

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

Luca Pieroni^{a,*}, Giorgio d'Agostino^b, Donatella Lanari^c

^a*Department of Political Science, University of Perugia*

^b*Department of Economics, Roma Tre University*

^c*Department of Medicine, University of Perugia*

Abstract

In this paper, we examine how the Italian language problems of immigrants affect their labour market performance in terms of employment using two hitherto unexploited immigration surveys published recently by the Italian Institute of Statistics. While these surveys collect data on the life conditions on the first and second generations of immigrants, and includes relevant information on several aspects regarding socio-economic conditions of each household member in the country of origin and in Italy, they complement employment information with wages. We use age at arrival to instrument the relationship between language skills and labour market outcomes of immigrants. With respect to immigrants with a good Italian proficiency, our findings suggest that language problems reduce by about 30% employment rate, and estimate an equivalent large job discrimination. Italian language skills also affect significantly wages of immigrants and are gender differentiated. The point estimates suggest a wage gap of about 20% between immigrants with a lack of Italian proficiency, a magnitude that increases to 25% for men. Robustness confirmed our estimates.

Keywords: Immigrants, Language skills, Employment, Wages

JEL codes: J24, J15

*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 the Southern Europe as migration flows have increased recently. In particular, it has been affected by the relatively high proportion of immigrants, a group that is more frequently exposed to a weak labour market position: employment rate was lower and unemployment rate was higher with respect to natives¹ and the job opportunities were registered mainly in the low-skilled (and often precarious) jobs, even if educational level were comparable with native speakers and across ethnic groups (Adserà and Pytliková, 2016).

The language skills is certainly an important dimension of individual's human capital of immigrants and a determinant of the employment success. A lack of language skills may induce the migrant to work in jobs that require a lower education level than the level achieved by the migrant and/or may lead to a lower performance on the job and, in turn, employment rate and wage gaps. While the literature makes indisputable that language skills have a strong impact on labour market outcomes and integration of immigrants into the labour market of their destination countries (Chiswick and Miller, 1995; Borjas, 2002; Hellerstein and Neumark, 2008), selection on employment (and/or job satisfaction) may be due to a lack of proficiency in the destination country's language.

In this paper, we examine how the language problems of immigrants affect their labour market performance in terms of employment, job discrimination and wages. Our laboratory is Italy where, differently from English or Spanish, the Italian language knowledge of immigrants is much lesser common². We refer this study to the determinants' literature of the migrant's language proficiency on labour market success in different destination countries. The sparse estimates we found are also controversial. Dustmann and Fabbri (2003), found a decrease of 22% on employment probability in UK for immigrants with a lack of English skills. Instead, Gonzales (2010) estimates significant negative effects of a lack on host lan-

¹Boeri and van Ours (2013).

²See Chiswick (2008) for a discussion on the importance to study less common languages.

guage skills on employment for Spain, while Yao and van Ours (2015) find that language problems affect wages, but not employment probability for immigrants in the Netherlands. On the other hand, there is unquestionable debate that the lack of language proficiency may facilitate discrimination with respect to natives, and contribute to social isolation and ghettoization in the workplace (Zschirnt and Ruedin, 2016). Although the language fluency of immigrants is just, although important, component of a larger integration process of migrants, a lack of proficiency is generally correlated to a perceived job non-satisfaction, as a result of a perceived job discrimination. However, the consequences of language proficiency that have received the most attention in the labour market concerned the effect on wages. Chiswick and Miller (2015) motivated it for the role of wages in summarizing economic status, as well the historical availability of data on wages and language proficiency of immigrants in several developed countries. Unanimously, the empirical results showed a significant relationship of the language proficiency of immigrants on wages, varying from 10 to 20 percent for the US, UK and Australia, and recording few points less for Germany and Spain, although these point estimates increased when unobserved heterogeneity and measurement errors were included³ (Yao and van Ours, 2015).

We contribute to the debate given the recent availability of 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 which has been collected between 2011 and 2012 - and published in 2014 (CSIFC 2011-2012), and the Income and Living Conditions of Households with Foreigners, which has been collected in 2009 and refers to the year 2008 for wages (ILCHF 2009). Both surveys include questions on social and economic conditions of interviewees and measures on language proficiency. The collected data in the two data sets are partly overlapping and complementary. For instance, CSIFC 2011-2012 survey includes relevant information of each household member of immigrants in the country

³See Table 5.5.of Chiswick and Miller (2015) for the extended analysis of the empirical estimates.

of origin and in Italy, but constraints to employment information about the labour market outcomes and individual perception of job discrimination. We also exploit the ILCHF 2009 - the first nation-wide survey on the socio-economic conditions of the foreign population living in Italy⁴, which contains information about employment and wages of immigrants, although it does not include items concerning job discrimination. Both datasets include self-reported language proficiency, even if information from interviewees do not overlap completely. In the CSIFC 2011-2012, the self-reported language proficiency is measured using ordered modalities which records writing and reading problems, whereas the ILCHF 2009 registers a general perception of knowledge of the Italian language. Despite this limitation, 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 language proficiency of migrants is age-different. Adult migrants typically make the decision to migrate after they have obtained their education in the country of origin, and mainly motivated from the absence of employment opportunities, while younger immigrants achieve the 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 the adult migrants in the labour market performance. Hence, the key challenge for causal estimates is that age at arrival may conceivably correlate with the language acquisition and it explains why generally the literature uses age at arrival as an instrumental variable (IV) to investigate the relationship between language skills of immigrants and labour market performance. We extend our analysis to account for unobserved heterogeneity due to the endogeneous choice of language acquisition

⁴It 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 immigrants are individuals who are born abroad and have moved to Italy, which may be the case also for Italian citizens. Although the survey refers to foreigners, in this paper we use the two terms, immigrants and foreigners, indistinctly.

and propose a propensity-score matching IV estimator (PSM-IV) which leads to estimate the "true effects" of language skills on labour market performance.

Based on these considerations, we 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 existing studies with employment's reduction which range from 20% to 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 the self-perceived job discrimination. Our findings support the evidence that reducing the lack of language proficiency in host country addresses to alleviate workplace discrimination. Third, we focus on the language proficiency on wages. We find that a good knowledge of the host country language increases significantly the expected wages of immigrants. However, part of the wage gap is persistent, also considering a scenario in which all immigrants speak a good Italian language, which leaves other channels to explain differential between Italian and immigrant wages. We also extend our analysis to investigate factors influencing the knowledge of the host country's language by the immigrant, such gender or ethnic composition. While there is a large literature showing different effects of gender on the relationship between language skill and labour market performance (e.g., Yao and van Ours (2015)), we emphasise that the composition of immigrants contribute differently on the effects of language proficiency on employment, discrimination and wages. Robustness analyses validated our main findings.

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

2. Background

2.1. The patterns of immigration in Italy

For almost a century from its unification, Italy was one of the leading European emigration countries. About twenty-six million of Italians went abroad in order to overcome 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 as a result of the restrictive policies implemented in the traditional receiving countries after the international oil crisis and identifying it as the moment of Italy's passage from emigration to immigration country. At the beginning, national returnees who came back to Italy characterized the immigration flows but the arrival of foreigners steadily increased during the 1980s, becoming the most important part of the country's international migration. These were 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), which attracted individuals seeking for higher economic opportunities, more jobs and, in turn, the promise of a better life. The initial tolerant attitude towards immigration on the public opinion and the weak answer to govern immigration contributed to consolidate regular immigration flows based on labour market quotas⁵ and to increase the share of irregular immigrants⁶(Bonifazi, 2009). In fact, the Italian government was called to set the maximum quota of foreigners allowed to enter for work reasons - such as the number of residence permits was 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 filling applications to sponsor an immigrant. The recruitment of foreign workers within this legal framework never worked

⁵Italy based immigration control on labour market quotas was introduced in 1995 and applied extensively starting from 1998 (under the Turco-Napolitano law).

⁶The 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.

properly and were mainly addressed to immigrants who were already living and working in Italy, implying that the number of applications was higher than quotas (Finotelli and Arango, 2011). The rationing of residence permits, which arose from an increasing rhetoric against immigration, favoured illegal immigration and, in turn, the need of a recurring use of the regularization programs (Mastrobuoni and Pinotti, 2015).

The first attempt to regulate the entry, residence and employment procedures of immigrants appeared in 1986 (Law 943). Successive regularisations in 1990, 1995 and 1998 had the same aim to regularise 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) made more restrictive eligibility rules linking permits to stay with work contracts, and making procedures for renewals more expensive. In fact, only those working in families or working in companies could apply while excluding self-employed, unemployed and family members. More than 705,000 applications were received and nearly 647,000 were accepted, which contributed to the growth in the size of the regular immigration from 1.8 million to almost 4.0 in the five-year period 2003-2008⁷.

Now, Italy is one of the main countries of immigration in Europe, with foreign residents. The majority of foreign nationals residing in Italy (more than five millions in 2017) are extra-European citizens, although immigrants arriving from Central and Eastern Europe from the second half of the 1990s onwards, like Romanian, represent the most important group. The immigration flow was favoured when Romania entered the European Union (January 2007), so that Romanian immigrants - along with Bulgarian - acquired the right to reside and work in Italy and "irregular" became "regular" immigrants. Thus, Romanian achieve the 23.3% of the total immigrants in Italy. Immigration from Northern Africa, especially from Morocco,

⁷There have been successively other 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 into the country for economic reasons, to be regularly employed (e.g. second Prodi decree law on immigration flows). In addition, two additional extraordinary regularizations were implemented in 2009 and 2012. During these regularizations, migration businesses often played the role of intermediary by providing immigrants with fake contracts so that they could obtain a residence permit and gain access to the formal labour market (e.g., Avallone, 2017).

represents a large part on the overall composition of the different groups according to country of origin. Moreover, data on immigrants in the last 20 years showed that immigration from Asia (China, India, Philippines) and Latin America (Peru, Ecuador) also increased sharply (IIS, various years).

2.2. Economic outcomes of 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 proficiency language on labour performance of immigrants because of its unique structural and institutional characteristics of labour market based on segmentation and rigidities. As described before, the needs of the labour force during the positive business cycles of the Italian economy and in an increasingly European integration process, have attracted different waves of migration in the last two decades. Here, we list some statistics of labour market outcomes of immigrants and natives in Italy and look for stylised theoretical explanation of these facts.

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

	Agriculture	Manufacture		Construction	Commerce	Services	Household services	Mean
		Low-tech	Hi-tech					
Foreign								
Employment (%)	4.97	13.67	6.52	9.13	9.22	37.50	18.96	
Wage (Euro)	918	1208	1382	1249	1104	1089	738	1069
Italian								
Employment (%)	3.82	11.43	9.32	6.03	14.46	53.99	0.92	
Wage (Euro)	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 observation 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 hi-tech and low-tech sectors (Pieroni and Pompei, 2015).

Table 1 lists employment prevalence and wage means of immigrants and Italian workers by sectors and, within manufacture, by hi-tech/low-tech sectors. Comparatively to Italian workers, a higher employment of immigrants is recorded in household services (e.g. domestic and care)⁸ (18.96%), construction (9.13%) and low-skill sectors (13.67%). In many of these

⁸The regularisation of Law 189/2002 accounted for a significant prevalence of women applicants form

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 the specific segments of the labour market, those that are partly rejected by natives (Ponzo and Salis, 2015).

Data also show a stable differential wage in favour to the native workers, irrespective of the employment sector. On average, Italian workers experimented in years 2015-2016 a supplementary income of 278 euro than immigrants (wages of 1368 euro for Italian workers against 1069 euro for immigrants). We also estimate formally an overall wage gap and the hourly wage gap: the Italian workers have wages and hourly wage higher over the 20% than immigrants in the year 2015-2016 and the wage gap of 18.5% is close to those estimated by National Institute of Statistics for the year 2014 (ISTAT, 2016), equal to 18.6%⁹. It is worth noting that in the hi-tech manufacture sectors, we find higher level of wages (1539 euro), 155 euro each month higher than immigrant wages.

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
Foreign									
Employment (E)	0.515	0.642	0.581	0.522	0.661	0.593	0.542	0.684	0.614
Wage (Euro)	999	1,103	1,058	973	1,062	1024	967	1,060	1,019
Italian									
Employment (E)	0.430	0.684	0.579	0.416	0.672	0.553	0.431	0.693	0.562
Wage (Euro)	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 for the tree periods. It is wort noting that the label "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). People are classified as "low educated" when the highest education degree is pre-primary, primary or lower secondary and with "high educated" when the highest education degree is upper-secondary, lower tertiary and upper tertiary.

Romania, Ukraine, Moldavia, Poland and Ecuador working in domestic help and carers.

⁹The indicator is calculated as: (wage natives-wages immigrants)/wages immigrants.

This may suggest that tasks and responsibilities performed by foreign workers, on average, are not completely equivalent to those of natives, although they may have equivalent educational backgrounds. Table 2 shows the labour market outcomes for high and low levels of education estimated for the average 2015-2016, and compared with those of the average 2011-2012 and 2009 corresponding to the years of interview of the surveys that we will investigate. The recent wave shows that Italian workers with a high education have more than 10 percentage points above the employment rate mean (68.4% with respect to 57.9%). In relative terms, also immigrant with high education are employed more than the means, even if the prevalence is decreased of 2.5 percentage points in six years. As expected, Italian low educated people are employed for the 43% in the years 2015-2016 against the 51.5% of immigrants. In addition, data suggest a mismatching between the potentiality of educational level of immigrants and their economic achievement. High educated immigrants have similar wages and lower hourly wages than Italian low-educated workers, although largely higher than wages of low-educated immigrants.

We are also interested to investigate whether gender affects differences of natives and immigrants in the observed employment and wage gap (Table 3). Employment rate by gender remained stable over years, although it is known that unemployment rate is increased with the Italian downturn economy¹⁰. In addition, using the year 2015-2016 as a reference, we can show that the wage gap mean between native and immigrant employees is 18% among men, but it raises to 25% among women, a gender pay-gap increased with respect to the previous years.

The data listed support the general idea that immigrant inequalities in Italian labour market is consequence of the selective allocation into low-status jobs, mismatching between immigrant education and job achievement, and gender pay gaps.

Some theories help to explain this disadvantage relative to natives and its persistence

¹⁰Immigrant men experimented a doubling of unemployment (e.g., from 7.8% in 2009 to 15.1% in 2015-2016).

Table 3: Differences in labour market performance by gender

	2015-2016			2011-2012			2008-2009		
	Men	Women	Average	Men	Women	Average	Men	Women	Average
Foreign									
Employment rate (E)	0.658	0.492	0.581	0.668	0.483	0.593	0.679	0.481	0.614
Wage (Euro)	1,220	898	1,058	1,177	853	1024	1,156	853	1,019
Italian									
Employment rate (E)	0.662	0.500	0.579	0.653	0.457	0.553	0.673	0.455	0.562
Wage (Euro)	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 label "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).

over time. For example, the segmented labour market theory suggests that the institutional framework differs widely between high and low skilled sectors, the first sector offering stable jobs, relatively high salaries, acceptable work conditions, and upward mobility, while the second sector is often characterized 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 significantly and largely explain differences in labour market performance. The limited knowledge about the functioning of the labour market or, more importantly, the lack of fluency in the host country's language may represent an obstacle for immigrants in the economic assimilation process and to finding better job opportunities¹¹ (Chiswick and Miller, 2003; Dustmann and Fabbri, 2003). This

¹¹This 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, 2009; Clark and Drinkwater, 2008; Friedberg, 2000).

central determinant is reinforced in Italy by the cyclical regularization and establishment of temporary resident permits which 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 (Kalter and Kogan, 2006).

Clearly, the Italian language fluency could explain not only differences in 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 make employers prefer applicants from their own culture and language or with higher affinity that is, lower social distance (Ebner and Helbling, 2016), suffering of a different treatment with respect to natives.

In order to test empirically the hypothesis that language abilities of immigrants in Italy contributes in explaining differences on labour market outcomes, we detail in next paragraphs sources and representativeness of our datasets which include unexploited variables of the Italian language skills.

3. Data

3.1. Data sources

Two national surveys were used to estimate the model parameters. The first survey was the CSIFC 2011-2012, published by the IIS in 2014. The reference population of this survey were immigrants that permanently or transitory living in Italy between May 2011 and December 2012¹². A sample of 12,000 households living in about 800 Italian cities has been then used for the interviews. The second is the ILCHF 2009, financed by the Italian Minister of Labour and Social Policies and conducted by IIS in 2009. This survey made use of the methodological framework of the survey on Income and Living Conditions, yearly carried out 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 thanks the long experience of EU-SILC, have been adapted to the specific needs implied by the foreigners' survey.

Final data sets were adjusted for some data issues. First, we considered only immigrants in age 15-64, excluding who was retired in 2011 and 2008, respectively for CSIFC and ILCHF. Second, although the survey included information on the Italian living in households with immigrants, we excluded them from the analysis since information on employment is missing, at least for CSIFC. Using this strategy, we obtained a CSIFC data set of 17,298 immigrants, in which 14,990 were of first generation immigrants, which we considered mainly in our investigation¹³, whereas the ILCHF data set included 8877 first generation of immigrants.

Table 4 lists the composition of samples by groups of immigrants' country of birth compared with the Population Census (2011) and adjusted for changes in 2012 (e.g., the average of resident immigrants between 2011 and 2012). It appears that both of surveys have immigrants' group composition similar to the resident population 2011 – 2012. Only the American

¹²The interviews have been extended until February 2013 to amply the sample of some big cities (e.g., Milan, Rome and Naples).

¹³Here, the second generation of immigrants is defined as those born in Italy by at least a foreign citizen or who was born abroad but completed a cycle of study in Italy (Di Bartolomeo and Strozza, 2014).

immigrant group appears to be slightly over-represented in the ILCHF survey (16.8% with respect to the 7.9% of the Population Census).

Table 4: Composition of samples by country of birth

Immigrant groups	CSIFC survey (2011-2012)		ILCHF survey (2009)		Pop. Census (2011-2012)	
	N.	%	N.	%	N.	%
F-Y*, Albany, Romania	5649	0.388	2930	0.330	1,613,015	0.374
Other-Europe	2961	0.203	1557	0.175	693,388	0.162
Asia	2118	0.145	1161	0.131	742,994	0.172
Cina & India	774	(0.053)	490	(0.055)	323,221	(0.075)
Africa	2840	0.195	1733	0.195	918,515	0.213
Tunisia & Morocco	1785	(0.123)	1030	(0.116)	525,189	(0.122)
America	976	0.067	1,496	0.168	342,718	0.079
Total	14,544		8,877		4,569,317	

Notes: The Table compares the conditions and social integration of foreign citizens survey (CSIFC survey 2011-2012), the income and living conditions of households with immigrants survey (ILCHI survey 2009) and the average of foreign population as extracted by the population census 2011 and population flows in the 2012. F-Y: Former Yugoslavia

3.2. Variables

Both data sets provide information concerning household members' socio-demographic characteristics in the country of origin and in Italy, outcomes of employment (E) 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 (differently treated with respect to the other workers) during your last job?*", while the ILCHF reports individual monthly wages [log wages (*WAGES*) and log hourly wages (*H_WAGE*)] that, as for EU-SILC, the amount refers to the previous calendar year (e.g., 2008). Table 5 shows the descriptive statistics for the outcomes of interest included in our analysis by gender and linguistic problems.

Concerning language abilities, in the CSIFC survey, the family member answered to the following questions: "*It is difficult for you to read in Italian?*" and "*It is difficult for you speak in Italian?*" in which the ordered choices are: often, sometimes, few times, never. Following Yao and van Ours (2015), we collected respondents with modalities often and sometimes to both questions about reading and speaking problems, and defined a dummy variable which equals one if the individual answered to this question statement that had

Table 5: Descriptive statistics, outcome variables and predictors

Variables	Any language problems	CSIFC survey (2011-2012)				ILCHF survey (2009)			
		Men		Women		Men		Women	
		No	Yes	No	Yes	No	Yes	No	Yes
<i>Labour Market Outcomes</i>									
Employment rate (E)		0.654	0.632	0.574	0.524	0.674	0.586	0.563	0.485
Observations		4922	1720	6383	1922	2633	1717	2787	1741
Job discrimination (JD)		0.816	0.806	0.839	0.798				
Observations		4692	1615	5123	1041				
Log wages (H_WAGE)						7.295	7.083	6.827	6.709
Observations						1799	1157	1455	815
Log hourly wages (WAGE)						2.205	2.038	1.963	1.854
Observations						1799	1157	1455	815
<i>Italian Language Knowledge</i>									
Speaking problems			26.95		24.05				
Reading problems			36.45		32.58				
	Poor						8.78		8.5
	Sufficient						30.691		29.951
	Discrete						37.22		42.691
	Good						23.311		18.86
Language problems (LP)*			26.95		24.05		0.394		0.384
Observations			1720		1922		1104		1738

Notes: The Table compares labour market outcomes of the CSIFC survey 2011-2012 and the ILCHF survey 2009. * LP is a dummy variable which equals one if the individual answer to this question statement that had problems either in speaking or reading and zero otherwise (LP). In the ILCHF survey, a four modality framework for language skills (LP) is grouped in whether immigrant has language problem (yes) or not.

problems either in speaking or reading and zero otherwise (LP). A four modality framework for language skills is also used in the ILCHF survey, although more generically is requested to evaluate whether the Italian language is *Poor*, *Sufficient*, *Discrete* or *Good*. Also in this case, we grouped a low level of language ability and proxy the binary aggregation (LP) whether immigrant has language problem (yes) or not.

Some important differences arise from employment outcomes among gender and language problem groups. Immigrant women have a lower probability to be employed with respect to men. In particular, when the performance is measured by employment rate, the probability of women without linguistic problems compared to the men with the same characteristic is 8 percentage point lower (65.4% for men and 57.4 for women). Within gender group, employment differences can be also extended for women with and without linguistic problems.

The probability to be employed for a women with linguistic problems is almost reduced of further 5 percentage points (from 57.4% to 52.4%). These patterns are even more important (e.g., 8 percentage points) using the ILCHF (2009) survey. We also find a large perception of job discrimination, irrespective of the language difficulties and gender differences. More than 80% of the responses suggest job discrimination of immigrants.

Table 5 also reports estimates by ILCHF of the hourly wages and log monthly wages (absolute values in parenthesis). By restricting the sample to the immigrants of first generation, we may underestimate mean wages with respect to the labour force statistics. However, if we calculate gender pay-gap, the representativeness of our sample is maintained; the estimates of the pay-gap using the ILCHF, says that men gains 35% more than women, a magnitude closes to those obtained using the labour force (i.e., 34.6)¹⁴. We extend this to the unconditional effect of language skills on wages. From the Table, we reveal that the wage gap between immigrants with hard Italian language problems and those with a good proficiency is about 15%.

Clearly, the groups of immigrants with language problems differ in terms of observable characteristics from those of immigrants without language problems. These individual and household differences concerns the immigrants' actual and past condition, if they never studied in country of origin, and the reason why he or she immigrated in Italy. In addition, other control variable 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 the childcare in the household, or whether woman in the household is an house-maker are included. To account for individual (unobserved) abilities obtained in the country of origin, which may affect its possibility of employment on the destination country net of the host country language channel, we include a proxy if immigrant has never worked into the country of origin. Descriptive statistics of these

¹⁴The gender wage gap is estimated from the following data in columns 7-8 of Table 3: $[(0.2188 - 0.1431)/0.2188]$.

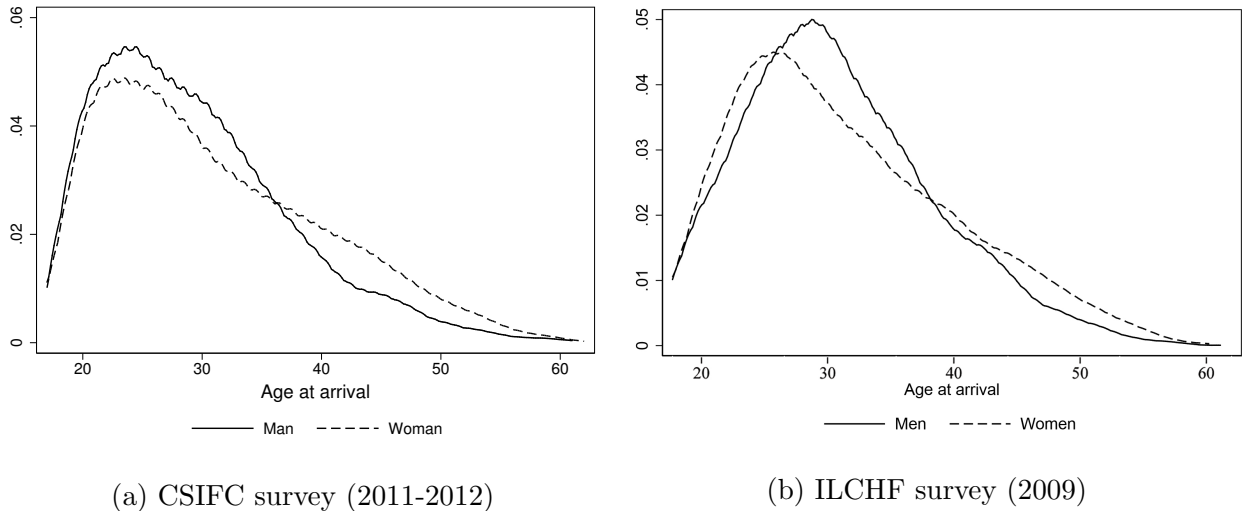
covariates by gender and linguistic problems are listed in Appendix A1.

3.3. Descriptive statistics of the key variables

Our data set is completed by creating the variable "age at arrival". This is obtained as a difference between the year in which he or she emigrates to Italy and the year of birth.

Kernel density plot in Figure 1 (a) shows the age at arrival of the first generation of immigrants for the CSIFC survey. The density appears to be enough different by gender, estimating a greater concentration for men in younger age which is overcome by women after age 35. A statistical distribution that follows the CSIFC survey is also obtained in the ILCHF survey (2009) (panel b).

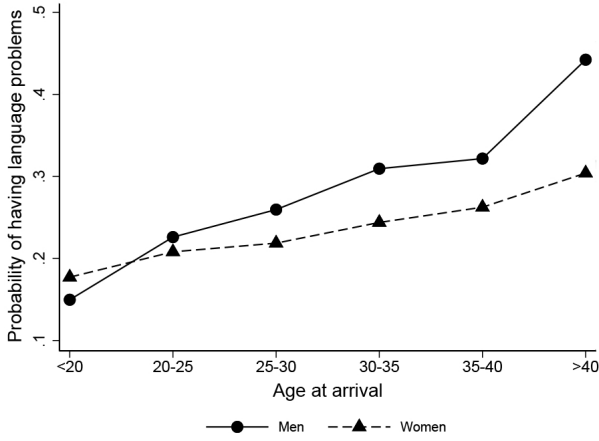
Figure 1: Kernel density plots of age at arrival



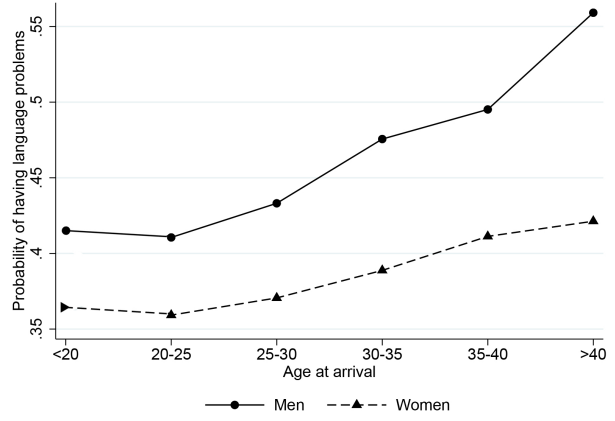
Although this evidence confirms the argument that women generally arrive later in host countries for family reunion, while men mainly immigrates before for work opportunities, it is worth noting that in Italy this delay may be overestimated by the cyclical regularisation discussed in the section 2. Indeed, the increasing demand of immigrant women from East-Europe, who from the end of the XX century have been working without permits to stay as caregivers, received support for regularisation after some years (Salmasi and Pieroni, 2015).

Figure 2 shows the probability to have language problems by age classes at arrival in

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

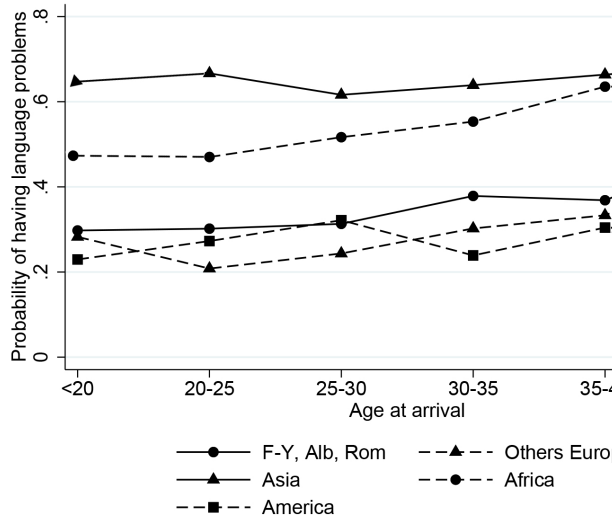


(a) CSIFC survey (2011-2012)

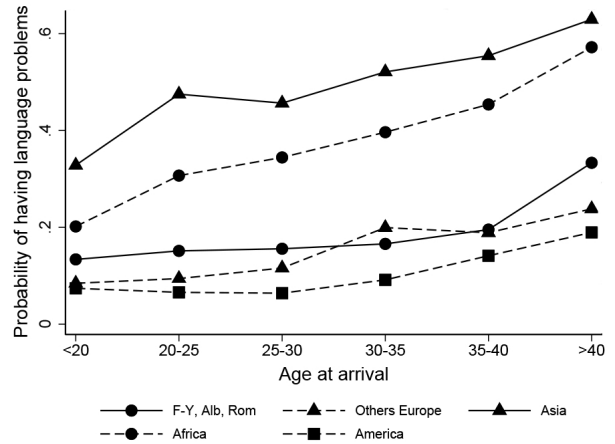


(b) ILCHF survey (2009)

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



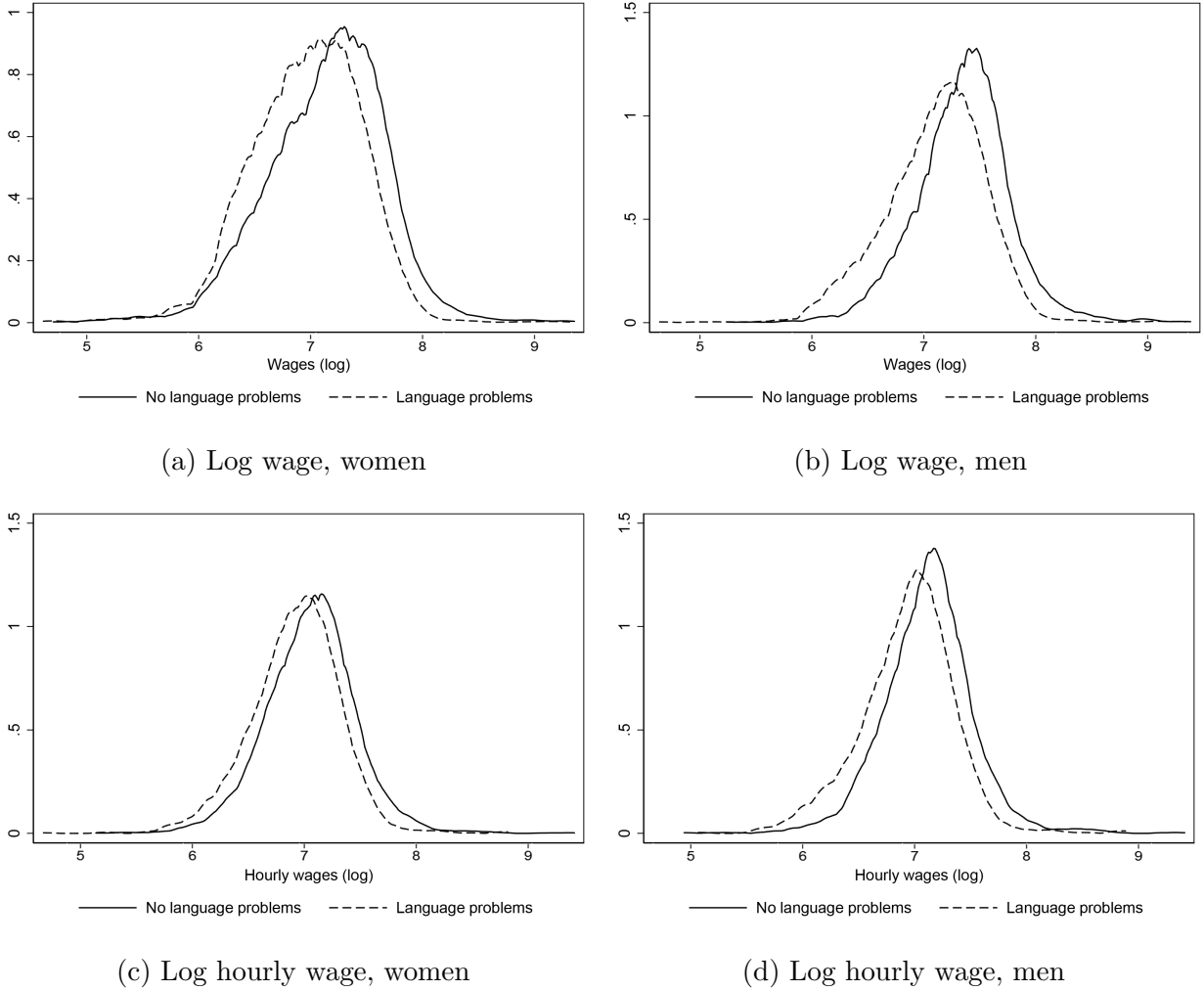
(a) CSIFC survey (2011-2012)



(b) ILCHF survey (2009)

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

Figure 4: Kernel density, ILCHF survey (2009)



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. Different groups of immigrants may also follow different patterns which may be hidden if we use the variable age at arrival of all immigrants, as an instrument to obtain estimates in these groups. Figure 3 suggests that all groups which we separate the sample 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 by the presence of 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.

4. Estimating the differential language proficiency on labour market outcomes

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

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

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

where i describes each immigrant in the dataset, y_i denotes an indicator of employment or wage performances, LP_i is dummy variable that is equal to one when the individual i has language problems and X is the vector that includes all individual and household characteristics listed in Appendix A1. The parameter of interest α_2 should measure how much more likely it is that employment, discrimination and wages is taken out in immigrants with language problems than immigrants with no-language problems.

Table 6 illustrates the OLS estimated coefficients for the two surveys separately and men and women. The coefficients of employment in the CSIFC survey (2011-2012) are of the expected sign, although in the great part close to zero in both of the surveys. Even when α_2 is significant at the usual 5% level, the magnitude of the differences in Italian language skills is up to 2.8 percentage points. More importantly, the inclusion of the conditional variables seems correcting the unobserved heterogeneity, as it is evident passing from the unconditional differences in sample women with language problems on employment rate (i.e., the difference is ten percentage points, see Table 5) to the point estimates in the conditional equation 1, where the effect of Italian language problems is smaller.

Also when we consider all sample of the ILCHF survey (2009), the coefficients for employment is smaller of 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 for immigrants may reduce by 3.8% and 4.9% the wages (log wages and log hourly wages, respectively). The relationship between language problems and wages appears driven by men sample, with point estimates 1.5%-2% higher than the average of the sample.

Table 6: Estimation results, OLS

	CSIFC survey (2011-2012)			ILCHF survey (2009)		
	All	Women	Men	All	Women	Men
Employment rate	-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
Job discrimination	0.007	0.022*	0.001			
	(0.009)	(0.013)	(0.013)			
<i>Observations</i>	11,056	7491	5637			
Log wages				-0.038***	-0.036*	-0.053***
				(0.011)	(0.015)	(0.015)
<i>Observations</i>				4432	2086	2346
Log hourly wages				-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 parenthesis. The asterisks stand for the *p-value* significance levels (* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$).

Overall, these correlations suggest that the 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 because, as anticipate, exist some sources of endogeneity which could downward or upward the bias of the estimated coefficient. Below, we discuss the strategy to account for these estimation biases.

4.1. The proposed empirical model

Endogeneity often arises by the unobserved individual selection and the presence of reverse causality. The set of conditioning variables available to us includes indicators that are

likely to be correlated with unobserved ability in the relationship between language skills and employment which sorts individuals into groups of those who do and who do not acquire the host country language, like education, ability tests and partner information. If these variables do not fully account for unobserved factors that select individuals into the group of those who are proficient and non-proficient in the Italian language, language effects may still be upward biased. We have proved it in the previous section.

To account for the language problem selection, we use a propensity-score matching (PSM) estimator of Lechner (2002), which balance covariates of immigrants having or not language problems. By 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 according to observable characteristics. We represent in Figure 5 the extent of balancing of 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, we obtain that in both CSIFC and ILCHF surveys have increased balancing in covariates. The standardised bias (%) is approximately around zero and the variance ratio of the residuals rely within the usual confidence intervals (Austin, 2009).

Thus, in order to eliminate the bias induced by differences in observable characteristics, we will estimate the coefficient α_2 by the ordinal least squares 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 a further explanation since self-reported language skill measures are subjected to substantial measurement error. Immigrants interviewed tend to over-estimate their language abilities for a incomplete knowledge of the language skills of the destination country, and for a positive propensity through the language abilities to show that they are integrated in the immigration country. This propensity leads to downward bias is only 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 instrumental variable estimator after applying PSM (IV-PSM), which should address to a causal interpretation to the language skill coefficient. An usual instrumental variable proposed by the literature is the age at the arrival in the host country by Bleakley and Chin (2004, Year 2010); Miranda and Zhu (Year 2013); Sweetman and van Ours (2008) and Yao and van Ours (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 problems in obtaining language skills. The assumption of a correct validity of the instrument for our dataset was confirmed empirically by the Figures 2 and 3. While other instrumental variables have been used in the literature¹⁵, the limited diffusion of Italian language outside Italy excludes that immigrants grew up speaking Italian language, limiting the threat to affect the relationship between age at arrival and language proficiency¹⁶.

5. Results

5.1. Baseline estimates

We first use the OLS-PSM estimator to study the effects on employment rate of the immigrants' differences in Italian language proficiency. Table 7 lists the estimated coefficients.

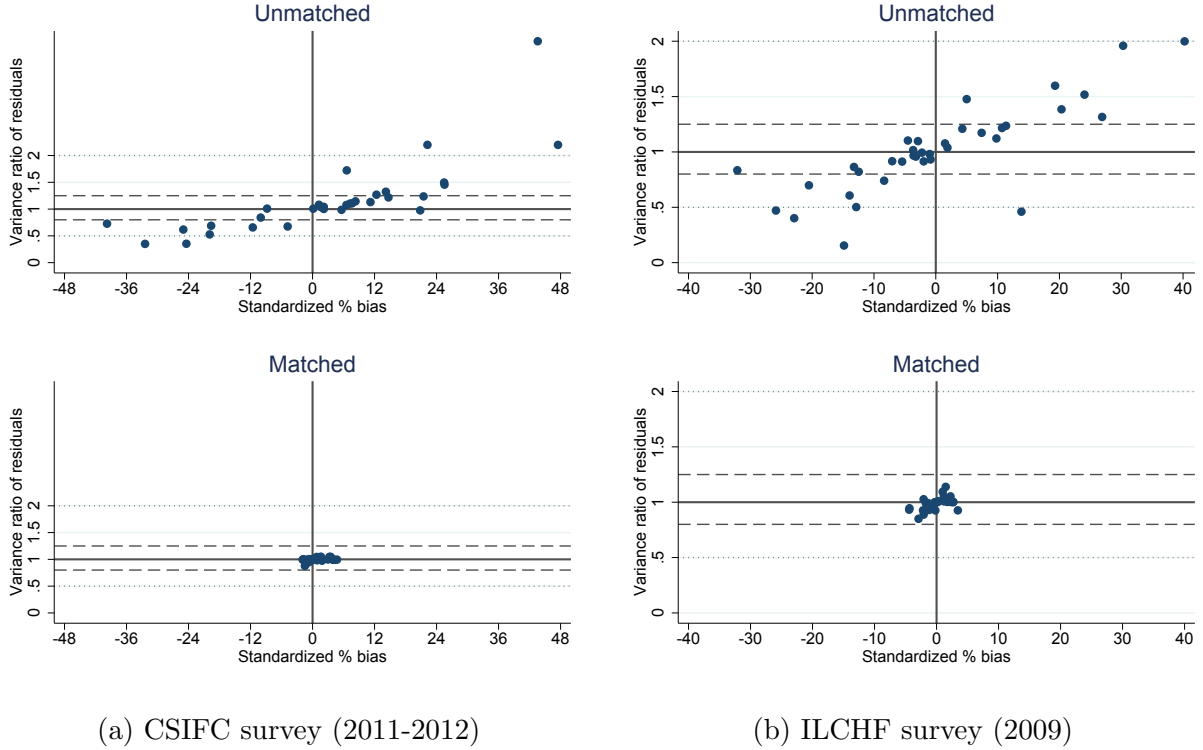
For the CSIFC 2011-2012 survey, the point estimates for employment are almost close to zero in the full sample suggesting that, if estimations would be corrected, the employment probability of immigrants with or without language problems may be substantially equivalent.

For the ILCHF survey (2009), the aggregate point estimate is instead significant, although small in magnitude, suggesting that the probability to be employed for immigrants with

¹⁵For 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).

¹⁶Estimates from CSIFC survey suggests that the fraction of immigrants which declares to spoke or read Italian language during childhood is small and is around to 0.5%.

Figure 5: Unmatching and matching of covariates



language problems decreases by 2.4 percentage points (s.e. 0.009). In addition, the point estimates by gender suggest that only men is significant in driving the divergence in the relationship between language proficiency and unemployment ($\alpha_2=0.019$; s.e.. 0.010).

Table 8 list the parameter estimates controlled by PSM for the effect of language problems on job discrimination and wage outcomes. The results show a non-significant effect of language skills on job discrimination, whereas significant negative effects are recorded on wage outcomes; in the latter case, the magnitude ranges from 4 – 7.5% and is consistent with OLS results in Table 6, irrespective of gender differences. This also implies that estimates obtained by OLS estimator with covariates is able to control for upward bias due to endogenous choice of learning the Italian language.

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' overestimate of the 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 (E)			Job discrimination (JD)		
	All	Women	Men	All	Women	Men
CSIFC survey (2011-2012)						
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
ILCHF survey (2009)						
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 parenthesis. 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 (WAGE)			Log hourly wages (H_WAGE)		
	All	Women	Men	All	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 and reading problems. Robust clustered standard errors are in parenthesis. 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 is larger and statistically significant, irrespective if estimated using the CSIFC survey (2011-2012) or ILCHF survey (2009). In both samples, we find that immigrants with language problems increased their employment rate significantly, with magnitude of the point estimates around 16 – 20

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

	Employment rate (E)			Job discrimination (JD)		
	All	Women	Men	All	Women	Men
CSIFC survey (2011-2012)						
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
ILCHF survey (2009)						
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 and reading problems. Robust clustered standard errors are in parenthesis. The asterisks stand for the *p-value* significance levels (* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$). At the bottom of the table we report the relevance of instrumental variable by F-statistics under the testing hypothesis of a weak instrument. We mark with The star indicates the rejection of the weak instrument associated to the F-statistic that exceeds 10 (Angrist and Pischke, 2009).

Table 10: Estimation results by IV matching, wages by ILCHF survey (2009)

	Log wages (WAGE)			Log hourly wages (H_WAGE)		
	All	Women	Men	All	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 and reading problems. Robust clustered standard errors are in parenthesis. The asterisks stand for the *p-value* significance levels (* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$). At the bottom of the table we report the relevance of instrumental variable by F-statistics under the testing hypothesis of a weak instrument. We mark with The star indicates the rejection of the weak instrument associated to the F-statistic that exceeds 10 (Angrist and Pischke, 2009).

percentage points. This suggests that measurement errors accounting for Italian language abilities in employment are very important and they do not vary systematically across sample surveys. In percentage, we calculate that these effects vary from 27% to 33% in the two samples¹⁷.

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 differently employment rate. Columns 2-3 of Table 9 show that all estimates are consistent with our expectations. By the CSIFC survey, men with language problems have 25 percentage points lesser to be in employment status compared to the immigrants with a good Italian proficiency. The magnitude is instead slightly smaller (minus 21.9 percentage points), when we estimate the model using the ILCHF survey. Although the confidence intervals overlaps estimates by gender, we conclude that men with a good Italian language proficiency increased the probability to be employed more than immigrants' women. The table also report the F-test for the relevance of instrumental variable, age at arrival. In fact, the estimates obtained may be biased with a weak instrument. In all estimates the F-statistics are very large than the rule of thumb (e.g., F-statistics exceeding 10), indicating that our estimates do not suffer from weak instruments¹⁸. 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 a large discrimination driven by Italian language problem exists (35 percentage points) and that this discrimination in the working place is emphasized for men (i.e., $\alpha_2 = 0.398$) mainly involved in the manufacture sectors, where a good knowledge of the Italian language is, for example, a necessary condition to achieve higher tasks within firms. Given the high percentage of responses of workers discriminated in the workplace (e.g., sample mean 0.85), our calculus suggest that workplace discrimination is increased more than 40%

¹⁷We calculate this percentage change using the mean value of immigrant employment rate of 0.6.

¹⁸We also perform a Durbin-Wu-Hausman test for endogeneity of language problems, where significant F-statistic suggests that the language indicator is endogenous. From 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 upon request by the corresponding author.

by the language proficiency.

Next, 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 language. The results show that the magnitude of these effects in decreasing wage - from 23% for the log wages to 27% for log hourly wages, are consistent with the large downward bias induced by measurement errors of Italian language abilities. As for the other labour outcomes, gender differences show that wages of immigrants with language problems are estimated to be reduced by 30% for men using the outcome of log hourly wage.

Finally, using the estimated IV coefficients, we document the magnitude of wage immigrants conditional to the presence or absence of language problems. Since the parameter identified by the instrumental variables framework measures how immigrant's wage responds to changes in Italian language abilities, we use the shares of immigrants with poor or good language skills and information from the Italian labour force surveys to estimate the specific (log) wage gap. We report results relating to immigrant wages by the two groups in Table 11 and separated by the different scenarios. Column 2-4, which refers to the parameters of language problems estimated in the year 2009, show that immigrants with a poor language proficiency have a mean wage by around 864 euro for month which increases in non-exposed immigrants to 1069 euro for month. This suggests that immigrants with a poor language ability are severely affected and, in turn, constrains their economic opportunities. The wage disparity was emphasised by gender. Women was estimated gains with 760 euro for month. In the successive columns, we also report the scenario for the year 2011-2012 and for the 2015-2016. The patterns are fully consistent with the scenario described in the sample 2009, confirming that Italian language abilities represents the main source of wage inequalities between immigrants.

In the last two lines of the Table 11, we list the contribution of language problems in explaining the wage variability of the people resident in Italy. The result that arises from

the Table indicates that using the ILCHF (2009) about 60% of wage inequality may be imputed to the problems in Italian language, while other unobserved channels may explain the remaining 40% of the wage differences between immigrants with a good proficiency in Italian language and Italian workers. The contribution of the Italian language skills to the wage inequality of immigrants appears to be driven by men. In the ILCHF (2009) survey, more than 70% of the wage variability of immigrants and natives was explained through the language problems of immigrants. These gender differences seem reducing in the successive years under the spinta of the financial crisis which, in turn, affected the real economy and made less important the expectation of employment and careers of immigrants.

Table 11: Conditional wage of immigrants by language proficiency and wage inequality: estimates

	Scenario 2009			Scenario 2011-2012			Scenario 2015-2016		
	All	Women	Men	All	Women	Men	All	Women	Men
ILCHF survey 2009									
Language problems (α_2)	-23.8	-16.2	-27.2	-23.8	-16.2	-27.2	-23.8	-16.2	-27.2
Shares of poor language	0.245	0.240	0.270	0.245	0.240	0.270	0.245	0.240	0.270
Shares of good language	0.755	0.760	0.730	0.755	0.760	0.730	0.755	0.760	0.730
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
Conditional estimates for immigrant wage									
Immigrant Wages (Poor language, LP=1)	*	864	760	965	868	760	982	942	800
Immigrant Wages (Good language, LP=0)	*	1069	883	1227	1075	884	1249	1166	929
Contribution of language skills to wage inequality (%)	**	59.60	41.30	71.70	54.20	36.43	66.15	58.50	36.40
Other channels contributing to wage inequality (%)	**	40.40	58.70	28.20	45.80	63.57	33.85	41.50	63.60

Notes: Language problems (α_2) are defined as having either speaking and reading problems.

5.2. Heterogeneous estimates

To be completed

5.3. Robustness

To be completed

6. Conclusion

To be completed

Appendix A1 - Descriptive statistics, covariates

Variables	Any language problems	CSIFC survey (2014)				ILCHF survey (2009)			
		Men		Women		Men		Women	
		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				
	Mother with children	0.033	0.017	0.116	0.081				
Family size	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				
Education degree obtained in Italy(No)						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
Housemaker		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				
	Municipalities (less 10.000)	0.5	0.537	0.516	0.507				
Area of residence	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

References

- ADSERÀ, A. and PYTLIKOVÁ, M. (2016). *Language and Migration*. The Palgrave Handbook of Economics and Language, 342-372.
- 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.
- BLEAKLEY, H. and CHIN, A. (2004). Language Skills and Earnings: Evidence from Childhood Immigrants. *The Review of Economics and Statistics*, **86** (2), 481–496.
- and — (Year 2010). Age at arrival, english proficiency, and social assimilation among u.s. immigrants. *American Economic Review, Applied Economics*, **2** (1), 165-192.
- BOERI, T. and VAN OURS, J. (2013). *The Economics of Imperfect Labor Markets: Second Edition*. Princeton University Press, stu - student edition edn.
- BONIFAZI, C. (2009). *The Italian transition from migration to immigration country*. Tech. rep., IRPPS Working Papers.
- BORJAS, G. J. (2002). Homeownership in the immigrant population. *Journal of Urban Economics*, **52** (3), 448 – 476.
- CHISWICK, B. and MILLER, P. (2003). The complementarity of language and other human capital: immigrant earnings in canada. *Economics of Education Review*, **22** (5), 469–480.
- and 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.
- CHISWICK, B. R. (2008). *The Economics of Language: An Introduction and Overview*. IZA Discussion Papers 3568, Institute for the Study of Labor (IZA).
- and MILLER, P. W. (1995). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics*, **13** (2), 246–288.
- CHSWICK, B. R. and MILLER, P. W. (2009). Earnings and occupational attainment among immigrants¹. *Industrial Relations: A Journal of Economy and Society*, **48** (3), 454–465.
- CLARK, K. and DRINKWATER, S. (2008). The labour-market performance of recent migrants. *Oxford Review of Economic Policy*, **24** (3), 495–516.
- DI BARTOLOMEO, A. and STROZZA, S. (2014). Immigrants living in the EU15 countries and their conditions of integration in the labour market. *RIEDS - Rivista Italiana di Economia, Demografia e Statistica - Italian Review of Economics, Demography and Statistics*, **0** (3-4), 95–102.

- DUSTMANN, C. and FABRI, F. (2003). Language proficiency and labour market performance of immigrants in the uk*. *The Economic Journal*, **113** (489), 695–717.
- and 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. and HELBLING, M. (2016). Social distance and wage inequalities for immigrants in switzerland. *Work, Employment and Society*, **30** (3), 436–454.
- FINOTELLI, C. and 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 Economa de la Educacin* 5, Asociacin de Economa de la Educacin, pp. 798–824.
- HEATH, A. and CHEUNG, S. (2007). *Unequal chances: Ethnic minorities in Western labour markets*. British Academy/Oxford University Press, Oxford.
- HELLERSTEIN, J. K. and NEUMARK, D. (2008). Workplace segregation in the united states: Race, ethnicity, and skill. *Review of Economics and Statistics*, **90** (3), 459–477.
- KALTER, F. and 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. and 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. and ZHU, Y. (Year 2013). English deficiency and the native immigrant wage gap. *Economics Letters*, **118** (1), 3841.
- PIERONI, L. and POMPEI, F. (2015). Evaluating innovation and labour market relationships: the case of italy. *Cambridge Journal of Economics*, **32**, 325347.
- PONZO, F. C. M. J. F. L., I. and SALIS, E. (2015). Is the economic crisis in southern europe turning into a migrant integration crisis?. *Politiche Sociali*, **1**, 5988.
- SALMASI, L. and PIERONI, L. (2015). Immigration policy and birth weight: Positive externalities in italian law. *Journal of Health Economics*, **43**, 128 – 139.
- SWEETMAN, A. and 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 11411193.

- YAO, Y. and VAN OURS, J. C. (2015). Language skills and labor market performance of immigrants in the netherlands. *Labour Economics*, **34**, 76 – 85.
- ZSCHIRNT, E. and 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.