RESEARCH PAPER - ITALY AND EUROPE

Do informal Networks Increase Migrants' Over-Education? Comparing Over-Education for Natives, Migrants and Second Generations in Italy and Assessing the Role of Networks in Generating It

Pierre Georges Van Wolleghem[1] · Marina De Angelis[2] · Sergio Scicchitano[2,3]

Abstract

Whilst migration has become a structural feature of most European countries, the integration of foreigners in the labour market continues to raise concerns. Evidence across countries shows that migrants are more often over-educated than natives. Over the last few years, scholarship has intended to capture the effect of informal networks on migrants' over-education. Interestingly, no study has looked into the Italian case, a country for which the effect of networks on education-occupation mismatch is well documented. This article has two objectives: it assesses the extent to which over-education affects migrants and it evaluates the role informal networks play in producing it. We find that migrants have a higher probability of being over-educated than natives and second-generation migrants. Likewise, we find little evidence of a differentiated effect of networks as they tend to increase migrants' over-education whilst decreasing it for natives and second-generation migrants. Empirical evidence is drawn from the application of causal inference modelling to PLUS 2018—Participation, Labour, Unemployment Survey.

✉ Pierre Georges Van Wolleghem
pierre.vanwolleghem@uib.no

[1] Department of Comparative Politics, University of Bergen, Postboks 7802, 5020 Bergen, Norway

[2] INAPP, Italian National Institute for the Analysis of Public Policies, Corso d'Italia, 33, 00198 Rome, Italy

[3] Global Labor Organization (GLO), Essen, Germany

 Springer

1 Introduction

Whilst migration has become a structural feature of most European countries, the integration of foreigners into the labour market continues to raise concerns as evidence across the board highlights the concentration of migrants in low-skilled occupations, despite their educational level (OECD 2018). In an ageing Europe, migration presents indubitable positive economic effects. By feeding the workforce, it alleviates the old-age dependency ratio (the number of workers compared to that of pensioners) and the risks looming over the European population's ability to sustain its economy. But migrants' contribution to their receiving country's economy is by no means immediate. Coming from different cultural, linguistic and institutional backgrounds, migrants need to adapt to a reasonable extent to the pre-existing structures of their receiving societies (Penninx and Garcés-Mascareñas 2016). On the other hand, migrants may also face resistance to their integration from their receiving countries; resistance that may outweigh their skills and competences in the labour market (Wrench et al. 2016; Fellini and Fullin 2018). Consequently, there may exist a mismatch between foreigners' education and their occupation; a misallocation of work forces that would, ideally, be temporary and end swiftly.

Scholarship tends to concur on the existence of a complementarity between domestic and foreign labour forces (Dustmann et al. 2005; Esposito et al. 2019). Yet, there appears to be a significant difference between education-occupation mismatch for foreigners and natives (Piracha and Vadean, 2013; McGowan and Andrews, 2015; Reyneri and Fullin 2010). The specialised literature has put forth a series of explanations such as imperfect information (Dolado et al. 2009), imperfect transferability of human capital across borders (Chiswick and Miller 2009) or even work experience—or mismatch—in the country of origin (Piracha et al. 2012). Over the years though, the attention has been moving towards the role of job-search channels in generating mismatch, with particular emphasis on informal ones (Chort 2016; Kalfa and Piracha 2018; Alaverdyan and Zaharieva 2019). Namely, resorting to family and friends to look for and find a job would be associated (or not) with over-qualification. Studies have thus far investigated specific migrant communities or specific countries. None of them has investigated the Italian case, despite it being a country where the relationship between informal channels and mismatch is well-documented (Mosca and Pastore 2008).

Given the foregoing, this article has two main objectives. On the one hand, it assesses the extent of the phenomenon; namely, are first- and second-generation migrants more over-educated than nationals, *ceteris paribus*? If yes, by how much? On the other hand, this article seeks to shed light on the role informal channels play in generating it: do they lead to more or less over-education? Are their effects different for foreigners and natives? In order to answer these questions, we rely on PLUS data—Participation, Labour, Unemployment Survey—a survey conducted by the Italian National Institute for Public Policy Analysis (INAPP).

Our empirical strategy is twofold. Firstly, we compute the probability of over-qualification for natives and migrants through probit regressions. Secondly, we borrow from causal modelling methods to reinforce our causal claims. We thus implement a propensity score matching (PSM) model to reduce the distance that separates natives

and foreigners and make them more comparable. We then rely on inverse-probability treatment weighting methods (IPW) to compare the probability of over-education for natives, foreigners who grew up predominantly in Italy and foreigners who migrated later on in life.

We find evidence of a difference in over-education between foreigners and natives; a difference that is clearer when we distinguish between the foreigners who grew up in Italy and those who migrated. The former are undistinguishable from natives whilst the latter tend to be significantly more over-educated. These findings are in line with the European (Portes et al. 2016; Crul et al. 2017) and North-American (Alba and Nee 1997) literature on the matter. As for the effect of networks, we show that there is (weak) evidence of a differentiated effect from one category to another: while they appear to increase the probability of mismatch for migrants, they play the opposite role for natives and second-generation migrants. This finding is somewhat in line with the results presented elsewhere; namely Kalfa and Piracha's (2018) for the Australian case, and Alaverdyan and Zaharieva's (2019) for the German one; but goes farther in studying their effect for second generations. Our article is divided as follows. The second section selectively reviews the specialised literature on migrants' over-education and on education-occupation mismatch. We also refine our contribution to said literature. The third section describes our data and outlines our empirical strategy. More information on method and robustness tests are provided in the appendix (Supplementary material). The fourth section presents our empirical results while the fifth discusses them and proposes ways to move forward with policy considerations. We conclude in a sixth section.

2 Over-Qualification of Migrants and Definitions of Mismatch

2.1 The Over-Qualification of Migrants

The adjustment of migrants into their receiving country's labour market has been extensively studied, with origins dating as far back as the early Chicago School, at the instigation of Park et al. (1921). Using natives as the gold standard, scholarship has intended to assess the extent to which immigrants become more similar to natives in terms of earnings and occupation (Constant and Zimmermann 2013). In this respect, the pioneering work of Chiswick (1978) has considerably influenced current scholarship. He posited that, due to imperfect transferability of human capital across borders, migrants would initially have lower earnings than natives. With the passage of time, migrants would gain information and adjust to the functioning of the labour market; they would eventually catch up with natives' earnings.

More recently, the attention has turned to another, related and important feature of labour market performance, that of the match between education level and occupation; more specifically, over-education. Whilst mismatch is an issue that generally affects labour markets in many economies (McGowan and Andrews 2015), including Italy (Caroleo and Pastore 2018; Gaeta et al. 2017; Pastore 2019), evidence tends to converge towards the existence of a significant difference between natives and foreigners: migrants are consistently more over-educated for the positions they occupy

than natives (Piracha and Vadean 2013; McGowan and Andrews 2015). According to European Social Survey data covering the years 2000–2009, 13% of the native respondents across 22 European countries were over-educated for their jobs whilst the figure rose to 22% for foreigners (Aleksynska and Tritah 2013).

Different explanations were put to the test. Drawing upon Chiswick's (1978) work on earnings, some studies have posited the role of information adjustment. Accordingly, migrants would need to become acquainted with the functioning of the labour market in order to fully transfer their human capital (Chiswick and Miller 2009). Differently, Mattoo et al. (2008) have pointed to the role of the quality of the human capital that is being transferred. They notably argue that a large part of over-education is due to the attributes of the country of origin, which affect the quality of the human capital transferred.[1] Instead, Piracha et al. (2012) have posited the effect of mismatch in the country of origin on mismatch in the receiving country. Yet other explanations have put forth the role played by the characteristics of the receiving society. Cultural proximity (especially regarding language[2]) and natives' attitudes towards foreigners (namely discrimination[3]) likely affect labour market integration. Over the years though, the attention has moved towards the role of informal job-search channels in generating over-education (Chort 2016; Kalfa and Piracha 2018; Alaverdyan and Zaharieva 2019). Informal channels can be conceived as an ensemble of resources accessible to the members of a more or less institutionalised network (Bourdieu 1986). As resources, networks can play a positive role in labour market integration, which is however conditioned by the resources networks actually possess. As not all networks have equal resources, their effect likely varies from one social group to another. Inspired by the vast literature on the effect of referral hiring on mismatch (Pistaferri 1999; Meliciani and Radicchia 2016), the few studies available thus far have produced no decisive conclusions. Kalfa and Piracha (2018) have demonstrated that social capital exacerbates migrants' education-occupation mismatch in the Australian labour market. Alaverdyan and Zaharieva (2019) present concurring results for the German case. Conversely, Chort (2016) studies the effect of the use of informal channels within the Senegalese community across four countries—France, Italy, Mauritania and Ivory Coast—and concludes to the positive effect of networks on education-occupation match. This article contributes to this growing strand in literature for several reasons. Firstly, we investigate the Italian case, a country for which the relationship between informal channels and mismatch is well-documented (Pistaferri 1999; Mosca and Pastore 2008; Meliciani and Radicchia 2016).

Secondly, we rely on respondents' declared use of informal networks. Whereas other studies have relied on composite and indirect indicators of network use (Kalfa and Piracha 2018), we consider the intensity of networks use by the respondents, which we compare to eleven other channels—amongst which are job centres, professional networks, and temporary work agencies—as well as the occurrences in which networks have actually led to employment.

---

[1] Mattoo et al. (2008) refer mostly to expenditure on tertiary education in countries of origin.

[2] Language proficiency was proven to significantly affect migrants' labour market outcomes (Dustmann and Van Soest 2002).

[3] In this regard, see Neumark 2013; Wrench et al. 2016.

Finally, another significant contribution lies with the decomposition of the foreign population into different groups; namely separating migrants with a migration background (i.e. who grew up predominantly in Italy from 0 to 18 years old) from people who migrated from their home country to Italy subsequently. We also provide an alternative definition of second-generation migrants based on whether they arrived in Italy before secondary school. The specialised literature has highlighted the processes of inter-generational changes (Alba and Nee 1997). Newly arrived migrants are socialised into the values and expectations of their country of origin whereas their children are brought up within the educational system of the country of destination. Foreigners who grew up and went to school in Italy are more likely to be acquainted with Italy's institutions, economy and labour market. Conversely, those who migrated to Italy have had to learn some Italian and gain information on the labour market's functioning.

2.2 Defining Mismatch and Over-Education

There exist three main ways to measure education-occupation mismatch (ILO 2018): the normative approach, workers' self-assessment and the statistical approach. All three approaches carry different information and do not necessarily overlap with one another (ILO 2018). In the first approach, mismatch is measured using a classification elaborated ex-ante, which specifies the level of educational attainment required for each occupation. Whilst this method is regarded as perhaps the most accurate (Green et al. 2007), it requires extensive data and is therefore discarded for the purpose of this article.

Differently, self-assessment provides the workers' perception of their own mismatch. Whilst an interesting dimension, it does not necessarily fit every research purpose[4] inasmuch as it can be affected by classification error as the researcher does not know how the respondent elaborated her/his judgment (Chevallier 2003). Similarly, this method is presumably little suitable if it comes to comparing different groups with intrinsic differences as they likely display very different characteristics underlying their perception of mismatch.[5]

Finally, the statistical approach is based on the distribution of workers' education levels within occupational groups. Whilst not as precise as the normative approach, its relying on statistical distributions, and the distance of a given worker from the latter, provides a relatively objective measure of education-occupation mismatch, all the more so if we intend to compare two groups such as migrants and natives. More precisely, since it is based on the distance of an individual from the mean, or the mode education level (depending on the method chosen; we chose the latter, see below), it allows for measuring whether this distance is statistically more often observed for natives or for foreigners, all other things being equal. This measurement is not exempt of drawbacks as it is sensitive to the aggregation level of the occupations' classification

---

[4] One of the issues associated with self-assessed mismatch lies with the possible confusion between vertical (education level) and horizontal (field of study) mismatch (see Chort 2016; Robst 2007).

[5] As Borjas (1988) argued, migrants are not sorted at random but self-selected; they rationally decide whether and where to migrate by comparing various opportunities.

as well as cohort effects (since it is based on observed distribution of education for a given occupation; see Chevallier 2003). Notwithstanding, and given the cross-section nature of our analysis, we consider that the third approach is here the best alternative to the normative method. Not as data-demanding as the latter, it is more suitable for our purposes than the subjective perception of mismatch, in that it allows comparison between different groups without considering underlying determinants of perceptions, as would be the case with the self-assessment method.

3 Over-Education in Italy: Data, Descriptive Statistics and Empirical Strategy

3.1 Empirical Strategy, Data and Variables

Our analysis relies on data collected by the Italian National Institute for the Analysis of Public Policy (INAPP) through PLUS (Participation, Labour and Unemployment Survey), a survey on the Italian population of working age, created to provide reliable information on labour market phenomena that are rare or less explored by other data sources, such as the use of networks. The sample we use is a subset of PLUS 2018 (for more detail, see online appendix). It counts a total of 31,600 respondents either in work or available for work, of which 2.4% are foreigners. The relatively low proportion of foreigners in the sample is in contrast with the percentage of foreigners in Italy, estimated at around 8.5% in 2018 (Blangiardo 2019, p. 21).[6] Resultantly, we opt for a twofold methodological approach that rests on probit regressions on the one hand, causal modelling on the other. Given the low share of foreigners, it is likely that the results we obtain through probit models on the whole sample are biased and over-determined by the portion of natives in the data. Resorting to causal models produces more robust results as they do not rely on sample representativeness but select and weigh the observations considered in order to facilitate comparison between foreigners and natives. We thus implement a propensity score matching model (PSM; Rosenbaum and Rubin 1983) where we consider being foreigner as the 'treatment' and create a control group of natives with similar characteristics. We further break down the effect of being foreign into two categories: those who prevalently grew up abroad and those who prevalently grew up in Italy. For this, we rely on the inverse-probability of treatment weighting method (IPW; Feng et al. 2012). A caveat is here in order. If causal models allow us to reduce the distance between foreigners and natives, they do not remove it completely as some features remain fundamentally different (education obtained abroad may have different standards from that obtained in Italy, as we point out in the discussion section; language skills may remain an obstacle to the

---

[6] Despite this drawback, we prefer PLUS to other databases in which the share of foreigners is more representative, such as the Labour Force Survey (LFS). This is because PLUS contains more usable information on network use. Firstly, the question in LFS includes unions along with friends and family, thus affecting the definition of informal networks. Secondly, the question on networks in LFS is only asked to a sub-sample of respondents; namely, those who are looking for a new job or a second job, and contains information for about 0.9% of the total number of respondents.

realisation of one's potential even years after migration; to give but a few examples). These differences shall be borne in mind whilst interpreting our results.

In accordance with the definition of mismatch introduced above, a worker is regarded as under- or over-educated if their education level is respectively lesser or greater than the mean or modal educational level (ILO 2018). Because we focus on over-education, we do not investigate under-education (our dependent variable is thus equal to one in case of over-education, zero otherwise). Whilst the original version of the statistical approach rests on the mean of the education distribution (Verdugo and Verdugo 1989), we here follow Kiker et al. measurement (1997) which compares the actual level of education of an individual worker to the modal level of education of all workers in their occupational group. This choice is notably due to the nature of the data at hand. Education levels and occupation groups are defined in accordance with international standard classifications; namely ISCED and ISCO one digit. Whereas ISCED one digit provides a fair idea of people's education levels, ISCO one digit is an aggregation levels that does not allow a fine-grained analysis. It is however good enough to study differentiated distributions for two to three groups. For the purpose of this article, we shall focus on over-qualification. Mismatch is calculated on the basis of the modal education level of those in employment for a given occupation category. Note that we consider mismatch with regard to current employment but also former employment in case the respondent is unemployed and looking for a job. The mode education for the employed is thus applied to the unemployed. Our dependent variable is binary with value 1 in case of over-education, 0 otherwise. Table 1 below summarises the data as to education, occupation and over-qualification for the two groups under scrutiny.

The use of informal networks is captured by two different questions in the questionnaire. Firstly, "among the following *job-search* methods, can you tell me which one you have used and to what extent?" Respondents were shown 12 items and had the possibility to indicate the intensity of their use on a scale from 0 to 10. The item that interests us most is the one reading: "friends, relatives and acquaintances". In order to take account of all 12 items, we ranked them according to their intensity of use in order to obtain a more precise idea of the way respondents look for jobs. We then multiplied the inverse rank by the declared intensity (on a 0–10 scale) to obtain a more precise measurement of their use, and then divided the result by 10 so as to have a more interpretable scale, ranging from 0 (little use compared to other job-hunt means, low intensity) to 12 (most used channel and one used intensely). Secondly, we considered the channel through which respondents *obtained* their current job as another indicator of the use of networks. Note that the two measurements are highly correlated (with a Pearson's polychoric coefficient of 0.71). The limitation of this variable lies in the fact that we do not have information regarding the composition of networks, whether they are principally made up of co-ethnics or if they are mixed networks (Neumark 2013; Battisti et al. 2021). We attempt to limit this source of uncertainty by running additional models that include the number of foreigners from the same country of origin residing in the same province as the respondent. This is intended as a proxy for the size of potential networks [see appendix(Supplementary material)]. In addition, the use of network likely depends on other observable and unobservable characteristics

Table 1 Education, Occupation and Over-qualification in Italy for Italians and foreigners available for work (%)

	Nationals	Foreigners	All
Education levels			
Elementary	1	2.1	1
Lower middle	16	17.8	16
High school	48.6	49.6	48.7
Bachelor	30.5	27.1	30.4
Post bachelor	3.9	3.4	3.8
Total	100	100	100
Occupation classification			
Chief executives, senior officials and legislators	4.5	3.5	4.5
Intellectual and science professionals	24.5	14.9	24.3
Technicians and associated professionals	16.7	13	16.7
Clerical support workers	20.5	10	20.3
Service and sales qualified workers	13.9	34.2	14.4
Skilled and agricultural workers, Craft and related trades workers	9.8	7.9	9.7
Plant and machine operators and assemblers	3.8	3.5	3.7
Elementary occupations	6.1	12.8	6.2
Armed forces occupations	0.3	0.2	0.3
Total	100	100	100
Over-education			
Chief executives, senior officials and legislators	10.5	6	10.4
Intellectual and science professionals	15.1	8.2	14.9
Technicians and associated professionals	32.6	26.1	32.4
Clerical support workers	28.9	17.9	28.6
Service and sales qualified workers	8.8	29.1	9.4
Skilled and agricultural workers, Craft and related trades workers	1.8	2.2	1.9
Plant and machine operators and assemblers	0.7	0.8	0.7
Elementary occupations	1.2	9.7	1.4
Armed forces occupations	0.4	0	0.4
Total	100	100	100

which may affect the results of our analyses. To keep the endogeneity issue in check, we run a series of robustness check on PLUS panel data.

As for the foreign population, we distributed people who were not born with Italian citizenship[7] into two categories: migrants and second-generation migrants. Whilst the definition of migrant is little problematic for our purposes (migrants are those who prevalently grew up abroad and migrated to Italy), the definition of second-generation migrant is less straightforward. In migration studies, the term refers either to the offspring of parents who were born and/or raised up in the country of residence (Schneider 2016). In our study, we consider as second-generation migrants those who were predominantly raised in Italy (from 0 to 18 years old). The latter category accounts for 32% of the non-natives in our dataset whilst the category migrant accounts for 68% of them.[8] In addition, we also provide an alternative definition of first- and second- generation migrants, which is based on the age at which our respondents arrived in Italy. In this second approach, migrants are those who arrived in Italy at 13 years old or older whilst second-generation migrants are those who arrived at 12 or younger; that is, roughly before the start of secondary school. More detail on this is provided in the appendix (Supplementary material). The data at hand does not allow us to distinguish migrants according to their residence motives (work, family or international protection), still a determinant of labour market outcomes (Connor 2010).

We control for a series of factors that may affect mismatch: area of residence[9] (North, Centre or South; see Ambrosini 2013), type of agglomeration in which the respondent resides,[10] gender (1 being male), whether the respondents have children (1 being yes; 0 being no), work status (0 being in occupation; 1 being unemployed), father's education[11] and sector of activity (whether public or private; public being 1).[12] Finally, a further control aims at capturing the period in which mismatch occurred. It thus consists of the year in which occupied respondents took up their current job and the year in which unemployed respondents had their last job.

---

[7] Thus excluding Italian nationals born abroad who returned to Italy.

[8] Foreigners from EU15 countries were excluded from the analyses as EU15 states tend to be destination countries among which their citizens can move and settle freely. Pressure to migrate (upstream) and to settle (downstream) are thus not as high as for other countries of origin. We provide more detail on this, on foreigners' characteristics and areas of origin in the online appendix.

[9] The areas of residence thus defined find support in Migration Studies literature. That being stated, further disaggregation at the local level would provide sounder results as the particularities of the local labour markets may influence the probability of over-education and the use of networks. We attempt to control for that with the areas of residence together with an indicator of rurality (the type of agglomeration).

[10] See appendix (Supplementary material) for more detail on categories.

[11] The education level of the respondent's father aims to capture the skills of the respondents and thus controls for the possibility that skills affect both the likelihood of being over-educated and the likelihood of being a migrant.

[12] See appendix (Supplementary material) for more detail on categories.

3.2 Methodology

This section briefly outlines the methodology used. More detail is available in the online appendix. As already stated, our approach is twofold. Firstly, we run a series of probit regressions and report the results of 10 models. The first four are run on the whole sample. M5 through M10 are partial regressions on three subsamples: migrants, second-generation migrants and natives. These allow comparing the effect of our variables across groups. M1, M3, M9 and additional models are run on panel data to control for endogeneity biases. Secondly, we resort to causal models, namely Propensity Score Matching (PSM) and Inverse Probability Weighting (IPW). For PSM, the scores are computed on the probability (logistic model) of a respondent being a foreigner considering gender, the presence of children, whether the respondent works in the public or private sector, the area of residence, age, city size, education, father's education and work status. The appendix (Supplementary material) details the approach followed, formulae and additional robustness tests. The matching methods used to associate observations in the treatment group with observations in the control group are: nearest neighbour, Kernel matching and radius matching (with radius size = 0.1). Reported standard errors are bootstrapped standard errors, 50 repetitions, or robust standard errors, as specified in the tables below. As for IPW, the approach corrects selection bias through attributing each observation a weight inversely proportional to its probability of selection (Cattaneo 2010). Thus, we calculate observations' weights according to their propensity to fall into any of our three categories—native, migrant or migration background. For the propensity model, we use the same covariates as those used to calculate propensity scores with the logistic model above, but we use a multinomial logistic model to account for our three categories. The appendix (Supplementary material) provides more detail on the method.

3.3 Robustness of Our Results: PSM and IPW

This section provides essential elements relating to the soundness of the method used. More results and tests are available in the appendix (Supplementary material). Starting with PSM, in order for the method to yield reliable results, the two subsamples in our analysis (natives and foreigners), must overlap significantly, so that any combination of characteristics present in one group is also observed in the other group (the common support condition). Meeting the common support condition therefore ensures the quality of matching processes. Figure 1 below produces the density histograms of the propensity scores for natives and foreigners.

As can be seen, there is significant overlap between the two groups at the lower end of the x-axis, less so at the higher end. Consequently, the application of the matching algorithms presented above restrict matching to those observations in the group of foreigners which have their counterpart in the group of natives, and vice versa. Turning to the quality of the matching analyses on overlapping groups; i.e. on the regions of common support, a look at the difference between the distribution of the propensity scores as calculated before matching for both groups under scrutiny, and the distribution of propensity scores once the observations in the two groups are
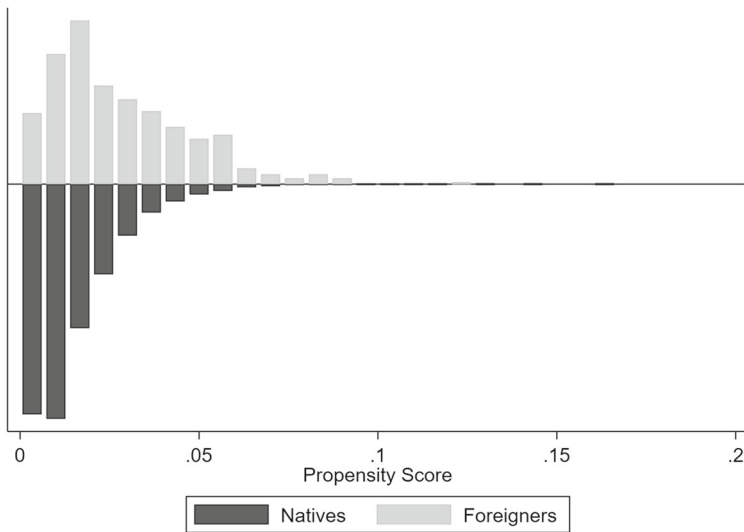
Fig.1 Density histograms of propensity scores for natives and foreigners

matched, reveals that the match between the two groups corrects the distribution of propensity scores in an adequate way (Fig. 2).

Moving to IPW, the main conditions for its validity is the overlap of the weights (see appendix (Supplementary material) for more on the size of the weights, another important feature to consider). Our weights appear to be fairly concentrated in the same area between the three groups, bearing evidence that the overlap assumption is met (Fig. 3).

4 Empirical Results

4.1 Probit Regressions

Table 2 below reports the average marginal effects yielded by the regressions run. Overall, there does not appear to be any effect for being a foreign citizen. If the coefficients are of relevant magnitude (M1 and M3), they are not statistically significant. Conversely, it appears that the foreigners who grew up in Italy are less likely (by 4.8–6.7 p.p.) to be over-qualified than natives whilst there is no evidence that migrants who prevalently grew up abroad are more over-qualified than natives (M2 and M4).

On a different note, the effect of informal networks is associated to a decrease in the probability of over-education. While looking for a job, a one-point increase of the intensity of use of informal networks translates into a 0.4 p.p. decrease in the probability of over-education (M1 and M2). Considering the fact of finding a job through informal networks, the probability of over-education decreases by 6.5 p.p. (M3 and M4).
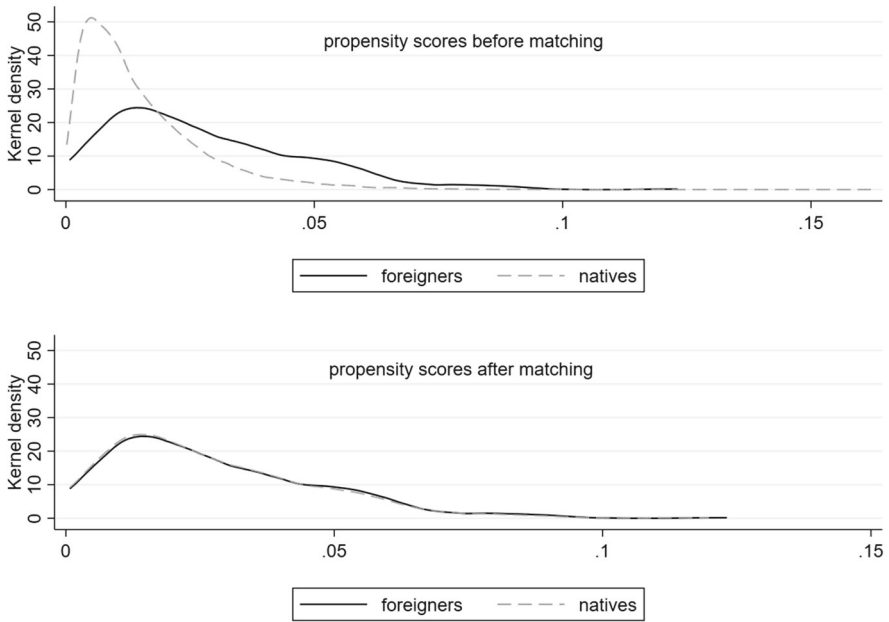
Fig. 2 Distribution of propensity scores before and after matching
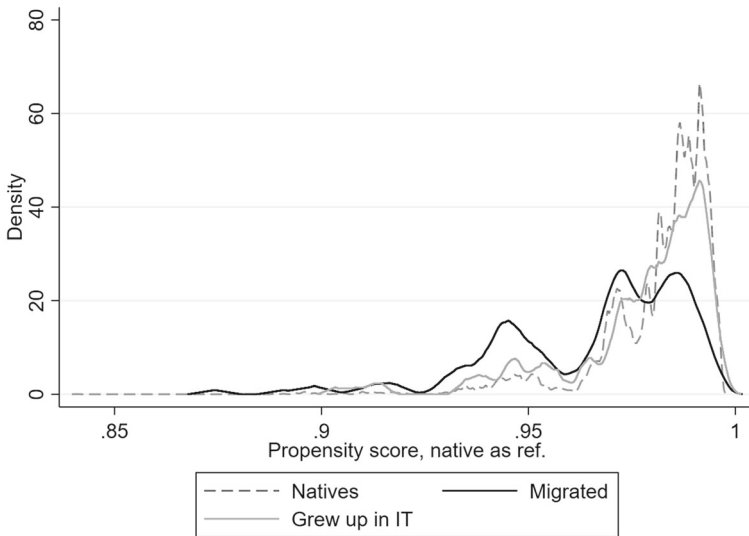
Fig. 3 Overlap of inverse-probability weighted propensity score

Table 2 Probit regressions, average marginal effects

	Model 1	Model 2	Model 3	Model 4	Model 5 Only migrants	Model 6 Only migrants	Model 7 Only 2nd gen.	Model 8 Only 2nd gen.	Model 9 Only natives	Model 10 Only natives
Foreign citizen	-0.0162 (0.012)		-0.0204 (0.016)							
Migrated		-0.0041 (0.014)		-0.0068 (0.018)						
Grew up in IT		-0.0476 *** (0.016)		-0.0670 * (0.032)						
Network-looking	-0.0041 *** (0.000)	-0.0041 *** (0.000)			-0.0021 (0.003)		-0.0001 (0.003)		-0.0042 *** (0.000)	
Age	0.0003 * (0.000)	0.0003 * (0.000)	0.0005 ** (0.000)	0.0005 ** (0.000)	0.0058 *** (0.002)	0.0053 *** (0.002)	0.0044 ** (0.002)	0.0053 * (0.003)	0.0002 (0.000)	0.0004 * (0.000)
Gender (male)	-0.0554 *** (0.004)	-0.0552 *** (0.004)	-0.0541 *** (0.004)	-0.0540 *** (0.004)	-0.0639 ** (0.027)	-0.0885 ** (0.041)	-0.0721 ** (0.036)	-0.0850 * (0.045)	-0.0549 *** (0.004)	-0.0537 *** (0.004)
Child(ren)	-0.0253 *** (0.004)	-0.0254 *** (0.004)	-0.0237 *** (0.005)	-0.0239 *** (0.005)	-0.0647 * (0.037)	-0.0946 ** (0.041)	-0.0810 * (0.047)	-0.0910 (0.058)	-0.0253 *** (0.004)	-0.0233 *** (0.004)
Area Centre[x]	0.0178 *** (0.005)	0.0178 *** (0.005)	0.0160 *** (0.005)	0.0159 *** (0.005)	0.0499 (0.045)	0.0782 * (0.045)	-0.0285 (0.032)	-0.0284 (0.051)	0.0174 *** (0.005)	0.0151 *** (0.005)
Area South[x]	0.0099 ** (0.004)	0.0099 ** (0.004)	0.0087 * (0.005)	0.0088 * (0.005)	-0.0575 * (0.030)	-0.0386 (0.052)	0.0322 (0.053)	0.0436 (0.054)	0.0107 *** (0.004)	0.0090 * (0.005)
Major cities[xx]	0.0267 *** (0.004)	0.0267 *** (0.004)	0.0296 *** (0.004)	0.0296 *** (0.004)	0.0389 *** (0.029)	0.0317 (0.039)	0.0216 (0.035)	0.0035 (0.045)	0.0264 *** (0.004)	0.0290 *** (0.004)
Father's education	0.1184 *** (0.005)	0.1185 *** (0.005)	0.1031 *** (0.004)	0.1031 *** (0.004)	0.0997 *** (0.033)	0.0968 *** (0.035)	0.2892 *** (0.000)	0.3795 *** (0.102)	0.1178 *** (0.005)	0.1027 *** (0.004)
Work status	-0.0601 *** (0.004)	-0.0601 *** (0.005)			0.0196 (0.039)		0.2309 *** (0.000)		-0.0614 *** (0.0036)	
Tenure	-0.0034 *** (0.000)	-0.0034 *** (0.000)	-0.0034 *** (0.000)	-0.0033 *** (0.000)	-0.0000 (0.003)	-0.0022 (0.003)	0.0069 ** (0.003)	0.0030 (0.003)	-0.0033 *** (0.000)	-0.0033 *** (0.000)
Public	-0.0445 *** (0.004)	-0.0447 *** (0.004)	-0.0430 *** (0.005)	-0.0430 *** (0.005)	-0.1156 ** (0.052)	-0.0915 *** (0.058)	-0.1332 *** (0.041)	-0.1443 ** (0.072)	-0.0439 *** (0.004)	-0.0432 *** (0.005)
Network-finding			-0.0649 *** (0.005)	-0.0652 *** (0.005)		-0.0739 ** (0.037)		-0.0298 (0.044)		-0.0649 *** (0.005)
Length of stay					-0.0022 (0.002)	0.0016 (0.003)				
Origin E. Europe					-0.0535 (0.061)	-0.1499 (0.093)				
Origin C. S. America					-0.0507 (0.065)	0.1688 * (0.097)				
Origin Africa					-0.0098 (0.069)	-0.0912 (0.102)				
Origin Asia Oceania					-0.0431 (0.125)	0.1687 (0.134)				
Origin other					-0.0193 (0.071)	-0.1079 (0.103)				
N	24304	24300	18408	18407	275	172	100	69	23929	18154

[x] With respect to Area North
[xx] More than 50,000 inhabitants

Looking at our estimates on subsamples, there appears to be no effect of the use of networks to look for work for foreign citizens (M5 and M7) while they seem to decrease over-education for natives (M9). Finding work through networks seemingly decreases the probability of over-education for migrants and natives (M6 and M10) but has no effect for second-generation migrants (M8). Interestingly, M5 and M6 indicate that there is no effect of the length of stay or the area of origin (in contrast with previous literature).

In order to control for the potential endogeneity bias introduced by the use of cross-sectional data in the regressions presented in the main text, we provide further results obtained through the application of the same models on PLUS panel data, which covers the years 2016 and 2018. PLUS panel data counts a total of 13,500 observations, fewer than for the cross-section data. Resultantly, the number of foreigners in the panel data is not suitable for all the analyses presented above. In addition, the way the question on citizenship is asked in PLUS 2016 differs than that in PLUS 2018, thus preventing us from splitting the foreign population between migrants and second-generation migrants. Namely, whilst PLUS 2018 allows the split between foreigners who grew up abroad and foreigners who grew up in Italy, PLUS 2016 only identifies those with and those without the Italian citizenship. This is problematic because some of those with the Italian citizenship may have acquired it over the years through naturalization. That being stated, the number of foreigners allows us to re-run M1, M3 and M9, presented in the table below as M1', M3' and M9'. We also add a model run on the foreigners in PLUS 2018 cross-sectional data (M11; not included in Table 2) and its equivalent in PLUS panel data (M11').

The models presented in Table 3 and their comparison with Table 2 suggest consistency in our results and, thus, a limited endogeneity bias. For M1, M1', M3 and M3', the coefficients for foreign citizens are not statistically significant. The use of networks in these models is captured by coefficients that are similar in the direction, the magnitude, and the statistical significance of the effect. M9 and M9' on the native population yield similar conclusions. M11 and M11' are also widely similar the results yielded for the use of networks (which is not statistically significant and, at any rate, with small magnitude), despite the decrease in number of observations. Note though that some control variables change from M11 to M11' in their statistical significance. The variables that are significant in both models are alike in terms of direction of the effect; less so in terms of magnitude of said effect.

To further test the results of the probit regressions presented thus far, we use matching techniques through propensity scores to correct the unbalances of our sample.

4.2 Matching Strategy

Firstly, we analyse the dichotomy foreigner-native through PSM. After having matched the observations as described above, the effect of being a foreigner on the probability of mismatch clearly appears. The probit regression on the matched sample yields a positive 12 p.p. difference between natives and foreigner (Table 4, first model). More specifically, whilst natives have, on average, a 12.7% probability of being over-educated; foreigners display a 24.7% probability. Matching the observations relating

Table 3 Probit regressions on panel data (M1', M3', M9', M11') and cross-section data (M11), average marginal effects

	M1'		M3'		M9' Only natives		M11 Only foreigners		M11' Only foreigners	
Foreign citizens	0.0489		0.0585							
	(0.034)		(0.040)							
Network-looking	-0.0048	***			-0.0048	***	-0.0011		0.0020	
	(0.001)				(0.001)		(0.002)		(0.007)	
Age	0.0002		0.0003		0.0001		0.0047	***	0.0025	
	(0.000)		(0.000)		(0.000)		(0.001)		(0.003)	
Gender (male)	-0.0422	***	-0.0403	***	-0.0434	***	-0.0694	***	0.0630	
	(0.007)		(0.008)		(0.007)		(0.023)		(0.090)	
Child(ren)	-0.0254	***	-0.0193	*	-0.0243	***	-0.0541	*	-0.0380	
	(0.008)		(0.010)		(0.008)		(0.031)		(0.076)	
Area Centre	0.0227	**	0.0252	**	0.0235	**	0.0321		-0.0619	
	(0.010)		(0.010)		(0.010)		(0.033)		(0.061)	
Area South	0.0125		0.0060		0.0128	*	-0.0485	*	0.1416	
	(0.008)		(0.009)		(0.008)		(0.025)		(0.124)	
Major cities	0.0231	***	0.0237	***	0.0228	***	0.0407	*	0.0918	
	(0.007)		(0.008)		(0.007)		(0.024)		(0.070)	
Father's education	0.1263	***	0.1080	***	0.1246	***	0.1191	***	0.2317	***
	(0.010)		(0.009)		(0.010)		(0.027)		(0.072)	
Work status	-0.0554	***			-0.0557	***	0.0031		0.0721	
	(0.008)				(0.008)		(0.035)		(0.109)	
Tenure	-0.0025	***	-0.0025	***	-0.0025	***	-0.0011		0.0013	
	(0.000)		(0.000)		(0.000)		(0.002)		(0.003)	
Public	-0.0456	***	-0.0415	***	-0.0437	***	-0.0954	***	-0.2699	***
	(0.008)		(0.009)		(0.008)		(0.036)		(0.066)	
Network-finding			-0.0512	***						
			(0.009)							
Year (2018)	0.0128	***	-0.0034		0.0126	***			-0.0191	
	(0.004)		(0.003)		(0.004)				(0.051)	
Area of origin – Fixed effects							Yes		Yes	
N	12483		9307		12373		275		104	

***p < 0.01, **p < 0.05, *p < 0.1. Robust standard errors reported in parenthesis

Table 4 Estimation of the average effect on the treated (ATT) of being a foreigner: logistic, nearest neighbour, Kernel and radius matching estimations

DV: over-education	Probit M1[xx]	Nearest neighbour[x]	Kernel matching[x]	Radius matching[x] (0.1)	Probit M2[xx]	Probit M3[xx]
Foreign citizen	0.120 ***	0.096 ***	0.061 **	0.117 ***	0.114 ***	0.124 ***
	(0.024)	(0.026)	(0.021)	(0.024)	(0.029)	(0.035)
Network-looking					-0.000	
					(0.003)	
Network-finding						-0.073 **
						(0.035)

[x]Bootstrap std.err. 50 repetitions
[xx]Robust std. err
***p < 0.01, **p < 0.05, *p < 0.1

to foreigners to their close native neighbours, irrespective of the matching method, reveals a statistically significant difference ranging between 6.1 and 11.7 p.p. from one group to another.

In a different fashion, considering the effect of informal networks on matched observations reveals that there is no significant (neither statistically nor substantively) effect of the intensity of informal network use on over-education even though there
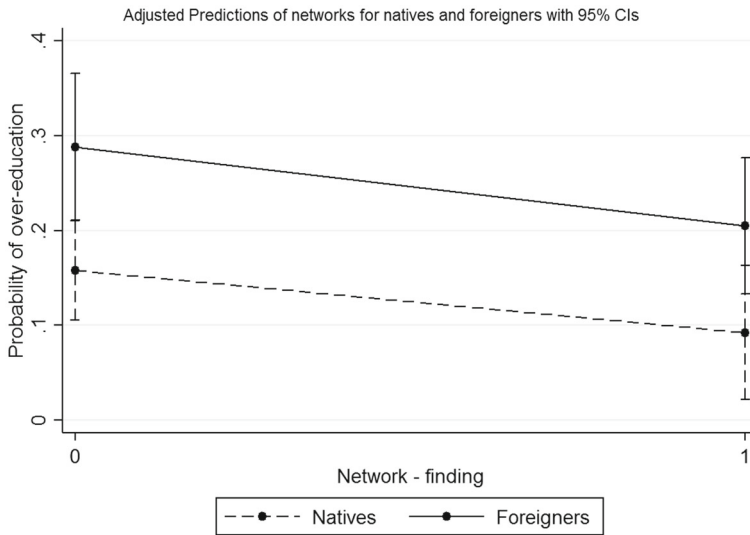
Fig. 4 Predicted effect of networks on mismatch between natives and foreigners (proportions)

does appear to be an effect of networks on mismatch when they lead up to employment. Interestingly, the effect is negative, meaning that finding employment through networks consistently decreases the probability of over-education by 7.3 p.p. Figure 4, however, shows the breadth of the effect is no different for natives and foreigners as the slopes are very similar.

Now turning to the difference between types of foreigners, breaking down the foreign category into migrants and foreigners who grew up in Italy points to a notable difference between these two groups[13] (Table 4). Namely, migrants are 8.3 p.p. more likely to be over educated than natives and 19.7 p.p. more likely to be so than foreigners who grew up in Italy. Reversely, there does not appear to be any difference between foreigners who grew up in Italy and natives (coefficient of limited magnitude and not statistically significant). This is an important finding as it suggests the existence of an upward mobility for children of foreigners, a phenomenon already established in literature with regard to old immigration countries such as Sweden, the UK, Germany or the Netherlands (Crul et al. 2017) but also with respect to newer immigration destination such as Spain (Portes et al. 2016).

The estimates provided thus far converged towards a greater probability of over-education for the foreigners who migrated than for natives or foreigners who grew up in Italy. It remains to be seen whether the use of informal networks plays a role in

---

[13] To test the robustness of our results on over-education, we run similar models on Statistics on Income and Living Conditions 2017 (Eurostat data; not reported for legibility purposes). Considering the differences in variables and sample, we obtain similar results in terms of statistical significance and direction of the effect. In terms of magnitude of the coefficients, the results are moderately different: the difference between migrants and natives is of 13.5 p.p. instead of 8.3 p.p. with PLUS; the difference between migrants and second-generation migrants is of 20 p.p. instead of 19.7 with PLUS. To control the aggregation effect of ISCO levels, we also run the models calculating over-education with ISCO 2 and 3 digits.

Table 5 Estimation of the average effect of migration and migration background: multinomial logistic estimation

Average treatment effect on the treated DV: over-education	Coefficients[x]	
Migrated vs. natives	0.0827	***
	(0.014)	
Migration background vs. natives	0.0334	
	(0.055)	
Migrated vs. migration background	0.1965	***
	(0.061)	

[x]Robust std. err
***p < 0.01, **p < 0.05, *p < 0.1

Table 6 Estimation of the effect of informal networks. Logistic regression with inverse probability weighting, average marginal effects

Average treatment effect on the treated DV: over-education	Probit M1[x]		Probit M2[x]
Migrated	0.0804	*	0.0775
	(0.047)		(0.032)
Grew up in IT	-0.0261		-0.0283
	(0.041)		(0.049)
Network-looking	-0.0036		
	(0.004)		
Network-finding			-0.0781
			(0.053)

[x]Robust std. err
***p < 0.01, **p < 0.05, *p < 0.1

this. Table 5 reports our estimates. With models breaking down foreigners between migrants and second generations, it appears that the use of informal networks to look for a job does not decrease over-education. Nor does it decrease it when people found their current job through informal networks (Table 6).

However, said effect does appear to vary from one category to another. When considering the intensity of network use to look for a job, the effect is positive for migrants while it is negative for natives and foreigners who grew up in Italy. The differentiated effect appears clearer where the use of networks is most intense. Figure 5 illustrates that.[14]

If, instead, we consider the instances in which respondents reported having found their current employment through networks, the effect of the latter does not vary from one category to another as the confidence intervals for our two categories of interest overlap significantly. Figure 6 illustrates this.[15]

---

[14] For a clearer idea of this effect, probit model 1 was run again with an interaction term. We do not report the coefficient for ease of reading. The interaction term in the probit regression is statistically significant at the 90% level but plotting the predicted probabilities shows that the effect is not different between natives and foreigners who grew up in Italy while it is with respect to the third category; all the more so as the use of networks increases.

[15] As for the previous figure, we re-run probit model 2 with an interaction effect, of which we do not report the results for ease of reading. Note that the interaction term is not statistically significant.
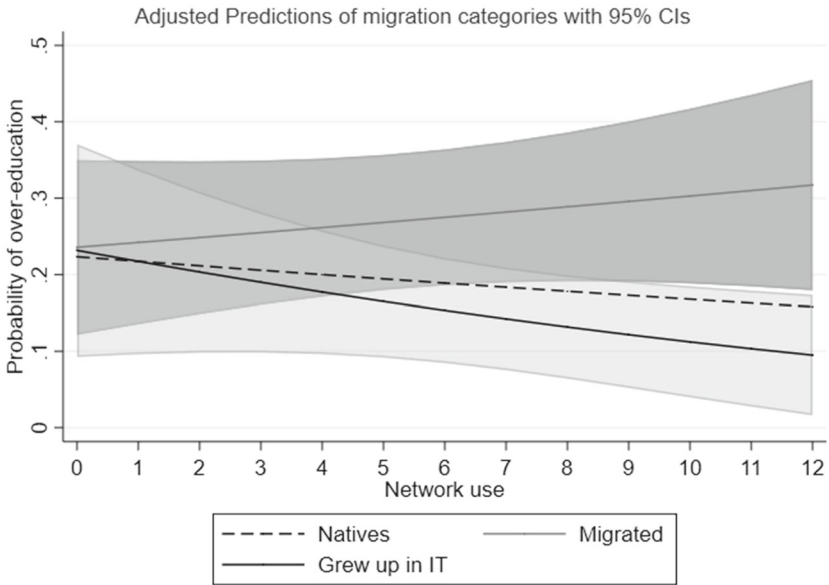
Fig. 5 Effect of migration categories at different levels of networks' use. Predicted probabilities. *NB* Confidence interval for natives is not plotted, but it is very narrow and thus very close to the dashed curve
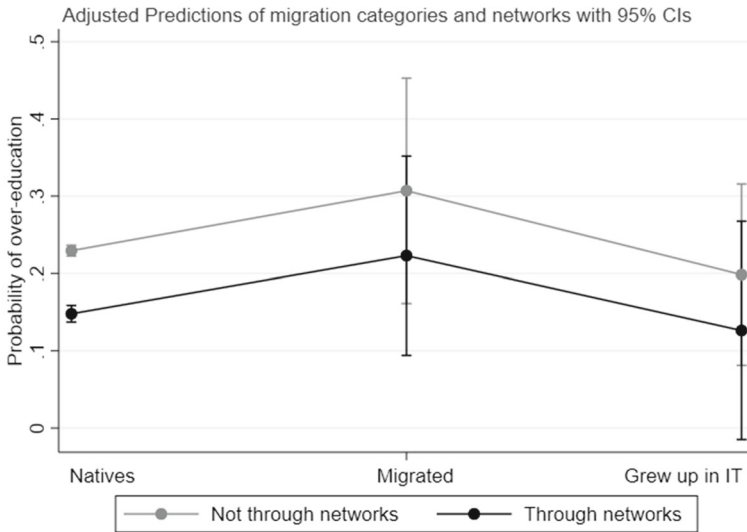
Fig. 6 Effect of migration categories and networks leading to employment. Predicted probabilities

5 Discussion: Policy Considerations to Move Past Foreigners' Mismatch

The results from our two empirical approaches (probit regressions and causal models) are coherent. The coefficients in M1–M4 (Table 2) are likely biased and inconsistent due to the unbalanced composition of our sample: they are over-determined by the large number of natives in the data (see section on data above). In contrast, in separating our three groups in sub-samples, M5–M10 provide estimates that are coherent with that of the causal models in terms of significance, direction and magnitude. However, because they are run on sub-samples, they do not allow direct comparison between groups. The causal models therefore allow us to correct the biases in our sample and to compare our three groups in a more reliable manner. The downside of it though, lies with the fact that causal models, if they ensure sounder comparison between groups by controlling for self-selection into one group or another, they seldom allow inferences on the population as a whole. Put differently, what is gained of internal validity is at a loss of external validity.

One of the most significant findings of this study lies with the difference between first- and second-generation migrants. Whilst migrants are significantly more prone to mismatch than natives, individuals with a migration background have labour market outcomes very similar to that of Italian citizens. Our analysis suggests that migrants' length of stay in the territory is not the reason why. Rather, our findings suggest that much of the phenomenon may have to do with the recognition of educational levels. Whilst some EU countries avail of consolidated education recognition systems (such as Sweden, Germany and the Netherlands), Italy features a system that is fragmented, complex, and the outcomes of which are uncertain for migrants with non-EU education (Lodigiani and Sarli 2017). Whilst important steps have been taken for the recognition of educational levels within the EU,[16] there is no unique and coherent framework to deal with third countries' education systems and, consequently, qualifications. Yet, evidence from other EU countries, notably from Germany, shows the relationship between the recognition of qualifications and migrants' labour market outcomes: three years after recognition, migrants' employment rate as well as their hourly wages significantly rose (Brücker et al. 2018). However, research on the topic is still scarce (particularly so in Italy) and would need to be expanded in order to turn the hypothesis of the effect of recognition schemes into evidence-based policy recommendation.

Regarding the effect of networks, it is interesting to see that one specification is statistically significant whilst the other is not (respectively use of networks and job found through networks). Even though some level of significance is found for the former, Fig. 2 displays sizable confidence intervals for migrants and people with a migration background so that the size of the effect itself is somewhat negligible. The results however suggests a different dynamic between groups, notably between migrants on the one hand, and natives and second-generation migrants on the other. A likely explanation lies with the composition of networks themselves. Whilst they may be constituted of co-ethnics at the bottom of the skills ladder for the former (as research on labour market segmentation demonstrates), they are likely mixed for

---

[16] Notably with Directive 2005/36/EC and 2013/55/EC.

second-generations and native for the Italian nationals (Battisti et al. 2021). Further research should be conducted to safely conclude to a differentiated effect of networks.

6 Conclusion

Education-occupation mismatch is an issue that undermines the efficient allocation of human resources in the labour market. Whilst it regards the population as a whole, evidence in a range of OECD countries tends to suggest that it affects migrants more than natives. Different explanations were put to the test over the years, with a growing attempt to capture the effect of informal networks on over-education. Presumably, resorting to those networks would undermine the ability of workers to match their qualification with their job. Interestingly, no study has aimed at studying the effect of networks on migrants' over-education in Italy. Yet, Italy is a country for which there is evidence on the relationship between informal networks and education-occupation mismatch.

This article is an attempt to fill this gap. It contributes to the specialised literature in three ways. First, we look into the Italian case, a country that has seen its foreign population growing at a fast pace over the recent years, but with a clear worsening of immigrants' relative position after the 2008 crisis (Ricci and Scicchitano 2021) and a recent step back in their integration process in the Italian society (Ceccarelli et al. 2014). Second, we use a unique database that allows us to rely on a clear, self-declared specification of the use of networks instead of building a proxy. Third, we propose to differentiate the effect of networks between foreigners who grew up in Italy and foreigners who migrated later on in life and who allegedly face more hardship in integrating into the Italian labour market. We rely on two definitions of second-generation migrants for more robust results. Through the use of causal models, we find evidence of a difference in over-education between foreigners and natives; a difference that is clearer when we distinguish between the foreigners who grew up in Italy and those who migrated later in life. The former are undistinguishable from natives whilst the latter tend to be significantly more over-educated. These results are in line with recent findings in other EU countries; namely Germany, the Netherlands, the United Kingdom and Sweden (Crul et al. 2017). It is also in line with similar findings concerning the integration of second generations in countries that have recently become immigration destinations (Portes et al. 2016). Whilst the picture proposed is encouraging in terms of integration for second-generation migrants, it remains worrisome if we consider the higher probability of over-education for migrants as this is symptomatic of a misallocation of human resources in the labour market. In addition, we provide (weak) evidence that resorting to informal networks to look for a job increases migrants' over-education whilst it decreases that of natives and second-generation migrants. More research in this regard needs to be conducted in order to investigate the types of networks migrants are involved in.

The results of our analysis suggest that part of the mismatch affecting migrants may stem from the lack of recognition of their educational attainment. If there is, in Italy, a framework intended to favour the recognition of foreign education, it is deemed cumbersome and inefficient. An improved recognition scheme could prove

effective in alleviating migrants' education-occupation mismatch, and thus correct some of the misallocation of human resources. That being stated, the data available on the topic is still scarce in Italy, which renders the establishment of solid causal links difficult. If our study leads us to this hypothesis, more research is needed in order to turn the hypothesis of the effect of recognition schemes into evidence-based policy recommendation.

References

Alaverdyan S, Zaharieva A (2019) Immigration, Social Networks and Occupational Mismatch. Bielefeld University, Center for Mathematical Economics Working Papers 619

Alba R, Nee V (1997) Rethinking assimilation theory for a new era of immigration. Int Migr Rev 31:826. https://doi.org/10.2307/2547416

Aleksynska M, Tritah A (2013) Occupation–education mismatch of immigrant workers in Europe: context and policies. Econ Educ Rev 36:229–244. https://doi.org/10.1016/j.econedurev.2013.06.001

Ambrosini M (2013) Immigration in Italy: between economic acceptance and political rejection. J Int Migr Integr 14:175–194. https://doi.org/10.1007/s12134-011-0231-3

Battisti M, Peri G, Romiti A (2021) Dynamic effects of co-ethnic networks on immigrants' economic success. Econ J. https://doi.org/10.1093/ej/ueab036

Blangiardo GC (2019) Statistical aspects. In: Cesareo V (ed) The Twenty-fourth Italian Report on Migrations 2018. Fondazione ISMU, Milan, pp 21–28

Borjas GJ (1988) Immigration And Self-Selection. National Bureau of Economic Research Working Papers 2566. https://doi.org/10.3386/w2566

Bourdieu P (1986) The forms of capital. In: Richardson J (ed) Handbook of theory and research for the sociology of education. Greenwood Press, Westport, pp 241–258

Brücker H, Glitz A, Lerche A, Romiti A (2018) Occupational Recognition and Immigrant Labor Market Outcomes. Institute for Employment Research

Caroleo FE, Pastore F (2018) Overeducation at a glance. determinants and wage effects of the educational mismatch based on AlmaLaurea data. Soc Indic Res 137:999–1032. https://doi.org/10.1007/s11205-017-1641-1

Cattaneo M (2010) Efficient semiparametric estimation of multi-valued treatment effects under ignorability. J Econom 155:138–154. https://doi.org/10.1016/j.jeconom.2009.09.023

Ceccarelli C, Giorgi GM, Guandalini A (2014) Is Italy a melting pot? Rivista Italiana Di Economia, Demografia e Statistica 68:23–30

Penninx R, Garcés-Mascareñas B (2016) The concept of integration as an analytical tool and as a policy concept. In: Garcés-Mascareñas B, Penninx R (eds) Integration processes and policies in Europe. Springer International Publishing, Cham, pp 11–29

Piracha M, Vadean F (2013) Migrant educational mismatch and the labor market. In: Constant A, Zimmermann KF (eds) International handbook on the economics of migration. Edward Elgar, Cheltenham, pp 176–192

Piracha M, Tani M, Vadean F (2012) Immigrant over- and under-education: the role of home country labour market experience. IZA J Migr 1:3. https://doi.org/10.1186/2193-9039-1-3

Pistaferri L (1999) Informal networks in the Italian labor market. Giornale Degli Economisti e Annali Di Economia 58:355–375

Portes A, Aparicio R, Haller WJ (2016) Spanish legacies: the coming of age of the second generation. University of California Press, Oakland

Reyneri E, Fullin G (2010) Labour market penalties of new immigrants in new and old receiving West European Countries. Int Migr 49:31–57. https://doi.org/10.1111/j.1468-2435.2009.00593.x

Ricci CA, Scicchitano S (2021) Decomposing changes in income polarization by population group: what happened during the crisis? Econ Polit J Anal Inst Econ 38:235–259

Robst J (2007) Education, College Major, and Job Match: Gender Differences in Reasons for Mismatch. Educ Econ 15:159–175. https://doi.org/10.1080/09645290701263070

Rosenbaum PR, Rubin D (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70:41–55. https://doi.org/10.1093/biomet/70.1.41

Schneider J (2016) First/Second Generation Immigrants. NESET II ad hoc question 4/2016. Network of Experts working on the Social dimension of Education and Training. https://nesetweb.eu/

Verdugo RR, Verdugo NT (1989) The impact of surplus schooling on earnings: some additional findings. J Hum Resour 24:629. https://doi.org/10.2307/145998

Wrench J, Rea A, Ouali N (2016) Migrants, ethnic minorities and the labour market: integration and exclusion in Europe. Palgrave Macmillan Limited, London