LUISS Guido Carli
School of European Political Economy

Does Low Skilled Immigration Cause Human Capital Polarization?

Giorgio Brunello
Elisabetta Lodigiani
Lorenzo Rocco

Working Paper
9/2017

LUISS
Academy

LUISS Guido Carli / School of European Political Economy
Working paper n. 9/2017
Publication date: September 2017
Does Low Skilled Immigration Cause Human Capital Polarization?

ISBN 978-88-6856-104-8

LUISS Academy is an imprint of
LUISS University Press – Pola Srl
Viale Pola 12, 00198 Roma
Tel. 06 85225485
e-mail universitypress@luiss.it
www.luissuniversitypress.it

Does Low Skilled Immigration Cause Human Capital Polarization?

Evidence from Italian Provinces

Giorgio Brunello (Padova, IZA, CESifo and ROA)

Elisabetta Lodigiani (Padova, LdA)

Lorenzo Rocco (Padova)

Abstract

While there is a vast literature considering the labor market effects of immigration, less has been done to investigate how immigration affects the educational choices of young natives**.** Using Italian provincial data and an instrumental variables strategy, we show that the recent increase in the immigration of low skilled labor has produced human capital polarization, i.e. an increase in the share of natives with less than high school and with college. This evidence is stronger for males than for females. We adapt the standard Card's model of educational choice and point out under what conditions human capital polarization occurs. We estimate wage equations by gender and find that these conditions are satisfied for Italian males.

Acknowledgements: we are grateful to Gaetano Basso and Frederic Docquier and to the audiences at seminars in Cosenza (AIEL), Maastricht (ROA) and Rome (La Sapienza) for comments and suggestions. Giorgio Brunello wishes to thank the School of European Political Economy at LUISS University, where some of this research was carried out, for its hospitality. The usual disclaimer applies.

1. Introduction.

Immigration is a very important economic and social phenomenon. According to the OECD, in 2015 about 244 million people were living outside their country of origin, more than half of them in G20 countries. In recent years, many of these countries have seen a significant rise in migration trends: between 2010 and 2015, net migration added about 10 million people to total population (OECD, 2017). The effects of immigration for the rich receiving countries have been widely investigated.

However, while there is a vast literature considering the labor market effects of immigration (e.g. Ottaviano and Peri, 2012, for the US; D'Amuri et al., 2010, for Germany; Manacorda et al., 2012, for the UK), less has been done to investigate how immigration affects the educational choices of young natives.

This paper looks at the impact of low skilled immigration on the educational choices of young natives in Italy during the period 2006-2016. We think that the case of Italy is interesting. First, the majority of immigrants originate from developing countries and are low skilled.[1] Second, immigration is a recent and rapidly growing phenomenon. During the period of interest, the average share of immigrants almost doubled, from 3.7 to 7 percent of the population. Yet, the literature on the relative impact of immigration in Italy is relatively scarce.

In one of the first studies in this area, Betts, 1998, provides a useful framework for thinking about the effects of immigration on the educational choices of natives. On the one hand, a large inflow of immigrants reduces the schooling resources available to natives. School quality declines and natives invest less. On the other hand, low skilled immigration increases the supply of low skilled labor. When low and high skills are complements in production, a higher

[1] The share of immigrants from less developed countries was 85.2 percent of the total in 2016. In that year, 53 percent of immigrants from less developed countries aged 25 to 59 had less than ISCED 3 education, and 55 percent worked in manual jobs (source: Italian Labor Force Survey). See also Sweetman and Van Ours, 2014. According to the International Organization for Migration, The share of immigrants employed in high skilled occupations is less than 10 percent in Italy, compared to more than 40 percent in the UK.

supply of low skills reduces the wages and employment probabilities of low skilled natives and increase the productivity and wages of high skilled workers. Returns to education raise and natives have stronger incentives to acquire additional schooling. Betts, 1998, presents evidence that the former effect prevails. He finds evidence of a negative link between immigration in the US and the probability that both blacks and Hispanics complete high school.

Later literature does not confirm Betts' results and generally finds a positive effect of immigration on the educational attainment of natives (e.g., Hunt, 2016), who typically select into jobs that involve communication - intensive tasks, where they have a comparative advantage (McHenry, 2015).

While recent research points to the positive effects of immigration on the human capital investment of natives, less is known about the heterogeneity of these effects. An important exception is Llull, 2014, who estimates a labor market equilibrium model using US data.[2] Her simulations suggest that, following immigration flows, some individuals are more likely to pursue a white - collar career and extend their stay in school. Others, however, become more detached from the labor market and, given the lower return to their investment, drop out from school earlier, reducing their education.

This pattern suggests that human capital polarization – or the contemporaneous increase in the share of natives with low and high education – is a possible outcome of immigration. In this paper, we document that polarization is not only a theoretical possibility but also an empirical fact. Our evidence for Italy – stronger for males than for females - indicates that the recent inflow of low skilled immigrants has increased the share of both low educated (with at most compulsory education) and higher educated (with college or enrolled in college) natives.

[2] See also Eberhard, 2012.

We also find that native males who choose not to invest in further education because of immigration are more likely to work in manual jobs and in the service sector. Native females are instead more likely to be inactive. In line with previous literature, we show that natives who invest more in education because of immigration are less likely to choose STEM fields and more likely to enroll in "communication-intensive" fields.

Our empirical analysis combines data on natives drawn from the Italian Labor Survey and data on legal immigrants drawn from other official sources. We recognize that the empirical relationship between the human capital choices of natives and immigration is influenced by the fact that immigrants are not randomly allocated across areas. We address the endogeneity of immigration by using the shift-share instrument proposed by Card (2001a), which exploits the tendency of immigrants to locate in areas where previous immigrants of the same ethnicity settled.

Our identifying assumption is that local economic shocks that attracted immigrants in the past (in our case, 1991) are uncorrelated with the educational choices of natives occurring more than 15 years later, conditional on a battery of controls. This assumption is likely to hold in our context, because 1991 was the year before the signing of the Maastricht Treaty and the establishment of a single European market. These events changed the economic and institutional landscape both in Italy and in Europe. The year 1991 was also a watershed in Italian immigration patterns, with the first large inflows of Albanians. In support of our identification strategy, we investigate the validity of the exclusion restriction by implementing the tests designed by Oster, 2016, and Conley, 2012.

We show that our empirical results can be explained using the standard Card's model of educational choice. We spell out the conditions for human capital polarization to hold, estimate wage equations by gender and find that these conditions are satisfied, especially for Italian males.

Polarization is the outcome of higher migration when the marginal benefits of education increase for higher education (college) but falls for intermediate education, including three and five years high school programs. When natives differ in their costs of education, these changes in marginal benefits increase the share of natives investing in college and reduce the share investing in high school. The remainder of the paper is organized as follows. Appendix 1 briefly reviews the literature. Section 2 describes the data. Section 3 presents the empirical model and the identification strategy; Section 4 illustrates the main results. An extension of the classical Card's model of rational educational choice that includes immigration as a driving variable is presented in Section 5. We test the conditions for human capital polarization in Section 6. Conclusions follow.

2. Data

We draw our data from two main sources, the Italian Labor Force Survey (LFS) for educational attainment and labor market variables and the Demographic Balance and Resident Population by Sex and Citizenship (*Bilancio demografico e popolazione residente per sesso e nazionalità*) for the stock of immigrants. Both datasets are managed by the Italian National Institute of Statistics (ISTAT). The LFS is a quarterly survey on labor market conditions covering a representative sample of almost 77,000 households and 175,000 individuals per quarter. We use the waves 2006 to 2016 for two reasons. First, the survey was substantially revised in 2004 and includes public information on the province of residence only from that date. Second, the identification of natives (defined in this paper as individuals born in Italy and with Italian citizenship) is only publicly available from 2006.

Since the number of provinces has changed during the selected sample period, from 103 in 2006 to 110 in 2012, we have reclassified local areas into 110 provinces for the entire sample period. We focus on the natives aged 19 to 27. In this age range, many Italians are in full time education. Typically, compulsory education ends at age 16, 3-year or 5-year high school at age 17 to 19, and college education – inclusive of Master level - at age 24. In practice, the actual age of completion can be higher than 24.[3]

We allocate natives into four groups: 1) those with less than high school, who are not enrolled in school or in other training activities (LN); 2) those with less than high school, who are enrolled during the same period (LY); 3) those with completed high school, who are not enrolled in education or training (HN); 4) those with completed high school and are enrolled in education, and those with a college degree (HY).

Table 1 shows how these four groups have changed over the sample period, separately for males and females. On average, and consistent with the broad increasing trend in education attainment, the group LN has declined from about 26 percent in 2006 to 18 percent in 2016 for males, and from 18 to 12 percent for females. On the other hand, the group HY has increased during the same period from 30 to 32 percent for males and from 44 to 47 percent for females.

Since 2002, the yearly Demographic Balance and Resident Population by Sex and Citizenship contains information on the stock of regular resident foreigners (by citizenship and sex) in each Italian municipality on January 1 of each year.[4] Unfortunately, these data do not include the education of immigrants, although it is plausible to assume that immigrants originating from developing countries are mainly unskilled. We define the share of unskilled immigrants as the ratio of the stock of immigrants from developing countries[5] to total resident population (including both natives and immigrants). During the period 2006-16, this ratio has almost doubled, from 3.7 to 7

[3] In 2001, the share of individuals graduating in the age ranges 22-24 and 25-27 was 38.6 and 34.2 respectively. See ANVUR, 2014.
[4] We prefer these data to those provided by the LFS, because LFS sample weights can be used to obtain provincial aggregates by sex and age groups, but not by immigrant status. For further details, see ISTAT (2006a, b).
[5] We classify developing countries following the 2016 World Bank income classification.

percent (see Table 2). We aggregate data by province and year. Since the variation in the share of immigrants is by province and year, we aggregate LFS data in a similar fashion, using sampling weights to reproduce provincial aggregates.

While the share of immigrants during the period under study increased in all provinces, the observed change was larger in the industrialized central and northern areas of the country. In our data, the observed increase ranges from less than one percentage point in some provinces of Sardinia (Ogliastra, Carbonia-Iglesias, Medio Campidano) to more than six percentage points in the provinces of Rome, Piacenza and Prato. Figure 1 illustrates these changes across provinces, with the dark blue areas indicating provinces with relatively large increases and with pale blue areas indicating provinces with small increases.

We also use data on residence permits in 1991 from the Italian Ministry of Interior, by province and country of origin, to construct the instrumental variable described in the next Section.[6]

3. The Empirical setup

3.1 Empirical Specification

We focus on the impact of the share of low skilled immigrants M in province p and time t on the educational choices of young natives aged 19-27 in the same province and time. We estimate the following empirical model separately by gender

$$Y_{pt} = \alpha_0 + \alpha_1 M_{pt} + \alpha_2 X_{pt-1} + \alpha_3 D_{pt} + \mu_p + \omega_t + \varepsilon_{pt} \qquad (1)$$

[6] Since in 1991 there were 95 provinces, we redistribute residence permits over the current 110 provinces using the 2011 Population Census. To illustrate with an example, a few municipalities separated from the Province of Milan in order to constitute the new Provinces of Lodi and Monza-Brianza. According to the 2011 Census, around 20 and 5 percent of the population of the province of Milan constituted the new provinces of Monza - Brianza and Lodi. We reallocate resident permits as follows: we take the 21 Afghan permits in the province of Milan and assign 4 to the Province of Monza-Brianza, 1 to the Province of Lodi and 16 to the province of Milan.

where the index p refers to province and t to time, Y is the share of natives in one of the four education groups defined above (LN, LY, HN, HY), μ_p is a vector of provincial dummies, ω_t a vector of time dummies and ε_{pt} a residual error term.

In specification (1), we regress Y by gender on the total share of immigrants, independently of gender. However, one could argue that Y for males depends on the share of male immigrants, and the same for females. To investigate whether this is the case, we have replaced M in (1) with the geometric average $M^{\alpha}M_m^{1-\alpha}$ for males and with $M^{\alpha}M_f^{1-\alpha}$ for females, where M_m and M_f are the share of male and female immigrants, and allowed α to vary between 0 (only the gender specific share matters) and 1 (only the total share matters). We find that the estimated elasticities of Y with respect to M vary only marginally with α, and that the mean root squared errors remain virtually unchanged. We conclude from this exercise that using the total share of immigrants rather than the gender specific shares does not alter our results.

The vector X_{pt-1} includes the following provincial confounders, which we lag once to alleviate endogeneity concerns: the log of provincial real GDP; the provincial unemployment rate for individuals aged 18-59, and the local share of workers employed in sectors with a higher than average proportion of manual workers in the pre-sample period (year 2004). This share is included to control for changes in the production structure that could affect the educational choices of natives independently of migration, for instance because of globalization or technical change.[7]

We allow for the effect of aggregate shocks to vary across provinces by adding in vector D the interactions of province dummies with a dummy for the recession years (2008, 2009, 2012 and 2013). We also control for average age and its square. In our estimates, we always cluster standard errors at the

[7] The aggregate effects of technical change and globalization are captures by the time dummies.

provincial level. We also weight regressions using the number of individuals in each cell. Table 3 shows the descriptive statistics of the variables used in the regressions.

3.2 Identification Strategy

We recognize that OLS estimates of Eq. (1) are biased by the endogeneity of the share of immigrants M. Clearly, immigrants are not randomly distributed across provinces, but self-select on local labor market characteristics, including the local industrial structure and the availability of low skill jobs. In addition, immigrants choose to locate in areas with lower house prices, where disadvantaged natives also live. This induces a positive correlation between the share of immigrants in an area and the share of low educated natives. Reverse causality is also an issue if immigrants with children tend to locate in provinces with high educational attainment (and potentially high income).

We address endogeneity by estimating equation (1) using instrumental variables. In selecting the instrument, we follow the approach proposed by Card (2001a) and followed by most of the relevant literature. Define Z in province p and at time t as

$$Z_{pt} = \frac{\sum_{c=1}^{N} \lambda_{pc91} IMM_{ct}}{Pop_{p91}}$$

where λ_{pc91} is the share of immigrants from developing country c in province p and year 1991, well before our sample starts (in 2006); IMM_{ct} is the national number of immigrants from developing country c in year t; and Pop_{p91} is the total resident population in province p and year 1991.[8]

The instrument Z exploits the fact that immigrants tend to locate in areas with a large share of immigrants of the same country of origin (enclave effect). The exclusion restriction requires that, conditional on province and time dummies and on the set of province by time controls, provincial shocks that attracted

[8] We use the population in 1991 rather than current population because the latter is potentially endogenous.

immigrants at least fifteen years before the start of the sample period are uncorrelated with current provincial characteristics and shocks, which influence the educational choices of natives.

We believe that this requirement is likely to hold in our data for at least two reasons. First, we control for local GDP, local unemployment and changes in the sectorial composition of employment. As discussed by Barone and Mocetti, 2011, these variables control for local employment opportunities and productivity.

Second, in the construction of Z, we predict the share of low skilled immigrants over the total resident population in a given province p and time t by redistributing immigrants from developing countries across provinces as of their distribution in 1991. Importantly, 1991 is the year before the signing of the Maastricht Treaty and the expected completion of the single market. These events changed the economic and institutional landscape in Italy and Europe. The year 1991 also predates by far the Eastern enlargements in 2004 and 2007, which brought some former communist countries, including Poland, Bulgaria and Romania, in the EU. These important historical changes indicate that past and current local shocks are unlikely to be correlated, which supports the validity of the exclusion restriction. We further investigate validity in the following Section.

Table 4 reports the first stage estimates, showing that the instrument has a statistically significant and positive effect on the share of immigrants.[9] Since the associated F-statistic is well above the critical values reported by Stock and Yogo, 2005, we reject the hypothesis that our instrument is weak.

4. Estimation results

4.1 Baseline results

Table 5 shows the IV estimates of equation (1), by gender. For males (Panel

[9] First stage estimates slightly differ by gender because of the inclusion of (cell specific) average age and its square and because cell sizes vary by gender.

A), we find that immigration induces human capital polarization, as the shares of both lower educated and higher educated natives increase. Consequently, the share of those who have completed high school but are not enrolled in education or training decreases. The size of the estimated effects is not small. We find that a one standard deviation increase in the share of immigrants (equal to 3.3 percentage points) increases the shares of lower educated and higher educated males by 4.1 and 6.5 percentage points respectively.

The results for females are less clear-cut. Table 5, panel B, shows that a one standard deviation increase in M raises the share of low educated females by 5.7 percentage points but have imprecise – and negative – effects on the share of high educated females. This negative effect results from two heterogeneous effects. On the one hand, higher migration reduces the share of those with high school who are enrolled in college (estimated effect: -0.537 with standard error 0.689). On the other hand, it increases the share of those who have completed college (estimated effect: 0.291, standard error: 0.359). As for males, the share of female with intermediate education but currently not in education or training declines following higher immigration.

4.2 Tests of instrument validity

One might argue that our results are driven by the fact that the instrument is invalid either because it correlates with local time-varying shocks which drive the decision to migrate to Italy from specific countries, or because past shocks which determined the distribution of immigrants across provinces in 1991 correlate with more recent local shocks, even after conditioning on observable province by time controls. Although we cannot dismiss a priori these possibilities, we support the exclusion restriction by means of two tests.

Our identification strategy could fail because the instrument Z is correlated with un-observables. We investigate this possibility by applying the method recently developed by Oster (2016) both to the reduced form and to the first stage estimates. The test establishes bounds to the true value of the first stage and reduced form parameter under two polar cases.

In the first case, there are no un-observables and the first stage and reduced form models corresponding to Equation (1) are correctly specified. The correct R squared is the one we estimate $(\widehat{R})$. In the latter case, there are un-observables but observables and un-observables are equally related to the treatment (δ=1). When un-observables are included, we conservatively assume that the R squared is equal to $R_{max} = \min(1.3\hat{R}, 1)$. If zero can be excluded from the bounding set, then accounting for un-observables would not change the direction of our estimates (showing evidence of polarization, at least for males).

In the case of the first stage, zero can be excluded from the bounding set, but the inclusion of un-observables would decrease the first stage estimate, and therefore increase the IV estimate in absolute value.[10] Table 6 presents the results of the test for reduced form. The table report the estimates of the reduced form coefficient associated to Equation (1), or $\hat{\beta}$, the associated standard error and R squared, the consistent parameter β^* and the bounding set for two educational options, quitting at the end of compulsory school, LN, and continuing after high school, HY.[11] Since in no case does the bounding set includes zero, the tests support the direction of the effects presented in Table 5.[12] A comparison of $\hat{\beta}$ and β^* indicates that, by omitting un-observables, we are likely to under-estimate the reduced form effects for males and females with low education (LN), and to over-estimate these effects for males with higher education (HY).

[10] Results are available upon request.

[11] Notice that $\beta^* = \hat{\beta} - [\beta^0 - \hat{\beta}] \frac{R\,max - \hat{R}}{\hat{R} - R^0}$, where β^0 and R^0 are, respectively, the reduced form coefficient and the R squared from the regression of Y on the instrument and time and province fixed effects.

[12] Not reported in the table, the same test applied to the first stage coefficients yields similar qualitative results.

We further assess the robustness of our results by adopting the approach proposed by Conley et al. (2012).[13] Suppose that the exclusion restriction does not hold precisely, so that Equation (1) can be written as

$$Y_{pt} = \alpha_0 + \alpha_1 M_{pt} + \alpha_2 X_{pt-1} + \alpha_3 D_{pt} + \gamma Z_{pt} + \mu_p + \omega_t + \varepsilon_{pt} \qquad (2)$$

where γ is small. Conley proposes to evaluate the key parameter α_1 when γ deviates marginally from zero, using a set of plausible values for γ. To implement this methodology we assume that the distribution of γ is normal and centered around a prior derived as follow. We consider individuals who completed their education before 1991, when there was virtually no immigration in Italy. For this sub-sample, M is equal to 0. Therefore, the eventual direct effect of Z on Y does not pass through changes in M and can be identified. This estimate is our prior for γ, that we use to set the mean of the distribution. We set the standard error at one third of the mean so that the sign of γ corresponds to that of the prior with probability equal to 0.99.

We implement the "local-to-zero approximation" approach and construct for α_1 95 percent confidence intervals. Table 7 presents the results for two groups, LN and HY, by gender. The table shows the estimated value of α_1 when γ=0, our estimated prior for γ and the upper and lower bounds of the 95 percent confidence interval obtained by 5000 independent draws of γ from its distribution. In no case do we find that the confidence interval includes zero. Therefore, the direction of our estimated results is robust to small deviations from instrument validity.

4.3 Educational choice, labor market status and field of study

In this sub-section, we look at the effects of immigration on the labor market status of those who end up with low education and on the selected field of study for those who continue into higher education. Consider first the group of low educated who are not in current education or training (LN). We can

[13] See Mitaritonna, Orefice and Peri, 2017, for a similar strategy.

classify this group into three sub-groups: the employed, the self-employed and the inactive. We can further classify the employed and self-employed by sector of activity.

Table 8 shows our estimates of the effects of immigration on the labor market status of low educated natives. For males, the share of low educated in employment increases significantly. For females, the share of low educated and inactive increases. For males, low-skilled immigration significantly increases the share of low educated who are employed in the service sector, building and other services. For females, there is no significant effect on employment or its composition.

Turning to the group with at least high school and currently in education or with a college degree, we find that, for males, immigration increases the share enrolled in or completing social sciences and the humanities and reduces the share enrolled in the hard sciences, in line with the finding that natives tend to specialize in "communication intensive" fields. For females, there is a negative effect on enrolling or completing medical sciences and no statistically significant effect on other fields (see Table 9)

We also compute the marginal effects of immigration on labor market outcomes by conditioning on those with low education and not currently in education or training and at the marginal effects on the choice of field of study by conditioning on those going into higher education. The results are shown in Appendix 3. As shown in Tables A1 and A2 in the Appendix, there is evidence that a higher share of immigrants increases the share of manual workers among low educated males and the share of inactive individuals among low educated females. There is also evidence that the share of those enrolled in or completing studies in the social sciences increases among higher educated individuals.

5. An Illustrative Model

We have found that an increase in the share of low skilled immigrants generates human capital polarization for males and increase the share of low educated native females, with imprecise effects on the share of higher educated females. Are these results consistent with the classical model of rational choice of education? In this section, we present an extended version of Card's model and show that polarization is one of the possible outcomes. We also spell out the necessary conditions for polarization to occur. In the next section, we verify in the data whether these conditions hold by estimating wage regressions by gender.

We adapt the by now standard Card model (Card, 2001b) by augmenting it with the share of immigrants in the relevant labor market. Let S be years of education and assume that $S = 0$ corresponds to compulsory education, attained by all (full compliance). Individuals investing in education compare expected benefits and costs, and choose a level of education higher than compulsory (S>0) if expected utility U(S) is higher than $\overline{U}$, the utility associated to compulsory education. Let the (log) benefits accruing to individual *i* be given by

$$\log w_i = \alpha_0 + \alpha_1 S_i + \alpha_2 S_i M + \alpha_3 S_i^2 + \alpha_4 S_i^2 M + \alpha_5 M \tag{3}$$

where w is for the wage and we assume that the marginal effects of S and M are constant across individuals. The presence in (3) of quadratic terms in S and of interactions between M and S allows for the possibility that the marginal returns to schooling vary both with S and with M.

We assume that the costs of education are convex in the years of schooling, a standard assumption in this literature

$$C_i = K_i + \frac{1}{2}\theta_i S_i^2 \tag{4}$$

where K is the cost of attaining compulsory education and θ varies across individuals, possibly reflecting differences in ability. Let $f(\theta)$ be the distribution of θ on the support $[\theta_{MIN}, \theta_{MAX}]$. Following empirical evidence suggesting that a higher share of immigrants in class has small effects on the school performance of European students (see Brunello and De Paola, 2017, for a review of this literature), we assume that changes in M have negligible effects on the costs of education.

Let individual utility be $U_i = \ln w_i - C_i$. Then optimal schooling above compulsory education is obtained by maximizing utility with respect to schooling S. Assuming that utility is concave in S ($\theta_i - 2\alpha_3 - 2\alpha_4 M > 0$) the optimal level of education $S_i^* > 0$ is

$$S_i^* = \frac{\alpha_1 + \alpha_2 M}{\theta_i - 2\alpha_3 - 2\alpha_4 M} \tag{5}$$

The higher the value of θ, the lower optimal schooling. We are interested in the effects of changes in the immigration rate M on individual schooling. Differentiating (5) with respect to S^* and M we obtain

$$\frac{\partial S_i^*}{\partial M} = \frac{\alpha_2 + 2\alpha_4 S^*}{\theta_i - 2\alpha_3 - 2\alpha_4 M} \tag{6}$$

Since the denominator is positive because of the concavity of U, the sign of (6) depends on the sign of the numerator, or on the sign of the marginal effect of a higher M on the returns to education.

When α_2 and α_4 are equal to zero, immigration has no effect on education. When α_2 and α_4 are both different from zero, we distinguish four cases:

1) $\alpha_4 > 0$ and $\alpha_2 > 0$: $\frac{\partial S^*}{\partial M}$ is always positive and a higher immigration rate increases education;

2) $\alpha_4 > 0$ and $\alpha_2 < 0$: $\frac{\partial S^*}{\partial M} < 0$ for low values of S* ($S^* < -\frac{\alpha_2}{2\alpha_4}$) and $\frac{\partial S^*}{\partial M} > 0$ for high values of S* ($S^* > -\frac{\alpha_2}{2\alpha_4}$). This configuration of parameters can lead to human capital polarization, i.e. an increase in the share of individuals with

lower and higher education;

3) $\alpha_4 < 0$ and $\alpha_2 > 0$: $\frac{\partial S^*}{\partial M} > 0$ if $S^* < \frac{\alpha_2}{2|\alpha_4|}$ and negative if $S^* > \frac{\alpha_2}{2|\alpha_4|}$. In this case, $\frac{\partial S^*}{\partial M}$ can switch from positive when S^* is relatively low to negative when S^* is relatively high. When this happens, a higher M increases education for those with low education and reduces it for those with higher education, the opposite of polarization.

4) $\alpha_4 < 0$ and $\alpha_2 < 0$: $\frac{\partial S^*}{\partial M}$ is always negative and a higher immigration rate reduces education.

Consider now the effect of M on the share of individuals who either attend college (π_C) or leave school at compulsory education S=0. Let $\tilde{S}$ be the level of education corresponding to an upper secondary degree, so that individuals with $S^* > \tilde{S}$ attend (some) college, and define $\tilde{\theta}$ as the value of θ such that $S^*(\tilde{\theta}, M) - \tilde{S} \equiv 0$. Then $\pi_C = \int_{\theta_{MIN}}^{\tilde{\theta}} f(\theta)d\theta$. By total differentiation, we have that

$$\frac{\partial \tilde{\theta}}{\partial M} = \frac{\alpha_2 + 2\alpha_4 \tilde{S}}{\tilde{S}}.$$

When the conditions compatible with polarization hold ($\alpha_4 > 0$ and $\alpha_2 < 0$), $\frac{\partial \tilde{\theta}}{\partial M}$ is positive and the share of the population attending (some) college increases if $\tilde{S} > -\frac{\alpha_2}{2\alpha_4}$.[14]

[14] Notice that π_C increases only if there are individuals with θ' such that $\theta' > \tilde{\theta}$ who react to an increase in *M* by raising their $S^*(\theta', M)$. In other words, π_C increases only if there are individuals who "initially" plan to attend upper secondary education and decide to increase

Turning to compulsory education, let individual utility at the optimal level of schooling be $U\left(S^*(\theta,M),\theta,M\right)$ and define $\bar{\theta}$ as the value of θ that makes an individual indifferent between compulsory and higher education. It follows that individuals with $\theta \geq \bar{\theta}$ choose compulsory education and individuals with $\theta < \bar{\theta}$ choose higher than compulsory schooling. Therefore, $\pi_L = \int_{\bar{\theta}}^{\theta_{MAX}} f(\theta)d\theta$.

The threshold value $\bar{\theta}$ is defined by the following condition

$$U\left(S^*(\bar{\theta},M),\bar{\theta},M\right) - (\alpha_0 + \alpha_5 M) = 0 \qquad (7)$$

and the effect of M on $\bar{\theta}$ is given by[15]

$$\frac{\partial\bar{\theta}}{\partial M} = \frac{\alpha_2 + \alpha_4 S^*(\bar{\theta},M)}{\frac{1}{2}S^*(\bar{\theta},M)} \qquad (8)$$

Notice that the numerator of (8) is similar to that in (6). When the necessary conditions for human capital polarization hold ($\alpha_4 > 0$ and $\alpha_2 < 0$), $\frac{\partial\bar{\theta}}{\partial M}$ is likely to be negative since $S^*(\bar{\theta},M)$ must be close to compulsory education. Therefore, when M increases, more individuals prefer compulsory to higher education.

We illustrate human capital polarization graphically in Figure 2. In the figure, the marginal benefits of education (continuous thick line) are assumed to be decreasing in schooling and the marginal costs (dotted thin lines) to be increasing. Furthermore, individuals are ordered from the individual with the highest to the lowest θ. In the figure, we show two threshold individuals, one who chooses HS years of schooling (high school) and another who chooses C years (college).

their schooling "after" an inflow of immigrants. If instead only individuals with $\theta < \tilde{\theta}$ react by increasing their education when M raises, π_C will decrease, although some people already at college may decide to further prolong their education.

[15] We use the fact that $\partial U/\partial S=0$ at $S=S^*$.

An increase in the share of immigrants rotates the marginal benefits of schooling in such a way that the marginal benefits of education decline for low education and increase for high education (dotted thick line). At the new equilibrium, the threshold individual choosing just HS years of schooling has a lower θ, and the threshold individual choosing just C years has a higher θ. Polarization occurs.[16]

6. The effect of immigration on wages

Table 5 shows evidence of human capital polarization: the shares of lower and higher educated young natives increase because of an increase in the share of immigrants. In this section, we test whether the necessary conditions for human capital polarization ($\alpha_4 > 0$ and $\alpha_2 < 0$) hold in our data, consistently with the results found in Table 5.

We estimate the following empirical model

$$logw_{ipt} = \alpha_0 + \alpha_1 S_{ipt} + \alpha_2 M_{pt} S_{ipt} + \alpha_3 S_{ipt}^2 + \alpha_4 M_{pt}^2 S_{ipt} + \alpha_5 M_{pt} + \alpha_6 X_{pt-1} + \alpha_7 C_{ipt} + \mu_p + \omega_t + \varepsilon_{ipt} \quad (9)$$

where *w* is the monthly net wage, *S* is years of schooling, the index *i* is for the individual and the vector *C* includes cohort dummies and their interactions with schooling *S*. We consider natives aged 35-55, who work at least 35 hours per week (full-time workers). This age group is a plausible reference group for individuals aged between 19 and 27, who make their educational decision using their expectations of future wages.

One problem with estimating equation (9) is that it includes two endogenous variables, years of schooling and the share of immigrants. We can instrument the latter with Z. Since it is difficult to find a credible instrument for education, we remove the endogenous component of M by means of the instrument Z, and obtain $\widehat{M}$, and we use the results by Nizalova and

[16] The same result would be obtained if the marginal effect of education were increasing in M, provided that the concavity of the objective function is preserved.

Murtazashvili, 2016, who show that, under some conditions, OLS estimates of the interaction terms involving one exogenous ($\widehat{M}$) and one endogenous variable (S) are consistent. We discuss in Appendix 2 the conditions required to apply this approach to the current setup.

The LFS includes information on monthly net wages for each individual. The available data are both bottom-coded at 250 euros (first percentile of net wage distribution) and top-coded at 3000 euros per month (99^{th} percentile). We estimate (9) using both OLS on the censored sample and IV Tobit. Since the estimates turn out to be qualitatively similar, we report in Table 10 only the Tobit estimates, that are not subject to the sample selection bias caused by the trimming.

Importantly, we find that the estimated values of α_2 and α_4 are negative and positive respectively. Therefore, the necessary conditions for polarization are satisfied. We also compute the threshold value of S* at which the sign of the marginal effect of M on S* switches. It turns out that this value is 2.8 years after compulsory schooling for males and 8.9 years for females.

These estimates help us in explaining why we find that higher immigration increases the share of higher educated native males but have no significant effect on the share of higher educated females. In the case of males, an increase in the share of immigrants reduces the marginal benefits of education below 2.8 years of schooling above compulsory level, and increases them above that level. In the case of females, the decline in the marginal benefits of education at low levels of schooling is much larger than for males, and persists until 8.9 years of schooling, just above college.

Conclusion

In contrast with the empirical literature showing that immigration induces natives to invest in higher education, we have found evidence of human capital polarization, with the shares of both high educated and low educated natives increasing as a consequence of higher immigration. This evidence is particularly strong for males.

We have argued that our results are perfectly consistent with human capital theory. An increase in the share of (low skilled) immigrants reduces low skill wages and increase high skill wages, motivating some individuals to switch into higher education and some others to quit high school because of the reduced expected returns.

A potential alternative explanation of our results is negative self-selection. When the share of immigrants in a province increases, school quality falls and the most talented natives move out to provinces with fewer immigrants. As a result, the share of natives with lower education increases. Yet this mechanism is more likely to apply within rather than between provinces, as natives move from a school with many immigrants to schools with fewer immigrants in the same city or in neighboring cities (see Betts and Fairlie, 2003).

Our results indicate that immigration is another source of polarization, as much as globalization and technological progress. The economic literature so far (see for instance Acemoglu and Autor, 2011) has emphasized the latter but has almost overlooked the former.

In our setup, the polarization of educational choices is privately optimal, but has unpalatable aggregate consequences, because it reduces social cohesion by reducing the size of the middle class. Importantly, the less privileged class is unlikely to consist only of immigrants. By pushing many natives out of school too early, low skilled immigration is contributing to expand a native underclass.

Our results raise questions about what sort of immigration policy a country should select. On the one hand, attracting cheap unskilled labor from abroad can help supporting an industrial structure that relies more on the price of labor than on technological innovation. On the other hand, by delaying innovation and by reducing the human capital investments of many natives, this policy can have negative consequences on long - term productivity and international competitiveness and contribute to economic decline.

References

Acemoglu, D. and D. Autor, 2011. "Skills, tasks and technologies: Implications for employment and earnings." Handbook of Labor Economics, 4: 1043–1171.

ANVUR, 2014, Rapporto sullo stato del Sistema universitario e della ricerca nel 2013, Rome.

Barone, G. and Mocetti, S., 2011. "With a litthe help from abroad: the effect of low-skilled immigration on female labour supply." Labour Economics, 18, 664-675.

Betts, Julian R. 1998. "Educational Crowding Out: Do Immigrants Affect the Educational Attainment of American Minorities?" *In Help or Hindrance? The Economic Implications of Immigration for African-Americans*, ed. Daniel S. Hamermesh and Frank D. Bean, 253–281. New York: Russell Sage Foundation.

Betts, J. R. and Lofstrom M., 2000. "The Educational Attainment of Immigrants: Trends and Implications." In *Issues in the Economics of Immigration*, ed. George J. Borjas, 51–116. Chicago: University of Chicago Press.

Betts, J.R: and Fairlie, R., 2003. "Does immigration induce a 'native flight' from public schools into private schools?", Journal of Public Economics, 87, 987-1012.

Brunello, G. and De Paola, M., 2017. "School Segregation of Immigrants and its Effects on Educational Outcomes in Europe." EENEE Analytical Report n.30.

Card, D. 2001a. "Immigrant Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration." Journal of Labor Economics*,* 19(1). 22–64.

Card, D. 2001b. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." Econometrica, 69 (5), 1127-1160.

Conley, T., Hansen, C.B., and Rossi P., 2012, "Plausibly Exogenous." The Review of Economics and Statistics, 94(1), 260-272.

D'Amuri, F., Ottaviano, G., Peri, G., 2008. "The labor market impact of immigration in Western Germany in the 1990s." European Economic Review 54, 550–570.

Eberhard, J. 2012. "Immigration, Human Capital and the Welfare of Natives." University of Southern California working paper.

Hunt, J., 2016. "The impact of immigration on the Educational Attainment of Natives." Journal of Human Resources, doi: 10.3368/jhr.52.4.0115-6913R1.

International Organization for Migration, 2012, Labor Market Inclusion of the Less Skilled Migrants in the European Union.

ISTAT (2006a), "La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione." Collana Metodi e norme n. 32 - 2006 – Roma.

ISTAT (2006b), "Gli stranieri nella rilevazione sulle forze di lavoro". Collana Metodi e norme n. 27 - 2006 – Roma.

Jackson, O., 2011. "Does immigration crowd natives into or out of higher education?", mimeo, Northeastern University.

Llull, J. 2014. "Immigration, Wages and Education: A Labor Market Equilibrium Structural Model." CEMFI working paper.

Manacorda, M., Manning, A. and Wadsworth, J., 2012. "The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain." Journal of the European Economic Association, 10, 120–151.

Mitaritonna, C., Orefice, G. and Peri, G., 2017, "Immigrants and Firms' Outcomes: Evidence from France." European Economic Review, 96, 1-82. .

McHenry, P., 2015. "Immigration and the Human Capital of Natives." Journal of Human Resources, 50(1), 34-71.

Neymotin, F., 2009. "Immigration and its effect on the college-going outcomes of natives." Economics of Education Review, 28, 538–50.

Nizalova, O.Y., and Murtazashvili, I., 2016, "Exogenous Treatment and Endogenous Factors: Vanishing of Omitted Variable Bias on the Interaction Term." Journal of Economic Methodology, 5(1), 71–77.

OECD (2017). G20 Global Displacement and Migration Trends Report 2017. OECD, Paris.

Orrenius, P., and Zavodny, M. , 2015, "Does Immigration Affect Whether US Natives Major in Science and Engineering?" Journal of Labor Economics, 33 (S1), S79-S108.

Oster, E., 2016, "Unobservable Selection and Coefficient Stability: Theory and Evidence." Journal of Business & Economic Statistics, forthcoming.

Ottaviano, G. I., and Peri, G., 2012. "Rethinking the effect of immigration on wages." Journal of the European Economic Association, 10, 152–197.

Peri, G. and Sparber, C. 2009. "Task Specialization, Immigration and Wages."

American Economic Journal: Applied Economics 1(3), 135–169.

Ranson. T., and Winters, J.V., 2016, "Do Foreigners Crowd Natives out of STEM Degrees and Occupations? Evidence from the U.S. Immigration Act of 1990." IZA discussion paper 9920.

Roed M and Schone P. (2016) "Impact of Immigration on Inhabitants' Educational Investments." Scandinavian Journal of Economics 118(3), 433-462.

Stock James H., and Yogo M. 2005. "Testing for Weak Instruments in Linear IV Regressions." In Identification and Inference for Econometric Models, ed. Andrews Donald W.K., Stock James H., 80–108. Cambridge: Cambridge University Press.

Sweetman A. and van Ours J. (2014), "Immigration: What About the Children and Grandchildren?" CentER Discussion Paper Series No. 2014-009.

Appendix 1. Literature Review.

The economic literature on the impact of immigration on the educational choices of young natives starts with Betts, 1998, who presents evidence of a negative link between immigration and the probability of high school graduation for US minorities. Betts argues that this evidence supports the "educational crowding out" hypothesis, whereby an influx of immigrants reduces the effectiveness of public education for minorities who attend the same schools, discouraging them from completing school. This effect appears to outweigh any increase in high school graduation which theory suggests could result among natives if immigration serves to boost the returns to education.[17]

Later literature does not confirm these negative results. Jackson, 2011, for instance, investigates the impact of immigration on the college enrollment of U.S. natives. Using U.S. Census data from 1970 to 2000, he shows that state - level increases in the number of immigrant college students do not significantly lower the enrollment rates of U.S. natives. On the contrary, state-level increases in the ratio of unskilled to skilled immigrant workers significantly raise native enrollment rates. Similarly, Neymotin, 2009, investigates the effects of immigration on the SAT-scores and college application patterns of high school students in California and Texas. She finds that the 1990s immigration did not harm, and possibly benefited the student outcomes of U.S. citizens.[18]

More recently, McHenry, 2015, uses the National Education Longitudinal Study and US Census data to show that low - skilled immigration to an area induces local natives to improve their performance in school, attain more years of schooling, and take jobs that involve communication - intensive tasks for which they (native English speakers) have a comparative advantage. Their

[17] See also and Betts and Lofstrom, 2000.

[18] Brunello and De Paola, 2017, review the extensive literature on the effects of the share of immigrants in classes and schools on the school performance of immigrants and natives.

findings complement Peri and Sparber, 2009, who show that low skilled immigrants in the U.S. induce low- skilled natives to specialize in communication-oriented job skills rather than manual jobs.

In a similar vein, Hunt, 2016, uses US state panel data from 1940–2010 to examine the impact of immigration on the high school completion of natives in the United States. Immigrant children could influence native children's educational experience as well as their expected future labor market. She finds evidence for both channels and a positive net effect. An increase of one percentage point in the share of immigrants in the population aged 11–64 increases the probability that natives aged 11–17 eventually complete 12 years of schooling by 0.3 percentage point.

Roed and Schone, 2016, use panel data of male students, aged 16-18, from 2001-2008 in Norway to examine the impact of immigration in the building and construction (BC) sector on the enrolment of natives in vocational programs specialized in BC skills. They find that an increase in the share of immigrants reduces the enrolment of natives in vocational programs specialized in BC skills.

Some recent studies have focused on specific human capital investments. Orrenius and Zavodny, 2015, for instance, consider whether immigration may affect the likelihood that US natives major in science or engineering. Using data on college majors from the 2009–11 American Community Surveys, they find some evidence that immigration adversely affects whether US born women who graduated from college majored in a science or engineering field.

Ranson and Winters, 2016, examine the effects of the U.S. Immigration Act of 1990 on STEM (science, technology, engineering, and mathematics) degree completion and labor market outcomes for native-born Americans. The Act increased the in-flow and stock of foreign STEM workers in the U.S., both by increasing green card allotments and by expanding temporary work visas via the H-1B visa program. These policy changes potentially altered the relative desirability of various college majors and occupations for natives. They find that the Immigration Act changed natives' skill investment and utilization in three ways: (1) it pushed black males out of STEM majors; (2) it pushed white male STEM graduates out of STEM occupations; and (3) it pushed white female STEM graduates out of the workforce.

Appendix 2

Consider the following model:

$$lw = \alpha_0 + \alpha_1 S + \alpha_2 SM + \alpha_3 S^2 M + \alpha_4 M + \varepsilon \tag{A1}$$

where lw are log wages, M is the share of immigrants and S is years of schooling. Both M and S are endogenous variables. We are interested in consistent estimates of α_2 and α_3, the parameters associated to the two interaction terms between M and S. In the current setup, there is one instrument, Z, to correct for the endogeneity of M, but no credible instrument for S. We show that, under some assumptions, it is possible to consistently estimate α_2 and α_3.

Let (M,S,Z) be jointly normally distributed. We assume

Assumption A M is independent of S and Z is independent of (S,ε).

Assumption B $E(M)=0$ and $E(Z)=0$

Ignore for the moment the endogeneity of S and imagine to proceed as in the usual two-stage procedure, by instrumenting M with Z and SM and S^2M with SZ and S^2Z. We run three separate first stage equations and obtain $\widehat{M}$, $\widehat{SM}$ and $\widehat{S^2M}$ as predicted values. Specifically, the first stage regressions are

$$M = b_0 + b_1 Z + \vartheta \tag{A2}$$

for M,

$$SM = c_0 + c_1 SZ + \nu \tag{A3}$$

for SM, and similarly for S^2M .

The OLS estimate of c_1 is:

$$\overline{c_1} = \frac{\mathrm{cov(SM, SZ)}}{\mathrm{V(SZ)}} = \frac{\mathrm{E(SMSZ) - E(SM)E(SZ)}}{\mathrm{E(S^2Z^2) - E(SZ)^2}} \tag{A4}$$

By the law of iterated expectations, (A4) can be written as

$$\overline{c_1} = \frac{\mathrm{E(S^2E(MZ \mid S^2)) - E(S \times E(M \mid S))E(S \times E(Z \mid S))}}{\mathrm{E(S^2 \times E(Z^2 \mid S^2)) - E(S \times E(Z \mid S))^2}} \tag{A5}$$

The second stage regression is

$$lw = a_0 + a_1 S + a_2 \widehat{SM} + a_3 \widehat{S^2M} + a_4 \widehat{M} + \varepsilon \tag{A6}$$

We show that, when Assumptions A and B hold, this two-step procedure is equivalent to running only the first stage regression for M and by interacting the predicted value $\widehat{M}$ with S and S^2 to obtain

$$lw = A_0 + a_1 S + a_2 S\widehat{M} + a_3 S^2\widehat{M} + a_4 \widehat{M} + \varepsilon \tag{A7}$$

Proof.

Under Assumption B the first stage for M is

$$M = b_1 Z + \vartheta \tag{A8}$$

By Assumption A, equation (A5) can be written as

$$\overline{c_1} = \frac{\mathrm{E(S^2)E(MZ) - E(S)^2 E(M)E(Z)}}{\mathrm{E(S^2)E(Z^2) - E(S)^2 E(Z)^2}} \tag{A9}$$

By Assumption B, it can be further simplified into

$$\overline{c_1} = \frac{E(S^2)E(MZ)}{E(S^2)E(Z^2)} = \frac{E(MZ)}{E(Z^2)} = \overline{b_1} \tag{A10}$$

Eq. (A10) shows that $\overline{c_1}$ is equal to $\overline{b_1}$. Therefore, the predicted value from the first stage for SM is

$$\widehat{SM} = c_0 + \bar{c}_1 SZ = c_0 + S \times \bar{b}_1 Z = c_0 + S\widehat{M} \tag{A11}$$

so that $\widehat{SM}$ and $S\widehat{M}$ differ only by a constant term. We can apply the same reasoning to show that $\widehat{S^2M}$ is closely related to $S^2\widehat{M}$ and conclude that equation (A6) is equivalent to (A7).

We observe that equation (A7) fits the case discussed by Nizalova and Murtazashvili, 2016, who show that the OLS estimate of the interaction term between an exogenous variable $\widehat{M}$ and an endogenous variable S is consistent if $\widehat{M}$ is jointly independent of (ε, S), a requirement implied by Assumption A.

To investigate whether Assumption A holds in our context, we regress S on M (instrumented by Z) and Z in the sample of individuals aged between 35 and 55. These individuals completed their schooling 10 to 30 years before the realization of M and Z. When they decided their schooling, they did not have information on the level of immigration prevailing later in the future. As expected, we find no correlation between S and either M or Z. Assumption B just requires the demeaning of M and Z.

Appendix 3

Let the share of employed individuals in the group LN be given by

$$P(N \mid E) = \frac{P(NE)}{P(E)} \tag{B.1}$$

Where *P(NE)* is the employed in group LN and *P(E)* is total LN. Equation (B.1) can be re-written as $P(NE) = P(N|E) * P(E)$.

Differentiating with respect to M yields

$$\frac{dP(N \mid E)}{dM} = \frac{1}{P(E)}\left(\frac{dP(NE)}{dM} - \frac{dP(E)}{dM}P(N \mid E)\right) \qquad \text{(B.2)}$$

where $\frac{dP(N|E)}{dM}$ is the marginal effect of an increase in the share of immigrants on the share employed in group LN.

Table A1 shows the marginal effects of M on P(N|E), where N can be employment, self-employment, inactivity or the sector of employment. Table A2 considers instead the group HY and presents the marginal effects of M on the share enrolling in or completing different fields of study.

Tables and Figures

Table 1. Average share of male and female Italians aged 19 to 27, by year and education group.

Year	Less than high school - not in education or training (LN)	Less than high school - in education or training (LY)	High school - not in education or training (HN)	High school in education - college (HY)
		Males		
2006	0.262	0.046	0.369	0.298
2007	0.251	0.0482	0.373	0.303
2008	0.246	0.0485	0.362	0.319
2009	0.228	0.0527	0.381	0.31
2010	0.222	0.0529	0.385	0.31
2011	0.219	0.0543	0.383	0.314
2012	0.211	0.0586	0.383	0.312
2013	0.208	0.053	0.387	0.315
2014	0.193	0.0499	0.405	0.313
2015	0.184	0.0537	0.411	0.311
2016	0.179	0.0548	0.401	0.321
		Females		
2006	0.177	0.0356	0.323	0.438
2007	0.174	0.0388	0.309	0.452
2008	0.172	0.0382	0.292	0.471
2009	0.158	0.0445	0.297	0.472
2010	0.149	0.045	0.311	0.464
2011	0.144	0.0434	0.317	0.466
2012	0.14	0.0453	0.313	0.467
2013	0.137	0.0415	0.321	0.464
2014	0.127	0.0392	0.336	0.463
2015	0.123	0.0431	0.336	0.461
2016	0.115	0.0396	0.334	0.471

Source: Italian Labor Force Survey

Table 2. Average share of immigrants (from developing countries over total population), by year

Year	Share of Immigrants.
2006	0.037
2007	0.041
2008	0.048
2009	0.055
2010	0.060
2011	0.065
2012	0.059
2013	0.063
2014	0.069
2015	0.070
2016	0.070

Source: ISTAT, Demographic Balance and resident Population by Sex and Nationality

Table 3. Descriptive Statistics

	Males				Females			
	Means	St. dev.	Min	Max	Means	St. dev.	Min	Max
Less than high school - not in education or training	0.221	0.088	0.026	0.564	0.149	0.081	0	0.406
Less than high school - in education and training	0.052	0.021	0	0.24	0.041	0.019	0	0.155
With high school - not in education or training	0.384	0.063	0.091	0.635	0.316	0.06	0.064	0.547
With high school - in education; with college	0.311	0.066	0.086	0.661	0.462	0.084	0.175	0.766
Less than high school - not in education or training :								
Inactive	0.094	0.069	0	0.382	0.096	0.08	0	0.327
Self-employed	0.023	0.018	0	0.139	0.006	0.008	0	0.068
Employee	0.104	0.053	0	0.307	0.047	0.033	0	0.23
Manual worker	0.117	0.056	0	0.375	0.048	0.032	0	0.23
Employed in manufacturing	0.035	0.029	0	0.159	0.01	0.015	0	0.097
Employed in construction	0.032	0.025	0	0.188	0	0.002	0	0.032
Employed in wholesale and trade; hotel and restaurants	0.035	0.02	0	0.155	0.023	0.021	0	0.157
Employed in collective and personal services	0.005	0.007	0	0.063	0.011	0.011	0	0.097
Employed in other sectors	0.033	0.02	0	0.14	0.054	0.036	0	0.2
With high school - in education; with college:								
Inactive	0.257	0.059	0.057	0.509	0.374	0.076	0.138	0.626
Hard sciences	0.123	0.04	0	0.386	0.085	0.033	0	0.287
Medical sciences	0.04	0.021	0	0.139	0.073	0.031	0	0.226
Social sciences	0.104	0.035	0	0.365	0.17	0.047	0.017	0.393
Humanities	0.032	0.017	0	0.124	0.12	0.039	0.006	0.332
Age	23.072	0.238	21.684	24.059	23.052	0.249	22.028	24.247
Immigrant share	0.053	0.034	0.003	0.156				
Log GDP lagged	9.488	0.998	6.77	11.989				
Unemployment rate lagged	0.062	0.032	0.011	0.18				
Sectorial index lagged	0.422	0.069	0.234	0.631				

Table 4. First stage regressions, by gender. Dependent variable: share of immigrants in a province M

	(1) Males	(2) Females
Z	0.232***	0.235***
	(0.035)	(0.036)
Observations	1,210	1,210
R^2	0.984	0.984
Province fixed effects	Y	Y
Year fixed effects	Y	Y
Controls	Y	Y
F-test	42.76	43.49

Note: Z is the instrument introduced in Section 3. Robust standard errors clustered by province within parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5. Effect of the share of immigrants onthe educational choices of natives, age 19-27. IV estimates

	(1)	(2)	(3)	(4)
	Less than high school - not in education or training	Less than high school – in education and training	With high school - not in education or training	With high school – in education; with college
Panel A		Males		
Immigrant Share	1.254**	0.000	-3.212***	1.980***
	(0.624)	(0.347)	(0.630)	(0.595)
Panel B		Females		
Immigrant Share	1.743***	0.425***	-1.694***	-0.436
	(0.468)	(0.162)	(0.616)	(0.907)
Observations	1,210	1,210	1,210	1,210
Province fixed effects	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y
Controls	Y	Y	Y	Y

Notes: additional controls include: the log of provincial real GDP; the provincial unemployment rate for individuals aged 18-59; the local share of workers employed in sectors with a higher than average proportion of manual workers in the pre-sample period (year 2004); province-by-recession fixed effects. Robust standard errors clustered by province within parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6. Oster tests applied to the reduced form equation.

Panel A	Less than high school - not in education or training				
	$\hat{\beta}$	Std. Error	$\widehat{R^2}$	β^*	Identified set
Males (Rmax= 1, δ = 1)					
Instrument	0.291	0.143	0.846	1.063	[0.291, 1.063]
Females (Rmax=1, δ = 1)					
Instrument	0.449	0.124	0.868	0.941	[0.449, 0.941]

Panel B	With high school – in education; with college				
	$\hat{\beta}$	Std. Error	$\widehat{R^2}$	β^*	Identified set
Males (Rmax= 0.85, δ = 1)					
Instrument	0.460	0.179	0.650	0.262	[0.460, 0.262]
Females (Rmax= 0.94, δ = 1)					
Instrument	-0.114	0.248	0.722	-0.455	[-0.114, -0.455]

Table 7. Conley tests.

	LN males	HY males	LN females	HY females
M	1.254**	1.980***	1.743***	-0.436
	(0.624)	(0.595)	(0.468)	(0.907)
Prior of γ	-0.037	-0.079	0.035	-0.161
Observations	1,210	1,210	1,210	1,210
R-squared	0.023	0.031	0.006	0.018
Province fixed effects	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y
Lower Bound normal distribution	0.200	1.147	0.668	-1.458
Upper Bound normal distribution	2.644	3.517	2.511	2.256

Note: see Table 5.

Table 8. Effects of the share of immigrants on the occupational choices of natives. Group LN (less than high education, not in education and training). IV estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Inactive	Self-employed	Employee	Manual worker	Employed in manufacturing	Employed in construction	Employed in wholesale and trade; hotel and restaurants	Employed in collective and personal services	Employed in other sectors
Panel A					Males				
Imm. Share	-0.165	0.416*	1.004**	1.373***	0.326	0.375**	0.477**	0.131**	0.228
	(0.340)	(0.226)	(0.453)	(0.477)	(0.454)	(0.174)	(0.207)	(0.060)	(0.151)
Panel B					Females				
Imm. Share	1.303***	0.061	0.379	0.519*	0.145	0.006	0.118	0.137	-0.380
	(0.318)	(0.094)	(0.366)	(0.305)	(0.260)	(0.016)	(0.197)	(0.191)	(0.234)
Observations	1,210	1,210	1,210	1,210	1,210	1,210	1,210	1,210	1,210
Province FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: See Table 5

Table 9. Effects of the share of immigrants on the selected field of study. Group HY (higher education – in education or college) . IV estimates

	(1)	(2)	(3)	(4)
	College - hard sciences	College -medical sciences	College -social sciences	College - humanities
Panel A		Males		
Immigrant Share	-0.165	0.244	0.906**	0.380***
	(0.599)	(0.224)	(0.352)	(0.144)
Panel B		Females		
Immigrant Share	0.037	-0.639***	0.429	-0.400
	(0.308)	(0.208)	(0.637)	(0.473)
Observations	1,210	1,210	1,210	1,210
Province fixed effects	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y
Controls	Y	Y	Y	Y

Notes: see Table 5

Table 10. Dependent variable: ln (wage). Sample: natives aged 35-55. IV Tobit estimates.

	(1) Males	(2) Females
Years of education	0.042***	0.091***
	(0.001)	(0.001)
Years of education sq.	0.002***	-0.002***
	(0.000)	(0.000)
Years of education*cohort dummies	-0.001***	-0.001***
	(0.000)	(0.000)
Immigrant Share	1.172***	4.070***
	(0.301)	(0.575)
Immigrant Share*years of education	-0.022*	-0.524***
	(0.013)	(0.025)
Immigrant Share*years of education sq.	0.004**	0.029***
	(0.002)	(0.002)
Observations	588,754	507,062
$\frac{\partial s_i^*}{\partial M} > 0\ if$	S_i^*>2.823	S_i^*>8.888

Note: additional controls include: the log of provincial real GDP; the provincial unemployment rate for individuals aged 18-59; the local share of workers employed in sectors with a higher than average proportion of manual workers in the pre-sample period (year 2004);, cohort dummies, province and year fixed effects. Bootstrapped standard errors (500 replications) within parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 1 Change in the share of immigrants 2006-2016, by province.

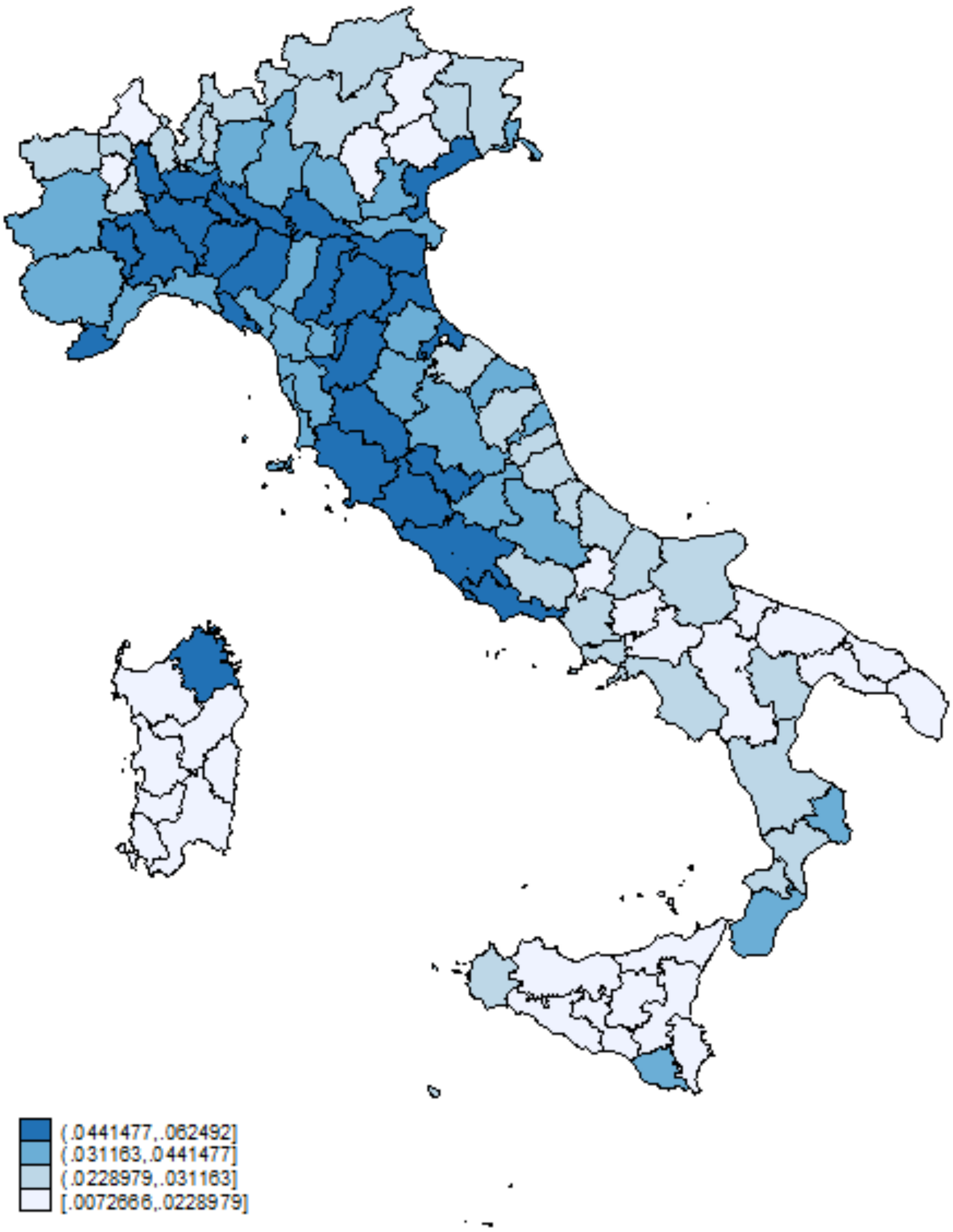

Figure 2. Human Capital Polarization

Figure 2. Human Capital Polarization

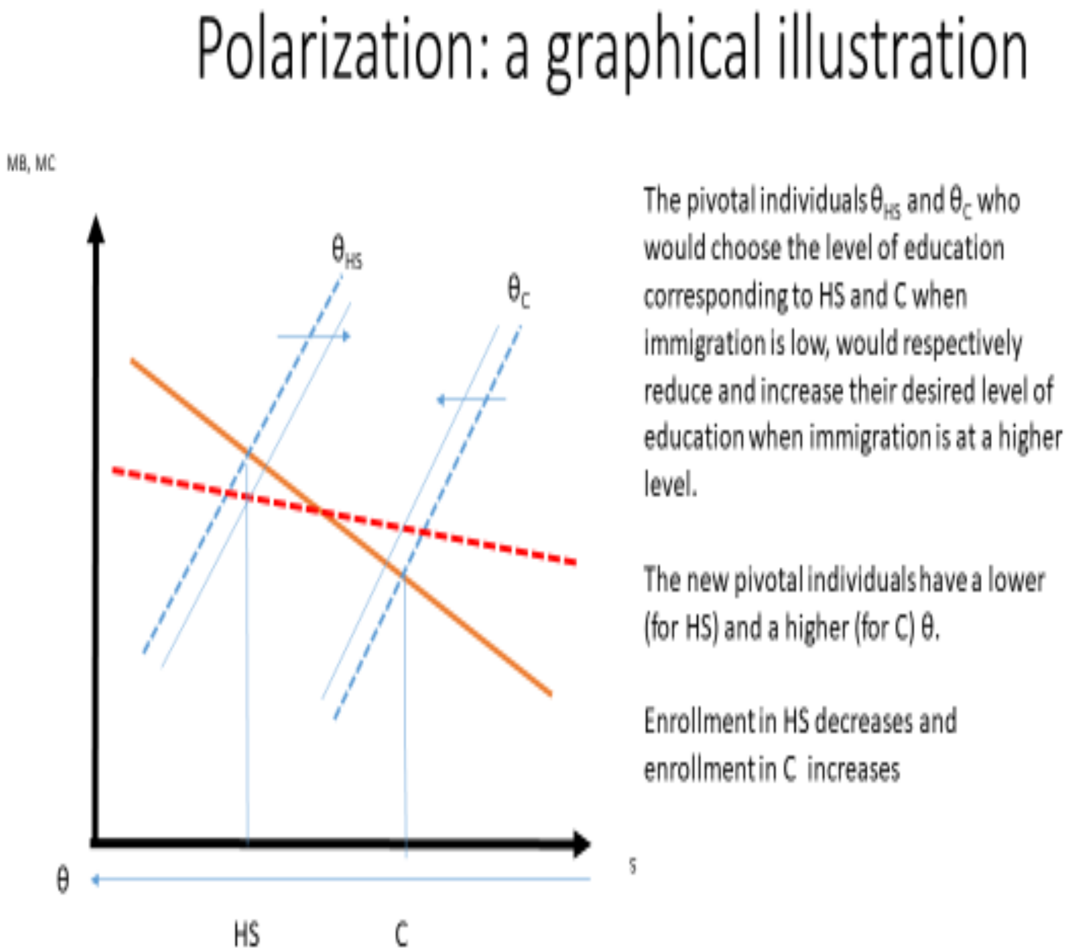

Table A1. Estimated effects of the share of immigrants on the occupational choices of natives, conditional on having less than high school – not in education or training.

	(1) Males	(2) Females
Inactive	-3.059***	2.537**
	(0.809)	(1.124)
Self-employed	1.321**	-0.278
	(0.568)	(0.526)
Employee	1.745**	-2.255**
	(0.845)	(1.108)
Manual worker	3.144***	-1.408
	(0.809)	(1.109)
Employed in manufacturing	0.431	-0.052
	(0.643)	(0.573)
Employed in construction	0.922	0.014
	(0.640)	(0.086)
Employed in wholesale and trade; hotel and restaurants	1.278*	-1.694*
	(0.712)	(0.924)
Employed in collective and personal services	0.484	-0.185
	(0.337)	(0.612)
Employed in other sectors	0.512	-3.717**
	(0.769)	(1.622)
Observations	269,583	255,754

Note: additional controls include: the log of provincial real GDP; the provincial unemployment rate for individuals aged 18-59; the local share of workers employed in sectors with a higher than average proportion of manual workers in the pre-sample period (year 2004); province-by-recession fixed effects. Bootstrapped standard errors within parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A2. Effect of the share of immigrants on the field of study chosen by those in higher education, conditional on belonging to group HY.

	(1) Males	(2) Females
Hard sciences	-3.088***	0.253
	(0.864)	(0.505)
Medical sciences	-0.038	-1.213**
	(0.543)	(0.497)
Social sciences	0.836	1.251*
	(0.819)	(0.655)
Humanities	0.578	-0.612
	(0.575)	(0.627)
Observations	269,583	255,754

Note: additional controls include: the log of provincial real GDP; the provincial unemployment rate for individuals aged 18-59; the local share of workers employed in sectors with a higher than average proportion of manual workers in the pre-sample period (year 2004); province-by-recession fixed effects. Bootstrapped standard errors within parentheses. *** p<0.01, ** p<0.05, * p<0.1