	Women	Men
<b>CSIFC survey (2011-2012)</b>						
Language problems	-0.164*** (0.040)	-0.104*** (0.040)	-0.248*** (0.074)	0.352*** (0.066)	0.258*** (0.076)	0.398*** (0.092)
Weak instrument test	342.711★	288.950★	131.158★	195.923★	125.861★	112.599★
Observations	13,809	7489	5600	11,055	5636	5419
<b>ILCHF survey (2009)</b>						
Language problems	-0.204*** (0.053)	-0.178** (0.080)	-0.219*** (0.063)			
Weak instrument test	113.800★	69.871★	75.507★			
Observations	6025	2907	3118			

*Notes:* Language problems are defined as having either speaking or reading problems. Robust clustered standard errors are in parentheses. The asterisks stand for the  $p$ -value significance levels (\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ ). We also report the relevance of the IV by F-statistics under the testing hypothesis of a weak instrument. The star indicates the rejection of the weak instrument associated with the F-statistic that exceeds 10 (Angrist and Pischke, 2009).

Table 10: Estimation results of wages by IV-PSM (ILCHF survey, 2009)

	Log wages			Log hourly wages		
	Total sample	Women	Men	Total sample	Women	Men
Language problems	-0.238*** (0.070)	-0.162** (0.069)	-0.272** (0.094)	-0.269*** (0.087)	-0.171** (0.077)	-0.304** (0.118)
Weak instrument test	110.500★	85.922★	51.882★	110.500★	85.922★	51.882★
Observations	4405	2072	2333	4405	2072	2333

*Notes:* Language problems are defined as having either speaking or reading problems. Robust clustered standard errors are in parentheses. The asterisks stand for the  $p$ -value significance levels (\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ ). We also report the relevance of the IV by F-statistics under the testing hypothesis of a weak instrument. The star indicates the rejection of the weak instrument associated with the F-statistic that exceeds 10 (Angrist and Pischke, 2009).

Further, given the representativeness of the sample, we apply the model to the subsamples of men and women and ascertain whether immigrants' language problems affected the employment rate differently. Columns 2-3 of Table 9 show that all estimates are consistent

with our expectations. Based on the CSIFC survey, male immigrants with language problems are less likely (by 25 percentage points) to be in employed compared to immigrants with good Italian proficiency. The magnitude is slightly smaller (minus 21.9 percentage points) when we estimate the model using the ILCHF survey. Although the confidence intervals overlap estimates by gender, we conclude that male immigrants with good Italian language proficiency have an increased probability of being employed compared to female immigrants. The table also reports the F-test for the relevance of the IV, age at arrival. In fact, the estimates obtained may be biased with a weak instrument. For all estimates the F-statistics are larger than the rule of thumb (e.g., F-statistics exceeding 10), indicating that our estimates do not suffer from weak instruments<sup>16</sup>. We also provide a quantitative analysis of the relationship between language problems and job discrimination. The IV-PSM estimates in the full sample suggest that significant discrimination driven by Italian language problem exists (35 percentage points) and that this discrimination in the workplace is stronger for men (i.e.,  $\alpha_2 = 0.398$ ) mainly involved in the manufacturing sectors, where good knowledge of the Italian language is necessary to achieve greater responsibility in firms. Given the high percentage of responses regarding discrimination in the workplace (e.g., sample mean 0.8), our estimates suggest that workplace discrimination increased more than 40% due to the lack of Italian language proficiency.

In Table 10, we employ the same model to evaluate the effect of language problems on wages (e.g. log wages and log hourly wages). Thus, we are effectively asking how much wages decrease when an immigrant has difficulty in speaking or reading Italian. The results show that the magnitude of these effects in decreasing wages - from 23% for the log wages to 27% for log hourly wages, are consistent with our expectations of a large downward bias induced by measurement errors associated with Italian language abilities. Regarding the other labour outcomes, gender differences show that the wages of immigrants with language problems are estimated to be reduced by 30% for men using the outcome of the log hourly wage.

## 5.2. Robustness

In this section, we present several robustness checks of our findings. Table 11 shows the point estimates of language problems when we include age at arrival as an exogenous variable (i.e. not as an instrument) in the OLS labour market outcome equations. This specification controls for direct effects of age at arrival on the relevant labour market variables that are independent from the Italian language skills of the immigrants. Although the estimated parameters of language problems may be affected by the inclusion of age at arrival as a right-hand side variable, we do not find any statistically significant differences with the point estimates presented in Tables 7 and 8. This reveals that the effect of age at arrival is almost fully captured by the language problems variable and that labour outcomes are limitedly affected through non-language channels.

---

<sup>16</sup>We also perform a Durbin-Wu-Hausman test for endogeneity of language problems, where significant F-statistics suggest that the language indicator is endogenous. Based on the data, we do not reject the null hypothesis of exogeneity for all IV estimates in this paper, although to save space we do not report it extensively in the tables. All estimates and tests are available from the corresponding author upon request.

Table 11: Robustness: OLS estimates including age at arrival as a right-hand side variable

	CSIFC survey (2011-2012)			ILCHF survey (2009)		
	Total sample	Women	Men	Total sample	Women	Men
<b>Employment rate</b>	-0.005 (0.008)	-0.015* (0.009)	-0.007 (0.012)	-0.019 (0.009)	-0.024 (0.014)	-0.015 (0.012)
Observations	13,809	7489	5600	6099	2907	3118
<b>Job discrimination</b>	0.013 (0.011)	0.016 (0.015)	0.010 (0.015)			
Observations	11,055	5636	5419			
<b>Log wages</b>				-0.037 ** (0.019)	-0.039** (0.019)	-0.054** (0.024)
Observations				4432	2072	2333
<b>Log hourly wages</b>				-0.043** (0.020)	-0.042** (0.021)	-0.071** (0.018)
Observations				4432	2072	2333

Notes: Robust clustered standard errors are in parentheses. The asterisks stand for the  $p$ -value significance levels (\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ ).

A second robustness check is performed by introducing a supplementary instrument in the labour market equations. We interact the age at arrival variable with a dummy variable indicating whatever or not an immigrant attended an Italian language course when he or she first arrived in Italy. This estimate is carried out for the ILCHF survey (2009), which records this information. Although the interaction may underestimate the true effect of language skills on labour market performance through the only 'age at immigration' channel, the small number of individuals who attended Italian language course (see Table 5) may not change the overall effect. Table 12 shows the estimates using in both instruments, which are close to the previous ones.

In addition, we use the CSIFC survey (2011-2012) to investigate whether restricting the language skill problems to reading led to different point estimates the impact on employment rate and workplace discrimination. We conducted this robustness check because the descriptive statistics showed that the percentage of immigrants with reading problems was significantly higher than the percentage of immigrants with speaking problems. We still show that the parameters of language problems listed in Table 13 are not significantly different with respect to the baseline estimations.

### 5.3. A focus on wage disparities of immigrants

In this section, we use the baseline IV coefficients to measure the contribution of immigrants' language problems in explaining the wage variability of people residing in Italy. While the ILCHF 2009 survey may be merged with the IT-SILC to extract wages of the Italian people conditional to the presence or absence of language problems, we have a lack of immigrant representativeness in the IT-SILC for the successive years. This leads to developing a different estimation strategy, integrating data and point estimates of the ILCHF 2009 survey with data from the Italian labour force surveys. Sections a) and b) of Table 14

Table 12: Robustness: IV-PSM estimates using an additional instrument

	Total sample	Women	Men
<b>Employment rate</b>	-0.323** (0.140)	-0.383** (0.192)	-0.243** (0.119)
Observations	5954	2871	3083
<b>Log wages</b>	-0.341** (0.162)	-0.375 (0.230)	-0.293** (0.142)
Observations	4359	2052	2307
<b>Log hourly wages</b>	-0.381* (0.210)	-0.480** (0.235)	-0.370** (0.211)
Observations	4359	2052	2307

*Notes:* Robust clustered standard errors are in parentheses. The asterisks stand for the *p-value* significance levels (\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ ).

Table 13: Robustness: IV-PSM estimates on labour performance restricted to reading problems

	Total sample	Women	Men
<b>Employment rate (E)</b>	-0.169*** (0.035)	-0.130*** (0.036)	-0.255*** (0.064)
Observations	12,970	7489	5600
<b>Job discrimination (JD)</b>	0.312*** (0.058)	0.209*** (0.067)	0.365*** (0.079)
Observations	10,985	5636	5419

*Notes:* Robust clustered standard errors are in parentheses. The asterisks stand for the *p-value* significance levels (\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ ).

summarise these information. We assume that the proportions of immigrants with poor ( $\eta_1$ ) or good language ( $\eta_2$ ) skills are approximately constant over time. We then calculate the different wages of immigrants using the estimated parameter  $\alpha_2$  in Table 10 and exploiting mean wage information ( $W_{mean}$ ) of immigrants (and Italians) from the LFS<sup>17</sup>.

We report the results related to immigrant wage estimates for the two language proficiency groups in sub-section c) of Table 14. In particular, column 2 of this table shows that immigrants with poor language proficiency have a mean wage of around 890 euros per month. This increases for immigrants without language problems to 1101 euros per month, suggesting that immigrants with poor language ability are severely affected in terms of economic opportunities. The immigrants' wage disparities are emphasised by gender (columns

<sup>17</sup>Formally, we obtain the wage estimates using the following equation:  $W_{mean} = \eta_1 * W_{LP_0} + \eta_2 * W_{LP_1}$ . We substitute the parameters from column 2 of Table 14, such that  $0.39 * W_{LP_1} + 0.61 * (1.23 * W_{LP_1}) = 1019$ . This also implies that  $W_{LP_1} * (0.39 + 0.61 * 1.23) = 1019$  and  $W_{LP_1} = 890$ . The wage for immigrants without language problems is  $W_{LP_0} = 1.238 * W_{LP_1} = 1101$ .

3 and 4). Estimated gains for immigrant women are reduced to 775 euros per month against 992 euros per month for men. In the successive columns, we also report the scenarios for the years 2011-2012 and 2015-2016. We found that the patterns are fully consistent with the scenario described in the 2009 sample, which confirms that Italian language abilities represent one of the main sources of wage inequalities between immigrants.

In the last two lines of Table 14, we list the contribution of immigrants' language problems in explaining the wage variability of people residing in Italy. We estimate it separating "within" contribution to wage inequality of immigrants determined by a lack of language skills with respect to other factors affecting wage differences between immigrants and natives (i.e. "between" contribution to wage variability of resident in Italy). The results indicate that in 2009 about 65% of the wage inequality could be ascribed to Italian language problems, while other unobserved channels could explain the remaining 35% of the wage differences between immigrants with good proficiency in Italian and Italian workers. The contribution of Italian language skills to the wage inequality of immigrants appears to be mostly driven by men. Indeed, more than 78% of the wage variability of immigrants was explained by the language problems of immigrants.

A potential mechanism to explain differences in language skill effects can be attributed to selection of immigrants into different occupations. For example, Berman et al. (2003) show that Hebrew language proficiency reduces the wage differentials (compared to native speakers) of programmers and technicians in Israel, while this does not hold for construction workers and gas station attendants. These results are also confirmed in Italy when the effects of language proficiency on wages are compared within immigrants' occupation by gender. Table 15 shows that immigrant women in Italy are mainly employed in the domestic sector as household workers (i.e. unskilled occupation), who receive the lowest wages, while the manufacturing and construction mainly employ immigrant men (56%). This gender structure of employment for women smooths the positive effect of a good command of Italian on wages and, in turn, limits the potential reduction of wage differentials compared to native speakers.

These differences in the impact of Italian language proficiency seem to decrease in the successive years under the boost of the financial crisis that, in turn, affected the real economy and made the expectation of employment and the careers of immigrants less important. Given the evidence that the composition by country of origin and gender of immigrants did not change recently (Appendix B), the mechanism based on the selection into different occupations may at least partly explain why a poor command of Italian impacts wages less. Data from the LFS (various years) show that the long term effects of the 2008 financial crisis reduced the proportion of immigrant men employed in the manufacturing sectors, where a good command of the host country language is required, particularly in the high-tech sector. This increased the mean wage differences between immigrants and natives as a consequence of a complementarity lack between language proficiency and further human capital emerged (Isphording, 2014). On the other hand, immigrant women experienced more stable household service employment, which supports the prediction that having a good command of Italian may not be a sufficient condition for immigrant women to reduce disparities with respect to



Table 14: Conditional wage of immigrants by language proficiency and wage inequality: estimates

	Estimates 2009			Scenario 2011-2012			Scenario 2015-2016			
	All	Women	Men	All	Women	Men	All	Women	Men	
a) ILCHF survey 2009										
Language problems ( $\alpha_2$ )	(23.8)	(16.2)	(27.2)	(23.8)	(16.2)	(27.2)	(23.8)	(16.2)	(27.2)	
Proportion of immigrants with poor language( $\eta_1$ )	0.39	0.384	0.394	0.39	0.384	0.394	0.39	0.384	0.394	
Proportion of immigrants with good language ( $\eta_2$ )	0.61	0.616	0.606	0.61	0.616	0.606	0.61	0.616	0.606	
b) Labour Force surveys: Years 2009, 2011-2012 and 2015-2016										
Wages of immigrant workers	1019	853	1156	1024	854	1177	1058	898	1220	
Wages of Italian workers	1235	1092	1349	1292	1148	1413	1360	1211	1491	
c) Conditional estimates for immigrant wage										
Wages of immigrants with poor language skills	*	890	775	992	894	777	1010	924	817	1047
Wages of immigrants with good language skills	*	1101	901	1262	1107	902	1285	1144	949	1332
Contribution of language skills to wage inequality (%)	**	0.65	0.44	0.78	0.58	0.38	0.72	0.55	0.38	0.68
Other factors contributing to wage inequality (%)		0.35	0.56	0.22	0.42	0.62	0.28	0.45	0.62	0.32

Notes: \* Wages of immigrants with poor or good language skills are obtained as described in footnote 17. \*\* The contribution of language problems to wage inequality is obtained as a share of within contribution due to the language problem effects to inequality of immigrants' wages with respect to the whole inequality of wages.

natives<sup>18</sup>.

Table 15: Employment composition (in percentage) by sector and gender

	Men		Women		Total	
	Italian	Foreign	Italian	Foreign	Italian	Foreign
Agriculture	2.3	7.1	1.7	2.9	2.0	5.2
Manufacturing	33.2	33.5	18.6	11.5	26.8	23.8
<i>Low-tech</i>	<i>23.7</i>	<i>31.5</i>	<i>13.2</i>	<i>9.9</i>	<i>19.1</i>	<i>21.8</i>
<i>High-tech</i>	<i>9.5</i>	<i>2.0</i>	<i>5.4</i>	<i>1.6</i>	<i>7.7</i>	<i>2.0</i>
Construction	9.5	24.6	1.8	1.3	6.1	14.4
Commerce	20.8	18.7	19.2	16.3	20.1	17.6
Services	32.1	7.3	56.1	15.6	42.7	11.0
Household services	2.0	9.0	2.7	52.4	2.3	28.1

Source: Italian Labour Force Survey (LFS), year 2009.

<sup>18</sup>Chiswick and Miller (2010) found equivalent results for workers with poor English skills employed in jobs that have very low English-language requirement in England.

## 6. Conclusion

Fluency in the language of the destination country can facilitate the transfer of migrants' skills to the new labour market, thus contributing to the global interchange of skills and stimulating economic growth. We first argued that migrants with a low proficiency in Italian had particularly reduced opportunities in the labour market and documented that they were significantly affected by decreasing employment and increasing workplace discrimination. We then showed that immigrants with Italian linguistic problems reduced their expected wages relative to immigrants with good Italian linguistic proficiency, suggesting that the tightening of the linguistic abilities of immigrants in host countries may significantly constraint the adverse impact on the individuals that were affected. In particular, we found that Italian language problems explained more than half of the wage inequalities between natives and immigrants, particularly for men. An important implication of this analysis for immigration policy is that immigrants proficient in Italian improve the opportunities in the economic, social and civic life of their new country.

The study is not without drawbacks. One such limitation is that we were unable to investigate the influence of different types of residence permits held by immigrant respondents given the absence of information about immigrants' declaration of having or not having a permanent residence permit. In addition, we do not have data on undocumented immigrants, who represent about 10% of the immigrant population. This implies that for undocumented immigrants facing increasing difficulty in realising their expectations in the host country, the lack of language skills exacerbates the difficulty of obtaining (good) employment. Therefore, our results most likely underestimate the potentially larger differences in labour market outcomes between immigrants and natives in the general population. Taking these limitations into account, our results are certainly a conservative test of the impact of language proficiency on employment, workplace discrimination and wage disadvantages in Italy.

## Appendix A. Descriptive statistics, covariates

		CSIFC survey (2014)				ILCHF survey (2009)			
		Men		Women		Men		Women	
Variables	Any language problems	No	Yes	No	Yes	No	Yes	No	Yes
Age		37.666	39.489	38.76	38.708	37.181	37.063	37.506	37.292
<i>Marital status</i>	Single	0.332	0.341	0.27	0.19	0.285	0.362	0.256	0.024
	Married	0.56	0.462	0.488	0.617	0.67	0.586	0.575	0.614
	Divorced	0.105	0.19	0.195	0.144	0.034	0.04	0.108	0.072
	Widowed	0.004	0.006	0.047	0.048	0.009	0.01	0.059	0.065
<i>Type of household</i>	Living alone	0.237	0.402	0.25	0.222				
	Couples with children	0.567	0.398	0.472	0.519				
	Couples without children	0.141	0.147	0.16	0.175				
	Father with children	0.022	0.037	0.002	0.002				
Family size	Mother with children	0.033	0.017	0.116	0.081				
	One					0.181	0.29	0.232	0.257
	Two					0.171	0.173	0.216	0.176
	Three					0.228	0.185	0.218	0.198
	Four					0.233	0.185	0.199	0.207
	>Four					0.185	0.166	0.132	0.159
Household relation	Head					0.748	0.762	0.618	0.496
	Spouse					0.108	0.078	0.287	0.414
	Head parents					0.008	0.008	0.023	0.021
	Head sons					0.055	0.048	0.031	0.023
	Head brother					0.06	0.071	0.028	0.031
	Other relative					0.018	0.029	0.009	0.012
<i>Education</i>	No education	0.061	0.216	0.039	0.189	0.045	0.166	0.028	0.115
	Primary	0.058	0.108	0.038	0.105	0.092	0.169	0.069	0.149
	Lower secondary	0.33	0.352	0.246	0.345	0.311	0.372	0.254	0.345
	Upper secondary	0.466	0.301	0.529	0.316	0.447	0.263	0.511	0.329
	Higher education	0.085	0.023	0.149	0.045	0.102	0.028	0.134	0.06
Number of children		1.176	1.206	1.232	1.55				
Foreign degrees						0.953	0.979	0.965	0.986
Illiterate		0.028	0.152	0.019	0.135				
Ever employed in the country of birth		0.442	0.401	0.482	0.713				
First time job seeker		0.024	0.019	0.023	0.03				
Never worked						0.01	0.019	0.149	0.277
Childcare		0.242	0.226	0.209	0.309	0.008	0.01	0.199	0.264
House-maker		0.003	0.003	0.241	0.44	0.003	0.014	0.241	0.33
Attending It.course						0.074	0.025	0.072	0.019
Area of residence	Metropolitan areas	0.252	0.223	0.267	0.214				
	Municipalities (more 10.000)	0.248	0.24	0.217	0.279				
Area of residence	Municipalities (less 10.000)	0.5	0.537	0.516	0.507				
	Big					0.399	0.419	0.43	0.415
	Intermediate					0.448	0.43	0.415	0.43
	Small					0.151	0.149	0.153	0.153
Macro-region	North	0.416	0.265	0.358	0.338	0.524	0.408	0.444	0.376
	Centre	0.185	0.122	0.191	0.118	0.21	0.165	0.232	0.17
	South	0.399	0.613	0.452	0.544	0.263	0.424	0.321	0.451
Reason to immigrate	Absence of work	0.443	0.5	0.309	0.285				
	A higher income	0.2	0.255	0.164	0.127				
	To improve quality of life	0.121	0.107	0.122	0.147				
	Family reunion	0.114	0.08	0.296	0.384				
	War	0.03	0.016	0.009	0.008				
	Other	0.092	0.041	0.099	0.048				
Reason to immigrate	Work					0.824	0.912	0.57	0.569
	Family					0.11	0.065	0.374	0.406
	To study					0.017	0.001	0.015	0.002
	Other					0.047	0.02	0.039	0.022
Region of birth	F-Y*, Albany, Romania	0.446	0.277	0.403	0.293	0.323	0.29	0.376	0.304
	Other Europe	0.13	0.093	0.302	0.163	0.095	0.077	0.289	0.209
	Asia	0.127	0.303	0.079	0.263	0.078	0.239	0.068	0.202
	Africa	0.24	0.305	0.114	0.247	0.199	0.331	0.096	0.213
	America	0.055	0.019	0.099	0.031	0.303	0.061	0.169	0.069

*Notes:* The Table compares the conditions and social integration of foreign citizens survey (CSIFC survey 2011-2012) and the income and living conditions of households with immigrants survey (ILCHF survey 2009). \*F-Y: Former Yugoslavia

## Appendix B. Composition by country of origin and gender of immigrants in Italy, 2009-2015

	2009		2010		2011		2012		2013		2014		2015	
	M	F	M	F	M	F	M	F	M	F	M	F	M	F
F-Y*, Albany, Romania	34.80	33.43	34.65	33.72	34.03	33.44	33.65	33.46	33.28	33.91	32.98	34.22	33.03	34.69
Other-Europe	14.01	24.73	13.99	24.56	14.01	24.98	13.98	24.4	13.74	23.97	13.43	23.75	13.06	23.45
Asia	17.50	14.22	18.11	14.45	18.91	14.79	20.31	15.52	20.99	15.96	21.71	16.42	22.29	16.68
China & India	(7.48)	(6.02)	(7.78)	(6.17)	(8.26)	(6.29)	(9.01)	(6.72)	(9.25)	(6.95)	(9.28)	(7.06)	(9.45)	(7.17)
Africa	27.39	17.55	26.89	17.37	26.61	16.92	25.59	16.97	25.67	16.76	25.53	16.33	25.36	16.11
Tunisia & Morocco	(15.57)	(10.41)	(15.09)	(10.31)	(14.65)	(9.98)	(14.34)	(10.20)	(13.84)	(9.89)	(13.22)	(9.42)	(12.75)	(9.18)
America	6.29	10.06	6.35	9.89	6.42	9.86	6.46	9.64	6.33	9.39	6.34	9.27	6.26	9.06
Total	100	100	v 100	100	100	100	100	100	100	100	100	100	100	100

Source: Italian Institute of Statistics, <http://istat.it/>. The value are expressed in percentage. Subsamples of immigrants by continents in parentheses.

**Acknowledgments:** The authors are grateful to the participants to the XXVIII International Population Conference (Cape Town, 2017) and to the XXXIII Italian Conference of Labour Economics (Ancona, 2018) for their helpful comments. The authors are also grateful to the participants of the internal seminars at the University of Maiorca and Perugia for their suggestions.

## References

- Adserà, A., Pytliková, M., 2016. Language and Migration. The Palgrave Handbook of Economics and Language, 342-372., London, UK.
- Angrist, J. D., Pischke, J.-S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, New Jersey.
- Austin, P., 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity score matched samples. *Statistics in Medicine* 28 (25), 3083–3107.
- Avallone, G., 2017. The land of informal intermediation. The social regulation of the migrant agricultural labour in Piana del Sele, Italy. Routledge, New York, Ch. Migration and Agriculture, pp. 217–230.
- Berman, E., Lang, K., Siniver, E., 2003. Language-skill complementarity: returns to immigrant language acquisition. *Labour Economics* 10 (3), 265–290.
- Bleakley, H., Chin, A., 2004. Language Skills and Earnings: Evidence from Childhood Immigrants. *The Review of Economics and Statistics* 86 (2), 481–496.
- Bleakley, H., Chin, A., 2010. Age at arrival, english proficiency, and social assimilation among U.S. immigrants. *American Economic Review, Applied Economics* 2 (1), 165–192.
- Bonifazi, C., 2009. The italian transition from migration to immigration country. Tech. rep., IRPPS Working Papers.
- Chiswick, B. R., Jun. 2008. The Economics of Language: An Introduction and Overview. IZA Discussion Papers 3568, Institute for the Study of Labor (IZA).
- Chiswick, B. R., Miller, P. W., 2003. The complementarity of language and other human capital: immigrant earnings in Canada. *Economics of Education Review* 22 (5), 469–480.
- Chiswick, B. R., Miller, P. W., 2009. Earnings and occupational attainment among immigrants<sup>1</sup>. *Industrial Relations: A Journal of Economy and Society* 48 (3), 454–465.
- Chiswick, B. R., Miller, P. W., 2010. Occupational language requirements and the value of English in the US labor market. *Journal of Population Economics* 23 (1), 353–372.
- Chiswick, B. R., Miller, P. W., 2015. International Migration and the Economics of Language, In: Chiswick, B.R., Miller, P.W. (Eds.). *Handbook of the Economics of International Migration: The Immigrants*, Chapter 5, 240-269. Elsevier., Amsterdam, Netherlands.
- Clark, K., Drinkwater, S., 2008. The labour-market performance of recent migrants. *Oxford Review of Economic Policy* 24 (3), 495–516.
- Dustmann, C., Fabbri, F., 2003. Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal* 113 (489), 695–717.

- Dustmann, C., van Soest, A., 2001. Language fluency and earnings: Estimation with misclassified language indicators. *The Review of Economics and Statistics* 83 (4), 663–674.
- Ebner, C., Helbling, M., 2016. Social distance and wage inequalities for immigrants in Switzerland. *Work, Employment and Society* 30 (3), 436–454.
- Finotelli, C., Arango, J., 2011. Regularisation of unauthorised immigrants in Italy and Spain: Determinants and effects. *Documents d’Analisi Geografica* 57 (3), 495–515.
- Friedberg, R. M., 2000. You can’t take it with you? Immigrant assimilation and the portability of human capital. *Journal of Labor Economics* 18 (2), 221–251.
- Gonzales, P. S., 2010. The impact of education and host language skills on the labor market outcomes of immigrants in Spain. In: *Investigaciones de Economía de la Educación 5. Asociación de Economía de la Educación*, pp. 798–824.
- Güven, C., Islam, A., 2015. Age at migration, language proficiency, and socioeconomic outcomes: Evidence from Australia. *Demography* 52 (2), 513–542.
- Heath, A., Cheung, S., 2007. *Unequal chances: Ethnic minorities in Western labour markets*. No. Oxford, UK. British Academy/Oxford University Press, Oxford.
- Hermansen, A. S., 2017. Age at arrival and life chances among childhood immigrants. *Demography* 54 (1), 201–229.
- Isphording, I. E., 2014. *Language and Labor Market Success*. IZA Discussion Papers 8572, Institute for the Study of Labor (IZA).
- Kalter, F., Kogan, I., 2006. Ethnic inequalities at the transition from school to work in Belgium and Spain: Discrimination or self-exclusion? *Research in Social Stratification and Mobility* 24 (3), 259 – 274.
- Mastrobuoni, G., Pinotti, P., 2015. Legal status and criminal activity of immigrants. *American Economic Journal: Applied Economics* 7 (2), 175–206.
- McGovern, P., 2007. Immigration, labour markets and employment relations: Problems and prospects. *British Journal of Industrial Relations* 45 (2), 217–235.
- Miranda, A., Zhu, Y., 2013. English deficiency and the native immigrant wage gap. *Economics Letters* 118 (1), 38–41.
- Pieroni, L., Pompei, F., 2008. Evaluating innovation and labour market relationships: the case of Italy. *Cambridge Journal of Economics* 32, 325–347.
- Ponzo, I., Finotelli, C., Malheiros, J., Fonseca, L., Salis, E., 2015. Is the economic crisis in southern Europe turning into a migrant integration crisis?. *Politiche Sociali* 1, 59–88.
- Salmasi, L., Pieroni, L., 2015. Immigration policy and birth weight: Positive externalities in Italian law. *Journal of Health Economics* 43, 128 – 139.

- Sweetman, A., van Ours, J., 2008. Immigration: what about the children and grandchildren?. In: Chiswick, B.R., Miller, P.W. (Eds.), Handbook of the Economics of Immigration. Handbook of the Economics of Immigration. Elsevier, 1141-1193, Amsterdam, Netherlands.
- Yao, Y., van Ours, J. C., 2015. Language skills and labor market performance of immigrants in the Netherlands. Labour Economics 34, 76– 85.
- Zschirnt, E., Ruedin, D., 2016. Ethnic discrimination in hiring decisions: a meta-analysis of correspondence tests 1990 - 2015. Journal of Ethnic and Migration Studies 42 (7), 1115 – 1134.